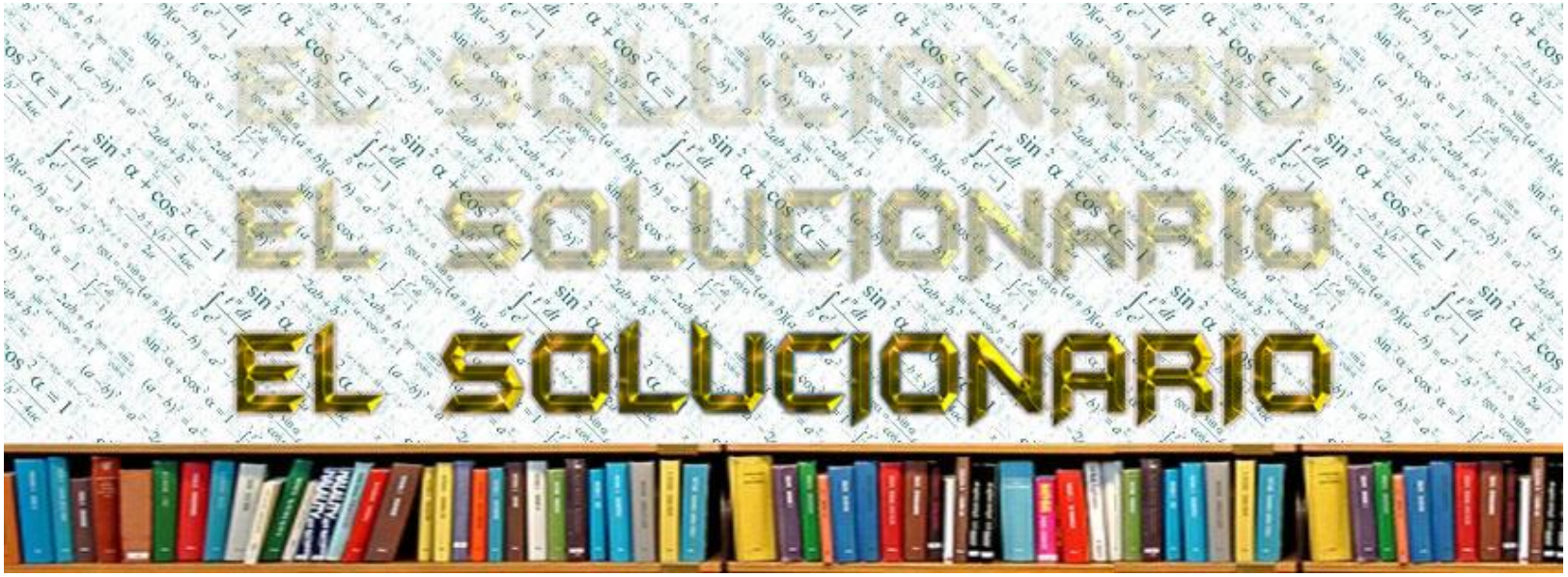
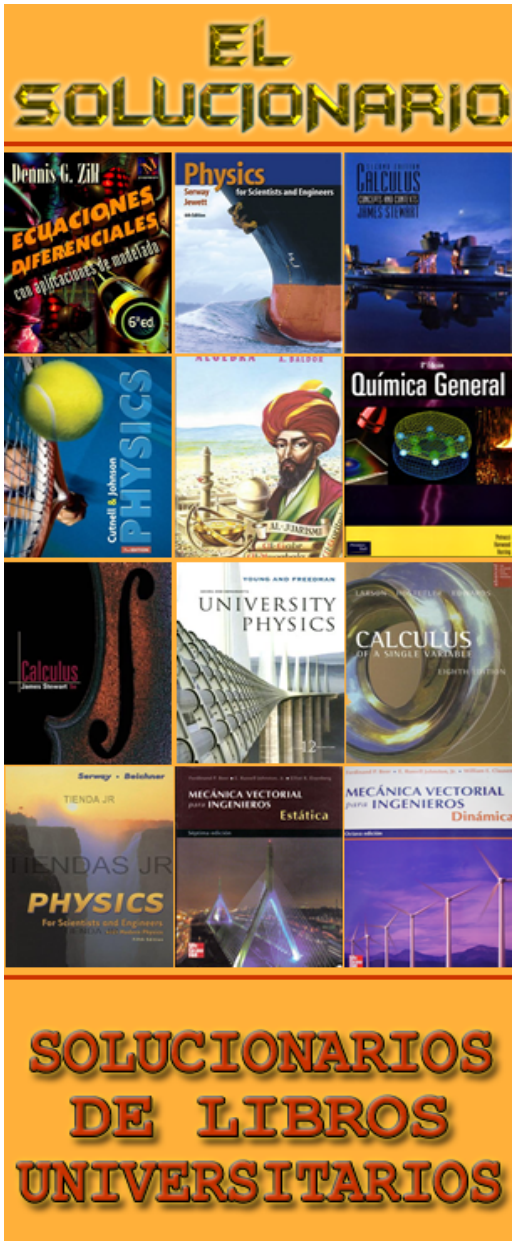


http://www.elsolucionario.blogspot.com



LIBROS UNIVERISTARIOS
Y SOLUCIONARIOS DE
MUCHOS DE ESTOS LIBROS

LOS SOLUCIONARIOS
CONTIENEN TODOS LOS
EJERCICIOS DEL LIBRO
RESUELTOS Y EXPLICADOS
DE FORMA CLARA

VISITANOS PARA
DESARGALOS GRATIS.

THOMSON

PROBABILIDAD Y ESTADÍSTICA

para ingeniería y ciencias

Jay L. DeVore

Sexta edición

Probabilidad y estadística para ingeniería y ciencias

SEXTA EDICIÓN

Jay L. Devore

California Polytechnic State University
San Luis Obispo

Revisión técnica

Leonardo Bañuelos Saucedo
UNAM-Facultad de ingeniería

THOMSON

Australia • Brasil • Canadá • España • Estados Unidos • México • Reino Unido • Singapur

THOMSON

Probabilidad y estadística para ingeniería y ciencias, 6ª. Ed.
Jay L. Devore

Director editorial y de producción:
José Tomás Pérez Bonilla

Editor de desarrollo:
Pedro de la Garza Rosales

Gerente de producción:
Luis Villanueva Rebollo

Editor de producción:
Alma Castrejón Alcocer

Supervisor de manufactura:
Israel Robles Martínez

Traducción:
Francisco Sánchez Fragoso

COPYRIGHT © 2005 por
International Thomson Editores,
S. A. de C. V., una división
de Thomson Learning, Inc.
Thomson Learning™ es una marca
registrada usada bajo permiso.

DERECHOS RESERVADOS.
Queda prohibida la reproducción o transmisión total o parcial del texto de la presente obra bajo cualesquiera formas, electrónica o mecánica, incluyendo fotocopiado, almacenamiento en algún sistema de recuperación de información, o grabado sin el consentimiento previo y por escrito del editor.

Impreso en México
Printed in Mexico
1 2 3 4 07 06 05

Para mayor información contáctenos en:
Séneca 53
Col. Polanco
México, D.F., 11560

Puede visitar nuestro sitio en
http://www.thomsonlearning.com.mx

Traducido del libro *Probability and Statistics for Engineering and the Sciences, 6th.*, publicado en inglés por Brooks Cole, ©2004
ISBN 0-534-39933-9
Datos para catalogación bibliográfica:
Devore, Jay L., *Probabilidad y estadística para ingeniería y ciencias, 6*a edición.
ISBN 970-686-457-1
1. Probabilidad y estadística para ingeniería y ciencias. 2. Estadística descriptiva; probabilidad; variables aleatorias y distribuciones de probabilidad discretas y continuas; muestras probabilísticas conjuntas; estimaciones puntuales; intervalos estadísticos; pruebas de hipótesis; inferencias; análisis de varianza; regresión lineal simple y correlación; regresión múltiple; bondad de las pruebas de ajuste; métodos de control de calidad.

División Iberoamericana

México y América Central:
Thomson Learning
Séneca 53
Col. Polanco
México, D. F., 11560
Tel. (52-55) 15 00 60 00
Fax (52-55) 5281 26 56
editor@thomsonlearning.com.mx

El Caribe:
Thomson Learning
598 Aldebarán
Altamira, San Juan
Puerto Rico
Zip Code: 00920
Tel. (787) 641-11 12
Fax (787) 641-11 19

Cono Sur
Buenos Aires, Argentina
thomson@thomsonlearning.com.ar

América del Sur:
Thomson Learning
Calle 39 No. 24-09
La Soledad
Bogotá, Colombia
Tel. (571) 340-9470
Fax (571) 340-9475
cliente@thomsonlearning.com.co

España:
Thomson Learning
Calle Magallanes 25
28015 Madrid
España
Tel. 34 (0) 91 446-3350
Fax 34 (0) 91 445-6218
clientes@paraninfo.es

Esta obra se terminó de imprimir en septiembre del 2005
en los talleres de Impresos y Acabados Marbeth
S.A. de C.V. Privada de Alamo No. 35
Col. Arenal C.P. 02980, México, D.F.

A mi esposa Carol:

Tu dedicación a la enseñanza

es para mí una inspiración continua.

A mis hijas Allison y Teresa:

El gran orgullo que siento por sus

logros es totalmente sincero.

Contenido

1 Panorama general y estadística descriptiva 1

Introducción 1
1.1 Poblaciones, muestras y procesos 3
1.2 Métodos gráfico y tabular en estadística descriptiva 11
1.3 Medidas de localización 28
1.4 Medidas de variabilidad 36
Ejercicios complementarios 48
Bibliografía 51

2 Probabilidad 52

Introducción 52
2.1 Espacios muestrales y eventos 53
2.2 Axiomas, interpretaciones y propiedades de la probabilidad 58
2.3 Técnicas de conteo 67
2.4 Probabilidad condicional 75
2.5 Independencia 86
Ejercicios complementarios 92
Bibliografía 96

3 Variables aleatorias discretas y distribuciones de probabilidad 97

Introducción 97
3.1 Variables aleatorias 98
3.2 Distribuciones de probabilidad para variables aleatorias discretas 101
3.3 Valores esperados de variables aleatorias discretas 111
3.4 Distribución de pobabilidad binomonial 120
3.5 Distribuciones hipergeométrica y binomial negativa 128
3.6 Distribución de probabilidad de Poisson 135
Ejercicios complementarios 140
Bibliografía 144

4 Variables aleatorias continuas y distribuciones de probabilidad 145

Introducción 145
4.1 Variables aleatorias continuas y funciones de densidad de probabilidad 146
4.2 Funciones de distribución acumulada y valores esperados 152
4.3 Distribución normal 160
4.4 Distribución gamma y sus relativos 174
4.5 Otras distribuciones continuas 181
4.6 Gráficas de probabilidad 188
Ejercicios complementarios 199
Bibliografía 204

5 Distribuciones de probabilidad conjunta y muestras aleatorias 205

Introducción 205
5.1 Variables aleatorias con distribución conjunta 206
5.2 Valores esperados, covarianza y correlación 219
5.3 Estadísticos y sus distribuciones 225
5.4 Distribución de la media muestral 237
5.5 Distribución de una combinación lineal 243
Ejercicios complementarios 249
Bibliografía 252

6 Estimación puntual 253

Introducción 253
6.1 Algunos conceptos generales de la estimación puntual 254
6.2 Métodos de estimación puntual 269
Ejercicios complementarios 278
Bibliografía 279

7 Intervalos estadísticos basados en una sola muestra 281

Introducción 281
7.1 Propiedades básicas de intervalos de confianza 282
7.2 Intervalos de confianza de muestras grandes para una media y proporción poblacionales 291
7.3 Intervalos basados en una distribución poblacional normal 299
7.4 Intervalos de confianza para la varianza y la desviación estándar de una población normal 308
Ejercicios complementarios 311
Bibliografía 314

8 Pruebas de hipótesis con base en una sola muestra 315

Introducción 315
8.1 Hipótesis y procedimientos de prueba 316
8.2 Pruebas acerca de una media poblacional 326

8.3 Pruebas relacionadas con una proporción poblacional 339
8.4 Valores *P* 344
8.5 Algunos comentarios acerca de seleccionar un procedimiento de prueba 352
Ejercicios complementarios 356
Bibliografía 359

9 Inferencias basadas en dos muestras 360

Introducción 360
9.1 Pruebas *z* e intervalos de confianza para una diferencia entre dos medias poblacionales 361
9.2 Prueba *t* de dos muestras e intervalo de confianza 372
9.3 Análisis de pares de datos 381
9.4 Inferencias en relación con una diferencia entre proporciones poblacionales 391
9.5 Inferencias relacionadas con dos varianzas poblacionales 399
Ejercicios complementarios 403
Bibliografía 409

10 Análisis de varianza 410

Introducción 410
10.1 ANOVA de un solo factor 411
10.2 Comparaciones múltiples en el ANOVA 422
10.3 Más cerca de ANOVA de una solo factor 428
Ejercicios complementarios 438
Bibliografía 440

11 Análisis de varianza con varios factores 441

Introducción 441
11.1 ANOVA de dos factores con $K_{ij} = 1$ 442
11.2 ANOVA de dos factores con $K_{ij} > 1$ 456
11.3 ANOVA de tres factores 465
11.4 Experimentos factoriales 2^p 476
Ejercicios complementarios 491
Bibliografía 494

12 Regresión lineal simple y correlación 496

Introducción 496
12.1 Modelo de regresión lineal simple 497
12.2 Estimación de los parámetros del modelo 505
12.3 Inferencias acerca del parámetro de la pendiente β_1 520
12.4 Inferencias relacionadas con $\mu_{Y \cdot x^*}$ y predicción de valores *Y* futuros 530
12.5 Correlación 538
Ejercicios complementarios 549
Bibliografía 554

13 Regresión no lineal y múltiple 555

Introducción 555
13.1 Pertinencia y comprobación del modelo 556
13.2 Regresión con variables transformadas 564
13.3 Regresión polinomial 576
13.4 Análisis de regresión múltiple 587
13.5 Otros temas de la regresión múltiple 612
Ejercicios complementarios 626
Bibliografía 632

14 Pruebas de bondad de ajuste y análisis de datos categóricos 633

Introducción 633
14.1 Pruebas de bondad de ajuste cuando se especifican por completo las probabilidades de categoría 634
14.2 Pruebas de bondad del ajuste para hipótesis compuestas 642
14.3 Tablas de contingencia bidireccionales 655
Ejercicios complementarios 663
Bibliografía 666

15 Procedimientos libres de distribución 667

Introducción 667
15.1 Prueba de Wilcoxon de los rangos con signo 668
15.2 Prueba de Wilcoxon de la suma de rangos 677
15.3 Intervalos de confianza sin distribución 684
15.4 ANOVA sin distribución 689
Ejercicios complementarios 693
Bibliografía 695

16 Métodos de control de calidad 696

Introducción 696
16.1 Comentarios generales acerca de las gráficas de color 697
16.2 Gráficas de control para localización de proceso 699
16.3 Gráficas de control para variación de proceso 708
16.4 Gráficas de control para atributos 713
16.5 Procedimientos CUSUM 718
16.6 Muestreo de aceptación 727
Ejercicios complementarios 733
Bibliografía 734

Apéndice de tablas 735

A.1 Probabilidades binomiales acumuladas 736
A.2 Probabilidades de Poisson acumuladas 738
A.3 Áreas de la curva normal estándar 740

A.4 Función gamma incompleta 742
A.5 Valores críticos para la distribución 743
A.6 Valores críticos de la tolerancia para las distribuciones normales de la población 744
A.7 Valores críticos para la distribución de chi cuadrada 745
A.8 Áreas de cola de la curva *t* 746
A.9 Valores críticos para la distribución de *F* 748
A.10 Valores críticos para la distribución del rango estudentizado 754
A.11 Áreas de colas de la curva chi cuadrada 755
A.12 Valores críticos para la prueba de normalidad de Ryan-Joiner 757
A.13 Valores críticos para la prueba de rangos con signo de Wilcoxon 758
A.14 Valores críticos para la prueba de suma de rangos de Wilcoxon 759
A.15 Valores críticos para el intervalo de rangos con signo de Wilcoxon 760
A.16 Valores críticos para el intervalo de suma de rangos de Wilcoxon 761
A.17 Curvas β para pruebas *t* 762

Respuestas a ejercicios con número impar 763
Índice 785

Prólogo

Objetivo

El uso de modelos probabilísticos y métodos estadísticos para analizar datos, se ha vuelto práctica común en casi todas las disciplinas científicas. Mediante esta obra se pretende proporcionar una introducción completa a los modelos y métodos que los estudiantes de ingeniería y ciencias naturales encontrarán en sus carreras. Aunque los ejemplos y ejercicios se diseñaron en especial para científicos e ingenieros, la mayor parte de los métodos estudiados son esenciales para el análisis estadístico en muchas otras áreas, de manera que los estudiantes de administración y ciencias sociales también enriquecerán sus conocimientos con esta obra.

Método

Los estudiantes de un curso de estadística diseñado como auxiliar de otras materias importantes pueden mostrar, en un inicio, su escepticismo con respecto al valor e importancia de esta materia, pero mi experiencia es que la estadística *atrae* a los estudiantes si se emplean buenos ejemplos y ejercicios que combinen sus experiencias diarias con sus intereses científicos. Por consiguiente, trabajé con ahínco para hallar ejemplos con información real, no ficticia, que alguien consideró valioso recopilar y analizar. Muchos de los métodos presentados, sobre todo en los últimos capítulos sobre inferencia estadística, se ilustran mediante el análisis de datos tomados de un trabajo ya publicado, y en diversos ejercicios también se requiere trabajar con esos datos. En ocasiones es posible que el lector desconozca el contexto de un problema en particular (como de hecho muchas veces me sucedió), pero encontré que los estudiantes se sienten más atraídos por problemas reales, dentro de un contexto poco familiar, que por problemas evidentemente irreales en un contexto conocido.

Nivel matemático

La exposición es relativamente modesta en términos de desarrollo matemático. Sólo en el capítulo 4 y en algunas secciones del 5 y 6 se utiliza el cálculo. En particular, con excepción de alguna observación ocasional o paréntesis, el cálculo sólo aparece en la parte de inferencia del libro en la segunda sección del capítulo 6. El álgebra de matrices no se usa en absoluto. Por lo tanto, casi toda la exposición debe ser accesible para quienes ya hayan cursado un semestre o dos trimestres de cálculo diferencial e integral.

Contenido

El capítulo 1 inicia con algunos conceptos básicos y terminología: población, muestra, estadística descriptiva e inferencial, estudios enumerativos contra analíticos, etcétera, y continúa con un examen de los métodos descriptivos numéricos y gráficos más importantes. En el capítulo 2, se plantea el enfoque tradicional de la probabilidad, seguido de distribuciones de probabilidad de variables aleatorias discretas y continuas en los capítulos 3 y 4, respectivamente. Las distribuciones conjuntas y sus propiedades se estudian en la primera parte del capítulo 5. Al final de éste se introducen los estadísticos y sus distribuciones de muestreo, que forman el puente entre la probabilidad y la inferencia. En los tres capítulos siguientes se tratan la estimación puntual, intervalos de confianza y prueba de hipótesis con base en una sola muestra. Los métodos de inferencia con dos muestras independientes y pares de datos se estudian en el capítulo 9. El análisis de varianza es el tema de los capítulos 10 y 11 (un solo factor y factores múltiples, respectivamente). La regresión aparece por primera vez en el capítulo 12 (modelo de regresión lineal simple y correlación), y se retoma para un análisis más profundo en el capítulo 13. Los métodos chi cuadrada, procedimientos no paramétricos y técnicas de control estadístico de calidad se desarrollan en los últimos tres capítulos.

Auxiliares en el aprendizaje de los estudiantes

Aunque el nivel matemático del libro representa muy pocas dificultades para la mayor parte de los estudiantes de ingeniería y ciencias, llegar a comprender los conceptos y apreciar el desarrollo lógico de la metodología podría requerir en muchas ocasiones un esfuerzo sustancial. Para ayudar a que los estudiantes adquieran este conocimiento y valoración, se incluyeron numerosos ejercicios que varían en grado de dificultad, desde los más sencillos, para los que se requiere la aplicación rutinaria del material del texto, hasta algunos que piden que el lector amplíe los conceptos estudiados en situaciones nuevas. Hay mucho más ejercicios de los que la mayor parte de los profesores quisieran asignar durante cualquier curso en particular, pero se recomienda pedir a los estudiantes que resuelvan un gran número de ellos. En una disciplina en que se solucionan problemas, la participación activa en dichas soluciones es la forma más segura de identificar y llenar huecos en el aprendizaje. Las respuestas de ejercicios de número impar aparecen en la sección de respuestas, al final del texto.

Lo nuevo de esta edición

Entre los cambios de la sexta edición se encuentran los siguientes:

- Se incluyen muchos ejemplos y ejercicios nuevos que se basan en información real. Algunos de ellos son menos estructurales que los característicos ejercicios de libro de texto. Por ejemplo, un ejercicio del capítulo 13, proporciona un conjunto de datos de regresión múltiple (que al principio se parece a un conjunto de datos de regresión lineal simple, excepto que también se incluye información acerca de un predictor categórico); al estudiante sólo se le pide que ajuste varios modelos y saque conclusiones apropiadas. Un ejercicio al final del capítulo 14 incluye partes que tienen que ver con la correlación, ANOVA, un análisis t con dos muestras y un análisis chi cuadrada, sin que de hecho se especifique. En general, aparte de hacer énfasis en la mecánica, se destaca también la interpretación. Por ejemplo, en el capítulo 1, se pide al estudiante que además de construir un diagrama de caja exprese lo que és-

te le sugiere. En algunos ejercicios se introducen nuevos conceptos o técnicas, por ejemplo, una gráfica seminormal en el capítulo 4 y la autocorrelación para series de tiempo en el capítulo 12.

- En una definición ampliada de una hipótesis al principio del capítulo 8, se incluye una aseveración de que la distribución de la población tiene una forma particular (las hipótesis dejan de ser sólo afirmaciones acerca de parámetros).
- Hay una descripción más detallada de procedimientos t combinados para proporcionar mejor equilibrio entre análisis no agrupados y agrupados.
- Se ofrece mayor información acerca de opciones para fórmulas comunes a fin de obtener estimaciones por intervalo de una sola proporción y una diferencia entre proporciones sin las fórmulas tradicionales.
- La sección de comparaciones múltiples en ANOVA se volvió a escribir con el fin de destacar mejor la naturaleza del nivel de confianza simultáneo y el procedimiento de Tukey.
- Se añadió un análisis de los cálculos de la potencia de MINITAB® en ANOVA de un solo factor.
- Se hace más énfasis en la naturaleza de la variación al estimar la pendiente en una regresión lineal simple.
- Se presenta el método de LOWESS para ajustar una función de regresión no lineal.
- Se ofrece otra interpretación de los coeficientes de regresión en la regresión múltiple por medio de las gráficas residuales parciales.
- Se reforzó y afinó la exposición siempre que se requirió.
- Prueba BCA, una herramienta de elaboración de pruebas y administración de curso.

Cobertura recomendada

Hay suficiente material en este libro para un curso de un año de duración. Si el curso es de menor duración, es necesario elegir con cuidado los temas a tratar. En Cal Poly se imparten dos trimestres seguidos (cuatro horas de clase por semana). Durante las diez primeras semanas se estudia gran parte del material de los capítulos 1 a 7, tratando un poco las distribuciones conjuntas y los detalles de la estimación por máxima verosimilitud y el método de momentos. El segundo trimestre comienza con las pruebas de hipótesis y luego se estudian inferencias con dos muestras, ANOVA, regresión y selecciones de los capítulos sobre chi cuadrada, métodos no paramétricos y control de calidad. Es evidente que el estudio del material en un curso de un semestre tiene limitaciones. Y, naturalmente, ¡nunca hay tiempo suficiente para enseñar a los alumnos todo lo que nos gustaría que supieran!

Complementos

Este libro cuenta con una serie de complementos para el profesor, los cuales están en inglés y sólo se proporcionan a los docentes que adopten la presente obra como texto para sus cursos. Para mayor información, favor de comunicarse con las oficinas de nuestros representantes o a los siguientes correos electrónicos.

Thomson México y Centroamérica clientes@thomsonlearning.com.mx.

Thomson América del Sur clithomson@andinet.com

Thomson Caribe amy.reyes@thomsonlearning.com

Los archivos virtuales para realizar algunos de los ejercicios de este libro, pueden encontrarse y ser descargados en el sitio http://www.thomson.com.mx

Agradecimientos

Mis colegas de Cal Poly me ofrecieron apoyo y estímulo de valor incalculable en estos años. También deseo expresar mi gratitud y reconocimiento a quienes estudiaron en las ediciones anteriores y me hicieron numerosas sugerencias a fin de mejorar (me señalaron también algunos errores). Gracias especialmente a Julie Seely.

Agradezco a los siguientes revisores su valiosos comentarios: Robert L. Armacost, University of Central Florida; Douglas M. Bates, University of Wisconsin–Madison; David Clark, California State Polytechnic University at Pomona; David M. Cresap, University of Portland; Don E. Deal, University of Houston; Charles E. Donaghey, University of Houston; Patrick J. Driscoll, U.S. Military Academy; Mark Duva, University of Virginia; Nasser S. Fard, Northeastern University; Celso Grebogi, University of Maryland; James J. Halavin, Rochester Institute of Technology; Jennifer Hoeting, Colorado State University; Wei-Min Huang, Lehigh University; Stephen Kokoska, Colgate University; Arnold R. Miller, University of Denver; John J. Millson, University of Maryland; Don Ridgeway, North Carolina State University; Larry J. Ringer, Texas A&M University; Richard M. Soland, The George Washington University; Clifford Spiegelman, Texas A&M University; Jery Stedinger, Cornell University; David Steinberg, Tel Aviv University; G. Geoffrey Vining, University of Florida; Bhutan Wadhwa, Cleveland State University, y Michael G. Zabetakis, University of Pittsburgh.

Aprecio en todo lo que valen los servicios de producción que me proporcionó Martha Emry (quien siempre se mostró amable, solidaria y oportuna en sus conversaciones conmigo). Gracias a la editora Pamela Rockwell. El personal de Duxbury y Brooks/Cole manifestó, como siempre, gran colaboración; agradezco en particular a Carolyn Crockett, Curt Hinrichs, Ann Day, Julie Bliss, Joseph Rogove y Karen Haga. También deseo manifestar mi admiración a los representantes de ventas de Thomson Learning por su duro trabajo, a través de los años, para obtener la opinión acerca de las ediciones anteriores de este libro y de otros que he escrito. Por último, no puedo expresar con palabras la gratitud que siento hacia mi esposa, Carol, por apoyarme a lo largo de los últimos 30 años en la difícil tarea de escribir.

Jay L. Devore

1

Panorama general y estadística descriptiva

Introducción

Los conceptos y los métodos estadísticos, además de ser útiles, también suelen ser indispensables para comprender el mundo que nos rodea. Proporcionan formas para reflexionar acerca del comportamiento de muchos fenómenos con los que se enfrenta el especialista que eligió como campo la ingeniería o las ciencias.

La disciplina de la estadística enseña cómo razonar de manera lógica y tomar decisiones informadas en presencia de incertidumbre y variación. Sin estas dos últimas, habría poca necesidad de los métodos estadísticos o de los profesionales de la estadística. Si cada pieza de un tipo específico tuviera exactamente la misma vida, si todos los resistores que produce determinado fabricante tuvieran la misma resistencia, si las determinaciones de pH en muestras de suelo de determinado lugar fueran idénticas, etc., entonces se conocería toda la información que se desea con una sola observación.

Una manifestación interesante de la variación surge cuando se efectúan los análisis de emisión de gases en los vehículos automotores. Los costos y el tiempo necesario para el *Federal Test Procedure*, FTP (Procedimiento federal de prueba) en Estados Unidos evitan la difusión de su uso en los programas de inspección vehicular. Como resultado, muchos organismos elaboraron pruebas más baratas y rápidas, con la esperanza de reproducir los resultados FTP. Según el artículo "Motor Vehicle Emissions Variability" (J. *of the Air and Waste Mgmt. Assoc.*, 1996: 667-675), la aceptación del FTP como patrón de excelencia ocasionó la creencia muy difundida de que las mediciones repetidas en el mismo vehículo darían resul-

tados idénticos o casi idénticos. Los autores del artículo aplicaron el FTP a siete vehículos caracterizados como "grandes emisores". A continuación se indican los resultados de uno de esos vehículos:

HC (gm/milla)	13.8	18.3	32.2	32.5
CO (gm/milla)	118	149	232	236

La variación apreciable en las mediciones de HC y CO genera grandes dudas acerca de la creencia común, y dificulta aún más la evaluación precisa de los niveles de emisión.

¿Cómo se aplican las técnicas estadísticas para reunir información y sacar conclusiones? Supóngase, por ejemplo, que un ingeniero de materiales ha desarrollado un revestimiento para retardar la corrosión de tubos metálicos en circunstancias específicas. Si se aplica dicho revestimiento a diferentes secciones de tubo, la variación en las condiciones ambientales y en las mismas partes revestidas, dará como resultado corrosión más intensa en algunos tramos que en otros. Se podrían usar los métodos del análisis estadístico con los datos obtenidos en ese experimento para decidir si la cantidad *promedio* de corrosión sobrepasa al límite superior establecido, o bien para pronosticar cuánta corrosión se producirá en un solo tramo del tubo.

Supóngase también que el ingeniero desarrolló la nueva fórmula suponiendo que es mejor que el revestimiento actual. Podría efectuar un experimento comparativo para investigar esta posibilidad, aplicando la fórmula actual a algunos tramos del tubo y la nueva a otros. Esto se debe hacer con cuidado, porque de otro modo se podría llegar a una conclusión errónea. Por ejemplo, quizá la cantidad promedio de corrosión sea idéntica para los dos revestimientos. Sin embargo, la nueva fórmula se podría aplicar a tramos que tienen una mayor aptitud para resistir la corrosión y se encuentran en condiciones ambientales menos rigurosas que los segmentos con la pintura actual que tienen condiciones distintas. En ese caso, es probable que el investigador observe una diferencia entre las dos pinturas que no se puede atribuir sólo a los revestimientos, sino a variaciones de otro tipo. La estadística proporciona métodos para analizar los resultados de experimentos efectuados, pero también indica cómo se pueden efectuar los experimentos de manera eficaz para disminuir los efectos de la variación y tener mayor probabilidad de llegar a conclusiones correctas.

1.1 Poblaciones, muestras y procesos

Los ingenieros y los científicos tratan de forma constante con conjuntos de hechos, o **datos**, tanto en su práctica profesional como en la vida cotidiana. La disciplina de la estadística proporciona métodos para organizar y resumir datos, y para sacar conclusiones con base en la información que contienen los datos.

Una investigación por lo común tiene que ver con una colección muy bien definida de objetos que forman una **población** de interés. En un estudio, la población podría consistir en todas las cápsulas de gelatina, de determinado tipo, producidas durante un periodo específico. En otra investigación se estudiaría la población formada por los individuos que recibieron una licenciatura en ingeniería en el año académico inmediato anterior. Cuando se cuenta con la misma información para todos los objetos de la población, se tiene lo que se llama **censo**. Las restricciones de tiempo, dinero y demás recursos insuficientes, hacen impráctico o imposible levantar un censo. En vez de eso, se selecciona de alguna manera prescrita un subconjunto de la población, una **muestra**. Por consiguiente, se podría obtener como base una muestra de rodamientos de una corrida de producción determinada, con el fin de investigar si éstos se apegan a las especificaciones de manufactura; o bien, se podría seleccionar una muestra de los ingenieros graduados el año anterior para obtener retroalimentación con respecto a la calidad de los planes de estudio en ingeniería.

Por lo regular, el interés se centra sólo en ciertas características de los objetos de una población: la cantidad de defectos en la superficie de cada tramo de tubo, el espesor de pared de cada cápsula, el género de un graduado en ingeniería, la edad a la que se graduó, entre otras. Una característica puede ser categórica, como el género o tipo de defecto, o bien puede ser de naturaleza numérica. En el primer caso, el *valor* de la característica es una categoría (como género femenino o soldadura insuficiente), y en el último caso, el valor es un número (p. ej., edad = 23 años, o diámetro = 0.502 cm). Una **variable** es cualquier característica cuyo valor cambia entre los objetos de la población. Las variables se representarán al inicio mediante las letras minúsculas del final del alfabeto, por ejemplo:

x = marca de la calculadora que posee un estudiante

y = cantidad de defectos notorios en un automóvil recién fabricado

z = distancia de frenado de un automóvil en condiciones específicas

Los datos son el resultado de observaciones efectuadas en una sola variable, o bien, de observar de modo simultáneo dos o más variables. Un conjunto de datos **univariados** consiste en observaciones de una sola variable. Por ejemplo, se podría determinar el tipo de transmisión automática(A) o manual (M) de cada uno de los 10 automóviles recién comprados en determinada agencia, lo cual generaría un conjunto de datos categóricos

M A A A M A A M A A

La siguiente muestra de la vida útil (en horas) de baterías de la marca D con determinado uso es un conjunto univariado de datos numéricos:

5.6 5.1 6.2 6.0 5.8 6.5 5.8 5.5

Los datos **bivariados** se obtienen cuando se observan dos variables. El conjunto de datos podría consistir en un par (p. ej., altura y peso) de características para cada jugador de un equipo de basquetbol; la primera observación podría ser (72, 168), la segunda sería (75, 212), y así sucesivamente. Si un ingeniero determina el valor tanto de x = vida útil de la pieza como de y =

causa de la falla de la pieza, el conjunto de datos que resulta es bivariado, en el que una variable es numérica y la otra categórica. Se tienen datos **multivariados** cuando se hacen observaciones en más de dos variables. Por ejemplo, un médico investigador podría determinar la presión arterial sistólica, la presión arterial diastólica y la concentración de colesterol en el suero de cada paciente que participa en un estudio. Cada observación produciría una tercia de cantidades, como (120, 80, 146). En muchos conjuntos de datos multivariados, algunas variables son numéricas y otras son categóricas. Por ejemplo, la edición anual de *Consumer Reports* sobre automóviles proporciona los valores de variables como tipo de vehículo (pequeño, deportivo, compacto, mediano, grande), rendimiento de combustible en ciudad (millas por galón), rendimiento de combustible en carretera (millas por galón), tipo de unidad motriz (tracción trasera, tracción delantera, tracción en cuatro ruedas), entre otros.

Ramas de la estadística

Quizá un investigador que reunió datos quiera sólo resumirlos y explicar las particularidades importantes de ellos. Estas actividades requieren usar métodos de la **estadística descriptiva**. Algunos de esos métodos son de naturaleza gráfica; el trazo de histogramas, diagramas de caja y de dispersión, es el ejemplo principal. Para otros métodos descriptivos se necesita el cálculo de medidas numéricas de resumen, como medias, desviaciones estándar y coeficientes de correlación. La abundante disponibilidad de paquetes de cómputo estadístico facilita mucho más el cálculo que antes. Las computadoras tienen mejor desempeño que los seres humanos para calcular y elaborar gráficas (¡una vez que el usuario proporcionó las instrucciones adecuadas!). Esto quiere decir que el investigador no tiene que dedicar muchos esfuerzos al "trabajo tedioso" y que cuenta con más tiempo para estudiar los datos y captar los mensajes importantes. En esta obra se proporcionan resultados obtenidos mediante diversos programas, como MINITAB, SAS y S-Plus.

Ejemplo 1.1 La tragedia que ocurrió al transbordador espacial *Challenger* y a sus astronautas, en 1986, dio origen a varios estudios para investigar las razones del fracaso. La atención se dirigió rápidamente al comportamiento de los anillos O del motor del cohete. Los datos de las observaciones para x = temperatura del anillo O (°F) en cada ignición de prueba o de despegue real del motor cohete del transbordador (*Presidential Commission on the Space Shuttle Challenger Accident,* vol. 1, 1986: 129-131) son los siguientes:

84	49	61	40	83	67	45	66	70	69	80	58
68	60	67	72	73	70	57	63	70	78	52	67
53	67	75	61	70	81	76	79	75	76	58	31

Sin organizarlos, es muy difícil captar lo que podría ser una temperatura normal o representativa, si los valores están muy concentrados en torno a un valor típico, o si están muy dispersos, o si faltan datos, o qué porcentaje de los valores caen entre 60 y 70, etcétera. En la figura 1.1 se muestra un *diagrama de tallo y hojas* de los datos, y también un *histograma*. Se explica en forma breve la construcción y la interpretación de estas representaciones concisas; por el momento, el autor espera que el lector aprecie cómo se comienza a indicar la distribución de los valores de temperatura a lo largo de la escala de medición. Algunos de estos lanzamientos o igniciones tuvieron éxito, y otros no. En el capítulo 13 se investiga si la temperatura influyó en la posibilidad de un buen lanzamiento.

```
Stem-and-leaf of temp N = 36
Leaf Unit = 1.0
 1   3   1
 1   3
 2   4   0
 4   4   59
 6   5   23
 9   5   788
13   6   0113
(7)  6   6777789
16   7   000023
10   7   556689
 4   8   0134
```

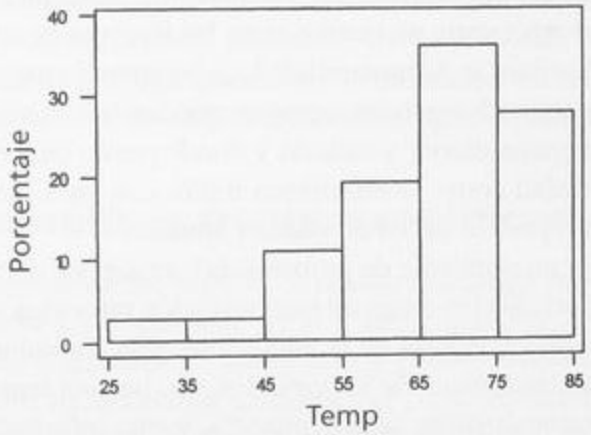


Figura 1.1 Diagrama de tallo y hojas con MINITAB e histograma de los datos de temperatura de los anillos O ■

Una vez obtenida la muestra de una población, el investigador querrá usar la información de la muestra para llegar a algún tipo de conclusión (hacer una inferencia de cierto tipo) acerca de la población. Es decir, la muestra es un medio hacia un fin, y no un fin. Las técnicas para hacer una generalización en toda la población a partir de una muestra se ubican dentro de la rama de la estadística llamada **estadística inferencial.**

Ejemplo 1.2 Las investigaciones sobre resistencia de materiales representan un área rica en aplicaciones de métodos estadísticos. El artículo "Effects of Aggregates and Microfillers on the Flexural Properties of Concrete" *(Magazine of Concrete Research,* 1997: 81-98) reseña un estudio de las propiedades de resistencia del concreto de alta resistencia obtenido con superplastificantes y ciertos aglutinantes. La resistencia a la compresión de ese concreto ya se había investigado, pero se desconocía mucho acerca de la resistencia a la flexión (medida de la capacidad de resistencia a la falla por flexión). Los datos siguientes sobre resistencia a la flexión están en megapascales, MPa (donde 1 Pa (pascal) = 1.45×10^{-4} psi) y aparecieron en el artículo citado:

5.9	7.2	7.3	6.3	8.1	6.8	7.0	7.6	6.8	6.5	7.0	6.3	7.9	9.0
8.2	8.7	7.8	9.7	7.4	7.7	9.7	7.8	7.7	11.6	11.3	11.8	10.7	

Supóngase que se desea una *estimación* del valor promedio de la resistencia a la flexión de todas las vigas que se podrían fabricar con este método (si se conceptualiza una población de todas esas vigas, se está tratando de estimar la media poblacional). Se puede demostrar que, con un alto grado de confianza, la resistencia promedio de la población está entre 7.48 MPa y 8.80 MPa; a este valor se le llama *intervalo de confianza* o *estimación del intervalo.* Otra posibili-

dad es que se podrían usar estos datos para predecir la resistencia a la flexión de *una sola* viga de este tipo. Con un alto grado de confianza, la resistencia de esa viga será mayor que 7.35 MPa; el número 7.35 recibe el nombre de *límite inferior de predicción.* ■

El objetivo fundamental de esta obra es presentar e ilustrar métodos de inferencia estadística que son útiles en trabajos científicos. Los tipos más importantes de procedimientos inferenciales, estimación puntual, prueba de hipótesis y estimación por intervalos de confianza, se tratan en los capítulos 6 a 8, y luego se utilizan en entornos más complicados en los capítulos 9 a 16. El resto de este capítulo se dedica a métodos de estadística descriptiva que se emplean más en el desarrollo de la inferencia.

En los capítulos 2 a 5 se encuentra material de la disciplina de la probabilidad. Este material representa un puente entre las técnicas descriptiva e inferencial. Cuando se logra un cierto dominio de la probabilidad, se comprende mejor cómo se desarrollan y utilizan los procedimientos inferenciales, cómo se pueden traducir e interpretar las conclusiones estadísticas en el lenguaje diario, y cuándo y dónde puede haber errores al aplicar los métodos. Tanto la probabilidad como la estadística tratan con cuestiones en donde intervienen poblaciones y muestras, pero lo hacen de manera opuesta.

En un problema de probabilidad, se supone que se conocen las propiedades de la población en estudio (es decir, en una población numérica se podría suponer alguna distribución específica de los valores de la población), y se formulan y responden preguntas en relación con una muestra tomada de la población. En un problema estadístico, el experimentador dispone de las características de una muestra, y esta información lo capacita para sacar conclusiones respecto a la población. La relación entre las dos disciplinas se puede resumir diciendo que el razonamiento en la probabilidad va de la población a la muestra (razonamiento deductivo), y en la estadística inferencial se parte de la muestra para llegar a la población (razonamiento inductivo). Este concepto se ilustra en la figura 1.2.

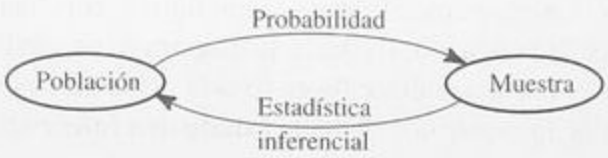


Figura 1.2 Relación entre probabilidad y estadística inferencial

Antes de comprender lo que indica una muestra particular acerca de la población, primero es necesario entender la incertidumbre que tiene relación con el hecho de tomar una muestra de una determinada población. Ésta es la razón de estudiar primero probabilidad y después estadística.

Como ejemplo del enfoque tan distinto de la probabilidad y la estadística inferencial, considérese a los conductores que usan cinturones de regazo manuales en automóviles equipados con sistemas automáticos de cinturones diagonales. (En el artículo, "Automobile Seat Belts: Usage Patterns in Automatic Belt Systems", *Human Factors,* 1998: 126-135, se resumen los datos de uso de estos cinturones.) En probabilidad se podría suponer que 50% de los conductores de automóviles con cinturones de seguridad en una área metropolitana determinada usan con regularidad sus cinturones de regazo (es una hipótesis sobre la población), así que se podría preguntar: "¿qué tan probable es que en una muestra de 100 conductores, al menos 70 de ellos usen su cinturón de regazo?", o bien, "En una muestra de tamaño 100, ¿cuántos de los conductores cabe esperar que usen con regularidad su cinturón de regazo?" Por otro lado, en estadística in-

ferencial, se cuenta con información muestral; por ejemplo, en una muestra de 100 conductores se encontró que 65 usan con regularidad su cinturón de regazo. Se podría preguntar entonces: "¿se tienen pruebas sustanciales para concluir que más de 50% de los conductores de esta zona usan con regularidad su cinturón de regazo?". En este último escenario, se pretende usar la información muestral para dar respuesta a una pregunta acerca de la estructura de la población total de la cual se seleccionó la muestra.

En el ejemplo del cinturón de regazo, se tiene una población bien definida y concreta: los conductores de automóviles equipados de cierta manera en una zona metropolitana determinada. Sin embargo, en el ejemplo 1 se dispone de una muestra de temperaturas de anillos O, pero proviene de una población que no existe en realidad. Por el contrario, es conveniente considerar que la población consiste en todas las mediciones de temperatura posibles que se podrían realizar en condiciones experimentales similares. A esta población se le denomina **población conceptual** o **hipotética**. Hay varias situaciones problemáticas en las que las preguntas se ajustan al marco de la estadística inferencial al conceptualizar una población.

Estudios enumerativos frente a analíticos

W. E. Deming, un estadounidense experto en estadística, quien promovió la revolución de la calidad en Japón durante las décadas de 1950 y 1960, introdujo la distinción entre *estudios enumerativos* y *estudios analíticos*. En los primeros, el interés se centra en un conjunto, finito, identificable e invariable, de individuos u objetos que constituyen una población. Un *marco de muestreo*, es decir, una lista de individuos u objetos con los que se lleva a cabo el muestreo, está disponible para un investigador o bien se puede construir. Por ejemplo, el marco podría estar formado por todas las firmas en relación con una petición a fin de calificar una cierta iniciativa para el voto en una elección próxima; por lo común, se elige una muestra para constatar si la cantidad de firmas *válidas* es mayor que un valor especificado. Otro ejemplo es la posibilidad de que el marco contenga números de serie de los hornos que fabrica una empresa durante cierto tiempo, y se puede seleccionar una muestra para inferir algo acerca de la duración promedio de estas unidades. El uso de métodos inferenciales que se desarrollan en este texto se acepta de forma razonable en situaciones de este tipo (aunque los expertos en estadística aún pueden argumentar sobre qué métodos específicos se deben usar).

Un estudio analítico se define en sentido amplio como uno que por naturaleza no es enumerativo. Estos estudios suelen llevarse a cabo con el fin de mejorar un producto futuro al emprender acciones en un proceso de determinado tipo (p. ej., calibrar de nuevo el equipo o ajustar el nivel de alguna entrada, como podría ser la cantidad de un catalizador). Es común que los datos se obtengan sólo de un proceso existente, uno que pudiera diferir en aspectos importantes respecto al proceso futuro. Así que no hay marco de muestreo que enumere los individuos u objetos de interés. Por ejemplo, una muestra de cinco turbinas con un nuevo diseño, se podría fabricar y probar de forma experimental a fin de investigar el desempeño. Estas cinco turbinas pueden ser consideradas como una muestra de la población conceptual de los prototipos que se fabricarían en condiciones similares, pero *no* necesariamente como representativas de la población de unidades fabricadas una vez iniciada la producción regular. Los métodos en los que se usa información muestral para sacar conclusiones en relación con una producción futura de unidades, pueden ser problemáticos. Se debe consultar a un experto en el área de diseño e ingeniería de turbinas (o del área pertinente) para que determine si la extrapolación es razonable. Una buena exposición de estas cuestiones se puede leer en el artículo "Assumptions for Statistical Inference", de Gerald Hahn y William Meeker (*The American Statistician*, 1993: 1-11).

Recopilación de datos

La estadística está relacionada no sólo con la organización y análisis de datos una vez que han sido recopilados, sino también con el desarrollo de técnicas para reunirlos. Si los datos no se reúnen de manera apropiada, es posible que un investigador no pueda contestar con un grado de confianza razonable las preguntas en consideración. Un problema común es que la población objetivo, es decir, aquella de la que se sacarán conclusiones, puede ser diferente de la que se muestrea en realidad. Por ejemplo, a los anunciantes les gustarían distintos tipos de información acerca de los hábitos de ver televisión de posibles clientes. La información más sistemática de esta clase proviene de colocar dispositivos de observación en un pequeño número de casas repartidas por todo el país. Se ha conjeturado que la instalación de estos aparatos per se altera el hábito de ver televisión, así que las características de la muestra podrían ser diferentes de las de la población objetivo.

Cuando la recopilación de datos conlleva a seleccionar individuos u objetos de un marco, el método más sencillo para asegurar una selección representativa es tomar una *muestra aleatoria simple*. Es decir, una en la que cualquier subconjunto particular de tamaño específico (p. ej., una muestra de tamaño 100) tiene la misma probabilidad de ser elegida. Por ejemplo, si el marco consiste en 1 000 000 de números en serie, los números 1, 2, . . . , hasta el número 1 000 000 se podrían escribir en tiras de papel idénticas. Después de colocar estas tiras en una caja y revolverlas bien, se podrían sacar una por una hasta obtener el tamaño de muestra requerido. Otra posibilidad más recomendable es que se pudiera usar una tabla de números aleatorios o un generador de números aleatorios de computadora.

A veces se pueden emplear otros métodos de muestreo para facilitar el proceso de selección, obtener información extra o aumentar el grado de confianza en las conclusiones. En uno de estos métodos, *muestreo estratificado*, es necesario separar las unidades de la población en grupos que no se traslapen y tomar una muestra de cada uno. Por ejemplo, es posible que un fabricante de videograbadoras desee información acerca de la satisfacción del cliente en relación con las unidades producidas durante el año anterior. Si se fabricaron y vendieron tres modelos diferentes, se podría seleccionar una muestra de cada uno de los tres estratos correspondientes. Esto produciría información de los tres modelos, y se tendría la seguridad de que ningún modelo tuvo una representación mayor o menor en toda la muestra.

Es común obtener una muestra "de conveniencia" al seleccionar individuos u objetos sin aleatorización sistemática. Por ejemplo, se podría apilar un conjunto de ladrillos de modo que se dificulte elegir los que están en el centro. Si los ladrillos de la pila que están arriba y a los lados fueran un poco diferentes de los demás, los datos de la muestra resultante no serían representativos de la población. Por lo común, un investigador supone que una muestra de conveniencia se aproxima a una muestra aleatoria, en cuyo caso se puede usar el repertorio de métodos inferenciales del especialista en estadística; sin embargo, ésta es una decisión basada en la experiencia. La mayor parte de los métodos estudiados aquí se basan en una variación del muestreo aleatorio simple descrito en el capítulo 5.

Los ingenieros y científicos suelen reunir datos al llevar a cabo cierta clase de experimento diseñado. Esto podría tener que ver con decidir cómo asignar varios tratamientos diferentes (p. ej., fertilizantes o capas para evitar la corrosión) a las diversas unidades experimentales (parcelas o tramos de tubo). Por otro lado, un investigador puede variar de modo sistemático los niveles o categorías de ciertos factores (p. ej., la presión o el tipo de material aislante) y observar el efecto en alguna variable de respuesta (como el rendimiento de un proceso de producción).

Ejemplo 1.3 En un artículo del *New York Times* (27 de enero de 1987) se informó que es posible reducir el riesgo de sufrir un infarto si se toma aspirina. Esta conclusión se basa en un experimento diseñado en el que participaron un grupo de control, con individuos que tomaron un placebo con apariencia de aspirina pero inerte, y un grupo de tratamiento que tomó aspirina según un régimen especificado. Los individuos fueron asignados al azar a los grupos a fin de evitar sesgos, y de esta manera poder usar métodos basados en la probabilidad para analizar los datos. De los 11 034 individuos del grupo de control, 189 experimentaron más tarde ataques al corazón, mientras que sólo 104 de los 11 037 del grupo de tratamiento tuvieron un ataque al corazón. La incidencia de ataques al corazón en el grupo de tratamiento fue sólo la mitad de la del grupo de control. Una explicación posible de este resultado es la variación de probabilidad, que en realidad la aspirina no tiene el efecto deseado y la diferencia observada es sólo la variación ordinaria en la misma forma que lanzar dos monedas idénticas normalmente produciría diferentes números de caras. Sin embargo, en este caso, los métodos inferenciales hacen pensar que la variación de probabilidad por sí misma no explica de manera adecuada la magnitud de la diferencia observada. ■

Ejemplo 1.4 Un ingeniero desea investigar los efectos del tipo de adhesivo y el material conductor en la resistencia de unión al montar un circuito integrado (CI) en un determinado sustrato. Están en consideración dos tipos de adhesivo y dos materiales conductores. Se hacen dos observaciones para cada combinación de tipo de adhesivo y material conductor, y se obtienen los datos siguientes:

Tipo de adhesivo	Material conductor	Resistencia de la unión	Promedio
1	1	82, 77	79.5
1	2	75, 87	81.0
2	1	84, 80	82.0
2	2	78, 90	84.0

Las resistencias de unión promedio resultantes se ilustran en la figura 1.3. Al parecer el adhesivo tipo 2 mejora la resistencia de unión en comparación con el tipo 1 en casi la misma cantidad, sin importar cuál material conductor se utilice, siendo la mejor combinación la 2, 2. De nuevo se pueden utilizar los métodos inferenciales para juzgar si estos efectos son reales o se deben simplemente a la variación de probabilidad.

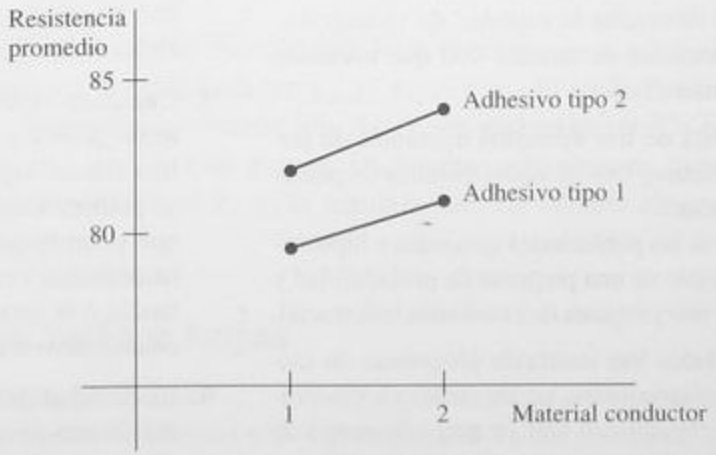


Figura 1.3 Resistencias de unión promedio del ejemplo 1.4

Además, suponga que hay en consideración dos tiempos de cura y también dos tipos de recubrimiento posterior del CI. Entonces hay $2 \cdot 2 \cdot 2 \cdot 2 = 16$ combinaciones de estos cuatro factores, y es posible que los recursos del ingeniero sean suficientes para hacer incluso una sola observación de cada una de estas combinaciones. En el capítulo 11, se estudia cómo la selección cuidadosa de una fracción de estas posibilidades, produce por lo regular la información deseada. ■

Ejercicios | Sección 1.1 (1 a 9)

1. Determine una posible muestra de tamaño 4 de cada una de las poblaciones siguientes:
   a. Los periódicos publicados en Estados Unidos.
   b. Las empresas listadas en la Bolsa de valores de Nueva York.
   c. Todos los estudiantes de su universidad.
   d. Los promedios de calificaciones de los alumnos de su universidad.

2. Para cada una de las poblaciones hipotéticas siguientes, determine una muestra plausible de tamaño:
   a. Todas las distancias que pudieran resultar al lanzar un balón de futbol.
   b. Número de páginas de los libros publicados en cinco años a partir de ahora.
   c. Las mediciones posibles de la intensidad de temblores (escala Richter) que pudieran registrarse en California durante el año próximo.
   d. Los rendimientos posibles, en gramos, de cierta reacción efectuada en un laboratorio.

3. Considere la población que consta de todas las videograbadoras de determinada marca y modelo, y trate el caso de si alguna necesita servicio mientras funciona la garantía.
   a. Redacte varias preguntas de probabilidad basadas en el hecho de seleccionar una muestra de 100 de esas videograbadoras.
   b. ¿Qué pregunta de la estadística inferencial se podría contestar si se determina la cantidad de videograbadoras en una muestra de tamaño 100 que necesiten servicio de garantía?

4. a. Elabore una lista de tres ejemplos diferentes de poblaciones concretas y tres ejemplos distintos de poblaciones hipotéticas.
   b. Para cada una de las poblaciones concretas e hipotéticas, dé un ejemplo de una pregunta de probabilidad y un ejemplo de una pregunta de estadística inferencial.

5. Muchas universidades han instituido programas de instrucción (PI) complementarios, en los cuales un moderador se reúne con regularidad con un pequeño grupo de alumnos inscritos en el curso para promover discusiones sobre el material del curso y aumentar el dominio del tema. Suponga que los alumnos de un grupo grande de estadística se distribuyen al azar en un grupo de control que no participa en el PI, y otro de tratamiento que sí participa. Al final del periodo escolar, se determina la calificación total de cada alumno en el curso.
   a. ¿Son las calificaciones del grupo PI una muestra de una población existente? En caso afirmativo, ¿cuál es? En caso contrario, ¿cuál es la población conceptual pertinente?
   b. ¿Cuál cree que sea la ventaja de dividir al azar a los alumnos en los dos grupos, en vez de dejar que cada alumno escoja a cuál grupo unirse?
   c. ¿Por qué los investigadores no colocaron a todos los alumnos en el grupo de tratamiento? *Nota:* En el artículo "Supplemental Instruction: An Effective Component of Student Affairs Programming" (*J. of College Student Devel.*, 1997: 577-586) se describe el análisis de datos generados de varios programas PI.

6. El sistema de la California State University (CSU) consta de 23 *campus*, desde San Diego State en el sur hasta Humboldt State, cerca de los límites con Oregon. Un administrador de la CSU desea hacer una inferencia acerca de la distancia promedio entre las ciudades de origen de los estudiantes y sus áreas de estudio. Describa y analice varios métodos de muestreo que puedan ser útiles. ¿Sería éste un estudio enumerativo o analítico? Explique su razonamiento.

7. Una cierta ciudad se divide en forma natural en diez distritos. ¿Cómo podría un perito evaluador de bienes inmuebles seleccionar una muestra de casas unifamiliares que se pudiera usar como base para obtener una ecuación que permita predecir el valor estimado a partir de características, como edad, tamaño, número de baños, distancia a la escuela más cercana, etc? ¿Este estudio es enumerativo o analítico?

8. La cantidad de flujo que pasa por una válvula solenoide del sistema de control de contaminantes de un automóvil, es una característica importante. Se llevó a cabo un expe-

rimento para estudiar cómo el flujo depende de tres factores: longitud de la armadura, carga del resorte y profundidad de la bobina. Se seleccionaron dos niveles (bajo y alto) de cada factor, y se hizo una sola observación del flujo por cada combinación de niveles.

a. ¿De cuántas observaciones consistió el conjunto de datos resultante?

b. ¿Es un estudio enumerativo o analítico? Explique su razonamiento.

9. En un famoso experimento realizado en 1882, Michelson y Newcomb obtuvieron 66 observaciones en el tiempo que tardó la luz en viajar entre dos lugares de Washington, D. C. Algunas de las mediciones (codificadas en cierto modo) fueron 31, 23, 32, 36, −2, 26, 27 y 31.

a. ¿Por qué no son idénticas estas mediciones?

b. ¿Se trata de un estudio enumerativo? ¿Por qué sí o por qué no?

1.2 Métodos gráfico y tabular en estadística descriptiva

La estadística descriptiva se divide en dos áreas. En esta sección se estudia la primera de ellas: representar un conjunto de datos por medio de técnicas visuales. En las secciones 1.3 y 1.4, se elaboran algunas medidas de resumen numérico para conjuntos de datos. Es posible que muchas técnicas visuales ya sean conocidas del estudiante: tablas de frecuencia, hojas de registro, histogramas, gráficas circulares, gráficas de barras, diagramas de dispersión y otras. Aquí se tratan algunas de las técnicas que son de mayor utilidad y pertinentes para la probabilidad y la estadística inferencial.

Notación:

Cierta notación general facilita la aplicación de métodos y fórmulas a una amplia variedad de problemas prácticos. El número de observaciones de una sola muestra, es decir, el *tamaño de muestra*, se denota por n, así que $n = 4$ para la muestra de universidades {Stanford, Iowa State, Wyoming, Rochester} y también para la muestra de mediciones de pH {6.3, 6.2, 5.9, 6.5}. Si se consideran dos muestras al mismo tiempo, se puede usar ya sea m y n o n_1 y n_2 para denotar los números de observaciones. De este modo, si {29.7, 31.6, 30.9} y {28.7, 29.5, 29.4, 30.3} son mediciones de rendimiento térmico para dos motores diesel diferentes, entonces $m = 3$ y $n = 4$.

Dado un conjunto de datos que consiste en n observaciones sobre una variable x, cada una de las observaciones se representa por $x_1, x_2, x_3, \ldots, x_n$. El subíndice no tiene relación con la magnitud de una determinada observación. Así, x_1 no será en general la observación más pequeña del conjunto, ni x_n será la más grande. En muchas aplicaciones, la primera observación que obtenga el experimentador es x_1, x_2 la segunda, etc. La i-ésima observación del conjunto de datos se denota por x_i.

Diagrama de tallo y hojas

Considere un conjunto de datos $x_1, x_2, \ldots, x_n$ para el cual cada x_i consta de por lo menos dos dígitos. Una forma rápida de obtener una representación visual ilustrativa del conjunto de datos es construir un *diagrama de tallo y hojas*.

Pasos para construir un diagrama de tallo y hojas

1. Seleccione uno o más dígitos directores para los valores del tallo. Los dígitos finales se convierten en las hojas.
2. Liste los posibles valores del tallo en una columna vertical.
3. Registre la hoja por cada observación junto al valor correspondiente del tallo.
4. Indique las unidades para tallos y hojas en algún lugar del diagrama.

Si el conjunto de datos consiste en las puntuaciones de los exámenes, cada una entre 0 y 100, la puntuación 83 tendría un tallo de 8 y una hoja de 3. Para un conjunto de datos de rendimientos de combustible en automóviles (millas por galón), todos entre 8.1 y 47.8, se podría usar el dígito de las decenas como el tallo, así que 32.6 tendría entonces una hoja de 2.6. En general, se recomienda un diagrama basado en tallos entre 5 y 20.

Ejemplo 1.5 La cantidad de alcohol que consumen los alumnos universitarios preocupa mucho no sólo a la comunidad académica, sino también a toda la sociedad debido a las posibles consecuencias de salud y seguridad. En el artículo "Health and Behavioral Consequences of Binge Drinking in College" (*J. of the Amer. Med. Assoc.*, 1994: 1672-1677), se describe un estudio detallado del alto consumo de alcohol en las universidades de Estados Unidos. Para los varones se definió una borrachera como cinco o más copas seguidas, y cuatro o más para las mujeres. En la figura 1.4 se muestra un diagrama de tallo y hojas de 140 valores de x = porcentaje de alumnos que consumen alcohol. (Estos valores no aparecen en el artículo citado, pero el diagrama concuerda con un cuadro de los datos que sí apareció).

La primera hoja del renglón 2 del tallo es 1, lo cual indica que 21% de los alumnos de una de las universidades de la muestra ingerían alcohol. Sin la identificación de los dígitos del tallo y de la hoja, no se podría saber si la observación del tallo 2, hoja 1 se debe leer como 21%, 2.1%, o 0.21%.

```
0 | 4
1 | 1345678889
2 | 1223456666777889999
3 | 011223334455566667777788899999
4 | 111222223344445566666677788888999
5 | 00111222233455666667777888899
6 | 01111244455666778
```

Tallo: dígito de las decenas
Hoja: dígito de las unidades

Figura 1.4 Diagrama de tallo y hojas para el porcentaje de bebedores en cada una de 140 universidades

Cuando se traza a mano un diagrama puede requerir mucho tiempo ordenar las hojas de la más pequeña a la más grande en cada renglón y, por lo general, este ordenamiento aporta poca información adicional. Suponga que las observaciones se listan en orden alfabético por nombre de escuela, como

16% 33% 64% 37% 31% . . .

Entonces el hecho de colocar en este orden estos valores en el diagrama produciría el renglón del tallo 1 con 6 como su primera hoja, y el comienzo del renglón del tallo 3 sería

3 | 371 . . .

El diagrama indica que hay un valor característico o representativo en el renglón del tallo 4, quizá a la mitad del intervalo de 40%. Las observaciones no se acumulan mucho respecto a este valor característico, como sería el caso si todos los valores estuvieran entre 20 y 49%. El diagrama crece hasta un solo pico al ir bajando, y luego disminuye; no hay huecos en él. La forma del diagrama no es del todo simétrica, sino que al parecer se extiende un poco más en la dirección de las hojas bajas que en la de las hojas altas. Por último, no hay observaciones que estén muy alejadas del grueso de los datos (*ningún valor atípico*), como sería el caso si en vez de uno de los valores de 26% hubiera sido de 86%. La particularidad más sorprendente de estos datos es que, en la mayor parte de las universidades de la muestra, al menos una cuarta parte de los alumnos son bebedores. El problema del alto consumo de alcohol en las escuelas está más extendido de lo que se piensa. ■

Un diagrama de tallo y hojas transmite información acerca de los siguientes aspectos de los datos:

- Identificación de un valor característico o representativo.
- Grado de dispersión respecto al valor característico.
- Presencia de algún hueco en los datos.
- Grado de simetría en la distribución de los valores.
- Cantidad y ubicación de picos.
- Presencia de cualquier valor atípico.

Ejemplo 1.6 En la figura 1. 5 se muestran diagramas de tallo y hojas para una muestra aleatoria de trayectos en yardas de canchas de golf que, según la designación de *Golf Magazine,* se encuentran entre las más difíciles de Estados Unidos. Entre la muestra de 40 trayectos en yardas, el más corto mide 6 433 yardas de largo, y el más largo mide 7 280 yardas. Al parecer, las distancias están distribuidas de manera casi uniforme (en yardas) en el intervalo de valores de la muestra. Observe que, en este caso, una selección de tallo, ya sea de un solo dígito (6 o 7) o tres dígitos (643, . . . , 728) proporcionaría un diagrama con poca información, primero por ser muy pocos tallos y segundo por ser demasiados.

Tallo: dígitos de unidades de millar y centenas
Hoja: dígitos de decenas y unidades

64	35	64	33	70			
65	26	27	06	83			
66	05	94	14				
67	90	70	00	98	70	45	13
68	90	70	73	50			
69	00	27	36	04			
70	51	05	11	40	50	22	
71	31	69	68	05	13	65	
72	80	09					

(a)

```
Stem-and-leaf of yardage  N = 40
Leaf Unit = 10
     4          64  3367
     8          65  0228
    11          66  019
    18          67  0147799
   (4)          68  5779
    18          69  0023
    14          70  012455
     8          71  013666
     2          72  08
```

(b)

Figura 1.5 Diagramas de tallo y hojas de longitudes de campos de golf, en yardas: (a) hojas de dos dígitos; (b) diagrama de MINITAB, con hojas truncadas de un dígito ■

Un diagrama de tallo y hojas no muestra el orden en que se obtuvieron las observaciones, y es posible que oculte información importante acerca del mecanismo que generó los datos. Por ejemplo, durante un determinado periodo, el ancho de corte de una fresadora podría tender a aumentar en relación con el valor objetivo, lo que indica que se tiene un proceso "fuera de control". Un *diagrama de puntos y dígitos* combina una representación de las observaciones con el tiempo y un diagrama de tallo y hojas.

Ejemplo 1.7 Cada observación en el diagrama de tallo y hojas de la figura 1.6 es el valor de la producción de cerveza en Estados Unidos (millones de barriles) para un trimestre diferente durante el periodo 1975-1982. En el diagrama se usan *tallos repetidos*; por ejemplo, el renglón 4L es para observaciones con una hoja "baja": 0, 1, 2, 3 o 4, y las observaciones con hojas altas se ponen en el renglón 4H. La *gráfica de series de tiempo* de la derecha muestra una tendencia creciente con el tiempo y también mayor producción en los trimestres segundo y tercero de cualquier año, que en los otros dos trimestres (efecto *estacional*).

Tallo: dígito de las decenas
Hoja: dígito de las unidades

5	5H
242330	5L
7688996	4H
21421414444	4L
9696656	3H

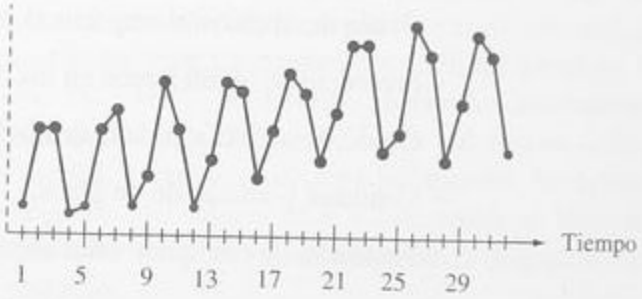


Figura 1.6 Diagrama de puntos y dígitos de la producción de cerveza en Estados Unidos ■

Diagramas de puntos

Un diagrama de puntos es un resumen atractivo de datos numéricos cuando el conjunto de datos es más o menos pequeño o hay relativamente pocos valores de datos distintos. Cada observación se representa por un punto arriba del lugar correspondiente en una escala de medición horizontal. Cuando un valor se presenta más de una vez, hay un punto para cada ocurrencia, y esos puntos se apilan verticalmente. Al igual que con el diagrama de tallo y hojas, un diagrama de puntos proporciona información acerca de la ubicación, dispersión, valores extremos y huecos.

Ejemplo 1.8 En la figura 1.7 se ilustra un diagrama de puntos para los datos de temperatura de los anillos O que se describieron en el ejemplo 1.1 de la sección anterior. Un valor de temperatura representativo es el que está a la mitad de los 60-70 (°F), y hay bastante dispersión de datos respecto al centro. Los datos se extienden más en el extremo inferior que en el extremo superior, y la observación más pequeña, 31, se puede describir de manera clara como valor atípico.

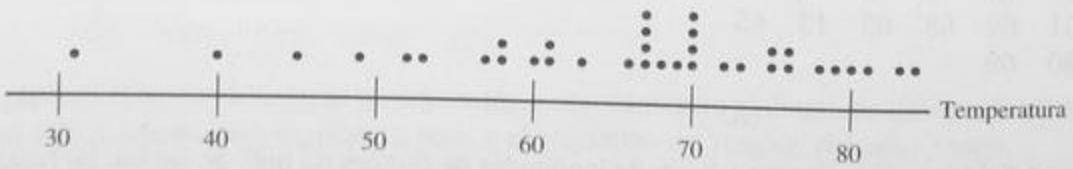


Figura 1.7 Diagrama de puntos de los datos de temperatura de los anillos O (°F) ■

Si el conjunto de datos analizado en el ejemplo 1.8 hubiera consistido en 50 o 100 observaciones de temperatura, cada una con exactitud de una décima de grado, hubiera sido más tedioso trazar un diagrama de puntos. La técnica siguiente es muy adecuada para esos casos.

Histogramas

Algunos datos numéricos se obtienen mediante conteo para determinar el valor de una variable (la cantidad de multas de tránsito que una persona recibió el año pasado, el número de personas que llegan para ser atendidas durante un periodo determinado), mientras que otros datos se obtienen tomando medidas (el peso de un individuo, el tiempo de reacción a un estímulo particular). En general, es distinta la técnica para trazar el histograma en estos dos casos.

DEFINICIÓN

Una variable es **discreta** si su conjunto de valores posibles es finito o se puede enumerar en una sucesión infinita (una en la que hay un primer número, un segundo número, etcétera). Una variable es **continua** si sus valores posibles consisten en un intervalo completo en la recta numérica.

Una variable discreta x casi siempre resulta de contar, en cuyo caso los valores posibles son 0, 1, 2, 3, . . . o algún subconjunto de estos enteros. Las variables continuas surgen de hacer mediciones. Por ejemplo, si x es el pH de una sustancia química, entonces en teoría x podría ser cualquier número entre 0 y 14: 7.0. 7.03, 7.032, etc. Claro que en la práctica hay limitaciones en el grado de exactitud de cualquier instrumento de medición, así que es posible que no se determine el pH, tiempo de reacción, altura y concentración hasta un número grande, asignado de manera arbitraria, de cifras decimales. Sin embargo, desde el punto de vista de crear modelos matemáticos para distribuciones de datos, es útil imaginar un continuo completo de valores posibles.

Considere los datos que consisten en observaciones sobre una variable discreta x. La **frecuencia** de cualquier valor particular de x es la cantidad de veces que se presenta ese valor en el conjunto de datos. La **frecuencia relativa** de un valor es la fracción o proporción de las veces que se presenta ese valor:

$$\text{frecuencia relativa de un valor} = \frac{\text{número de veces que aparece el valor}}{\text{número de observaciones en el conjunto de datos}}$$

Suponga, por ejemplo, que el conjunto de datos consiste en 200 observaciones en x = cantidad de defectos principales de un automóvil nuevo de cierto tipo. Si 70 de esos valores son 1, entonces

$$\text{frecuencia del valor 1 de } x\text{: } 70$$

$$\text{frecuencia relativa del valor 1 de } x\text{: } \frac{70}{200} = 0.35$$

Al multiplicar por 100 una frecuencia relativa, se obtiene un porcentaje; en el ejemplo de los defectos, 35% de los automóviles de la muestra sólo tuvo un defecto importante. Las frecuencias relativas, o los porcentajes, suelen ser de más interés que las frecuencias. En teoría, la suma de las frecuencias relativas debe ser igual a la unidad, pero en la práctica podría ser un poco diferente de uno, debido al redondeo. Una **distribución de frecuencias** es una tabla de las frecuencias o de las frecuencias relativas, o ambas.

Construcción de un histograma para datos discretos

Primero, determine la frecuencia y la frecuencia relativa de cada valor de x. Luego, marque los valores posibles de x en una escala horizontal. Arriba de cada valor trace un rectángulo cuya altura sea la frecuencia relativa (o como alternativa, la frecuencia) de ese valor.

Con esta construcción se asegura que el *área* de cada rectángulo sea proporcional a la frecuencia relativa del valor. Así, si las frecuencias relativas de $x = 1$ y $x = 5$ son 0.35 0.07, respectivamente, entonces el área del rectángulo arriba de 1 es cinco veces mayor que el área del rectángulo arriba de 5.

Ejemplo 1.9 ¿Qué tan inusual es que se dé un juego sin *hits* o de un solo *hit* en la liga mayor de béisbol, y con qué frecuencia un equipo produce más de 10, 15 o incluso 20 *hits*? La tabla 1.1 es una distribución de frecuencias de la cantidad de *hits* por equipo y por juego para los juegos de nueve entradas que se realizaron entre 1989 y 1993.

Tabla 1.1 Distribución de frecuencias para *hits* en juegos de nueve entradas

Hits/Juego	Número de juegos	Frecuencia relativa	*Hits*/Juego	Número de juegos	Frecuencia relativa
0	20	0.0010	14	569	0.0294
1	72	0.0037	15	393	0.0203
2	209	0.0108	16	253	0.0131
3	527	0.0272	17	171	0.0088
4	1048	0.0541	18	97	0.0050
5	1457	0.0752	19	53	0.0027
6	1988	0.1026	20	31	0.0016
7	2256	0.1164	21	19	0.0010
8	2403	0.1240	22	13	0.0007
9	2256	0.1164	23	5	0.0003
10	1967	0.1015	24	1	0.0001
11	1509	0.0779	25	0	0.0000
12	1230	0.0635	26	1	0.0001
13	834	0.0430	27	1	0.0001
				19 383	1.0005

El histograma correspondiente de la figura 1.8 sube con bastante uniformidad hasta un máximo y después baja. Se extiende un poco más a la derecha (hacia valores grandes) que a la izquierda; un ligero "sesgo positivo".

Ya sea de la información tabulada o del histograma, se determina lo siguiente:

$$\begin{aligned}\text{proporción de juegos con a lo sumo dos } \textit{hits} &= \begin{matrix}\text{frecuencia}\\ \text{relativa}\\ \text{para } x = 0\end{matrix} + \begin{matrix}\text{frecuencia}\\ \text{relativa}\\ \text{para } x = 1\end{matrix} + \begin{matrix}\text{frecuencia}\\ \text{relativa}\\ \text{para } x = 2\end{matrix}\\ &= 0.0010 + 0.0037 + 0.0108 = 0.0155\end{aligned}$$

De manera similar,

$$\begin{pmatrix}\text{proporción de juegos con}\\ \text{entre 5 y 10 } \textit{hits} \text{ (inclusive)}\end{pmatrix} = 0.0752 + 0.1026 + \cdots + 0.1015 = 0.6361$$

Es decir, alrededor de 64% de todos esos juegos produjeron entre 5 y 10 *hits* (inclusive).

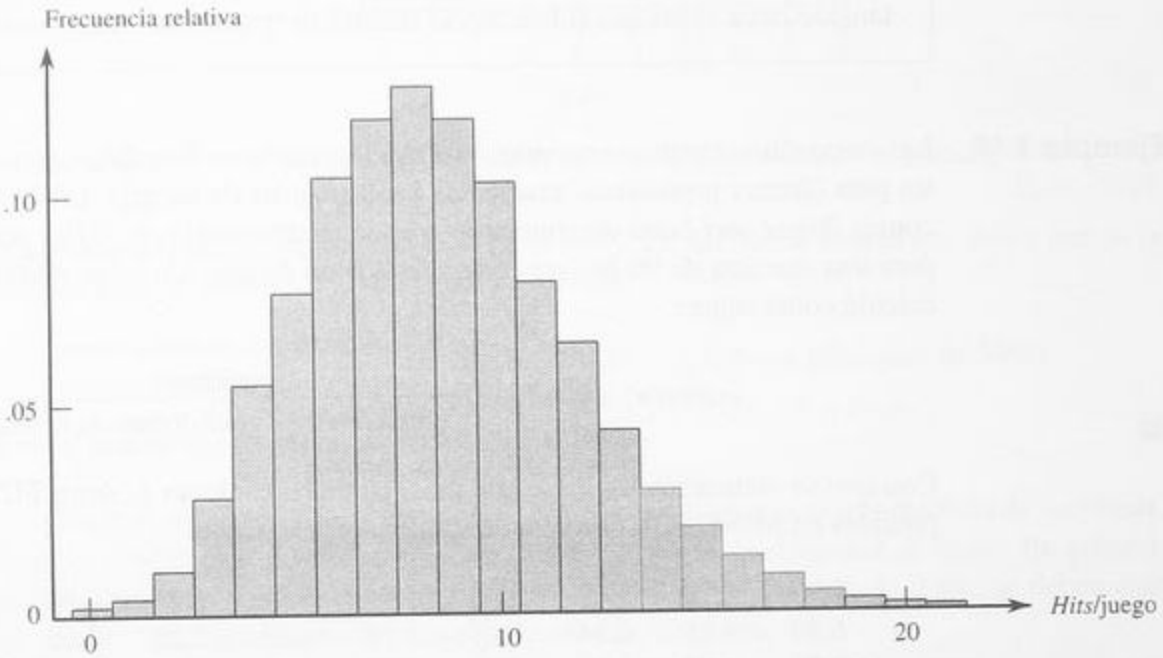
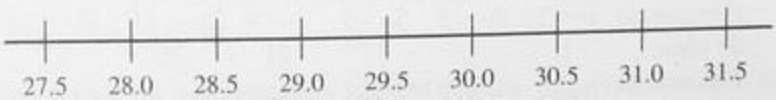


Figura 1.8 Histograma del número de *hits* por juego de nueve entradas ■

Construir un histograma para datos (mediciones) continuos, requiere subdividir el eje de las mediciones en una cantidad adecuada de **intervalos de clase** o **clases**, de modo que cada observación esté contenida en exactamente una clase. Supóngase, por ejemplo, que se tienen 50 observaciones de x = rendimiento de combustible de un automóvil (mpg), de las cuales la más pequeña es 27.8 y la más grande es 31.4. En este caso se podrían usar los límites de clase 27.5, 28.0, 28.5, . . . , y 31.5 como se muestra a continuación:



Una posible dificultad es que a veces una observación queda en un límite de clase, así que no está exactamente en un intervalo; por ejemplo, 29.0. Una forma de superar esta dificultad es usar límites como 27.55, 28.05, . . . , 31.55. Añadir un dígito de centésimos a los límites de clase evita que las observaciones caigan en los límites resultantes. Otro método es usar las clases 27.5–<28.0, 28.0–<28.5, . . . , 31.0–<31.5. Entonces, 29.0 cae en la clase 29.0–<29.5 y no en la clase 28.5–<29.0. En otras palabras, con esta convención, una observación sobre un lí-

mite se coloca en el intervalo de la derecha del límite. Ésta es la forma como MINITAB elabora un histograma.

Construcción de un histograma para datos continuos: anchos de clase iguales

Determine la frecuencia y la frecuencia relativa para cada clase. Marque los límites de clase en un eje de medición horizontal. Arriba de cada intervalo de clase, trace un rectángulo cuya altura sea la frecuencia relativa (o frecuencia) correspondiente.

Ejemplo 1.10 Las compañías eléctricas necesitan información acerca de los hábitos de consumo de los clientes para obtener pronósticos exactos de las demandas de energía. Los investigadores de *Wisconsin Power and Light* determinaron consumos de energía (en BTU) durante cierto periodo para una muestra de 90 hogares con calefacción de gas. Un valor ajustado del consumo se calculó como sigue:

$$\text{consumo ajustado} = \frac{\text{consumo}}{(\text{clima, en días grado})(\text{área de la casa})}$$

Con esto se obtuvieron los datos adjuntos (parte del conjunto de datos FURNACE.MTW, disponibles en MINITAB), que se ordenan de menor a mayor.

2.97	4.00	5.20	5.56	5.94	5.98	6.35	6.62	6.72	6.78
6.80	6.85	6.94	7.15	7.16	7.23	7.29	7.62	7.62	7.69
7.73	7.87	7.93	8.00	8.26	8.29	8.37	8.47	8.54	8.58
8.61	8.67	8.69	8.81	9.07	9.27	9.37	9.43	9.52	9.58
9.60	9.76	9.82	9.83	9.83	9.84	9.96	10.04	10.21	10.28
10.28	10.30	10.35	10.36	10.40	10.49	10.50	10.64	10.95	11.09
11.12	11.21	11.29	11.43	11.62	11.70	11.70	12.16	12.19	12.28
12.31	12.62	12.69	12.71	12.91	12.92	13.11	13.38	13.42	13.43
13.47	13.60	13.96	14.24	14.35	15.12	15.24	16.06	16.90	18.26

Se deja que MINITAB seleccione los intervalos de clase. La característica más notable del histograma de la figura 1.9 es su semejanza con una curva en forma de campana (y por lo tanto simétrica), con el eje de simetría aproximadamente en 10.

Frecuencia de clase	1–<3	3–<5	5–<7	7–<9	9–<11	11–<13	13–<15	15–<17	17–<19
	1	1	11	21	25	17	9	4	1
Frecuencia relativa	0.011	0.011	0.122	0.233	0.278	0.189	0.100	0.044	0.011

Del histograma,

$$\text{proporción de observaciones menores que 9} \approx 0.01 + 0.01 + 0.12 + 0.23 = 0.37 \quad \left(\text{valor exacto} = \frac{34}{90} = 0.378\right)$$

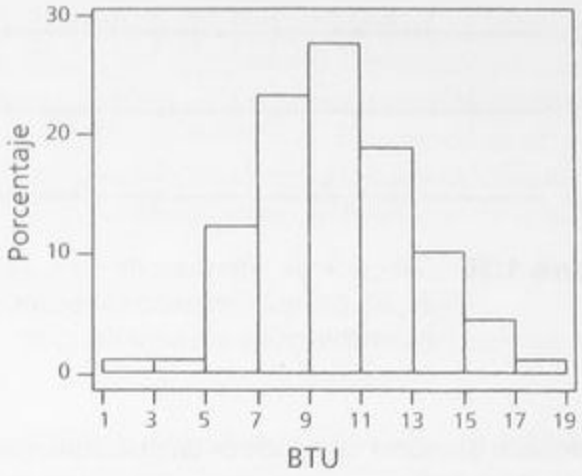


Figura 1.9 Histograma de los datos de consumo de energía del ejemplo 1.10

La frecuencia relativa para la clase 9–<11 es casi 0.27, así que se estima que más o menos la mitad de esto, es decir 0.135, está entre 9 y 10. Así,

$$\text{proporción de observaciones menores que 10} \approx 0.37 + 0.135 = 0.505 \text{ (un poco más de 50\%)}$$

El valor exacto de esta proporción es 47/90 = 0.522. ■

No hay reglas invariables en relación con el número de clases o la elección de las clases. Entre 5 y 20 clases son satisfactorias para la mayor parte de conjuntos de datos. En general, mientras más grande sea el número de observaciones en un conjunto de datos, se deben usar más clases. Una regla empírica razonable es número de clases

$$\text{número de clases} \approx \sqrt{\text{número de observaciones}}$$

Es posible que las clases de igual amplitud no sean una elección atinada si un conjunto de datos se "extiende" a un lado u otro. En la figura 1.10 (página 20) se ilustra un diagrama de puntos de tal conjunto de datos. Con una cantidad pequeña de clases de igual amplitud se obtienen casi todas las observaciones que caen en sólo una o dos de las clases. Si se utiliza una gran cantidad de clases del mismo ancho, muchas tienen frecuencia cero. Una segunda elección es usar algunos intervalos más anchos cerca de observaciones extremas e intervalos más reducidos en la región de alta concentración.

Construcción de un histograma para datos continuos: anchos de clase desiguales

Después de determinar las frecuencias y las frecuencias relativas, calcule la altura de cada rectángulo con la fórmula:

$$\text{altura del rectángulo} = \frac{\text{frecuencia relativa de la clase}}{\text{ancho de clase}}$$

Las alturas de rectángulo que resultan de denominan *densidades*, y la escala vertical es la **escala de densidad**. Esto también funciona cuando los anchos de clase son iguales.

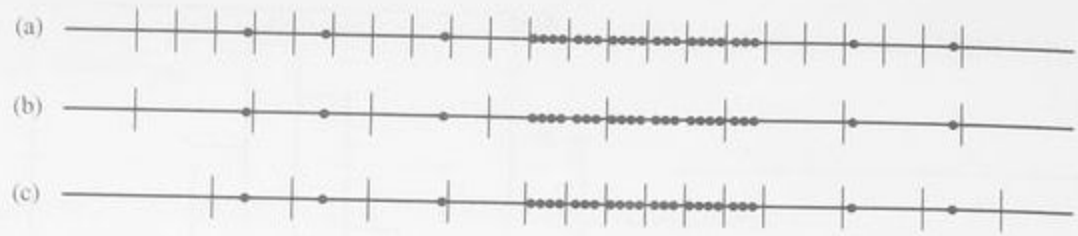


Figura 1.10 Selección de intervalos de clase para puntos alejados: (a) muchos intervalos cortos de igual anchura; (b) pocos intervalos largos de igual longitud; (c) intervalos de anchura desigual

Ejemplo 1.11 La corrosión del acero reforzado es un problema grave en las estructuras de concreto dentro de ambientes afectados por condiciones climatológicas extremas. Por esta razón, los investigadores estudian el empleo de varillas de refuerzo fabricadas con materiales compuestos. Se llevó a cabo un estudio a fin de establecer normas para fijar al concreto varillas de refuerzo plásticas reforzadas con fibra de vidrio ("Design Recommendations for Bond of GFRP Rebars to Concrete", *J. of Structural Engr.*, 1996: 247-254). Considere las 48 observaciones siguientes en la resistencia de enlace medida:

11.5	12.1	9.9	9.3	7.8	6.2	6.6	7.0	13.4	17.1	9.3	5.6
5.7	5.4	5.2	5.1	4.9	10.7	15.2	8.5	4.2	4.0	3.9	3.8
3.6	3.4	20.6	25.5	13.8	12.6	13.1	8.9	8.2	10.7	14.2	7.6
5.2	5.5	5.1	5.0	5.2	4.8	4.1	3.8	3.7	3.6	3.6	3.6

Frecuencia de clase	2–<4 9	4–<6 15	6–<8 5	8–<12 9	12–<20 8	20–<30 2
Frecuencia relativa	0.1875	0.3125	0.1042	0.1875	0.1667	0.0417
Densidad	0.094	0.156	0.052	0.047	0.021	0.004

El histograma resultante se muestra en la figura 1.11. El extremo derecho o superior se prolonga mucho más que el extremo izquierdo o inferior: una desviación sustancial respecto a la simetría.

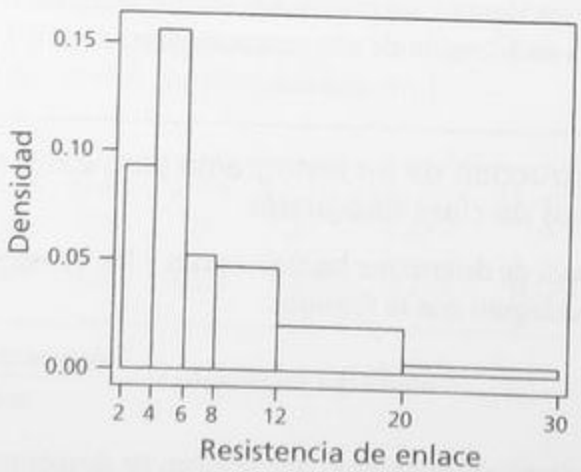


Figura 1.11 Histograma de densidad obtenido con MINITAB para los datos de resistencia del enlace del ejemplo 1.11 ■

Cuando los anchos de clase son distintos, el hecho de no usar una escala de densidad, da como resultado una figura con áreas distorsionadas. Para anchos de clase iguales, el divisor es el mismo en cada cálculo de densidad, y las operaciones aritméticas adicionales sólo dan como resultado una nueva escala del eje vertical (es decir, el histograma de frecuencia relativa y el de densidad tienen la misma apariencia. Un histograma de densidades tiene una propiedad interesante. Al multiplicar por el ancho de clase ambos lados de la fórmula de la densidad, se obtiene

$$\begin{aligned}\text{frecuencia relativa} &= (\text{ancho de clase})(\text{densidad}) \\ &= (\text{ancho del rectángulo})(\text{altura del rectángulo}) \\ &= \text{área del rectángulo}\end{aligned}$$

Es decir, *el área de cada rectángulo es la frecuencia relativa de la clase correspondiente*. Además, puesto que la suma de las frecuencias relativas debe ser 1.0 (salvo en caso de redondeo), *el área total de los rectángulos en un histograma de densidad es igual a* 1. Siempre es posible trazar un histograma de manera que el área sea igual a la frecuencia relativa (esto también es válido para un histograma de datos discretos; sólo se usa la escala de densidad). Esta propiedad desempeña un papel importante en la elaboración de modelos para distribuciones en el capítulo 4.

Formas de histogramas

Los histogramas tienen diversas formas. Un histograma **unimodal** es el que sube hasta un solo pico y después baja. Un histograma **bimodal** tiene dos picos distintos. La bimodalidad resulta cuando el conjunto de datos consiste en observaciones de dos clases bastante distintas de individuos u objetos. Por ejemplo, considere un conjunto grande de datos que consiste en los tiempos de conducción para automóviles que viajan entre San Luis Obispo, California y Monterey, California (sin incluir el tiempo que toma echar un vistazo, comer, etcétera). En este histograma aparecerían dos picos, uno para los automóviles que toman la ruta tierra adentro (alrededor de 2.5 horas) y otro para los automóviles que viajan por la costa (3.5 a 4 horas). Sin embargo, en este tipo de situaciones la bimodalidad no se sigue de manera automática. La bimodalidad ocurre en el histograma de datos combinados sólo si los dos histogramas separados están "muy apartados" con respecto a sus dispersiones. Así, un conjunto de datos grande que consiste en las estaturas de estudiantes universitarios no debe producir un histograma bimodal porque la estatura representativa de los varones de casi 69 pulgadas no está muy arriba de la estatura característica de las mujeres, alrededor de 64 a 65 pulgadas. Se dice que un histograma con más de dos picos es **multimodal**. Por supuesto, el número de picos podría depender de la elección de los intervalos de clase, en particular con una cantidad pequeña de observaciones.

Un histograma es **simétrico** si la mitad izquierda es imagen especular de la mitad derecha. Un histograma unimodal tiene **sesgo positivo** si el extremo derecho o superior se extiende comparado con el extremo izquierdo o inferior, y tiene **sesgo negativo** si se extiende a la izquierda. En la figura 1.12 se muestran histogramas "suavizados", obtenidos al superponer una curva suavizada sobre los rectángulos, que ilustran las diversas posibilidades.

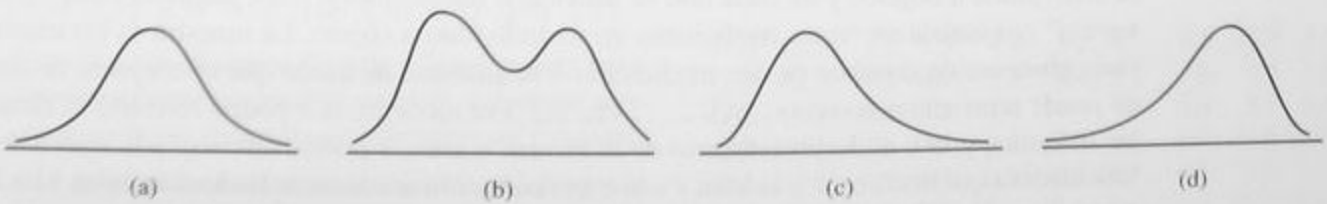


Figura 1.12 Histogramas suavizados: (a) unimodal simétrico, (b) bimodal, (c) con sesgo positivo y (d) con sesgo negativo

Datos cualitativos

Tanto una distribución de frecuencias como un histograma se pueden construir cuando el conjunto de datos es de naturaleza *cualitativa* (categórica). En algunos casos habrá un ordenamiento natural de clases; por ejemplo, estudiantes de primer año, segundo, penúltimo, último y graduados, en tanto que en otros casos el orden será arbitrario, por ejemplo, católicos, judíos, protestantes y otros. Con estos datos categóricos, los intervalos anteriores cuyos rectángulos se construyen, deben tener igual amplitud.

Ejemplo 1.12 A cada miembro de una muestra de 120 individuos dueños de motocicletas se le pidió el nombre del fabricante de su máquina. La distribución de frecuencia para los datos resultantes se da en la tabla 1.2, y el histograma se ilustra en la figura 1.13.

Tabla 1.2 Distribución de frecuencias para los datos de las motocicletas

Fabricante	Frecuencia	Frecuencia relativa
1. Honda	41	0.34
2. Yamaha	27	0.23
3. Kawasaki	20	0.17
4. Harley-Davidson	18	0.15
5. BMW	3	0.03
6. Otras	11	0.09
	120	1.01

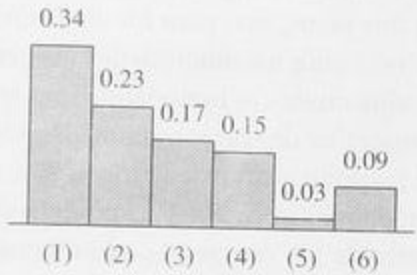


Figura 1.13 Histograma para datos de motocicletas ■

Datos multivariados

Las técnicas presentadas hasta el momento han sido sólo para situaciones en las que cada observación de un conjunto de datos es ya sea un solo número o una sola categoría. Sin embargo, es común que los datos sean de naturaleza *multivariada*. Es decir, si se obtiene una muestra de individuos u objetos y en cada uno se hacen dos mediciones o más, entonces cada "observación" consistiría en varias mediciones en un individuo u objeto. La muestra es bivariada si cada observación consiste en dos mediciones o respuestas, de modo que el conjunto de datos se puede representar como $(x_1, y_1), \ldots, (x_n, y_n)$. Por ejemplo, la x podría referirse al tamaño de máquina, y la y al desplazamiento de la misma, o bien, x podría referirse a la marca de la calculadora que uno posee y la letra y a la especialización académica. En los capítulos 11 a 14, se analizan conjuntos de datos multivariados de este tipo, así que se pospone hasta ese momento una explicación detallada.

Ejercicios | Sección 1.2 (10 a 32)

10. Considere los datos de resistencia para las vigas del ejemplo 1.2.

a. Construya un diagrama de tallo y hojas con los datos. ¿Cuál parece ser un valor de resistencia representativo? ¿Según parece las observaciones están muy concentradas respecto al valor representativo o bastante dispersas?

b. ¿El diagrama tiene una apariencia razonablemente simétrica respecto a un valor representativo, o se describiría su forma de algún otro modo?

c. ¿Observa algunos valores de resistencia alejados?

d. ¿Qué proporción de las observaciones de resistencia en esta muestra sobrepasan los 10 MPa?

11. Cada puntuación en el siguiente lote de calificaciones de exámenes se encuentra en 60 y 69, 70 y 79, 80 y 89 o 90 y más. Un diagrama de tallo y hojas con sólo los cuatro tallos 6, 7, 8 y 9 no proporcionaría una descripción detallada de la distribución de calificaciones. En este tipo de situaciones, es recomendable usar tallos repetidos. Aquí se podría repetir el tallo 6 dos veces, usando 6L para las calificaciones bajas de 60 (hojas 0, 1, 2, 3 y 4) y 6H para las calificaciones altas de 60 (hojas 5, 6, 7, 8 y 9). De manera similar, los otros tallos se pueden repetir dos veces para obtener una muestra que consiste en ocho renglones. Construya un diagrama para las puntuaciones suministradas. ¿Qué característica de los datos se remarca en este diagrama?

74 89 80 93 64 67 72 70 66 85 89 81 81
71 74 82 85 63 72 81 81 95 84 81 80 70
69 66 60 83 85 98 84 68 90 82 69 72 87
88

12. Los valores de densidad relativa adjuntos para varios tipos de madera utilizados en la construcción se publicaron en el artículo "Bolted Connection Design Values Based on European Yield Model" (*J. of Structural Engr.*, 1993: 2169-2186):

0.31 0.35 0.36 0.36 0.37 0.38 0.40 0.40 0.40
0.41 0.41 0.42 0.42 0.42 0.42 0.42 0.43 0.44
0.45 0.46 0.46 0.47 0.48 0.48 0.48 0.51 0.54
0.54 0.55 0.58 0.62 0.66 0.66 0.67 0.68 0.75

Construya un diagrama de tallo y hojas usando tallos repetidos (véase el ejercicio anterior), y comente acerca de cualquier característica interesante del diagrama.

13. Las propiedades mecánicas permisibles para diseño estructural de vehículos aeroespaciales metálicos, requieren un método aprobado para analizar estadísticamente los datos de prueba empíricos. En el artículo "Establishing Mechanical Property Allowables for Metals" (*J. of Testing and Evaluation*, 1998: 293-299), se emplearon los datos adjuntos de resistencia última a la tensión (ksi) como base para superar las dificultades en el desarrollo del método.

122.2 124.2 124.3 125.6 126.3 126.5 126.5 127.2 127.3
127.5 127.9 128.6 128.8 129.0 129.2 129.4 129.6 130.2
130.4 130.8 131.3 131.4 131.4 131.5 131.6 131.6 131.8
131.8 132.3 132.4 132.4 132.5 132.5 132.5 132.5 132.6
132.7 132.9 133.0 133.1 133.1 133.1 133.1 133.2 133.2
133.2 133.3 133.3 133.5 133.5 133.5 133.8 133.9 134.0
134.0 134.0 134.0 134.1 134.2 134.3 134.4 134.4 134.6
134.7 134.7 134.7 134.8 134.8 134.8 134.9 134.9 135.2
135.2 135.2 135.3 135.3 135.4 135.5 135.5 135.6 135.6
135.7 135.8 135.8 135.8 135.8 135.8 135.9 135.9 135.9
135.9 136.0 136.0 136.1 136.2 136.2 136.3 136.4 136.4
136.6 136.8 136.9 136.9 137.0 137.1 137.2 137.6 137.6
137.8 137.8 137.8 137.9 137.9 138.2 138.2 138.3 138.3
138.4 138.4 138.4 138.5 138.5 138.6 138.7 138.7 139.0
139.1 139.5 139.6 139.8 139.8 140.0 140.0 140.7 140.7
140.9 140.9 141.2 141.4 141.5 141.6 142.9 143.4 143.5
143.6 143.8 143.8 143.9 144.1 144.5 144.5 147.7 147.7

a. Construya un diagrama de tallo y hojas de los datos eliminando (truncando) primero el dígito de las decenas y luego repitiendo cada valor del tallo cinco veces (una vez para las hojas 1 y 2, una segunda vez para las hojas 3 y 4, etcétera). ¿Por qué es relativamente fácil identificar un valor de resistencia representativo?

b. Construya un histograma usando las clases de igual amplitud con la primera clase con un límite inferior de 122 y un límite superior de 124. Luego, comente acerca de cualquier característica interesante del histograma.

14. El conjunto de datos adjunto consiste en observaciones acerca del flujo de la regadera (L/min) para una muestra de $n = 129$ casas en Perth, Australia ("An Application of Bayes Methodology to the Analysis of Diary Records in a Water Use Study", *J. Amer. Stat. Assoc.*, 1987: 705-711):

4.6 12.3 7.1 7.0 4.0 9.2 6.7 6.9 11.5 5.1
11.2 10.5 14.3 8.0 8.8 6.4 5.1 5.6 9.6 7.5
7.5 6.2 5.8 2.3 3.4 10.4 9.8 6.6 3.7 6.4
8.3 6.5 7.6 9.3 9.2 7.3 5.0 6.3 13.8 6.2
5.4 4.8 7.5 6.0 6.9 10.8 7.5 6.6 5.0 3.3
7.6 3.9 11.9 2.2 15.0 7.2 6.1 15.3 18.9 7.2
5.4 5.5 4.3 9.0 12.7 11.3 7.4 5.0 3.5 8.2
8.4 7.3 10.3 11.9 6.0 5.6 9.5 9.3 10.4 9.7

5.1	6.7	10.2	6.2	8.4	7.0	4.8	5.6	10.5	14.6
10.8	15.5	7.5	6.4	3.4	5.5	6.6	5.9	15.0	9.6
7.8	7.0	6.9	4.1	3.6	11.9	3.7	5.7	6.8	11.3
9.3	9.6	10.4	9.3	6.9	9.8	9.1	10.6	4.5	6.2
8.3	3.2	4.9	5.0	6.0	8.2	6.3	3.8	6.0	

a. Construya un diagrama de tallo y hojas de los datos.
b. ¿Cuál es un flujo representativo?
c. ¿Al parecer el diagrama está muy concentrado o disperso?
d. ¿La distribución de valores parece ser razonablemente simétrica? Si no es así, ¿cómo describiría la desviación respecto a la simetría?
e. ¿Describiría alguna observación como alejada del resto de los datos (un valor atípico)?

15. En un artículo de *Consumer Reports* acerca de la mantequilla de maní (septiembre de 1990) aparecen las puntuaciones siguientes para varias marcas:

Cremosa	56	44	62	36	39	53	50	65	45	40
	56	68	41	30	40	50	56	30	22	
Crujiente	62	53	75	42	47	40	34	62	52	
	50	34	42	36	75	80	47	56	62	

Construya un diagrama *comparativo* de tallo y hojas listando los tallos a la mitad de la página y luego muestre las hojas para la mantequilla cremosa a la derecha y las hojas para la mantequilla crujiente a la izquierda. Describa similitudes y diferencias para los dos tipos.

16. En el artículo citado del ejemplo 1.2 también se proporcionan las observaciones de resistencia adjuntas para los cilindros:

6.1	5.8	7.8	7.1	7.2	9.2	6.6	8.3	7.0	8.3
7.8	8.1	7.4	8.5	8.9	9.8	9.7	14.1	12.6	11.2

a. Construya un diagrama comparativo de tallo y hojas (véase el ejercicio anterior) de los datos de la viga y el cilindro, y luego conteste las preguntas de los incisos (b) a (d) del ejercicio 10 para las observaciones de los cilindros.
b. ¿En qué forma los dos lados del diagrama son similares? ¿Hay diferencias obvias entre las observaciones de la viga y las del cilindro?
c. Construya una gráfica de puntos de los datos del cilindro.

17. Transductores de temperatura de un cierto tipo se envían en lotes de 50. Se seleccionó una muestra de 60 lotes, y se determinó el número de transductores de cada lote que difieren de las especificaciones de diseño, dando como resultado los siguientes datos:

2	1	2	4	0	1	3	2	0	5	3	3	1	3	2	4	7	0	2	3
0	4	2	1	3	1	1	3	4	1	2	3	2	2	8	4	5	1	3	1
5	0	2	3	2	1	0	6	4	2	1	6	0	3	3	3	6	1	2	3

a. Determine las frecuencias y las frecuencias relativas para los valores observados de x = número de transductores que incumplen las especificaciones en un lote.
b. ¿Qué proporción de lotes en la muestra tienen a lo sumo cinco transductores discordantes? ¿Qué proporción tiene menos de cinco? ¿Qué proporción tiene al menos cinco unidades discordantes?
c. Trace un histograma de los datos con la frecuencia relativa en la escala vertical, y comente acerca de las características.

18. En un estudio de productividad de autores ("Lotka's Test", *Collection Mgmt.*, 1982: 111-118), se clasificó a un gran número de autores según la cantidad de artículos que publicaron durante cierto periodo. Los resultados se presentan en la distribución de frecuencias adjunta:

Número de artículos	1	2	3	4	5	6	7	8
Frecuencia	784	204	127	50	33	28	19	19

Número de artículos	9	10	11	12	13	14	15	16	17
Frecuencia	6	7	6	7	4	4	5	3	3

a. Construya un histograma que corresponda a esta distribución de frecuencias. ¿Cuál es la característica más interesante de la forma de la distribución?
b. ¿Qué proporción de estos autores publicaron por lo menos cinco artículos? ¿Por lo menos diez artículos? ¿Más de diez artículos?
c. Suponga que los cinco números 15, los tres 16 y los tres 17 se hubieran agrupado en una sola categoría mostrada como "≥15". ¿Se podría trazar un histograma? Explique.
d. Suponga que en lugar de que los valores 15, 16 y 17 se listen por separado, se hubieran combinado en una categoría 15 a 17 con frecuencia 11. ¿Se podría trazar un histograma? Explique.

19. Se determinó el número de partículas contaminantes en una oblea de silicio antes de cierto proceso de enjuague para cada oblea de una muestra de tamaño 100, obteniéndose las frecuencias siguientes:

Número de partículas	0	1	2	3	4	5	6	7
Frecuencia	1	2	3	12	11	15	18	10

Número de partículas	8	9	10	11	12	13	14
Frecuencia	12	4	5	3	1	2	1

a. ¿Qué proporción de las obleas muestreadas tuvo por lo menos una partícula? ¿Por lo menos cinco partículas?

b. ¿Qué proporción de las obleas muestreadas tuvo entre cinco y diez partículas, inclusive? ¿Estrictamente entre cinco y diez partículas?

c. Dibuje un diagrama usando frecuencia en el eje central. ¿Cómo describiría la forma del diagrama?

20. En el artículo "Determination of Most Representative Subdivision" (*J. of Energy Engr.*, 1993: 43-55) se muestran los datos de varias características de subdivisiones que se podrían usar para decidir si se suministra energía eléctrica por medio de líneas aéreas o subterráneas. A continuación se dan los valores de la variable x = longitud total de calles dentro de una subdivisión:

1280	5320	4390	2100	1240	3060	4770
1050	360	3330	3380	340	1000	960
1320	530	3350	540	3870	1250	2400
960	1120	2120	450	2250	2320	2400
3150	5700	5220	500	1850	2460	5850
2700	2730	1670	100	5770	3150	1890
510	240	396	1419	2109		

a. Construya un diagrama de tallo y hojas usando el dígito de las unidades de millar como el tallo y el dígito de las centenas como la hoja, y comente acerca de las distintas características del diagrama.

b. Construya un histograma con los límites de clase 0, 1000, 2000, 3000, 4000, 5000 y 6000. ¿Qué proporción de las subdivisiones tiene longitud total menor de 2000? ¿Entre 2000 y 4000? ¿Cómo describiría la forma del histograma?

21. El artículo citado en el ejercicio 20 también proporciona los valores siguientes de las variables y = número de calles sin salida y z = número de intersecciones:

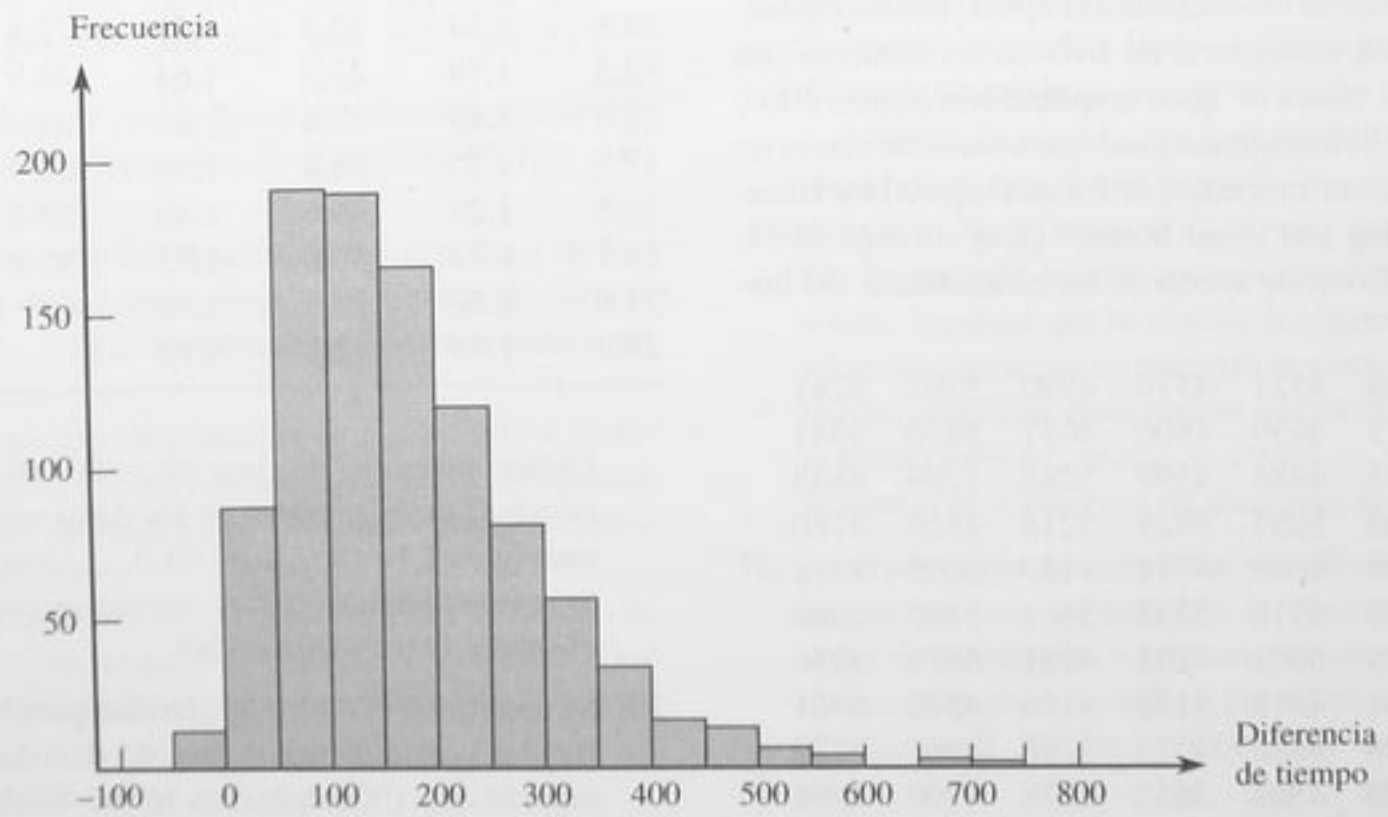
y	1	0	1	0	0	2	0	1	1	1	2	1	0	0	1	1	0	1	1	
z	1	8	6	1	1	5	3	0	0	4	4	0	0	1	2	1	4	0	4	
y	1	1	0	0	0	1	1	2	0	1	2	2	1	1	0	2	1	1	0	
z	0	3	0	1	1	0	1	3	2	4	6	6	0	1	1	8	3	3	5	
y	1	5	0	3	0	1	1	0	0											
z	0	5	2	3	1	0	0	0	3											

a. Construya un histograma para los datos de y. ¿Qué proporción de estas subdivisiones no tuvo calles sin salida? ¿Por lo menos una calle sin salida?

b. Construya un histograma para los datos de z. ¿Qué proporción en estas subdivisiones tuvo a lo sumo cinco intersecciones? ¿Menos de cinco intersecciones?

22. ¿Cómo varía la velocidad de un corredor en el curso de un maratón (una distancia de 42.195 km)? Considere determinar el tiempo de carrera de los primeros cinco km y el tiempo de carrera entre los puntos 35 km y 40 km, y luego reste el primer tiempo del último. Un valor positivo de esta diferencia corresponde a un corredor que disminuye su velocidad hacia el final de la carrera. El histograma adjunto se basa en tiempos de corredores que participaron en varios maratones realizados en Japón ("Factors Affecting Runners' Marathon Performance", *Chance*, otoño de 1993: 24-30).

Histograma para el ejercicio 22



¿Cuáles son algunas características interesantes de este histograma? ¿Cuál es un valor de diferencia característico? ¿Más o menos qué proporción de los corredores recorre más rápido la distancia posterior que la primera?

23. En un estudio de ruptura por torsión durante el tejido de tela (*Technometrics*, 1982: 63), se probaron 100 muestras de hilo. Se determinó el número de ciclos de tensión a la ruptura para cada muestra de hilo, obteniéndose los datos siguientes:

86	146	251	653	98	249	400	292	131	169
175	176	76	264	15	364	195	262	88	264
157	220	42	321	180	198	38	20	61	121
282	224	149	180	325	250	196	90	229	166
38	337	65	151	341	40	40	135	597	246
211	180	93	315	353	571	124	279	81	186
497	182	423	185	229	400	338	290	398	71
246	185	188	568	55	55	61	244	20	284
393	396	203	829	239	236	286	194	277	143
198	264	105	203	124	137	135	350	193	188

a. Construya un histograma de frecuencias relativas con base en los intervalos de clase 0–<100, 100–<200, . . . , y comente acerca de las características del histograma.

b. Construya un histograma con base en los intervalos de clase siguientes: 0–<50, 50–<100, 100–<150, 150–<200, 200–<300, 300–<400, 400–<500, 500–<600 y 600–<900.

c. Si las especificaciones de tejido requieren una resistencia a la ruptura de por lo menos 100 ciclos, ¿qué proporción de los especímenes de hilo de esta muestra sería considerada satisfactoria?

24. El conjunto de datos adjunto consiste en observaciones de resistencia al corte (lb) de soldaduras de punto ultrasónicas hechas en un determinado tipo de lámina Alclad. Construya un histograma de frecuencias relativas con base en diez clases de igual amplitud con límites 4000, 4200, ... [El histograma tiene que concordar con el de "Comparison of Properties of Joints Prepared by Ultrasonic Welding and Other Means" (*J. of Aircraft*, 1983: 552-556).] Comente acerca de las caraterísticas del histograma.

5434	4948	4521	4570	4990	5702	5241
5112	5015	4659	4806	4637	5670	4381
4820	5043	4886	4599	5288	5299	4848
5378	5260	5055	5828	5218	4859	4780
5027	5008	4609	4772	5133	5095	4618
4848	5089	5518	5333	5164	5342	5069
4755	4925	5001	4803	4951	5679	5256
5207	5621	4918	5138	4786	4500	5461
5049	4974	4592	4173	5296	4965	5170
4740	5173	4568	5653	5078	4900	4968
5248	5245	4723	5275	5419	5205	4452
5227	5555	5388	5498	4681	5076	4774
4931	4493	5309	5582	4308	4823	4417
5364	5640	5069	5188	5764	5273	5042
5189	4986					

25. Una transformación de datos por medio de alguna función matemática, como raíz de $\sqrt{x}$ o $1/x$, suele ofrecer un conjunto de números que tienen propiedades estadísticas "más bonitas" que los datos originales. En particular, se podría encontrar una función para la cual el histograma de valores transformados es más simétrico (o, incluso mejor, más parecido a una curva en forma de campana) que el de los datos originales. Por ejemplo, en el artículo "Time Lapse Cinematographic Analysis of Beryllium–Lung Fibroblast Interactions" (*Environ. Research*, 1983: 34-43) se muestran los resultados de experimentos diseñados para estudiar el comportamiento de determinadas células que hubieran sido expuestas a berilio. Una característica importante de la célula es su tiempo de interdivisión (TID). Los TID se determinaron para una gran cantidad de células tanto en condiciones expuestas (tratamiento) como inexpuestas (control). Los autores del artículo utilizaron una transformación logarítmica, es decir, valor transformado = log(valor original). Considere los siguientes datos representativos de TID.

TID	$\log_{10}$(TID)	TID	$\log_{10}$(TID)	TID	$\log_{10}$(TID)
28.1	1.45	60.1	1.78	21.0	1.32
31.2	1.49	23.7	1.37	22.3	1.35
13.7	1.14	18.6	1.27	15.5	1.19
46.0	1.66	21.4	1.33	36.3	1.56
25.8	1.41	26.6	1.42	19.1	1.28
16.8	1.23	26.2	1.42	38.4	1.58
34.8	1.54	32.0	1.51	72.8	1.86
62.3	1.79	43.5	1.64	48.9	1.69
28.0	1.45	17.4	1.24	21.4	1.33
17.9	1.25	38.8	1.59	20.7	1.32
19.5	1.29	30.6	1.49	57.3	1.76
21.1	1.32	55.6	1.75	40.9	1.61
31.9	1.50	25.5	1.41		
28.9	1.46	52.1	1.72		

Utilice intervalos de clase 10–<20, 20–<30, . . . para construir un histograma de los datos originales. Utilice intervalos 1.1–<1.2, 1.2–<1.3, . . . con el fin de hacer lo mismo para los datos transformados. ¿Cuál es el efecto de la transformación?

26. Se determinó el índice de claridad para los cielos sobre Bagdad para cada uno de los 365 días durante un determinado año ("Contribution to the Study of the Solar

Radiation Climate of the Baghdad Environment", *Solar Energy*, 1990: 7-12). En la tabla siguiente se dan los resultados.

Clase	Frecuencia
0.15–<0.25	8
0.25–<0.35	14
0.35–<0.45	28
0.45–<0.50	24
0.50–<0.55	39
0.55–<0.60	51
0.60–<0.65	106
0.65–<0.70	84
0.70–<0.75	11

a. Determine las frecuencias relativas y trace el histograma correspondiente.

b. Los días nublados son los que tienen un índice de claridad más pequeño que 0.35. ¿Qué porcentaje de los días fueron nublados?

c. Los días claros son aquellos para los que el índice es por lo menos 0.65. ¿Qué porcentaje de los días fueron claros?

27. En el artículo "Study on the Life Distribution of Microdrills" (*J. of Engr. Manufacture*, 2002: 301-305) aparecen las observaciones siguientes, listadas en orden creciente, del tiempo de vida del taladro (número de agujeros que maquina un taladro antes de romperse) cuando se hacen agujeros en una cierta aleación de latón.

11	14	20	23	31	36	39	44	47	50
59	61	65	67	68	71	74	76	78	79
81	84	85	89	91	93	96	99	101	104
105	105	112	118	123	136	139	141	148	158
161	168	184	206	248	263	289	322	388	513

a. ¿Por qué no se puede basar una distribución de frecuencia en los intervalos de clase 0-50, 50-100, 100-150, etcétera?

b. Construya una distribución de frecuencias y un histograma de los datos usando los límites de clase 0, 50, 100, . . . , y comente acerca de características interesantes.

c. Construya una distribución de frecuencias e histograma de logaritmos naturales de las observaciones de tiempo de vida, y comente acerca de características interesantes.

d. ¿Qué proporción de las observaciones de tiempo de vida en esta muestra son menores que 100? ¿Qué proporción de las observaciones son por lo menos 200?

28. Construya un diagrama de puntos y dígitos para la serie de tiempo adjunta. Los datos son mensuales y se obtuvieron durante el periodo 1985-1989. Cada valor es la radiación solar promedio en la banda de 385 a 530 nm como porcentaje de la radiación total ("Global Energy in the Different Spectral Bands at Dhahran, Saudi Arabia", *J. Solar Energy Engr.*, 1991: 290-294). Comente acerca de características importantes de los datos.

20.9	19.6	20.4	20.3	20.8	20.6	20.5	20.4
19.9	19.8	19.5	20.2	16.5	18.3	18.7	19.6
20.0	20.0	19.5	19.6	19.1	18.8	18.3	17.6
17.2	17.8	18.7	19.0	19.0	18.6	18.8	19.0
18.5	18.3	17.5	16.9	17.0	17.8	18.1	18.8
18.9	18.9	19.1	18.8	18.4	17.8	17.0	16.8
17.9	18.4	19.0	19.4	19.7	19.5	19.5	19.5
19.0	18.7	18.1	17.9				

29. Considere los datos siguientes acerca del tipo de dolencia (J = inflamación de las articulaciones, F = fatiga, B = dolor de espalda, M = debilidad muscular, C = tos, N = irritación de la nariz, O = otra) de sembradores de árboles. Obtenga las frecuencias y las frecuencias relativas para las distintas categorías y trace un histograma. (Los datos son consistentes con los porcentajes proporcionados en el artículo "Physiological Effects of Work Stress and Pesticide Exposure in Tree Planting by British Columbia Silviculture Workers", *Ergonomics*, 1993: 951-961.)

O	O	N	J	C	F	B	B	F	O	J	O	O	M	
O	F	F	O	O	N	O	N	J	F	J	B	O	C	
J	O	J	J	F	N	O	B	M	O	J	M	O	B	
O	F	J	O	O	B	N	C	O	O	O	M	B	F	
J	O	F	N											

30. Un **diagrama de Pareto** es una variación de un histograma para datos categóricos que resultan de un estudio de control de calidad. Cada categoría representa un tipo diferente de producto que incumple las especificaciones o problema de producción. Las categorías están ordenadas de modo que la que tienen la frecuencia más grande aparezcan en el extremo izquierdo, luego la categoría con la segunda frecuencia más grande, y así sucesivamente. Suponga que se obtiene la siguiente información sobre discordancias en paquetes de circuitos: componente con fallas, 126; componente incorrecto, 210; soldadura insuficiente, 67; exceso de soldadura, 54; componente faltante, 131. Construya un diagrama de Pareto.

31. La **frecuencia acumulada** y la frecuencia relativa acumulada para un intervalo de clase particular son la suma de frecuencias y frecuencias relativas, respectivamente, para ese intervalo y los intervalos que queden debajo de él. Si, por ejemplo, hay cuatro intervalos

con frecuencias 9, 16, 13 y 12, entonces las frecuencias acumuladas son 9, 25, 38 y 50, y las frecuencias relativas acumuladas son 0.18, 0.50, 0.76 y 1.00. Calcule las frecuencias acumuladas y las frecuencias relativas acumuladas para los datos del ejercicio 24.

32. La carga de incendio (MJ/m^2) es la energía calorífica que la combustión del contenido y la estructura misma podrían liberar por metro cuadrado de área de piso. En el artículo "Fire Loads in Office Buildings" (*J. of Structural Engr.*, 1997: 365-368) se muestran los siguientes porcentajes acumulados (leídos de una gráfica) para cargas de incendio en una muestra de 388 habitaciones:

Valor	0	150	300	450	600
% acumulado	0	19.3	37.6	62.7	77.5
Valor	750	900	1050	1200	1350
% acumulado	87.2	93.8	95.7	98.6	99.1
Valor	1500	1650	1800	1950	
% acumulado	99.5	99.6	99.8	100.0	

a. Construya un histograma de frecuencias relativas y comente acerca de características importantes.

b. ¿Qué proporción de las cargas de incendio son menores que 600? ¿Por lo menos 1200?

c. ¿Qué proporción de las cargas está entre 600 y 1200?

1.3 Medidas de localización

Los resúmenes visuales de datos son herramientas excelentes para obtener impresiones y puntos de vista preliminares. El análisis más formal de los datos requiere calcular e interpretar medidas de resumen numéricas; es decir, a partir de los datos, se intenta extraer varias cantidades de resumen: números que puedan servir para caracterizar el conjunto de datos y mostrar algunas de sus propiedades sobresalientes. El interés fundamental tiene que ver con los datos numéricos, aunque al final de esta sección aparecen algunos comentarios en relación con datos categóricos.

Supóngase, entonces, que el conjunto de datos es de la forma $x_1, x_2, \ldots, x_n$, donde cada x_i es un número. ¿Qué características de este conjunto de números son más interesantes y merecen mayor atención? Una característica importante de un conjunto de números es su localización y, en particular, su centro. En esta sección se presentan métodos para describir la localización de un conjunto de datos; en la sección 1.4 se vuelve a los métodos para medir la variabilidad de un conjunto de números.

La media

Para un determinado conjunto de números $x_1, x_2, \ldots, x_n$, la medida de tendencia central más familiar y útil es la *medida*, o promedio aritmético del conjunto. Debido a que casi siempre se considera a las x_i como partes constituyentes de una muestra, se hace referencia al promedio aritmético como la *media muestral* y se denota por $\bar{x}$.

DEFINICIÓN

La **media muestral** $\bar{x}$ de un conjunto de observaciones $x_1, x_2, \ldots, x_n$ está dada por

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

El numerador de $\bar{x}$ se puede escribir de manera más formal como $\sum x_i$, donde la suma es sobre las observaciones de la muestra.

Para informar el valor de $\bar{x}$, se recomienda usar exactitud decimal de un dígito decimal más que para la exactitud de las x_i. Así, si las observaciones son las distancias de frenado $x_1 = 125$, $x_2 = 131$, etcétera, se podría tener $\bar{x} = 127.3$ pies.

Ejemplo 1.13 Se ha estudiado el agrietamiento que causa la corrosión por tensión cáustica del hierro y del acero como resultado de fallas en torno a los remaches en calderas de acero y en los rotores de máquinas de vapor. Examine las observaciones adjuntas sobre x = longitud de la grieta (μm) como resultado de pruebas de corrosión por tensión con carga constante en muestras flexibles de barras lisas durante un tiempo determinado. (Los datos son consistentes con un histograma y las cantidades de resumen del artículo "On the Role of Phosphorus in the Caustic Stress Corrosion Cracking of Low Alloy Steels", *Corrosion Science*, 1989: 53-68.)

$x_1 = 16.1$ $x_2 = 9.6$ $x_3 = 24.9$ $x_4 = 20.4$ $x_5 = 12.7$ $x_6 = 21.2$ $x_7 = 30.2$
$x_8 = 25.8$ $x_9 = 18.5$ $x_{10} = 10.3$ $x_{11} = 25.3$ $x_{12} = 14.0$ $x_{13} = 27.1$ $x_{14} = 45.0$
$x_{15} = 23.3$ $x_{16} = 24.2$ $x_{17} = 14.6$ $x_{18} = 8.9$ $x_{19} = 32.4$ $x_{20} = 11.8$ $x_{21} = 28.5$

En la figura 1.14 se muestra un diagrama de tallo y hojas de los datos; al parecer es "normal" una longitud de grieta cuyo valor se sitúa entre 20 y 25.

0H	96	89				
1L	27	03	40	46	18	
1H	61	85				
2L	49	04	12	33	42	
2H	58	53	71	85		
3L	02	24				
3H						
4L						
4H	50					

Tallo: dígito de la decenas
Hoja: dígito de las unidades y de los décimos

Figura 1.14 Diagrama de tallo y hojas de los datos de longitud de grieta

Con $\sum x_i = 444.8$, la media muestral es

$$\bar{x} = \frac{444.8}{21} = 21.18$$

un valor que es congruente con la información que presenta el diagrama de tallo y hojas. ■

Una interpretación física de $\bar{x}$ demuestra cómo ésta mide la localización (centro) de una muestra. Considere de medición horizontal, y luego represente cada observación muestral por un peso de 1 libra que se coloca en el punto correspondiente sobre el eje. El único punto en el que se puede poner un fulcro para equilibrar el sistema de pesas, es el que corresponde al valor de $\bar{x}$ (véase la figura 1.15).

Del mismo modo que $\bar{x}$ representa el valor promedio de las observaciones en una muestra, se puede calcular el promedio de los valores de la población. Este promedio se llama **media poblacional** y se denota mediante la letra griega μ Cuando hay N valores en la población

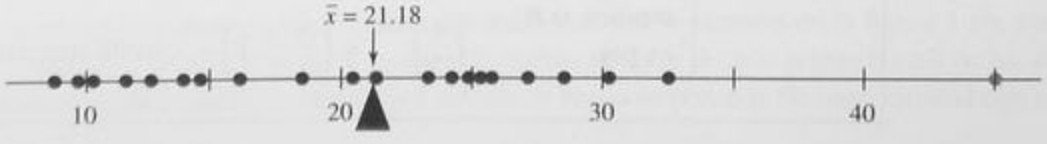


Figura 1.15 La media como punto de equilibrio para un sistema de pesas

(una población finita), entonces μ = (suma de los N valores de la población)/N. En los capítulos 3 y 4 se proporciona una definición más general para μ que se aplica tanto a poblaciones finitas como infinitas (desde un punto de vista conceptual). Así como $\bar{x}$ es una medida interesante e importante de la ubicación de la muestra, μ es una característica interesante e importante de una población (a veces la más importante). En los capítulos sobre inferencia estadística, se presentan métodos basados en la media muestral para sacar conclusiones acerca de una media poblacional. Por ejemplo, se podría utilizar la media muestral $\bar{x} = 21.18$ calculada en el ejemplo 1.13, como *estimación puntual* (un solo número que es la "mejor" suposición) de μ = la longitud de grieta promedio real para todas las muestras que recibieron el tratamiento descrito.

La media muestral adolece de una deficiencia que la convierte en una medida inapropiada del centro bajo algunas circunstancias: su valor se puede ver afectado en gran medida incluso por la presencia de un solo valor atípico (por lo general una observación grande o pequeña). En el ejemplo 1.13, el valor $x_{14} = 45.0$ es, claro está, un valor atípico. Sin esta observación $\bar{x} = 399.8/20 = 19.99$; el valor atípico aumenta a la media en más de 1 μm. Si el valor catastrófico de 45.0 μm un valor atípico en realidad extremo, sustituye a la observación de 295.0 μm, entonces, $\bar{x} = 694.8/21 = 33.09$, que es mayor que todas las observaciones, excepto una.

Una muestra de ingresos suele producir algunos valores atípicos (los pocos afortunados que obtienen ingresos muy altos), y el uso del ingreso promedio como medida de localización resulta engañoso. Estos ejemplos hacen pensar en buscar una medida que sea menos sensible a valores atípicos que $\bar{x}$; y por el momento se propondrá una. Sin embargo, si bien $\bar{x}$ tiene este posible defecto, aún es la medida más utilizada, en gran parte porque hay muchas poblaciones para las que un valor inusual en la muestra sería muy poco probable. Cuando se lleva a cabo un muestreo de una población como ésta (donde una población normal o en forma de campana es el ejemplo más importante), la media muestral tiende a ser estable y bastante representativa de la muestra.

La mediana

La palabra *mediana* es sinónimo de "mitad", y la mediana muestral es de hecho el valor medio cuando las observaciones se ordenan de menor a mayor. Si las observaciones se denotan por $x_1, \ldots, x_n$, se emplea el símbolo $\tilde{x}$ para representar la mediana muestral.

DEFINICIÓN

La **mediana muestral** se obtiene al ordenar primero las n observaciones de menor a mayor (incluso valores repetidos de manera que cada observación muestral aparezca en la lista ordenada). Entonces,

$$\tilde{x} = \begin{cases} \text{Único valor medio si } n \text{ es impar} & = \left(\dfrac{n+1}{2}\right)^{\text{ésimo}} \text{ valor ordenado} \\ \text{Promedio de los dos valores medios si } n \text{ es par} & = \text{promedio de los valores ordenados } \left(\dfrac{n}{2}\right)^{\text{ésimo}} \text{ y } \left(\dfrac{n}{2}+1\right)^{\text{ésimo}} \text{ valor ordenado} \end{cases}$$

Ejemplo 1.14 El riesgo de manifestar deficiencia de hierro en algún momento es alto, en particular durante el embarazo. El problema con la detección de esta deficiencia, es que algunos métodos para cuantificar el hierro se ven afectados por el estado de embarazo. Considere los siguientes datos en relación con la concentración del receptor de transferrina para una muestra de mujeres con pruebas de laboratorio de anemia explícita por deficiencia de hierro ("Serum Transferrin Receptor for the Detection of Iron Deficiency in Pregnancy," *Amer. J. of Clinical Nutrition*, 1991: 1077-1081):

$$x_1 = 15.2 \quad x_2 = 9.3 \quad x_3 = 7.6 \quad x_4 = 11.9 \quad x_5 = 10.4 \quad x_6 = 9.7$$
$$x_7 = 20.4 \quad x_8 = 9.4 \quad x_9 = 11.5 \quad x_{10} = 16.2 \quad x_{11} = 9.4 \quad x_{12} = 8.3$$

La lista de valores ordenados es

7.6 8.3 9.3 9.4 9.4 9.7 10.4 11.5 11.9 15.2 16.2 20.4

Como $n = 12$ es par, se promedia $n/2 =$ valores ordenados sexto y séptimo:

$$\text{mediana de la muestra} = \frac{9.7 + 10.4}{2} = 10.05$$

Nótese que si la observación mayor, 20.4, no estuviera en la muestra, la mediana muestral resultante para las $n = 11$ observaciones habría sido el único valor medio, 9.7 (la $(n + 1)/2 =$ sexto valor ordenado). La media muestral es $\bar{x} = \sum x_i/n = 139.3/12 = 11.61$, que es un poco mayor que la mediana debido a los valores atípicos, 15.2, 16.2 y 20.4. ■

Los datos del ejemplo 1.14 ilustran una propiedad importante de $\tilde{x}$ en contraste con $\bar{x}$: la mediana muestral es bastante insensible a una cantidad de valores de datos muy pequeña o muy grande. Si, por ejemplo, se incrementan las dos x_i más grandes de 16.2 y 20.4 a 26.2 y 30.4, respectivamente, no resultaría afectada $\tilde{x}$. Así, en el tratamiento de valores atípicos de los datos, $\bar{x}$ y $\tilde{x}$ están en los extremos opuestos de un espectro: $\bar{x}$ es sensible incluso a uno de estos valores, en tanto que $\tilde{x}$ es insensible a una gran cantidad de valores atípicos.

Debido a que los valores grandes de la muestra del ejemplo 1.14 afectan a $\bar{x}$ más que a $\tilde{x}$, $\tilde{x} < \bar{x}$ para esos datos. Si bien $\bar{x}$ y $\tilde{x}$ proporcionan una medida para el centro de un conjunto de datos, en general no son iguales debido a que se concentran en aspectos diferentes de la muestra.

Análogo a $\tilde{x}$ como el valor medio de la muestra es un valor medio de la población, la **mediana poblacional**, denotada por $\tilde{\mu}$. Al igual que con $\bar{x}$ y μ, se puede pensar en usar la mediana muestral $\tilde{x}$ para hacer una inferencia acerca de $\tilde{\mu}$. En el ejemplo 1.14, se podría usar $\tilde{x} = 10.05$ como una estimación de la concentración mediana en la población total de la que se tomó la muestra. Una mediana suele emplearse para describir datos de ingreso o salario (debido a que en ella no influyen en gran medida pocos salarios grandes). Si el salario mediano para una muestra de ingenieros fuera $\tilde{x} = 66\,416$ dólares, se podría utilizar esto como base para concluir que el salario mediano para los ingenieros es mayor de 60 000 dólares.

La media μ y la mediana $\tilde{\mu}$ de la población en general no son idénticas. Si la distribución de la población tiene un sesgo positivo o negativo, como se observa en la figura 1.16, entonces $\mu \neq \tilde{\mu}$. Cuando éste sea el caso, al hacer inferencias se debe decidir primero cuál de las dos características poblacionales es de mayor interés, y luego se procede de conformidad con eso.

(a) Sesgo negativo (b) Simétrica (c) Sesgo positivo

Figura 1.16 Tres formas diferentes para una distribución poblacional

Otras medidas de localización: cuartiles, percentiles y medias recortadas

La mediana (poblacional o muestral) divide el conjunto de datos en dos partes de igual tamaño. Para obtener medidas de localización más finas, se podrían dividir los datos en más de dos partes. En términos generales, los cuartiles dividen el conjunto de datos en cuatro partes iguales, donde las observaciones arriba del tercer cuartil son el cuarto superior del conjunto de datos, el segundo cuartil es idéntico a la mediana y el primer cuartil separa el cuarto inferior de los tres cuartos superiores. De manera similar, un conjunto de datos (muestral o poblacional) se puede dividir incluso de manera más fina por medio de percentiles; el percentil 99 separa al 1% más alto del 99% restante, etc. A menos que el número de observaciones sea un múltiplo de 100, se debe tener cuidado al obtener percentiles. En el capítulo 4 se emplean percentiles en relación con ciertos modelos para poblaciones infinitas, así que se pospone el análisis hasta ese punto.

La media y la mediana de la muestra se ven afectadas por valores atípicos en una forma muy diferente: la media en gran medida y la mediana nada en absoluto. Puesto que podría ser indeseable el comportamiento extremo de cualquier tipo, se consideran de manera breve medidas opcionales que no son tan sensibles como $\bar{x}$ ni tan insensibles como $\tilde{x}$. Para motivar estas opciones, observe que $\bar{x}$ y $\tilde{x}$ están en extremos opuestos de la misma "familia" de medidas. Después de ordenar el conjunto de datos, $\tilde{x}$ se calcula descartando tantos valores de cada extremo como sea posible sin eliminar todo (dejando sólo uno o dos valores medios) y se promedia lo que queda. Por otro lado, para calcular $\bar{x}$ no se desecha nada antes de promediar. En otras palabras, en el cálculo de la media es necesario recortar 0% de cada extremo de la muestra, en tanto que para la mediana se recorta de cada extremo la cantidad máxima posible. Una **media recortada** es un término medio entre $\bar{x}$ y $\tilde{x}$. Una media recortada al 10%, por ejemplo, se calcularía eliminando 10% más pequeño y 10% más grande de la muestra y luego se promediaría el resto.

Ejemplo 1.15 Considere las 20 observaciones siguientes, ordenadas de menor a mayor, donde cada una representa la duración (en horas) de un cierto tipo de lámpara incandescente:

612	623	666	744	883	898	964	970	983	1003
1016	1022	1029	1058	1085	1088	1122	1135	1197	1201

El promedio de las 20 observaciones es $\bar{x} = 965.0$, en tanto que $\tilde{x} = 1009.5$. La media recortada al 10% se obtiene eliminando las dos observaciones más pequeñas (612 y 623) y las dos más grandes (1197 y 1201) y luego se promedian las 16 restantes para obtener $\bar{x}_{\text{rec}(10)} = 979.1$. El efecto de recortar aquí es obtener un "valor central" que está un poco arriba de la media (unos cuantos tiempos de vida pequeños reducen a $\bar{x}$) y, todavía, bastante abajo de la mediana. De manera similar, la media recortada al 20% promedia los 12 valores medios para obtener $\bar{x}_{\text{rec}(20)} = 999.9$, incluso más cerca de la mediana (véase la figura 1.17.)

Figura 1.17 Diagrama de puntos de los tiempos vida (en horas) de lámparas incandescentes ■

En términos generales, usar una media recortada con una proporción de recorte moderada (entre 5 y 25%), proporciona una medida que no es tan sensible a valores atípicos como la media ni tan insensible como la mediana. Por esta razón, las medias recortadas reciben cada vez más atención de los estadísticos, tanto para fines descriptivos como inferenciales. Se amplía la información acerca de las medias recortadas cuando se analiza la estimación puntual en el capítulo 6. Como punto final, si la proporción de recorte se denota por α y $n\alpha$ no es un entero, entonces no es obvio cómo se debe calcular la media recortada al $100\alpha\%$. Por ejemplo, si $\alpha = 0.10$ (10%) y $n = 22$, entonces $n\alpha = (22)(0.10) = 2.2$, y no es posible eliminar 2.2 observaciones de cada extremo de la muestra ordenada. En este caso, la media recortada al 10% se obtendría al eliminar primero dos observaciones de cada extremo y calcular $\bar{x}_{rec}$, luego eliminar tres y calcular $\bar{x}_{rec}$, y, por último, se interpola entre los dos valores para obtener $\bar{x}_{rec(10)}$.

Datos categóricos y proporciones muestrales

Cuando los datos son categóricos, una distribución de frecuencias o distribución de frecuencias relativas proporciona un resumen efectivo de los datos. Las cantidades de resumen numéricas, naturales, en esta situación son cada una de las frecuencias y las frecuencias relativas. Por ejemplo, si se lleva a cabo una encuesta para estudiar la preferencia de marcas de individuos que poseen receptores estéreo, entonces cada individuo de la muestra identificaría la marca del receptor que posee, de donde se podría contar el número de personas que poseen Sony, Marantz, Pioneer, etc. Considere el muestreo de una población dicotómica (una que consta de sólo dos categorías, por ejemplo, votó o no votó en las últimas elecciones, posee o no posee un receptor estéreo, etc.). Si x denota el número de observaciones de la muestra que cae en la categoría uno, entonces el número en la categoría dos es $n - x$. La frecuencia relativa o *proporción muestral* en la categoría uno es x/n y la proporción muestral en la categoría dos es $1 - x/n$. Sea 1 el número que denota una respuesta que cae en la categoría 1, y 0 el número para una respuesta que cae en la categoría dos. Un tamaño muestral de $n = 10$ produce entonces las respuestas, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1. La media muestral para esta muestra numérica es (puesto que el número de unos es $= x = 7$)

$$\frac{x_1 + \cdots + x_n}{n} = \frac{1 + 1 + 0 + \cdots + 1 + 1}{10} = \frac{7}{10} = \frac{x}{n} = \text{proporción muestral}$$

Este resultado se puede generalizar y resumir como sigue: *si en una situación de datos categóricos se centra la atención en una determinada categoría y se codifican los resultados muestrales de modo que se registre un 1 para un individuo que está en la categoría y un 0 para un individuo que no está en la categoría, entonces la proporción muestral de individuos en la categoría es la media muestral de la secuencia de unos y ceros.* Así que una media muestral se puede emplear para resumir los resultados de una muestra categórica. Estas observaciones también se aplican a situaciones en donde las categorías se definen al agrupar los valores

en una muestra o población numérica (p. ej., es posible que interese saber si las personas han tenido su automóvil actual por lo menos 5 años, en vez de estudiar desde cuando poseen el automóvil).

De manera análoga a la proporción muestral x/n de individuos que caen en una categoría en particular, sea p la proporción de individuos de la población total que caen en la categoría. Al igual que con x/n, p es una cantidad entre 0 y 1, y en tanto que x/n es una característica de la muestra, p es una característica de la población. La relación entre las dos se asemeja a la relación entre $\tilde{x}$ y $\tilde{\mu}$ y entre $\bar{x}$ y μ. En particular, de aquí en adelante se usa x/n para hacer inferencias acerca de p. Si, por ejemplo, de una muestra de 100 propietarios de automóvil se encuentra que 22 tuvieron su automóvil por lo menos 5 años, entonces se podría usar $22/100 = 0.22$ como estimación puntual de la proporción de propietarios que tuvieron su automóvil por lo menos 5 años. Más adelante se estudian las propiedades de x/n como un estimador de p y se ve cómo se puede usar x/n para responder otras preguntas inferenciales. Con k categorías ($k > 2$), se emplean las k proporciones muestrales para responder preguntas acerca de las proporciones poblacionales $p_1, \ldots, p_k$.

Ejercicios | Sección 1.3 (33 a 43)

33. En el artículo "The Pedaling Technique of Elite Endurance Cyclists" (*Int. J. of Sport Biomechanics*, 1991: 29-53) aparecen los siguientes datos de energía en una sola pierna con una alta carga de trabajo:

244 191 160 187 180 176 174
205 211 183 211 180 194 200

a. Calcule e interprete la media y la mediana de la muestra.

b. Suponga que la primera observación fue 204 y no 244. ¿Cómo cambiarían la media y la mediana?

c. Calcule una media recortada eliminando las observaciones mínima y máxima de la muestra. ¿Cuál es el porcentaje correspondiente de recorte?

d. En el artículo también se proporcionan los valores de la energía de una sola pierna para poca carga de trabajo. La media muestral para $n = 13$ observaciones fue $\bar{x} = 119.8$ (en realidad, 119.7692), y la observación decimocuarta, algo apartada, fue 159. ¿Cuál es el valor de $\bar{x}$ para la muestra total?

34. Considere las siguientes observaciones acerca de la resistencia al corte, en MPa, de una unión pegada de cierta manera (tomadas de una gráfica del artículo "Diffusion of Silicon Nitride to Austenitic Stainless Steel without Interlayers," *Metallurgical Trans.*, 1993: 1835-1843):

22.2 40.4 16.4 73.7 36.6 109.9
30.0 4.4 33.1 66.7 81.5

a. Determine el valor de la media muestral.

b. Determine el valor de la mediana muestral. ¿Por qué es tan diferente de la media?

c. Calcule una media recortada eliminando las observaciones menor y mayor. ¿Cuál es el porcentaje recortado correspondiente? ¿Cómo se compara este valor con la media y la mediana?

35. Se determinó la presión mínima de inyección (psi) para ocho muestras distintas de moldeo por inyección de maíz con alto contenido de amilosa (la presión mayor corresponde a mayor dificultad de procesamiento), y se obtuvieron las siguientes observaciones (de "Thermoplastic Starch Blends with a Polyethylene-Co-Vinyl Alcohol: Processability and Physical Properties", *Polymer Engr. and Science*, 1994: 17-23):

15.0 13.0 18.0 14.5 12.0 11.0 8.9 8.0

a. Determine los valores de la media, mediana y media recortada al 12.5%, y compárelos.

b. ¿Por cuánto se podría incrementar la observación mínima de la muestra, que en la actualidad es 8.0, sin afectar el valor de la mediana muestral?

c. Suponga que se quieren obtener los valores de la media y la mediana muestrales cuando las observaciones se expresan en kilogramos por pulgada cuadrada (ksi), y no en psi. ¿Es necesario expresar de nuevo cada observación en ksi, o se pueden usar en forma directa los valores calculados en el inciso (a)? *Sugerencia:* 1 kg = 2.2 lb.

36. Una muestra de 26 trabajadores petroleros de una plataforma marina tomaron parte en un simulacro de evacuación, y se obtuvieron los siguientes datos de tiempo (segundos) para terminar la evacuación ("Oxygen Consumption and Ventilation During Escape from an Offshore Platform", *Ergonomics,* 1997: 281-292):

389 356 359 363 375 424 325 394 402
373 373 370 364 366 364 325 339 393
392 369 374 359 356 403 334 397

a. Trace un diagrama de tallo y hojas con los datos. ¿De qué manera esto hace pensar que están correlacionadas la media y la mediana muestrales?
b. Calcule los valores de la media y la mediana muestrales. *Sugerencia:* $\sum x_i = 9638$.
c. ¿Por cuánto se podría aumentar el tiempo máximo, que actualmente es 424, sin afectar el valor de la mediana muestral? ¿En cuánto se podría disminuir este tiempo sin afectar el valor de la mediana muestral?
d. ¿Cuáles son los valores de $\bar{x}$ y $\tilde{x}$ cuando las observaciones se expresan en minutos?

37. En el artículo "Snow Cover and Temperature Relationships in North America and Eurasia" (*J. Climate and Applied Meteorology,* 1983: 460-469) se emplearon técnicas estadísticas para relacionar la cantidad de capa de nieve en cada continente con la temperatura continental promedio. Los datos presentados incluyen las siguientes diez observaciones de la capa de nieve en octubre para Eurasia, durante los años 1970-1979 (en millones de km^2):

6.5 12.0 14.9 10.0 10.7 7.9 21.9 12.5 14.5 9.2

¿Qué reportaría como valor representativo, o característico, de la capa de nieve en octubre para este periodo, y qué motivó su elección?

38. Es común informar los valores de tensión arterial hasta los 5 mmHg más cercanos (100, 105, 110, etc.). Suponga que los valores de tensión arterial reales para nueve individuos seleccionados al azar son

118.6 127.4 138.4 130.0 113.7 122.0 108.3 131.5 133.2

a. ¿Cuál es la mediana de los valores de tensión arterial reportados?
b. Suponga que la tensión arterial del segundo individuo es 127.6 en lugar de 127.4 (un pequeño cambio en un solo valor). ¿Cómo afecta esto a la mediana de los valores reportados? ¿Qué indica lo anterior acerca de la sensibilidad de la mediana para redondear o agrupar los datos?

39. La propagación de grietas por fatiga, en diversas partes de las aeronaves, ha sido objeto de estudio amplio en años recientes. Los datos que aparecen a continuación, consisten en tiempos de propagación (horas de vuelo/10^4) para alcanzar un determinado tamaño de grieta en agujeros de sujeción propios de aeronaves militares ("Statistical Crack Propagation in Fastener Holes under Spectrum Loading," *J. Aircraft,* 1983: 1028-1032):

0.736 0.863 0.865 0.913 0.915 0.937 0.983 1.007
1.011 1.064 1.109 1.132 1.140 1.153 1.253 1.394

a. Calcule y compare los valores de la media y la mediana muestrales.
b. ¿En cuánto se podría reducir la observación muestral más grande sin afectar el valor de la mediana?

40. Calcule la mediana muestral, la media recortada al 25%, la media recortada al 10% y la media muestral para los datos de concentración del ejercicio 27, y compare estas mediciones.

41. Se seleccionó una muestra de $n = 10$ automóviles, y cada uno se sometió a una prueba de choque a una velocidad de 5 millas por hora. Si se denota por S (éxito) un automóvil sin daños visibles y por F uno con daños, los resultados serían los siguientes:

S S F S S S F F S S

a. ¿Cuál es el valor de la proporción muestral de éxitos x/n?
b. Sustituya cada S con un 1 y cada F con un 0. Luego, calcule $\bar{x}$ para esta muestra codificada numéricamente. ¿Cómo se compara $\bar{x}$ con x/n?
c. Suponga que se decide incluir otros 15 automóviles en el experimento. ¿Cuántos de éstos tendrían que ser S para obtener $x/n = 0.80$ para la muestra total de 25 automóviles?

42. a. Si se agrega una constante c a cada x_i en una muestra, obteniéndose $y_i = x_i + c$, ¿cómo se relacionan la media y la mediana muestrales de las y_i con la media y la mediana de las x_i? Compruebe sus conjeturas.

b. Si cada x_i se multiplica por una constante c, con lo que se obtiene $y_i = cx_i$, conteste la pregunta del inciso (a). De nuevo, compruebe sus conjeturas.

43. En un experimento para estudiar la vida útil (en horas) de determinado componente, se pusieron en funcionamiento diez componentes y se tuvieron en observación por 100 horas. Durante este periodo fallaron ocho de los componentes, y se registraron los tiempos de vida. Denote los tiempos de vida de los dos componentes que aún funcionan después de 100 horas por 100+. Las observaciones muestrales obtenidas fueron

48 79 100+ 35 92 86 57 100+ 17 29

¿Cuál de las medidas de tendencia central analizadas en esta sección se pueden calcular? (*Nota:* se dice que los datos de este experimento están "censurados a la derecha".)

1.4 Medidas de variabilidad

Describir una medida de tendencia central sólo proporciona información parcial acerca de un conjunto de datos o distribución. Las distintas muestras o poblaciones pueden tener medidas idénticas de tendencia central, pero diferir entre sí en otros aspectos importantes. En la figura 1.18 se muestran los diagramas de puntos de tres muestras con la misma media y mediana; sin embargo, el grado de dispersión en torno al centro es distinto para las tres muestras. La primera muestra tiene la mayor cantidad de variabilidad, la tercera tiene la cantidad más pequeña y la segunda es intermedia con las otras dos en este respecto.

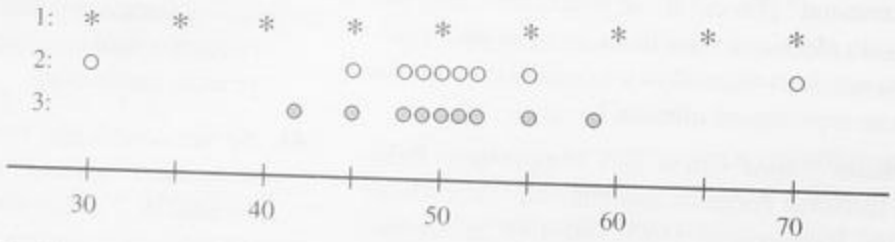


Figura 1.18 Muestras con medidas de tendencia central idénticas, pero cantidades distintas de variabilidad

Medidas de variabilidad para datos muestrales

La medida más sencilla de la variabilidad en una muestra es el **rango o recorrido**, que es la diferencia entre los valores máximo y mínimo en la muestra. Observe que el valor del recorrido para la muestra 1 en la figura 1.18 es mucho mayor que para la muestra 3, lo que refleja más variabilidad en la primera muestra que en la tercera. Sin embargo, una desventaja del recorrido es que depende sólo de las dos observaciones más extremas, y no toma en cuenta las posiciones de los $n - 2$ valores restantes. Las muestras 1 y 2, en la figura 1.18, tienen recorridos idénticos, pero cuando se toman en cuenta las observaciones entre los dos extremos, hay mucho menos variabilidad o dispersión en la segunda muestra que en la primera.

Las principales medidas de variabilidad tienen que ver con las **desviaciones a partir de la media,** $x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x}$. Es decir, las desviaciones respecto a la media se obtienen restando $\bar{x}$ de cada una de las n observaciones de la muestra. Una desviación es positiva si la observación es mayor que la media (está a la derecha de la media en el eje de medición), y

negativa si es menor que la media. Si es pequeña la magnitud de las desviaciones, entonces todas las x_i están cerca de la media y hay poca variabilidad. Por otro lado, si algunas de las desviaciones son de magnitud grande, entonces algunas de las x_i quedan lejos de $\bar{x}$, lo que indica una mayor cantidad de variabilidad. Una forma simple de combinar las desviaciones en una sola cantidad es promediarlas (sumarlas y dividirlas entre n). Sin embargo, hay un gran problema con esta sugerencia:

$$\text{suma de las desviaciones} = \sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

de modo que la desviación promedio siempre es cero. Para la comprobación se usan varias reglas de la suma y el hecho de que $\sum \bar{x} = \bar{x} + \bar{x} + \cdots + \bar{x} = n\bar{x}$:

$$\sum (x_i - \bar{x}) = \sum x_i - \sum \bar{x} = \sum x_i - n\bar{x} = \sum x_i - n\left(\frac{1}{n}\sum x_i\right) = 0$$

¿Cómo cambiar las desviaciones a cantidades no negativas para que las desviaciones positivas y negativas no se contrarresten entre sí cuando se combinan? Una posibilidad es trabajar con valores absolutos de las desviaciones y calcular la desviación absoluta promedio, $\sum |x_i - \bar{x}|/n$. Debido a que la operación del valor absoluto origina varias dificultades teóricas, considere en su lugar las desviaciones elevadas al cuadrado $(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \ldots, (x_n - \bar{x})^2$. En vez de usar la desviación promedio elevada al cuadrado, $\sum (x_i - \bar{x})^2/n$, por varias razones se divide mejor la suma de las desviaciones al cuadrado entre $n - 1$ y no entre n.

DEFINICIÓN

La **varianza muestral**, que se denota por s^2, se define por

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{S_{xx}}{n - 1}$$

La **desviación estándar muestral**, representada por s, es la raíz cuadrada (positiva) de la varianza:

$$s = \sqrt{s^2}$$

La unidad para s es igual que la de cada una de las x_i. Si, por ejemplo, las observaciones son rendimientos de gasolina, en millas por galón, entonces se podría tener que $s = 2.0$ mpg. Una interpretación aproximada de la desviación estándar muestral es que es el tamaño de una desviación característica, o representativa, con respecto a la media muestral dentro de la muestra. Así, si $s = 2.0$ mpg, entonces algunas de las x_i de la muestra se acercan más a $\bar{x}$, que 2.0, mientras que otras están bastante lejos; 2.0 es una desviación representativa (o "estándar") respecto al rendimiento medio de la gasolina. Si $s = 3.0$ para una segunda muestra de automóviles de otro tipo, una desviación característica en esta muestra es, más o menos, una y media veces la de la primera, una indicación de que en la segunda muestra hay más variabilidad.

Ejemplo 1.16 La resistencia es una característica importante de los materiales que se emplean en casas prefabricadas. Cada uno de $n = 11$ elementos de placa prefabricados se sometió a una prueba intensa de esfuerzo y se registró el ancho máximo (mm) de las grietas resultantes. Los datos de

Tabla 1.3 Datos para el ejemplo 1.16

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
0.684	−0.9841	0.9685
2.540	0.8719	0.7602
0.924	−0.7441	0.5537
3.130	1.4619	2.1372
1.038	−0.6301	0.3970
0.598	−1.0701	1.1451
0.483	−1.1851	1.4045
3.520	1.8519	3.4295
1.285	−0.3831	0.1468
2.650	0.9819	0.9641
1.497	−0.1711	0.0293
$\sum x_i = 18.349$	$\sum (x_i - \bar{x}) = -0.0001$	$S_{xx} = \sum (x_i - \bar{x})^2 = 11.9359$

$$\bar{x} = \frac{18.349}{11} = 1.6681$$

la tabla 1.3 aparecieron en el artículo "Prefabricated Ferrocement Ribbed Elements for Low-Cost Housing" (*J. Ferrocement*, 1984: 347–364). Los efectos de redondeo explican por qué la suma de las desviaciones no es exactamente cero. El numerador de s^2 es 11.9359; por lo tanto, $s^2 = 11.9359/(11 - 1) = 11.9359/10 = 1.19359$ y $s = \sqrt{1.19359} = 1.0925$ mm. ■

Motivación para s^2

Para explicar por qué para medir variabilidad se utiliza s^2 y no la desviación cuadrada promedio, observe primero que mientras s^2 mide variabilidad muestral, hay una medida de variabilidad en la población llamada *varianza poblacional*. Se utiliza σ^2 o (el cuadrado de la letra griega minúscula sigma) para denotar la varianza poblacional y σ para denotar la desviación estándar poblacional (la raíz cuadrada de σ^2). Cuando la población es finita y consta de N valores,

$$\sigma^2 = \sum_{i=1}^{N} (x_i - \mu)^2/N$$

que es el promedio de todas las desviaciones cuadradas con respecto a la media poblacional μ (para la población, el divisor es N y no $N - 1$). En los capítulos 3 y 4 se dan definiciones más generales de σ^2.

Del mismo modo que $\bar{x}$ se utiliza para hacer inferencias en relación con la media poblacional μ, se debe definir la varianza muestral, de modo que se pueda emplear para hacer inferencias con respecto a σ^2. Ahora, observe que σ^2 implica desviaciones cuadradas alrededor de la media poblacional μ. Si en realidad se conociera el valor de μ, entonces se podría definir la varianza muestral como la desviación cuadrada promedio de las x_i muestrales respecto a μ. Sin embargo, casi nunca se conoce el valor de μ así que se debe usar la suma de las desviaciones cuadradas respecto a $\bar{x}$. Pero, *las x_i tienden a aproximarse más a su promedio $\bar{x}$ que a la media poblacional μ, así que para compensar esto se utiliza el divisor $n - 1$ en lugar de n.* En otras palabras, si se utiliza un divisor n en la varianza muestral, entonces la cantidad resultante ten-

dería a subestimar σ^2 (se obtienen valores estimados que son demasiado pequeños en promedio), en tanto que dividir entre el valor ligeramente menor $n - 1$ corrige esta subestimación.

Es común mencionar que s^2 se basa en $n - 1$ **grados de libertad** (gl). Esta terminología resulta del hecho de que si bien s^2 se basa en n cantidades $x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x}$, éstas suman 0, así que especificar los valores de cualquier $n - 1$ de las cantidades determina el valor restante. Por ejemplo, si $n = 4$ y $x_1 - \bar{x} = 8$, $x_2 - \bar{x} = -6$ y $x_4 - \bar{x} = -4$, entonces se tiene que $x_3 - \bar{x} = 2$, así que sólo tres de los cuatro valores de $x_i - \bar{x}$ se determinan de manera libre (3 gl).

Fórmula para calcular s^2

Calcular y elevar al cuadrado las desviaciones puede ser tedioso, sobre todo si se está usando suficiente exactitud decimal en $\bar{x}$ para contrarrestar los efectos del redondeo. Una fórmula opcional para calcular el numerador de s^2 evita realizar todas las restas necesarias en el cálculo de las desviaciones. En la fórmula interviene tanto $(\sum x_i)^2$, sumar y después elevar al cuadrado, como $\sum x_i^2$, elevar al cuadrado y después sumar.

Una expresión opcional para el numerador de s^2 es

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

Demostración Debido a que $\bar{x} = \sum x_i/n$, $n\bar{x}^2 = (\sum x_i)^2/n$. Entonces,

$$\begin{aligned}\sum (x_i - \bar{x})^2 &= \sum (x_i^2 - 2\bar{x} \cdot x_i + \bar{x}^2) = \sum x_i^2 - 2\bar{x}\sum x_i + \sum (\bar{x})^2 \\ &= \sum x_i^2 - 2\bar{x} \cdot n\bar{x} + n(\bar{x})^2 = \sum x_i^2 - n(\bar{x})^2\end{aligned}$$

Ejemplo 1.17 La cantidad de luz reflejada en las hojas de las plantas se usa para varios fines, entre otros la evaluación del color del follaje, estimación del contenido de nitrógeno y medición de biomasa. En el artículo "Leaf Reflectance–Nitrogen–Chlorophyll Relations in Buffel-Grass" (*Photogrammetric Engr. and Remote Sensing*, 1985: 463-466) aparecen las siguientes observaciones, obtenidas por espectrofotogrametría, acerca de la reflexión bajo ciertas condiciones experimentales.

Observación	x_i	x_i^2	Observación	x_i	x_i^2
1	15.2	231.04	9	12.7	161.29
2	16.8	282.24	10	15.8	249.64
3	12.6	158.76	11	19.2	368.64
4	13.2	174.24	12	12.7	161.29
5	12.8	163.84	13	15.6	243.36
6	13.8	190.44	14	13.5	182.25
7	16.3	265.69	15	12.9	166.41
8	13.0	169.00			
				$\sum x_i = 216.1$	$\sum x_i^2 = 3168.13$

Con la fórmula ahora se obtiene

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 3168.13 - \frac{(216.1)^2}{15}$$
$$= 3168.13 - 3113.28 = 54.85$$

El método breve produce valores de s^2 y s que difieren de los valores calculados por medio de las definiciones. Estas diferencias se deben a efectos de redondeo, y no son importantes en la mayor parte de las muestras. Para reducir los efectos de redondeo cuando se utilice la fórmula breve, los cálculos intermedios se deben hacer con varios dígitos significativos más que se retienen en la respuesta final. Debido a que el numerador de s^2 es la suma de cantidades no negativas (desviaciones cuadradas), se garantiza que s^2 no es negativa. Sin embargo, si se utiliza el método breve, en particular con datos que tienen poca variabilidad, un ligero error numérico puede originar un numerador negativo ($\sum x_i^2$ menor que $(\sum x_i)^2/n$). Si el valor de s^2 es negativo, se ha cometido un error de cálculo.

Varias propiedades de s^2 facilitan su cálculo.

PROPOSICIÓN

Sea $x_1, x_2, \ldots, x_n$ una muestra y c cualquier constante diferente de cero.

1. Si $y_1 = x_1 + c, y_2 = x_2 + c, \ldots, y_n = x_n + c$, entonces $s_y^2 = s_x^2$, y
2. Si $y_1 = cx_1, \ldots, y_n = cx_n$, entonces $s_y^2 = c^2 s_x^2, s_y = |c| s_x$,

donde s_x^2 es la varianza muestral de las x y s_y^2 es la varianza muestral de las y.

En palabras, el resultado 1 dice que si se suma una constante c a (o se resta de) cada dato, la varianza no cambia. Esto es intuitivo, puesto que sumar o restar c cambia la ubicación del conjunto de datos, pero deja sin cambio las distancias entre los valores de los datos. Según el resultado 2, la multiplicación de cada x_i por c origina que s^2 sea multiplicada por un factor c^2. Estas propiedades se demuestran si se observa en el resultado 1 que $\bar{y} = \bar{x} + c$ y en el resultado 2 que $\bar{y} = c\bar{x}$.

Diagramas de caja

Los diagramas de tallo y hojas y los histogramas transmiten impresiones bastante generales acerca de un conjunto de datos, mientras que un resumen simple, como la media o la desviación estándar, se centra en sólo un aspecto de los datos. En años recientes se ha empleado con éxito un resumen gráfico, llamado *diagrama de caja*, para describir varias de las características más destacadas de un conjunto de datos. Entre estas características están (1) centro, (2) dispersión, (3) naturaleza y magnitud de cualquier desviación respecto a la simetría y (4) identificación de "valores atípicos", observaciones bastante alejadas del grueso de los datos. Debido a que incluso un solo punto atípico afecta de manera drástica los valores de $\bar{x}$ y s, un diagrama de caja se basa en medidas que son "resistentes" a la presencia de algunos valores atípicos: la mediana y una medida de dispersión llamada *cuarta dispersión*.

DEFINICIÓN

Ordene las n observaciones de menor a mayor y separe la mitad más pequeña de la mitad más grande; la mediana $\tilde{x}$ se incluye en ambas mitades si n es impar. Entonces, el **cuarto inferior** es la mediana de la mitad más pequeña y el **cuarto superior** es la mediana de la mitad más grande. Una medida de la dispersión que es resistente a los valores atípicos es la **cuarta dispersión** f_s, dada por

$$f_s = \text{cuarto superior} - \text{cuarto inferior}$$

En general, las posiciones de las observaciones en el 25% más pequeño o el 25% más grande de los datos no afectan a la cuarta dispersión.

El diagrama de caja más sencillo se basa en el siguiente resumen de cinco cantidades:

x_i mínima cuarto inferior mediana cuarto superior x_i máxima

Primero, trace una escala horizontal de medición. A continuación, coloque un rectángulo sobre este eje; el lado izquierdo del rectángulo está en el cuarto inferior y el derecho en el cuarto superior (así, ancho de la caja = f_s). Coloque un segmento de recta vertical o algún otro símbolo dentro del rectángulo en el lugar de la mediana; la posición del símbolo de la mediana en relación con los dos lados transmite información acerca del sesgo en el 50% intermedio de los datos. Por ultimo, trace unos "bigotes" desde cualquier extremo del rectángulo hasta las observaciones mínima y máxima. También se puede trazar un diagrama de caja con orientación vertical, haciendo las modificaciones obvias en el proceso de construcción.

Ejemplo 1.18 Por medio de ultrasonido se obtuvieron los datos de corrosión adjuntos en el espesor de la placa de piso de un tanque elevado para almacenar petróleo crudo ("Statistical Analysis of UT Corrosion Data from Floor Plates of a Crude Oil Aboveground Storage Tank", *Materials Eval.*, 1994: 846-849); cada observación es la máxima profundidad del hoyo en la placa, expresada en milésimas de pulgada.

40 52 55 60 70 75 85 85 90 90 92 94 94 95 98 100 115 125 125

El resumen de cinco números es el siguiente:

x_i mínima = 40 cuarto inferior = 72.5 $\tilde{x} = 90$ cuarto superior = 96.5
x_i máxima = 125

En la figura 1.19 se ilustra el diagrama de caja resultante. El lado derecho de esta caja se aproxima más a la mediana que el izquierdo, lo que indica un sesgo sustancial en la mitad intermedia de los datos. El ancho de la caja, f_s, también es razonablemente grande en relación con el intervalo de los datos (distancia entre las puntas de los bigotes).

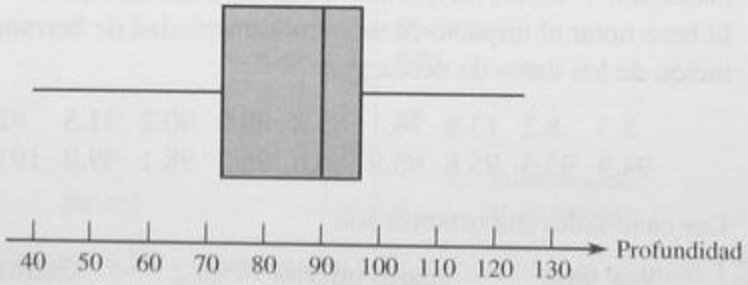


Figura 1.19 Diagrama de caja de los datos de corrosión

En la figura 1.20 se ilustra el resultado de MINITAB de una solicitud para describir los datos de corrosión. La media recortada es el promedio de las 17 observaciones que quedan después de eliminar los valores mínimo y máximo (porcentaje de recorte $\approx 5\%$). Q1 y Q3 son los cuartiles inferior y superior; éstos son similares a los cuartos, pero se calculan de una manera un poco distinta. El término *SE Mean* (error estándar promedio) es $s/\sqrt{n}$; ésta será una cantidad importante en el trabajo posterior en relación con las inferencias acerca de μ.

Variable	N	Mean	Median	TrMean	StDev	SE Mean
depth	19	86.32	90.00	86.76	23.32	5.35

Variable	Minimum	Maximum	Q1	Q3
depth	40.00	125.00	70.00	98.00

Figura 1.20 Descripción de MINITAB de los datos de profundidad de hoyos ■

Diagramas de caja que muestran valores atípicos

Se puede mejorar un diagrama de caja para indicar de manera explícita la presencia de valores atípicos.

DEFINICIÓN

Cualquier observación más allá de $1.5f_s$ desde el cuarto más cercano es un **valor atípico**. Un valor atípico es **extremo** si está a más de $3f_s$ del cuarto más cercano, y es **moderado** en cualquier otro caso.

Muchos procedimientos de inferencia se basan en la hipótesis de que la muestra provino de una distribución normal. Hasta un solo valor extremo en la muestra advierte al investigador que no se deben usar estos procedimientos, y la presencia de varios valores moderados conlleva el mismo mensaje.

Ahora se modifica la construcción anterior de un diagrama de caja, dibujando un bigote desde cada extremo de la caja hasta las observaciones mínima y máxima que no sean valores atípicos. Cada valor atípico moderado se representa con un círculo lleno, y cada valor atípico extremo con un círculo vacío. Algunos paquetes de cómputo estadístico no diferencian entre valores atípicos moderados y extremos.

Ejemplo 1.19 Los efectos de descargas parciales en la degradación del material de la cavidad de aislamiento tienen implicaciones importantes en la duración de los componentes de alto voltaje. Considere la siguiente muestra de $n = 25$ anchos de impulso de descargas lentas en una cavidad cilíndrica hecha de polietileno. (Estos datos concuerdan con un histograma de 250 observaciones que aparece en el artículo "Assessment of Dielectric Degradation by Ultrawide-band PD Detection", *IEEE Trans. on Dielectrics and Elec. Insul.*, 1995: 744-760.) El autor del artículo hace notar el impacto de una amplia variedad de herramientas estadísticas sobre la interpretación de los datos de descarga).

5.3 8.2 13.8 74.1 85.3 88.0 90.2 91.5 92.4 92.9 93.6 94.3 94.8
94.9 95.5 95.8 95.9 96.6 96.7 98.1 99.0 101.4 103.7 106.0 113.5

Las cantidades importantes son

$\tilde{x} = 94.8$ cuarto inferior = 90.2 cuarto superior = 96.7
$f_s = 6.5$ $1.5f_s = 9.75$ $3f_s = 19.50$

Así, cualquier observación menor que $90.2 - 9.75 = 80.45$ o mayor que $96.7 + 9.75 = 106.45$ es un valor atípico. Hay un valor de este tipo en el extremo superior de la muestra, y cuatro en el extremo inferior. Como $90.2 - 19.5 = 70.7$, las tres observaciones 5.3, 8.2 y 13.8 son valores extremos; los otros dos valores atípicos son moderados. Los bigotes se prolongan hasta 85.3 y 106.0, las observaciones más extremas que no son valores atípicos. El diagrama de caja resultante se muestra en la figura 1.21. Hay mucho sesgo negativo en la mitad intermedia de la muestra, así como en toda la muestra.

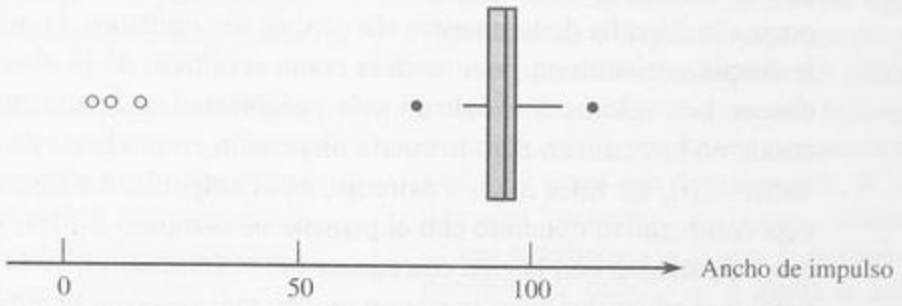


Figura 1.21 Diagrama de caja de los datos de ancho de impulso en el que se muestran los valores atípicos moderados y extremos ■

Diagramas de caja comparativos

Un diagrama de caja comparativo o lado a lado es una forma muy eficaz de mostrar semejanzas y diferencias entre dos o más conjuntos de datos que consisten en observaciones de la misma variable.

Ejemplo 1.20 En años recientes, ciertas evidencias hacen pensar que la alta concentración de radón en interiores puede estar relacionada con el desarrollo de cáncer infantil, pero muchos profesionales de la salud aún no están convencidos. Un artículo reciente ("Indoor Radon and Childhood Cancer", *The Lancet*, 1991: 1537-1538) presentó los siguientes datos acerca de la concentración de radón (Bq/m^3) en dos muestras diferentes de casas. La primera muestra consistió en casas en las que había estado residiendo un niño al que se le diagnosticó cáncer. Las casas de la segunda muestra no tenían casos registrados de cáncer infantil. En la figura 1.22 se muestra un diagrama de tallo y hojas para los datos.

1. Con cáncer		2. Sin cáncer
9683795	0	95768397678993
86071815066815233150	1	12271713114
12302731	2	99494191
8349	3	839
5	4	
7	5	55
	6	
	7	Tallo: decenas
HI: 210	8	5 Hoja: unidades

Figura 1.22 Diagrama de tallo y hojas para el ejemplo 1.20

A continuación se resumen las cantidades numéricas:

	$\bar{x}$	$\tilde{x}$	s	f_s
Con cáncer	22.8	16.0	31.7	11.0
Sin cáncer	19.2	12.0	17.0	18.0

Los valores de la media y la mediana hacen pensar que la muestra de cáncer está centrada un poco a la derecha de la muestra sin cáncer. Sin embargo, la media exagera la magnitud de este desplazamiento, en gran medida como resultado de la observación 210 en la muestra con cáncer. Los valores de s indican más variabilidad en la muestra con cáncer que en la muestra donde no hay cáncer, pero la cuarta dispersión contradice esta dispersión. De nuevo, la observación 210, un valor atípico extremo, es el culpable. La figura 1.23 muestra un diagrama de caja comparativo obtenido con el paquete de cómputo S-Plus. La caja sin cáncer está alargada en comparación con la caja con cáncer ($f_s = 18$ contra $f_s = 11$), y las posiciones de las líneas medianas en las dos cajas muestran mucho más sesgo en la mitad central de la muestra sin cáncer que en la muestra con cáncer. Los segmentos horizontales de recta representan los valores atípicos, y no hay distinción entre los valores atípicos moderados y extremos.

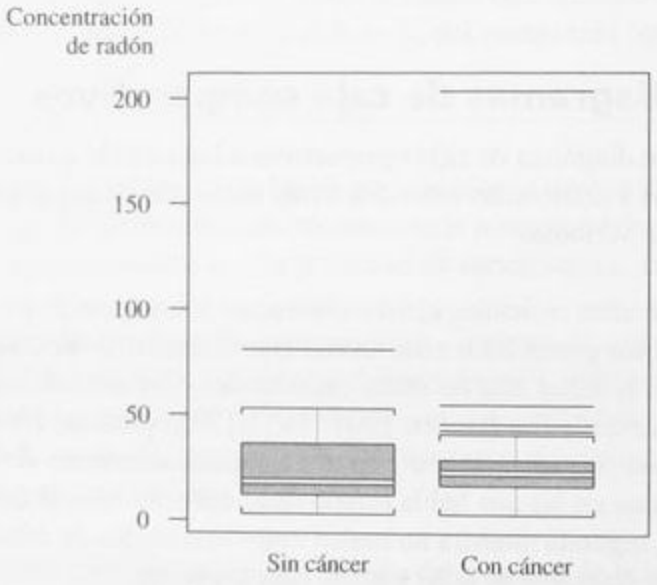


Figura 1.23 Diagrama de caja obtenido con S-Plus para los datos del ejemplo 1.20 ■

Ejercicios | Sección 1.4 (44 a 61)

44. En el artículo "Oxygen Consumption During Fire Suppression: Error of Heart Rate Estimation" (*Ergonomics*, 1991: 1469–1474) aparecen los datos siguientes en relación con el consumo de oxígeno (mL/kg/min) para una muestra de diez bomberos que realizan un simulacro de combate de incendio:

29.5 49.3 30.6 28.2 28.0 26.3 33.9 29.4 23.5 31.6

Calcule lo siguiente:

a. El intervalo de la muestra

b. La varianza muestral s^2 a partir de la definición (es decir, calcular primero las desviaciones, luego elevarlas al cuadrado, etcétera.)

c. La desviación estándar muestral

d. s^2 usando el método breve.

45. Se determinó el valor del módulo de elasticidad, en GPa, de placas coladas que constan de ciertos sustratos intermetálicos, y se obtuvieron las siguientes observaciones muestrales ("Strength and Modulus of a Molybdenum-Coated Ti-25Al-10Nb-3U-1Mo Intermetallic", *J. of Materials Engr. and Performance*, 1997: 46-50):

116.4 115.9 114.6 115.2 115.8

a. Calcule $\bar{x}$ y las desviaciones respecto a la media.
b. Use las desviaciones que calculó en el inciso (a) para obtener la varianza y la desviación estándar muestrales.
c. Calcule s^2 con la fórmula de cálculo para el numerador S_{xx}.
d. Reste 100 de cada observación para obtener una muestra de valores transformados. Ahora calcule la varianza muestral de los valores transformados y compárela con s^2 para los datos originales.

46. Las siguientes mediciones de viscosidad estabilizada (cP) para especímenes de determinado grado de asfalto con 18% de hule se tomaron del artículo "Viscosity Characteristics of Rubber-Modified Asphalts" (*J. of Materials in Civil Engr.*, 1996: 153-156):

2781 2900 3013 2856 2888

a. ¿Cuáles son los valores de la media y la mediana muestrales?
b. Calcule la varianza de la muestra con la fórmula de cálculo. (*Sugerencia:* primero reste un número adecuado a cada observación.)

47. Calcule e interprete los valores de la media y la desviación estándar de la muestra para las siguientes observaciones de resistencia a la fractura (en MPa, tomados de una gráfica en "Heat-Resistant Active Brazing of Silicon Nitride: Mechanical Evaluation of Braze Joints", *Welding J.*, agosto de 1997):

87 93 96 98 105 114 128 131 142 168

48. En el ejercicio 36 de la sección 1.3, se presentó una muestra de 26 tiempos de evacuación de trabajadores petroleros en un simulacro de emergencia. Calcule e interprete la desviación estándar de la muestra. (*Sugerencia:* $\sum x_i = 9638$ y $\sum x_i^2 = 3\,587\,566$.)

49. En un estudio de la relación entre la edad y varias funciones visuales, por ejemplo, agudeza y percepción de profundidad, se informan las siguientes observaciones sobre el área de la lámina esclerótica (mm^2) de cabezas de nervios ópticos humanos ("Morphometry of Nerve Fiber Bundle Pores in the Optic Nerve Head of the Human", *Experimental Eye Research*, 1988: 559-568):

2.75 2.62 2.74 3.85 2.34 2.74 3.93 4.21 3.88
4.33 3.46 4.52 2.43 3.65 2.78 3.56 3.01

a. Calcule $\sum x_i$ y $\sum x_i^2$.
b. Utilice los valores calculados en el inciso (a) para determinar la varianza s^2 y la desviación estándar s de la muestra.

50. En 1997, una mujer demandó a una compañía fabricante de teclados para computadora, con el cargo de que el teclado era el causante de las lesiones repetitivas por fatiga (*Genessy v. Digital Equipment Corp.*). Se adjudicó una cantidad de alrededor de 3.5 millones de dólares por dolor y sufrimiento a causa de la lesión, pero la corte revocó el fallo por ser una compensación poco razonable. Al tomar esta determinación, la corte identificó un grupo "normativo" de 27 casos similares y especificó una recompensa razonable como una dentro de dos desviaciones estándar de la media de las recompensas en los 27 casos. Las 27 recompensas fueron (en miles de dólares) 37, 60, 75, 115, 135, 140, 149, 150, 238, 290, 340, 410, 600, 750, 750, 750, 1050, 1100, 1139, 1150, 1200, 1200, 1250, 1576, 1700, 1825 y 2000, de donde $\sum x_i = 20\,179$, $\sum x_i^2 = 24\,657\,511$. ¿Cuál es la cantidad máxima posible que se podría otorgar bajo la regla de dos desviaciones estándar?

51. En el artículo "A Thin-Film Oxygen Uptake Test for the Evaluation of Automotive Crankcase Lubricants" (*Lubric. Engr.*, 1984: 75-83) aparecen los siguientes datos del tiempo de oxidación-inducción (en minutos) para varios aceites comerciales:

87 103 130 160 180 195 132 145 211 105 145
153 152 138 87 99 93 119 129

a. Calcule la varianza muestral y la desviación estándar.
b. Si las observaciones se volvieran a expresar en horas, ¿cuáles serían los valores resultantes de la varianza y la desviación estándar de la muestra? Conteste sin volver a expresar en realidad las observaciones.

52. Las primeras cuatro desviaciones con respecto a la media en una muestra de $n = 5$ tiempos de reacción fueron 0.3, 0.9, 1.0 y 1.3. ¿Cuál es la quinta desviación respecto a la media? Dé una muestra para la cual éstas sean las cinco desviaciones en relación con la media.

53. Considere de nuevo los datos del ejercicio 49 acerca del área de lámina esclerótica.
a. Determine los cuartos inferior y superior.
b. Calcule el valor de la cuarta dispersión.
c. Si los dos valores más grandes de la muestra, 4.33 y 4.52 hubieran sido 5.33 y 5.52, ¿cómo afectaría esto a f_s? Explique.

d. ¿Cuánto se puede aumentar la observación 2.34 sin afectar f_s? Explique.

e. Si se agrega a la muestra una decimoctava observación, $x_{18} = 4.60$, ¿cuál es el valor de f_s?

54. Considere de nuevo las observaciones de resistencia al corte (en MPa) que se mencionaron en el ejercicio 34 de este capítulo:

22.2	40.4	16.4	73.7	36.6	109.9
30.0	4.4	33.1	66.7	81.5	

a. ¿Cuáles son los valores de los cuartos, y cuál es el valor de f_s?

b. Trace un diagrama de caja con base en el resumen de cinco números, y comente acerca de sus características.

c. ¿Qué tan grande o pequeña tiene que ser una observación para considerarla como valor atípico? ¿Como valor atípico extremo?

d. ¿En cuánto se podría disminuir la observación más grande sin afectar a f_s?

55. El siguiente diagrama de tallo y hojas se construyó con los tiempos de evasión mostrados en el ejercicio 36 de este capítulo.

32	55
33	49
34	
35	6699
36	34469
37	03345
38	9
39	2347
40	23
41	
42	4

a. Determine el valor de la cuarta dispersión.

b. ¿Hay algunos valores atípicos en la muestra? ¿Algún valor atípico extremo?

c. Trace un diagrama de caja y comente acerca de sus características.

d. ¿En cuánto se podría disminuir la observación más grande, actualmente 424, sin afectar el valor de la cuarta dispersión?

56. Se determinó la cantidad de aluminio contaminante (ppm) en un determinado tipo de plástico para una muestra de 26 muestras, y se obtuvieron los datos siguientes ("The Lognormal Distribution for Modeling Quality Data when the Mean Is Near Zero", *J. of Quality Technology*, 1990: 105-110):

30	30	60	63	70	79	87	90	101
102	115	118	119	119	120	125	140	145
172	182	183	191	222	244	291	511	

Trace un diagrama de caja que muestre los valores atípicos y comente acerca de sus características.

57. Se seleccionó una muestra de 20 botellas de vidrio de determinado tipo, y se midió la resistencia de cada una a la presión interna. Tome en cuenta la siguiente información parcial de la muestra:

mediana = 202.2 cuarto inferior = 196.0
cuarto superior = 216.8

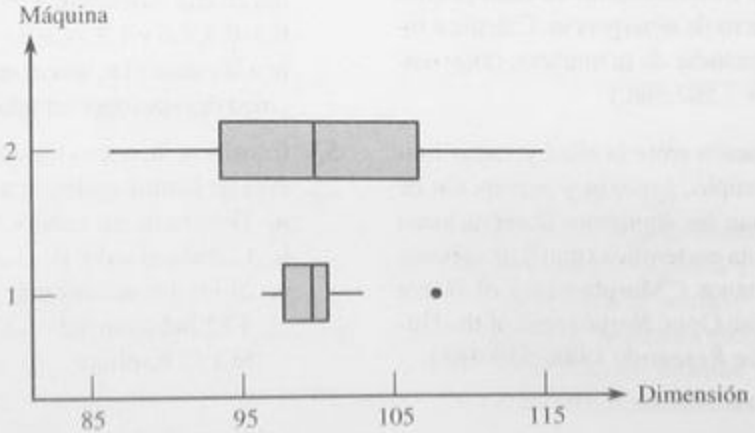
Tres observaciones mínimas	125.8	188.1	193.7
Tres observaciones máximas	221.3	230.5	250.2

a. ¿Hay valores atípicos en la muestra? ¿Hay valores atípicos extremos?

b. Construya un diagrama de caja que muestre los valores atípicos y comente acerca de características interesantes.

58. En una empresa se usan dos máquinas distintas para fabricar partes de cierto tipo. Durante un solo turno, se obtuvo una muestra de $n = 20$ partes que produce cada máquina, se determinó el valor de una dimensión crítica determinada. Con estos datos se construyó el diagra-

Diagrama de caja comparativo para el ejercicio 58



ma de caja comparativo. Compare y describa las diferencias de las dos muestras.

59. Se determinó la concentración de cocaína en la sangre (mg/L) en una muestra de individuos que murieron de delirio excitado (DE) inducido por cocaína y en una muestra de quienes murieron por sobredosis de cocaína sin delirio excitado; el tiempo de supervivencia en ambos grupos fue a lo sumo seis horas. Los datos adjuntos se obtuvieron de un diagrama de caja deel artículo "Fatal Excited Delirium Following Cocaine Use" (*J. of Forensic Sciences*, 1997: 25-31).

DE 0 0 0 0 0.1 0.1 0.1 0.1 0.2 0.2 0.3 0.3
0.3 0.4 0.5 0.7 0.8 1.0 1.5 2.7 2.8
3.5 4.0 8.9 9.2 11.7 21.0

Sin *DE* 0 0 0 0 0 0.1 0.1 0.1 0.1 0.2 0.2 0.2
0.3 0.3 0.3 0.4 0.5 0.5 0.6 0.8 0.9 1.0
1.2 1.4 1.5 1.7 2.0 3.2 3.5 4.1
4.3 4.8 5.0 5.6 5.9 6.0 6.4 7.9
8.3 8.7 9.1 9.6 9.9 11.0 11.5
12.2 12.7 14.0 16.6 17.8

a. Determine las medianas, los cuartos y las cuartas dispersiones en las dos muestras.

b. ¿Hay valores atípicos en alguna muestra? ¿Hay valores atípicos extremos?

c. Trace un diagrama de caja comparativo y utilícelo como base para comparar y confrontar las muestras con DE y sin DE.

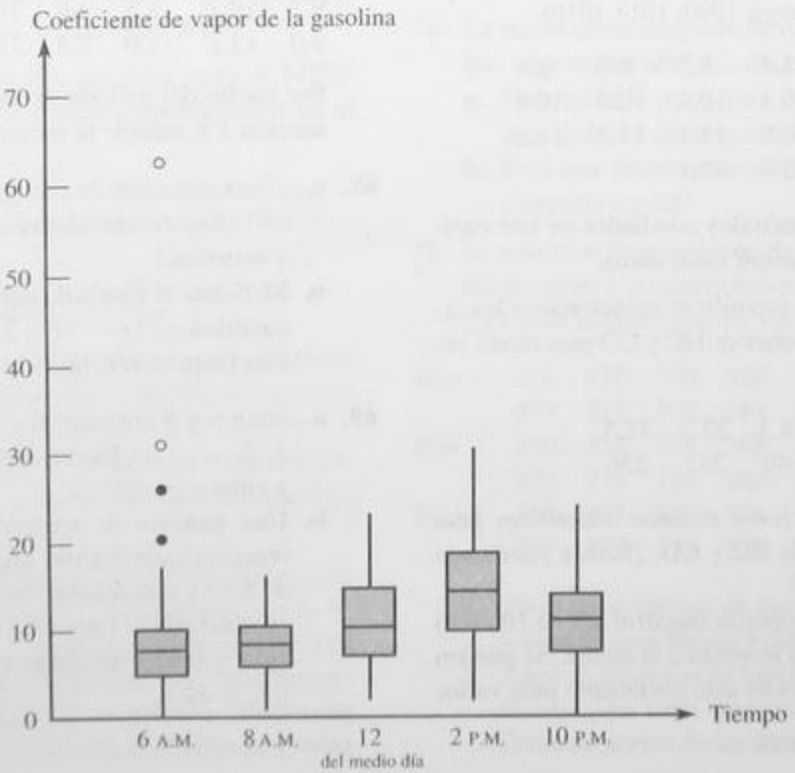
60. Se realizaron mediciones de resistencia a la explosión (lb/pulg^2) para soldaduras de cierre de toberas de prueba y para soldaduras de toberas de recipientes de producción ("Proper Procedures Are the Key to Welding Radioactive Waste Cannisters", *Welding J.*, agosto de 1997: 61-67).

Prueba	7200	6100	7300	7300	8000	7400
	7300	7300	8000	6700	8300	
Recipiente	5250	5625	5900	5900	5700	6050
	5800	6000	5875	6100	5850	6600

Trace un diagrama de caja comparativo y comente acerca de sus características interesantes. (En el artículo citado no se incluye la figura, pero los autores dijeron haber examinado una.)

61. El siguiente diagrama de caja comparativo de los coeficientes de vapor de gasolina para vehículos de Detroit apareció en el artículo "Receptor Modeling Approach to VOC Emission Inventory Validation" (*J. of Envir. Engr.*, 1995: 483-490). Describa cualquier característica interesante.

Diagrama de caja comparativo para el ejercicio 61



Ejercicios complementarios (62 a 83)

62. Examine la siguiente información acerca de la resistencia última a la tensión (lb/pulg) para una muestra de $n = 4$ especímenes de alambre duro de cobre y circonio (tomado de "Characterization Methods for Fine Copper Wire", *Wire J. Intl.*, agosto de 1997: 74-80):

$\bar{x} = 76\,831$ $s = 180$ mínima $x_i = 76\,683$
máxima $x_i = 77\,048$

Determine los valores de las dos observaciones intermedias de la muestra (¡no lo haga por tanteo!).

63. Se examinaron tres flujos de C_2F_6 diferentes en un experimento para investigar el efecto del flujo en la uniformidad (%) del grabado al ácido en una oblea de silicio utilizada en la fabricación de circuitos integrados, y se obtuvieron los datos siguientes:

Flujo						
125	2.6	2.7	3.0	3.2	3.8	4.6
160	3.6	4.2	4.2	4.6	4.9	5.0
200	2.9	3.4	3.5	4.1	4.6	5.1

Compare las observaciones de uniformidad que resultan de estos tres diferentes flujos.

64. La cantidad de radiación recibida en un invernadero juega un papel importante en la determinación de rapidez de la fotosíntesis. Las observaciones siguientes sobre radiación solar entrante se leyeron de una gráfica del artículo "Radiation Components over Bare and Planted Soils in a Greenhouse" (*Solar Energy*, 1990: 1011-1016).

6.3	6.4	7.7	8.4	8.5	8.8	8.9
9.0	9.1	10.0	10.1	10.2	10.6	10.6
10.7	10.7	10.8	10.9	11.1	11.2	11.2
11.4	11.9	11.9	12.2	13.1		

Utilice algunos de los métodos estudiados en este capítulo para describir y resumir estos datos.

65. En la introducción del capítulo se mencionaron los siguientes datos de emisiones de HC y CO para cierto vehículo.

HC (gm/mi)	13.8	18.3	32.2	32.5
CO (gm/mi)	118	149	232	236

a. Calcule las desviaciones estándar muestrales para las observaciones de HC y CO. ¿Parece justificada la creencia general?

b. El *coeficiente de variación muestral* $s/\bar{x}$ (o $100\,s/\bar{x}$) mide la variabilidad respecto a la media. Se pueden comparar los valores de este coeficiente para varios conjuntos de datos a fin de determinar cuáles presentan mayor o menor variación. Realice esta comparación con los datos del problema.

66. La distribución adjunta de frecuencias en mediciones de resistencia a la fractura (en MPa) para barras de cerámica quemadas en un horno apareció en el artículo "Evaluating Tunnel Kiln Performance" (*Amer. Ceramic Soc. Bull.*, agosto de 1997: 59-63).

Clase	81–<83	83–<85	85–<87	87–<89	89–<91
Frecuencia	6	7	17	30	43
Clase	91–<93	93–<95	95–<97	97–<99	
Frecuencia	28	22	13	3	

a. Trace un histograma con base en las frecuencias relativas y comente acerca de algunas propiedades interesantes.

b. ¿Qué proporción de las observaciones son por lo menos 85? ¿Menores de 95?

c. ¿Más o menos qué proporción de las mediciones fueron menores que 90?

67. Se obtuvieron quince muestras de aire de una cierta región, y para cada una se determinó la concentración de monóxido de carbono. Los resultados (en ppm) fueron

9.3	10.7	8.5	9.6	12.2	15.6	9.2	10.5
9.0	13.2	11.0	8.8	13.7	12.1	9.8	

Por medio del método de interpolación sugerido en la sección 1.3, calcule la media recortada al 10%.

68. **a.** ¿Para qué valor de c se minimiza la cantidad $\sum(x_i - c)^2$? (*Sugerencia:* derive con respecto a c, iguale a 0 y resuelva.)

b. Mediante el resultado del inciso (a), ¿cuál de las dos cantidades $\sum(x_i - \bar{x})^2$ y $\sum(x_i - \mu)^2$ es menor que la otra (suponiendo que $\bar{x} \neq \mu$)?

69. **a.** Sean a y b constantes y sea $y_i = ax_i + b$ para $i = 1, 2, \ldots, n$. ¿Cuáles son las relaciones entre $\bar{x}$ y $\bar{y}$ y entre s_x^2 y s_y^2?

b. Una muestra de temperaturas para iniciar cierta reacción química dio una promedio muestral (°C) de 87.3 y una desviación estándar muestral de 1.04. ¿Cuáles son el promedio de la muestra y la desviación estándar medidas en °F? (*Sugerencia:* $F = \frac{9}{5}C + 32$.)

70. El elevado consumo de energía durante el ejercicio continúa después que termina el trabajo físico. Debido a que las calorías quemadas después del ejercicio contribuyen a la pérdida de peso y tienen otras consecuencias, es importante entender este proceso. En el artículo "Effect of Weight Training Exercise and Treadmill Exercise on Post-Exercise Oxygen Consumption" (*Medicine and Science in Sports and Exercise,* 1998: 518-522) aparecen los datos adjuntos de un estudio en que se midió de forma continua el consumo de oxígeno (litros) durante 30 minutos para cada uno de 15 individuos después de un ejercicio de entrenamiento con pesas y después de un ejercicio en una cinta rodante.

Individuo	1	2	3	4	5	6	7	8	9
	10	11	12	13	14	15			
Pesas (x)	14.6	14.4	19.5	24.3	16.3	22.1			
	23.0	18.7	19.0	17.0	19.1	19.6			
	23.2	18.5	15.9						
Cinta rodante (y)	11.3	5.3	9.1	15.2	10.1	19.6			
	20.8	10.3	10.3	2.6	16.6	22.4			
	23.6	12.6	4.4						

a. Trace un diagrama de caja comparativo de las observaciones para el ejercicio con pesas y cinta rodante, y comente acerca de lo que considere importante.

b. Debido a que los datos están en la forma de pares (x, y) con las mediciones x y y en la misma variable en dos condiciones distintas, es natural centrarse en diferencias dentro de los pares: $d_1 = x_1 - y_1, \ldots, d_n = x_n - y_n$. Construya un diagrama de caja de las diferencias muestrales. ¿Qué indica esto?

71. A continuación se da una descripción a partir de MINITAB de los datos de resistencia proporcionados en el ejercicio 13.

```
Variable       N     Mean  Median  TrMean  StDev  SE Mean
strength     153   135.39  135.40  135.41   4.59     0.37

Variable  Minimum  Maximum      Q1      Q3
strength   122.20   147.70  132.95  138.25
```

a. Haga comentarios acerca de algunas características interesantes (en este caso los cuartiles y los cuartos son casi idénticos).

b. Construya un diagrama de caja de los datos con base en los cuartiles y comente acerca de características importantes.

72. Los trastornos y síntomas de ansiedad suelen tratarse de manera eficaz con medicaciones de benzodiazepina. Se sabe que los animales expuestos al estrés exhiben una disminución en el enlace del receptor de bezodiazepinas en la corteza frontal. En el artículo "Decreased Benzodiazepine Receptor Binding in Prefrontal Cortex in Combat-Related Posttraumatic Stress Disorder" (*Amer. J. of Psychiatry,* 2000: 1120-1126) se describe el primer estudio de enlace de receptor de benzodiazepina en individuos con trastorno de estrés postraumático (PTSD). Los datos adjuntos acerca de una medida del enlace del receptor (volumen de distribución ajustado) se obtuvieron de una gráfica del artículo.

PTSD: 10, 20, 25, 28, 31, 35, 37, 38, 38, 39, 39, 42, 46

Saludable: 23, 39, 40, 41, 43, 47, 51, 58, 63, 66, 67, 69, 72

Utilice los distintos métodos de este capítulo para describir y resumir los datos.

73. En el artículo "Can We Really Walk Straight?" (*Amer. J. of Physical Anthropology,* 1992: 19-27) se describe un experimento en el que se pidió a cada uno de 20 hombres sanos caminar tan derecho como fuera posible hasta un lugar a 60 m a velocidad normal. Considere las siguientes observaciones en relación con la cadencia (número de pasos por segundo):

0.95 0.85 0.92 0.95 0.93 0.86 1.00 0.92 0.85 0.81
0.78 0.93 0.93 1.05 0.93 1.06 1.06 0.96 0.81 0.96

Utilice los métodos de este capítulo para resumir los datos; incluya una interpretación, o análisis, siempre que sea apropiado. (*Nota:* el autor del artículo utilizó un análisis estadístico complejo para llegar a la conclusión de que las personas no pueden caminar en línea recta y sugirió varias explicaciones para esto.)

74. La **moda** de un conjunto de datos numéricos es el valor que se presenta con más frecuencia en el conjunto.

a. Determine la moda para el conjunto de datos de cadencia dados en el ejercicio 73.

b. Para una muestra categórica, ¿cómo se definiría la categoría modal?

75. Se seleccionaron muestras de tres tipos distintos de cable de acero y se determinó el límite de fatiga, en MPa, para cada muestra con los resultados siguientes:

Tipo 1	350	350	350	358	370	370	370	371
	371	372	372	384	391	391	392	
Tipo 2	350	354	359	363	365	368	369	371
	373	374	376	380	383	388	392	
Tipo 3	350	361	362	364	364	365	366	371
	377	377	377	379	380	380	392	

a. Trace un diagrama de caja comparativo y comente acerca de semejanzas y diferencias.

b. Trace un diagrama de caja comparativo (un diagrama de caja para cada muestra con una escala común). Comente acerca de las semejanzas y diferencias.

c. ¿El diagrama de caja comparativo del inciso (a) proporciona una evaluación informativa de las semejanzas y diferencias? Explique por qué.

76. Las tres medidas de tendencia central presentadas en este capítulo son media, mediana y media recortada. Dos medidas más de este tipo que se utilizan a veces son el *recorrido medio*, que es el promedio de las observaciones mínima y máxima, y *el cuarto medio*, que es el promedio de los dos cuartos. ¿Cuáles de estas cinco medidas de tendencia central son resistentes a los efectos de los valores atípicos, y cuáles no? Explique por qué.

77. Considere los siguientes datos sobre tiempo activo de reparación (horas) para una muestra de $n = 46$ receptores de comunicaciones transmitidas por aire:

0.2	0.3	0.5	0.5	0.5	0.6	0.6	0.7	0.7	0.7	0.8	0.8
0.8	1.0	1.0	1.0	1.0	1.1	1.3	1.5	1.5	1.5	1.5	2.0
2.0	2.2	2.5	2.7	3.0	3.0	3.3	3.3	4.0	4.0	4.5	4.7
5.0	5.4	5.4	7.0	7.5	8.8	9.0	10.3	22.0	24.5		

Construya lo siguiente:

a. Un diagrama de tallo y hojas en la que los dos valores más grandes se muestren por separado en un renglón identificado con HI

b. Un histograma con base en seis intervalos de clase con 0 como límite inferior del primer intervalo y anchos de intervalo de 2, 2, 2, 4, 10 y 10, respectivamente.

78. Considere una muestra $x_1, x_2, \ldots, x_n$ y suponga que ya se calcularon los valores de $\bar{x}$, s^2 y s.

a. Sea $y_i = x_i - \bar{x}$ para $i = 1, \ldots, n$. ¿Cómo se comparan los valores de s^2 y s para las y_i con los valores correspondientes para las x_i? Explique.

b. Sea $z_i = (x_i - \bar{x})/s$ para $i = 1, \ldots, n$. ¿Cuáles son los valores de la varianza y desviación estándar de la muestra para las z_i?

79. Sean $\bar{x}_n$ y s_n^2 la media y la varianza para la muestra $x_1, \ldots, x_n$ y sean $\bar{x}_{n+1}$ y s_{n+1}^2 estas cantidades cuando se añade a la muestra una observación adicional x_{n+1}.

a. Demuestre cómo se puede calcular $\bar{x}_{n+1}$ a partir de $\bar{x}_n$ y x_{n+1}.

b. Demuestre que

$$ns_{n+1}^2 = (n-1)s_n^2 + \frac{n}{n+1}(x_{n+1} - \bar{x}_n)^2$$

de modo que s_{n+1}^2 se pueda calcular de x_{n+1}, $\bar{x}_n$, y s_n^2.

c. Suponga que una muestra de 15 hebras de hilo de paño dio como resultado una elongación media del hilo de la muestra de 12.58 mm y una desviación estándar muestral de 0.512 mm. Una decimosexta hebra da un valor de elongación de 11.8. ¿Cuáles son los valores de la media y la desviación estándar muestrales para las 16 observaciones de elongación?

80. Las longitudes de las rutas de autobús para un sistema de tránsito particular varían por lo común de una ruta a otra. En el artículo "Planning of City Bus Routes" (*J. of the Institution of Engineers*, 1995: 211-215) aparece la siguiente información acerca de las longitudes (en km) para un determinado sistema:

Longitud	6– <8	8– <10	10– <12	12– <14	14– <16
Frecuencia	6	23	30	35	32
Longitud	16– <18	18– <20	20– <22	22– <24	24– <26
Frecuencia	48	42	40	28	27
Longitud	26– <28	28– <30	30– <35	35– <40	40– <45
Frecuencia	26	14	27	11	2

a. Trace el histograma correspondiente a estas frecuencias.

b. ¿Qué proporción de estas longitudes de ruta son menores que 20? ¿Qué proporción de estas rutas tienen longitudes de por lo menos 30?

c. ¿Más o menos cuál es el valor del percentil 90 de la distribución de longitudes de ruta?

d. De manera aproximada, ¿cuál es la longitud de ruta mediana?

81. En un estudio realizado para investigar la distribución del tiempo de frenado total (tiempo de reacción más tiempo para pasar del acelerador al freno, en ms) en condiciones reales de manejo a 60 km/h se obtuvo el siguiente resumen de la distribución de los tiempos ("A Field Study on Braking Responses during Driving," *Ergonomics*, 1995: 1903-1910):

media = 535 mediana = 500 moda = 500
Desv. Est. = 96 mínimo = 220 máximo = 925

percentil 5 = 400 percentil 10 = 430
percentil 90 = 640 percentil 95 = 720

¿Qué se concluye en relación con la forma de un histograma de estos datos? Explique su razonamiento.

82. Los datos muestrales $x_1, x_2, \ldots, x_n$ representan a veces una **serie de tiempo**, donde x_t = valor observado de una variable x de respuesta en el tiempo t. Con frecuencia, la serie observada presenta una gran cantidad de variación aleatoria, lo que dificulta estudiar el comportamiento a largo plazo. En estos casos, es conveniente producir una

versión suavizada de la serie. Una técnica para llevar a cabo lo anterior utiliza **suavizamiento exponencial**. Se escoge el valor de una constante de suavizamiento α $(0 < \alpha < 1)$. Entonces, con $\bar{x}_t$ = valor suavizado en el tiempo t, se establece $\bar{x}_1 = x_1$, y para $t = 2, 3, \ldots, n$, $\bar{x}_t = \alpha x_t + (1 - \alpha)\bar{x}_{t-1}$.

a. Considere la siguiente serie de tiempo en la cual x_t = temperatura (°F) de emanaciones en una planta de tratamiento de aguas negras en el día t: 47, 54, 53, 50, 46, 46, 47, 50, 51, 50, 46, 52, 50, 50. Trace la gráfica de cada x_t en función de t en un sistema coordenado de dos dimensiones (gráfica de serie de tiempo). ¿Aparece algún patrón?

b. Calcule las $\bar{x}_t$ usando $\alpha = 0.1$. Repita el cálculo con $\alpha = 0.5$. ¿Con cuál valor de α se obtiene una serie más suavizada?

c. Sustituya $\bar{x}_{t-1} = \alpha x_{t-1} + (1 - \alpha)\bar{x}_{t-2}$ en el lado derecho de la ecuación para $\bar{x}_t$, luego sustituya $\bar{x}_{t-2}$ en términos de x_{t-2} y $\bar{x}_{t-3}$, etc. ¿De cuántos de los valores $x_t, x_{t-1}, \ldots, x_1$ depende $\bar{x}_t$? ¿Qué sucede con el coeficiente en x_{t-k} cuando aumenta k?

d. Refiérase al inciso (c). Si t es grande, ¿qué tan sensible es $\bar{x}_t$ a la inicialización $\bar{x}_1 = x_1$? Explique.

(*Nota:* una referencia importante es el artículo "Simple Statistics for Interpreting Environmental Data", *Water Pollution Control Fed. J.*, 1981: 167-175.)

83. Considere las observaciones numéricas $x_1, \ldots, x_n$. A menudo interesa saber si las x_i están (al menos de manera aproximada) distribuidas en forma simétrica respecto a algún valor. Si n es por lo menos moderadamente grande, la magnitud de la simetría se puede evaluar a partir de un diagrama de tallo y hojas o un histograma. Sin embargo, si n no es muy grande, las figuras no contienen mucha información. Considere la siguiente opción. Sea y_1 la x_i más pequeña, y_2 la segunda x_i más pequeña, etc. Luego, trace la gráfica de los siguientes pares como puntos en un sistema coordenado de dos dimensiones: $(y_n - \tilde{x}, \tilde{x} - y_1)$, $(y_{n-1} - \tilde{x}, \tilde{x} - y_2)$, $(y_{n-2} - \tilde{x}, \tilde{x} - y_3)$, ... Hay $n/2$ puntos cuando n es par y $(n - 1)/2$ cuando n es impar.

a. ¿Qué aspecto tiene esta gráfica cuando existe simetría perfecta en los datos? ¿Qué aspecto tiene cuando las observaciones se alargan más arriba de la mediana que abajo de ésta (un extremo superior largo)?

b. Los siguientes datos sobre lluvias (acre-pies) de 26 nubes "sembradas" se tomaron del artículo "A Bayesian Analysis of a Multiplicative Treatment Effect in Weather Modification" (*Technometrics*, 1975: 161-166). Construya la gráfica y comente acerca del grado de simetría o naturaleza de la desviación con respecto a la simetría.

4.1	7.7	17.5	31.4	32.7	40.6	92.4
115.3	118.3	119.0	129.6	198.6	200.7	242.5
255.0	274.7	274.7	302.8	334.1	430.0	489.1
703.4	978.0	1656.0	1697.8	2745.6		

Bibliografía

Chambers, John, William Cleveland, Beat Kleiner y Paul Tukey, *Graphical Methods for Data Analysis*, Brooks/Cole, Pacific Grove, CA (1983). Una presentación muy recomendada de metodología gráfica y pictórica en estadística, tanto de años anteriores como recientes.

Devore, Jay y Roxy Peck, *Statistics: The Exploration and Analysis of Data* (4a. ed.), Duxbury Press, Pacific Grove, CA, 2001. En los primeros capítulos se presenta un estudio no matemático de métodos para describir y resumir datos.

Freedman, David, Robert Pisani y Roger Purves, *Statistics* (3a. ed.), Norton, Nueva York (1998). Un estudio excelente, no muy matemático, de razonamiento y metodología estadística básica.

Hoaglin, David, Frederick Mosteller y John Tukey, *Understanding Robust and Exploratory Data Analysis*, Wiley, Nueva York, 1983. Estudia por qué, y también cómo, se deben emplear métodos exploratorios; es bueno en detalles de diagramas de tallo y hojas y diagramas de caja.

Hoaglin, David y Paul Velleman, *Applications, Basics, and Computing of Exploratory Data Analysis*, Duxbury Press, Boston (1980). Un buen análisis de algunos métodos exploratorios básicos.

Moore, David, *Statistics: Concepts and Controversies* (5a. ed.), Freeman, San Francisco, 2001. Una obra muy amena y divertida que contiene un análisis intuitivo de problemas en relación con el muestreo y diseño de experimentos.

Tanur, Judith, *et al.* (eds.), *Statistics: A Guide to the Unknown* (3a. ed.), Duxbury Press, Belmont, CA, 1988. Contiene diversos artículos breves no técnicos que describen varias aplicaciones de la estadística.

2

Probabilidad

Introducción

El término **probabilidad** se refiere al estudio de la aleatoriedad y la incertidumbre. En cualquier situación donde podría ocurrir uno de varios resultados posibles, la teoría de la probabilidad proporciona métodos para cuantificar las probabilidades relacionadas con varios resultados. El lenguaje de la probabilidad se utiliza por lo común de una manera informal, tanto en contextos escritos como verbales. Algunos ejemplos son expresiones como "Es posible que el promedio Dow-Jones aumente al final del año", "hay una probabilidad 50-50 de que el titular busque la reelección", "es probable que el próximo año se ofrezca por lo menos una sección de ese curso", "las posibilidades favorecen un arreglo rápido de la huelga", y "se espera que por lo menos se vendan 20 000 boletos para el concierto". En este capítulo se introducen algunos conceptos elementales de probabilidad, se indica cómo se interpretan las probabilidades y se muestra cómo se pueden aplicar las reglas de la probabilidad para calcular las posibilidades de muchos eventos interesantes. La metodología de la probabilidad permite expresar en lenguaje preciso declaraciones informales como las mencionadas.

El estudio de la probabilidad como rama de las matemáticas se remonta a más de 300 años, donde tuvo su origen junto con preguntas relacionadas con juegos de azar. Muchos libros se dedican sólo a la probabilidad, pero el objetivo aquí es abarcar la parte del tema que tiene relación más directa con problemas de inferencia estadística.

2.1 Espacios muestrales y eventos

Un **experimento** es cualquier acción o proceso cuyo resultado está sujeto a incertidumbre. Aunque la palabra *experimento* normalmente hace pensar en una situación de prueba de laboratorio planeada o controlada con cuidado, aquí se utiliza en un sentido mucho más amplio. Así, los experimentos que pueden ser de interés, son lanzar al aire una moneda una vez o varias veces, seleccionar una o diversas cartas de una baraja, pesar una barra de pan, averiguar el tiempo para ir de la casa al trabajo en una mañana en particular, obtener los tipos de sangre de un grupo de individuos, o medir las resistencias compresivas de diferentes vigas de acero.

El espacio muestral de un experimento

DEFINICIÓN

El **espacio muestral** de un experimento, denotado por $\mathcal{S}$, es el conjunto de los resultados posibles de ese experimento.

Ejemplo 2.1 El experimento más sencillo al que se aplica la probabilidad, es aquel con dos resultados posibles. Un experimento de este tipo consiste en examinar un solo fusible para ver si está defectuoso. El espacio muestral para este experimento, se puede abreviar como $\mathcal{S} = \{N, D\}$, donde N se utiliza para indicar que el elemento es no defectuoso, D defectuoso, y las llaves se emplean para encerrar los elementos de un conjunto. En otro de estos experimentos se podría lanzar al aire una tachuela y observar si cae con la punta hacia arriba o hacia abajo, con espacio muestral $\mathcal{S} = \{U, D\}$, y otro más consistiría en observar el género del siguiente niño nacido en el hospital local, con $\mathcal{S} = \{M, F\}$. ■

Ejemplo 2.2 Si se examinan tres fusibles en secuencia y se observa el resultado de cada examen, entonces un resultado para todo el experimento es cualquier secuencia de letras N y D de tamaño 3, por lo tanto

$$\mathcal{S} = \{NNN, NND, NDN, NDD, DNN, DND, DDN, DDD\}$$

Si se hubiera lanzado al aire una tachuela tres veces, el espacio muestral se obtendría sustituyendo N por U en $\mathcal{S}$ con un cambio de notación que produce el mismo espacio muestral para el experimento en el que se observa el género de tres niños recién nacidos. ■

Ejemplo 2.3 Dos gasolineras se ubican en una cierta intersección. Cada una tiene seis bombas. Considere el experimento en el que para cada gasolinera se determina el número de bombas en uso en determinado momento del día. Un resultado experimental especifica cuántas bombas están en uso en la primera gasolinera y cuántas en la segunda. Un posible resultado es (2, 2), otro es (4, 1), y otro más (1, 4). Los 49 resultados en $\mathcal{S}$ se muestran en la siguiente tabla. El espacio muestral para el experimento en el que se lanza dos veces un dado, resulta de borrar el renglón 0 y la columna 0 de la tabla, obteniéndose 36 resultados.

		Segunda gasolinera						
		0	**1**	**2**	**3**	**4**	**5**	**6**
Primera gasolinera	**0**	(0, 0)	(0, 1)	(0, 2)	(0, 3)	(0, 4)	(0, 5)	(0, 6)
	1	(1, 0)	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
	2	(2, 0)	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
	3	(3, 0)	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
	4	(4, 0)	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
	5	(5, 0)	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
	6	(6, 0)	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

■

Ejemplo 2.4 Si una nueva batería tipo D para linterna tiene un voltaje que está fuera de ciertos límites, esa batería se caracteriza como una falla (F); si la batería tiene un voltaje dentro de los límites prescritos, es un éxito (S). Suponga que un experimento consiste en probar cada batería cuando sale de la línea de ensamble hasta que se observa un éxito. Aunque podría no ser muy probable, un resultado posible de este experimento es que las primeras 10 (o 100 o 1000 o . . .) sean F y la siguiente sea S. Es decir, para cualquier entero positivo n, es posible que haya que examinar n baterías antes de ver la primera S. El espacio muestral es $\mathcal{S} = \{S, FS, FFS, FFFS, \ldots\}$, que contiene un número infinito de resultados posibles. La misma forma abreviada del espacio muestral es apropiada para un experimento donde, comenzando en un momento específico, se registra el género de cada infante recién nacido hasta observar el nacimiento de un varón. ■

Eventos

Cuando se estudia probabilidad, se está interesado no sólo en cada uno de los resultados de $\mathcal{S}$, sino también en cualquier conjunto de resultados de $\mathcal{S}$.

DEFINICIÓN

> Un **evento** es cualquier colección (subconjunto) de resultados contenida en el espacio muestral $\mathcal{S}$. Se dice que un evento es **simple** si consiste en exactamente un resultado, y **compuesto** si consta de más de uno.

Cuando se lleva a cabo un experimento, se dice que ocurre un evento particular A si el resultado experimental está contenido en A. En general, ocurrirá exactamente un evento simple, pero muchos eventos compuestos ocurrirán en forma simultánea.

Ejemplo 2.5 Considere un experimento en el que cada uno de tres automóviles que toma una determinada salida de autopista da vuelta a la izquierda (L) o a la derecha (R) al final de la rampa de salida. Los ocho resultados posibles que comprenden el espacio muestral son LLL, RLL, LRL, LLR, LRR, RLR, RRL y RRR. Así que hay ocho eventos simples, entre los cuales están $E_1 = \{LLL\}$ y $E_5 = \{LRR\}$. Algunos eventos compuestos son

$A = \{RLL, LRL, LLR\}$ = el evento en el que exactamente uno de los tres automóviles da vuelta a la derecha.

$B = \{LLL, RLL, LRL, LLR\}$ = el evento en el que a lo sumo uno de los automóviles da vuelta a la derecha.

$C = \{LLL, RRR\}$ = el evento en el que los tres automóviles dan vuelta en la misma dirección.

Suponga que cuando se realiza el experimento el resultado es LLL. Entonces ocurrió el evento simple E_1 y, por lo tanto, también los eventos B y C (pero no A). ■

Ejemplo 2.6 (continuación del ejemplo 2.3)

Cuando se observa el número de bombas en uso en cada una de dos gasolineras provistas con seis bombas, hay 49 resultados posibles, así que hay 49 eventos simples: $E_1 = \{(0, 0)\}$, $E_2 = \{(0, 1)\}, \ldots, E_{49} = \{(6, 6)\}$. Ejemplos de eventos compuestos son:

$A = \{(0, 0), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$ = el evento en el que el número de bombas en uso es el mismo para ambas gasolineras.

$B = \{(0, 4), (1, 3), (2, 2), (3, 1), (4, 0)\}$ = el evento en el que el número total de bombas en uso es cuatro.

$C = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ = el evento en el que a lo sumo una bomba está en uso en cada gasolinera. ■

Ejemplo 2.7 (continuación del ejemplo 2.4)

El espacio muestral para el experimento de examen de baterías contiene un número infinito de resultados y, por lo tanto, hay un número infinito de eventos simples. Los eventos compuestos son

$A = \{S, FS, FFS\}$ = el evento en el que a lo sumo se examinan tres baterías.

$E = \{FS, FFFS, FFFFFS, \ldots\}$ = el evento en el que se examina un número par de baterías. ■

Algunas relaciones de teoría de conjuntos

Un evento no es otra cosa que un conjunto, así que las relaciones y resultados de la teoría elemental de conjuntos se puede usar para estudiar eventos. Las siguientes operaciones se usan para construir eventos nuevos a partir de eventos determinados.

DEFINICIÓN

1. La **unión** de dos eventos A y B denotada por $A \cup B$ y que se lee "*A unión B*", es el evento que consiste en los resultados que están *ya sea en A o en B o en ambos eventos* (así que la unión incluye resultados para los que ocurren tanto A como B, así como resultados para los que ocurre exactamente uno). Es decir, los resultados en por lo menos uno de los eventos.
2. La **intersección** de dos eventos A y B, que se denota como $A \cap B$ y se lee "*A intersección B*", es el evento que consiste en los resultados que están tanto en A como en B.
3. El **complemento** de un evento A, denotado por A', es el conjunto de todos los resultados en $\mathcal{S}$.

Ejemplo 2.8 (continuación del ejemplo 2.3)

Para el experimento donde se observa el número de bombas en uso en una sola gasolinera con seis bombas, sea $A = \{0, 1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$ y $C = \{1, 3, 5\}$. Entonces

$$A \cup B = \{0, 1, 2, 3, 4, 5, 6\} = \mathcal{S}, \quad A \cup C = \{0, 1, 2, 3, 4, 5\},$$

$$A \cap B = \{3, 4\}, \quad A \cap C = \{1, 3\}, \quad A' = \{5, 6\}, \quad \{A \cup C\}' = \{6\}$$ ■

Ejemplo 2.9 (continuación del ejemplo 2.4) En el experimento de la batería, defina A, B y C mediante

$$A = \{S, FS, FFS\}$$
$$B = \{S, FFS, FFFFS\}$$

y

$$C = \{FS, FFFS, FFFFFS, \ldots\}$$

Entonces

$$A \cup B = \{S, FS, FFS, FFFFS\}$$
$$A \cap B = \{S, FFS\}$$
$$A' = \{FFFS, FFFFS, FFFFFS, \ldots\}$$

y

$$C' = \{S, FFS, FFFFS, \ldots\} = \{\text{un número impar de baterías examinadas}\}$$ ■

A veces A y B no tienen resultados en común, así que la intersección de A y B no contiene resultados.

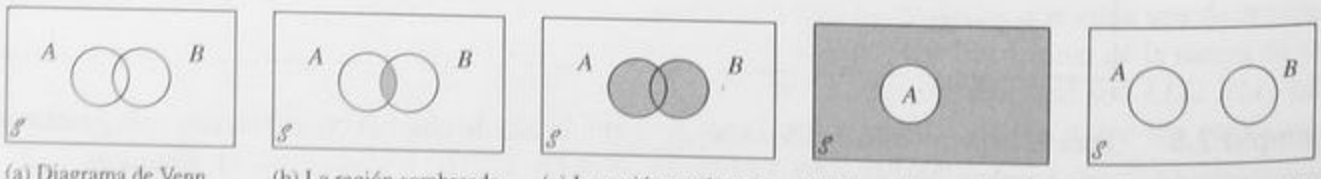
DEFINICIÓN

Cuando A y B no tienen resultados en común, se dice que son eventos **mutuamente excluyentes o disjuntos**.

Ejemplo 2.10 Una ciudad pequeña tiene tres distribuidores de automóviles: un distribuidor de GM que vende Chevrolets, Pontiacs y Buicks; uno de Ford que vende Fords y Mercurys, y el de Chrysler que vende Plymouths y Chryslers. Si un experimento consiste en observar la marca del siguiente automóvil vendido, entonces los eventos $A = \{\text{Chevrolet, Pontiac, Buick}\}$ y $B = \{\text{Ford, Mercury}\}$ son mutuamente excluyentes, porque el siguiente automóvil vendido no puede ser ni un producto de GM ni de Ford. ■

Las operaciones de unión e intersección se pueden extender a más de dos eventos. Para tres eventos A, B y C, el evento $A \cup B \cup C$ es el conjunto de resultados contenidos en los tres eventos. Dados los eventos $A_1, A_2, A_3, \ldots$, se dice que éstos son mutuamente excluyentes (o disjuntos por pares) si ninguno de los dos eventos tiene resultados en común.

Por medio de diagramas de Venn se obtiene una representación gráfica de eventos y manipulaciones con eventos. Para construir un diagrama de Venn, se dibuja un rectángulo cuyo interior representa el espacio muestral $\mathcal{S}$. Después, cualquier evento A se representa como el interior de una curva cerrada (a veces un círculo) contenida en $\mathcal{S}$. En la figura 2.1 aparecen ejemplos de diagramas de Venn.



(a) Diagrama de Venn de los eventos A y B (b) La región sombreada es $A \cap B$ (c) La región sombreada es $A \cup B$ (d) La región sombreada es A' (e) Eventos mutuamente excluyentes

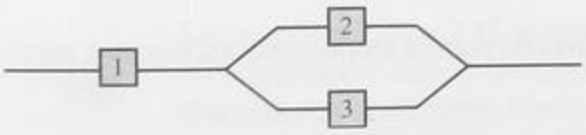
Figura 2.1 Diagramas de Venn

Ejercicios | Sección 2.1 (1 a 10)

1. Cuatro universidades, 1, 2, 3 y 4, participarán en un torneo vacacional de basquetbol. En la primera ronda, 1 jugará contra 2 y 3 contra 4. Los dos ganadores jugarán por el campeonato, y también jugarán los dos perdedores. Un posible resultado se puede representar por 1324 (1 vence a 2 y 3 le gana a 4 en la primera ronda, y después 1 derrota a 3 y 2 vence a 4).
   a. Liste los resultados de $\mathcal{S}$.
   b. Sea A el evento en el que 1 gana el torneo. Enumere los resultados de A.
   c. Sea B el evento en el que 2 llega a la final. Liste los resultados de B.
   d. ¿Cuáles son los resultados de $A \cup B$ y de $A \cap B$? ¿Cuáles son los resultados de A'?

2. Suponga que los vehículos que toman determinada salida de la autopista dan vuelta a la derecha (R), a la izquierda (L) o siguen derecho (S). Considere el hecho de observar la dirección de tres vehículos sucesivos.
   a. Liste los resultados del evento A donde los tres vehículos van en la misma dirección.
   b. Enumere los resultados del evento B donde los tres vehículos toman direcciones distintas.
   c. Registre los resultados del evento C en el que exactamente dos de los tres vehículos dan vuelta a la derecha.
   d. Liste los resultados del evento D en el que exactamente dos vehículos van en la misma dirección.
   e. Enumere los resultados de D', $C \cup D$ y $C \cap D$.

3. Tres componentes se conectan para formar un sistema, como se indica en el siguiente diagrama. Como los componentes en el subsistema 2-3 están conectados en paralelo, ese subsistema trabaja si por lo menos funciona uno de los dos componentes. Para que funcione todo el sistema, debe trabajar el componente 1 y el subsistema 2-3.



   El experimento consiste en determinar la condición de cada componente [S (éxito) para un componente que funciona y F (falla) para uno que no funciona].
   a. ¿Cuáles resultados están contenidos en el evento A en el que exactamente funcionan dos de los tres componentes?
   b. ¿Qué resultados están contenidos en el evento B en el que por lo menos funcionan dos de los componentes?
   c. ¿Qué resultados están contenidos en el evento C en el que funciona el sistema?
   d. Liste los resultados de C', $A \cup C$, $A \cap C$, $B \cup C$ y $B \cap C$.

4. Cada muestra de cuatro hipotecas para vivienda está clasificada como de tasa fija (F) o variable (V).
   a. ¿Cuáles son los 16 resultados de $\mathcal{S}$?
   b. ¿Cuáles resultados están en el evento en el que exactamente tres de las hipotecas seleccionadas son de tasa fija?
   c. ¿Cuáles resultados están en el evento en el que las cuatro hipotecas son del mismo tipo?
   d. ¿Cuáles resultados están en el evento en el que a lo sumo una de las cuatro hipotecas es de tasa variable?
   e. ¿Cuál es la unión de los eventos de los incisos (c) y (d), y cuál es la intersección de estos dos eventos?
   f. ¿Cuáles son la unión e intersección de los dos eventos de los incisos (b) y (c)?

5. Una familia que consta de tres personas, A, B y C, pertenece a una clínica médica que siempre tiene un médico en cada uno de los consultorios 1, 2 y 3. Durante cierta semana, cada miembro de la familia visita una vez la clínica y se le asigna al azar un consultorio. El experimento consiste en registrar el número de consultorio asignado a cada miembro de la familia. Un resultado es (1, 2, 1) en el que A se asigna al consultorio 1, B al 2 y C al 1.
   a. Enumere los 27 resultados del espacio muestral.
   b. Liste los resultados del evento en el que las tres personas de la familia van al mismo consultorio.
   c. Mencione los resultados del evento en el que todos los miembros de la familia vayan a distintos consultorios.
   d. Registre los resultados del evento en el que ningún miembro de la familia vaya al consultorio 2.

6. La biblioteca de una universidad tiene en reserva cinco ejemplares de cierto texto. Dos ejemplares (1 y 2) son primeras impresiones y los otros tres (3, 4 y 5) son segundas impresiones. Un estudiante examina estos libros en orden aleatorio, y se detiene sólo cuando se selecciona una segunda impresión. Un resultado posible es 5, y otro es 213.
   a. Liste los resultados de $\mathcal{S}$.
   b. Sea A el evento en el que exactamente se debe examinar un libro. ¿Qué resultados están en A?
   c. Sea B el evento en el que se selecciona el libro 5, ¿Qué resultados están en B?
   d. Sea C el evento en el que no se examina el libro 1. ¿Qué resultados están en C?

7. Un departamento académico acaba de terminar una votación secreta para elegir jefe de departamento. La urna contiene cuatro papeletas con votos para el candidato A y tres papeletas con votos para el candidato B. Suponga que las papeletas se sacan una por una de la caja.
   a. Liste los resultados posibles.
   b. Suponga que se realiza un conteo a medida que se sacan las papeletas. ¿En qué resultados A se mantiene adelante de B en todo el conteo?

8. Una compañía de ingenieros constructores está trabajando actualmente en plantas eléctricas en tres lugares diferentes. Sea A_i el evento en el que la planta del lugar i se termina en la fecha del contrato. Utilice las operaciones de unión, intersección y complemento para describir cada uno de los siguientes eventos, en términos de A_1, A_2 y A_3, dibuje un diagrama de Venn y sombree la región correspondiente a cada uno.
   a. Por lo menos una planta se termina en la fecha del contrato.
   b. Todas las plantas se terminan en la fecha del contrato.
   c. Sólo se termina la planta del sitio 1 en la fecha del contrato.
   d. Exactamente se termina una planta en la fecha del contrato.
   e. Se termina ya sea la planta del lugar 1 o las otras dos en la fecha del contrato.

9. Utilice un diagrama de Venn a fin de comprobar las dos relaciones siguientes para cualquiera de los eventos A y B (éstas se llaman leyes de De Morgan):
   a. $(A \cup B)' = A' \cap B'$
   b. $(A \cap B)' = A' \cup B'$

10. a. En el ejemplo 2.10, identifique tres eventos que sean mutuamente excluyentes.
    b. Suponga que no hay un resultado común para los tres eventos A, B y C. ¿Estos tres eventos son necesariamente mutuamente excluyentes? Si la respuesta es afirmativa, explique por qué; si es negativa, dé un ejemplo en contra por medio del experimento del ejemplo 2.10.

2.2 Axiomas, interpretaciones y propiedades de la probabilidad

Dados un experimento y un espacio muestral $\mathcal{S}$, *el objetivo de la probabilidad es asignar a cada evento A un número P(A), denominado probabilidad del evento A, que dará una medida más precisa de la probabilidad de que ocurra el evento A.* Para asegurar que las asignaciones de probabilidad concuerden con las nociones intuitivas de probabilidad, las asignaciones deben satisfacer los axiomas siguientes (propiedades básicas) de probabilidad.

AXIOMA 1 Para cualquier evento A, $P(A) \geq 0$.

AXIOMA 2 $P(\mathcal{S}) = 1$.

AXIOMA 3
   a. Si $A_1, A_2, \ldots, A_k$ es una colección finita de eventos mutuamente excluyentes, entonces
   $$P(A_1 \cup A_2 \cup \cdots \cup A_k) = \sum_{i=1}^{k} P(A_i)$$
   b. Si $A_1, A_2, A_3, \ldots$ es una colección infinita de eventos mutuamente excluyentes, entonces
   $$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i)$$

El axioma 1 refleja la noción intuitiva de que la probabilidad de que ocurra A debe ser no negativa. El espacio muestral es, por definición, el evento que debe ocurrir cuando se realiza el experimento ($\mathcal{S}$ contiene todos los posibles resultados), así que el axioma 2 señala que la

probabilidad máxima posible de 1 se asigna a $\mathcal{S}$. El tercer axioma formaliza la idea de que si se desea la probabilidad de que ocurra al menos uno de varios eventos, y no ocurran dos al mismo tiempo, entonces la probabilidad de que ocurra al menos uno es la suma de las probabilidades de cada uno de los eventos.

Ejemplo 2.11 En el experimento en el que se lanza una sola moneda, el espacio muestral es $\mathcal{S} = \{H, T\}$. Los axiomas especifican $P(\mathcal{S}) = 1$, así que para completar la asignación de probabilidad, sólo resta determinar $P(H)$ y $P(T)$. Puesto que H y T son eventos disjuntos y $H \cup T = \mathcal{S}$, el axioma 3 implica que

$$1 = P(\mathcal{S}) = P(H) + P(T)$$

Esto significa que $P(T) = 1 - P(H)$. La única libertad que permiten los axiomas en este experimento es la probabilidad asignada a H. Una posible asignación de probabilidades es $P(H) = 0.5$, $P(T) = 0.5$, en tanto que otra asignación posible es $P(H) = 0.75$, $P(T) = 0.25$. De hecho, si p representa cualquier número fijo entre 0 y 1, $P(H) = p$ y $P(T) = 1 - p$ es una asignación consistente con los axiomas. ■

Ejemplo 2.12 Considere el experimento del ejemplo 2.4, en el que las baterías que salen de una línea de ensamble se prueban una por una hasta encontrar una que tiene un voltaje dentro de límites prescritos. Los eventos simples son $E_1 = \{S\}$, $E_2 = \{FS\}$, $E_3 = \{FFS\}$, $E_4 = \{FFFS\}$, ... Suponga que la probabilidad de cualquier batería que es satisfactoria es 0.99. Entonces se puede demostrar que $P(E_1) = 0.99$, $P(E_2) = (0.01)(0.99)$, $P(E_3) = (0.01)^2(0.99)$, ... es una asignación de probabilidades para los eventos simples que satisface los axiomas. En particular, debido a que las E_i son disjuntas y $\mathcal{S} = E_1 \cup E_2 \cup E_3 \cup \ldots$, debe ser el caso de que

$$\begin{aligned} 1 = P(\mathcal{S}) &= P(E_1) + P(E_2) + P(E_3) + \cdots \\ &= 0.99[1 + 0.01 + (0.01)^2 + (0.01)^3 + \cdots] \end{aligned}$$

La validez de esta igualdad es una consecuencia de un resultado matemático que tiene que ver con la suma de una serie geométrica.

Sin embargo, se obtiene otra asignación de probabilidad legítima (según los axiomas) del mismo tipo "geométrico" al reemplazar 0.99 por cualquier otro número p entre 0 y 1 (y 0.01 por $1 - p$). ■

Cómo interpretar la probabilidad

Los ejemplos 2.11 y 2.12 muestran que los axiomas no determinan por completo una asignación de probabilidades a los eventos. Los axiomas sólo sirven para eliminar asignaciones inconsistentes con el concepto intuitivo de probabilidad. En el experimento en el que se lanza una moneda al aire, en el ejemplo 2.11, fueron sugeridas dos asignaciones en particular. La asignación correcta o adecuada depende de la manera cómo se realiza un experimento, y también de la interpretación que se haga de la probabilidad. La interpretación utilizada con más frecuencia y que se entiende con más facilidad, está basada en la noción de frecuencias relativas.

Considere un experimento que se puede llevar a cabo una y otra vez de una manera idéntica e independiente, y sea A un evento que consiste en un conjunto fijo de resultados del experimento. Los experimentos estudiados antes, como lanzar una moneda y tirar dados, son ejemplos simples de estos experimentos repetibles. Si el experimento se realiza n veces, en algunas de las repeticiones ocurre el evento A (el resultado se encuentra en el conjunto A), y en otros no ocurre A. Sea $n(A)$ el número de repeticiones en las que ocurre A. Entonces, la rela-

ción $n(A)/n$ se denomina *frecuencia relativa* de ocurrencia del evento A en la secuencia de n repeticiones. La evidencia empírica, basada en los resultados de estas secuencias de experimentos reproducibles, indica que, a medida que crece n, se estabiliza la frecuencia relativa $n(A)/n$, como se ilustra en la figura 2.2. Es decir, a medida que n se hace arbitrariamente grande, la frecuencia relativa se aproxima a un valor límite que se llama *frecuencia relativa limitante* del evento A. La interpretación objetiva de probabilidad identifica esta frecuencia relativa limitante con $P(A)$.

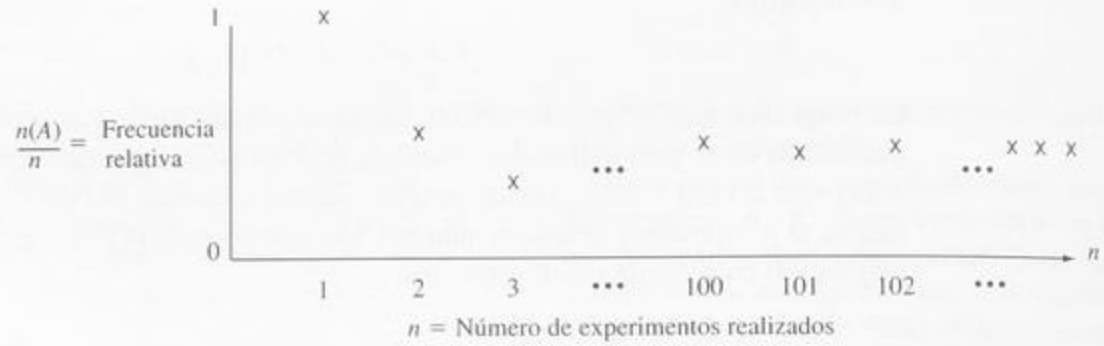


Figura 2.2 Estabilización de la frecuencia relativa

Si se asignan las probabilidades a los eventos de acuerdo con sus frecuencias relativas limitantes, entonces el enunciado "la probabilidad de esa moneda que caiga con la cara hacia arriba cuando es lanzada es 0.5", se puede interpretar como que en un gran número de lanzamientos aparecerá una cara en aproximadamente la mitad de los lanzamientos y una cruz en la otra mitad.

Se dice que esta interpretación de frecuencia relativa para la probabilidad es objetiva, porque se apoya en una propiedad del experimento, y no en cualquier individuo particular relacionado con el experimento. Por ejemplo, dos observadores diferentes de una secuencia de lanzamientos de una moneda, deben utilizar las mismas asignaciones de probabilidad, ya que los observadores no tienen nada que ver con la frecuencia relativa limitante. En la práctica, esta interpretación no es tan objetiva como podría parecer, puesto que se desconoce la frecuencia relativa limitante de un evento. Entonces, se tienen que asignar probabilidades basadas en creencias personales acerca de la frecuencia relativa limitante de eventos bajo estudio. Afortunadamente, hay muchos experimentos para los cuales habrá un consenso con respecto a asignaciones de probabilidad. Cuando se habla de una moneda imparcial, se quiere decir que $P(H) = P(T) = 0.5$, y un dado no cargado es aquel para el que las frecuencias relativas limitantes de los seis resultados son $\frac{1}{6}$, lo que hace pensar en asignaciones de probabilidad $P(\{1\}) = \cdots = P(\{6\}) = \frac{1}{6}$.

Debido a que la interpretación objetiva de probabilidad se basa en el concepto de frecuencia límite, su aplicación se limita a situaciones experimentales repetibles. Sin embargo, el lenguaje de la probabilidad se utiliza con frecuencia en relación con situaciones que de manera inherente son irrepetibles. Algunos ejemplos son "las probabilidades son buenas para un acuerdo de paz", "es probable que se adjudique el contrato a nuestra empresa", y "debido a que está lesionado su mariscal de campo, espero que anoten no más de 10 puntos contra nosotros". En estos casos sería preferible, como antes, asignar probabilidades numéricas a varios resultados y eventos (por ejemplo, la probabilidad de que se consiga el contrato es 0.9). Por lo tanto, se debe adoptar otra interpretación de estas probabilidades. Debido a que diferentes observa-

dores pueden tener información y opiniones previas diferentes en relación con este tipo de situaciones experimentales, las asignaciones de probabilidad podrían diferir ahora de un individuo a otro. Así que las interpretaciones en estas situaciones se conocen como *subjetivas*. En el libro de Robert Winkler citado en las referencias del capítulo se proporciona un estudio descifrable de varias interpretaciones subjetivas.

Propiedades de la probabilidad

PROPOSICIÓN

Para cualquier evento A, $P(A) = 1 - P(A')$.

Demostración:
En el axioma 3a, sea $k = 2$, $A_1 = A$ y $A_2 = A'$. Puesto que por definición de A', $A \cup A' = \mathcal{S}$ en tanto que A y A' son disjuntos, $1 = P(\mathcal{S}) = P(A \cup A') = P(A) + P(A')$, de donde se deduce el resultado deseado. ■

Esta proposición es sorprendentemente útil, porque hay situaciones en las que $P(A')$ se obtiene con más facilidad por métodos directos que $P(A)$.

Ejemplo 2.13 Considere un sistema de cinco componentes idénticos conectados en serie, como se ilustra en la figura 2.3.

Figura 2.3 Sistema de cinco componentes conectados en serie

Denote con F un componente que falla y con S uno que no falla. Sea A el evento de que falla el *sistema*. Para que A ocurra, por lo menos debe fallar uno de los componentes individuales. Los resultados de A incluyen *SSFSS* (1, 2, 4 y 5 trabajan, pero 3 no), *FFSSS*, etc. Hay de hecho 31 resultados diferentes en A. Sin embargo, A', el evento en el que funciona el sistema, consiste en un solo resultado *SSSSS*. En la sección 2.5 se ve que si no falla 90% de los componentes y componentes diferentes fallan en forma independiente entre sí, entonces $P(A') = P(SSSSS) = 0.9^5 = 0.59$. Así que $P(A) = 1 - 0.59 = 0.41$; Por lo tanto, entre un gran número de estos sistemas, fallará alrededor de 41 por ciento. ■

En general, la proposición anterior es útil cuando el evento de interés se puede expresar como "al menos . . . ," ya que entonces podría ser más fácil trabajar con el complemento "menor que . . ." (en algunos problemas, es más fácil tratar con "más que . . ." que con "a lo sumo . . ."). Cuando se tiene dificultad para calcular $P(A)$ de manera directa, considere determinar $P(A')$.

PROPOSICIÓN

Si A y B son mutuamente excluyentes, entonces $P(A \cap B) = 0$.

Demostración:
Debido a que $A \cap B$ no contiene resultados $(A \cap B)' = \mathcal{S}$. Entonces, se tiene que $1 = P[(A \cap B)'] = 1 - P(A \cap B)$, lo cual implica que $P(A \cap B) = 1 - 1 = 0$. ■

Cuando los eventos A y B son mutuamente excluyentes, el axioma 3 da $P(A \cup B) = P(A) + P(B)$. Cuando A y B no son mutuamente excluyentes, la probabilidad de la unión se obtiene del siguiente resultado.

PROPOSICIÓN

Para dos eventos cualesquiera A y B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Observe que la proposición es válida incluso si A y B son mutuamente excluyentes, puesto que entonces $P(A \cap B) = 0$. La idea clave es que, al sumar $P(A)$ y $P(B)$, la probabilidad de la intersección $A \cap B$ en realidad se cuenta dos veces, así que se debe restar $P(A \cap B)$.

Demostración:
Observe primero que $A \cup B = A \cup (B \cap A')$, como se ilustra en la figura 2.4. Dado que A y $(B \cap A')$ son mutuamente excluyentes, $P(A \cup B) = P(A) + P(B \cap A')$. Pero $B = (B \cap A) \cup (B \cap A')$ (la unión de esa parte de B en A y la parte de B que no está en A), con $(B \cap A)$ y $(B \cap A')$ que son mutuamente excluyentes; por lo tanto, $P(B) = P(B \cap A) + P(B \cap A')$. Al combinar estos resultados se obtiene

$$\begin{aligned} P(A \cup B) = P(A) + P(B \cap A') &= P(A) + [P(B) - P(A \cap B)] \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

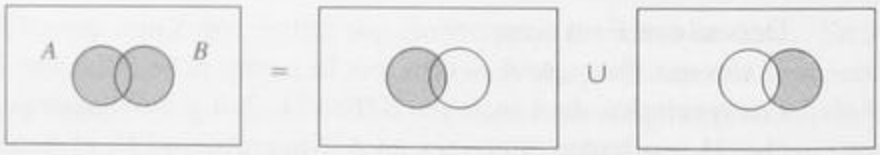


Figura 2.4 Representación de $A \cup B$ como una unión de eventos disjuntos ■

Ejemplo 2.14 En un determinado suburbio residencial, 60% de los hogares se suscriben al periódico metropolitano publicado en una ciudad cercana, 80% se suscriben al periódico local y 50% se suscriben a ambos periódicos. Si se selecciona al azar una familia, ¿cuál es la probabilidad de que esté suscrita 1) al menos a uno de los dos periódicos y 2) exactamente a uno de los dos periódicos?

Con $A = \{$se suscribe al periódico metropolitano$\}$ y $B = \{$se suscribe al periódico local$\}$, la información proporcionada implica que $P(A) = 0.6$, $P(B) = 0.8$ y $P(A \cap B) = 0.5$. Luego se aplica la proposición previa y se obtiene

$P($se suscribe por lo menos a uno de los dos periódicos$)$

$$= P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.6 + 0.8 - 0.5 = 0.9$$

El evento de que una familia se suscriba sólo al periódico local se puede escribir como $A' \cap B$ [(no metropolitano) y local]. Ahora la figura 2.4 implica que

$$0.9 = P(A \cup B) = P(A) + P(A' \cap B) = 0.6 + P(A' \cap B)$$

de donde $P(A' \cap B) = 0.3$. De forma similar, $P(A \cap B') = P(A \cup B) - P(B) = 0.1$. Todo esto se ilustra en la figura 2.5, en la cual se observa que

$$P(\text{exactamente uno}) = P(A \cap B') + P(A' \cap B) = 0.1 + 0.3 = 0.4$$

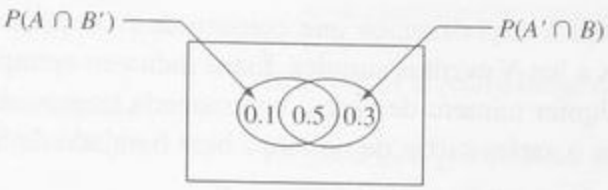


Figura 2.5 Probabilidades para el ejemplo 2.14 ■

La probabilidad de la unión de más de dos eventos se puede calcular de manera análoga. Para tres eventos A, B y C, el resultado es

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Esto se puede ver al examinar el diagrama de Venn de $A \cup B \cup C$, que se muestra en la figura 2.6. Cuando se suman $P(A)$, $P(B)$ y $P(C)$ ciertas intersecciones se cuentan dos veces, así que se deben restar, pero esto da como resultado que $P(A \cap B \cap C)$ se reste una vez con demasiada frecuencia.

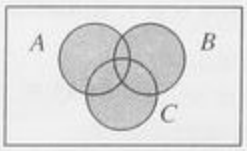


Figura 2.6 $A \cup B \cup C$

Determinación de probabilidades en forma sistemática

Cuando el número de resultados posibles (eventos simples) es grande, habrá muchos eventos compuestos Una forma simple de determinar las probabilidades para estos eventos que evita violar los axiomas y propiedades derivadas es establecer primero las probabilidades $P(E_i)$ para todos los eventos simples. Éstos deben satisfacer $P(E_i) \geq 0$ y $\sum_{\text{toda } i} P(E_i) = 1$. Entonces, la probabilidad de cualquier evento compuesto A se calcula al sumar las $P(E_i)$ para todas las E_i en A:

$$P(A) = \sum_{\text{todas las } E_i \text{ en } A} P(E_i)$$

Ejemplo 2.15 Denote con $\{1\}, \ldots, \{6\}$ los seis eventos elementales $E_1, \ldots, E_6$. relacionados con lanzar una vez un dado de seis lados. Si el dado está construido de modo que cualquiera de los tres resultados pares tenga el doble de probabilidad de ocurrir que cualquiera de los tres resultados impares, entonces una asignación apropiada de probabilidades para eventos elementales es $P(E_1) = P(E_3) = P(E_5) = \frac{1}{9}$, $P(E_2) = P(E_4) = P(E_6) = \frac{2}{9}$. Entonces para el evento $A = \{\text{el resultado es par}\} = E_2 \cup E_4 \cup E_6$, $P(A) = P(E_2) + P(E_4) + P(E_6) = \frac{6}{9} = \frac{2}{3}$; para $B = \{\text{el resultado es} \leq 3\} = E_1 \cup E_2 \cup E_3$, $P(B) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$. ■

Resultados equiprobables

En muchos experimentos que constan de N resultados, es razonable asignar probabilidades iguales a los N eventos simples. Éstos incluyen ejemplos obvios, como lanzar una o dos veces (o cualquier número de veces) una moneda imparcial o un dado no cargado, o bien, seleccionar una o varias cartas de un mazo bien barajado de 52. Con $p = P(E_i)$ así que i,

$$1 = \sum_{i=1}^{N} P(E_i) = \sum_{i=1}^{N} p = p \cdot N \qquad \text{por lo tanto } p = \frac{1}{N}$$

Es decir, si hay N resultados posibles, entonces la probabilidad asignada a cada uno es $1/N$.

Ahora considere un evento A, con $N(A)$ que indica el número de resultados contenidos en A. Entonces

$$P(A) = \sum_{E_i \text{ en } A} P(E_i) = \sum_{E_i \text{ en } A} \frac{1}{N} = \frac{N(A)}{N}$$

Una vez contado el número N de resultados del espacio muestral, para calcular la probabilidad de cualquier evento, se debe contar el número de resultados contenidos en ese evento y tomar la relación entre los dos números. Entonces, cuando los resultados son equiprobables, calcular las probabilidades se reduce a contar.

Ejemplo 2.16 Cuando se lanzan por separado dos dados, hay $N = 36$ resultados (elimine el primer renglón y la primera columna de la tabla del ejemplo 2.3). Si ninguno de los dados está cargado, los 36 resultados son equiprobables, así que $P(E_i) = \frac{1}{36}$. Entonces el evento $A = \{\text{suma de dos números} = 7\}$ consiste en los seis resultados (1, 6), (2, 5), (3, 4), (4, 3), (5, 2) y (6, 1), por lo tanto

$$P(A) = \frac{N(A)}{N} = \frac{6}{36} = \frac{1}{6}$$ ■

Ejercicios | Sección 2.2 (11 a 28)

11. Una compañía de fondo mutuo ofrece a sus clientes varios fondos: uno del mercado monetario, tres fondos de bonos diferentes (corto, mediano y largo plazo), dos de acciones (riesgo moderado y alto) y uno equilibrado. Entre los clientes que poseen acciones de un solo fondo, los porcentajes de clientes en los diferentes fondos son como sigue:

Mercado monetario	20%	Acciones de alto riesgo	18%
Bono a corto plazo	15%	Acciones de riesgo moderado	25%
Bono intermedio	10%	Fondo equilibrado	7%
Bono a largo plazo	5%		

Se selecciona al azar un cliente que tiene acciones en un solo fondo.

a. ¿Cuál es la probabilidad de que el individuo elegido tenga acciones del fondo equilibrado?

b. ¿Cuál es la probabilidad de que el individuo posea acciones de un fondo de bonos?

c. ¿Cuál es la probabilidad de que el individuo seleccionado no posea acciones del fondo de acciones?

12. Considere elegir al azar un alumno de cierta universidad, y sea A el evento de que el individuo seleccionado tenga una tarjeta de crédito Visa y B el evento análogo para una MasterCard. Suponga que $P(A) = 0.5$, $P(B) = 0.4$ y $P(A \cap B) = 0.25$.

a. Calcule la probabilidad de que el individuo seleccionado tenga al menos una de las dos tarjetas (es decir, la probabilidad del evento $A \cup B$).

b. ¿Cuál es la probabilidad de que el individuo elegido no tenga ninguna de esas tarjetas?

c. Describa, en términos de A y B, el evento de que el alumno seleccionado tenga una tarjeta Visa, pero no una MasterCard, y luego calcule la probabilidad de este evento.

13. Una consultoría de computadoras ha licitado en tres proyectos. Sea $A_i = \{\text{proyecto } i \text{ otorgado}\}$, para $i = 1, 2, 3$, y suponga que $P(A_1) = 0.22$, $P(A_2) = 0.25$, $P(A_3) = 0.28$, $P(A_1 \cap A_2) = 0.11$, $P(A_1 \cap A_3) = 0.05$, $P(A_2 \cap A_3) = 0.07$, $P(A_1 \cap A_2 \cap A_3) = 0.01$. Exprese con palabras cada uno de los siguientes eventos y calcule la probabilidad de cada uno:

a. $A_1 \cup A_2$

b. $A_1' \cap A_2'$ [*Sugerencia:* $(A_1 \cup A_2)' = A_1' \cap A_2'$]

c. $A_1 \cup A_2 \cup A_3$

d. $A_1' \cap A_2' \cap A_3'$

e. $A_1' \cap A_2' \cap A_3$

f. $(A_1' \cap A_2') \cup A_3$

14. Una compañía eléctrica ofrece una tasa subsidiada a cualquier familia cuyo consumo de electricidad sea menor de 240 kWh durante un determinado mes. Sea A el evento en el que una familia elegida al azar, en cierta comunidad no rebasa el consumo subsidiado durante enero, y sea B el evento análogo para el mes de julio (A y B se refieren a la misma familia). Supóngase que $P(A) = 0.8$, $P(B) = 0.7$ y $P(A \cup B) = 0.9$. Calcule lo siguiente:

a. $P(A \cap B)$.

b. La probabilidad de que el consumo subsidiado sea rebasado en exactamente uno de los dos meses. Describa este evento en términos de A y B.

15. Considere el tipo de secadora de ropa (de gas o eléctrica) que compran cada cinco clientes en cierta tienda.

a. Si la probabilidad de que uno de estos clientes compre una secadora eléctrica es 0.428, ¿cuál es la probabilidad de que por lo menos dos compren una secadora eléctrica?

b. Si $P(\text{los cinco compran secadoras de gas}) = 0.116$ y $P(\text{los cinco compran secadoras eléctricas}) = 0.005$, ¿cuál es la probabilidad de que por lo menos sea vendida una de cada tipo?

16. Un individuo tiene ante sí tres diferentes vasos con refresco de cola, marcados C, D y P. Se le pide que pruebe los tres y luego los enumere en orden de preferencia. Supongamos que en realidad los tres vasos tienen el mismo tipo de refresco de cola.

a. ¿Cuáles son los eventos simples en este experimento de clasificación, y qué probabilidad se asignaría a cada uno?

b. ¿Cuál es la probabilidad de que se clasifique primero a C?

c. ¿Cuál es la probabilidad de que se clasifique primero a C y por último a D?

17. Sea A el evento en el que el siguiente requisito para ayuda de un asesor de *software* de estadística se relacione con el paquete SPSS, y sea B el evento en el que el siguiente requisito es para ayuda con SAS. Supóngase que $P(A) = 0.30$ y $P(B) = 0.50$.

a. ¿Por qué no es posible que $P(A) + P(B) = 1$?

b. Calcule $P(A')$.

c. Calcule $P(A \cup B)$.

d. Calcule $P(A' \cap B')$.

18. Una caja contiene cuatro focos de 40 W, cinco de 60 W y seis de 75 W. Si se seleccionan los focos uno por uno en orden aleatorio, ¿cuál es la probabilidad de que al menos se deba elegir dos focos para obtener uno de 75 W?

19. La inspección visual humana de uniones de soldadura en tarjetas de circuito impreso puede ser muy subjetiva. Parte del problema surge de los distintos tipos de defectos de soldadura (p. ej., no humedecer la zona terminal, visibilidad de codos, huecos) e incluso el grado en el que la unión posee uno o más de estos defectos. En consecuencia, aun los inspectores más experimentados pueden estar en desacuerdo con la disposición de determinada unión. En un lote de 10 000 uniones, el inspector A encontró 724 que estaban defectuosas, el inspector B encontró 751, y por lo menos uno de los inspectores consideró que 1159 de las uniones eran defectuosas. Suponga que se selecciona al azar una de las 10 000 uniones.

a. ¿Cuál es la probabilidad de que ninguno de los inspectores considere defectuosa la unión seleccionada?

b. ¿Cuál es la probabilidad de que el inspector A, pero no el B, juzgue defectuosa la unión elegida?

20. En una fábrica los obreros trabajan tres turnos distintos. En el último año ocurrieron 200 accidentes. Algunos de éstos se pueden atribuir por lo menos en parte a las condiciones inseguras de trabajo, en tanto que otros se re-

lacionan con condiciones de trabajo. En la tabla siguiente se proporciona el porcentaje de accidentes que caen en cada tipo de categoría accidente-turno.

		Condiciones inseguras	Condiciones no relacionadas
Turno	*Diurno*	10%	35%
	Vespertino	8%	20%
	Nocturno	5%	22%

Suponga que se elige al azar uno de los 200 informes de accidentes de un archivo, y se determina el turno y tipo de accidente.

a. ¿Cuáles son los eventos simples?

b. ¿Cuál es la probabilidad de que el accidente seleccionado se atribuya a condiciones inseguras?

c. ¿Cuál es la probabilidad de que no ocurra en el turno diurno?

21. Una compañía de seguros ofrece cuatro niveles diferentes de deducibles —ninguno, bajo, medio y alto— para sus asegurados propietarios de casa; y tres niveles distintos —bajo, medio y alto— para quienes poseen un seguro de automóvil. En la tabla siguiente se dan las proporciones para las diversas categorías de asegurados que tienen ambos tipos de seguro. Por ejemplo, la proporción de individuos que tienen deducibles bajos de casa y automóvil es 0.06 (6%).

	Dueño de casa			
Automóvil	**N**	**B**	**M**	**A**
B	0.04	0.06	0.05	0.03
M	0.07	0.10	0.20	0.10
A	0.02	0.03	0.15	0.15

Suponga que se elige al azar un individuo que tiene ambos tipos de pólizas.

a. ¿Cuál es la probabilidad de que el individuo tenga un deducible medio de automóvil y un deducible alto de casa habitación?

b. ¿Cuál es la probabilidad de que tenga un deducible de automóvil bajo? ¿Un deducible de casa habitación bajo?

c. ¿Cuál es la probabilidad de que el individuo esté en la misma categoría de deducible para automóvil y casa?

d. Con base en las respuestas del inciso (c), ¿cuál es la probabilidad de que las dos categorías sean diferentes?

e. ¿Cuál es la probabilidad de que el individuo tenga por lo menos un nivel de deducible bajo?

f. Con la respuesta del inciso (e), ¿cuál es la probabilidad de que ningún nivel de deducible sea bajo?

22. La ruta que utiliza cierto automovilista para ir al trabajo tiene dos cruceros con semáforo. La probabilidad de que pare en el primer semáforo es 0.4, la probabilidad análoga para el segundo semáforo es 0.5, y la probabilidad de que se detenga por lo menos en uno de los semáforos es 0.6. ¿Cuál es la probabilidad de que se detenga

a. en ambos semáforos?

b. en el primero, pero no en el segundo?

c. en exactamente un semáforo?

23. Se van a reemplazar las computadoras de seis miembros de la facultad en cierto departamento. Dos de los miembros de la facultad eligieron *laptops* y los otros cuatro eligieron computadoras de escritorio. Supóngase que sólo se pueden llevar a cabo dos configuraciones en un día establecido, y las dos computadoras por configurar se eligen al azar de seis (así que hay 15 resultados equiprobables; si se enumera a las computadoras como 1, 2, . . . , 6, entonces el resultado consiste en las computadoras 1 y 2, otro consiste en las computadoras 1 y 3, etcétera).

a. ¿Cuál es la probabilidad de que ambas configuraciones seleccionadas sean para las computadoras *laptop*?

b. ¿Cuál es la probabilidad de que ambas configuraciones elegidas sean para máquinas de escritorio?

c. ¿Cuál es la probabilidad de que por lo menos una configuración seleccionada sea para una computadora de escritorio?

d. ¿Cuál es la probabilidad de que por lo menos se elija una computadora de cada tipo para configuración?

24. Utilice los axiomas para demostrar que si un evento A está contenido en otro evento B (es decir, A es un subconjunto de B), entonces $P(A) \leq P(B)$. [*Sugerencia:* para A y B, A y $B \cap A'$ son disjuntos y $B = A \cup (B \cap A')$, como puede verse en un diagrama de Venn.] Para A y B, ¿qué significa lo anterior acerca de la relación entre $P(A \cap B)$, $P(A)$ y $P(A \cup B)$?

25. Las tres opciones principales en cierto tipo de automóvil nuevo son transmisión automática (A), techo corredizo (B) y radio con lector de CD (C). Si 70% de los compradores piden A, 80% B, 75% C, 85% A o B, 90% A o C, 95% B o C y 98% A o B o C, calcule las probabilidades de los siguientes eventos. [*Sugerencia:* "A o B" es el evento de que se requiera por lo menos una de las dos opciones. Intente trazar un diagrama de Venn y marque todas las regiones.]

a. El siguiente comprador solicita una de las tres opciones.

b. El siguiente comprador no selecciona ninguna de las tres opciones.

c. El siguiente comprador sólo selecciona un automóvil con transmisión automática y ninguna de las otras dos opciones.

d. El siguiente comprador selecciona exactamente una de las tres opciones.

26. Un determinado sistema puede experimentar tres tipos de defectos. Sea A_i $(i = 1, 2, 3)$ el evento en el que el sistema tiene un defecto de tipo i. Suponga que

$$P(A_1) = 0.12 \quad P(A_2) = 0.07 \quad P(A_3) = 0.05$$
$$P(A_1 \cup A_2) = 0.13 \quad P(A_1 \cup A_3) = 0.14$$
$$P(A_2 \cup A_3) = 0.10 \quad P(A_1 \cap A_2 \cap A_3) = 0.01$$

a. ¿Cuál es la probabilidad de que el sistema no tenga un defecto tipo 1?

b. ¿Cuál es la probabilidad de que el sistema tenga tanto el defecto tipo 1 como el 2?

c. ¿Cuál es la probabilidad de que el sistema tenga los defectos tipo 1 y tipo 2, pero no un tipo 3?

d. ¿Cuál es la probabilidad de que el sistema tenga a lo sumo dos de esos defectos?

27. Un departamento académico con cinco miembros de la facultad: Anderson, Box, Cox, Cramer y Fisher, debe seleccionar a dos de sus miembros para formar parte de un comité de revisión de personal. Debido a que el trabajo será lento, nadie tiene la menor intención de prestar ese servicio, así que se decide que el representante sea seleccionado colocando cinco tiras de papel en una caja, mezclándolas y seleccionando dos.

a. ¿Cuál es la probabilidad de que sean seleccionados Anderson y Box? (*Sugerencia:* liste los resultados equiprobables.)

b. Cuál es la probabilidad de que se seleccione a por lo menos uno de los dos miembros cuyo apellido comienza con *C*?

c. Si los cinco miembros de la facultad han impartido cátedra en la universidad durante 3, 6, 7, 10 y 14 años, respectivamente, ¿cuál es la probabilidad de que los dos representantes seleccionados tengan por lo menos 15 años de experiencia docente en la universidad?

28. En el ejercicio 5, suponga que cualquier individuo tiene las mismas probabilidades de ser asignado a alguno de los tres consultorios, sin considerar dónde han sido asignados los otros individuos. ¿Cuál es la probabilidad de que

a. Los tres miembros de la familia sean asignados al mismo consultorio?

b. A lo sumo dos miembros de la familia sean asignados al mismo consultorio?

c. Cada uno de los miembros de la familia sea asignado a un consultorio diferente?

2.3 Técnicas de conteo

Cuando los diversos resultados de un experimento son equiprobables (se asigna la misma probabilidad a cada evento simple), la tarea de calcular probabilidades se reduce a contar. En particular, si N es el número de resultados en un espacio muestral y $N(A)$ el número de resultados contenidos en un evento A, entonces

$$P(A) = \frac{N(A)}{N} \tag{2.1}$$

Si se tiene, o es fácil construir, una lista de resultados y N es pequeño, entonces se puede obtener el numerador y el denominador de la ecuación (2.1) sin el beneficio de ningún principio de conteo general.

Sin embargo, hay muchos experimentos para los que el esfuerzo requerido para construir la lista es prohibitivo, porque N es bastante grande. Si se aprovechan algunas de las reglas generales de conteo, es posible calcular probabilidades de la forma (2.1) sin enumerar los resultados. Estas reglas también son útiles en problemas que tienen que ver con resultados que no son equiprobables. Varias de las reglas desarrolladas aquí se usan para estudiar distribuciones de probabilidad en el capítulo siguiente.

Regla del producto para pares ordenados

La primera regla de conteo se aplica a cualquier situación en la que un conjunto (evento) consiste en pares ordenados de objetos y se desea contarlos. Por par ordenado, se entiende que si

O_1 y O_2 son objetos, entonces el par (O_1, O_2) es diferente del par (O_2, O_1). Por ejemplo, si un individuo selecciona una línea aérea para un viaje de Los Ángeles a Chicago y (después de hacer negocios en Chicago) una segunda para continuar a Nueva York, una posibilidad es (American, United), otra es (United, American), y otra más (United, United).

PROPOSICIÓN

Si el primer elemento u objeto de un par ordenado se puede seleccionar en n_1 formas, y para cada una de estas n_1 formas se puede seleccionar el segundo elemento del par en n_2 formas, entonces el número de pares es n_1n_2.

Ejemplo 2.17 El propietario de una casa que realiza algunas remodelaciones, requiere los servicios de un contratista plomero y de un contratista electricista. Si hay 12 contratistas plomeros y 9 contratistas electricistas en la zona, ¿de cuántas formas se puede seleccionar a los contratistas? Si se utiliza $P_1, \ldots, P_{12}$ para representar a los plomeros y $Q_1, \ldots, Q_9$, para los electricistas, entonces se desea el número de pares de la forma (P_i, Q_j). Con $n_1 = 12$ y $n_2 = 9$, la regla de producto produce $N = (12)(9) = 108$ formas posibles de elegir los dos tipos de contratistas. ■

En el ejemplo 2.17, la elección del segundo elemento del par no depende de qué primer elemento se elija u ocurra. Mientras que haya el mismo número de opciones del segundo elemento para cada primer elemento, la regla del producto es válida aun cuando el conjunto de posibles segundos elementos dependa del primer elemento.

Ejemplo 2.18 Una familia recién llegada a una ciudad, requiere los servicios de un obstetra y un pediatra. Hay dos clínicas de fácil acceso, cada una con dos obstetras y tres pediatras. La familia obtiene los máximos beneficios del seguro de salud si se asocia con una clínica y selecciona ambos médicos de allí. ¿De cuántas formas se puede llevar a cabo lo anterior? Utilice O_1, O_2, O_3 y O_4 para representar a los obstetras y $P_1, \ldots, P_6$ para los pediatras. Así, se quiere el número de pares (O_i, P_j) para los que O_i y P_j se relacionan con la misma clínica. Debido a que hay cuatro obstetras, $n_1 = 4$, y para cada uno hay tres opciones de pediatra, entonces $n_2 = 3$. Al aplicar la regla del producto se obtiene $N = n_1n_2 = 12$ elecciones posibles. ■

Diagramas de árbol

En muchos problemas de conteo y probabilidad, se puede usar una configuración denominada *diagrama de árbol* para representar en forma gráfica todas las posibilidades. El diagrama de árbol relacionado con el ejemplo 2.18 se ilustra en la figura 2.7. Comenzando en un punto en el lado izquierdo del diagrama, para cada posible primer elemento de un par sale a la derecha un segmento de recta. Cada una de estas líneas se conoce como rama de primera generación. Ahora bien, para cualquier rama de primera generación, se construye otro segmento de recta que sale de la punta de la rama para cada elección posible de un segundo elemento del par. Cada uno de estos segmentos de recta es una rama de segunda generación. Puesto que hay cuatro obstetras, hay cuatro ramas de primera generación, y con tres pediatras por cada obstetra, se obtienen tres ramas de segunda generación que salen de cada rama de primera generación.

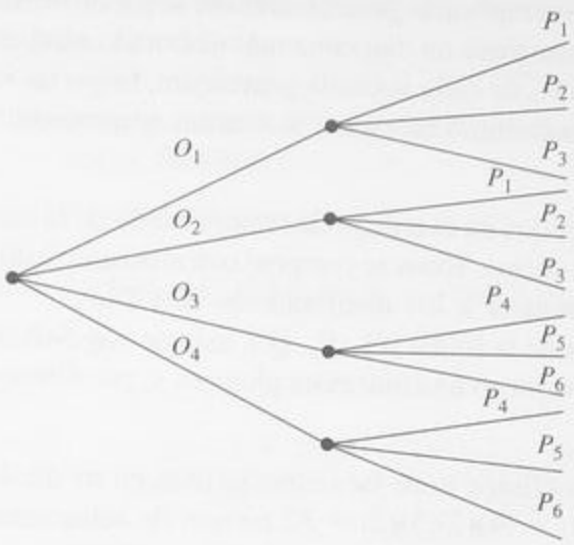


Figura 2.7 Diagrama de árbol para el ejemplo 2.18

Para generalizar, suponga que hay n_1 ramas de primera generación, y por cada una de éstas, hay n_2 ramas de segunda generación. Entonces el número total de ramas de segunda generación es n_1n_2. Puesto que el final de cada rama de segunda generación corresponde a exactamente un par posible (al elegir un primer elemento y, luego, un segundo se está al final de una rama de segunda generación), hay n_1n_2 pares, de modo que se comprueba la regla del producto.

La construcción de un diagrama de árbol no depende de tener el mismo número de ramas de segunda generación que salen de cada rama de primera generación. Si la segunda clínica tenía cuatro pediatras, entonces sólo habría tres ramas que salen de las dos ramas de primera generación y cuatro que salen de cada una de las otras dos ramas de primera generación. Por consiguiente, se puede usar un diagrama de árbol para representar en forma gráfica experimentos distintos a aquellos en los que se aplica la regla del producto.

Una regla de producto más general

Si un dado de seis lados se lanza cinco veces sucesivas en vez de dos, entonces cada resultado posible es un conjunto ordenado de cinco números, por ejemplo, (1, 3, 1, 2, 4) o (6, 5, 2, 2, 2). Una colección ordenada de k objetos se llama ***k-tupla*** (así que un par es una 2-tupla y una tripleta es una 3-tupla). Cada resultado del experimento de lanzar un dado es entonces una quíntupla.

Regla del producto para k-tuplas

Suponga que un conjunto consiste en colecciones ordenadas de k elementos (k-tuplas) y que hay n_1 opciones posibles para el primer elemento; para cada elección del primer elemento, hay n_2 elecciones posibles del segundo elemento; . . .; para cada elección posible de los primeros $k - 1$ elementos, hay n_k elecciones del k-ésimo elemento. Entonces hay $n_1n_2 \cdot \cdots \cdot n_k$ k-tuplas posibles.

Esta regla más general también se puede ilustrar mediante un diagrama de árbol; simplemente construya un diagrama más elaborado añadiendo ramas de tercera generación que salen de la punta de cada segunda generación, luego las ramas de cuarta generación, etcétera, hasta añadir por último las ramas de k-ésima generación.

Ejemplo 2.19 (continuación del ejemplo 2.17) Suponga que en el trabajo de remodelación de la casa se requiere comprar primero varios aparatos de cocina. Todos se compran con el mismo distribuidor y hay cinco distribuidores en el área. Si se designa a los distribuidores con $D_1, \ldots, D_5$, hay $N = n_1n_2n_3 = (5)(12)(9) = 540$ 3-tuplas de la forma (D_i, P_j, Q_k), así que hay 540 formas de elegir primero un distribuidor de aparatos, luego un contratista plomero y, por último, un contratista electricista. ■

Ejemplo 2.20 (continuación del ejemplo 2.18) Si cada clínica tiene tres especialistas en medicina interna y dos cirujanos generales, hay $n_1n_2n_3n_4 = (4)(3)(3)(2) = 72$ formas de seleccionar un médico de cada tipo, de modo que todos los médicos trabajen en la misma clínica. ■

Permutaciones

Hasta este punto, los elementos sucesivos de una k-tupla se eligieron de conjuntos distintos por completo (p. ej., distribuidores de aparatos, luego plomeros y, por último, electricistas). En varios lanzamientos de un dado, el conjunto a partir del cual se eligen los elementos sucesivos es siempre $\{1, 2, 3, 4, 5, 6\}$, pero las elecciones se hacen "con sustitución", para que el mismo elemento pueda aparecer más de una vez. Ahora se considera un conjunto fijo de n elementos distintos y se supone que una k-tupla se forma al seleccionar de manera sucesiva de este conjunto *sin sustitución* para que un elemento pueda aparecer en a lo sumo una de las k posiciones.

DEFINICIÓN

Cualquier secuencia ordenada de k objetos tomada de un conjunto de n objetos distintos se llama **permutación** de tamaño k de los objetos. El número de permutaciones de tamaño k que se puede construir a partir de n objetos se indica por medio de $P_{k,n}$.

El número de permutaciones de tamaño k se obtiene de inmediato de la regla general del producto. El primer elemento se puede seleccionar de n formas, para cada n formas el segundo elemento se puede elegir en $n - 1$ formas, y así sucesivamente; por último, para cada forma de elegir los primeros $k - 1$ elementos, el k-ésimo se puede elegir de $n - (k - 1) = n - k + 1$ formas, así que

$$P_{k,n} = n(n - 1)(n - 2) \cdot \cdots \cdot (n - k + 2)(n - k + 1)$$

Ejemplo 2.21 Hay diez asistentes de enseñanza disponibles para calificar los exámenes de un determinado curso. El primer examen consiste en cuatro preguntas, y el profesor desea seleccionar un asistente distinto para calificar cada pregunta (sólo un asistente por pregunta). ¿De cuántas maneras se puede elegir a los asistentes para calificar el examen? Aquí n = número de asistentes = 10 y k = número de preguntas = 4. Entonces el número de asignaciones de calificación diferente es $P_{4,10} = (10)(9)(8)(7) = 5040$. ■

El uso de la notación factorial permite expresar a $P_{k,n}$ en forma más compacta.

DEFINICIÓN

Para cualquier entero positivo m, $m!$ se lee "m factorial" y está definido por $m! = m(m-1)\cdot\cdots\cdot(2)(1)$. También, $0! = 1$.

Por medio de la notación factorial $(10)(9)(8)(7) = (10)(9)(8)(7)(6!)/6! = 10!/6!$ En términos generales,

$$P_{k,n} = n(n-1)\cdot\cdots\cdot(n-k+1) = \frac{n(n-1)\cdot\cdots\cdot(n-k+1)(n-k)(n-k-1)\cdot\cdots\cdot(2)(1)}{(n-k)(n-k-1)\cdot\cdots\cdot(2)(1)}$$

que se convierte en

$$P_{k,n} = \frac{n!}{(n-k)!}$$

Por ejemplo, $P_{3,9} = 9!/(9-3)! = 9!/6! = 9\cdot 8\cdot 7\cdot 6!/6! = 9\cdot 8\cdot 7$. Observe también que debido a que $0! = 1$, $P_{n,n} = n!/(n-n)! = n!/0! = n!/1 = n!$, como debe ser.

Combinaciones

Hay muchos problemas de conteo en los que se tiene un conjunto de n objetos distintos y se desea contar el número de subconjuntos *no ordenados* de tamaño k. Por ejemplo, en el *bridge* lo que importa son las 13 cartas de una mano y no el orden en que se reparten; al formar un comité, suele carecer de importancia el orden en que se listan los miembros.

DEFINICIÓN

Dado un conjunto de n objetos distintos, cualquier subconjunto no ordenado de tamaño k de los objetos se llama **combinación.** El número de combinaciones de tamaño k que se puede formar a partir de n objetos distintos se denota mediante $\binom{n}{k}$. (Esta notación es más común en probabilidad que $C_{k,n}$, que debe ser análoga a la notación para permutaciones.)

El número de combinaciones de tamaño k de un conjunto particular es más pequeño que el número de permutaciones porque, cuando no se considera el orden, un número de permutaciones corresponde a la misma combinación. Por ejemplo, considere el conjunto $\{A, B, C, D, E\}$ que consiste en cinco elementos. Hay $5!/(5-3)! = 60$ permutaciones de tamaño 3. Hay seis permutaciones de tamaño 3 que consisten en los elementos A, B y C puesto que estos tres se pueden ordenar en $3\cdot 2\cdot 1 = 3! = 6$ formas: (A, B, C), (A, C, B), (B, A, C), (B, C, A), (C, A, B) y (C, B, A). Estas seis permutaciones son equivalentes a la simple combinación $\{A, B, C\}$. De manera similar, para cualquier otra combinación de tamaño 3, hay 3! permutaciones, cada una obtenida al ordenar los tres objetos. Así,

$$60 = P_{3,5} = \binom{5}{3}\cdot 3!; \quad \text{así} \quad \binom{5}{3} = \frac{60}{3!} = 10$$

Estas diez combinaciones son

$$\{A, B, C\}\ \{A, B, D\}\ \{A, B, E\}\ \{A, C, D\}\ \{A, C, E\}\ \{A, D, E\},\ \{B, C, D\}$$
$$\{B, C, E\}\ \{B, D, E\}\ \{C, D, E\}$$

Cuando hay n objetos distintos, cualquier permutación de tamaño k se obtiene ordenando los k objetos no ordenados de una combinación en una de $k!$ formas, así que el número de permutaciones es el producto de $k!$ y el número de combinaciones. Esto da

$$\binom{n}{k} = \frac{P_{k,n}}{k!} = \frac{n!}{k!(n-k)!}$$

Observe que $\binom{n}{n} = 1$ y $\binom{n}{0} = 1$ ya que sólo hay una forma de elegir un conjunto de (todos los) n elementos o de ningún elemento, y $\binom{n}{1} = n$ puesto que hay n subconjuntos de tamaño 1.

Ejemplo 2.22 Una mano de *bridge* consiste en 13 cartas elegidas de una baraja de 52, sin considerar el orden. Hay $\binom{52}{13} = 52!/(13!39!)$ manos distintas de *bridge*, que son alrededor de 635 mil millones. Puesto que hay 13 cartas en cada palo, el número de manos que consisten por completo en tréboles o picas, o ambos (ninguna carta roja) es $\binom{26}{13} = 26!/(13!13!) = 10\,400\,600$. Una de estas $\binom{26}{13}$ manos consiste por completo en picas, y una consta nada más de tréboles, así que hay $[\binom{26}{13} - 2]$ manos que consisten por completo en tréboles y picas, con ambos palos representados en la mano. Suponga que se reparte una mano de *bridge* de un juego de naipes bien barajado (es decir, se eligen al azar 13 cartas de entre las 52 posibilidades) y sea

A = {la mano que consiste por completo en picas y tréboles con ambos palos representados}

B = {la mano que consiste en exactamente dos palos}

Los $N = \binom{52}{13}$ resultados posibles son equiprobables, por lo tanto

$$P(A) = \frac{N(A)}{N} = \frac{\binom{26}{13} - 2}{\binom{52}{13}} = 0.0000164$$

Debido a que hay $\binom{4}{2} = 6$ combinaciones que consisten en dos palos, de los que picas y tréboles son una combinación,

$$P(B) = \frac{6\left[\binom{26}{13} - 2\right]}{\binom{52}{13}} = 0.0000983$$

Es decir, una mano que consiste por completo en cartas de exactamente dos de los cuatro palos ocurrirá más o menos una vez en cada 10 000 manos. Si el lector juega *bridge* una vez al mes, es probable que nunca reciba esa mano. ■

Ejemplo 2.23 Un almacén de una universidad recibió un envío de 25 impresoras, de las cuales 10 son modelos láser y 15 de inyección de tinta. Si un técnico elige al azar seis de estas 25 impresoras para comprobar su funcionamiento, ¿cuál es la probabilidad de que exactamente tres de las elegidas sean impresoras láser (de modo que las otras tres sean de inyección de tinta)?

Sea D_3 = {exactamente tres de las seis elegidas son impresoras de inyección de tinta}. Suponiendo que cualquier conjunto particular de seis impresoras tiene la misma probabilidad de ser elegido que cualquier otro conjunto de seis, se tienen resultados equiprobables, por consiguiente $P(D_3) = N(D_3)/N$, donde N es el número de formas de elegir seis impresoras de las 25 y $N(D_3)$ es el número de formas de elegir tres impresoras láser y tres de inyección de tinta. Así, $N = \binom{25}{6}$. Para obtener $N(D_3)$, considere elegir primero tres de los 15 modelos de inyección de tinta y luego tres de las impresoras láser. Hay $\binom{15}{3}$ formas de elegir tres modelos de inyección de tinta y $\binom{10}{3}$ formas de escoger las tres impresoras láser; ahora $N(D_3)$ es el producto de estos dos números (imagine un diagrama de árbol, en realidad aquí se está usando un argumento de la regla del producto), por lo tanto

$$P(D_3) = \frac{N(D_3)}{N} = \frac{\binom{15}{3}\binom{10}{3}}{\binom{25}{6}} = \frac{\frac{15!}{3!12!} \cdot \frac{10!}{3!7!}}{\frac{25!}{6!19!}} = 0.3083$$

Sea D_4 = {exactamente 4 de las 6 impresoras seleccionadas son modelos de inyección de tinta} y defina D_5 y D_6 de una manera análoga. Entonces la probabilidad de que por lo menos sean elegidas tres impresoras de inyección de tinta es

$$P(D_3 \cup D_4 \cup D_5 \cup D_6) = P(D_3) + P(D_4) + P(D_5) + P(D_6)$$
$$= \frac{\binom{15}{3}\binom{10}{3}}{\binom{25}{6}} + \frac{\binom{15}{4}\binom{10}{2}}{\binom{25}{6}} + \frac{\binom{15}{5}\binom{10}{1}}{\binom{25}{6}} + \frac{\binom{15}{6}\binom{10}{0}}{\binom{25}{6}} = 0.8530$$

■

Ejercicios | Sección 2.3 (29 a 44)

29. El consejo de estudiantes de ingeniería de cierta universidad, tiene un representante en cada una de las cinco especializaciones de ingeniería (civil, eléctrica, industrial, de materiales y mecánica). ¿En cuántas formas se puede

- **a.** Seleccionar presidente y vicepresidente del consejo?
- **b.** Seleccionar un presidente, vicepresidente y secretario?
- **c.** Seleccionar dos miembros para el consejo del decano?

30. Un amigo ofrece una cena. Su provisión actual de vino son ocho botellas de zinfandel, 10 de merlot y 12 de cabernet (él sólo bebe vino tinto), todas de diferentes vinaterías.

- **a.** Si desea servir tres botellas de zinfandel y es importante la orden de servicio, ¿cuántas formas hay para hacer esto?
- **b.** Si se eligen al azar seis botellas de vino de las 30 para servir, ¿cuántas formas hay para hacer esto?
- **c.** Si se seleccionan al azar seis botellas, ¿cuántas formas hay para obtener dos botellas de cada variedad?
- **d.** Si se eligen al azar seis botellas, ¿cuál es la probabilidad de que esto dé como resultado que se elijan dos botellas de cada variedad?
- **e.** Si se seleccionan al azar seis botellas, ¿cuál es la probabilidad de que todas sean de la misma variedad.

31. **a.** Beethoven escribió nueve sinfonías y Mozart escribió 27 conciertos para piano. Si el locutor de una estación de radio universitaria desea reproducir primero una sinfonía de Beethoven y después un concierto de Mozart, ¿en cuántas formas se puede hacer esto?

- **b.** El gerente de la radiodifusora determina que en cada noche sucesiva (siete días por semana), se transmita una sinfonía de Beethoven, seguida de un concierto para piano de Mozart y después un cuarteto para cuerdas de Schubert (de los cuales hay 15). ¿Durante cuántos años podría continuar esta políti-

ca antes que se tenga que repetir el mismo programa?

32. Una cadena de tiendas de aparatos estereofónicos ofrece un precio especial por un conjunto completo de componentes (receptor, reproductor de discos compactos, bocinas y tocacintas). Para cada componente un comprador puede elegir el fabricante:

Receptor:	Kenwood, Onkyo, Pioneer, Sony, Sherwood
Reproductor de discos compactos:	Onkyo, Pioneer, Sony, Technics
Bocinas:	Boston, Infinity, Polk
Tocacintas:	Onkyo, Sony, Teac, Technics

Una pantalla de conmutador en la tienda permite al cliente reunir una selección de componentes (que consiste en uno de cada tipo). Use las reglas del producto para contestar las siguientes preguntas:

a. ¿De cuántas maneras se puede seleccionar un componente de cada tipo?
b. ¿De cuántas maneras se puede elegir a los componentes si el receptor y el reproductor de discos compactos deben ser Sony?
c. ¿De cuántas maneras se puede seleccionar a los componentes si ninguno tiene que ser Sony?
d. ¿De cuántas maneras se puede hacer una selección si por lo menos un componente debe ser Sony?
e. Si alguien acciona los interruptores del conmutador en forma aleatoria, ¿cuál es la probabilidad de que el sistema seleccionado contenga por lo menos un componente Sony? ¿Exactamente un componente Sony?

33. Poco después de comenzar a prestar servicio, algunos de los autobuses que fabricó cierta compañía presentan grietas en la parte inferior del bastidor principal; suponga que determinada ciudad tiene 25 de estos autobuses, y las grietas aparecieron en realidad en ocho de ellos.

a. ¿Cuántas formas hay de seleccionar una muestra de cinco autobuses de los 25 para una inspección completa?
b. ¿En cuántas formas una muestra de cinco autobuses contiene exactamente cuatro con grietas visibles?
c. Si se escoge al azar una muestra de cinco autobuses, ¿cual es la probabilidad de que exactamente cuatro de los cinco tengan grietas visibles?
d. Si se seleccionan los autobuses como en el inciso (c), ¿cuál es la probabilidad de que al menos cuatro de los autobuses seleccionados tengan grietas visibles?

34. Una planta de producción emplea 20 trabajadores en el turno diurno, 15 en el segundo turno y 10 en el de la noche. Un consultor de control de calidad selecciona seis de estos trabajadores para hacerles entrevistas exhaustivas. Supóngase que la selección se hace de tal forma que un determinado grupo de seis trabajadores tiene las mismas probabilidades de ser seleccionado que cualquier otro grupo (sacar seis trozos de papel sin sustitución de entre 45).

a. ¿Cuántas selecciones dan como resultado seis trabajadores del turno diurno? ¿Cuál es la probabilidad de que los seis trabajadores seleccionados sean del turno diurno?
b. ¿Cuál es la probabilidad de que los seis trabajadores seleccionados sean del mismo turno?
c. ¿Cuál es la probabilidad de que por lo menos dos turnos diferentes estén representados entre los trabajadores seleccionados?
d. ¿Cuál es la probabilidad de que por lo menos uno de los turnos no esté representado en la muestra de trabajadores?

35. Un departamento académico con cinco miembros de la facultad redujo su opción para jefe de departamento al candidato A o al candidato B. Luego, cada uno de los miembros votó en una papeleta por uno de los candidatos. Suponga que en realidad hay tres votos para A y dos para B. Si se eligen las papeletas para llevar el conteo en orden aleatorio, ¿cuál es la probabilidad de que A esté adelante de B todo el conteo de votos (p. ej., este evento ocurre si el orden seleccionado es $AABAB$, pero no para $ABBAA$)?

36. Un investigador estudia los efectos de la temperatura, presión y el tipo de catalizador en el rendimiento de cierta reacción química. Están en consideración tres temperaturas diferentes, cuatro presiones y cinco catalizadores distintos.

a. Si cualquier ejecución del experimento tiene que ver con el uso de una sola temperatura, presión y catalizador, ¿cuántas ejecuciones del experimento son posibles?
b. ¿En cuántas ejecuciones del experimento se usa la menor temperatura y las dos presiones más bajas?

37. Con referencia al ejercicio 36, suponga que se llevan a cabo cinco corridas diferentes el primer día de experimentación. Si las cinco se seleccionan al azar de entre todas las posibilidades, de modo que cualquier grupo de cinco tiene la probabilidad de selección, ¿cuál es la probabilidad de que se utilice un catalizador diferente en cada corrida?

38. Una caja de determinado almacén contiene cuatro focos de 40 W, cinco de 60 W y seis de 75 W. Suponga que se seleccionan al azar tres focos.

a. ¿Cuál es la probabilidad de que exactamente dos de los focos seleccionados sean de 75 W?

b. ¿Cuál es la probabilidad de que los tres focos seleccionados tengan la misma capacidad nominal?

c. ¿Cuál es la probabilidad de seleccionar un foco de cada capacidad en watts?

d. Ahora suponga que se eligen los focos uno por uno hasta encontrar uno de 75 W. ¿Cuál es la probabilidad necesaria para examinar por lo menos seis focos?

39. Se acaban de recibir 15 teléfonos en un centro de servicio autorizado. Cinco son celulares, cinco inalámbricos y los otros cinco son de cable. Suponga que a estos elementos se les asignan al azar los números 1, 2, . . . , 15 a fin de establecer el orden en que se van a reparar.

a. ¿Cuál es la probabilidad de que los teléfonos inalámbricos estén entre los diez primeros que se van a reparar?

b. ¿Cuál es la probabilidad de que después de revisar diez de los teléfonos, sólo queden pendientes teléfonos de dos de los tres tipos?

c. ¿Cuál es la probabilidad de que haya dos teléfonos de cada tipo entre los seis primeros que se van a reparar?

40. Tres moléculas tipo A, tres B, tres C, y tres D se deben combinar para formar una molécula en cadena. Una de estas moléculas en cadena es $ABCDABCDABCD$, y otra es $BCDDAAABDBCC$.

a. ¿Cuántas moléculas en cadena hay? (*Sugerencia*: si se pudieran distinguir entre sí las tres letras A, A_1, A_2, A_3 y también las letras B, C y D, ¿cuántas moléculas habría? ¿Cómo se reduce este número cuando se eliminan de las letras A los subíndices?

b. Suponga que se elige al azar una molécula del tipo descrito. ¿Cuál es la probabilidad de que las tres moléculas de cada tipo terminen una junto a otra (como en $BBBAAADDDCCC$)?

41. Una profesora de matemáticas desea programar una cita con cada uno de sus ocho asistentes de enseñanza, cuatro hombres y cuatro mujeres, para analizar su curso de cálculo. Suponga que los ordenamientos posibles de citas tengan las mismas probabilidades de ser seleccionados.

a. ¿Cuál es la probabilidad de que por lo menos una asistente esté entre los tres primeros con quien se reúne la profesora?

b. ¿Cuál es la probabilidad de que después de las cinco primeras citas, la maestra se haya reunido con todas las asistentes?

c. Supóngase que la maestra tiene los mismos ocho asistentes el semestre siguiente y otra vez programa citas sin considerar el orden durante el primer semestre. ¿Cuál es la probabilidad de que el orden de las citas sea diferente?

42. Tres parejas de casados compraron boletos para el teatro y se sientan en una fila que consiste sólo en seis asientos. Si eligen sus asientos de un modo totalmente aleatorio (orden aleatorio), ¿cuál es la probabilidad de que Jim y Paula (marido y mujer) elijan los dos asientos del extremo izquierdo? ¿Cuál es la probabilidad de que Jim y Paula se sienten juntos? ¿Cuál es la probabilidad de que, por lo menos, una de las esposas termine sentada junto a su marido?

43. En el póquer de cinco cartas, una escalera consiste en cinco cartas con denominaciones adyacentes (p. ej., 9 de tréboles, 10 de corazones, sota de corazones, reina de picas y rey de tréboles). Si se supone que los ases pueden ser altos o bajos, y recibe una mano de cinco cartas, ¿cuál es la probabilidad de que sea una escalera con carta alta de 10? ¿Cuál es la probabilidad de que sea una escalera? ¿Cuál es la probabilidad de que sea una escalera con cartas del mismo palo?

44. Demuestre que $\binom{n}{i} = \binom{n}{n-i}$. Interprete en términos de subconjuntos.

2.4 Probabilidad condicional

Las probabilidades asignadas a distintos eventos dependen de lo que se sepa acerca de la situación experimental cuando se hace la asignación. Después de la asignación inicial podría estar disponible información parcial pertinente o relacionada con el resultado del experimento. Es posible que la información dé lugar a una revisión de las asignaciones de probabilidad. Para un evento particular A, se emplea $P(A)$ para representar la probabilidad asignada a A; ahora se considera que $P(A)$ es la probabilidad original o incondicional del evento A.

En esta sección, se examina cómo la información "ocurrió un evento B" afecta a la probabilidad asignada a A. Por ejemplo, A se podría referir a un individuo que tiene una determinada enfermedad en presencia de ciertos síntomas. Si se lleva a cabo una prueba de sangre en el individuo y el resultado es negativo (B = prueba de sangre negativa), entonces cambia la

probabilidad de tener la enfermedad (debe disminuir, pero por lo común no a cero, puesto que las pruebas de sangre no son infalibles). Se usa la notación $P(A \mid B)$ para representar la **probabilidad condicional de A, dado que ocurrió el evento B.**

Ejemplo 2.24 En una planta en la que utilizan dos líneas de ensamble diferentes, A y A', se ensamblan componentes complejos. En la línea A se utiliza equipo más antiguo que en A', así que es un poco más lenta y menos confiable. Suponga que en un determinado día en la línea A se ensamblan ocho componentes, de los cuales dos se identifican como defectuosos (B) y seis como no defectuosos (B'), en tanto que en A' se obtuvieron un componente defectuoso y nueve no defectuosos. Esta información se resume en la tabla siguiente.

		Condición	
		B	B'
Línea	A	2	6
	A'	1	9

Sin conocer esta información, el gerente de ventas elige al azar uno de estos 18 componentes para una demostración. Antes de la demostración

$$P(\text{de elegir un componente de la línea}) = P(A) = \frac{N(A)}{N} = \frac{8}{18} = 0.44$$

Sin embargo, si resulta defectuoso el componente elegido, entonces ocurrió el evento B, así que el componente debe haber sido uno de los tres de la columna B de la tabla. Puesto que estos tres componentes son equiprobables entre sí después que ocurrió,

$$P(A \mid B) = \frac{2}{3} = \frac{\frac{2}{18}}{\frac{3}{18}} = \frac{P(A \cap B)}{P(B)} \tag{2.2}$$

■

En la ecuación (2.2), la probabilidad condicional se expresa como una relación de probabilidades incondicionales: el numerador es la probabilidad de la intersección de los dos eventos, mientras que el denominador es la probabilidad del evento condicional B. Esta relación se ilustra en un diagrama de Venn (figura 2.8).

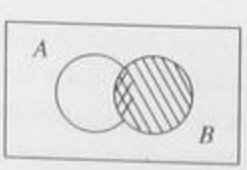


Figura 2.8 Motivación de la definición de probabilidad condicional

Dado que ocurrió B, el espacio muestral pertinente ya no es $\mathcal{S}$ sino que consiste en los resultados de B; A ocurre si y sólo si ocurre uno de los resultados de la intersección, así que la probabilidad condicional de A dada B es proporcional a $P(A \cap B)$. La constante de proporcio-

nalidad $1/P(B)$ se usa para asegurar que la probabilidad $P(B \mid B)$ del nuevo espacio muestral B sea igual a 1.

Definición de probabilidad condicional

En el ejemplo 2.24 se demuestra que cuando los resultados son equiprobables, el cálculo de las probabilidades condicionales se puede basar en la intuición. Sin embargo, cuando los experimentos son más complicados podría fallar la intuición, así que se desea tener una definición general de probabilidad condicional que produzca respuestas intuitivas en problemas simples. El diagrama de Venn y la ecuación (2.2) hacen pensar en la definición apropiada.

DEFINICIÓN

Para dos eventos cualesquiera A y B con $P(B) > 0$, la **probabilidad condicional de A dado que ocurrió B** se define mediante la ecuación

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{2.3}$$

Ejemplo 2.25 Suponga que de todos los individuos que compran determinada cámara digital, 60% incluyen una tarjeta de memoria opcional en su compra, 40% incluyen una batería y 30% una tarjeta y una batería. Considere elegir al azar un comprador y sea $A = \{\text{compró una tarjeta de memoria}\}$ y $B = \{\text{compró una batería}\}$. Entonces, $P(A) = 0.60$, $P(B) = 0.40$ y $P(\text{compró ambas cosas}) = P(A \cap B) = 0.30$. Dado que los individuos seleccionados compraron una batería extra, la probabilidad de que también sea comprada una tarjeta opcional es

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.30}{0.40} = 0.75$$

Es decir, de todos los que compran una batería extra, 75% compraron una tarjeta de memoria adicional. De manera similar,

$$P(\text{batería} \mid \text{tarjeta de memoria} = P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.30}{0.60} = 0.50$$

Observe que $P(A \mid B) \neq P(A)$ y $P(B \mid A) \neq P(B)$. ■

Ejemplo 2.26 Una revista de noticias publica tres columnas tituladas "Art" (A), "Books" (B) y "Cinema" (C). Los hábitos de lectura de un lector elegido al azar con respecto a estas columnas son

Probabilidad de leer con regularidad	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
	0.14	0.23	0.37	0.08	0.09	0.13	0.05

(Véase la figura 2.9.)

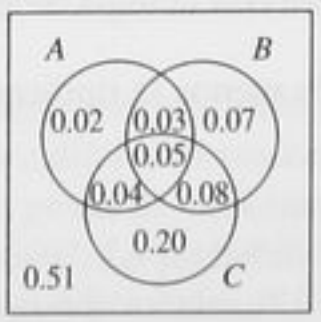


Figura 2.9 Diagrama de Venn para el ejemplo 2.26

Por lo tanto, se tiene

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.08}{0.23} = 0.348$$

$$P(A \mid B \cup C) = \frac{P(A \cap (B \cup C))}{P(B \cup C)} = \frac{0.04 + 0.05 + 0.03}{0.47} = \frac{0.12}{0.47} = 0.255$$

$$P(A \mid \text{lee por lo menos una}) = P(A \mid A \cup B \cup C) = \frac{P(A \cap (A \cup B \cup C))}{P(A \cup B \cup C)}$$

$$= \frac{P(A)}{P(A \cup B \cup C)} = \frac{0.14}{0.49} = 0.286$$

y

$$P(A \cup B \mid C) = \frac{P((A \cup B) \cap C)}{P(C)} = \frac{0.04 + 0.05 + 0.08}{0.37} = 0.459$$

■

Regla de la multiplicación para $P(A \cap B)$

La definición de probabilidad condicional produce el resultado siguiente, obtenido al multiplicar ambos lados de la ecuación (2.3) por $P(B)$.

Regla de la multiplicación

$$P(A \cap B) = P(A \mid B) \cdot P(B)$$

Esta regla es importante, porque por lo común se desea $P(A \cap B)$ en tanto que $P(B)$ y $P(A \mid B)$ se pueden especificar a partir de la descripción del problema. Al tomar en cuenta $P(B \mid A)$ se obtiene $P(A \cap B) = P(B \mid A) \cdot P(A)$.

Ejemplo 2.27 Cuatro individuos respondieron a una petición de un banco de sangre para hacer donaciones. Ninguno de ellos ha sido donador antes, así que se desconoce su tipo de sangre. Suponga que sólo se desea el tipo O+ y en realidad sólo uno de los cuatro tiene este tipo. Si se eligen al azar los posibles donadores para la tipificación, ¿cuál es la probabilidad de que por lo menos se deba tipificar a tres individuos para obtener el tipo deseado?

Con la identificación $B = \{\text{primer tipo no O+}\}$ y $A = \{\text{segundo tipo no O+}\}$, $P(B) = \frac{3}{4}$. Dado que el primer tipo no es O+, dos de los tres individuos restantes no son O+, así que $P(A \mid B) = \frac{2}{3}$. Ahora con la regla de la multiplicación, se obtiene

$$P(\text{por lo menos se tipifican tres individuos}) = P(A \cap B)$$
$$= P(A \mid B) \cdot P(B)$$
$$= \frac{2}{3} \cdot \frac{3}{4} = \frac{6}{12}$$
$$= 0.5$$

■

La regla de la multiplicación es más útil cuando el experimento consiste en varias etapas en sucesión. El evento condicional B describe entonces el resultado de la primera etapa, y A el resultado de la segunda, así que por lo común se conoce $P(A \mid B)$ a condición de lo que ocurra primero. La regla se amplía sin dificultad a experimentos que tienen que ver con más de dos etapas. Por ejemplo,

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_3 \mid A_1 \cap A_2) \cdot P(A_1 \cap A_2) \\ &= P(A_3 \mid A_1 \cap A_2) \cdot P(A_2 \mid A_1) \cdot P(A_1) \end{aligned} \tag{2.4}$$

donde A_1 ocurre primero, seguida de A_2, y, por último, A_3.

Ejemplo 2.28 Para el experimento de tipificación de sangre del ejemplo 2.27,

$$\begin{aligned} P(\text{el tercer tipo es O+}) &= P(\text{el tercero es} \mid \text{el primero no es} \cap \text{el segundo no es}) \\ &\quad \cdot P(\text{el segundo no es} \mid \text{el primero no es}) \cdot P(\text{el primero no es}) \\ &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{4} = 0.25 \end{aligned}$$

■

Cuando el experimento de interés consiste en una secuencia de varias etapas, es conveniente representarlas con un diagrama de árbol. Una vez que se tiene un diagrama de árbol apropiado, las probabilidades y las probabilidades condicionales se introducen en las distintas ramas; con esto el uso repetido de la regla de la multiplicación es bastante directo.

Ejemplo 2.29 Una cadena de tiendas de video vende tres marcas de videograbadoras (VCR). De sus ventas de VCR, 50% son de la marca 1 (la menos cara), 30% son de la marca 2 y 20% de la marca 3. Cada fabricante ofrece un año de garantía en las partes y la mano de obra. Se sabe que 25% de las VCR de la marca 1 requieren trabajo de garantía, en tanto que los porcentajes correspondientes para las marcas 2 y 3 son 20% y 10%, respectivamente.

1. ¿Cuál es la probabilidad de que un comprador elegido al azar haya comprado una VCR de la marca 1 que necesite reparación mientras está en garantía?
2. ¿Cuál es la probabilidad de que un comprador elegido al azar tenga una VCR que requiera reparación mientras esté en garantía?
3. Si un cliente vuelve a la tienda con una VCR que necesite se haga válida la garantía, ¿cuál es la probabilidad de que sea una VCR de la marca 1? ¿Una VCR de la marca 2? ¿Una VCR de la marca 3?

La primera etapa del problema tiene que ver con un cliente que elige una de las tres marcas de VCR. Sea $A_i = \{\text{se compra la marca } i\}$, para $i = 1, 2$ y 3. Luego $P(A_1) = 0.50$, $P(A_2) = 0.30$ y $P(A_3) = 0.20$. Una vez que se elige una marca de VCR, la segunda etapa requiere observar si la VCR elegida requiere hacer válida la garantía. Con $B = \{\text{necesita reparación}\}$ y $B' = \{\text{no requiere reparación}\}$, la información suministrada implica que $P(B \mid A_1) = 0.25$, $P(B \mid A_2) = 0.20$ y $P(B \mid A_3) = 0.10$.

El diagrama de árbol que representa esta situación experimental se muestra en la figura 2.10. Las ramas iniciales corresponden a diferentes marcas de videograbadoras; hay dos ramas de segunda generación que salen de la punta de cada rama inicial, una para "requiere reparación", y otra para "no necesita reparación". La probabilidad $P(A_i)$ aparece en la i-ésima rama inicial, en tanto que las probabilidades condicionales $P(B \mid A_i)$ y $P(B' \mid A_i)$ aparecen en las ramas de segunda generación. A la derecha de cada rama de segunda generación que correspon-

de a la ocurrencia de B, se muestra el producto de probabilidades en las ramas que salen para ese punto. Esto es simplemente la regla de la multiplicación en acción. Por lo tanto, la respuesta a la pregunta planteada en 1 es $P(A_1 \cap B) = P(B \mid A_1) \cdot P(A_1) = 0.125$. La respuesta a la pregunta dos es

$$\begin{aligned} P(B) &= P[\text{(marca 1 y reparación) o (marca 2 y reparación) o (marca 3 y reparación)}] \\ &= P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) \\ &= 0.125 + 0.060 + 0.020 = 0.205 \end{aligned}$$

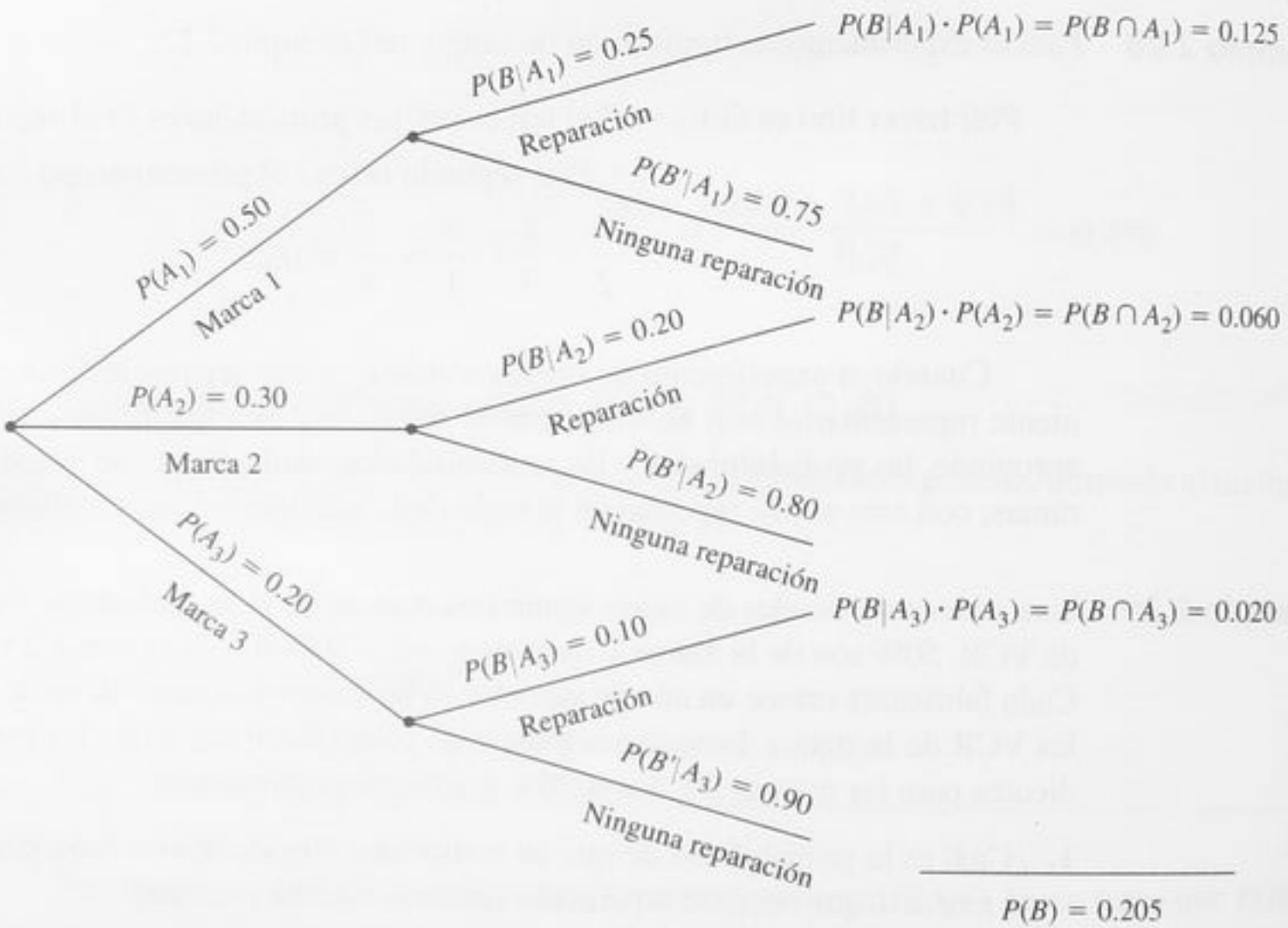


Figura 2.10 Diagrama de árbol para el ejemplo 2.29

Por último,

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{0.125}{0.205} = 0.61$$

$$P(A_2 \mid B) = \frac{P(A_2 \cap B)}{P(B)} = \frac{0.060}{0.205} = 0.29$$

y

$$P(A_3 \mid B) = 1 - P(A_1 \mid B) - P(A_2 \mid B) = 0.10$$

Observe que la probabilidad inicial o *probabilidad anterior* de la marca 1 es 0.50, en tanto que una vez que se sabe que la VCR elegida requiere reparación, la *probabilidad posterior* de la marca 1 se incrementa a 0.61. Esto se debe a que las VCR de la marca 1 tienen más pro-

babilidades de requerir reparación que las otras marcas. La probabilidad posterior de la marca 3 es $P(A_3 \mid B) = 0.10$, que es mucho menor que la probabilidad anterior $P(A_3) = 0.20$. ■

Teorema de Bayes

El cálculo de la probabilidad posterior $P(A_j \mid B)$ a partir de probabilidades anteriores $P(A_i)$ y probabilidades condicionales $P(B \mid A_i)$ ocupa una posición central en la probabilidad elemental. La regla general para estos cálculos, que en realidad es sólo una aplicación simple de la regla de la multiplicación, se remonta al reverendo Thomas Bayes, quien vivió en el siglo XVIII. Para expresarla se requiere primero otro resultado. Recuerde que los eventos $A_1, \ldots, A_k$ son mutuamente excluyentes si no hay dos con resultados comunes. Los eventos son *exhaustivos* si debe ocurrir una A_i, así que $A_1 \cup \cdots \cup A_k = \mathcal{S}$.

Ley de la probabilidad total

Sean $A_1, \ldots, A_k$ eventos mutuamente excluyentes y exhaustivos. Entonces para cualquier otro evento B,

$$P(B) = P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_k)P(A_k) = \sum_{i=1}^{k} P(B \mid A_i)P(A_i) \tag{2.5}$$

Demostración

Debido a que las A_i son mutuamente excluyentes y exhaustivas, si ocurre B debe ser junto con exactamente una de las A_i Es decir, $B = (A_1 \text{ y } B) \text{ o } \ldots \text{ o } (A_k \text{ y } B) = (A_1 \cap B) \cup \cdots \cup (A_k \cap B)$, donde los eventos $(A_i \cap B)$ son mutuamente excluyentes. Esta "partición de B" se ilustra en la figura 2.11. Así,

$$P(B) = \sum_{i=1}^{k} P(A_i \cap B) = \sum_{i=1}^{k} P(B \mid A_i)P(A_i)$$

como se deseaba.

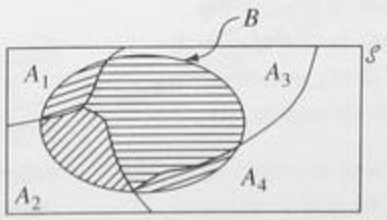


Figura 2.11 Partición de B mediante las A_i mutuamente excluyentes y exhaustivas ■

Un ejemplo del uso de la ecuación (2.5) apareció en la respuesta de la pregunta dos del ejemplo 2.29, donde $A_1 = \{\text{marca 1}\}$, $A_2 = \{\text{marca 2}\}$, $A_3 = \{\text{marca 3}\}$ y $B = \{\text{reparación}\}$.

Teorema de Bayes

Sea $A_1, A_2, \ldots, A_k$ una colección de k eventos mutuamente excluyentes y exhaustivos con $P(A_i) > 0$ para $i = 1, \ldots, k$. Entonces para cualquier otro evento B para el cual $P(B) > 0$,

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B \mid A_j)P(A_j)}{\sum_{i=1}^{k} P(B \mid A_i) \cdot P(A_i)} \qquad j = 1, \ldots, k \tag{2.6}$$

La transición de la segunda a la tercer expresión en (2.6) se basa en usar la regla de la multiplicación en el numerador y la ley de la probabilidad total en el denominador.

La proliferación de eventos y subíndices en (2.6) puede ser un poco intimidante para los recién llegados a la probabilidad. Mientras haya relativamente pocos eventos en la partición, se puede usar un diagrama de árbol (como en el ejemplo 2.29) como base para calcular las probabilidades posteriores sin referirse alguna vez de forma explícita al teorema de Bayes.

Ejemplo 2.30 *Incidencia de una enfermedad rara.* Sólo 1 de 1000 adultos se ve afectado por una enfermedad rara para la cual se ha desarrollado una prueba diagnóstica. La prueba es tal, que cuando un individuo en realidad tiene la enfermedad, ocurre un resultado positivo 99% de las veces, en tanto que un individuo sin la enfermedad presenta un resultado positivo sólo 2% de las veces. Si se aplica la prueba a un individuo seleccionado al azar y el resultado es positivo, ¿cuál es la probabilidad de que el individuo tenga la enfermedad?

Para usar el teorema de Bayes, sea $A_1 = \{\text{el individuo tiene la enfermedad}\}$, $A_2 = \{\text{el individuo no tiene la enfermedad}\}$ y $B = \{\text{resultado de prueba positivo}\}$. Entonces $P(A_1) = 0.001$, $P(A_2) = 0.999$, $P(B \mid A_1) = 0.99$ y $P(B \mid A_2) = 0.02$. El diagrama de árbol para este problema se ilustra en la figura 2.12.

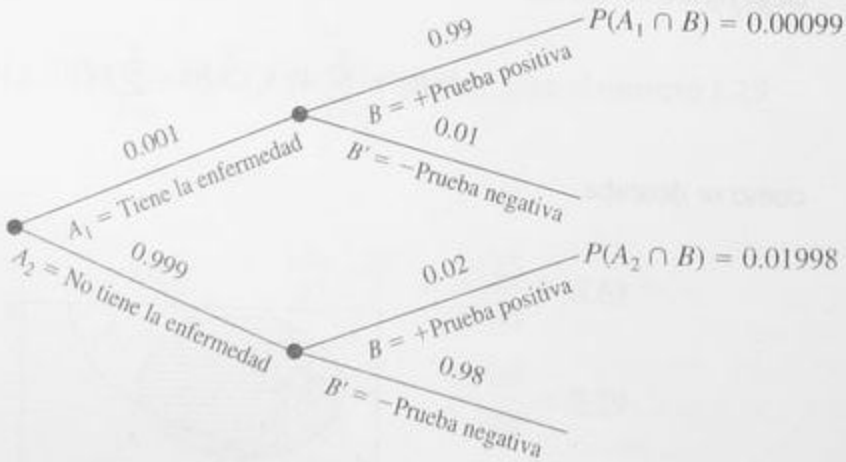


Figura 2.12 Diagrama de árbol para el problema de la enfermedad rara

Junto a cada rama que corresponde a un resultado de prueba positivo, la regla de la multiplicación produce las probabilidades registradas. Por lo tanto, $P(B) = 0.00099 + 0.01998 = 0.02097$, de donde se tiene

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{0.00099}{0.02097} = 0.047$$

Este resultado parece ser contraintuitivo; al parecer la prueba diagnóstica es tan precisa, que se espera que alguien con un resultado positivo tenga altas probabilidades de padecer la enfermedad, en tanto que la probabilidad condicional calculada sólo es 0.047. Sin embargo, debido a que la enfermedad es rara y la prueba sólo tiene una confiabilidad moderada, la mayor parte de los resultados de prueba positivos surgen de errores y no de individuos enfermos. La probabilidad de tener la enfermedad se incrementó por un factor multiplicativo de 47 (del 0.001 anterior al 0.047 posterior); pero para obtener un incremento más en la probabilidad posterior, se requiere una prueba diagnóstica con índices de error mucho más pequeños. Si la enfermedad no fuera tan rara (p. ej., incidencia de 25% en la población), entonces los índices de error para la prueba actual proporcionarían buenos diagnósticos. ■

Ejercicios | Sección 2.4 (45 a 67)

45. La población de determinado país consiste en tres grupos étnicos. Cada individuo pertenece a uno de los cuatro grupos sanguíneos principales. En la tabla siguiente de *probabilidad conjunta* se dan las proporciones de individuos de las distintas combinaciones de grupo étnico y grupo sanguíneo.

		Grupo sanguíneo			
		O	A	B	AB
Grupo étnico	1	0.082	0.106	0.008	0.004
	2	0.135	0.141	0.018	0.006
	3	0.215	0.200	0.065	0.020

Suponga que se elige al azar un individuo de la población, y defina los eventos con A = {se eligió el tipo A}, B = {se seleccionó el tipo B y C = {se eligieron tres grupos étnicos}.

a. Calcule $P(A)$, $P(C)$ y $P(A \cap C)$.

b. Calcule $P(A \mid C)$ y $P(C \mid A)$, y explique lo que representa cada una de estas probabilidades.

c. Si el individuo elegido no tiene sangre tipo B, ¿cuál es la probabilidad de que sea del grupo étnico 1?

46. Suponga que se selecciona al azar un individuo de la población de los adultos varones que viven en Estados Unidos. Sea A el evento de que el individuo seleccionado tenga una estatura de más de 6 pies, y sea B el evento de que el individuo seleccionado sea un jugador profesional de baloncesto. ¿Cuál considera que es mayor, $P(A \mid B)$ o $P(B \mid A)$? ¿Por qué?

47. Regrese al escenario de las tarjetas de crédito del ejercicio 12 (sección 2.2), donde A = {Visa}, B = {MasterCard}, $P(A) = 0.5$, $P(B) = 0.4$ y $P(A \cap B) = 0.25$. Calcule e interprete cada una de las probabilidades siguientes (un diagrama de Venn podría servir).

a. $P(B \mid A)$ **b.** $P(B' \mid A)$

c. $P(A \mid B)$ **d.** $P(A' \mid B)$

e. Debido a que el individuo seleccionado tiene al menos una tarjeta, ¿cuál es la probabilidad de que tenga una tarjeta Visa?

48. Considere de nuevo la situación de defectos del sistema descrita en el ejercicio 26 (sección 2.2).

a. Dado que el sistema tiene un defecto tipo 1, ¿cuál es la probabilidad de que tenga un defecto tipo 2?

b. Dado que el sistema tiene un defecto tipo 1, ¿cuál es la probabilidad de que tenga los tres tipos de defectos?

c. Dado que el sistema tiene al menos un tipo de defecto, ¿cuál es la probabilidad de que tenga exactamente un tipo de defecto?

d. Dado que el sistema tiene los dos primeros tipos de defectos, ¿cuál es la probabilidad de que no tenga el tercer tipo de defecto?

49. Si se eligen al azar dos focos de la caja del ejercicio 38 (sección 2.3) y se encuentra que por lo menos uno de ellos es de 75 W, ¿cuál es la probabilidad de que los dos sean de 75 W? Dado que por lo menos uno de los dos focos elegidos no es de 75 W, ¿cuál es la probabilidad de que ambos tengan el mismo valor nominal?

50. Una tienda de departamentos vende camisas deportivas en tres tallas (pequeña, mediana y grande), en tres modelos (a cuadros, estampadas y de franjas) y con dos largos de manga (corta y larga). En las tablas siguientes se dan las proporciones de camisas vendidas de las distintas combinaciones de categorías.

Manga corta

	Modelo		
Talla	**Cuadros**	**Estampada**	**Franjas**
Pequeña	0.04	0.02	0.05
Mediana	0.08	0.07	0.12
Grande	0.03	0.07	0.08

Manga larga

	Modelo		
Talla	**Cuadros**	**Estampada**	**Franjas**
Pequeña	0.03	0.02	0.03
Mediana	0.10	0.05	0.07
Grande	0.04	0.02	0.08

a. ¿Cuál es la probabilidad de que la siguiente camisa vendida sea mediana, de manga larga y estampada?
b. ¿Cuál es la probabilidad de que la siguiente camisa vendida sea mediana y estampada?
c. ¿Cuál es la probabilidad de que la siguiente camisa vendida sea de manga corta? ¿Una de manga larga?
d. ¿Cuál es la probabilidad de que la talla de la siguiente camisa vendida sea mediana? ¿Que el modelo de la siguiente camisa vendida sea estampado?
e. Dado que la camisa recién vendida era a cuadros y de manga corta, ¿cuál es la probabilidad de que su talla fuera mediana?
f. Dado que la camisa recién vendida era a cuadros y mediana, ¿cuál es la probabilidad de que fuera de manga corta? ¿De manga larga?

51. Una caja contiene seis bolas rojas y cuatro verdes, y una segunda caja contiene siete bolas rojas y tres verdes. Se escoge al azar una bola de la primera caja y se coloca en la segunda caja. Después, se selecciona al azar una bola de la segunda caja y se pone en la primera caja.
a. ¿Cuál es la probabilidad de que se seleccione una bola roja de la primera caja y una bola roja de la segunda?
b. En la conclusión del proceso de selección, ¿cuál es la probabilidad de que los números de bolas rojas y verdes de la primera caja sean idénticos a los números del comienzo?

52. Un sistema de bombeo consta de dos bombas idénticas, #1 y #2. Si falla una bomba, el sistema todavía funciona. Sin embargo, debido al esfuerzo, ahora la bomba restante tiene más probabilidades de fallar que al principio. Es decir, $r = P(\#2 \text{ falla} \mid \#1 \text{ falla}) > P(\#2 \text{ falla}) = q$. Si por lo menos falla una bomba al final de su vida útil de diseño en 7% de los sistemas y ambas bombas fallan durante ese periodo en sólo 1%, ¿cuál es la probabilidad de que la bomba #1 falle durante su vida útil de diseño?

53. Cierto taller repara componentes de audio y video. Sea A el evento en el cual el siguiente componente llevado a reparación sea uno de audio, y sea B el evento en el cual el siguiente componente sea un reproductor de discos compactos (así que el evento B está contenido en A). Suponga que $P(A) = 0.6$ y $P(B) = 0.05$. ¿Cuál es $P(B \mid A)$?

54. En el ejercicio 13, $A_i = \{\text{proyecto } i \text{ otorgado}\}$, para $i = 1, 2, 3$. Utilice las probabilidades proporcionadas allí para calcular las siguientes probabilidades:
a. $P(A_2 \mid A_1)$
b. $P(A_2 \cap A_3 \mid A_1)$
c. $P(A_2 \cup A_3 \mid A_1)$
d. $P(A_1 \cap A_2 \cap A_3 \mid A_1 \cup A_2 \cup A_3)$.
Exprese en palabras la probabilidad que calculó.

55. En el ejercicio 42, seis personas (tres parejas de casados) escogen asientos al azar en una fila que consta de seis asientos.
a. Utilice la regla de la multiplicación para calcular la probabilidad de que Jim y Paula se sienten juntos en el extremo izquierdo (evento A) y que John y Mary Lou (marido y mujer) se sienten juntos en medio (evento B).
b. Dado que John y Mary Lou se sientan juntos en medio, ¿cuál es la probabilidad de que los otros dos esposos se sienten junto a sus esposas?
c. Dado que John y Mary Lou se sientan juntos, ¿cuál es la probabilidad de que todos los esposos se sienten junto a sus esposas?

56. Si $P(B \mid A) > P(B)$, demuestre que $P(B' \mid A) < P(B')$. (*Sugerencia*: sume $P(B' \mid A)$ a ambos lados de la desigualdad y luego use el resultado del ejercicio 57.)

57. Para dos eventos cualesquiera A y B con $P(B) > 0$, demuestre que $P(A \mid B) + P(A' \mid B) = 1$.

58. Demuestre que para tres eventos cualesquiera A, B y C con $P(C) > 0$, $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$.

59. En cierta gasolinera, 40% de los clientes utilizan gasolina regular sin plomo (A_1), 35% usan gasolina extra sin plomo (A_2), y 25% gasolina premium sin plomo (A_3). De los clientes que consumen gasolina regular, sólo 30% llenan sus tanques (evento B). De los que compran gasolina extra, 60% llenan sus tanques, en tanto que quienes llevan gasolina premium, 50% llenan sus tanques.

a. ¿Cuál es la probabilidad de que el siguiente cliente pida gasolina extra sin plomo y llene su tanque ($A_2 \cap B$)?

b. ¿Cuál es la probabilidad de que el siguiente cliente llene el tanque?

c. Si el siguiente cliente llena el tanque, ¿cuál es la probabilidad de que pida gasolina regular? ¿Extra? ¿Premium?

60. Setenta por ciento de los aviones ligeros que desaparecen mientras vuelan en cierto país son descubiertos después. De las aeronaves descubiertas, 60% tienen localizador de emergencia, mientras que 90% de las no descubiertas no tienen un localizador. Suponga que desaparece un avión ligero.

a. Si tiene localizador de emergencia, ¿cuál es la probabilidad de que no sea localizado?

b. Si no tiene localizador de emergencia, ¿cuál es la probabilidad de que sea localizado?

61. Se envían componentes de cierto tipo a un proveedor en lotes de diez. Suponga que 50% de los lotes no contienen componentes defectuosos, 30% contienen un componente defectuoso y 20% dos componentes defectuosos. Se eligen al azar y se prueban dos componentes de un lote. ¿Cuáles son las probabilidades relacionadas con 0, 1 y 2 componentes defectuosos que están en el lote bajo cada una de las siguientes condiciones?

a. Ningún componente probado está defectuoso.

b. Uno de los dos componentes probados es defectuoso. (*Sugerencia*: dibuje un diagrama de árbol con tres ramas de primera generación para los tres tipos de lotes.)

62. Una empresa que fabrica cámaras de video produce un modelo básico y uno de lujo. El año pasado, 40% de las cámaras vendidas fueron del modelo básico. De los compradores del modelo básico, 30% compran una garantía ampliada, en tanto que 50% de los compradores del modelo de lujo hacen lo mismo. Si se sabe que un comprador seleccionado al azar tiene una garantía ampliada, ¿qué tan probable es que tenga un modelo básico?

63. Para clientes que compran un juego completo de neumáticos en cierta distribuidora, considere los eventos siguientes:

A = {los neumáticos son fabricados en Estados Unidos}

B = {el comprador solicita de inmediato el balanceo de los neumáticos}

C = {el comprador pide alineación de neumáticos delanteros}

junto con A', B' y C'. Suponga las siguientes probabilidades incondicionales y condicionales:

$P(A) = 0.75$ $P(B \mid A) = 0.9$ $P(B \mid A') = 0.8$

$P(C \mid A \cap B) = 0.8$ $P(C \mid A \cap B') = 0.6$

$P(C \mid A' \cap B) = 0.7$ $P(C \mid A' \cap B') = 0.3$

a. Construya un diagrama de árbol que consta de ramas de primera generación, segunda y tercera y ponga una marca de evento y la probabilidad adecuada junto a cada rama.

b. Calcule $P(A \cap B \cap C)$.

c. Calcule $P(B \cap C)$.

d. Calcule $P(C)$.

e. Calcule $P(A \mid B \cap C)$, la probabilidad de una compra de neumáticos fabricados en Estados Unidos, dado que se pidió alineación y balanceo.

64. En el ejemplo 2.30, suponga que la tasa de incidencia para la enfermedad es 1 en 25 y no 1 en 1000. ¿Cuál es la probabilidad de un resultado de prueba positivo? Dado que el resultado de prueba es positivo, ¿cuál es la probabilidad de que el individuo tenga la enfermedad? Con un resultado de prueba negativo, ¿cuál es la probabilidad de que el individuo no tenga la enfermedad?

65. En una gran universidad, en la interminable búsqueda de un libro de texto satisfactorio, el departamento de estadística ha probado un texto diferente en cada uno de los tres últimos trimestres. Durante el trimestre de otoño, 500 estudiantes emplearon el texto del profesor Mean; durante el trimestre de invierno, 300 utilizaron el texto del profesor Median; y durante el trimestre de primavera, 200 usaron el texto del profesor Mode. Un estudio al final de cada trimestre, demostró que 200 estudiantes estuvieron satisfechos con el libro de Mean, 150 con el libro de Median y 160 con el de Mode. Si se selecciona al azar un estudiante que tomó estadística durante uno de estos trimestres y admite estar satisfecho con el texto, ¿es más probable que el estudiante haya empleado el libro de Mean, Median o Mode? ¿Quién

es el autor menos probable? (*Sugerencia*: dibuje un diagrama de árbol o utilice el teorema de Bayes.)

66. Una amiga que vive en Los Ángeles realiza viajes frecuentes de consultoría a Washington, D.C.; 50% de las veces viaja en la aerolínea #1, 30% en la aerolínea #2 y el 20% restante de las veces en la aerolínea #3. Para la aerolínea #1, los vuelos llegan con retraso a Washington D.C. 30% de las veces y a L.A. 10% de las veces. Para la aerolínea #2, estos porcentajes son 25 y 20%, mientras que para la aerolínea #3 los porcentajes son 40 y 25%. Si se sabe que en un determinado viaje ella llega tarde a exactamente uno de los dos destinos, ¿cuáles son las probabilidades posteriores de haber volado en las aerolíneas #1, #2 y #3? (*Sugerencia*: de la punta de cada rama de primera generación en un diagrama de árbol, dibuje tres ramas de segunda generación marcadas, respectivamente, 0 retrasado, 1 retrasado y 2 retrasado.)

67. En el ejercicio 59, considere la siguiente información adicional acerca del uso de las tarjetas de crédito:

70% de los clientes que llenan su tanque con gasolina regular, usan una tarjeta de crédito.

50% de los clientes que consumen gasolina regular y no llenan su tanque, usan tarjeta de crédito.

60% de los clientes que llenan su tanque con gasolina extra, usan tarjeta de crédito.

50% de los clientes que compran gasolina extra pero no llenan su tanque, usan tarjeta de crédito.

50% de los clientes que llenan su tanque con gasolina premium, usan tarjeta de crédito.

40% de los clientes que consumen gasolina premium, pero no llenan su tanque, usan tarjeta de crédito.

Calcule la probabilidad de cada uno de los siguientes eventos para el siguiente cliente que llegue (un diagrama de árbol puede ser útil).

a. {extra, tanque lleno, tarjeta de crédito}
b. {premium, no llena el tanque y usa tarjeta de crédito}
c. {premium y tarjeta de crédito}
d. {tanque lleno y tarjeta de crédito}
e. {tarjeta de crédito}
f. Si el siguiente cliente usa una tarjeta de crédito, ¿cuál es la probabilidad de que pida gasolina premium?

2.5 Independencia

La definición de probabilidad condicional permite revisar la probabilidad $P(A)$ que al principio se asigna a A cuando después se sabe que ocurrió otro evento B; la nueva probabilidad de A es $P(A \mid B)$. En los ejemplos, solía suceder que $P(A \mid B)$ era distinta a la probabilidad incondicional $P(A)$, indicación de que la información "ocurrió B" produjo un cambio en la probabilidad de la ocurrencia de A. Sin embargo, hay otras situaciones en las que la probabilidad de que ocurra, o ya haya ocurrido, A no resultara afectada si se sabe que ocurrió B, así que $P(A \mid B) = P(A)$. Entonces es natural pensar en A y B como eventos independientes, lo que significa que la ocurrencia o no ocurrencia de un evento no tiene nada que ver con la probabilidad de que ocurra el otro.

DEFINICIÓN

Dos eventos A y B son **independientes** si $P(A \mid B) = P(A)$ de lo contrario son **dependientes**.

La definición de independencia podría parecer "asimétrica", porque no se requiere que también $P(B \mid A) = P(B)$ Sin embargo, si se emplea la definición de probabilidad condicional y la regla de la multiplicación,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A)} \tag{2.7}$$

El lado derecho de la ecuación (2.7) es $P(B)$ si y sólo si $P(A \mid B) = P(A)$ (independencia), así que la igualdad en la definición implica la otra igualdad (y viceversa). También es directo demostrar que si A y B son independientes, entonces también lo son los siguientes pares de eventos: (1) A' y B, (2) A y B' y (3) A' y B'.

Ejemplo 2.31 Considere lanzar un dado no cargado y defina los eventos $A = \{2, 4, 6\}$, $B = \{1, 2, 3\}$ y $C = \{1, 2, 3, 4\}$. Entonces se tiene $P(A) = \frac{1}{2}$, $P(A \mid B) = \frac{1}{3}$ y $P(A \mid C) = \frac{1}{2}$. Es decir, los eventos A y B son dependientes, mientras que los eventos A y C son independientes. Por intuición, si se lanza el dado y se sabe que el resultado fue 1, 2, 3 o 4 (ocurrió C), entonces la probabilidad de que ocurra A es $\frac{1}{2}$, como fue al principio, ya que dos de los cuatro eventos importantes son pares y los resultados aún son equiprobables. ■

Ejemplo 2.32 Sea A y B dos eventos cualesquiera mutuamente excluyentes con $P(A) > 0$. Por ejemplo, para un automóvil elegido al azar, sea $A = \{$el automóvil es de cuatro cilindros$\}$ y $B = \{$el automóvil tiene seis cilindros$\}$. Puesto que los eventos son mutuamente excluyentes, si ocurre B, entonces es posible que no ocurra A, así que $P(A \mid B) = 0 \neq P(A)$. El mensaje aquí es que si *dos eventos son mutuamente excluyentes, no pueden ser independientes*. Cuando A y B son mutuamente excluyentes, la información de que ocurrió A dice algo acerca del evento B (éste no pudo haber ocurrido), así que se imposibilita la independencia. ■

$P(A \cap B)$ cuando los eventos son independientes

Con frecuencia, la naturaleza de un experimento hace pensar que dos eventos A y B se deben suponer independientes. Por ejemplo, éste es el caso si un fabricante recibe una tarjeta de circuitos de cada uno de dos distribuidores diferentes, cada tarjeta se prueba al recibirla, y $A = \{$la primera es defectuosa$\}$ y $B = \{$la segunda es defectuosa$\}$. Si $P(A) = 0.1$, también debe ser el caso que $P(A \mid B) = 0.1$; el hecho de saber la condición de la segunda tarjeta no debe proporcionar información acerca de la condición de la primera. En el resultado siguiente se muestra cómo calcular $P(A \cap B)$ cuando los eventos son independientes.

PROPOSICIÓN

A y B son independientes si y sólo si

$$P(A \cap B) = P(A) \cdot P(B) \tag{2.8}$$

En otras palabras, A y B son eventos independientes si y sólo si la probabilidad de que ocurran ambos $(A \cap B)$ es el producto de cada una de las probabilidades. La comprobación es como sigue:

$$P(A \cap B) = P(A \mid B) \cdot P(B) = P(A) \cdot P(B) \tag{2.9}$$

donde la segunda igualdad de la ecuación (2.9) es válida si y sólo si A y B son independientes. Como resultado de la equivalencia de independencia con la ecuación (2.8), esta última se puede usar como una definición de independencia.

Ejemplo 2.33 Se sabe que 30% de las lavadoras de cierta compañía requieren servicio mientras está vigente la garantía, en tanto que sólo 10% de sus secadoras necesitan este servicio. Si alguien compra una lavadora y una secadora de esta compañía, ¿cuál es la probabilidad de que ambas máquinas requieran servicio de garantía?

Sea A el evento en el cual la lavadora necesite servicio mientras está vigente la garantía y sea B el evento definido de manera análoga para la secadora. Entonces, $P(A) = 0.30$ y $P(B) = 0.10$. Suponiendo que las dos máquinas funcionan de modo independiente, la probabilidad deseada es

$$P(A \cap B) = P(A) \cdot P(B) = (0.30)(0.10) = 0.03$$

La probabilidad de que ninguna máquina requiera servicio es

$$P(A' \cap B') = P(A') \cdot P(B') = (0.70)(0.90) = 0.63$$

■

Ejemplo 2.34 Cada día, de lunes a viernes, un lote de componentes que envía un primer proveedor, llega a cierta instalación de inspección. Dos días después, también llega un lote de un segundo proveedor. Ochenta por ciento de los lotes del proveedor 1 pasan la inspección, y también 90% de los lotes del proveedor 2. ¿Cuál es la probabilidad de que, en un día elegido al azar, dos lotes pasen la inspección? Para responder esta pregunta, se supone que en los días cuando se prueban dos lotes, si pasa el primer lote es independiente de si también pasa el segundo. En la figura 2.13 se muestra la información pertinente.

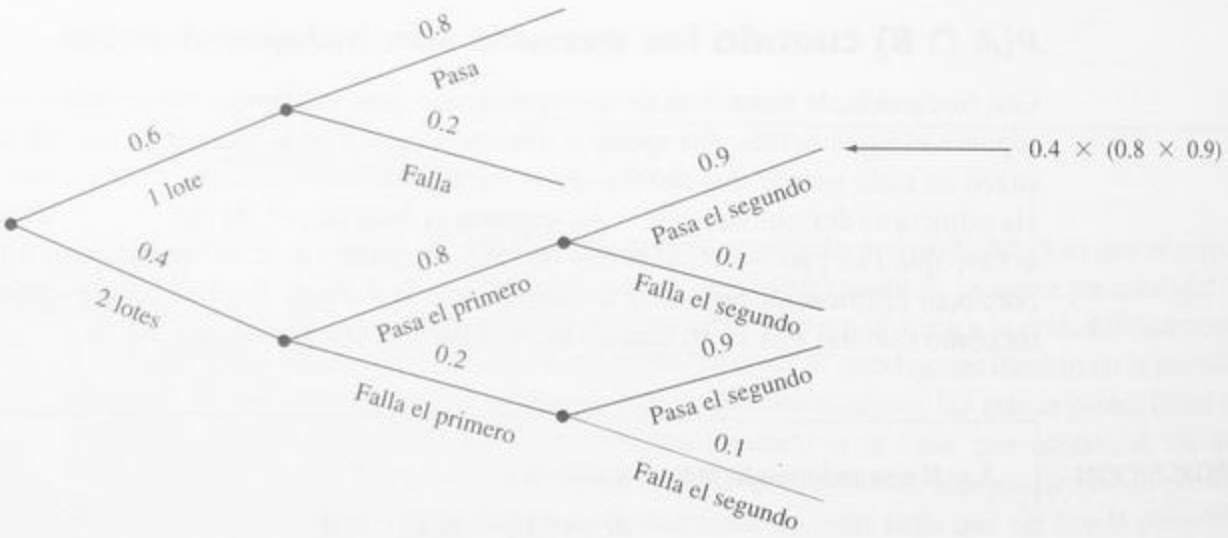


Figura 2.13 Diagrama de árbol para el ejemplo 2.34

$$\begin{aligned} P(\text{pasan dos}) &= P(\text{dos recibidos} \cap \text{ambos pasan}) \\ &= P(\text{ambos pasan} \mid \text{dos recibidos}) \cdot P(\text{dos recibidos}) \\ &= [(0.8)(0.9)](0.4) = 0.288 \end{aligned}$$

■

Independencia de más de dos eventos

El concepto de independencia de dos eventos se puede ampliar a colecciones de más de dos eventos. Aunque es posible ampliar la definición para dos eventos independientes trabajando en términos de probabilidades condicionales e incondicionales, es más directo y menos problemático proceder de acuerdo con lo que se establece en la última proposición.

DEFINICIÓN

Los eventos $A_1, \ldots, A_n$ son **mutuamente independientes** si para toda k ($k = 2, 3, \ldots, n$) y todo subconjunto de índices $i_1, i_2, \ldots, i_k$,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \cdots \cdot P(A_{i_k}).$$

Para parafrasear la definición, los eventos son mutuamente independientes si la probabilidad de intersección de cualquier subconjunto de n eventos es igual al producto de cada una de las probabilidades. Como fue el caso con dos eventos, es común especificar al comienzo de un problema la independencia de ciertos eventos. Así, la definición se puede usar para calcular la probabilidad de una intersección.

Ejemplo 2.35 En el artículo "Reliability Evaluation of Solar Photovoltaic Arrays" (*Solar Energy*, 2002: 129-141) se presentan varias configuraciones de sistemas fotovoltaicos solares que consisten en celdas solares de silicio cristalino. Considere primero el sistema ilustrado en la figura 2.14(a). Hay dos subsistemas conectados en paralelo, cada uno con tres celdas. A fin de que funcione el sistema, debe funcionar por lo menos uno de los dos subsistemas en paralelo. Dentro de cada subsistema, las tres celdas están conectadas en serie, así que un subsistema funciona sólo si trabajan las celdas del subsistema. Considere un determinado valor de tiempo de vida t_0. Sea A_i el evento en el que el tiempo de vida de la celda i excede a t_0 ($i = 1, 2, \ldots, 6$). Se supone que los A_i son eventos independientes (si determinada celda dura más de t_0 horas no condiciona que pase lo mismo con otra celda o no) y que $P(A_i) = 0.9$ para toda i puesto que son idénticas las celdas. Entonces,

$$\begin{aligned} P(\text{el tiempo de vida del sistema exceda a } t_0) &= P[(A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_5 \cap A_6)] \\ &= P(A_1 \cap A_2 \cap A_3) + P(A_4 \cap A_5 \cap A_6) \\ &\quad - P[(A_1 \cap A_2 \cap A_3) \cap (A_4 \cap A_5 \cap A_6)] \\ &= (0.9)(0.9)(0.9) + (0.9)(0.9)(0.9) \\ &\quad - (0.9)(0.9)(0.9)(0.9)(0.9)(0.9) = 0.927 \end{aligned}$$

De manera alternativa,

$$\begin{aligned} P(\text{el tiempo de vida del sistema exceda a } t_0) &= 1 - P(\text{ambas vidas de los subsistemas sean} \le t_0) \\ &= 1 - [P(\text{la vida del subsistema sea} \le t_0)]^2 \\ &= 1 - [1 - P(\text{la vida del subsistema sea} > t_0)]^2 \\ &= 1 - [1 - (0.9)^3]^2 = 0.927 \end{aligned}$$

A continuación considere el sistema de enlaces cruzados total que se ilustra en la figura 2.14(b), obtenido del sistema serie-paralelo al conectar cada columna de uniones. Ahora el sistema

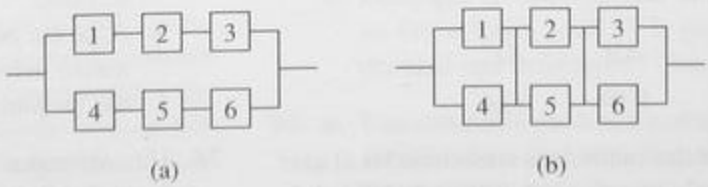


Figura 2.14 Configuraciones del sistema para el ejemplo 2.35: (a) serie-paralelo; (b) enlace cruzado total

falla siempre que falle toda la columna, y el tiempo de vida del sistema excede a t_0 sólo si lo hace cada columna. Para esta configuración,

$$\begin{aligned}P(\text{el tiempo de vida del sistema sea por lo menos } t_0) &= [P(\text{el tiempo de vida de la columna sea mayor que } t_0)]^3\\ &= [1 - P(\text{el tiempo de vida de la columna sea} \le t_0)]^3\\ &= [1 - P(\text{ambas celdas de una columna tengan un tiempo de vida} \le t_0)]^3\\ &= [1 - (1 - 0.9)^2]^3 = 0.970\end{aligned}$$

Ejercicios | Sección 2.5 (68 a 87)

68. Considere de nuevo la situación de las tarjetas de crédito del ejercicio 47, sección 2.4, y demuestre que A y B son dependientes, primero con la definición de independencia y después comprobando que no se cumple la propiedad de la multiplicación.

69. Una compañía de prospección petrolera tiene dos proyectos activos, uno en Asia y otro en Europa. Sea A el evento donde el proyecto asiático tiene éxito y B el evento donde el proyecto europeo sea exitoso. Suponga que A y B son eventos independientes con $P(A) = 0.4$ y $P(B) = 0.7$.

- **a.** Si fracasa el proyecto asiático, ¿cuál es la probabilidad de que también fracase el proyecto europeo? Explique su razonamiento.
- **b.** ¿Cuál es la probabilidad de que por lo menos uno de los proyectos tenga éxito?
- **c.** Dado que por lo menos uno de los dos proyectos es exitoso, ¿cuál es la probabilidad de que sólo el proyecto asiático tenga éxito?

70. En el ejercicio 13, ¿alguna A_i es independiente de cualquier otra A_j? Conteste utilizando la propiedad de la multiplicación para eventos independientes.

71. Si A y B son eventos independientes, demuestre que también lo son A' y B. [*Sugerencia*: establezca primero la relación entre $P(A' \cap B)$, $P(B)$ y $P(A \cap B)$.]

72. Suponga que las proporciones de fenotipos sanguíneos en determinada población son como se indica a continuación:

A	B	AB	O
0.42	0.10	0.04	0.44

Si los fenotipos de dos individuos seleccionados al azar son independientes entre sí, ¿cuál es la probabilidad de que ambos fenotipos sean O? ¿Cuál es la probabilidad de que los fenotipos de dos individuos elegidos al azar sean iguales?

73. Una de las hipótesis que sustentan la teoría de gráficas de control (véase el capítulo 16) es que los puntos sucesivos son independientes entre sí. Cada punto graficado indica que un proceso de manufactura está trabajando de manera correcta o que hay algún mal funcionamiento. Aun cuando un proceso esté trabajando bien, hay una pequeña probabilidad de que determinado punto indique un problema en el proceso. Suponga que esta probabilidad es 0.05. ¿Cuál es la probabilidad de que por lo menos uno de 10 puntos sucesivos indique un problema cuando en realidad el proceso trabaja en forma correcta? Conteste esta pregunta para 25 puntos sucesivos.

74. La probabilidad de que un alumno cometa un error de marcación en una pregunta de un examen de opción múltiple es 0.1. Si hay diez preguntas, y éstas se marcan de manera independiente, ¿cuál es la probabilidad de que no se cometan errores? ¿De que al menos se cometa un error? Si hay n preguntas y la probabilidad de cometer un error es p y no 0.1, dé expresiones para estas dos probabilidades.

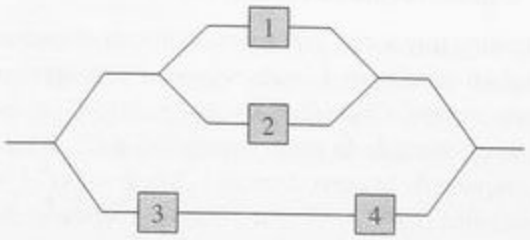
75. La junta de un avión requiere 25 remaches. La junta se tiene que volver a hacer si está defectuoso cualquiera de los remaches. Suponga que cada uno de los remaches tiene la misma probabilidad de estar defectuoso y que no depende de los demás.

- **a.** Si 20% de las juntas se tiene que volver a hacer, ¿cuál es la probabilidad de que un remache esté defectuoso?
- **b.** ¿Qué tan pequeña debe ser la probabilidad de un remache defectuoso para asegurar que sólo 10% de todas las juntas se tengan que volver a realizar?

76. Un calentador tiene cinco válvulas de alivio idénticas. La probabilidad de que en algún momento se abra determinada válvula es 0.95. Si la operación de las válvulas es independiente, calcule P(por lo menos se abre una válvula) y P(por lo menos no se abre una válvula).

77. Dos bombas conectadas en paralelo fallan en determinado día sin que haya dependencia mutua. La probabilidad de que sólo falle la bomba más vieja es 0.10, y la probabilidad de que sólo falle la bomba más nueva es 0.05. ¿Cuál es la probabilidad de que el sistema de bombeo falle cualquier día (esto sucede si fallan ambas bombas)?

78. Considere el sistema de componentes conectados como en la figura siguiente. Los componentes 1 y 2 están conectados en paralelo, de modo que el subsistema trabaja si y sólo si funciona 1 o 2; puesto que 3 y 4 están conectados en serie, ese subsistema trabaja si y sólo si trabajan 3 y 4. Si los componentes funcionan sin depender entre sí y P(el componente funciona) = 0.9, calcule P(el sistema funciona).



79. Refiérase de nuevo a la configuración del sistema serie-paralelo presentado en el ejemplo 2.35, y suponga que sólo hay dos celdas en vez de tres en cada subsistema en paralelo [en la figura 2.14(a) elimine las celdas 3 y 6, y reenumere las celdas 4 y 5 como 3 y 4]. Con $P(A_i) = 0.9$, es claro que la probabilidad de que el tiempo de vida del sistema exceda a t_0 es 0.9639. ¿Hasta qué valor se tendría que cambiar 0.9 a fin de incrementar la confiabilidad del tiempo de vida del sistema de 0.9639 a 0.99? (*Sugerencia*: sea $P(A_i) = p$, exprese la confiabilidad del sistema en términos de p, y luego sea $x = p^2$.)

80. Considere lanzar de manera independiente dos dados, uno rojo y otro verde. Sea A el evento donde el dado rojo muestra 3 puntos, B el evento donde el dado verde muestra 4 puntos y C el evento donde el número total de puntos que aparece en los dos dados es 7. ¿Son estos eventos independientes por pares (es decir, A y B son eventos independientes, A y C son independientes y B y C son independientes)? ¿Los tres eventos son mutuamente independientes?

81. Dos inspectores comprueban si tienen defectos los componentes que llegan a un distribuidor (ambos inspectores comprueban cada componente). El primer inspector detecta 90% de los componentes defectuosos, al igual que el segundo inspector. Por lo menos un inspector no detecta un defecto en 20% de los componentes defectuosos?
   a. ¿Cuál es la probabilidad de que sólo el primer inspector detecte un componente defectuoso? ¿Cuál es la probabilidad de que exactamente uno de los dos inspectores detecte un componente defectuoso?
   b. ¿Cuál es la probabilidad de que ambos inspectores no detecten los tres componentes defectuosos de un lote (suponiendo que las inspecciones de componentes diferentes son independientes entre sí)?

82. Sesenta por ciento de los automóviles examinados en un centro de verificación pasan la prueba. Si se supone que automóviles sucesivos pasan la prueba o no sin dependencia mutua, calcule las siguientes probabilidades:
   a. P(pasan los tres automóviles contiguos).
   b. P(por lo menos es rechazado uno de tres automóviles contiguos inspeccionados).
   c. P(exactamente pasa uno de los tres automóviles contiguos inspeccionados).
   d. P(a lo sumo pasa uno de los tres automóviles contiguos inspeccionados).
   e. Dado que por lo menos uno de los tres automóviles contiguos pasa la verificación, ¿cuál es la probabilidad de que pasen los tres (una probabilidad condicional)?

83. Un inspector de control de calidad busca defectos en artículos recién producidos. El inspector revisa un artículo en una serie de etapas independientes, cada una con duración fija. Si hay un defecto, sea p la probabilidad de que la falla sea detectada durante cualquiera de las etapas (este modelo se estudia en "Human Performance in Sampling Inspection", *Human Factors*, 1979: 99-105).
   a. Si se supone que un artículo tiene un defecto, ¿cuál es la probabilidad de que sea detectado al final de la segunda etapa (una vez detectada una falla termina la secuencia de etapas)?
   b. Dé una expresión para la probabilidad de que se detecte una falla al final de la n-ésima etapa.
   c. Si cuando en tres etapas no ha sido detectada una falla, el artículo pasa, ¿cuál es la probabilidad de que un artículo con falla pase la inspección?
   d. Suponga que 10% de los artículos tienen un defecto [P(el artículo escogido al azar está defectuoso) = 0.1]. Con la suposición del inciso (c), ¿cuál es la probabilidad de que un artículo seleccionado al azar pase la inspección (pasa de forma automática si no está defectuoso, pero también podría pasar si está defectuoso)?
   e. Dado que un artículo pasó la inspección (sin fallas en tres etapas), ¿cuál es la probabilidad de que en realidad esté defectuoso? Calcule para $p = 0.5$.

84. a. Una compañía maderera acaba de recibir un lote de 10 000 tablas de 2 × 4. Suponga que 20% de estas tablas (2000) están demasiado verdes para ser utilizadas en construcción de primera calidad. Se eligen al azar dos tablas, una después de la otra. Sea A = {la primera tabla está verde} y B = {la segunda tabla está verde}. Calcule $P(A)$, $P(B)$ y $P(A \cap B)$ (un

diagrama de árbol podría ser útil). ¿Son independientes A y B?

b. Con A y B independientes y $P(A) = P(B) = 0.2$, ¿cuál es $P(A \cap B)$? ¿Cuánta diferencia hay entre esta respuesta y $P(A \cap B)$ del inciso (a)? Para calcular $P(A \cap B)$ ¿se puede suponer que A y B del inciso (a) son independientes para obtener en esencia la probabilidad correcta?

c. Supóngase que el lote consta de diez tablas, de las cuales dos están verdes. ¿La suposición de independencia produce ahora más o menos la respuesta correcta para $P(A \cap B)$? ¿Cuál es la diferencia crítica entre la situación aquí y la del inciso (a)? ¿Cuándo se considera que una suposición de independencia es válida para obtener una respuesta más o menos correcta para $P(A \cap B)$?

85. Refiérase a las suposiciones indicadas en el ejercicio 78 y conteste la pregunta formulada para el sistema de la figura siguiente. ¿Cómo cambiaría la probabilidad si se tratara de un subsistema conectado en paralelo al subsistema mostrado en la figura 2.14(a)?

86. El profesor Stan der Deviation puede tomar una de dos rutas en su camino de regreso a casa desde el trabajo. En la primera ruta hay cuatro cruces de ferrocarril. La probabilidad de que se detenga ante el paso de un tren en cualquiera de los cruces es 0.1, y los trenes operan de manera independiente en los cuatro cruces. La otra ruta es más larga pero sólo hay dos cruces, independientes entre sí, con la misma probabilidad de detención para cada uno que en la primera ruta. En determinado día, el profesor Deviation tiene una junta en casa programada para cierta hora. Sin importar qué ruta tome, calcula que llegará tarde si el tren se cruza en su camino por lo menos en la mitad de los cruces.

a. ¿Cuál ruta debe tomar para reducir la probabilidad de llegar tarde a la junta?

b. Si lanza al aire una moneda para determinar la ruta y llega tarde, ¿cuál es la probabilidad de que haya tomado la ruta con cuatro cruces?

87. Suponga que a una zorra se le colocan etiquetas idénticas en las orejas. Luego, se suelta a la zorra durante cierto tiempo. Considere los dos eventos $C_1 = \{$se pierde la etiqueta de la oreja izquierda$\}$ y $C_2 = \{$se pierde la etiqueta de la oreja derecha$\}$. Sea $\pi = P(C_1) = P(C_2)$ y suponga que C_1 y C_2 son eventos independientes. Obtenga una expresión (donde aparezca π) para tener la probabilidad de que exactamente se pierda una etiqueta dado que a lo sumo se pierde una ("Ear Tag Loss in Red Foxes", *J. Wildlife Mgmt.*, 1976: 164-167). (*Sugerencia*: trace un diagrama de árbol en el que las dos ramas iniciales se refieran a si se pierde la etiqueta de la oreja izquierda.)

Ejercicios complementarios (88 a 111)

88. Una pequeña empresa manufacturera comenzará a operar un turno nocturno. La compañía emplea a 20 maquinistas.

a. Si una cuadrilla consta de 3 maquinistas, ¿cuántas cuadrillas diferentes son posibles?

b. Si se utiliza $1, 2, \ldots, 20$ para clasificar a los maquinistas en orden de competencia, ¿cuántas de estas cuadrillas no tendrían al mejor maquinista?

c. ¿Cuántas de las cuadrillas tendrían por lo menos uno de los 10 mejores maquinistas?

d. Si se selecciona al azar una de estas cuadrillas para trabajar en determinada noche, ¿cuál es la probabilidad de que el mejor maquinista no trabaje esa noche?

89. Una fábrica utiliza tres líneas de producción para fabricar latas de cierto tipo. En la tabla siguiente se dan los porcentajes de latas que no cumplen con las especificaciones, para cada una de las tres líneas durante un periodo particular.

	Línea 1	Línea 2	Línea 3
Mancha	15	12	20
Fisura	50	44	40
Problema con la pestaña	21	28	24
Defecto superficial	10	8	15
Otro	4	8	2

Durante este periodo, en la línea 1 se produjeron 500 latas defectuosas, 400 en la línea 2 y 600 en la línea 3. Suponga que se elige al azar una de estas 1500 latas.

a. ¿Cuál es la probabilidad de que la lata sea de la línea 1? ¿De que la razón de que no cumpla con las especificaciones sea una fisura?
b. Si la lata seleccionada proviene de la línea 1, ¿cuál es la probabilidad de que tuviera una mancha?
c. Dado que la lata elegida tuvo un defecto superficial, ¿cuál es la probabilidad de que proviniera de la línea 1?

90. Un empleado de la oficina de registros de cierta universidad tiene diez formas en su escritorio para ser procesadas. Seis de éstas son solicitudes de cancelación, y las otras cuatro son peticiones de cambio de curso.
a. Si el empleado elige al azar seis de esas formas para pasarlas a un subordinado, ¿cuál es la probabilidad de que quede en su escritorio sólo uno de los dos tipos de formas?
b. Suponga que tiene tiempo para procesar sólo cuatro de las formas antes de salir. Si se selecciona al azar una por una, ¿cuál es la probabilidad de que cada una de las formas sucesivas sea distinta a la anterior?

91. Se tiene programado lanzar un satélite desde Cabo Cañaveral Florida y otro desde una base de la Fuerza Aérea en Vandenberg, California. Consideré A para marcar que el lanzamiento en Vandenberg fue según lo programado y B para representar que el evento en Cabo Cañaveral también fue a tiempo. Si A y B son eventos independientes con $P(A) > P(B)$ y $P(A \cup B) = 0.626$, $P(A \cap B) = 0.144$, determine los valores de $P(A)$ y $P(B)$.

92. Un transmisor está enviando un mensaje mediante un código binario, es decir, una secuencia de ceros y unos. Cada bit transmitido (0 o 1) debe pasar por tres relevadores para llegar al receptor. En cada relevador se tiene una probabilidad 0.20 de que el bit enviado sea diferente del bit recibido (una inversión). Suponga que los relevadores operan de forma independiente entre sí.

Transmisor $\longrightarrow$ relevador 1 $\longrightarrow$ relevador 2 $\longrightarrow$ relevador 3 $\longrightarrow$ receptor

a. Si se envía un 1 desde el transmisor, ¿cuál es la probabilidad de que los tres relevadores envíen un 1?
b. Si se envía un 1 desde el transmisor, ¿cuál es la probabilidad de que el receptor reciba un 1? (*Sugerencia*: los ocho resultados experimentales se pueden mostrar en un diagrama de árbol con tres generaciones de ramas, una generación por cada relevador.)
c. Suponga que 70% de los bits enviados desde el transmisor son unos. Si el receptor recibe un 1, ¿cuál es la probabilidad de que se haya enviado un 1?

93. El individuo A tiene un círculo de cinco amigos cercanos (B, C, D, E y F). A oyó por casualidad un rumor e invitó a sus cinco amigos a una reunión para hacer circular el rumor. Para comenzar, A selecciona al azar a uno de los cinco y le comenta el rumor. Luego, ese individuo selecciona al azar a uno de los cuatro individuos restantes y repite el rumor. Continuando, el individuo que acaba de escuchar elige a un nuevo individuo de los que no han oído el rumor hasta que todos lo han oído.
a. ¿Cuál es la probabilidad de que el rumor sea repetido en el orden B, C, D, E y F?
b. ¿Cuál es la probabilidad de que F sea la tercera persona de la reunión a quien se cuente el rumor?
c. ¿Cuál es la probabilidad de que F sea la última persona en escuchar el rumor?

94. Refiérase al ejercicio 93. Si en cada etapa la persona que actualmente "tiene" el rumor no sabe quién ya lo escuchó y selecciona al azar al siguiente receptor de entre los cinco posibles individuos, ¿cuál es la probabilidad de que F aún no haya oído el rumor después de haberlo contado diez veces en la reunión?

95. Un ingeniero químico está interesado en determinar si cierta impureza está presente en un producto. Un experimento tiene una probabilidad de 0.80 de detectar la impureza si está presente. La probabilidad de no detectar la impureza si está ausente es 0.90. Las probabilidades previas de la impureza presente y ausente son 0.40 y 0.60, respectivamente. Tres experimentos separados dan como resultado sólo dos detecciones. ¿Cuál es la probabilidad posterior de que esté presente la impureza?

96. A cada concursante de un programa de preguntas se le pide que especifique una de seis posibles categorías de las cuales se harán las preguntas. Suponga que P(el concursante pide la categoría i) $= \frac{1}{6}$ y concursantes sucesivos escogen sus categorías de manera independiente entre sí. Si hay tres concursantes en cada programa y los tres concursantes de determinado programa seleccionan diferentes categorías, ¿cuál es la probabilidad de que exactamente uno haya seleccionado la categoría 1?

97. Los sujetadores que se utilizan en la fabricación de aviones tienen unas cuantas estrías para que queden seguros y no se aflojen por la vibración. Suponga que 95% de los sujetadores pasa una inspección inicial. Del 5% que falla, 20% están tan defectuosos, que se deben desechar. Los demás sujetadores se mandan a estriar de nuevo, donde 40% no se pueden salvar, y se eliminan. El otro 60% de estos sujetadores se corrige al hacerle nuevas estrías, y después pasa la inspección.

a. ¿Cuál es la probabilidad de que un sujetador entrante seleccionado al azar pase la inspección, ya sea al inicio o después de volverlo a estriar?

b. Dado que un sujetador pasó la inspección, ¿cuál es la probabilidad de que pase la inspección inicial y no requiera un nuevo estriado?

98. Uno por ciento de los individuos de cierta población es portador de determinada enfermedad. Una prueba diagnóstica para esta enfermedad, tiene una tasa de detección de 90% para portadores y una de 5% para no portadores. Suponga que la prueba se aplica de forma independiente a dos muestras sanguíneas diferentes del mismo individuo seleccionado al azar.

a. ¿Cuál es la probabilidad de que ambas pruebas den el mismo resultado?

b. Si ambas pruebas son positivas, ¿cuál es la probabilidad de que el individuo seleccionado sea un portador?

99. Un sistema consiste en dos componentes. La probabilidad de que el segundo componente funcione de manera satisfactoria durante su vida de diseño es 0.9, la probabilidad de que esto suceda por lo menos con uno de los dos componentes es 0.96 y la probabilidad de que los dos componentes funcionen del mismo modo es 0.75. Dado que el primer componente funciona de manera satisfactoria toda su vida de diseño, ¿cuál es la probabilidad de que pase lo mismo con el segundo componente?

100. Cierta empresa envía 40% de sus paquetes de correo nocturnos vía el servicio de correo exprés E_1. De estos paquetes, 2% llega después del tiempo de entrega establecido (denote con L el evento "entrega tardía"). Si del archivo de la compañía se selecciona al azar un registro de envíos nocturnos, ¿cuál es la probabilidad de que el paquete sea enviado vía E_1 y llegue tarde?

101. Refiérase al ejercicio 100. Suponga que 50% de los paquetes nocturnos son enviados vía el servicio de correo exprés E_2 y el 10% restante se envía vía E_3. De los enviados vía E_2 sólo 1% llega tarde, mientras que 5% de los paquetes que maneja E_3 llega tarde.

a. ¿Cuál es la probabilidad de que un paquete seleccionado al azar llegue tarde?

b. Si un paquete seleccionado al azar llega a tiempo, ¿cuál es la probabilidad de que no haya sido enviado vía E_1?

102. Una compañía utiliza tres líneas de ensamble diferentes —A_1, A_2 y A_3— para fabricar un componente. De los componentes fabricados en la línea A_1, 5% se regresan para corregir un defecto, en tanto que para la línea A_2 se devuelven 8% de los componentes y para A_3 10%. Suponga que 50% de los componentes se producen en la línea A_1, 30% en la línea A_2 y 20% en la A_3. Si un componente seleccionado al azar debe ser devuelto a la línea de ensamble, ¿cuál es la probabilidad de que provenga de la línea A_1? ¿De la línea A_2? ¿De la línea A_3?

103. Si se desecha la posibilidad de un nacimiento en un 29 de febrero, supóngase que un individuo escogido al azar tiene las mismas probabilidades de haber nacido en cualquier otro de los 365 días.

a. Si elige al azar a diez personas, ¿cuál es la probabilidad de que tengan cumpleaños diferentes? ¿De que por lo menos dos tengan el mismo cumpleaños?

b. Si k sustituye a diez en el inciso (a), ¿cuál es la k mínima para la cual hay por lo menos una probabilidad 50-50 de que dos o más personas tengan el mismo cumpleaños?

c. Si se elige al azar a diez personas, ¿cuál es la probabilidad de que por lo menos dos tengan el mismo cumpleaños o al menos dos tengan los mismos últimos tres dígitos de sus números del seguro social? [*Nota:* en el artículo "Methods for Studying Coincidences" (F. Mosteller y P. Diaconis, *J. Amer. Stat. Assoc.*, 1989: 853-861) se analizan problemas de este tipo.]

104. Un método empleado para distinguir entre rocas graníticas (G) y basálticas (B) consiste en examinar una porción del espectro infrarrojo de la energía solar reflejada en la superficie de la roca. Sean R_1, R_2 y R_3 las intensidades del espectro en tres longitudes de onda diferentes; por lo común, para el granito $R_1 < R_2 < R_3$, mientras que para el basalto $R_3 < R_1 < R_2$. Cuando las mediciones se hacen a distancia (por medio de aviones) pueden surgir varios ordenamientos de las R_i si la roca es basalto o granito. En vuelos sobre regiones de composición conocida se obtuvo la siguiente información:

	Granito	Basalto
$R_1 < R_2 < R_3$	60%	10%
$R_1 < R_3 < R_2$	25%	20%
$R_3 < R_1 < R_2$	15%	70%

Suponga que para una roca seleccionada al azar en cierta región, $P(\text{granito}) = 0.25$ y $P(\text{basalto}) = 0.75$.

a. Demuestre que $P(\text{granito} \mid R_1 < R_2 < R_3) > P(\text{basalto} \mid R_1 < R_2 < R_3)$. Si el resultado de las medi-

ciones fuera $R_1 < R_2 < R_3$, ¿se clasificaría la roca como granito o basalto?

b. Si las mediciones produjeran $R_1 < R_3 < R_2$, ¿cómo se clasificaría la roca? Conteste la misma pregunta para $R_3 < R_1 < R_2$.

c. Con las reglas de clasificación indicadas en los incisos (a) y (b), cuando se seleccione una roca de esta región, ¿cuál es la probabilidad de una clasificación errónea? [*Sugerencia*: G se podría clasificar como B o B como G, y se conocen $P(B)$ y $P(G)$.]

d. Si $P(\text{granito}) = p$ en lugar de 0.25, ¿hay valores de p(diferentes de 1) para los cuales una roca siempre se clasificaría como granito?

105. A una persona se le permite una secuencia de miradas para detectar un blanco. Sea G_i = {el blanco es detectado en la i-ésima mirada}, con $p_i = P(G_i)$. Suponga que las G_i son eventos independientes y escriba una expresión para hallar la probabilidad de que el blanco sea detectado al final de la n-ésima mirada. (*Nota*: este modelo se analiza en "Predicting Aircraft Detectability", *Human Factors*, 1979: 277-291.)

106. En un juego de béisbol de la liga infantil, el lanzador del equipo A lanza un *strike* 50% del tiempo y una bola el otro 50%, los lanzamientos sucesivos son independientes entre sí, y el lanzador nunca golpea a un bateador. Al saber esto, el manager del equipo B indica al primer bateador no conectar ningún lanzamiento. Calcule la probabilidad de que

a. El bateador reciba base por bola en el cuarto lanzamiento.

b. El bateador reciba base por bola en el sexto lanzamiento (dos de los primeros cinco deben haber sido *strikes*), con un argumento de conteo o por medio de un diagrama de árbol.

c. El bateador reciba base por bola.

d. El primer bateador anote una carrera sin que haya ningún *out* (suponiendo que cada bateador sigue la estrategia de no conectar ningún lanzamiento).

107. Cuatro ingenieros, A, B, C y D, han sido programados para entrevistas de trabajo a las 10:00 A.M. del viernes 13 de enero en la empresa Random Sampling, Inc. El gerente de personal ha programado a los cuatro para los salones de entrevista 1, 2, 3 y 4, respectivamente. Sin embargo, la secretaria del gerente ignora esto, así que los asigna a los cuatro salones de modo completamente aleatorio. ¿Cuál es la probabilidad de que

a. Los cuatro terminen en el salón que les corresponde?

b. Ninguno de los cuatro esté en el salón correcto?

108. Una aerolínea tiene vuelos a las 10:00 A.M. de Chicago a Nueva York, Atlanta y Los Ángeles. Sea A el evento en el que el vuelo de Nueva York está completo y defina los eventos B y C de manera análoga para los otros dos vuelos. Suponga que $P(A) = 0.6$, $P(B) = 0.5$, $P(C) = 0.4$ y que los tres eventos son independientes

a. ¿Cuál es la probabilidad de que los tres vuelos estén completos? ¿De que por lo menos un vuelo no esté completo?

b. ¿Cuál es la probabilidad de que sólo el vuelo de Nueva York esté completo? ¿De que exactamente uno de los tres vuelos esté completo?

109. Un gerente de personal va a entrevistar a cuatro candidatos para un trabajo. Se utiliza 1, 2, 3 y 4 para clasificarlos en orden de preferencia y la entrevista se realiza en orden aleatorio. Sin embargo, al término de cada entrevista, el gerente sabrá cómo se compara el candidato actual con los entrevistados antes. Por ejemplo, el orden de entrevista 3, 4, 1, 2, no genera información después de la primera entrevista, deja ver que el segundo candidato es peor que el primero, y que el tercero es mejor que los dos primeros. Pero el orden 3, 4, 2, 1 generaría la misma información después de cada una de las primeras tres entrevistas. El gerente desea contratar al mejor candidato, pero debe tomar una decisión irrevocable de contratar o no después de cada entrevista. Considere la siguiente estrategia: rechazar de manera automática a los primeros s candidatos, y después contratar al primer candidato posterior que sea el mejor que los ya entrevistados (si no aparece este candidato, se contrata al último entrevistado).

Por ejemplo, con $s = 2$, el orden 3, 4, 1, 2 daría como resultado la mejor contratación, no así el orden 3, 1, 2, 4. De los cuatro posibles valores de s (0, 1, 2 y 3), ¿cuál maximiza P(se contrata al mejor)? (*Sugerencia*: escriba los 24 órdenes de entrevista equiprobables: $s = 0$ significa que se contrata de manera automática al primer candidato.)

110. Considere cuatro eventos independientes A_1, A_2, A_3 y A_4 y sea $p_i = P(A_i)$ para $i = 1, 2, 3, 4$. Exprese la probabilidad de que por lo menos ocurra uno de estos cuatro eventos en términos de las p_i y haga lo mismo para la probabilidad de que por lo menos ocurran dos de los eventos.

111. Una caja contiene las siguientes cuatro papeletas, cada una con exactamente las mismas dimensiones: (1) gana el premio 1; (2) gana el premio 2; (3) gana el premio 3, (4) gana los premios 1, 2 y 3. Se selecciona al azar una papeleta. Sea A_1 = {gana el premio 1}, A_2 = {gana el premio 2} y A_3 = {gana el premio 3}. Demuestre que A_1 y A_2 son independientes, que A_1 y A_3 son independientes y que A_2 y A_3 también son independientes (independencia *por pares*). Sin embargo, demuestre que $P(A_1 \cap A_2 \cap A_3) \neq P(A_1) \cdot P(A_2) \cdot P(A_3)$, de modo que los tres eventos no son mutuamente independientes.

Bibliografía

Durrett, Richard, *The Essentials of Probability,* Duxbury Press, Belmont, CA, 1993. Una presentación concisa con un nivel un poco mayor que el de este texto.

Mosteller, Frederick, Robert Rourke y George Thomas, *Probability with Statistical Applications* (2a. ed.), Addison-Wesley, Reading, MA, 1970. Una buena introducción de las bases de cálculo para la probabilidad, con ejemplos amenos; en particular es bueno en reglas de conteo y sus aplicaciones.

Olkin, Ingram, Cyrus Derman y Leon Gleser, *Probability Models and Application* (2a. ed.), Macmillan, Nueva York, 1994. Una introducción completa a la probabilidad, escrita con un nivel matemático un poco mayor que el de este texto, pero que contiene muchos buenos ejemplos.

Ross, Sheldon, *A First Course in Probability* (6a. ed.), Macmillan, Nueva York, 2002. Escrito en forma concisa y con mayor rigor matemático que este texto, pero con gran cantidad de ejemplos y ejercicios interesantes.

Winkler, Robert, *Introduction to Bayesian Inference and Decision,* Holt, Rinehart & Winston, Nueva York, 1972. Una muy buena introducción a la probabilidad subjetiva.

3

Variables aleatorias discretas y distribuciones de probabilidad

Introducción:

Si un experimento produce resultados cualitativos o cuantitativos, los métodos estadísticos de análisis requieren que se centre la atención en ciertos aspectos numéricos de los datos (p. ej., una proporción muestral x/n, media $\bar{x}$ o desviación estándar s). El concepto de variable aleatoria permite pasar de los resultados experimentales a una función numérica de los resultados. Hay dos tipos de variables aleatorias, variables aleatorias discretas y variables aleatorias continuas. En este capítulo, se examinan las propiedades básicas y se analizan los ejemplos más importantes de variables discretas. En el capítulo 4 se tratan las variables aleatorias continuas.

3.1 Variables aleatorias

En cualquier experimento, hay numerosas características que se pueden observar o medir, pero en la mayor parte de los casos un experimentador se centra en algún aspecto específico o aspectos de una muestra. Por ejemplo, en un estudio de patrones de desplazamiento en un área metropolitana, cada individuo de una muestra podría ser interrogado acerca de la distancia de desplazamiento y el número de personas que se desplazan en el mismo vehículo, pero no acerca del CI, ingreso, tamaño de la familia y otras características. Otra posibilidad es que un investigador podría probar una muestra de componentes y registrar sólo la cantidad de los que fallan en 1000 horas, en vez de registrar cada uno de los tiempos de falla.

En general, cada resultado de un experimento se puede asociar con un número si se especifica una regla de asociación (p. ej., la cantidad entre la muestra de diez componentes que no duran 1000 horas, o el peso total de equipaje de una muestra de 25 pasajeros de una aerolínea). Esta regla de asociación se llama **variable aleatoria**, una variable porque son posibles diferentes valores numéricos, y aleatoria porque el valor observado depende de cuál de los resultados experimentales posibles resulte (figura 3.1).

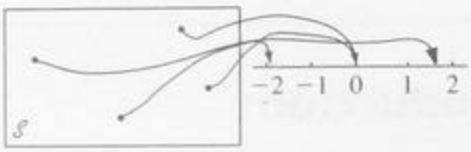


Figura 3.1 Una variable aleatoria.

DEFINICIÓN

Para un determinado espacio muestral S de algún experimento, una **variable aleatoria (va)** es cualquier regla que relaciona un número con cada resultado en S. En el lenguaje matemático, una variable aleatoria es una cuyo dominio es el espacio muestral y cuyo recorrido es el conjunto de los números reales.

Por lo común, las variables aleatorias se denotan con letras mayúsculas, como X y Y, del final de alfabeto. En contraste con el uso previo de una letra minúscula, como x, para denotar una variable, las letras minúsculas se utilizan ahora para representar determinado valor de la variable aleatoria correspondiente. La notación $X(s) = x$ significa que x es el valor relacionado con el resultado s mediante la variable aleatoria X.

Ejemplo 3.1 Cuando un estudiante intenta conectarse a un sistema de computadoras de tiempo compartido, uno de dos puertos están ocupados (F), en cuyo caso el estudiante no logra tener acceso, o bien hay por lo menos un puerto libre (S), en cuyo caso el estudiante logra conectarse al sistema. Con $S = \{S, F\}$, defina una variable aleatoria X mediante

$$X(S) = 1 \qquad X(F) = 0$$

La variable aleatoria X indica si (1) el estudiante se puede conectar o no (0). ■

En el ejemplo 3.1, la variable X se especificó al listar de forma explícita cada elemento de S y el número relacionado. Si S contiene más que algunos resultados, el listado resulta tedioso, pero por lo común se puede evitar.

Ejemplo 3.2 Considere el experimento en el que un número telefónico de un cierto código de área se marca por medio de un marcador de números aleatorios (estos dispositivos tienen un uso extendido en las organizaciones donde se llevan a cabo sondeos de opinión), y defina una variable aleatoria Y mediante

$$Y = \begin{cases} 1 & \text{si no está listado el número seleccionado} \\ 0 & \text{si el número elegido está listado en el directorio} \end{cases}$$

Por ejemplo, si el número 5282966 aparece en el directorio telefónico, entonces $Y(5282966) = 0$, en tanto que $Y(7727350) = 1$ indica que no está listado el número 7727350. Una descripción expresada de esta clase es más económica que un listado completo, así que siempre que sea posible, se usará una descripción como ésta. ■

En los ejemplos 3.1 y 3.2, los únicos valores posibles de la variable aleatoria fueron 0 y 1. Esta clase de variable aleatoria surge con suficiente frecuencia como para darle un nombre especial, en honor al individuo que la estudió primero.

DEFINICIÓN

> Cualquier variable aleatoria cuyos únicos valores posibles son 0 y 1 se llama **variable aleatoria de Bernoulli.**

Con frecuencia se querrá definir y estudiar varias variables aleatorias distintas del mismo espacio muestral.

Ejemplo 3.3 En el ejemplo 2.3 se describió un experimento en el que se determinó el número de bombas en uso en cada una de dos gasolineras. Defina las variables aleatorias X, Y y U mediante

X = número total de bombas en uso en las dos gasolineras

Y = la diferencia entre el número de bombas en uso en la gasolinera 1 y la cantidad de bombas en uso en la gasolinera 2

U = el máximo de los números de bombas en uso en las dos gasolineras

Si se lleva a cabo este experimento y resulta $s = (2, 3)$, entonces, $X((2, 3)) = 2 + 3 = 5$; por lo tanto, se dice que el valor observado de X fue $x = 5$. De manera similar, el valor observado de Y sería $y = 2 - 3 = -1$, y el valor observado de U sería $u = \text{máx}(2, 3) = 3$. ■

Cada una de las variables aleatorias de los ejemplos 3.1 y 3.3 asumen sólo un número finito de valores posibles. Aunque no siempre es así.

Ejemplo 3.4 En el ejemplo 2.4, se consideró el experimento en el que se examinaron las baterías hasta encontrar una buena (S). El espacio muestral fue $\mathcal{S} = \{S, FS, FFS, \ldots\}$. Defina una variable aleatoria X mediante

X = número de baterías examinadas antes que termine el experimento

Entonces, $X(S) = 1$, $X(FS) = 2$, $X(FFS) = 3$, . . . , $X(FFFFFFS) = 7$, y así sucesivamente. Cualquier entero positivo es un valor posible de X, así que el conjunto de valores posibles es infinito. ■

Ejemplo 3.5 Supóngase que de cierta forma aleatoria, se elige un lugar (latitud y longitud) de Estados Unidos continental. Defina una variable aleatoria Y mediante

Y = la altitud por arriba del nivel del mar en el lugar elegido

Por ejemplo, si el lugar elegido estuviera a (39°50′N, 98°35′W), entonces se podría tener Y((39°50′N, 98°35′W)) = 1748.26 pies. El valor posible más grande de Y es 14 494 (montaña Whitney), y el valor posible más pequeño es −282 (Valle de la Muerte). El conjunto de los valores posibles de Y es el conjunto de los números en el intervalo entre −282 y 14 494, es decir,

$$\{y: y \text{ es un número}, -282 \leq y \leq 14\,494\}$$

y hay una cantidad infinita de números en este intervalo. ■

Dos tipos de variables aleatorias

En la sección 1.2, se hizo una distinción entre dos tipos diferentes de variables numéricas, las que son *discretas* y las que son *continuas*. La misma distinción se traslada a las variables aleatorias.

DEFINICIÓN

Una variable aleatoria **discreta** es una variable aleatoria cuyos valores posibles constituyen un conjunto finito, o bien se pueden listar en una secuencia infinita en la que hay un primer elemento, un segundo elemento, etcétera. Una variable aleatoria es **continua** si su subconjunto de valores posibles consiste en un intervalo completo en la recta numérica.

Aunque cualquier intervalo en la recta numérica contiene una cantidad infinita de números, se puede demostrar que no hay manera de crear una lista infinita de todos estos valores, sólo hay demasiados.

Ejemplo 3.6 Las variables aleatorias de los ejemplos 3.1 a 3.4 son discretas. Como otro ejemplo, supóngase que se elige al azar parejas casadas y se hace una prueba de sangre en cada persona hasta que se encuentra una pareja de marido y mujer que tienen el mismo factor Rh. Con X = número de pruebas de sangre realizadas, los valores posibles de X son $D = \{2, 4, 6, 8, \ldots\}$. Puesto que los valores posibles han sido listados en secuencia, X es una variable aleatoria discreta. ■

Para estudiar las propiedades básicas de las variables aleatorias discretas, sólo se requieren las herramientas de las matemáticas discretas, suma y resta. El estudio de las variables aleatorias continuas requiere las matemáticas continuas del cálculo, integrales y derivadas.

Ejercicios | Sección 3.1 (1 a 10)

1. Una viga de concreto podría fallar ya sea por esfuerzo de corte (S) o flexión (F). Supóngase que se eligen al azar tres vigas con falla y se determina el tipo de falla para cada una. Sea X = número de vigas entre las tres seleccionadas que fallaron por esfuerzo de corte. Enumere cada resultado del espacio muestral junto con el valor relacionado de X.

2. Dé tres ejemplos de variable aleatoria (*va*) de Bernoulli (distintos a los del texto).

3. Por medio del experimento del ejemplo 3.3, defina otras dos variables aleatorias y enumere los posibles valores de cada una.

4. Sea X = número de dígitos diferentes de cero de un código postal seleccionado al azar. ¿Cuáles son los valores posibles de X? Dé tres posibles resultados y sus valores X relacionados.

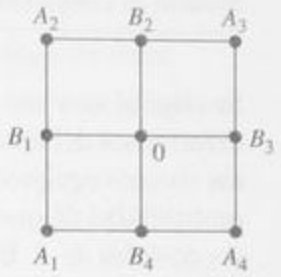
5. Si el espacio muestral $\mathcal{S}$ es un conjunto infinito, ¿esto implica necesariamente que cualquier *va* X definida a partir de $\mathcal{S}$ tendrá un conjunto infinito de valores posibles? Si es así, diga por qué. Si no, dé un ejemplo.

6. Comenzando en determinado momento, cada automóvil que llega a un crucero es observado para ver si da vuelta a la izquierda (L), a la derecha (D) o sigue derecho (A). El experimento termina cuando se observa que un automóvil da vuelta a la izquierda. Sea X = número de automóviles observados. ¿Cuáles son los valores posibles de X? Enumere cinco resultados y sus valores X relacionados.

7. Para cada variable aleatoria definida aquí, describa el conjunto de valores posibles de la variable e indique si la variable es discreta.
   a. X = número de huevos que no están rotos en un cartón de huevo elegido al azar.
   b. Y = número de estudiantes en la lista de un curso que no asistieron el primer día de clases.
   c. U = número de veces que un novato intenta golpear una pelota de golf antes de lograrlo.
   d. X = longitud de una serpiente de cascabel seleccionada al azar.
   e. Z = cantidad de regalías obtenidas de la venta de una primera edición de 10 000 libros de texto.
   f. Y = pH de una muestra de suelo seleccionada al azar
   g. X = tensión (en libras por pulgada cuadrada) del encordado de una raqueta de tenis seleccionada al azar.
   h. X = número total de lanzamientos de una moneda requeridos para que tres individuos obtengan una coincidencia (HHH o TTT).

8. Cada vez que se prueba un componente, el ensayo es un éxito (S) o fracaso (F). Supongamos que el componente se prueba varias veces hasta que ocurre un éxito en tres ensayos seguidos. Sea Y el número de ensayos necesarios para lograr esto. Enumere los resultados que corresponden a los cinco valores posibles mínimos de Y, e indique cuál valor de Y se relaciona con cada uno.

9. Un individuo llamado Claudio se localiza en el punto 0 del siguiente diagrama.



   Por medio de un dispositivo de aleatorización adecuado (p. ej., un dado tetraédrico, uno que tiene cuatro lados), Claudio se mueve primero a uno de los cuatro lugares B_1, B_2, B_3, B_4. Una vez que está en uno de estos lugares, se utiliza otro dispositivo de aleatorización para determinar si Claudio regresa a 0, o bien, visita uno de los otros dos puntos adyacentes. Este proceso continúa; después de cada movimiento, se determina otro movimiento a uno de los (nuevos) puntos adyacentes al lanzar un dado o moneda adecuados.
   a. Sea X = número de movimientos que realiza Claudio antes de regresar primero a 0. ¿Cuáles son los valores posibles de X? ¿X es discreta o continua?
   b. Si se permiten los movimientos a lo largo de trayectorias diagonales que enlacen al 0 con A_1, A_2, A_3, y A_4, respectivamente, conteste las preguntas del inciso (a).

10. Se determinará el número de bombas en uso en una gasolinera de seis bombas y en una de cuatro. Dé los valores posibles para cada una de las siguientes variables aleatorias:
    a. T = número total de bombas en uso.
    b. X = diferencia entre el número de bombas en uso en las gasolineras 1 y 2.
    c. U = número máximo de bombas en uso en cualquier gasolinera.
    d. Z = número de gasolineras que tienen exactamente dos bombas en uso.

3.2 Distribuciones de probabilidad para variables aleatorias discretas

Cuando se asignan probabilidades a varios resultados en $\mathcal{S}$, éstos a su vez determinan las probabilidades relacionadas con los valores de cualquier *va* X particular. La *distribución de probabilidad de* X establece cómo se distribuye la probabilidad total de 1 entre (asignada a) los valores posibles de X.

Ejemplo 3.7 Seis lotes de componentes están listos para que cierto proveedor los envíe. El número de componentes defectuosos de cada lote es el siguiente:

Lote	1	2	3	4	5	6
Número de componentes defectuosos	0	2	0	1	2	0

Se elige al azar uno de estos lotes para enviarlo a un cliente. Sea X el número de componentes defectuosos del lote seleccionado. Los tres valores posibles de X son 0, 1 y 2. De los seis eventos simples equiprobables, tres producen $X = 0$, uno $X = 1$, y los otros dos $X = 2$. Sea $p(0)$ la probabilidad de que $X = 0$ y $p(1)$ y $p(2)$ representan las probabilidades de los otros dos valores posibles de X. Entonces

$$p(0) = P(X = 0) = P(\text{se envía el lote 1, 3 o 6}) = \frac{3}{6} = 0.500$$
$$p(1) = P(X = 1) = P(\text{se envía el lote 4}) = \frac{1}{6} = 0.167$$
$$p(2) = P(X = 2) = P(\text{se envía el lote 2 o 5}) = \frac{2}{6} = 0.333$$

Es decir, se distribuye una probabilidad de 0.500 al valor 0 de X, se coloca una probabilidad de 0.167 en el valor 1 de X, y la probabilidad restante, 0.333, se relaciona con el valor 2 de X. Los valores de X junto con sus probabilidades especifican de forma colectiva la distribución de probabilidad o *función masa de probabilidad de X*. Si se repitiera una y otra vez este experimento, a la larga $X = 0$ ocurriría la mitad del tiempo, $X = 1$ un sexto del tiempo y $X = 2$ un tercio del tiempo. ■

DEFINICIÓN

La **distribución de probabilidad** o **función masa de probabilidad** (fmp) de una variable aleatoria discreta se define para todo número x mediante $p(x) = P(X = x) = P(\text{toda } s \in \mathcal{S}: X(s) = x)$.*

En palabras, para todo valor posible x de la variable aleatoria, la fmp especifica la probabilidad de observar ese valor cuando se lleva a cabo el experimento. Se requieren las condiciones $p(x) \geq 0$ y $\sum_{\text{todas las } x \text{ posibles}} p(x) = 1$ de cualquier fmp.

Ejemplo 3.8 Suponga que va a una librería universitaria durante la primera semana de clases, y observa si la siguiente persona que compra una computadora adquiere una *laptop* o un modelo de escritorio. Sea

$$X = \begin{cases} 1 & \text{si el cliente compra una computadora } \textit{laptop} \\ 0 & \text{si el cliente compra una computadora de escritorio} \end{cases}$$

Si 20% de los compradores durante esa semana eligen una *laptop*, la fmp para X es

$$p(0) = P(X = 0) = P(\text{el siguiente cliente compra un modelo de escritorio}) = 0.8$$
$$p(1) = P(X = 1) = P(\text{el siguiente cliente compra una computadora } \textit{laptop}) = 0.2$$
$$p(x) = P(X = x) = 0 \quad \text{para } x \neq 0 \text{ o } 1$$

*$P(X = x)$ se lee "la probabilidad de que la *va* X tome el valor x". Por ejemplo, $P(X = 2)$ denota la probabilidad de que el valor X resultante sea 2.

Una descripción equivalente es

$$p(x) = \begin{cases} 0.8 & \text{si } x = 0 \\ 0.2 & \text{si } x = 1 \\ 0 & \text{si } x \neq 0 \text{ o } 1 \end{cases}$$

La figura 3.2 es una ilustración de esta fmp, llamada *gráfica de línea.*

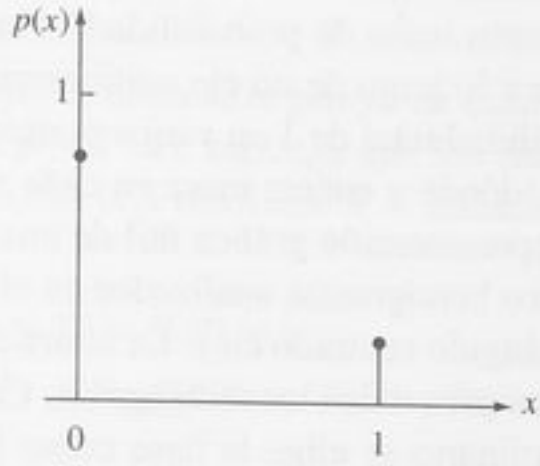


Figura 3.2 Gráfica de línea para la fmp del ejemplo 3.8 ■

Ejemplo 3.9 Considere un grupo de cinco posibles donadores de sangre —A, B, C, D y E— de los cuales, sólo A y B tienen sangre tipo O+. Cinco muestras de sangre, una de cada individuo, se tipifican en orden aleatorio hasta que se identifica un individuo O+. Sea la variable aleatoria $Y =$ número de tipificaciones necesarias para identificar un individuo O+. Entonces la fmp de Y es

$$p(1) = P(Y = 1) = P(\text{se tipifica primero A o B}) = \frac{2}{5} = 0.4$$

$$p(2) = P(Y = 2) = P(\text{primero C, D o E y luego A o B})$$

$$= P(\text{primero C, D o E}) \cdot P(\text{luego A o B} \mid \text{primero C, D o E}) = \frac{3}{5} \cdot \frac{2}{4} = 0.3$$

$$p(3) = P(Y = 3) = P(\text{C, D o E primero y segundo, y luego A o B})$$

$$= \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{2}{3}\right) = 0.2$$

$$p(4) = P(Y = 4) = P(\text{C, D y E todas tipificadas primero}) = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right) = 0.1$$

$$p(y) = 0 \quad \text{si } y \neq 1, 2, 3, 4$$

La fmp se puede presentar bastante bien en forma tabular:

y	1	2	3	4
$p(y)$	0.4	0.3	0.2	0.1

donde cualquier valor de y no listado recibe probabilidad cero. Esta fmp se puede mostrar también en una gráfica de línea (figura 3.3).

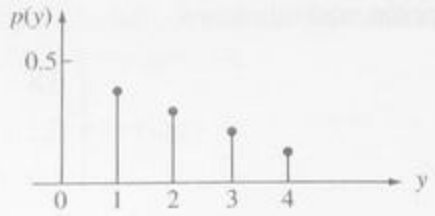


Figura 3.3 Gráfica de línea para la fmp del ejemplo 3.9 ■

Un modelo empleado en física para un sistema de "masas puntuales" hace pensar en el nombre "función masa de probabilidad". En este modelo, las masas están distribuidas en varios lugares x a lo largo de un eje unidimensional. La fmp describe cómo se distribuye la masa de probabilidad total de 1 en varios puntos a lo largo del eje de valores posibles de la variable aleatoria (dónde y cuánta masa en cada x).

Otra representación gráfica útil de una fmp, denominada **histograma de probabilidad**, es similar a los histogramas analizados en el capítulo 1. Arriba de cada y con $p(y) > 0$, construya un rectángulo centrado en y. La altura de cada rectángulo es proporcional a $p(y)$, y la base es la misma para todos los rectángulos. Cuando los valores posibles están igualmente espaciados, de ordinario se elige la base como la distancia entre valores sucesivos de y (aunque podría ser más pequeña). En la figura 3.4 se muestran dos diagramas de probabilidad.

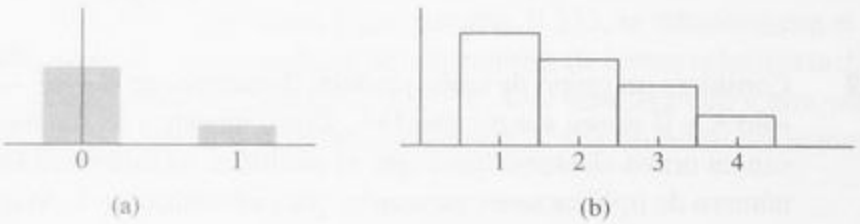


Figura 3.4 Histogramas de probabilidad: (a) Ejemplo 3.8; (b) Ejemplo 3.9

Parámetro de una distribución de probabilidad

En el ejemplo 3.8, se tuvo $p(0) = 0.8$ y $p(1) = 0.2$ porque 20% de los compradores eligieron una computadora *laptop*. En otra librería, se podría dar el caso de que $p(0) = 0.9$ y $p(1) = 0.1$. De manera más general, la fmp de cualquier variable aleatoria de Bernoulli se puede expresar en la forma $p(1) = \alpha$ y $p(0) = 1 - \alpha$, donde $0 < \alpha < 1$. Debido a que la fmp depende del valor particular de α, por lo común se escribe $p(x; \alpha)$ en vez de sólo $p(x)$:

$$p(x; \alpha) = \begin{cases} 1 - \alpha & \text{si } x = 0 \\ \alpha & \text{si } x = 1 \\ 0 & \text{en caso contrario} \end{cases} \tag{3.1}$$

Entonces cada elección de α en la expresión (3.1) produce una fmp diferente.

DEFINICIÓN

Suponga que $p(x)$ depende de una cantidad a la que se le puede asignar cualquiera de varios valores posibles, donde cada valor diferente determina una distribución de probabilidad distinta. A esta cantidad se le llama **parámetro** de la distribución. La colección de las distribuciones de probabilidad para diferentes valores del parámetro se llama **familia** de distribuciones de probabilidad.

La cantidad α de la expresión (3.1) es un parámetro. Cada número diferente α entre 0 y 1 determina un miembro diferente de una familia de distribuciones; dos de estos miembros son

$$p(x; 0.6) = \begin{cases} 0.4 & \text{si } x = 0 \\ 0.6 & \text{si } x = 1 \\ 0 & \text{en caso contrario} \end{cases} \quad \text{y} \quad p(x; 0.5) = \begin{cases} 0.5 & \text{si } x = 0 \\ 0.5 & \text{si } x = 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Toda distribución de probabilidad para una variable aleatoria de Bernoulli tiene la forma de la expresión (3.1), así que se denomina *familia de distribuciones de Bernoulli.*

Ejemplo 3.10 Comenzando en un tiempo fijo, se observa el género de cada recién nacido en un hospital hasta que nace un niño (B). Sea $p = P(B)$, suponga que los nacimientos sucesivos son independientes y defina la variable aleatoria X mediante X = número de nacimientos observados. Entonces,

$$p(1) = P(X = 1) = P(B) = p$$
$$p(2) = P(X = 2) = P(GB) = P(G) \cdot P(B) = (1 - p)p$$

y

$$p(3) = P(X = 3) = P(GGB) = P(G) \cdot P(G) \cdot P(B) = (1 - p)^2 p$$

Al continuar de esta manera, se obtiene una fórmula general:

$$p(x) = \begin{cases} (1 - p)^{x-1} p & x = 1, 2, 3, \ldots \\ 0 & \text{en caso contrario} \end{cases} \tag{3.2}$$

La cantidad p en la expresión (3.2) representa un número entre 0 y 1 y es un parámetro de la distribución de probabilidad. En el ejemplo del género, $p = 0.51$ podría ser apropiado, pero si se estuviera buscando a los primeros niños con sangre de Rh positivo, entonces se podría establecer que $p = 0.85$. ■

Función de distribución acumulada

Para algún valor fijo x, por lo común se desea calcular la probabilidad de que el valor observado de X sea a lo sumo x. Por ejemplo, la fmp del ejemplo 3.7 fue

$$p(x) = \begin{cases} 0.500 & x = 0 \\ 0.167 & x = 1 \\ 0.333 & x = 2 \\ 0 & \text{en caso contrario} \end{cases}$$

Entonces, la probabilidad de que X sea a lo sumo 1 es

$$P(X \le 1) = p(0) + p(1) = 0.500 + 0.167 = 0.667$$

En este ejemplo, $X \le 1.5$ si y sólo si $X \le 1$, entonces $P(X \le 1.5) = P(X \le 1) = 0.667$. De manera similar, $P(X \le 0) = P(X = 0) = 0.5$, y $P(X \le 0.75) = 0.5$ también. Puesto que 0 es el valor posible más pequeño de X, $P(X \le -1.7) = 0$, $P(X \le -0.0001) = 0$, etcétera. El valor posible más grande de X es 2, así que $P(X \le 2) = 1$, y si x es cualquier número más grande que 2, $P(X \le x) = 1$; es decir, $P(X \le 5) = 1$, $P(X \le 10.23) = 1$, y así sucesivamente. Observe que $P(X < 1) = 0.5 \ne P(X \le 1)$, ya que la probabilidad del valor 1 de X está inclui-

do en la última probabilidad, pero no en la anterior. Cuando X es una variable aleatoria discreta y x es un valor posible de X, $P(X < x) < P(X \le x)$.

DEFINICIÓN

La **función de distribución acumulada** (fda) $F(x)$ de una variable aleatoria discreta X con fmp $p(x)$ se define para todo número x mediante

$$F(x) = P(X \le x) = \sum_{y:\, y \le x} p(y)$$

Para cualquier número x, $F(x)$ es la probabilidad de que el valor observado de X sea a lo sumo x.

Ejemplo 3.11 La fmp de Y (el número de tipificaciones de sangre) en el ejemplo 3.9 fue

y	1	2	3	4
$p(y)$	0.4	0.3	0.2	0.1

Primero se determina $F(y)$ para cada valor del conjunto $\{1, 2, 3, 4\}$ de valores posibles:

$$F(1) = P(Y \le 1) = P(Y = 1) = p(1) = 0.4$$
$$F(2) = P(Y \le 2) = P(Y = 1 \text{ o } 2) = p(1) + p(2) = 0.7$$
$$F(3) = P(Y \le 3) = P(Y = 1 \text{ o } 2 \text{ o } 3) = p(1) + p(2) + p(3) = 0.9$$
$$F(4) = P(Y \le 4) = P(Y = 1 \text{ o } 2 \text{ o } 3 \text{ o } 4) = 1$$

Ahora, para cualquier otro número y, $F(y)$ es igual al valor de F en el valor posible más cercano de Y a la izquierda de y. Por ejemplo, $F(2.7) = P(Y \le 2.7) = P(Y \le 2) = 0.7$, y $F(3.999) = F(3) = 0.9$. Por lo tanto, la fda es

$$F(y) = \begin{cases} 0 & \text{si } y < 1 \\ 0.4 & \text{si } 1 \le y < 2 \\ 0.7 & \text{si } 2 \le y < 3 \\ 0.9 & \text{si } 3 \le y < 4 \\ 1 & \text{si } 4 \le y \end{cases}$$

En la figura 3.5 se muestra una gráfica de $F(y)$.

Figura 3.5 Gráfica de la fda del ejemplo 3.11 ■

Para una variable aleatoria discreta X, la gráfica de $F(x)$ tiene un salto en todo valor posible de X y es plana entre los valores posibles. A esta gráfica se le denomina **función escalón.**

Ejemplo 3.12 En el ejemplo 3.10, cualquier entero positivo fue un valor posible de X, y la fmp fue

$$p(x) = \begin{cases} (1-p)^{x-1}p & x = 1, 2, 3, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

Para cualquier entero positivo x,

$$F(x) = \sum_{y \le x} p(y) = \sum_{y=1}^{x} (1-p)^{y-1}\, p = p \sum_{y=0}^{x-1} (1-p)^{y} \tag{3.4}$$

Para evaluar esta suma, se usa el hecho de que la suma parcial de una serie geométrica es

$$\sum_{y=0}^{k} a^{y} = \frac{1 - a^{k+1}}{1 - a}$$

Al usar esta ecuación en la ecuación (3.4), con $a = 1 - p$ y $k = x - 1$, se obtiene

$$F(x) = p \cdot \frac{1 - (1-p)^{x}}{1 - (1-p)} = 1 - (1-p)^{x} \quad x \text{ entero positivo}$$

Puesto que F es constante entre enteros positivos,

$$F(x) = \begin{cases} 0 & x < 1 \\ 1 - (1-p)^{[x]} & x \ge 1 \end{cases} \tag{3.5}$$

donde $[x]$ es el entero más grande $\le x$ (e.g., $[2.7] = 2$). Así, si $p = 0.51$ como en el ejemplo de los nacimientos, entonces la probabilidad de tener que examinar a lo sumo cinco nacimientos para ver el primer niño es $F(5) = 1 - (0.49)^5 = 1 - 0.0282 = 0.9718$, en tanto que $F(10) \approx 1.0000$. La gráfica de esta fda se muestra en la figura 3.6.

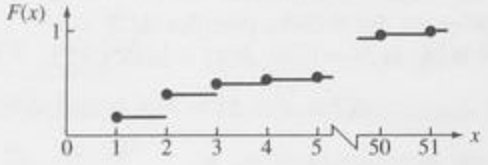


Figura 3.6 Gráfica de $F(x)$ para el ejemplo 3.12 ■

En los ejemplos dados hasta ahora, la fda se ha derivado de la fmp. Este proceso se puede invertir para obtener la fmp a partir de la fda siempre que se tenga esta última función. Suponga, por ejemplo, que X representa el número de componentes defectuosos en un envío que consiste en seis componentes, así que los valores posibles de X son $0, 1, \ldots, 6$. Entonces,

$$\begin{aligned} p(3) &= P(X = 3) \\ &= [p(0) + p(1) + p(2) + p(3)] - [p(0) + p(1) + p(2)] \\ &= P(X \le 3) - P(X \le 2) \\ &= F(3) - F(2) \end{aligned}$$

De manera más general, la probabilidad de que X caiga en un intervalo especificado se obtiene sin dificultad de la fda. Por ejemplo,

$$\begin{aligned} P(2 \le X \le 4) &= p(2) + p(3) + p(4) \\ &= [p(0) + \cdots + p(4)] - [p(0) + p(1)] \\ &= P(X \le 4) - P(X \le 1) \\ &= F(4) - F(1) \end{aligned}$$

Observe que $P(2 \le X \le 4) \ne F(4) - F(2)$. Esto es porque el valor 2 de X está incluido en $2 \le X \le 4$, así que no se desea restar su probabilidad. Sin embargo, $P(2 < X \le 4) = F(4) - F(2)$ porque $X = 2$ no está incluida en el intervalo $2 < X \le 4$.

PROPOSICIÓN

Para dos números cualesquiera a y b con $a \le b$,

$$P(a \le X \le b) = F(b) - F(a-)$$

donde "$a-$" representa el valor posible más grande de X que es estrictamente menor que a. En particular, si los únicos valores posibles son enteros y si a y b son enteros, entonces

$$\begin{aligned} P(a \le X \le b) &= P(X = a \text{ o } a + 1 \text{ o } \ldots \text{ o } b) \\ &= F(b) - F(a - 1) \end{aligned}$$

se toma $a = b$ se obtiene $P(X = a) = F(a) - F(a - 1)$ en este caso.

La razón de restar $F(a-)$ y no $F(a)$ es que se desea incluir $P(X = a)$, con $F(b) - F(a)$ se obtiene $P(a < X \le b)$. Esta proposición se usa de manera extensa al calcular probabilidades binomiales y de Poisson en las secciones 3.4 y 3.6.

Ejemplo 3.13 Sea X = número de días de incapacidad que toma cierto empleado elegido al azar de una gran compañía durante determinado año. Si el número máximo de días de incapacidad permisibles por año es 14, los valores posibles de X son 0, 1, . . . , 14. Con $F(0) = 0.58$, $F(1) = 0.72$, $F(2) = 0.76$, $F(3) = 0.81$, $F(4) = 0.88$ y $F(5) = 0.94$,

$$P(2 \le X \le 5) = P(X = 2, 3, 4, \text{ o } 5) = F(5) - F(1) = 0.22$$

y

$$P(X = 3) = F(3) - F(2) = 0.05$$ ■

Otro punto de vista para las funciones masa de probabilidad

En muchos casos, es útil considerar que una fmp especifica un modelo matemático para una población discreta.

Ejemplo 3.14 Considere elegir al azar un estudiante de entre los 15 000 para el semestre actual en Mega University. Sea X = número de cursos en los que se registra el alumno elegido, y suponga que X tiene la fmp siguiente:

x	1	2	3	4	5	6	7
$p(x)$	0.01	0.03	0.13	0.25	0.39	0.17	0.02

Una forma de ver esta situación, es considerar que la población consiste en 15 000 individuos, cada uno con su propio valor de X; $p(x)$ da la proporción con cada valor de X. Otro punto de vista es hacer caso omiso de los estudiantes y considerar que la población consiste en los valores de X: hay algunos unos en la población, algunos números dos y, por último, algunos sietes. Así, la población consiste en los números 1, 2, . . . , 7 (por lo tanto es discreta), y $p(x)$ proporciona un modelo para la distribución de valores de la población. ■

Una vez que se tiene un modelo poblacional, se emplea para calcular características de la población (p. ej., la media μ) y se hacen deducciones acerca de estas características.

Ejercicios | Sección 3.2 (11 a 27)

11. Un taller automotriz que se especializa en ajustes de motor sabe que 45% de los ajustes se efectúa en automóviles de cuatro cilindros, 40% en automóviles de seis cilindros y 15% en automóviles de ocho cilindros. Sea X = número de cilindros del siguiente automóvil al que se le hace un ajuste de motor.

a. ¿Cuál es la fmp de X?

b. Trace una gráfica de línea y un histograma de probabilidad para la fmp del inciso (a).

c. ¿Cuál es la probabilidad de que el siguiente automóvil ajustado tenga por lo menos cuatro cilindros? ¿Más de seis cilindros?

12. A veces las aerolíneas registran más pasajeros del cupo normal de los vuelos. Supóngase que para un avión con 50 asientos, 55 pasajeros tienen boleto. Defina la variable aleatoria Y como el número de pasajeros con boleto que en realidad se presenta para el vuelo. En la tabla siguiente se ilustra la función masa de probabilidad de Y.

y	45	46	47	48	49	50	51	52	53	54	55
$p(y)$	0.05	0.10	0.12	0.14	0.25	0.17	0.06	0.05	0.03	0.02	0.01

a. ¿Cuál es la probabilidad de que se pueda acomodar a todos los pasajeros con boleto que se presentan para realizar el vuelo?

b. ¿Cuál es la probabilidad de que no todos los pasajeros con boleto que se presentan para realizar el vuelo puedan ser acomodados?

c. Si el lector es la primera persona en lista de espera (lo que significa que será el primero en abordar el avión si hay asientos disponibles después que han sido acomodados todos los pasajeros con boleto), ¿cuál es la probabilidad de que pueda realizar el vuelo? ¿Cuál es la probabilidad si es la tercera persona en la lista de espera?

13. Un negocio de computadoras que atiende pedidos por correo, tiene seis líneas telefónicas. Sea X el número de líneas en uso en un momento específico. Supongamos que la fmp de X es como la que se proporciona en la tabla siguiente.

x	0	1	2	3	4	5	6
$p(x)$	0.10	0.15	0.20	0.25	0.20	0.06	0.04

Calcule la probabilidad de cada uno de los siguientes eventos.

a. {A lo sumo tres líneas están en uso}.

b. {Menos de tres líneas están en uso}.

c. {Por lo menos tres líneas están en uso}.

d. {Entre dos y cinco líneas, inclusive, están en uso}.

e. {Entre dos y cuatro líneas, inclusive, no están en uso}.

f. {Por lo menos cuatro líneas no están en uso}.

14. Un departamento de planeación de un municipio requiere a un contratista para que remita una, dos, tres, cuatro o cinco formas (dependiendo de la naturaleza del proyecto) a fin de solicitar un permiso de construcción. Sea Y = número de formas requeridas del siguiente solicitante. Se sabe que la probabilidad de que se requieran y formas es proporcional a y, es decir, $p(y) = ky$ para $y = 1, \ldots, 5$.

a. ¿Cuál es el valor de k? [*Sugerencia:* $\sum_{y=1}^{5} p(y) = 1$.]

b. ¿Cuál es la probabilidad de que a lo sumo se requieran tres formas?

c. ¿Cuál es la probabilidad de que se requieran entre dos y cuatro formas (inclusive)?

d. ¿Podría ser $p(y) = y^2/50$ para $y = 1, \ldots, 5$ la fmp de Y?

15. Muchos fabricantes tienen programas de control de calidad que incluyen la revisión de defectos en materiales recibidos. Supóngase que un fabricante de computadoras recibe tarjetas en lotes de cinco. De cada lote se seleccionan dos tarjetas para inspeccionarlas. Se pueden representar los posibles resultados del proceso de selección mediante pares. Por ejemplo, el par (1, 2) representa la selección de las tarjetas 1 y 2 para inspección.

a. Liste los diez resultados posibles diferentes.

b. Supóngase que las tarjetas 1 y 2 son las únicas defectuosas de un lote de cinco y se escogerán al azar dos tarjetas. Sea X el número de tarjetas defectuosas observado entre las inspeccionadas. Determine la distribución de probabilidad de X.

c. Sea $F(x)$ la fda de X. Determine primero $F(0) = P(X \le 0)$, $F(1)$ y $F(2)$, y después obtenga $F(x)$ para las otras x.

16. Algunas regiones de California son particularmente propensas a sismos. Suponga que en un área, 30% de los propietarios de casa está asegurado contra daños por sismos. Se elige al azar a cuatro propietarios de casa; sea X el número de propietarios de entre los cuatro que tienen seguro contra sismos.

a. Determine la distribución de probabilidad de X. [*Sugerencia:* Sea S un propietario de casa asegurado y F uno sin seguro. Entonces un resultado posible es $SFSS$, con probabilidad $(0.3)(0.7)(0.3)(0.3)$ y valor de X relacionado 3. Hay otros 15 resultados.]

b. Trace el histograma de probabilidad correspondiente.

c. ¿Cuál es el valor más probable para X?

d. ¿Cuál es la probabilidad de que por lo menos dos de los cuatro seleccionados tenga seguro contra sismos?

17. El voltaje de una batería nueva puede ser aceptable (A) o inaceptable (U). Cierta linterna necesita dos baterías, así que se eligen de manera independiente y se prueban hasta encontrar dos aceptables. Supóngase que 90% de las baterías tienen voltaje aceptable. Sea Y el número de baterías que deben ser probadas.

a. ¿Cuál es el valor de $p(2)$, es decir, $P(Y = 2)$?

b. ¿Cuál es el valor de $p(3)$? (*Sugerencia:* hay dos resultados diferentes que producen $Y = 3$.)

c. Para tener $Y = 5$, ¿qué debe ser cierto de la quinta batería seleccionada? Liste los cuatro resultados para los cuales $Y = 5$ y luego determine $p(5)$.

d. Utilice el modelo en sus respuestas de los incisos (a) a (c), a fin de obtener una fórmula general para $p(y)$.

18. Se lanzan dos dados no cargados de manera independiente. Sea M = el máximo de los dos lanzamientos (así que $M(1, 5) = 5$, $M(3, 3) = 3$, etcétera).

a. ¿Cuál es la fmp de M? [*Sugerencia:* primero determine $p(1)$, después $p(2)$, etcétera.]

b. Determine la fda de M y trace su gráfica.

19. En el ejemplo 3.9 suponga que sólo hay cuatro posibles donadores de sangre, de los cuales sólo uno tiene sangre tipo O+. Calcule la fmp de Y.

20. Una biblioteca se suscribe a dos semanarios, cada uno de los cuales se supone llega en el correo del miércoles. En realidad, cada uno podría llegar el miércoles, jueves, viernes o sábado. Suponga que los dos llegan de manera independiente y para cada uno $P(\text{miércoles}) = 0.3$, $P(\text{jueves}) = 0.4$, $P(\text{viernes}) = 0.2$ y $P(\text{sábado}) = 0.1$. Sea Y = número de días que pasan después del miércoles hasta que llegan los semanarios (así que los valores posibles de Y son 0, 1, 2 o 3). Calcule la fmp de Y. [*Sugerencia:* hay 16 resultados posibles; $Y(W, W) = 0$, $Y(F, Th) = 2$, etcétera.]

21. Refiérase al ejercicio 13 y obtenga y trace la gráfica de la fda $F(x)$. Después, utilícela para calcular la probabilidad de los eventos proporcionados en los incisos (a) a (d) de ese problema.

22. Una organización de consumidores que evalúa automóviles nuevos, suele informar el número de defectos importantes de cada automóvil examinado. Sea X el número de defectos importantes en un automóvil de cierto tipo seleccionado al azar. La fda de X es como sigue:

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.06 & 0 \le x < 1 \\ 0.19 & 1 \le x < 2 \\ 0.39 & 2 \le x < 3 \\ 0.67 & 3 \le x < 4 \\ 0.92 & 4 \le x < 5 \\ 0.97 & 5 \le x < 6 \\ 1 & 6 < x \end{cases}$$

Calcule las siguientes probabilidades directamente de la fda:

a. $p(2)$, es decir, $P(X = 2)$

b. $P(X > 3)$

c. $P(2 \le X \le 5)$

d. $P(2 < X < 5)$

23. Una compañía de seguros ofrece a sus asegurados varias opciones para el pago de primas. Para un asegura-

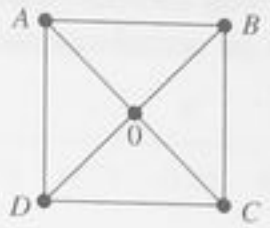
do seleccionado al azar, sea X = número de meses entre pagos sucesivos. La fda de X es como sigue:

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.30 & 1 \le x < 3 \\ 0.40 & 3 \le x < 4 \\ 0.45 & 4 \le x < 6 \\ 0.60 & 6 \le x < 12 \\ 1 & 12 \le x \end{cases}$$

a. ¿Cuál es la fmp de X?
b. Sólo con la fda, calcule $P(3 \le X \le 6)$ y $P(4 \le X)$.

24. En el ejemplo 3.10, sea Y = número de niñas que nacen antes que termine el experimento. Con $p = P(B)$ y $1 - p = P(G)$, ¿cuál es la fmp de Y? (*Sugerencia*: primero enumere los valores posibles de Y, comenzando con el mínimo, y continúe hasta tener una fórmula general.)

25. Alvie Singer vive en el punto 0 del siguiente diagrama y tiene cuatro amigos que viven en A, B, C y D. Un día, Alvie decide ir de visita, lanza una moneda al aire dos veces para decidir a cuál de los cuatro visitar. Una vez en la casa de un amigo, regresa a casa o se dirige a una de las dos casas adyacentes (p. ej., 0, A o C cuando está en B), y cada una de las tres posibilidades tiene una probabilidad de $\frac{1}{3}$. De este modo, Alvie continúa visitando amigos hasta que regresa a casa.



a. Sea X = número de veces que Alvie visita a un amigo. Deduzca la fmp de X.
b. Sea Y = número de segmentos de recta que Alvie cruza (incluyendo los que llevan a 0 y salen de allí), ¿cuál es la fmp de Y?
c. Suponga que las amigas de Alvie viven en A y C y los amigos en B y D. Si Z = número de visitas a las amigas, ¿cuál es la fmp de Z?

26. Después que los estudiantes abandonan el salón de clases, un maestro de estadística observa que debajo de los pupitres hay cuatro libros de texto. Al comenzar la siguiente clase, el maestro distribuye al azar los cuatro libros a cada uno de los cuatro estudiantes (1, 2, 3 y 4) que dicen haber olvidado su libro. Un resultado posible es que 1 reciba el libro de 2, 2 el libro de 4, 3 su propio libro y 4 el libro de 1. Este resultado se puede abreviar como (2, 4, 3, 1).
a. Enumere los otros 23 resultados posibles.
b. Represente con X el número de estudiantes que reciben su propio libro. Determine la fmp de X.

27. Demuestre que la fda $F(x)$ es una función no decreciente; es decir; si, $x_1 < x_2$ significa que $F(x_1) \le F(x_2)$. ¿En qué condición se tiene que $F(x_1) = F(x_2)$?

3.3 Valores esperados de variables aleatorias discretas

En el ejemplo 3.14 se consideró una universidad que tiene 15 000 estudiantes y X = número de cursos a los que se inscribió un estudiante elegido al azar. Se deduce la fmp de X. Puesto que $p(1) = 0.01$, se sabe que $(0.01) \cdot (15\,000) = 150$ de los estudiantes se inscribieron a un curso, y de manera similar para los otros valores de x.

x	1	2	3	4	5	6	7
$p(x)$	0.01	0.03	0.13	0.25	0.39	0.17	0.02
Número de inscritos	150	450	1950	3750	5850	2550	300

(3.6)

Para calcular el número promedio de cursos por estudiante, o el valor promedio de X en la población, se debe calcular el número de cursos y dividir entre el número total de estudiantes. Puesto que cada uno de los 150 estudiantes está tomando un curso, éstos 150 contribuyen

con 150 cursos al total. De manera similar, 450 estudiantes contribuyen con 2(450) cursos, y así sucesivamente. Así, el valor promedio poblacional es

$$\frac{1(150) + 2(450) + 3(1950) + \cdots + 7(300)}{15\,000} = 4.57 \tag{3.7}$$

Puesto que $150/15\,000 = 0.01 = p(1)$, $450/15\,000 = 0.03 = p(2)$, etcétera, una expresión alterna parar (3.7) es

$$1 \cdot p(1) + 2 \cdot p(2) + \cdots + 7 \cdot p(7) \tag{3.8}$$

En la expresión (3.8) se muestra que para calcular el valor promedio poblacional de X, sólo se requieren los valores posibles de X junto con sus probabilidades (proporciones). En particular, el tamaño de la población es irrelevante siempre que la ecuación (3.6) dé la fmp. El promedio o valor medio de X es entonces un promedio *ponderado* de los valores posibles $1, \ldots, 7$, donde las ponderaciones son las probabilidades de esos valores.

Valor esperado de X

DEFINICIÓN

Sea X una variable aleatoria discreta con un conjunto D de valores posibles y una fmp $p(x)$. El **valor esperado** o **valor medio** de X, denotado por $E(X)$ o μ_X, es

$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)$$

Cuando está claro a qué X se refiere el valor esperado, suele utilizarse μ y no μ_X.

Ejemplo 3.15 Para la fmp en (3.6),

$$\begin{aligned}\mu &= 1 \cdot p(1) + 2 \cdot p(2) + \cdots + 7 \cdot p(7)\\ &= (1)(0.01) + 2(0.03) + \cdots + (7)(0.02)\\ &= 0.01 + 0.06 + 0.39 + 1.00 + 1.95 + 1.02 + 0.14 = 4.57\end{aligned}$$

Si se considera que la población consiste en los valores $1, 2, \ldots, 7$ de X, entonces $\mu = 4.57$ es la media poblacional. Como resultado, se hará referencia a μ como la *media poblacional* y no como la media de X en la población. ■

En el ejemplo 3.15, el valor esperado μ fue 4.57, que no es un valor posible de X. La palabra *esperado* se debe interpretar con precaución, porque no se esperaría ver un valor 4.57 para X cuando se elige un solo estudiante.

Ejemplo 3.16 Justo después del nacimiento, cada recién nacido se evalúa según una escala llamada escala de Apgar. Las posibles evaluaciones son $0, 1, \ldots, 10$, donde el color, tomo muscular, esfuerzo respiratorio, latidos del corazón, irritabilidad de reflejos, determinan la clasificación del niño

(la mejor puntuación posible es 10). Sea X la puntuación de Apgar de un niño elegido al azar que nace en cierto hospital durante el siguiente año, y suponga que la fmp de X es

x	0	1	2	3	4	5	6	7	8	9	10
$p(x)$	0.002	0.001	0.002	0.005	0.02	0.04	0.18	0.37	0.25	0.12	0.01

Entonces el valor medio de X es

$$\begin{aligned} E(X) = \mu &= 0(0.002) + 1(0.001) + 2(0.002) \\ &\quad + \cdots + 8(0.25) + 9(0.12) + 10(0.01) \\ &= 7.15 \end{aligned}$$

De nuevo, μ no es un valor posible de la variable X. Asimismo, debido a que la variable se refiere a un futuro niño, no hay población existente concreta a la que se refiera μ. En vez de eso, se considera a la fmp como modelo para una población conceptual que consiste en los valores 0, 1, 2, . . . , 10. Así, el valor medio de esta población conceptual es $\mu = 7.15$. ■

Ejemplo 3.17 Sea $X = 1$ si un componente elegido al azar requiere servicio de garantía y $= 0$ en caso contrario. Entonces X es una variable aleatoria de Bernoulli con fmp

$$p(x) = \begin{cases} 1 - p & x = 0 \\ p & x = 1 \\ 0 & x \neq 0, 1 \end{cases}$$

en la que $E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0(1 - p) + 1(p) = p$. Es decir, el valor esperado de X es justo la probabilidad de que X tome el valor 1. Si se conceptualiza una población que consiste en ceros, en proporción $1 - p$ y en unos en proporción p, entonces el promedio de la población es $\mu = p$. ■

Ejemplo 3.18 La forma general para la fmp de $X =$ número de niños nacidos hasta e incluso el primer niño es

$$p(x) = \begin{cases} p(1 - p)^{x-1} & x = 1, 2, 3, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

De la definición,

$$E(X) = \sum_D x \cdot p(x) = \sum_{x=1}^{\infty} xp(1 - p)^{x-1} = p\sum_{x=1}^{\infty}\left[-\frac{d}{dp}(1 - p)^x\right] \qquad (3.9)$$

Si se intercambia el orden de tomar la derivada y la suma, la suma es la de una serie geométrica. Después que se calcula la suma, se toma la derivada y el resultado final es $E(X) = 1/p$. Si p es cercana a 1, se espere ver muy pronto a un niño, en tanto que si p es cercana a 0, se esperan muchos nacimientos antes del primer niño. Para $p = 0.5$, $E(X) = 2$. ■

Hay otra interpretación de μ que se usa con mucha frecuencia. Considere la fmp

$$p(x) = \begin{cases} (0.5) \cdot (0.5)^{x-1} & \text{si } x = 1, 2, 3, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

Ésta es la fmp de X = número de lanzamientos de una moneda necesarios para obtener la primera H (un caso especial del ejemplo 3.18). Supóngase que se observa un valor de x de esta fmp (lanzar una moneda hasta que aparezca una H), luego se observa de forma independiente otro valor (continuar lanzando la moneda), luego otro, etcétera. Si después de observar una cantidad muy grande de valores de x, se promedian, el promedio muestral resultante será muy cercano a $\mu = 2$. Es decir, μ se puede interpretar como el valor promedio de X observado a la larga cuando el experimento se lleva a cabo muchas veces.

Ejemplo 3.19 Sea X el número de entrevistas que realiza un estudiante antes de conseguir un trabajo, con fmp

$$p(x) = \begin{cases} k/x^2 & x = 1, 2, 3, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

donde k se elige de modo que $\sum_{x=1}^{\infty} (k/x^2) = 1$. (En un curso de matemáticas finitas, se demuestra que $\sum_{x=1}^{\infty} (1/x^2) < \infty$, lo que significa que esta k existe; sin embargo, no interesa su valor exacto). El valor esperado de X es

$$\mu = E(X) = \sum_{x=1}^{\infty} x \cdot \frac{k}{x^2} = k \sum_{x=1}^{\infty} \frac{1}{x} \tag{3.10}$$

La suma del lado derecho de la ecuación (3.10) es la famosa serie armónica de las matemáticas, y se demuestra que es igual a ∞. $E(X)$ no es finita aquí porque $p(x)$ no disminuye con rapidez suficiente cuando se incrementa x; los estadísticos dicen que la distribución de probabilidad de X tiene "cola larga". Si se elige una secuencia de valores de X por medio de esta distribución, el promedio de la muestra no se establece en algún número finito, sino que tiende a crecer sin límite.

Los estadísticos emplean la frase "cola larga" junto con cualquier distribución que tiene una gran cantidad de probabilidad lejos de μ (así que las colas largas no requieren que μ sea igual a infinito). Las colas largas dificultan sacar conclusiones acerca de μ. ■

Valor esperado de una función

Por lo regular, se tiene interés en el valor esperado de alguna función $h(X)$ y no en X.

Ejemplo 3.20 Suponga que una librería compra diez copias de un libro a seis dólares cada uno para venderlo a 12 dólares, con el entendido de que al final de un periodo de tres meses, cualquier copia no vendida se devuelve por dos dólares. Si X es el número de copias vendidas, entonces el ingreso neto es igual a $h(X) = 12X + 2(10 - X) - 60 = 10X - 40$. ■

El ejemplo siguiente indica una forma fácil de calcular el valor esperado de $h(X)$.

Ejemplo 3.21 Sea X = número de cilindros del motor del siguiente automóvil por ajustar en un determinado taller. El costo de un ajuste se relaciona con X mediante $h(X) = 20 + 3X + 0.5X^2$. Puesto que X es una variable aleatoria, también lo es $h(X)$; denote a esta última variable aleatoria con Y. Las fmp de X y Y son como sigue:

x	4	6	8	y	40	56	76
$p(x)$	0.5	0.3	0.2	$p(y)$	0.5	0.3	0.2

D^* se utiliza para denotar los valores posibles de Y,

$$E(Y) = E[h(X)] = \sum_{D^*} y \cdot p(y) \tag{3.11}$$
$$= (40)(0.5) + (56)(0.3) + (76)(0.2)$$
$$= h(4) \cdot (0.5) + h(6) \cdot (0.3) + h(8) \cdot (0.2)$$
$$= \sum_{D} h(x) \cdot p(x)$$

De acuerdo con la ecuación (3.11), era innecesario determinar la fmp de Y para obtener $E(Y)$; en cambio, el valor esperado deseado es un promedio ponderado de los valores posibles de $h(x)$ (en vez de x). ■

PROPOSICIÓN

Si la variable aleatoria X tiene un conjunto de valores posibles D y una fmp $p(x)$, entonces el valor esperado de cualquier función $h(X)$, que se denota mediante $E[h(X)]$ o $\mu_{h(X)}$, se calcula con

$$E[h(X)] = \sum_{D} h(x) \cdot p(x)$$

Según esta proposición, $E[h(X)]$ se calcula de la misma forma que $E(X)$ excepto que en lugar de x se sustituye $h(x)$.

Ejemplo 3.22 Una tienda de computadoras compró tres máquinas de cierto tipo a 500 dólares cada una. Las venderá a 1000 dólares la pieza. El fabricante está de acuerdo en volver a comprar las computadoras que no se vendan después de un periodo especificado, a 200 dólares la pieza. Sea X el número de computadoras vendidas, y suponga que $p(0) = 0.1$, $p(1) = 0.2$, $p(2) = 0.3$ y $p(3) = 0.4$. Si $h(X)$ representa la ganancia obtenida con la venta de X unidades, la información proporcionada indica que $h(X) = \text{ingreso} - \text{costo} = 1000X + 200(3 - X) - 1500 = 800X - 900$. Entonces la ganancia esperada es

$$E[h(X)] = h(0) \cdot p(0) + h(1) \cdot p(1) + h(2) \cdot p(2) + h(3) \cdot p(3)$$
$$= (-900)(0.1) + (-100)(0.2) + (700)(0.3) + (1500)(0.4)$$
$$= \$700$$

■

Reglas del valor esperado

La función de interés $h(X)$ es con bastante frecuencia una función lineal $aX + b$. En este caso, $E[h(X)]$ se calcula de manera fácil a partir de $E(X)$.

PROPOSICIÓN

$$E(aX + b) = a \cdot E(X) + b$$

(O bien, con otra notación, $\mu_{aX+b} = a \cdot \mu_X + b$.)

En palabras, el valor esperado de una función lineal es igual a la función lineal evaluada en el valor esperado $E(X)$. Puesto que $h(X)$ en el ejemplo (3.22) es lineal y $E(X) = 2$, $E[h(X)] = 800(2) - 900 = 700$ dólares como antes.

Proposición:

$$E(aX + b) = \sum_D (ax + b) \cdot p(x) = a\sum_D x \cdot p(x) + b\sum_D p(x)$$
$$= aE(X) + b$$

■

Dos casos especiales de la proposición producen reglas importantes del valor esperado.

1. Para cualquier constante a, $E(aX) = a \cdot E(X)$ (tome $b = 0$).
2. Para cualquier constante b, $E(X + b) = E(X) + b$ (tome $a = 1$). (3.12)

La multiplicación de X por una constante a cambia la unidad de medida (de dólares a centavos, donde $a = 100$, pulgadas a cm, donde $a = 2.54$, etcétera). La regla 1 establece que el valor esperado en las nuevas unidades es igual al valor esperado en las antiguas unidades multiplicado por el factor de conversión a. De manera similar, si se suma una constante b a cada valor posible de X, entonces el valor esperado cambia por esa misma cantidad constante.

Varianza de X

El valor esperado de X describe dónde se centra la distribución de probabilidad. Usando esta analogía física de colocar la masa puntual $p(x)$ en el valor x en un eje unidimensional, si el eje estuviera apoyado en un fulcro colocado en μ, el eje no tendería a inclinarse. Esto se ilustra para dos distribuciones diferentes en la figura 3.7.

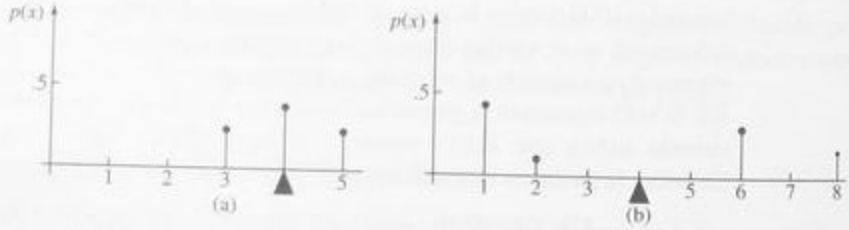


Figura 3.7 Dos distribuciones de probabilidad diferentes con $\mu = 4$

Aunque ambas distribuciones ilustradas en la figura 3.7 tienen la misma μ central, la distribución de la figura 3.7(b) tiene mayor dispersión o variabilidad que la de la figura 3.7(a). Se usa la varianza de X para evaluar la cantidad de variabilidad en (la distribución de) X, del mismo modo que s^2 se utilizó en el capítulo 1 para medir la variabilidad en una muestra.

DEFINICIÓN

Sea X con una fmp $p(x)$ y valor esperado μ. Entonces la **varianza** de X, denotada por $V(X)$ o σ_X^2, justo nada más σ^2, es

$$V(X) = \sum_D (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$$

La **desviación estándar** (DE) de X es

$$\sigma_X = \sqrt{\sigma_X^2}$$

La cantidad $h(X) = (X - \mu)^2$ es la desviación cuadrada de X desde su media, y σ^2 es la desviación cuadrada esperada. Si la mayor parte de la distribución de probabilidad está cerca de μ, entonces σ^2 es relativamente pequeña. Sin embargo, si hay valores de x lejos de μ que tienen una $p(x)$ grande, entonces σ^2 será bastante grande.

Ejemplo 3.23 Si X es el número de cilindros en el siguiente automóvil que será ajustado en un taller mecánico, con la fmp del ejemplo 3.21 [$p(4) = 0.5$, $p(6) = 0.3$, $p(8) = 0.2$, de donde $\mu = 5.4$], entonces

$$V(X) = \sigma^2 = \sum_{x=4}^{8} (x - 5.4)^2 \cdot p(x)$$

$$= (4 - 5.4)^2(0.5) + (6 - 5.4)^2(0.3) + (8 - 5.4)^2(0.2) = 2.44$$

La desviación estándar de X es $\sigma = \sqrt{2.44} = 1.562$. ■

Cuando la fmp $p(x)$ especifica un modelo matemático para la distribución de valores de la población, tanto σ^2 como σ miden la dispersión de valores en la población; σ^2 es la varianza poblacional y σ es la desviación estándar de la población.

Fórmula abreviada para obtener σ^2

La cantidad de operaciones aritméticas necesarias para calcular σ^2 se puede reducir si se usa una fórmula de cálculo opcional.

PROPOSICIÓN

$$V(X) = \sigma^2 = \left[\sum_D x^2 \cdot p(x)\right] - \mu^2 = E(X^2) - [E(X)]^2$$

Al usar esta fórmula, $E(X^2)$ se calcula primero sin ninguna resta; luego, se calcula $E(X)$, se eleva al cuadrado y se resta (una vez) de $E(X^2)$.

Ejemplo 3.24 La fmp del número de cilindros X en el siguiente automóvil por afinar en un taller mecánico se dio en el ejemplo 3.23 como $p(4) = 0.5$, $p(6) = 0.3$ y $p(8) = 0.2$, de donde $\mu = 5.4$ y

$$E(X^2) = (4^2)(0.5) + (6^2)(0.3) + (8^2)(0.2) = 31.6$$

Por consiguiente, $\sigma^2 = 31.6 - (5.4)^2 = 2.44$ como en el ejemplo 3.23. ■

Demostración de la fórmula abreviada

Desarrolle $(x - \mu)^2$ en la definición de σ^2 para obtener $x^2 - 2\mu x + \mu^2$, y luego aplique $\sum$ a cada uno de los tres términos:

$$\sigma^2 = \sum_D x^2 \cdot p(x) - 2\mu \cdot \sum_D x \cdot p(x) + \mu^2 \sum_D p(x)$$

$$= E(X^2) - 2\mu \cdot \mu + \mu^2 = E(X^2) - \mu^2$$ ■

Reglas de varianza

La varianza de $h(X)$ es el valor esperado de la diferencia cuadrada entre $h(X)$ y su valor esperado:

$$V[h(X)] = \sigma^2_{h(X)} = \sum_D \{h(x) - E[h(X)]\}^2 \cdot p(x) \qquad (3.13)$$

Cuando $h(X)$ es una función lineal, $V[h(X)]$ se relaciona fácilmente con $V(X)$.

PROPOSICIÓN

$$V(aX + b) = \sigma^2_{aX+b} = a^2 \cdot \sigma^2_X \quad \text{y} \quad \sigma_{aX+b} = |a| \cdot \sigma_X$$

Este resultado establece que la suma de la constante b no afecta la varianza, lo cual es intuitivo, porque la suma de b cambia la ubicación (valor medio) pero no la dispersión de valores. En particular,

$$\begin{aligned} &1.\ \sigma^2_{aX} = a^2 \cdot \sigma^2_X,\ \sigma_{aX} = |a| \cdot \sigma_X \\ &2.\ \sigma^2_{X+b} = \sigma^2_X \end{aligned} \tag{3.14}$$

La razón del valor absoluto en σ_{aX} es que a podría ser negativa, en tanto que una desviación estándar no puede ser negativa; a^2 resulta cuando a sale del término que se eleva al cuadrado en la ecuación (3.13).

Ejemplo 3.25 En el problema de ventas de computadoras del ejemplo 3.22, $E(X) = 2$ y

$$E(X^2) = (0)^2(0.1) + (1)^2(0.2) + (2)^2(0.3) + (3)^2(0.4) = 5$$

así que $V(X) = 5 - (2)^2 = 1$. Por lo tanto, la función de ganancia $h(X) = 800X - 900$ tiene una varianza $(800)^2 \cdot V(X) = (640\,000)(1) = 640\,000$ y desviación estándar 800. ■

Ejercicios | Sección 3.3 (28 a 43)

28. La fmp para X = número de defectos importantes en un aparato de cierta clase seleccionado al azar, es

x	0	1	2	3	4
$p(x)$	0.08	0.15	0.45	0.27	0.05

Calcule lo siguiente:

a. $E(X)$
b. $V(X)$ directamente de la definición
c. La desviación estándar de X
d. $V(X)$ con la fórmula abreviada

29. Se selecciona al azar a un individuo que tiene un seguro automotriz de cierta compañía. Sea Y la cantidad de infracciones de circulación por las que el individuo ha sido citado durante los últimos tres años. La fmp de Y es

y	0	1	2	3
$p(y)$	0.60	0.25	0.10	0.05

a. Calcule $E(Y)$.
b. Suponga que un individuo con Y multas obtiene un recargo de $\$100Y^2$. Calcule la cantidad esperada del recargo.

30. Refiérase al ejercicio 12 y calcule $V(Y)$ y σ_Y. A continuación, determine la probabilidad de que Y se encuentre dentro de 1 desviación estándar de su valor medio.

31. Un distribuidor de aparatos electrodomésticos vende tres modelos diferentes de congeladores verticales con capacidad de 13.5, 15.9 y 19.1 pies cúbicos de espacio de almacenaje, respectivamente. Sea X = cantidad de espacio de almacenaje que compra el siguiente cliente. Suponga que X tiene la siguiente fmp

x	13.5	15.9	19.1
$p(x)$	0.2	0.5	0.3

a. Calcule $E(X)$, $E(X^2)$ y $V(X)$.

b. Si el precio de un congelador con capacidad de X pies cúbicos es $25X - 8.5$, ¿cuál es el precio esperado que paga el siguiente cliente por un congelador?

c. ¿Cuál es la varianza del precio $25X - 8.5$ que paga el siguiente cliente?

d. Suponga que si bien la capacidad nominal de un congelador es X, la capacidad real es $h(X) = X - 0.01X^2$. ¿Cuál es la capacidad real esperada del congelador que compra el siguiente cliente?

32. Sea X una variable aleatoria de Bernoulli con la fmp del ejemplo 3.17.

a. Calcule $E(X^2)$.

b. Demuestre que $V(X) = p(1 - p)$.

c. Calcule $E(X^{79})$.

33. Suponga que el número de plantas de un tipo particular localizadas en una región rectangular (que los ecologistas llaman cuadrante) de cierta área geográfica es una variable aleatoria X con fmp

$$p(x) = \begin{cases} c/x^3 & x = 1, 2, 3, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

¿$E(X)$ es finita? Justifique su respuesta (ésta es otra distribución a la que los estadísticos llamarían de cola larga).

34. Una pequeña farmacia solicita ejemplares de una revista cada semana para su anaquel de revistas. Sea X = demanda de la revista, con fmp

x	1	2	3	4	5	6
$p(x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$

Suponga que el propietario de la farmacia paga en realidad un dólar por cada ejemplar de la revista y el precio al cliente es de dos dólares. Si las revistas que quedan al final de la semana no tienen valor de recuperación, ¿es mejor pedir tres o cuatro ejemplares de la revista? (*Sugerencia*: para tres o cuatro ejemplares solicitados, exprese el ingreso neto en función de la demanda X y después calcule el ingreso esperado).

35. Sea X el daño incurrido (en dólares) en cierto tipo de accidente durante determinado año. Los posibles valores de X son 0, 1000, 5000 y 10 000, con probabilidades 0.8, 0.1, 0.08 y 0.02, respectivamente. Una compañía ofrece una póliza deducible de 500 dólares. Si la compañía desea que su ganancia esperada sea de 100 dólares, ¿qué prima debe cobrar?

36. Los n candidatos para un trabajo han sido clasificados como 1, 2, 3, . . . , n. Sea X = la clasificación de un candidato seleccionado al azar, de modo que X tiene una fmp

$$p(x) = \begin{cases} 1/n & x = 1, 2, 3, \ldots, n \\ 0 & \text{en caso contrario} \end{cases}$$

(ésta recibe el nombre de *distribución uniforme discreta*). Calcule $E(X)$ y $V(X)$ con la fórmula abreviada. [*Sugerencia*: la suma de los primeros n enteros positivos es $n(n + 1)/2$, mientras que la suma de sus cuadrados es $n(n + 1)(2n + 1)/6$.]

37. Sea X = el resultado cuando se lanza una vez un dado no cargado. Si antes de lanzar el dado se ofrecen al tirador $(1/3.5)$ dólares o $h(X) = 1/X$ dólares, ¿aceptaría la cantidad garantizada o jugaría? [*Nota*: por lo general no se cumple que $1/E(X) = E(1/X)$.]

38. Una compañía proveedora de productos químicos tiene en existencia 100 libras de un producto que vende a los clientes en lotes de 5 libras. Sea X = número de lotes que pide un cliente seleccionado al azar y suponga que X tiene la siguiente fmp

x	1	2	3	4
$p(x)$	0.2	0.4	0.3	0.1

Calcule $E(X)$ y $V(X)$. Luego, calcule el número esperado de libras sobrantes después que se envía el pedido del siguiente cliente, y la varianza del número de libras restantes. (*Sugerencia*: el número de libras que quedan es una función lineal de X.)

39. **a.** Trace una gráfica lineal de la fmp de X del ejercicio 34. Luego, determine la fmp de $-X$ y dibuje su gráfica de línea. De estas dos gráficas, ¿qué se puede decir en relación con $V(X)$ y $V(-X)$?

b. Utilice la proposición en la que interviene $V(aX + b)$ para establecer una relación general entre $V(X)$ y $V(-X)$.

40. Utilice la definición de la expresión (3.13) para demostrar que $V(aX + b) = a^2 \cdot \sigma_X^2$. [*Sugerencia*: con $h(X) = aX + b$, $E[h(X)] = a\mu + b$ donde $\mu = E(X)$.]

41. Suponga que $E(X) = 5$ y $E[X(X - 1)] = 27.5$. ¿Cuál es

a. $E(X^2)$? [*Sugerencia*: $E[X(X - 1)] = E[X^2 - X] = E(X^2) - E(X)$]?

b. $V(X)$?

c. La relación general entre las cantidades $E(X)$, $E[X(X - 1)]$ y $V(X)$?

42. Escriba una regla general para $E(X - c)$ donde c es una constante. ¿Qué pasa cuando se hace $c = \mu$, el valor esperado de X?

43. Un resultado llamado **desigualdad de Chebyshev** establece que para cualquier distribución de probabilidad de una variable aleatoria X y cualquier número k que sea por lo menos 1, $P(|X - \mu| \geq k\sigma) \leq 1/k^2$. En otras palabras, la probabilidad de que el valor de X se ubique a por lo menos k desviaciones estándar de su media es a lo sumo $1/k^2$.

a. ¿Cuál es el valor del límite superior para $k = 2$? $k = 3$? $k = 4$? $k = 5$? $k = 10$?

b. Calcule μ y σ para la distribución del ejercicio 13. Luego evalúe $P(|X - \mu| \geq k\sigma)$ para los valores de k dados en el inciso (a). ¿Qué indica esto acerca del límite superior en relación con la probabilidad correspondiente?

c. Si X tiene tres valores posibles, -1, 0 y 1, con probabilidades $\frac{1}{18}$, $\frac{8}{9}$ y $\frac{1}{18}$, respectivamente, ¿cuál es $P(|X - \mu| \geq 3\sigma)$, y cómo se compara con el límite correspondiente?

d. Dé una distribución para la cual $P(|X - \mu| \geq 5\sigma) = 0.04$.

3.4 Distribución de probabilidad binomial

Hay varios experimentos que cumplen de manera exacta o aproximada con la siguiente lista de requerimientos:

1. El experimento consiste en una secuencia de n experimentos más pequeños llamados *ensayos*, donde n se fija antes del experimento.
2. Cada ensayo produce uno de los mismos dos resultados posibles (ensayos dicótomos), que se denotan mediante éxito (S) o fracaso (F).
3. Los ensayos son independientes, así que el resultado de cualquier ensayo particular no influye en el resultado de ningún otro.
4. La probabilidad de éxito es constante de un ensayo a otro; esta probabilidad se denota con p.

DEFINICIÓN

Un experimento para el cual se cumplen las condiciones 1 a 4, se llama **experimento binomial.**

Ejemplo 3.26 La misma moneda se lanza de manera sucesiva e independiente n veces. Se emplea S en forma arbitraria para denotar el resultado H(caras) y F para denotar el resultado T (cruces). Entonces este experimento satisface las condiciones 1 a 4. Lanzar una tachuela n veces, con S = punta hacia arriba y F = punta hacia abajo, también produce un experimento binomial. ■

Muchos experimentos tienen que ver con una secuencia de ensayos independientes, para los cuales hay más de dos resultados posibles en cualquier ensayo. Entonces se puede crear un experimento binomial al dividir los resultados posibles en dos grupos.

Ejemplo 3.27 El color de las semillas de chícharo se determina mediante un solo lugar genético. Si los dos alelos de este lugar son AA o Aa (el genotipo), entonces el chícharo será amarillo (el fenotipo), y si el alelo es aa, el chícharo será verde. Suponga que se forman parejas con 20 semillas Aa y se cruzan las dos semillas de cada uno de los diez pares para obtener diez nuevos genotipos. Llame a cada nuevo genotipo un éxito S si es aa o un fracaso en caso contrario. Entonces, con esta identificación de S y F, el experimento es binomial con $n = 10$ y $p = P$(genotipo aa). Si cada miembro del par tiene las mismas probabilidades de contribuir con a o A, entonces $p = P(\text{a}) \cdot P(\text{a}) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$. ■

Ejemplo 3.28 Suponga que cierta ciudad tiene cincuenta restaurantes autorizados, de los cuales 15 tienen en la actualidad por lo menos una violación grave del código sanitario y los otros 35 no tienen violaciones graves. Hay cinco inspectores, y cada uno inspeccionará un restaurante durante la

próxima semana. El nombre de cada restaurante se escribe en una tira de papel distinta y luego se mezclan las tiras, cada inspector saca a su vez una de las tiras *sin reemplazo*. Marque el i-ésimo intento como un éxito si el i-ésimo restaurante elegido ($i = 1, \ldots, 5$) no tiene violaciones graves. Entonces,

$$P(S \text{ en el primer intento}) = \frac{35}{50} = 0.70$$

y

$$\begin{aligned} P(S \text{ en el segundo intento}) &= P(SS) + P(FS) \\ &= P(S \text{ en el segundo intento} \mid S \text{ en el primer intento}) \\ &\quad + P(S \text{ en el segundo intento} \mid F \text{en el primer intento}) \\ &\quad P(F \text{ en el primer intento}) \\ &= \frac{34}{49} \cdot \frac{35}{50} + \frac{35}{49} \cdot \frac{15}{50} = \frac{35}{50}\left(\frac{34}{49} + \frac{15}{49}\right) = \frac{35}{50} = 0.70 \end{aligned}$$

De manera similar, se puede demostrar que $P(S \text{ en el } i\text{-ésimo intento}) = 0.70$ para $i = 3, 4, 5$. Sin embargo,

$$P(S \text{ en el quinto intento} \mid SSSS) = \frac{31}{46} = 0.67$$

mientras que

$$P(S \text{ en el quinto intento} \mid FFFF) = \frac{35}{46} = 0.76$$

El experimento no es binomial, porque los ensayos no son independientes. En general, si el muestreo es sin reemplazo, el experimento no produce ensayos independientes. Si se hubiera sustituido cada tira después de haberla extraído, entonces los ensayos habrían sido independientes, pero esto podría haber producido que más de un inspector inspeccionara el mismo restaurante. ■

Ejemplo 3.29 Suponga que cierto estado tiene 500 000 conductores autorizados, de los cuales 400 000 están asegurados. Se elige al azar una muestra de 10 conductores sin reemplazo. El i-ésimo ensayo se marca con S si está asegurado el i-ésimo conductor. Aunque esta situación parecería idéntica a la del ejemplo 3.28, la diferencia importante es que el tamaño de la población en la que se realiza el muestreo es muy grande en comparación con el tamaño de la muestra. En este caso

$$P(S \text{ en el segundo intento} \mid S \text{ en el primer intento}) = \frac{399\,999}{499\,999} = 0.80000$$

y

$$P(S \text{ en el décimo intento} \mid S \text{ en el noveno intento}) = \frac{399\,991}{499\,991} = 0.799996 \approx 0.80000$$

Estos cálculos indican que aunque los intentos no son del todo independientes, las probabilidades condicionales difieren tan poco entre sí, que para fines prácticos se puede considerar que los ensayos son independientes con constante $P(S) = 0.8$. Así, para una muy buena aproximación, el experimento es binomial con $n = 10$ y $p = 0.8$. ■

A fin de decidir si un experimento "sin reemplazo" se puede tratar como binomial, se usa la siguiente regla empírica:

REGLA

> Considere el muestreo sin reemplazo de una población dicótoma de tamaño N. Si el tamaño de la muestra (número de ensayos) n es al sumo 5% del tamaño de la población, el experimento se puede analizar como si fuera un experimento binomial.

Con la palabra "analizar", se indica que las probabilidades basadas en las hipótesis del experimento binomial serán bastante cercanas a las probabilidades reales "sin reemplazo", que por lo común son más difíciles de calcular. En el ejemplo 3.28, $n/N = 5/50 = 0.1 > 0.05$, así que el experimento binomial no es una buena aproximación, pero en el ejemplo 3.29, $n/N = 10/500\,000 < 0.05$.

Variable y distribución aleatoria binomial

En la mayor parte de los experimentos, lo que interesa es el número total de éxitos, en vez de saber con exactitud qué ensayos produjeron S.

DEFINICIÓN

> Dado un experimento binomial que consiste en n ensayos, la **variable aleatoria binomial** X relacionada con este experimento se define como
>
> $$X = \text{número de éxitos } (S) \text{ entre los } n \text{ ensayos.}$$

Supóngase, por ejemplo, que $n = 3$. Entonces hay ocho resultados posibles para el experimento:

$$SSS\ SSF\ SFS\ SFF\ FSS\ FSF\ FFS\ FFF$$

De la definición de X, $X(SSF) = 2$, $X(SFF) = 1$, etcétera. Los valores posibles para X en un experimento de n-ensayos son $x = 0, 1, 2, \ldots, n$. En muchos casos se escribirá $X \sim \text{Bin}(n, p)$ para indicar que X es una variable aleatoria binomial basada en n ensayos con probabilidad de éxito p.

NOTACIÓN

> Debido a que la fmp de una variable aleatoria binomial X depende de los dos parámetros n y p, la fmp se denota mediante $b(x; n, p)$.

Considere primero el caso $n = 4$ para el cual cada resultado, su probabilidad y valor de x correspondiente, se listan en la tabla 3.1. Por ejemplo,

$$\begin{aligned} P(SSFS) &= P(S) \cdot P(S) \cdot P(F) \cdot P(S) \qquad \text{(ensayos independientes)} \\ &= p \cdot p \cdot (1 - p) \cdot p \qquad (P(S) \text{ constante}) \\ &= p^3 \cdot (1 - p) \end{aligned}$$

Tabla 3.1 Resultados y probabilidades para un experimento binomial con cuatro ensayos

Resultado	x	Probabilidad	Resultado	x	Probabilidad
SSSS	4	p^4	*FSSS*	3	$p^3(1-p)$
SSSF	3	$p^3(1-p)$	*FSSF*	2	$p^2(1-p)^2$
SSFS	3	$p^3(1-p)$	*FSFS*	2	$p^2(1-p)^2$
SSFF	2	$p^2(1-p)^2$	*FSFF*	1	$p(1-p)^3$
SFSS	3	$p^3(1-p)$	*FFSS*	2	$p^2(1-p)^2$
SFSF	2	$p^2(1-p)^2$	*FFSF*	1	$p(1-p)^3$
SFFS	2	$p^2(1-p)^2$	*FFFS*	1	$p(1-p)^3$
SFFF	1	$p(1-p)^3$	*FFFF*	0	$(1-p)^4$

En este caso especial, se desea $b(x; 4, p)$ para $x = 0, 1, 2, 3$ y 4. Para $b(3; 4, p)$, se identifica cuál de los 16 resultados produce un valor de x de 3 y se suman las probabilidades relacionadas con cada resultado:

$$b(3; 4, p) = P(FSSS) + P(SFSS) + P(SSFS) + P(SSSF) = 4p^3(1-p)$$

Hay cuatro resultados con $x = 3$ y cada uno tiene probabilidad $p^3(1-p)$ (carece de importancia el orden de las letras S y F, sólo importa el número de letras S), así

$$b(3; 4, p) = \begin{Bmatrix}\text{número de resultados}\\ \text{con } X = 3\end{Bmatrix} \cdot \begin{Bmatrix}\text{probabilidad de cualquier}\\ \text{resultado con } X = 3\end{Bmatrix}$$

De manera similar, $b(2; 4, p) = 6p^2(1-p)^2$, que también es el producto del número de resultados con $X = 2$ y la probabilidad de cualquiera de esos resultados.

En general,

$$b(x; n, p) = \begin{Bmatrix}\text{número de secuencia de longitud}\\ n \text{ que consiste en } x \text{ letras } S\end{Bmatrix} \cdot \begin{Bmatrix}\text{probabilidad de cualquiera}\\ \text{de esas secuencias}\end{Bmatrix}$$

Puesto que no importa el orden de las letras S y F, el segundo factor en la ecuación anterior es $p^x(1-p)^{n-x}$ (p. ej., los primeros x ensayos producen S y los últimos $n - x$ producen F). El primer factor es el número de formas de elegir x de los n ensayos que serán éxitos (S); es decir, el número de combinaciones de tamaño x que se pueden construir a partir de n objetos distintos (en este caso ensayos).

TEOREMA

$$b(x; n, p) = \begin{cases}\binom{n}{x} p^x(1-p)^{n-x} & x = 0, 1, 2, \ldots n\\ 0 & \text{en caso contrario}\end{cases}$$

Ejemplo 3.30 A cada una de seis personas elegidas al azar que toman refresco de cola, se les da un vaso que contiene cola S y uno que contiene cola F. Los vasos tienen la misma apariencia excepto por un código en el fondo para identificar el refresco de cola. Supóngase que en realidad no hay tendencia entre los bebedores de cola a preferir entre una bebida u otra. Entonces, $p = P$(un individuo elegido prefiere S) $= 0.5$, de modo que con $X =$ el número entre los seis que prefieren S, $X \sim \text{Bin}(6, 0.5)$.

Por lo tanto,

$$P(X = 3) = b(3; 6, 0.5) = \binom{6}{3}(0.5)^3(0.5)^3 = 20(0.5)^6 = 0.313$$

La probabilidad de que por lo menos tres prefieran S es

$$P(3 \le X) = \sum_{x=3}^{6} b(x; 6, 0.5) = \sum_{x=3}^{6} \binom{6}{x}(0.5)^x(0.5)^{6-x} = 0.656$$

y la probabilidad de que a lo sumo uno prefiera S es

$$P(X \le 1) = \sum_{x=0}^{1} b(x; 6, 0.5) = 0.109$$

■

Cómo usar las tablas binomiales

Incluso para un valor de n relativamente pequeño, puede ser tedioso el cálculo de las probabilidades binomiales. En la tabla A.1 del apéndice se tabula la fda $F(x) = P(X \le x)$ para $n = 5, 10, 15, 20, 25$ en combinación con valores seleccionados de p. Algunas otras probabilidades se pueden calcular por medio de la proposición acerca de las fda de la sección 3.2.

NOTACIÓN

Para $X \sim \text{Bin}(n, p)$, la fda se denota mediante

$$P(X \le x) = B(x; n, p) = \sum_{y=0}^{x} b(y; n, p) \qquad x = 0, 1, \ldots, n$$

Ejemplo 3.31 Supóngase que 20% de las copias de determinado libro de texto no pasan cierta prueba de resistencia de pegado. Sea X el número entre las 15 copias elegidas al azar que no pasan la prueba. Entonces X tiene una distribución binomial con $n = 15$ y $p = 0.2$.

1. La probabilidad de que a lo sumo 8 no pasen la prueba es

$$P(X \le 8) = \sum_{y=0}^{8} b(y; 15, 0.2) = B(8; 15, 0.2)$$

que es el elemento del renglón $x = 8$ y la columna $p = 0.2$ de la tabla binomial $n = 15$. De la tabla A.1 del apéndice, la probabilidad es $B(8; 15, 0.2) = 0.999$.

2. La probabilidad de que exactamente 8 no pasen la prueba es

$$P(X = 8) = P(X \le 8) - P(X \le 7) = B(8; 15, 0.2) - B(7; 15, 0.2)$$

que es la diferencia entre dos elementos consecutivos en la columna $p = 0.2$. El resultado es $0.999 - 0.996 = 0.003$.

3. La probabilidad de que por lo menos 8 no pasen la prueba es

$$\begin{aligned} P(X \ge 8) &= 1 - P(X \le 7) = 1 - B(7; 15, 0.2) \\ &= 1 - \begin{pmatrix}\text{elemento del renglón } x = 7 \\ \text{de la columna } p = 0.2\end{pmatrix} \\ &= 1 - 0.996 = 0.004 \end{aligned}$$

4. Por último, la probabilidad de que entre 4 y 7, inclusive, no pasen la prueba es

$$P(4 \le X \le 7) = P(X = 4, 5, 6, \text{o } 7) = P(X \le 7) - P(X \le 3)$$
$$= B(7; 15, 0.2) - B(3; 15, 0.2) = 0.996 - 0.648 = 0.348$$

Observe que esta última probabilidad es la diferencia entre elementos de los renglones $x = 7$ y $x = 3$ *no* los renglones $x = 7$ y $x = 4$. ■

Ejemplo 3.32 Un fabricante de aparatos electrónicos afirma que a lo sumo 10% de sus unidades de suministro de energía necesitan servicio durante el periodo de garantía. Para investigar la afirmación, los técnicos del laboratorio de pruebas compran 20 unidades y someten cada una a comprobación acelerada para simular el uso durante el periodo de garantía. Sea p la probabilidad de una fuente de energía que requiera ser reparada durante el periodo (la proporción de las unidades que necesitan reparación). Los técnicos del laboratorio deben decidir si los datos obtenidos del experimento sustentan la afirmación de que $p \le .10$. Sea X el número entre las 20 fuentes de energía de la muestra que necesitan reparación, así que $X \sim \text{Bin}(20, p)$. Considere la regla de decisión.

Rechace la afirmación de que $p \le 0.10$ en favor de la conclusión de que $p > 0.10$ si $x \ge 5$ (donde x es el valor observado de X), y considere la afirmación plausible si $x \le 4$.

La probabilidad de que la afirmación sea rechazada cuando $p = 0.10$ (una conclusión incorrecta) es

$$P(X \ge 5 \text{ cuando } p = 0.10) = 1 - B(4; 20, 0.1) = 1 - 0.957 = 0.043$$

La probabilidad de que la afirmación no sea rechazada cuando $p = 0.20$ (un tipo diferente de conclusión incorrecta) es

$$P(X \le 4 \text{ cuando } p = 0.2) = B(4; 20, 0.2) = 0.630$$

La primera probabilidad es bastante pequeña, pero la segunda es demasiado grande. Cuando $p = 0.20$, de modo que el fabricante subestimó con mucho el porcentaje de unidades que necesitan servicio, y se emplea la regla de decisión expresada, ¡63% de las muestras dan como resultado que la afirmación del fabricante se considere plausible!

Se podría pensar que la probabilidad de este segundo tipo de conclusión errónea se pudo hacer más pequeña al cambiar el valor de corte 5 en la regla de decisión por algún otro. Sin embargo, aunque sustituir 5 por un número más pequeño produciría una probabilidad más pequeña que 0.630, se incrementaría entonces la otra probabilidad. La única forma de hacer pequeñas ambas "probabilidades de error" es basar la regla de decisión en un experimento con muchas más unidades. ■

Observe que un elemento de 0 en la tabla significa únicamente que una probabilidad es 0 hasta tres dígitos significativos, porque en realidad los elementos de la tabla son positivos. Los paquetes de computadora para cálculos estadísticos, como MINITAB generan ya sea $b(x; n, p)$ o $B(x; n, p)$ para valores especificados de n y p. En el capítulo 4, se presenta un método para obtener aproximaciones rápidas y precisas de las probabilidades binomiales cuando n es grande.

Media y varianza de *X*

Para $n = 1$, la distribución binomial se convierte en la distribución de Bernoulli. Del ejemplo 3.17, el valor medio de una variable de Bernoulli es $\mu = p$, así que el valor esperado de las S en cualquier ensayo simple es p. Puesto que un experimento binomial consiste en n ensayos,

la intuición hace pensar que para $X \sim \text{Bin}(n, p)$, $E(X) = np$, el producto del número de ensayos y la probabilidad de éxito en un solo ensayo. La expresión para $V(X)$ no es tan intuitiva.

PROPOSICIÓN

Si $X \sim \text{Bin}(n, p)$, entonces $E(X) = np$, $V(X) = np(1 - p) = npq$, y $\sigma_X = \sqrt{npq}$ (donde $q = 1 - p$).

Así, calcular la media y la varianza de una variable aleatoria binomial no requiere evaluar sumas. La demostración del resultado para $E(X)$ se ilustra en el ejercicio 60.

Ejemplo 3.33 Si 75% de las compras en cierta tienda se hacen con tarjeta de crédito y X es el número de compras hechas con una tarjeta de crédito entre diez compras elegidas al azar, entonces $X \sim \text{Bin}(10, 0.75)$. Por consiguiente, $E(X) = np = (10)(0.75) = 7.5$, $V(X) = npq = 10(0.75)(0.25) = 1.875$ y $\sigma = \sqrt{1.875}$. De nuevo, aunque X sólo puede tomar valores enteros, $E(X)$ no necesita ser un entero. Si se lleva a cabo un número grande de experimentos binomiales independientes, cada uno con $n = 10$ ensayos y $p = 0.75$, entonces el número promedio de éxitos (S) por experimento será cercano a 7.5. ■

Ejercicios | Sección 3.4 (44 a 63)

44. Calcule las siguientes probabilidades binomiales directamente de la fórmula para $b(x; n, p)$:
- **a.** $b(3; 8, 0.6)$
- **b.** $b(5; 8, 0.6)$
- **c.** $P(3 \le X \le 5)$ cuando $n = 8$ y $p = 0.6$
- **d.** $P(1 \le X)$ cuando $n = 12$ y $p = 0.1$

45. Utilice la tabla A.1 del apéndice para obtener las siguientes probabilidades:
- **a.** $B(4; 10, 0.3)$
- **b.** $b(4; 10, 0.3)$
- **c.** $b(6; 10, 0.7)$
- **d.** $P(2 \le X \le 4)$ cuando $X \sim \text{Bin}(10, 0.3)$
- **e.** $P(2 \le X)$ cuando $X \sim \text{Bin}(10, 0.3)$
- **f.** $P(X \le 1)$ cuando $X \sim \text{Bin}(10, 0.7)$
- **g.** $P(2 < X < 6)$ cuando $X \sim \text{Bin}(10, 0.3)$

46. Cuando se prueban tarjetas de circuitos empleadas en la manufactura de reproductores de discos compactos, a la larga el porcentaje de partes defectuosas es de 5%. Sea X = número de tarjetas defectuosas en una muestra aleatoria de tamaño $n = 25$, entonces $X \sim \text{Bin}(25, 0.05)$.
- **a.** Determine $P(X \le 2)$.
- **b.** Determine $P(X \ge 5)$.
- **c.** Determine $P(1 \le X \le 4)$.
- **d.** ¿Cuál es la probabilidad de que ninguna de las 25 tarjetas esté defectuosa?
- **e.** Calcule el valor esperado y desviación estándar de X.

47. Una compañía que produce cristal fino sabe por experiencia que 10% de sus copas tienen imperfecciones y se deben clasificar como "de segunda".
- **a.** Entre seis copas seleccionadas al azar, ¿qué tan probable es que sólo una sea de segunda?
- **b.** Entre seis copas seleccionadas al azar, ¿cuál es la probabilidad de que por lo menos dos sean de segunda?
- **c.** Si se examina una por una a las copas, ¿cuál es la probabilidad de que a lo sumo se deban seleccionar cinco para encontrar cuatro que no sean de segunda?

48. Suponga que sólo 25% de los automovilistas se detienen por completo en un cruce con luces rojas intermitentes en todas direcciones cuando no está visible ningún otro automóvil. ¿Cuál es la probabilidad de que, de 20 automovilistas seleccionados al azar que llegan a la intersección en estas condiciones,
- **a.** A lo sumo seis se detengan por completo?
- **b.** Exactamente seis se detengan por completo?
- **c.** Al menos seis se detengan por completo?
- **d.** ¿Cuántos de los siguientes 20 automovilistas espera el lector que se detengan por completo?

49. En el ejercicio 29 (sección 3.3) se dio la fmp de Y, la cantidad de infracciones de tránsito para un individuo seleccionado al azar que cuenta con un seguro de cierta compañía. ¿Cuál es la probabilidad de que entre 15 individuos seleccionados al azar

a. Por lo menos 10 no tengan infracciones?
b. Menos de la mitad por lo menos tengan una infracción?
c. La cantidad de quienes tienen por lo menos una infracción esté entre 5 y 10, inclusive?*

50. Un tipo de raqueta de tenis se fabrica en dos tamaños: mediano y extragrande. Sesenta por ciento de los clientes de cierta tienda prefieren la versión extragrande.
a. Entre 10 clientes seleccionados al azar que prefieren este tipo de raqueta, ¿cuál es la probabilidad de que por lo menos seis prefieran la versión extragrande?
b. Entre 10 clientes seleccionados al azar, ¿cuál es la probabilidad de que el número que prefiere la versión extragrande esté dentro de una desviación estándar del valor medio?
c. La tienda tiene en existencia siete raquetas de cada modelo. ¿Cuál es la probabilidad de que los siguientes 10 clientes que prefieren este tipo de raqueta obtengan la versión que desean de las existencias actuales?

51. Veinte por ciento de todos los teléfonos de cierto tipo son llevados a reparación cuando todavía está vigente la garantía. De éstos, 60% se reparan, en tanto que el otro 40% se debe sustituir por nuevas unidades. Si una compañía compra 10 de estos teléfonos, ¿cuál es la probabilidad de que exactamente dos terminen siendo reemplazados dentro del periodo de garantía?

52. El consejo universitario informa que 2% de los 2 millones de estudiantes de secundaria que se presentan al SAT (*scholastic aptitude test*) cada año reciben lugares especiales como resultado de incapacidades documentadas (*Los Angeles Times*, 16 de julio de 2002). Considere una muestra aleatoria de 25 estudiantes que ha realizado el examen en fechas recientes.
a. ¿Cuál es la probabilidad de que exactamente uno recibió un lugar especial?
b. ¿Cuál es la probabilidad de que por lo menos uno recibió un lugar especial?
c. ¿Cuál es la probabilidad de que por lo menos dos recibieron un lugar especial?
d. ¿Cuál es la probabilidad de que el número entre los 25 que recibieron un lugar especial esté dentro de dos desviaciones estándar del número que el lector esperaría pudieran ser acomodados?
e. Suponga que a un estudiante que no recibe un lugar especial se le dan tres horas para el examen, mientras que a un estudiante con lugar especial se le dan 4.5 horas. ¿Cuál esperaría que fuera el tiempo promedio permitido a los 25 estudiantes seleccionados?

53. Suponga que 90% de las baterías de cierto proveedor tienen voltajes aceptables. Cierto tipo de linterna requiere dos baterías tipo D, y la linterna trabaja sólo si ambas baterías tienen voltajes aceptables. Entre diez linternas elegidas al azar, ¿cuál es la probabilidad de que por lo menos funcionen nueve? ¿Qué suposiciones hizo al momento de contestar la pregunta planteada?

54. Un distribuidor recibe un lote muy grande de componentes. El lote se puede clasificar como aceptable sólo si la proporción de componentes defectuosos es a lo sumo 0.10. El distribuidor decide seleccionar al azar 10 componentes y aceptar el lote sólo si el número de componentes defectuosos en la muestra es a lo sumo 2.
a. ¿Cuál es la probabilidad de que el lote sea aceptado cuando la proporción real de piezas defectuosas es 0.01? 0.05? 0.10? 0.20? 0.25?
b. Sea p la proporción real de piezas defectuosas en el lote. Una gráfica de P(lote aceptado) en función de p, con p en el eje horizontal y P(lote aceptado) en el eje vertical, se llama *curva característica de operación* para el plan de muestreo de aceptación del lote. Utilice los resultados del inciso (a) a fin de construir esta curva en $0 \le p \le 1$.
c. Repita los incisos (a) y (b) con "1" en lugar de "2" en el plan de muestreo de aceptación.
d. Repita los incisos (a) y (b) con "15" en lugar de "10" en el plan de muestreo de aceptación.
e. ¿Cuál de los tres planes de muestreo, incisos (a), (c) o (d) es al parecer más satisfactorio y por qué?

55. Un decreto que requiere la instalación de un detector de humo en todas las casas construidas con anterioridad ha estado en vigor en una ciudad durante un año. Al departamento de bomberos le preocupa que muchas casas aún no tengan detectores. Sea p = la proporción real de casas que tienen detectores, y supóngase que se inspecciona una muestra aleatoria de 25 casas. Si la muestra indica de manera enfática que menos de 80% de las casas tienen detector, el departamento de bomberos promoverá una campaña para llevar a cabo un programa de inspección obligatorio. Como resultado del alto precio del programa, el departamento prefiere no solicitar las inspecciones a menos que la evidencia muestral justifique que es necesario. Sea X el número de casas con detectores entre las 25 de la muestra. Considere rechazar la afirmación de que $p \ge 0.8$ si $x \le 15$.
a. ¿Cuál es la probabilidad de que se rechace la afirmación cuando el valor real de p es 0.8?
b. ¿Cuál es la probabilidad de no rechazar la afirmación cuando $p = 0.7$? ¿Cuándo $p = 0.6$?

*"Entre a y b, inclusive" es equivalente a $(a \le X \le b)$.

c. ¿Cómo cambian las "probabilidades de error" de los incisos (a) y (b) si se sustituye por 14 el valor 15 de la regla de decisión?

56. Un puente de peaje cobra un dólar a los automóviles de pasajeros y 2.50 a otros vehículos. Supóngase que durante las horas diurnas, 60% de los automóviles son de pasajeros. Si 25 vehículos cruzan el puente durante determinado periodo diurno, ¿cuál es el ingreso por peaje esperado? [*Sugerencia*: sea X = número de automóviles de pasajeros, entonces el ingreso por peaje $h(X)$ es una función lineal de X.]

57. Un estudiante que está tratando de escribir un trabajo para un curso tiene la opción de dos temas, A y B. Si elige el tema A, pedirá dos libros por medio de préstamo interbibliotecario, pero si selecciona el B, pedirá cuatro libros. El estudiante cree que un buen trabajo requiere por lo menos la mitad de los libros solicitados para cualquier tema elegido. Si la probabilidad de que un libro solicitado por medio de préstamo interbibliotecario en realidad llegue a tiempo es 0.9 y los libros llegan de modo independiente entre sí, ¿cuál tema debe seleccionar el estudiante para maximizar la probabilidad de hacer un buen trabajo? ¿Cuál si la probabilidad de llegada es sólo 0.5 en lugar de 0.9?

58. **a.** Para n fija, ¿hay valores de p $(0 \le p \le 1)$ para los cuales $V(X) = 0$? Explique por qué esto es así.

b. ¿Para qué valor de p se maximiza $V(X)$? [*Sugerencia*: trace la gráfica de $V(X)$ en función de p, o bien obtenga la derivada.]

59. **a.** Demuestre que $b(x; n, 1 - p) = b(n - x; n, p)$.

b. Demuestre que $B(x; n, 1 - p) = 1 - B(n - x - 1; n, p)$. [*Sugerencia*: a lo sumo x éxitos (S) es equivalente a por lo menos $(n - x)$ fracasos (F).]

c. ¿Qué implican los incisos (a) y (b) acerca de la necesidad de incluir valores de p mayores que 0.5 en la tabla A.1 del apéndice?

60. Demuestre que $E(X) = np$ cuando X es una variable aleatoria binomial. [*Sugerencia*: primero exprese $E(X)$ como una suma con límite inferior $x = 1$. Luego factorice np, y sea $y = x - 1$ de modo que la suma sea de $y = 0$ a $y = n - 1$, y demuestre que la suma es igual a 1].

61. Los clientes de una gasolinera pagan con tarjeta de crédito (A), tarjeta de débito (B) o en efectivo (C). Suponga que clientes sucesivos eligen su forma de pago de manera independiente, con $P(A) = 0.5$, $P(B) = 0.2$ y $P(C) = 0.3$.

a. Entre los siguientes 100 clientes, ¿cuáles son la media y varianza del número de quienes pagan con tarjeta de crédito? Explique su razonamiento.

b. Conteste el inciso (a) para el número entre los 100 que no pagan en efectivo.

62. La limosina de un aeropuerto tiene espacio para cuatro pasajeros en cualquier viaje. La compañía acepta un máximo de seis reservaciones por viaje, y un pasajero debe tener una reservación. De registros anteriores, 20% de quienes hacen reservaciones no se presentan para el viaje. Conteste las siguientes preguntas, asumiendo independencia siempre que sea apropiado.

a. Si se hacen seis reservaciones, ¿cuál es la probabilidad de que, por lo menos, un individuo con reservación no tenga cupo para el viaje?

b. Si se hacen seis reservaciones, ¿cuál es el número esperado de lugares disponibles cuando parte la limosina?

c. Suponga que la distribución de probabilidad del número de reservaciones hechas es la que se muestra en la siguiente tabla.

Número de reservaciones	3	4	5	6
Probabilidad	0.1	0.2	0.3	0.4

Sea X el número de pasajeros en un viaje seleccionado al azar. Obtenga la función de masa de la probabilidad de X.

63. Refiérase a la desigualdad de Chebyshev dada en el ejercicio 43. Calcule $P(|X - \mu| \ge k\sigma)$ para $k = 2$ y $k = 3$ cuando $X \sim \text{Bin}(20, 0.5)$, y compare con el límite superior correspondiente. Repita para $X \sim \text{Bin}(20, 0.75)$.

3.5 Distribuciones hipergeométrica y binomial negativa

Las distribuciones hipergeométrica y binomial negativa tienen una relación muy estrecha con la distribución binomial. Si bien la distribución binomial es el modelo de la probabilidad aproximado para el muestreo sin reemplazo a partir de una población dicótoma finita (S-F), la dis-

tribución hipergeométrica es el modelo de probabilidad exacto para el número éxitos (S) en la muestra. La variable aleatoria binomial X es el número de éxitos cuando se fija el número n de ensayos, en tanto que la distribución binomial negativa surge de fijar el número deseado de éxitos y permitir que sea aleatorio el número de ensayos.

Distribución hipergeométrica

Las suposiciones que dan lugar a la distribución hipergeométrica son las siguientes:

1. La población o conjunto por muestrear consiste en N individuos, objetos o elementos (una población *finita*).
2. Cada individuo se caracteriza como un éxito (S) o fracaso (F), y hay M éxitos en la población.
3. Se elige una muestra de n individuos sin reemplazo, de manera que cada subconjunto de tamaño n tenga las mismas probabilidades de ser elegido.

La variable aleatoria de interés es X = número de éxitos en la muestra. La distribución de probabilidad de X depende de los parámetros n, M y N, así que se desea obtener $P(X = x) = h(x; n, M, N)$.

Ejemplo 3.34 Durante un periodo particular, la oficina de información tecnológica de una universidad recibió 20 órdenes de servicio para problemas con impresoras, de las cuales ocho eran láser y 12 eran de inyección de tinta. Se seleccionará una muestra de 5 de estas órdenes de servicio para inclusión en una encuesta de satisfacción del cliente. Suponga que se eligen 5 en un modo completamente aleatorio, de modo que cualquier subconjunto de tamaño 5 tenga la misma probabilidad de ser elegido que cualquier otro subconjunto (considere colocar los números 1, 2, . . . , 20 o 20 en tiras idénticas de papel, mezclarlas y elegir 5 de ellas). ¿Cuál es la probabilidad de que exactamente x ($x = 0, 1, 2, 3, 4$ o 5) de las órdenes de servicio elegidas fueran para impresoras láser?

En este ejemplo, el tamaño de la población es $N = 20$, el tamaño de la muestra es $n = 5$, y el número de éxitos (impresora de inyección de tinta = S) y fracasos (F) en la población son $M = 12$ y $N - M = 8$, respectivamente. Considere el valor $x = 2$. Debido a que los resultados (cada uno consta de 5 órdenes particulares) tienen las mismas probabilidades,

$$P(X = 2) = h(2; 5, 12, 20) = \frac{\text{número de resultados que tienen } X = 2}{\text{número de resultados posibles}}$$

El número de resultados posibles en el experimento es el número de formas de elegir 5 de los 20 objetos sin considerar el orden; es decir, $\binom{20}{5}$. Para contar el número de resultados que tienen $X = 2$, observe que hay $\binom{12}{2}$ formas de elegir dos órdenes de impresoras de inyección de tinta, y para cada una de estas formas hay $\binom{8}{3}$ formas de elegir las 3 órdenes de impresoras láser para llenar la muestra. Así, la regla del producto del capítulo 2 da $\binom{12}{2}\binom{8}{3}$ como el número de resultados con $X = 2$, por lo tanto

$$h(2; 5, 12, 20) = \frac{\binom{12}{2}\binom{8}{3}}{\binom{20}{5}} = \frac{77}{323} = 0.238 \quad \blacksquare$$

En general, si el tamaño n de la muestra es menor que el número de éxitos en la población (M), entonces el valor de X más grande posible es n. Sin embargo, si $M < n$ (p. ej., un tamaño de

muestra de 25 y sólo 15 éxitos en la población), entonces X puede ser a lo sumo M. De manera similar, si el número de fracasos en la población $(N - M)$ es mayor que el tamaño de la muestra, el valor más pequeño posible de X es 0 (puesto que entonces los individuos muestreados podrían ser fracasos). Sin embargo, si $N - M < n$, el valor de X más pequeño posible es $n - (N - M)$. En resumen, los valores posibles de X satisfacen la restricción $\text{máx}(0, n - (N - M)) \le x \le \text{mín}(n, M)$. Un argumento semejante al del ejemplo previo proporciona la fmp de X.

PROPOSICIÓN

Si X es el número de éxitos (S) en una muestra completamente aleatoria de tamaño n extraídos de una población que consiste en M éxitos y $(N - M)$ fracasos, entonces la distribución de probabilidad de X, denominada **distribución hipergeométrica**, está dada por

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N - M}{n - x}}{\binom{N}{n}} \tag{3.15}$$

para un entero x que satisface $\text{máx}(0, n - N + M) \le x \le \text{mín}(n, M)$.

En el ejemplo 3.34, $n = 5$, $M = 12$ y $N = 20$, así que $h(x; 5, 12, 20)$ para $x = 0, 1, 2, 3, 4, 5$ se puede obtener sustituyendo estos números en la ecuación (3.15).

Ejemplo 3.35 Cinco individuos de una población animal que se cree está cerca de la extinción en cierta región, fueron capturados, marcados y liberados para mezclarse con la población. Después que tuvieron la oportunidad de mezclarse, se seleccionó una muestra aleatoria de 10 de estos animales. Sea X el número de animales marcados en la segunda muestra. Si en realidad hay 25 animales de este tipo en la región, ¿cuál es la probabilidad de que (a) $X = 2$ y (b) $X \le 2$?

Los valores de los parámetros son $n = 10$, $M = 5$ (animales marcados en la población), y $N = 25$, así que

$$h(x; 10, 5, 25) = \frac{\binom{5}{x}\binom{20}{10 - x}}{\binom{25}{10}} \quad x = 0, 1, 2, 3, 4, 5$$

Para el inciso (a),

$$P(X = 2) = h(2; 10, 5, 25) = \frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} = 0.385$$

Para el inciso (b),

$$P(X \le 2) = P(X = 0, 1, \text{o } 2) = \sum_{x=0}^{2} h(x; 10, 5, 25)$$
$$= 0.057 + 0.257 + 0.385 = 0.699$$

■

Se cuenta con las tablas completas de la distribución hipergeométrica, pero debido a que la distribución tiene tres parámetros, estas tablas requieren mucho más espacio que las tablas de la distribución binomial. MINITAB y otros paquetes estadísticos de software, generan con facilidad probabilidades hipergeométricas.

Como en el caso binomial, hay expresiones simples para $E(X)$ y $V(X)$ para variables aleatorias hipergeométricas.

PROPOSICIÓN

La media y la varianza de la variable aleatoria hipergeométrica X que tiene fmp $h(x; n, M, N)$ son

$$E(X) = n \cdot \frac{M}{N} \qquad V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$$

La relación M/N es la proporción de éxitos (S) en la población. Si se reemplaza a M/N por p en $E(X)$ y $V(X)$, se obtiene

$$E(X) = np$$
$$V(X) = \left(\frac{N-n}{N-1}\right) \cdot np(1-p) \tag{3.16}$$

En la expresión (3.16) se observa que las medias de las variables aleatorias binomial e hipergeométrica son iguales, en tanto que las varianzas de las dos variables aleatorias difieren por el factor $(N - n)/(N - 1)$, más conocido como **factor de corrección por población finita**. Este factor es menor que 1, así que la variable hipergeométrica tiene varianza más pequeña que la variable aleatoria binomial. El factor de corrección se puede escribir como $(1 - n/N)/(1 - 1/N)$, cuyo valor aproximado es 1 cuando n es relativamente pequeña con respecto a N.

Ejemplo 3.36 (continuación del ejemplo 3.35)

En el ejemplo del marcado de animales, $n = 10$, $M = 5$ y $N = 25$, así que $p = \frac{5}{25} = 0.2$ y

$$E(X) = 10(0.2) = 2$$
$$V(X) = \frac{15}{24}(10)(0.2)(0.8) = (0.625)(1.6) = 1$$

Si el muestreo se realizó sin reemplazo, $V(X) = 1.6$.

Suponga que en realidad no se sabe el tamaño de la población, N, así que se observa el valor x y se desea estimar N. Es razonable igualar la proporción muestral observada de éxitos, x/n, con la proporción poblacional, M/N, con lo que se obtiene la estimación

$$\hat{N} = \frac{M \cdot n}{x}$$

Si $M = 100$, $n = 40$ y $x = 16$, entonces $\hat{N} = 250$. ■

La regla empírica general de la sección 3.4 expresa que si el muestreo fue sin reemplazo, pero n/N fue a lo sumo 0.05, entonces se podría usar la distribución binomial para calcular probabilidades aproximadas en las que interviene el número de éxitos de la muestra. Un enunciado más preciso es el siguiente: sea N el tamaño de la población y M el número de éxitos de la población, que aumentan cuando la relación M/N tiende a p. Entonces $h(x; n, M, N)$ se apro-

xima a $b(x; n, p)$; así que para n/N pequeña, las dos son aproximadamente iguales siempre que p no esté demasiado cerca de 0 o 1. Esto es lo racional para la regla empírica.

Distribución binomial negativa

La variable aleatoria binomial negativa y la distribución se basan en un experimento que satisface las siguientes condiciones:

1. El experimento consiste en una secuencia de ensayos independientes.
2. Cada ensayo produce ya sea un éxito (S) o un fracaso (F).
3. La probabilidad de éxito es constante de un ensayo a otro, así que $P(S \text{ en el ensayo } i) = p$ para $i = 1, 2, 3, \ldots$.
4. El experimento continúa (se llevan a cabo los ensayos) hasta que se observa un total de r éxitos observados, donde r es un entero positivo especificado.

La variable de interés es X = número de fracasos que preceden al r-ésimo éxito. X se llama **variable aleatoria binomial negativa** porque, en contraste con la variable aleatoria binomial, el número de éxitos es fijo y el número de ensayos es aleatorio.

Los valores posibles de X son $0, 1, 2, \ldots$. Sea $nb(x; r, p)$ la fmp de X. El evento $\{X = x\}$ es equivalente a $\{r - 1$ éxitos (S) en los primeros $(x + r - 1)$ ensayos y una S en el $(x + r)$-ésimo ensayo$\}$ (p. ej., si $r = 5$ y $x = 10$, entonces debe haber cuatro letras S en los primeros 14 ensayos y el ensayo 15 debe ser una S). Puesto que los ensayos son independientes,

$$\begin{aligned} nb(x; r, p) &= P(X = x) \\ &= P(r - 1 \text{ letras } S \text{ en los primeros } x + r - 1 \text{ ensayos}) \cdot P(S) \end{aligned} \tag{3.17}$$

La primera probabilidad en el lado derecho de la expresión (3.17) es la probabilidad binomial

$$\binom{x + r - 1}{r - 1} p^{r-1}(1 - p)^x \qquad \text{donde } P(S) = p$$

PROPOSICIÓN

La fmp de la variable aleatoria binomial negativa X con parámetros r = número de éxitos (S) y $p = P(S)$ is

$$nb(x; r, p) = \binom{x + r - 1}{r - 1} p^r(1 - p)^x \qquad x = 0, 1, 2, \ldots$$

Ejemplo 3.37 Un pediatra desea reclutar a 5 parejas, cada una en espera de su primer niño, para participar en un nuevo régimen de parto natural. Sea $p = P(\text{una pareja elegida al azar decide participar})$. Si $p = 0.2$, ¿cuál es la probabilidad de que se les deba preguntar a 15 parejas antes de encontrar a 5 que estén de acuerdo en participar? Es decir, con $S = \{\text{está de acuerdo en participar}\}$, cuál es la probabilidad de que ocurran 10 fracasos antes del quinto éxito? Sustituyendo $r = 5, p = 0.2$ y $x = 10$ en $nb(x; r, p)$ se obtiene

$$nb(10; 5, 0.2) = \binom{14}{4}(0.2)^5(0.8)^{10} = 0.034$$

La probabilidad de que a lo sumo se observen 10 fracasos (F) (se les pregunte a lo sumo a 15 parejas) es

$$P(X \le 10) = \sum_{x=0}^{10} nb(x; 5, 0.2) = (0.2)^5 \sum_{x=0}^{10} \binom{x+4}{4} (0.8)^x = 0.164 \quad \blacksquare$$

En algunas fuentes, la variable aleatoria binomial negativa se toma como el número de ensayos $X + r$ en vez del número de fracasos.

En el caso especial $r = 1$, la fmp es

$$nb(x; 1, p) = (1 - p)^x p \qquad x = 0, 1, 2, \ldots \tag{3.18}$$

En el ejemplo 3.10, se dedujo la fmp para el número de ensayos necesarios a fin de obtener la primera S, y la fmp es similar a la expresión (3.18). Tanto X = número de fracasos como Y = número de ensayos ($= 1 + X$) se denominan **variables aleatorias geométricas**, y la fmp de la expresión (3.18) se llama **distribución geométrica.**

En el ejemplo 3.18, se demostró que el número esperado de ensayos hasta la primera S es $1/p$, así que el número esperado de fracasos hasta el primer éxito es $(1/p) - 1 = (1 - p)/p$. De manera intuitiva, se esperaría ver $r \cdot (1 - p)/p$ fracasos entes del r-ésimo éxito, y de hecho esto es $E(X)$. Hay una fórmula simple para $V(X)$.

PROPOSICIÓN

Si X es una variable aleatoria binomial negativa con fmp $nb(x; r, p)$, entonces

$$E(X) = \frac{r(1-p)}{p} \qquad V(X) = \frac{r(1-p)}{p^2}$$

Por último, al desarrollar el coeficiente binomial de $p^r(1 - p)^x$ y hacer alguna cancelación, se puede ver que $nb(x; r, p)$ está bien definida cuando r no es un entero. Se ha encontrado que esta *distribución binomial negativa generalizada* ajusta bastante bien los datos observados en una amplia variedad de aplicaciones.

Ejercicios | Sección 3.5 (64 a 74)

64. Cierto tipo de cámara digital viene en una versión de tres megapixeles y una de cuatro megapixeles. Una tienda de cámaras recibió un envío de 15 de estas cámaras, de las cuales seis tienen una resolución de tres megapixeles. Suponga que se eligen al azar cinco de estas cámaras para colocarlas en el mostrador; las otras diez se colocan en un almacén. Sea X = número de cámaras de tres megapixeles, entre las cinco elegidas para almacenarlas en el mostrador.

a. ¿Qué clase de distribución tiene X (nombre y valores de los parámetross)?

b. Calcule $P(X = 2)$, $P(X \le 2)$, and $P(X \ge 2)$.

c. Calcule el valor medio y la desviación estándar de X.

65. Cada uno de los 12 refrigeradores de cierto tipo se devolvió a un distribuidor debido a un sonido agudo oscilante cuando el refrigerador está funcionando. Suponga que siete de estos refrigeradores tienen un compresor defectuoso y los otros cinco tienen problemas menos graves. Si se examinan al azar, sea X el número entre los primeros seis examinados que tienen un compresor defectuoso. Calcule lo siguiente:

a. $P(X = 5)$

b. $P(X \le 4)$

c. La probabilidad de que X exceda su valor medio en más de una desviación estándar.

d. Considere un gran envío de 400 refrigeradores, de los cuales 40 tienen compresores defectuosos. Si X es el número entre 15 refrigeradores elegidos al azar que tienen compresores defectuosos, describa una forma menos tediosa para calcular (por lo menos de manera aproximada) $P(X \leq 5)$ que usar la fmp hipergeométrica.

66. Un profesor que enseñó dos secciones del último trimestre de estadística en ingeniería, el primero con 20 alumnos y el segundo con 30, decidió asignar un proyecto. Después de recibir todos los proyectos, el profesor los ordenó al azar antes de calificarlos. Considere los 15 primeros proyectos calificados.

a. ¿Cuál es la probabilidad de que exactamente 10 de éstos provengan de la segunda sección?

b. ¿Cuál es la probabilidad de que por lo menos 10 de éstos vengan de la segunda sección?

c. ¿Cuál es la probabilidad de que al menos 10 de éstos sean de la misma sección?

d. ¿Cuáles son el valor medio y la desviación estándar del número entre estos 15 que son de la segunda sección?

e. ¿Cuáles son el valor medio y la desviación estándar del número de proyectos que no están entre los 15 primeros que son de la segunda sección?

67. Un geólogo reunió 10 especímenes de roca basáltica y 10 de granito. El geólogo instruye a un asistente de laboratorio a seleccionar al azar 15 de los especímenes para análisis.

a. ¿Cuál es la fmp del número de especímenes de granito seleccionados para el análisis?

b. ¿Cuál es la probabilidad de que los especímenes de uno de los dos tipos de roca sean seleccionados para el análisis?

c. ¿Cuál es la probabilidad de que la cantidad de especímenes de granito seleccionados para su análisis esté dentro de 1a desviación estándar de su valor promedio?

68. Un gerente de personal que entrevista a 11 ingenieros para cuatro vacantes, ha programado seis entrevistas para el primer día de entrevistas y cinco para el segundo. Suponga que los candidatos son entrevistados al azar.

a. ¿Cuál es la probabilidad de que x de los cuatro mejores candidatos sean entrevistados el primer día?

b. ¿Cuántos de los cuatro mejores candidatos se espera que sean entrevistados el primer día?

69. Veinte pares de individuos que juegan en un torneo de *bridge* han sido clasificados del 1 al 20. En la primera parte del torneo, los 20 se dividen al azar en 10 parejas este-oeste y 10 parejas norte-sur.

a. ¿Cuál es la probabilidad de que x de las 10 mejores parejas terminen jugando este-oeste?

b. ¿Cuál es la probabilidad de que las cinco mejores parejas terminen jugando en la misma dirección?

c. Si hay $2n$ parejas, ¿cuál es la fmp de X = número entre las n mejores parejas que terminan jugando este-oeste? ¿Cuáles son $E(X)$ y $V(X)$?

70. En cierta área del condado de Los Ángeles entró en vigor la segunda etapa de contingencia ambiental, en la que hay 50 empresas industriales. Un inspector visitará 10 empresas seleccionadas al azar para comprobar violaciones a las normas.

a. Si 15 de las empresas en realidad violan al menos un reglamento, ¿cuál es la fmp del número de empresas que visita el inspector y violan por lo menos una norma?

b. Si hay 500 empresas en la zona, de las cuales 150 no cumplen con las normas, aproxime la fmp del inciso (a) mediante una fmp más sencilla.

c. Para X = número entre las 10 empresas visitadas que violan las normas, calcule $E(X)$ y $V(X)$ tanto para la fmp exacta, como para la fmp aproximada del inciso (b).

71. Suponga que $p = P$(de que nazca un varón) = 0.5. Una pareja desea tener exactamente dos niñas en su familia. Tendrán hijos hasta que se satisfaga esta condición.

a. ¿Cuál es la probabilidad de que la familia tenga x hijos varones?

b. ¿Cuál es la probabilidad de que la familia tenga cuatro hijos?

c. ¿Cuál es la probabilidad de que la familia tenga a lo sumo cuatro hijos?

d. ¿Cuántos varones se esperaría que tuviera esta familia? ¿Cuántos hijos se esperaría que tuviera esta familia?

72. Una familia decide tener hijos hasta tener tres del mismo género. Si se supone que $P(B) = P(G) = 0.5$, ¿cuál es la fmp de X = número de hijos en la familia?

73. Tres hermanos y sus esposas deciden tener hijos hasta que cada familia tenga dos niñas. ¿Cuál es la fmp de X = número total de varones que tienen los hermanos? ¿Cuál es $E(X)$ y cómo se compara con el número esperado de varones que tiene cada hermano?

74. El individuo A tiene un dado rojo y el B un dado verde (ambos no cargados). Si cada uno tira su dado hasta obtener "pares" (1–1, . . . , 6–6), ¿cuál es la fmp de X = número total de veces que se tira un dado? ¿Cuáles son $E(X)$ y $V(X)$?

3.6 Distribución de probabilidad de Poisson

Las distribuciones binomial, hipergeométrica y binomial negativa se obtuvieron comenzando con un experimento que consiste en ensayos o eliminaciones y aplicar las leyes de probabilidad a varios resultados del experimento. No hay experimento sencillo en que se base la distribución de Poisson, aunque en breve se describe cómo se puede obtener mediante ciertas operaciones limitantes.

DEFINICIÓN

Se dice que una variable aleatoria X tiene una **distribución de Poisson** con parámetro λ ($\lambda > 0$) si la pmf de X es

$$p(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

El valor de λ es por lo común una tasa por unidad de tiempo o por unidad de área. La letra e en $p(x; \lambda)$ representa la base del sistema de logaritmos naturales; su valor numérico aproximado es 2.71828. Debido a que λ debe ser positiva, $p(x; \lambda) > 0$ para todos los valores posibles de x. El hecho de que $\sum_{x=0}^{\infty} p(x; \lambda) = 1$ es una consecuencia del desarrollo en serie infinita de Maclaurin de e^{λ}, que aparece en la mayoría de los libros de cálculo:

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \tag{3.19}$$

Si los dos términos extremos de la expresión (3.19) se multiplican por $e^{-\lambda}$ y después $e^{-\lambda}$ se coloca dentro de la suma, el resultado es

$$1 = \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!}$$

lo que demuestra que $p(x; \lambda)$ satisface la segunda condición necesaria para especificar una fmp.

Ejemplo 3.38 Sea X la cantidad de criaturas de cierto tipo capturadas en una trampa durante determinado periodo. Supóngase que X tiene una distribución de Poisson con $\lambda = 4.5$, así que en promedio las trampas contienen 4.5 criaturas. [En el artículo "Dispersal Dynamics of the Bivalve *Gemma Gemma* in a Patchy Environment" (*Ecological Monographs*, 1995: 1-20) se sugiere este modelo; el bivalvo *Gemma gemma* es una pequeña almeja.] La probabilidad de que una trampa contenga exactamente cinco criaturas es

$$P(X = 5) = \frac{e^{-4.5}(4.5)^5}{5!} = 0.1708$$

La probabilidad de que una trampa tenga a lo sumo cinco criaturas es

$$P(X \le 5) = \sum_{x=0}^{5} \frac{e^{-4.5}(4.5)^x}{x!} = e^{-4.5}\left[1 + 4.5 + \frac{(4.5)^2}{2!} + \cdots + \frac{(4.5)^5}{5!}\right] = 0.7029 \qquad \blacksquare$$

Distribución de Poisson como límite

La siguiente proposición proporciona el fundamento para usar la distribución de Poisson en muchas situaciones.

PROPOSICIÓN

Suponga que en la fmp binomial $b(x; n, p)$, sea $n \to \infty$ y $p \to 0$ de tal manera que np tiende a un valor $\lambda > 0$. Entonces $b(x; n, p) \to p(x; \lambda)$.

De acuerdo con esta proposición, *en cualquier experimento binomial en el que n es grande y p es pequeña*, $b(x; n, p) \approx p(x; \lambda)$, *donde* $\lambda = np$. Como regla empírica, esta aproximación se puede aplicar con toda seguridad si $n \geq 100$, $p \leq 0.01$ y $np \leq 20$.

Ejemplo 3.39 Si un editor de novelas se esfuerza por asegurar que sus libros están libres de errores tipográficos, de modo que la probabilidad de que alguna página contenga por lo menos un error es 0.005 y los errores son independientes de una página a otra, ¿cuál es la probabilidad de que una de sus novelas de 400 páginas contenga exactamente una página con errores? ¿A lo sumo tres páginas con errores?

Si S denota una página que contiene por lo menos un error y F una página sin errores, el número X de páginas que contienen por lo menos un error es una variable aleatoria binomial con $n = 400$ y $p = 0.005$, así que $np = 2$. Se desea

$$P(X = 1) = b(1; 400, 0.005) \approx p(1; 2) = \frac{e^{-2}(2)^1}{1!} = 0.271$$

De manera similar,

$$P(X \leq 3) \approx \sum_{x=0}^{3} p(x; 2) = \sum_{x=0}^{3} e^{-2}\frac{2^x}{x!} = 0.135 + 0.271 + 0.271 + 0.180$$

$$= 0.857 \qquad \blacksquare$$

En la tabla A.2 del apéndice se ilustra la fda $F(x; \lambda)$ para $\lambda = 0.1, 0.2, \ldots, 1, 2, \ldots, 10, 15$ y 20. Por ejemplo, si $\lambda = 2$, entonces $P(X \leq 3) = F(3; 2) = 0.857$ como en el ejemplo 3.39, en tanto que $P(X = 3) = F(3; 2) - F(2; 2) = 0.180$. De manera alternativa, muchos paquetes de computadora para estadística generan $p(x; \lambda)$ y $F(x; \lambda)$ cuando se solicita.

Media y varianza de *X*

Puesto que $b(x; n, p) \to p(x; \lambda)$ cuando $n \to \infty$, $p \to 0$, $np \to \lambda$, la media y la varianza de una variable binomial deben aproximarse a las de una variable de Poisson. Estos límites son $np \to \lambda$ y $np(1 - p) \to \lambda$.

PROPOSICIÓN

Si X tiene una distribución de Poisson con parámetro λ, entonces $E(X) = V(X) = \lambda$.

Estos resultados también se pueden obtener de manera directa de las definiciones de media y varianza.

Ejemplo 3.40 (continuación del ejemplo 3.38) Tanto la cantidad esperada de almejas atrapadas como la varianza de esta cantidad son iguales a 4.5 y $\sigma_X = \sqrt{\lambda} = \sqrt{4.5} = 2.12$. ■

Proceso de Poisson

Una aplicación muy importante de la distribución de Poisson surge en relación con la ocurrencia de eventos de un tipo particular con el tiempo. Como ejemplo, suponga que empezando en un tiempo $t = 0$, se tiene interés en contar el número de impulsos radiactivos que registra un contador Geiger. Se hacen las suposiciones siguientes acerca de la forma en que ocurren los impulsos:

1. Existe un parámetro $\alpha > 0$ tal que para cualquier intervalo pequeño de tiempo de longitud Δt, la posibilidad de que se reciba exactamente un impulso es $\alpha \cdot \Delta t + o(\Delta t)$.*
2. La probabilidad de que se reciba más de un impulso durante Δt es $o(\Delta t)$ [que junto con la suposición 1, implica que la probabilidad de ausencia de impulsos durante Δt es $1 - \alpha \cdot \Delta t - o(\Delta t)$].
3. El número de impulsos recibidos durante el intervalo de tiempo Δt es independiente del número recibido antes de este intervalo de tiempo.

De manera informal, la suposición 1 establece que para un intervalo corto de tiempo, la probabilidad de recibir un solo impulso es más o menos proporcional a la longitud del intervalo de tiempo, donde α es la constante de proporcionalidad. Ahora, sea $P_k(t)$ la probabilidad de que el contador reciba k impulsos durante cualquier intervalo de tiempo de longitud t.

PROPOSICIÓN

$P_k(t) = e^{-\alpha t} \cdot (\alpha t)^k / k!$, de modo que el número de impulsos durante un intervalo de tiempo de longitud t es una variable aleatoria de Poisson con parámetro $\lambda = \alpha t$. Por lo tanto, el número esperado de impulsos durante cualquier intervalo de tiempo es αt, así que el número esperado durante un intervalo de tiempo unitario es α.

Ejemplo 3.41 Suponga que llegan impulsos al contador a una tasa promedio de seis por minuto, así que $\alpha = 6$. Para hallar la probabilidad de que en un intervalo de medio minuto se reciba por lo menos un impulso, observe que el número de impulsos en tal intervalo tiene una distribución de Poisson con parámetro $\alpha t = 6(0.5) = 3$ (se utiliza medio minuto porque α se expresa como una proporción por minuto). Entonces, con X = número de impulsos recibidos en el intervalo de 30 segundos,

$$P(1 \le X) = 1 - P(X = 0) = 1 - \frac{e^{-3}(3)^0}{0!} = 0.950$$ ■

Si en las suposiciones 1 a 3 se sustituye "impulso" por "evento", entonces el número de eventos que ocurren durante un intervalo fijo de tiempo de longitud t tiene una distribución de Poisson con parámetro αt. Cualquier proceso que tiene esta distribución se llama **proceso**

*Una cantidad es $o(\Delta t)$ (que se lee "o minúscula de delta t") si, cuando Δt tiende a 0, también lo hace $o(\Delta t)/\Delta t$. Es decir, $o(\Delta t)$ es incluso más insignificante que la misma Δt. La cantidad $(\Delta t)^2$ tiene esta propiedad, pero no sen(Δt).

de Poisson, y α se llama *tasa del proceso*. Otros ejemplos de situaciones que dan lugar a un proceso de Poisson son el monitoreo del estado de un sistema de computadoras a través del tiempo, con interrupciones que constituyen los eventos de interés, el registro del número de accidentes en una planta industrial con el paso del tiempo, contestar llamadas en un conmutador telefónico y observar el número de rayos cósmicos desde un observatorio con el tiempo.

En vez de observar eventos con el transcurso del tiempo, considere eventos de cierto tipo que ocurren en una región de dos o tres dimensiones. Por ejemplo, en un mapa se podría seleccionar cierta región R de un bosque, ir a esa región y contar el número de árboles. Cada árbol representaría un evento que ocurre en un punto particular del espacio. Bajo suposiciones similares, como de la 1 a la 3, se puede demostrar que el número de eventos que ocurren en la región R tiene una distribución de Poisson con parámetro $\alpha \cdot a(R)$, donde $a(R)$ es el área de R. La cantidad α es el número esperado de eventos por unidad de área o volumen.

Ejercicios | Sección 3.6 (75 a 89)

75. Sea X la cantidad de fracturas en la superficie de una caldera de cierto tipo, seleccionada al azar, que tiene una distribución de Poisson con parámetro $\lambda = 5$. Utilice la tabla A.2 del apéndice para calcular las probabilidades siguientes:

a. $P(X \le 8)$ **b.** $P(X = 8)$ **c.** $P(9 \le X)$

d. $P(5 \le X \le 8)$ **e.** $P(5 < X < 8)$

76. Suponga que el número X de tornados observados en determinada región durante un año tiene una distribución de Poisson con $\lambda = 8$.

a. Calcule $P(X \le 5)$.

b. Calcule $P(6 \le X \le 9)$.

c. Calcule $P(10 \le X)$.

d. ¿Cuál es la probabilidad de que el número observado de tornados sea mayor que el número esperado más de una desviación estándar?

77. Suponga que la cantidad de conductores que viajan entre cierto origen y destino durante un periodo determinado, tiene una distribución de Poisson con parámetro $\lambda = 20$ (sugerido en el artículo "Dynamic Ride Sharing: Theory and Practice", *J. of Transp. Engr.*, 1997: 308-312). ¿Cuál es la probabilidad de que la cantidad de conductores

a. sea a lo sumo 10?

b. sea mayor que 20?

c. esté entre 10 y 20, inclusive? ¿Esté estrictamente entre 10 y 20?

d. esté dentro de dos desviaciones estándar respecto de la media?

78. Considere escribir en un disco de computadora y luego enviarlo a través de un certificador que cuenta el número de impulsos faltantes. Suponga que este número X tiene una distribución de Poisson con parámetro $\lambda = 0.2$. (Sugerido en "Average Sample Number for Semi-Curtailed Sampling Using the Poisson Distribution". *J. Quality Technology*, 1983: 126-129.)

a. ¿Cuál es la probabilidad de que en el disco falte exactamente un impulso?

b. ¿Cuál es la probabilidad de que en el disco falten por lo menos dos impulsos?

c. Si se eligen de forma independiente dos discos, ¿cuál es la probabilidad de que ninguno contenga un impulso faltante?

79. En un artículo de *Los Angeles Times* (3 de diciembre de 1993) se informa que una de cada 200 personas es portadora del gen defectuoso que ocasiona cáncer de colon hereditario. En una muestra de 1000 personas, ¿cuál es la distribución aproximada del número de quienes son portadores del gen? Utilice esta distribución para calcular la probabilidad aproximada de que:

a. Entre 5 y 8 (inclusive) sean portadores del gen.

b. Por lo menos 8 lleven el gen.

80. Suponga que sólo 0.10% de las computadoras de cierto tipo experimentan falla del CPU durante el periodo de garantía. Considere una muestra de 10 000 computadoras.

a. ¿Cuál es el valor esperado y la desviación estándar del número de computadoras de la muestra que tienen el defecto?

b. ¿Cuál es la probabilidad (aproximada) de que más de 10 computadoras muestreadas tengan el defecto?

c. ¿Cuál es la probabilidad (aproximada) de que ninguna de las computadoras muestreadas tengan el defecto?

81. Suponga que aviones pequeños llegan a cierto aeropuerto según un proceso de Poisson con tasa $\alpha = 8$ aviones por hora, de modo que el número de llegadas durante t horas es una variable aleatoria de Poisson con parámetro $\lambda = 8t$.

a. ¿Cuál es la probabilidad de que exactamente seis aviones pequeños lleguen durante una hora? ¿Por lo menos seis? ¿Por lo menos 10?

b. ¿Cuáles son el valor esperado y la desviación estándar del número de aviones pequeños que llegan durante un periodo de 90 minutos?

c. ¿Cuál es la probabilidad de que por lo menos lleguen 20 aviones pequeños durante 2 horas y media? ¿De que a lo sumo lleguen 10 durante este periodo?

82. La cantidad de personas que llegan a una sala de urgencias se puede modelar mediante un proceso de Poisson con un parámetro de número de llegadas de cinco por hora.

a. ¿Cuál es la probabilidad de que ocurran exactamente cuatro llegadas durante una determinada hora?

b. ¿Cuál es la probabilidad de que lleguen por lo menos cuatro personas durante cierta hora?

c. ¿Cuántas personas espera que lleguen durante un periodo de 45 minutos?

83. El número de solicitudes de auxilio que recibe un servicio de remolque de vehículos es un proceso de Poisson con tasa $\alpha = 4$ por hora.

a. Calcule la probabilidad de que se reciban exactamente diez solicitudes durante un periodo de dos horas.

b. Si los operadores de las grúas toman un receso de 30 minutos para almorzar, ¿cuál es la probabilidad de que no pierdan ninguna llamada de auxilio?

c. ¿Cuántas llamadas se esperarían durante el descanso?

84. En una prueba de tarjetas de circuitos, la probabilidad de que falle algún diodo es 0.01. Suponga que una tarjeta contiene 200 diodos.

a. ¿Cuántos diodos esperaría que fallen y cuál es la desviación estándar del número de diodos que se espera fallen?

b. ¿Cuál es la probabilidad (aproximada) de que por lo menos fallen cuatro diodos en una tarjeta seleccionada al azar?

c. Si se envían cinco tarjetas a un cliente, ¿cuál es la probabilidad de que por lo menos cuatro funcionen bien? (Una tarjeta trabaja bien sólo si todos sus diodos funcionan).

85. En el artículo "Reliability-Based Service-Life Assessment of Aging Concrete Structures" (*J. Structural Engr.*, 1993: 1600-1621) se indica que se puede emplear un proceso de Poisson para representar la ocurrencia de cargas estructurales con el tiempo. Suponga que el tiempo promedio entre ocurrencias de cargas es medio año.

a. ¿Cuántas cargas se espera que ocurran durante dos años?

b. ¿Cuál es la probabilidad de que ocurran más de cinco cargas durante dos años?

c. ¿Cuán largo debe ser un periodo para que la probabilidad de que no ocurra ninguna carga durante ese tiempo sea a lo sumo 0.1?

86. Sea que X tiene una distribución de Poisson con parámetro λ. Demuestre que $E(X) = \lambda$ directamente de la definición de valor esperado. (*Sugerencia*: el primer término de la suma es igual a 0 y después se puede cancelar x. Ahora factorice λ y demuestre que la suma del remanente es igual a 1.)

87. Suponga que los árboles de un bosque están distribuidos según un proceso de Poisson bidimensional con parámetro α, el número esperado de árboles por acre, igual a 80.

a. ¿Cuál es la probabilidad de que en una parcela de un cuarto de acre haya a lo sumo 16 árboles?

b. Si el bosque cubre 85 000 acres, ¿cuál es el número esperado de árboles en el bosque?

c. Suponga que se selecciona un punto del bosque y se construye un círculo cuyo radio es un décimo de milla. Sea X = número de árboles dentro de esa región circular. ¿Cuál es la fmp de X? (*Sugerencia*: una milla cuadrada = 640 acres.)

88. Llegan automóviles a una estación de inspección de equipo vehicular de acuerdo con un proceso de Poisson con tasa $\alpha = 10$ por hora. Suponga que con probabilidad 0.5 en un vehículo que llega no se encuentran violaciones de equipo.

a. ¿Cuál es la probabilidad de que exactamente lleguen diez automóviles durante la hora, y todos cumplan con los requerimientos de equipo?

b. Para cualquier y fija ≥ 10, ¿cuál es la probabilidad de que y automóviles lleguen durante la hora, de los cuales diez cumplen con los requerimientos de equipo?

c. ¿Cuál es la probabilidad de que diez automóviles que cumplen con los requerimientos lleguen durante la siguiente hora? [*Sugerencia*: sume las probabilidades del inciso (b) de $y = 10$ hasta ∞.]

89. **a.** En un proceso de Poisson, ¿qué tiene que pasar tanto en el intervalo de tiempo $(0, t)$ como en el $(t, t + \Delta t)$ para que no ocurran eventos en todo el intervalo $(0, t + \Delta t)$? Utilice esto y las suposiciones 1 a 3 para escribir una relación entre $P_0(t + \Delta t)$ y $P_0(t)$.

b. Utilice el resultado del inciso (a) y escriba una expresión para la diferencia $P_0(t + \Delta t) - P_0(t)$. Después, divida entre Δt y permita que $\Delta t \to 0$ para obtener una ecuación donde aparezca $(d/dt)P_0(t)$, la derivada de $P_0(t)$ con respecto a t.

c. Compruebe que $P_0(t) = e^{-\alpha t}$ satisface la ecuación del inciso (b).

d. Se puede demostrar de manera similar a los incisos (a) y (b) que las $P_k(t)$ deben satisfacer el sistema de ecuaciones diferenciales

$$\frac{d}{dt} P_k(t) = \alpha P_{k-1}(t) - \alpha P_k(t)$$

$$k = 1, 2, 3, \ldots$$

Compruebe que $P_k(t) = e^{-\alpha t}(\alpha t)^k/k!$ satisface el sistema. (En realidad ésta es la única solución.)

Ejercicios complementarios (90 a 116)

90. Considere una baraja que consiste en siete cartas, marcadas $1, 2, \ldots, 7$. Se eligen al azar tres de estas cartas. Defina una variable aleatoria W por W = suma de los números resultantes y calcule la fmp de W. Después, calcule μ y σ^2. [*Sugerencia:* considere los resultados sin orden, de modo que (1, 3, 7) y (3, 1, 7) no son diferentes. Entonces hay un total de 35 resultados que se pueden enumerar. (Este tipo de variable aleatoria en realidad surge en relación con una prueba de hipótesis llamada prueba de la suma de rangos de Wilcoxon, donde hay una muestra x y una muestra y, y W es la suma de los rangos de las x en la muestra combinada).]

91. Después de barajar 52 cartas, un jugador distribuye cinco. Sea X = número de palos representados en la mano de cinco cartas.

a. Demuestre que la fmp de X es

x	1	2	3	4
$p(x)$	0.002	0.146	0.588	0.264

[*Sugerencia:* $p(1) = 4P$(todas son picas), $p(2) = 6P$(sólo picas y corazones con por lo menos una de cada una), y $p(4) = 4P$(2 picas $\cap$ una de cada uno de los otros palos).]

b. Calcule μ, σ^2, y σ.

92. La variable aleatoria binomial negativa X se definió como el número de las F que preceden a la r-ésima S. Sea Y = número de intentos necesarios para obtener la r-ésima S. De la misma forma en que se obtuvo la fmp de X, obtenga la fmp de Y.

93. De todos los clientes que compran abridores automáticos para puerta de cochera, 75% compran un modelo accionado por cadena. Sea X = el número entre los siguientes 15 compradores que seleccionan un modelo accionado por cadena.

a. ¿Cuál es la fmp de X?

b. Calcule $P(X > 10)$.

c. Calcule $P(6 \le X \le 10)$.

d. Calcule μ y σ^2.

e. Si en la actualidad la tienda tiene en existencia 10 modelos accionados por cadena y ocho accionados por flecha, ¿cuál es la probabilidad de que se cumplan las solicitudes de estos 15 clientes con las existencias actuales?

94. Un amigo planeó en fechas recientes un viaje de campamento. Él tenía dos linternas de mano, una que requería una sola batería de 6V y otra que usaba dos baterías tamaño D. Había empacado con anticipación dos baterías de 6V y 4 tamaño D en su vehículo. Suponga que la probabilidad de que cualquier batería funcione es p y que las baterías funcionan o fallan de manera independiente. El amigo desea tomar sólo una linterna. ¿Para qué valores de p debe tomar la linterna de 6V?

95. Un sistema denominado k-*de*-n funciona si y sólo si por lo menos funcionan k de los n componentes del sistema. Si los componentes individuales funcionan sin depender entre sí, cada uno con probabilidad 0.9, ¿cuál es la probabilidad de que funcione un sistema 3 de 5?

96. Un fabricante de baterías para linterna desea controlar la calidad de su producto rechazando cualquier lote en el que la proporción de baterías con voltaje inaceptable parezca demasiado alto. Con este fin, de cada lote grande (10 000 baterías), se eligen 25 baterías y se prueban. Si por lo menos cinco de éstas generan un voltaje ina-

ceptable, se rechaza todo el lote. ¿Cuál es la probabilidad de que un lote sea rechazado si

a. 5% de las baterías del lote tienen voltajes inaceptables?

b. los voltajes son inadmisibles en 10% de las baterías del lote?

c. 20% de las baterías del lote tienen voltajes inaceptables?

d. ¿Qué sucedería con las probabilidades de los incisos (a) a (c) si el número crítico de rechazo aumentara de 5 a 6?

97. De las personas que pasan por un detector de metales en un aeropuerto, 0.5% lo activan; sea X = número entre un grupo de 500 individuos seleccionados al azar que activan el detector.

a. ¿Cuál es la fmp (aproximada) de X?

b. Calcule $P(X = 5)$.

c. Calcule $P(5 \leq X)$.

98. Una firma consultora educativa intenta determinar si los estudiantes de secundaria que nunca han utilizado una calculadora de mano pueden resolver de modo más fácil cierto tipo de problema con una calculadora que utiliza lógica polaca inversa, o con una que no emplea esta lógica. Se elige una muestra de 25 estudiantes y se les permite practicar con ambas calculadoras. Después, a cada estudiante se le pide que resuelva un problema con la calculadora polaca inversa y un problema similar con la otra. Sea $p = P(S)$, donde S indica que un estudiante resolvió más rápido el problema por medio de lógica polaca inversa que sin ésta, y sea X = número de éxitos (S).

a. Si $p = 0.5$, ¿cuál es $P(7 \leq X \leq 18)$?

b. Si $p = 0.8$, ¿cuál es $P(7 \leq X \leq 18)$?

c. Si se debe rechazar la afirmación $p = 0.5$ cuando $X \leq 7$ o $X \geq 18$, ¿cuál es la probabilidad de rechazar la afirmación cuando en realidad es correcta?

d. Si se toma la decisión de rechazar la afirmación $p = 0.5$ como en el inciso (c), ¿cuál es la probabilidad de que la afirmación no sea rechazada cuando $p = 0.6$? ¿Cuándo $p = 0.8$?

e. ¿Qué regla de decisión elegiría para rechazar la afirmación $p = 0.5$ si se desea que la probabilidad del inciso (c) sea a lo sumo 0.01?

99. Considere una enfermedad cuya presencia se puede detectar mediante una prueba de sangre. Sea p la probabilidad de que un individuo seleccionado al azar tenga la enfermedad. Suponga que se elige de forma independiente a n individuos para la prueba. Una forma de proceder es realizar una prueba separada en cada una de las n muestras de sangre. Un método con posibilidades más económicas, prueba de grupo, se introdujo durante la Segunda Guerra Mundial para identificar a varones sifilíticos entre los reclutas. Primero, se toma una parte de cada muestra de sangre, se combinan estos especímenes y se realiza una sola prueba. Si nadie tiene la enfermedad, el resultado será negativo y sólo se necesitó una prueba. Si por lo menos un individuo tiene la enfermedad, la prueba de la muestra combinada produce un resultado positivo, en cuyo caso se lleva a cabo cada una de las n pruebas. Si $p = 0.1$ y $n = 3$, ¿cuál es el número esperado de pruebas por medio de este procedimiento? ¿Cuál es el número esperado cuando $n = 5$? [En el artículo [The article "Random Multiple-Access Communication and Group Testing" (*IEEE Trans. on Commun.*, 1984: 769-774) se aplican estas ideas a un sistema de comunicación donde la dicotomía fue usuario activo/inactivo en lugar de persona enferma/no enferma].]

100. Sea p_1 la probabilidad de que cualquier símbolo de determinado código se transmita de manera errónea a través de un sistema de comunicaciones. Suponga que en diferentes símbolos, los errores ocurren sin depender entre sí. Suponga también que con probabilidad p_2 un símbolo erróneo se corrige al recibirlo. Sea X el número de símbolos correctos en un bloque de mensajes que consiste en n símbolos (después que terminó el proceso de corrección). ¿Cuál es la distribución de probabilidad de X?

101. El comprador de una unidad generadora de energía requiere c arranques consecutivos exitosos antes de aceptar la unidad. Suponga que los resultados de arranques individuales son independientes entre sí. Sea p la probabilidad de que cualquier arranque sea satisfactorio. La variable aleatoria de interés es X = número de arranques que se deben llevar a cabo antes de aceptar. Dé la fmp de X para el caso $c = 2$. If $p = 0.9$, ¿cuál es $P(X \leq 8)$? [*Sugerencia*: para $x \geq 5$, exprese $p(x)$ "en forma recursiva" en términos de la fmp evaluada en los valores más pequeños $x - 3, x - 4, \ldots, 2$.] (Este problema fue sugerido en el artículo "Evaluation of a Start-Up Demonstration Test", *J. Quality Technology*, 1983: 103-106.)

102. Una línea aérea desarrolló un plan para un club de viajeros ejecutivos bajo la premisa de que 10% de sus clientes actuales reunirían las condiciones para obtener la membresía.

a. Si se asume la validez de esta premisa, entre 25 clientes actuales seleccionados al azar, ¿cuál es la

probabilidad de que entre 2 y 6 (inclusive) reúnan las condiciones para obtener la membresía?

b. De nuevo, suponiendo la validez de la premisa, ¿cuál es el número esperado de clientes que reúnen las condiciones para obtener la membresía y la desviación estándar del número de quienes reúnen las condiciones en una muestra aleatoria de 100 clientes actuales?

c. Sea X el número de clientes actuales que reúnen las condiciones para conseguir la membresía en una muestra aleatoria de 25 clientes. Considere rechazar la premisa de la compañía a favor de la afirmación de que $p > 0.10$ si $x \geq 7$. ¿Cuál es la probabilidad de que la premisa de la compañía sea rechazada cuando en realidad es válida?

d. Refiérase a la regla de decisión presentada en el inciso (c). ¿Cuál es la probabilidad de que la premisa de la compañía no sea rechazada aun cuando $p = 0.20$ (es decir, 20% reúnen las condiciones)?

103. Cuarenta por ciento de las mazorcas de maíz (maíz actual) tienen espículas sencillas y el otro 60% espículas apareadas. Una semilla con espículas sencillas produce mazorcas con espículas sencillas 29% de las veces, mientras que una con espículas apareadas produce mazorcas con espículas sencillas 26% de las veces. Considere elegir al azar 10 semillas.

a. ¿Cuál es la probabilidad de que exactamente cinco de esas semillas tengan una sola espícula y produzcan mazorcas con una espícula sencilla?

b. ¿Cuál es la probabilidad de que exactamente cinco de las mazorcas producidas con esas semillas tengan espículas sencillas? ¿Cuál es la probabilidad de que a lo sumo cinco mazorcas tengan espículas sencillas?

104. Un juicio dio como resultado una decisión dividida de los jueces debido a que ocho estuvieron a favor de un veredicto de culpabilidad y cuatro votaron por la absolución. Si los miembros del jurado abandonan la sala en orden aleatorio y un reportero busca entrevistar a cada uno de los primeros cuatro, ¿cuál es la fmp de X = número de los miembros del jurado a favor de la absolución entre los entrevistados? ¿Cuántos de los que están a favor de la absolución se esperaría que fueran entrevistados?

105. Un servicio de reservación emplea a cinco operadoras que reciben solicitudes de información sin depender entre sí, cada una según un proceso de Poisson con tasa $\alpha = 2$ por minuto.

a. ¿Cuál es la probabilidad de que durante un periodo de un minuto la primera operadora no reciba solicitudes?

b. ¿Cuál es la probabilidad de que durante un periodo de un minuto exactamente cuatro de las cinco operadoras no reciban solicitudes?

c. Escriba una expresión para la probabilidad de que durante un minuto todas las operadoras reciban exactamente el mismo número de solicitudes.

106. En un gran campo están distribuidos al azar saltamontes según una distribución de Poisson con parámetro $\alpha = 2$ por yarda cuadrada. ¿Qué longitud debe tener el radio R de una región circular de muestreo para que la probabilidad de hallar por lo menos uno en la región sea igual a 0.99?

107. Un puesto de periódicos pidió cinco ejemplares de cierta edición de una revista de fotografía. Sea X = número de individuos que entran a comprar esta revista. Si X tiene una distribución de Poisson con parámetro $\lambda = 4$, ¿cuál es el número esperado de ejemplares que se venden?

108. Los individuos A y B comienzan a jugar una serie de juegos de ajedrez. Sea $S = \{\text{A gana un juego}\}$, y suponga que los resultados de los juegos sucesivos son independientes con $P(S) = p$ y $P(F) = 1 - p$ (nunca empatan). Jugarán hasta que uno de ellos gane diez juegos. Sea X = número de juegos jugados (con valores posibles 10, 11, . . . , 19).

a. Para $x = 10, 11, \ldots, 19$, obtenga una expresión para $p(x) = P(X = x)$.

b. Si es posible un empate, con $p = P(S)$, $q = P(F)$, $1 - p - q = P(\text{empate})$, ¿cuáles son los valores posibles de X? ¿Cuál es $P(20 \leq X)$? [*Sugerencia:* $P(20 \leq X) = 1 - P(X < 20)$.]

109. Una prueba para detectar cierta enfermedad tiene probabilidad 0.20 de dar una lectura positiva-falsa (lo cual indica que un individuo tiene la enfermedad cuando no es el caso) y probabilidad 0.10 de dar un resultado negativo-falso. Suponga que se aplica la prueba a diez individuos, de los cuales cinco tienen la enfermedad y cinco no. Sea X = número de lecturas positivas resultantes.

a. ¿Tiene X una distribución binomial? Explique su razonamiento.

b. ¿Cuál es la probabilidad de que exactamente tres de los diez resultados de prueba sean positivos?

110. La fmp binomial negativa generalizada está dada por

$$nb(x; r, p) = k(r, x) \cdot p^r(1 - p)^x$$

$$x = 5\ 0, 1, 2, \ldots$$

Sea que X, el número de plantas de cierta especie localizadas en determinada región, tiene esta distribución con $p = 0.3$ y $r = 2.5$. ¿Cuál es $P(X = 4)$? ¿Cuál es la probabilidad de que por lo menos se encuentre una planta?

111. Defina una función $p(x; \lambda, \mu)$ mediante

$$p(x; \lambda, \mu) = \begin{cases} \frac{1}{2}e^{-\lambda}\frac{\lambda^x}{x!} + \frac{1}{2}e^{-\mu}\frac{\mu^x}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

a. Demuestre que $p(x; \lambda, \mu)$ satisface las dos condiciones necesarias para especificar una fmp. [*Nota*: si una empresa emplea a dos mecanógrafas, de las cuales una comete errores tipográficos a razón de λ por página y la otra a razón de μ por página, y cada una mecanografía la mitad de los trabajos de la empresa, entonces $p(x; \lambda, \mu)$ es la fmp de X = número de errores en una página seleccionada al azar.]

b. Si la primera mecanógrafa (tasa λ) mecanografía 60% de las páginas, ¿cuál es la fmp de X del inciso (a)?

c. ¿Cuál es $E(X)$ para $p(x; \lambda, \mu)$ dada por la expresión mostrada antes?

d. ¿Cuál es σ^2 para $p(x; \lambda, \mu)$ dada por la expresión?

112. La *moda* de una variable aleatoria discreta X con fmp $p(x)$ es el valor x^* para el cual $p(x)$ es máxima (el valor más probable de x).

a. Sea $X \sim \text{Bin}(n, p)$. Considere la relación $b(x + 1; n, p)/b(x; n, p)$, y demuestre que $b(x; n, p)$ aumenta con x siempre que $x < np - (1 - p)$. Concluya que la moda x^* es el entero que satisface $(n + 1)p - 1 \le x^* \le (n + 1)p$.

b. Demuestre que si X tiene una distribución de Poisson con parámetro λ, la moda es el entero más grande menor que λ. Si λ es un entero, demuestre que tanto $\lambda - 1$ como λ son modas.

113. Un dispositivo de almacenamiento de disco de computadora tiene diez pistas concéntricas, numeradas 1, 2, . . . , 10 de la más alejada a la interior y un solo brazo de acceso. Sea p_i = la probabilidad de que cualquier petición de datos lleve el brazo a la pista i ($i = 1, \ldots, 10$). Suponga que las pistas a las cuales se tiene acceso en búsquedas sucesivas son independientes. Sea X = número de pistas sobre las que el brazo de acceso pasa durante dos peticiones sucesivas (excluyendo la pista que acaba de dejar el brazo, así que los valores posibles de X son $x = 0, 1, \ldots, 9$). Calcule la fmp de X. [*Sugerencia:* P(el brazo ahora está en la pista i y $X = j$) = $P(X = j\Omega$el brazo ahora está en la pista $i) \cdot p_i$. Después de escribir la probabilidad condicional en términos de $p_1, \ldots, p_{10}$, por la ley de probabilidad total, la probabilidad deseada se obtiene al sumar sobre i.]

114. Si X es una variable aleatoria hipergeométrica, demuestre directamente de la definición que $E(X) = nM/N$ (considere sólo el caso $n < M$). [*Sugerencia*: factorice nM/N de la suma para $E(X)$, y demuestre que los términos dentro de la suma son de la forma $h(y; n - 1, M - 1, N - 1)$, donde $y = x - 1$.]

115. Utilice el hecho de que

$$\sum_{\text{toda } x} (x - \mu)^2 p(x) \ge \sum_{x:\,|x-\mu| \ge k\sigma} (x - \mu)^2 p(x)$$

para toda x para demostrar la desigualdad de Chebyshev dada en el ejercicio 43.

116. El proceso simple de Poisson de la sección 3.6 se caracteriza por una tasa constante α a la que ocurren los eventos por unidad de tiempo. Una generalización de esto es suponer que la probabilidad de que exactamente ocurra un evento en el intervalo $[t, t + \Delta t]$ es $\alpha(t) \cdot \Delta t + o(\Delta t)$. Entonces, se puede demostrar que el número de eventos que ocurren durante un intervalo $[t_1, t_2]$ tiene una distribución de Poisson con parámetro

$$\lambda = \int_{t_1}^{t_2} \alpha(t)\, dt$$

La ocurrencia de eventos con el tiempo en esta situación recibe el nombre de *proceso de Poisson no homogéneo*. En el artículo "Inference Based on Retrospective Ascertainment". *J. Amer. Stat. Assoc.*, 1989: 360-372, se considera que la función de intensidad

$$\alpha(t) = e^{a+bt}$$

es apropiada para eventos en los que interviene la transmisión del VIH (el virus del SIDA) vía transfusiones de sangre. Suponga que $a = 2$ y $b = 0.6$ (valores cercanos a los sugeridos en el artículo), con el tiempo en años.

a. ¿Cuál es el número esperado de eventos en el intervalo [0, 4]? ¿En [2, 6]?

b. ¿Cuál es la probabilidad de que a lo sumo ocurran 15 eventos en el intervalo [0, 0.9907]?

Bibliografía

Johnson, Norman, Samuel Kotz y Adrienne Kemp, *Discrete Univariate Distributions,* Wiley, Nueva York, 1992. Una enciclopedia de información acerca de distribuciones discretas.

Olkin, Ingram, Cyrus Derman y Leon Gleser, *Probability Models and Applications* (2a. ed.), Macmillan, Nueva York, 1994. Contiene un análisis profundo de propiedades generales de distribuciones discretas y continuas y resultados para distribuciones específicas.

Ross, Sheldon, *Introduction to Probability Models* (7a. ed.), Academic Press, Nueva York, 2003. Una buena fuente de material acerca del proceso de Poisson y generalizaciones y una introducción agradable a otros temas de probabilidad aplicada.

4

Variables aleatorias continuas y distribuciones de probabilidad

Introducción

Como se mencionó al comienzo del capítulo 3, no todas las variables aleatorias son discretas. En este capítulo, se estudia el segundo tipo general de variable aleatoria que surge en muchos problemas aplicados. En las secciones 4.1 y 4.2, se presentan definiciones y propiedades básicas de variables aleatorias continuas y sus distribuciones de probabilidad. En la sección 4.3, se estudia en detalle la variable aleatoria normal y la distribución, sin lugar a dudas, la más importante y útil en probabilidad y estadística. En las secciones 4.4 y 4.5, se analizan algunas otras distribuciones continuas que suelen usarse en el trabajo aplicado. En la sección 4.6, se introduce un método para evaluar si determinados datos muestrales son consistentes con una distribución específica.

4.1 Variables aleatorias continuas y funciones de densidad de probabilidad

Una variable aleatoria (*va*) discreta es una cuyos valores posibles constituyen un conjunto finito, o bien, se pueden listar en una secuencia infinita (una lista en la que hay un primer elemento, un segundo elemento, etc.). Una variable aleatoria cuyo conjunto de valores posibles es un intervalo completo de números, no es discreta.

Variables aleatorias continuas

DEFINICIÓN

Se dice que una variable aleatoria X es **continua** si su conjunto de valores posibles es un intervalo completo de números; es decir, si para alguna $A < B$, cualquier número x entre A y B es posible.

Ejemplo 4.1 Si en el estudio de la ecología de un lago, se hacen mediciones profundas en lugares elegidos al azar, entonces X = profundidad en algún lugar es una *va* continua. Aquí A es la profundidad mínima en la región que está siendo muestreada, y B es la profundidad máxima. ∎

Ejemplo 4.2 Si se elige al azar un compuesto químico y se determina su pH X, entonces X es una variable aleatoria continua, porque es posible cualquier valor de pH entre 0 y 14. Si se sabe más acerca del compuesto elegido para el análisis, entonces el conjunto de valores posibles podría ser un subintervalo de [0, 14], por ejemplo $5.5 \leq x \leq 6.5$, pero X aún es continua. ∎

Si es posible subdividir la escala de medición de X a cualquier grado deseado, entonces la variable es continua; si no se puede, la variable es discreta. Por ejemplo, si la variable es la altura o la longitud, entonces ésta se puede medir en kilómetros, metros, centímetros, milímetros, etcétera, de modo que la variable es continua. Sin embargo, si X = la cantidad facturada en una declaración mensual de gas natural elegida al azar, entonces la unidad más pequeña de medición es centavos, así que cualquier valor de X es un múltiplo de \$0.01 y X es una variable discreta.

Se podría argumentar que aunque, en principio, variables como la altura, peso y temperaturas son continuas, en la práctica las limitaciones de los instrumentos de medición restringen a un mundo discreto (aunque a veces subdividido de manera muy fina). No obstante, los modelos continuos suelen aproximar muy bien situaciones de mundo real, y suele ser más fácil trabajar con las matemáticas continuas (el cálculo) que con las matemáticas de variables y distribuciones discretas.

Distribuciones de probabilidad para variables continuas

Suponga que la variable X de interés es la profundidad de un lago en un punto elegido al azar en la superficie. Sea M = la profundidad máxima (en metros), de modo que cualquier número en el intervalo $[0, M]$ es un valor posible de X. Si se hace discreta a X al medir la profundidad hasta el metro más próximo, entonces los valores posibles son enteros no negativos menores que o iguales a M. La distribución discreta resultante de la profundidad se puede ilus-

trar por medio de un histograma de probabilidad. Si se dibuja el histograma de modo que el área del rectángulo arriba de cualquier entero posible k sea la proporción del lago cuya profundidad es (hasta el metro más próximo) k, entonces el área total de los rectángulos es 1. En la figura 4.1(a) se ilustra un histograma posible.

Si se mide la profundidad hasta el centímetro más próximo y se usa el mismo eje de medición que en la figura 4.1(a), cada rectángulo del histograma de probabilidad resultante es mucho más estrecho, aunque el área total de los rectángulos aún es 1. Un histograma posible se ilustra en la figura 4.1(b); éste tiene una apariencia mucho más uniforme que el histograma de la figura 4.1(a). Si se continúa de esta forma para medir la profundidad de manera cada vez más fina, la secuencia resultante de histogramas se aproxima a una curva lisa, como la que se ilustra en la figura 4.1(c). Debido a que para cada histograma el área total de los rectángulos es igual a 1, el área total bajo la curva lisa también es 1. La probabilidad de que la profundidad en un punto elegido al azar esté entre a y b es justo el área bajo la curva lisa entre a y b. Es exactamente una curva lisa del tipo ilustrado en la figura 4.1(c) la que especifica una distribución de probabilidad continua.

Figura 4.1 (a) Histograma de probabilidad de la profundidad medida hasta el metro más cercano; (b) histograma de probabilidad de la profundidad medida hasta el centímetro más próximo; (c) un límite de un secuencia de histogramas discretos

DEFINICIÓN

Sea X una *va* continua. Entonces una **distribución de probabilidad** o **función de densidad de probabilidad** (fdp) de X es una función $f(x)$ tal que para dos números cualesquiera a y b con $a \le b$,

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

Es decir, la probabilidad de que X tome un valor en el intervalo $[a, b]$ es el área arriba de este intervalo y bajo la gráfica de la función de densidad, como se ilustró en la figura 4.2. La gráfica de $f(x)$ se llama *curva de densidad.*

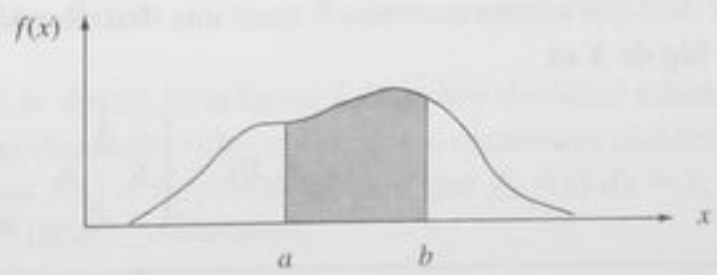


Figura 4.2 $P(a \le X \le b)$ = área bajo la curva de densidad entre a y b

Para que $f(x)$ sea una fdp legítima, se deben satisfacer las dos condiciones siguientes:

1. $f(x) \geq 0$ para toda x
2. $\int_{-\infty}^{\infty} f(x)\,dx$ = área bajo la gráfica completa de $f(x)$ $= 1$

Ejemplo 4.3 La dirección de una imperfección con respecto a una línea de referencia en un objeto circular como un neumático, el rotor del freno o el volante está, en general, sujeta a incertidumbre. Considere la línea de referencia que une el vástago de la válvula en una llanta con el punto central, y sea X el ángulo medido en el sentido de las manecillas del reloj hasta el lugar de una imperfección. Una fdp posible para X es

$$f(x) = \begin{cases} \dfrac{1}{360} & 0 \leq x < 360 \\ 0 & \text{en caso contrario} \end{cases}$$

La fdp se dibuja en la figura 4.3. Resulta claro que $f(x) \geq 0$. El área bajo la curva de densidad es justo el área de un rectángulo: (altura)(base) $= (\frac{1}{360})(360) = 1$. La probabilidad de que el ángulo esté entre 90° y 180° es

$$P(90 \leq X \leq 180) = \int_{90}^{180} \frac{1}{360}\,dx = \frac{x}{360}\Big|_{x=90}^{x=180} = \frac{1}{4} = 0.25$$

La probabilidad de que el ángulo de ocurrencia esté dentro de 90° de la línea de referencia es

$$P(0 \leq X \leq 90) + P(270 \leq X < 360) = 0.025 + 0.025 = 0.50$$

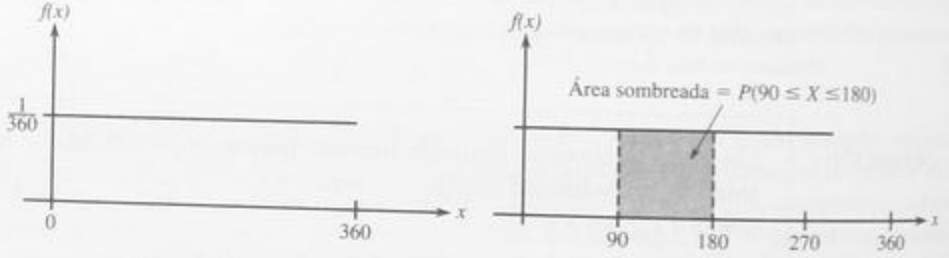


Figura 4.3 La fdp y la probabilidad a partir del ejemplo 4.3 ■

Debido a que siempre que $0 \leq a \leq b \leq 360$ en el ejemplo 4.3, $P(a \leq X \leq b)$ depende sólo de la amplitud $b - a$ del intervalo, se dice que X tiene una distribución uniforme.

DEFINICIÓN

Se dice que una *va* continua X tiene una **distribución uniforme** en el intervalo $[A, B]$ si la fdp de X es

$$f(x; A, B) = \begin{cases} \dfrac{1}{B - A} & A \leq x \leq B \\ 0 & \text{en caso contrario} \end{cases}$$

La gráfica de cualquier fdp se asemeja a la gráfica de la figura 4.3, excepto que el intervalo de densidad positiva es $[A, B]$ y no $[0, 360]$.

En el caso discreto, una función masa de probabilidad (fmp) indica cómo están distribuidas pequeñas "manchas" de masa de probabilidad de varias magnitudes a lo largo del eje de medición. En el caso continuo, la densidad de probabilidad se "dispersa" de un modo continuo a lo largo del intervalo de valores posibles. Cuando la densidad se dispersa de manera uniforme en el intervalo, se obtiene una fdp uniforme, como en la figura 4.3.

Cuando X es una *va* discreta, a cada valor posible se le asigna probabilidad positiva, pero éste no es el caso cuando X es continua.

PROPOSICIÓN

Si X es una *va* continua, entonces para cualquier número c, $P(X = c) = 0$. Además, para dos números cualesquiera a y b con $a < b$,

$$\begin{aligned} P(a \le X \le b) &= P(a < X \le b) \\ &= P(a \le X < b) = P(a < X < b) \end{aligned} \tag{4.1}$$

En palabras, la probabilidad asignada a algún valor en particular es cero, y la probabilidad de un intervalo no depende de si está incluido cualquiera de sus puntos terminales. Estas propiedades se deducen de los hechos de que el área bajo la gráfica de $f(x)$ y arriba del valor único c es cero y de que la exclusión o inclusión de los puntos terminales del intervalo no afectan el área bajo la gráfica arriba de un intervalo. Si X es discreta y hay una masa de probabilidad positiva tanto en $X = a$ como en $X = b$, entonces las cuatro probabilidades de la expresión (4.1) diferirán entre sí.

El hecho de que una distribución continua asigna probabilidad cero a cada valor tiene una analogía física. Considere una varilla circular sólida con área de sección transversal = 1 pulgada^2. Coloque la varilla a lo largo de un eje de medición y suponga que la densidad de la varilla en cualquier punto x está dada por el valor $f(x)$ de una función de densidad. Entonces, si se corta la varilla en los puntos a y b y se elimina este segmento, la cantidad de masa eliminada es $\int_a^b f(x)\,dx$; si se corta la varilla justo en el punto c, no se elimina ninguna masa. La masa se asigna a los segmentos de intervalo de la varilla, pero no a puntos individuales.

Ejemplo 4.4

"Avance del tiempo" en flujo de tránsito es el tiempo transcurrido entre el tiempo en que un automóvil termina de pasar un punto fijo y el instante en que el siguiente automóvil comienza a pasar ese punto. Sea X = el avance del tiempo para dos automóviles consecutivos elegidos al azar en una autopista durante un periodo de flujo intenso. La siguiente fdp de X es en esencia la que se indica en "The Statistical Properties of Freeway Traffic" (*Transp. Research*, vol. 11: 221-228):

$$f(x) = \begin{cases} 0.15e^{-0.15(x-0.5)} & x \ge 0.5 \\ 0 & \text{en caso contrario} \end{cases}$$

La gráfica de $f(x)$ se ilustra en la figura 4.4; no hay densidad relacionada con los tiempos de avance menores que 0.5 y la densidad del progreso disminuye rápidamente como x aumenta de 0.5. Resulta claro que $f(x) \ge 0$; para demostrar que $\int_{-\infty}^{\infty} f(x)\,dx = 1$, se utiliza el resultado de cálculo $\int_a^{\infty} e^{-kx}\,dx = (1/k)e^{-k\cdot a}$. Entonces,

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \int_{0.5}^{\infty} 0.15e^{-0.15(x-0.5)}\,dx = 0.15e^{.075}\int_{0.5}^{\infty} e^{-0.15x}\,dx \\ &= 0.15e^{0.075}\cdot\frac{1}{0.15}e^{-(0.15)(0.5)} = 1 \end{aligned}$$

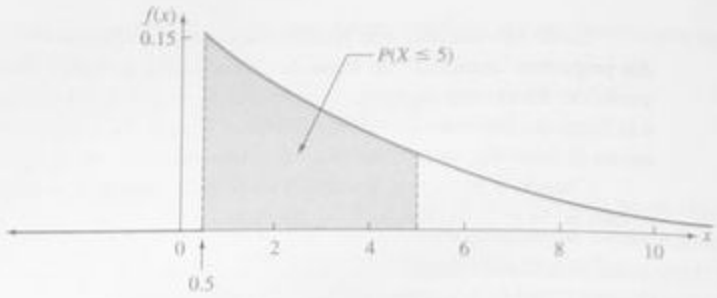


Figura 4.4 Curva de densidad para el tiempo de avance del ejemplo 4.4

La probabilidad de que el tiempo de avance sea a lo sumo 5 s es

$$\begin{aligned} P(X \le 5) &= \int_{-\infty}^{5} f(x)\,dx = \int_{0.5}^{5} 0.15e^{-0.15(x-0.5)}\,dx \\ &= 0.15e^{0.075}\int_{0.5}^{5} e^{-0.15x}\,dx = 0.15e^{0.075} \cdot \frac{1}{0.15} e^{-0.15x}\Big|_{x=0.5}^{x=5} \\ &= e^{0.075}(-e^{-0.75} + e^{-0.075}) = 1.078(-0.472 + 0.928) = 0.491 \\ &= P(\text{menos de 5 s}) = P(X < 5) \end{aligned}$$

■

A diferencia de las distribuciones discretas como la binomial, hipergeométrica y la binomial negativa, la distribución de alguna *va* continua por lo general no se deduce mediante argumentos probabilísticos simples. En vez de eso, se debe hacer una elección juiciosa de la fdp con base en el conocimiento previo y datos disponibles. Por fortuna, hay algunas familias generales de fdp que se ajustan bien a una amplia variedad de situaciones experimentales; varias de éstas se analizan después en el capítulo.

Al igual que en el caso discreto, suele ser útil considerar que la población de interés consiste en valores de X y no en individuos u objetos. Entonces, la fdp es un modelo para la distribución de valores en esta población numérica, y a partir de este modelo es posible calcular varias características de la población (p. ej., la media).

Ejercicios | Sección 4.1 (1 a 10)

1. Sea X la cantidad de tiempo durante el cual un estudiante seleccionado al azar saca en préstamo un libro de la reserva de dos horas en una biblioteca universitaria, y suponga que X tiene la función de densidad

$$f(x) = \begin{cases} 0.5x & 0 \le x \le 2 \\ 0 & \text{en caso contrario} \end{cases}$$

Calcule las probabilidades siguientes:

a. $P(X \le 1)$
b. $P(0.5 \le X \le 1.5)$
c. $P(1.5 < X)$

2. Suponga que la temperatura de reacción X (en °C) de cierto proceso químico tiene una distribución uniforme con $A = -5$ y $B = 5$.

a. Calcule $P(X < 0)$.
b. Calcule $P(-2.5 < X < 2.5)$.
c. Calcule $P(-2 \le X \le 3)$.
d. Para k que satisface $-5 < k < k + 4 < 5$, calcule $P(k < X < k + 4)$.

3. Suponga que el error que se comete al hacer cierta medición es una *va* continua X con fdp

$$f(x) = \begin{cases} 0.09375(4 - x^2) & -2 \le x \le 2 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Trace la gráfica de $f(x)$.
b. Calcule $P(X > 0)$.
c. Calcule $P(-1 < X < 1)$.
d. Calcule $P(X < -0.5 \text{ o } X > 0.5)$.

4. Sea X el esfuerzo vibratorio (psi) en la paleta de una turbina de viento a una determinada velocidad en un túnel de viento. En el artículo "Blade Fatigue Life Assessment with Application to VAWTS" (*J. Solar Energy Engr.*, 1982: 107-111) se propone la distribución de Rayleigh, con fdp

$$f(x;\theta)=\begin{cases}\dfrac{x}{\theta^2}\cdot e^{-x^2/(2\theta^2)} & x>0\\ 0 & \text{en caso contrario}\end{cases}$$

como modelo para la distribución de X.

a. Compruebe que $f(x;\theta)$ es una fdp legítima.

b. Suponga que $\theta=100$ (un valor sugerido en una gráfica del artículo). ¿Cuál es la probabilidad de que X sea a lo sumo 200? ¿Menor de 200? ¿Por lo menos 200?

c. ¿Cuál es la probabilidad de que X esté entre 100 y 200 (de nuevo suponga que $\theta=100$)?

d. Dé una expresión para $P(X\le x)$.

5. Un maestro de una universidad nunca termina su clase antes que termine la hora y siempre termina su clase dentro de dos minutos después de la hora. Sea $X=$ tiempo que transcurre entre el final de la hora y el fin de la clase, y suponga que la fdp de X es

$$f(x)=\begin{cases}kx^2 & 0\le x\le 2\\ 0 & \text{en caso contrario}\end{cases}$$

a. Encuentre el valor de k. [*Sugerencia*: el área total bajo la gráfica de $f(x)$ es 1.]

b. ¿Cuál es la probabilidad de que la clase termine dentro de un minuto después que termina la hora?

c. ¿Cuál es la probabilidad de que la clase continúe más allá de la hora entre 60 y 90 s?

d. ¿Cuál es la probabilidad de que la clase continúe durante por lo menos 90 s más allá del final de la hora?

6. El peso de arrastre real de una pastilla estéreo que se fija en 3 g en un tocadiscos se puede considerar como una *va* continua X con fdp

$$f(x)=\begin{cases}k[1-(x-3)^2] & 2\le x\le 4\\ 0 & \text{en caso contrario}\end{cases}$$

a. Dibuje la gráfica de $f(x)$.

b. Encuentre el valor de k.

c. ¿Cuál es la probabilidad de que el peso de arrastre real sea mayor que el peso establecido?

d. ¿Cuál es la probabilidad de que el peso real esté dentro de 0.25 g del peso prescrito?

e. Cuál es la probabilidad de que el peso real difiera del peso prescrito en más de 0.5 g?

7. Se cree que el tiempo X (minutos) para que un asistente de laboratorio prepare el equipo para cierto experimento tiene una distribución uniforme con $A=25$ y $B=35$.

a. Escriba la fdp de X y trace su gráfica.

b. ¿Cuál es la probabilidad de que el tiempo de preparación sea mayor que 33 minutos?

c. ¿Cuál es la probabilidad de que el tiempo de preparación esté dentro de dos minutos del tiempo promedio? [*Sugerencia*: identifique a m en la gráfica de $f(x)$.]

d. Para cualquier a tal que $25<a<a+2<35$, ¿cuál es la probabilidad de que el tiempo de preparación esté entre a y $a+2$ minutos?

8. Para ir al trabajo una persona primero debe abordar un autobús cerca de casa y después transbordar otro. Si el tiempo de espera (en minutos) en cada parada tiene una distribución uniforme con $A=0$ y $B=5$, entonces se puede demostrar que el tiempo total de espera Y tiene la fdp

$$f(y)=\begin{cases}\dfrac{1}{25}y & 0\le y<5\\ \dfrac{2}{5}-\dfrac{1}{25}y & 5\le y\le 10\\ 0 & y<0 \text{ o bien } y>10\end{cases}$$

a. Trace una gráfica de la fdp de Y.

b. Verifique que $\int_{-\infty}^{\infty} f(y)\,dy=1$.

c. ¿Cuál es la probabilidad de que el tiempo total de espera sea a lo sumo de 3 minutos?

d. ¿Cuál es la probabilidad de que el tiempo total de espera sea a lo sumo de 8 minutos?

e. ¿Cuál es la probabilidad de que el tiempo total de espera esté entre 3 y 8 minutos?

f. ¿Cuál es la probabilidad de que el tiempo total de espera sea menos de 2 minutos o más de 6 minutos?

9. Considere otra vez la fdp de $X=$ tiempo de avance proporcionado en el ejemplo 4.4. ¿Cuál es la probabilidad de que el tiempo de avance

a. sean a lo sumo 6 s?

b. sean más de 6 s? ¿Por lo menos 6 s?

c. esté entre 5 y 6 s?

10. Una familia de fdp que ha sido utilizada para aproximar la distribución del ingreso, la del tamaño de la población de la ciudad y la del tamaño de empresas es la familia de Pareto. La familia tiene dos parámetros, k y θ, ambos mayores que 0, y la fdp es

$$f(x;k,\theta)=\begin{cases}\dfrac{k\cdot\theta^k}{x^{k+1}} & x\ge\theta\\ 0 & x<\theta\end{cases}$$

a. Trace la gráfica de $f(x;k,\theta)$.

b. Verifique que el área total bajo la gráfica sea igual a 1.

c. Si la *va* X tiene fdp $f(x;k,\theta)$, para cualquier valor fijo $b>\theta$, obtenga una expresión para $P(X\le b)$.

d. Para $\theta<a<b$, obtenga una expresión para la probabilidad $P(a\le X\le b)$.

4.2 Funciones de distribución acumulada y valores esperados

Varios de los conceptos importantes introducidos en el estudio de distribuciones discretas, también desempeñan un papel importante para las distribuciones continuas. Las definiciones análogas a las del capítulo 3, requieren sustituir la suma por la integración.

Función de distribución acumulada

La función de distribución acumulada (fda) $F(x)$ para una *va* discreta X da, para algún número específico x, la probabilidad $P(X \le x)$. Se obtiene al sumar la fmp $p(y)$ sobre los valores posibles de y que satisfacen $y \le x$. La fda de una *va* continua da las mismas probabilidades $P(X \le x)$ y se obtiene al integrar la fdp $f(y)$ entre los límites $-\infty$ y x.

DEFINICIÓN

La **función de distribución acumulada** $F(x)$ para una *va* continua X está definida para todo número x mediante

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)\,dy$$

Para cada x, $F(x)$ es el área bajo la curva de densidad a la izquierda de x. Esto se ilustra en la figura 4.5, donde $F(x)$ se incrementa de manera uniforme cuando aumenta x.

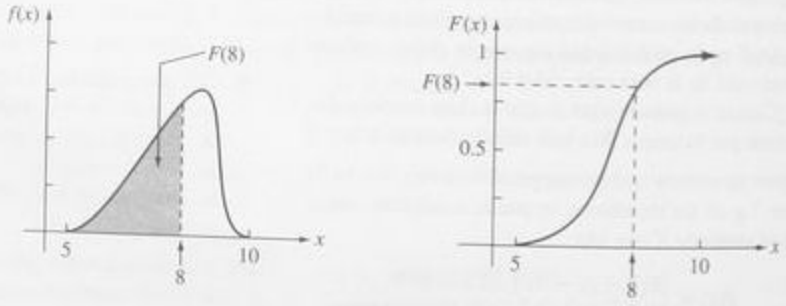


Figura 4.5 Una fdp y la fda relacionada

Ejemplo 4.5 Sea que X, el espesor de cierta placa de metal, tiene una distribución uniforme en $[A, B]$. La función de densidad se muestra en la figura 4.6. Para $x < A$, $F(x) = 0$, puesto que no hay área bajo la gráfica de la función de densidad a la izquierda de tal x. Para $x \ge B$, $F(x) = 1$, puesto que toda el área se acumula a la izquierda de tal x. Por último, para $A \le x \le B$,

$$F(x) = \int_{-\infty}^{x} f(y)\,dy = \int_{A}^{x} \frac{1}{B-A}\,dy = \frac{1}{B-A} \cdot y \Big|_{y=A}^{y=x} = \frac{x-A}{B-A}$$

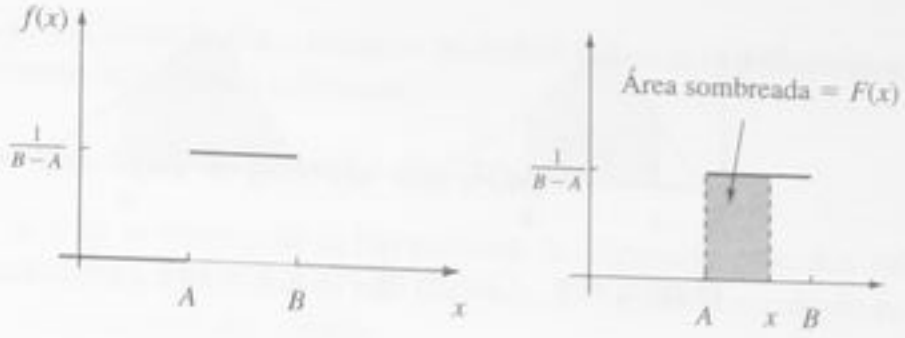


Figura 4.6 Función de densidad de probabilidad para una distribución uniforme

La fda completa es

$$F(x) = \begin{cases} 0 & x < A \\ \dfrac{x - A}{B - A} & A \le x < B \\ 1 & x \ge B \end{cases}$$

La gráfica de esta fda se ilustra en la figura 4.7.

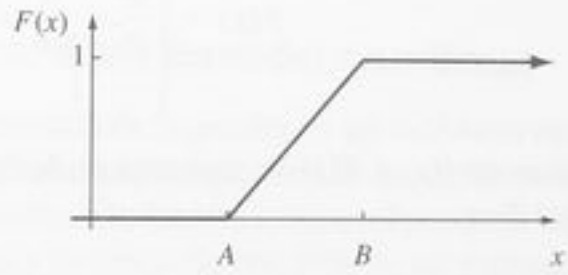


Figura 4.7 Función de densidad acumulada para una distribución uniforme ■

Cómo usar *F*(*x*) para calcular probabilidades

Aquí, la importancia de la fda, al igual que para las variables aleatorias discretas, es que las probabilidades de varios intervalos se pueden calcular a partir de una fórmula para $F(x)$ o de una tabla.

PROPOSICIÓN

Sea X una *va* aleatoria continua con fdp $f(x)$ y fda $F(x)$. Entonces para cualquier número a,

$$P(X > a) = 1 - F(a)$$

y para dos números cualesquiera a y b con $a < b$,

$$P(a \le X \le b) = F(b) - F(a)$$

En la figura 4.8 se ilustra la segunda parte de esta proposición; la probabilidad deseada es el área bajo la curva de densidad entre a y b, y es igual a la diferencia entre las dos áreas acumuladas sombreadas.

Figura 4.8 Cálculo de $P(a \le X \le b)$ a partir de probabilidades acumuladas

Ejemplo 4.6 Suponga que la fdp de la magnitud X de una carga dinámica en un puente (en newtons) está dada por

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \le x \le 2 \\ 0 & \text{en caso contrario} \end{cases}$$

Para cualquier número x entre 0 y 2,

$$F(x) = \int_{-\infty}^{x} f(y)\,dy = \int_{0}^{x} \left(\frac{1}{8} + \frac{3}{8}y\right) dy = \frac{x}{8} + \frac{3}{16}x^2$$

Por lo tanto,

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{8} + \frac{3}{16}x^2 & 0 \le x \le 2 \\ 1 & 2 < x \end{cases}$$

Las gráficas de $f(x)$ y $F(x)$ se muestran en la figura 4.9. La probabilidad de que la carga esté entre 1 y 1.5 es

$$\begin{aligned} P(1 \le X \le 1.5) &= F(1.5) - F(1) \\ &= \left[\frac{1}{8}(1.5) + \frac{3}{16}(1.5)^2\right] - \left[\frac{1}{8}(1) + \frac{3}{16}(1)^2\right] \\ &= \frac{19}{64} = 0.297 \end{aligned}$$

La probabilidad de que la carga sea mayor que 1 es

$$\begin{aligned} P(X > 1) &= 1 - P(X \le 1) = 1 - F(1) = 1 - \left[\frac{1}{8}(1) + \frac{3}{16}(1)^2\right] \\ &= \frac{11}{16} = 0.688 \end{aligned}$$

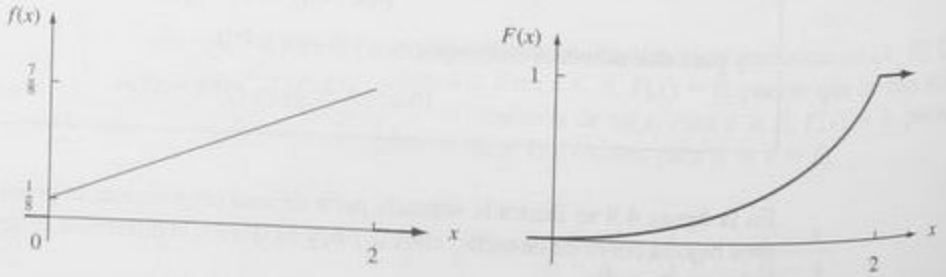


Figura 4.9 La fdp y la fda para el ejemplo 4.6

Una vez que se tiene la fda, cualquier probabilidad en la que interviene X se calcula fácilmente sin ninguna integración adicional.

Obtención de $f(x)$ a partir de $F(x)$

Para X discreta, la fmp se obtiene de la fda tomando la diferencia entre dos valores de $F(x)$. La analogía continua de una diferencia es una derivada. El siguiente resultado es una consecuencia del Teorema fundamental del cálculo.

PROPOSICIÓN

Si X es una variable aleatoria continua con fdp $f(x)$ y fda $F(x)$, entonces en toda x en la derivada $F'(x)$ existen, $F'(x) = f(x)$.

Ejemplo 4.7 (continuación del ejemplo 4.5) Cuando X tiene una distribución uniforme, $F(x)$ es diferenciable excepto en $x = A$ y $x = B$, donde la gráfica de $F(x)$ tiene esquinas ahusadas. Puesto que $F(x) = 0$ para $x < A$ y $F(x) = 1$ para $x > B$, $F'(x) = 0 = f(x)$ para tal x. Para $A < x < B$,

$$F'(x) = \frac{d}{dx}\left(\frac{x - A}{B - A}\right) = \frac{1}{B - A} = f(x)$$

■

Percentiles de una distribución continua

Cuando se dice que la puntuación de la prueba de un individuo estuvo en el percentil 85 de la población, se indica que 85% de las puntuaciones de la población estuvieron abajo de esa puntuación y 15% estuvieron arriba. De manera similar, el percentil 40 es la puntuación que excede 40% de las puntuaciones y es superado por 60% de las puntuaciones.

DEFINICIÓN

Sea p un número entre 0 y 1. El **percentil (100p)** de la distribución de una *va* continua X, denotado por $\eta(p)$, está definido por

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)\, dy \tag{4.2}$$

De acuerdo con la expresión (4.2), $\eta(p)$ es el valor en la escala de medición tal que $100p\%$ del área bajo la gráfica de $f(x)$ queda a la izquierda de $\eta(p)$ y $100(1 - p)\%$ queda a la derecha. Así, $\eta(0.75)$, el percentil 75, es tal que el área bajo la gráfica de $f(x)$ a la izquierda de $\eta(0.75)$ es 0.75. En la figura 4.10 se ilustra la definición.

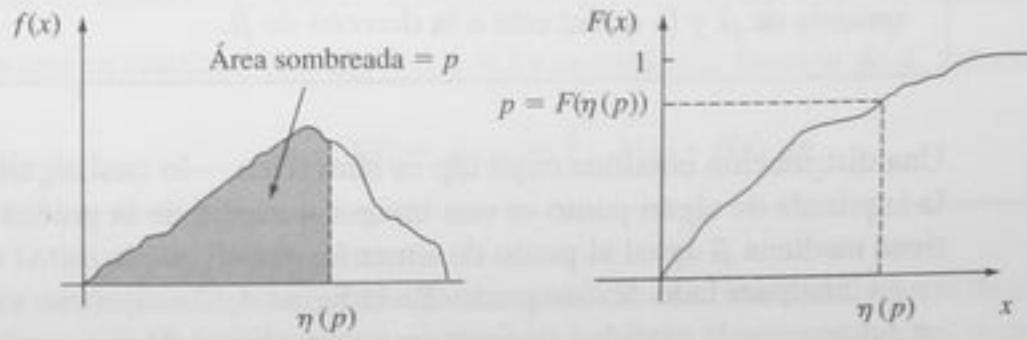


Figura 4.10 Percentil (100p) de una distribución continua

Ejemplo 4.8 La distribución de la cantidad de grava (en toneladas) que vende una casa de materiales para construcción en una semana específica es una *va* continua X con fdp

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

La fda de las ventas para cualquier x entre 0 y 1 es

$$F(x) = \int_0^x \frac{3}{2}(1 - y^2)\,dy = \frac{3}{2}\left(y - \frac{y^3}{3}\right)\Bigg|_{y=0}^{y=x} = \frac{3}{2}\left(x - \frac{x^3}{3}\right)$$

Las greaficas de $f(x)$ y $F(x)$ aparecen cn la figura 4.11. El percentil $(100p)$ de esta distribución, satisface la ecuación

$$p = F(\eta(p)) = \frac{3}{2}\left[\eta(p) - \frac{(\eta(p))^3}{3}\right]$$

es decir,

$$(\eta(p))^3 - 3\eta(p) + 2p = 0$$

Para el percentil 50, $p = 0.5$, y la ecuación por resolver es $\eta^3 - 3\eta + 1 = 0$; la solución es $\eta = \eta(0.5) = 0.347$. Si la distribución es la misma de una semana a otra, entonces a la larga, 50% de las semanas producen ventas menores de 0.347 toneladas y 50% más de 0.347 toneladas.

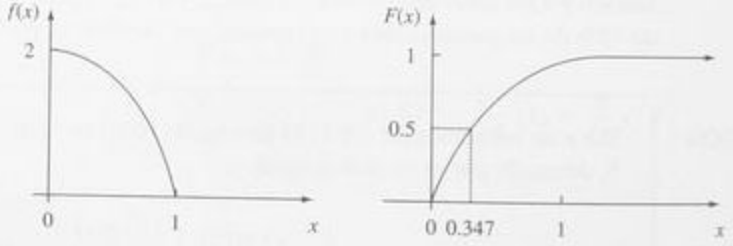


Figura 4.11 Fdp y fda para el ejemplo 4.8 ∎

DEFINICIÓN

> La **mediana** de una distribución continua, denotada por $\tilde{\mu}$, es el percentil 50, así que $\tilde{\mu}$ satisface $0.5 = F(\tilde{\mu})$. Es decir, la mitad del área bajo la curva de densidad está a la izquierda de $\tilde{\mu}$ y la mitad está a la derecha de $\tilde{\mu}$.

Una distribución continua cuya fdp es **simétrica** —lo cual significa que la gráfica de la fdp a la izquierda de algún punto es una imagen a espejo de la gráfica a la derecha de ese punto— tiene mediana $\tilde{\mu}$ igual al punto de simetría, puesto que la mitad del área bajo la curva se ubica en cualquier lado de este punto. En la figura 4.12 se ilustran varios ejemplos. Por lo común se supone que la cantidad de error en una medición de una cantidad física tiene una distribución simétrica.

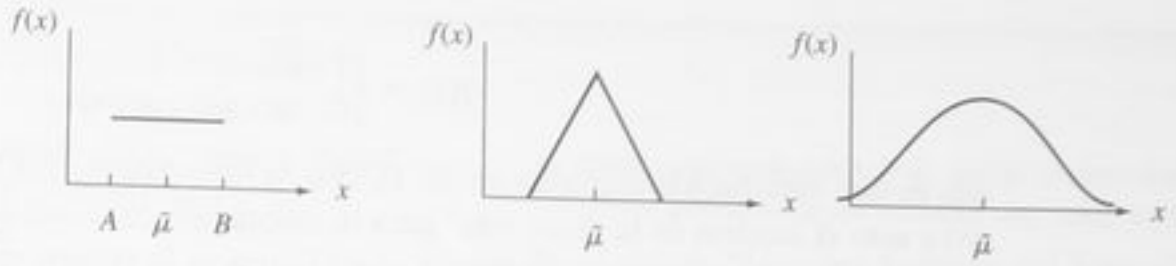


Figura 4.12 Medianas de distribución simétrica

Valores esperados para variables aleatorias continuas

Para una variable aleatoria discreta X, $E(X)$ se obtiene sumando $x \cdot p(x)$ sobre varios valores posibles de X. Aquí la integración sustituye a la suma y la fdp a la fmp para obtener un promedio ponderado continuo.

DEFINICIÓN

El **valor esperado** o **promedio** de una *va* continua X con fdp $f(x)$ es

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Ejemplo 4.9 (continuación del ejemplo 4.8)

La fdp de las ventas semanales de grava X fue

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

por lo tanto,

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_0^1 x \cdot \frac{3}{2}(1 - x^2)\, dx$$

$$= \frac{3}{2}\int_0^1 (x - x^3)\, dx = \frac{3}{2}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Bigg|_{x=0}^{x=1} = \frac{3}{8}$$

■

Cuando la fdp $f(x)$ especifica un modelo para la distribución de valores en una población numérica, entonces μ es la media poblacional, que es la medida de localización poblacional o centro que se emplea con más frecuencia.

A menudo se desea calcular el valor esperado de alguna función $h(X)$ de la variable aleatoria X. Si se considera a $h(X)$ como una nueva *va* Y, se pueden emplear técnicas de la estadística matemática para obtener la fdp de Y, y $E(Y)$ se calcula a partir de la definición. Por fortuna, como en el caso discreto, hay una forma más fácil de calcular $E[h(X)]$.

PROPOSICIÓN

Si X es una *va* continua con fdp $f(x)$ y $h(X)$ es cualquier función de X, entonces

$$E[h(X)] = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x) \cdot f(x)\, dx$$

Ejemplo 4.10 Dos especies compiten en una región para control de una cantidad limitada de cierto recurso. Sea X = proporción del recurso controlado por la especie 1 y suponga que X tiene una fdp

$$f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

que es una distribución uniforme en [0, 1]. (En su libro *Ecological Diversity*, E. C. Pielou llama a esto el modelo de la "vara rota" para la asignación de recursos, puesto que es análogo a romper una vara en un punto elegido al azar.) Entonces la especie que controla la mayor parte de este recurso controla la cantidad

$$h(X) = \text{máx}(X, 1 - X) = \begin{cases} 1 - X & \text{si } 0 \le X < \frac{1}{2} \\ X & \text{si } \frac{1}{2} \le X \le 1 \end{cases}$$

La cantidad esperada controlada por la especie que tiene el control de la mayor parte es entonces

$$E[h(X)] = \int_{-\infty}^{\infty} \text{máx}(x, 1 - x) \cdot f(x)\, dx = \int_0^1 \text{máx}(x, 1 - x) \cdot 1\, dx$$
$$= \int_0^{1/2} (1 - x) \cdot 1\, dx + \int_{1/2}^1 x \cdot 1\, dx = \frac{3}{4}$$

Para una función lineal $h(X)$, $E[h(X)] = E(aX + b) = aE(X) + b$.

Varianza de una variable aleatoria continua

DEFINICIÓN

La **varianza** de una variable aleatoria continua X con fdp $f(x)$ y valor promedio μ es

$$\sigma_X^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)\, dx = E[(X - \mu)^2]$$

La **desviación estándar** (DE) de X es $\sigma_X = \sqrt{V(X)}$.

Como en el caso discreto, σ_X^2 es la esperada o promedio elevado al cuadrado respecto a la media μ y da una medida de cuánta dispersión hay en la distribución de la población de valores de x. La forma más fácil de calcular σ^2 es usar de nuevo una fórmula abreviada.

PROPOSICIÓN

$$V(X) = E(X^2) - [E(X)]^2$$

Ejemplo 4.11 (continuación del ejemplo 4.9)

Para X = ventas semanales de grava, se calcula $E(X) = \frac{3}{8}$. Desde

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx = \int_0^1 x^2 \cdot \frac{3}{2}(1 - x^2)\, dx$$
$$= \int_0^1 \frac{3}{2}(x^2 - x^4)\, dx = \frac{1}{5}$$
$$V(X) = \frac{1}{5} - \left(\frac{3}{8}\right)^2 = \frac{19}{320} = 0.059 \qquad \text{y} \qquad \sigma_X = 0.244$$

Cuando $h(X)$ es una función lineal $V(X) = \sigma^2$, $V[h(X)] = V(aX + b) = a^2 \cdot \sigma^2$ y $\sigma_{aX+b} = |a| \cdot \sigma$.

Ejercicios | Sección 4.2 (11 a 25)

11. La fda de la duración X de tiempo de préstamo de un libro, como se describió en el ejercicio 1 es

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^2}{4} & 0 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

Utilícela para calcular lo siguiente:

a. $P(X \le 1)$

b. $P(0.5 \le X \le 1)$

c. $P(X > 0.5)$

d. La duración mediana de tiempo de préstamo $\tilde{\mu}$ [resuelva $0.5 = F(\tilde{\mu})$]

e. $F'(x)$ para obtener la función de densidad $f(x)$

f. $E(X)$

g. $V(X)$ y σ_X

h. Si al solicitante del préstamo se le carga una cantidad $h(X) = X^2$ cuando la duración del préstamo es X, calcule el cargo esperado $E[h(X)]$.

12. La fda para X (= error de medición) del ejercicio 3 es

$$F(x) = \begin{cases} 0 & x < -2 \\ \dfrac{1}{2} + \dfrac{3}{32}\left(4x - \dfrac{x^3}{3}\right) & -2 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

a. Calcule $P(X < 0)$.

b. Calcule $P(-1 < X < 1)$.

c. Calcule $P(0.5 < X)$.

d. Verifique que $f(x)$ es la que se da en el ejercicio 3 obteniendo $F'(x)$.

e. Verifique que $\tilde{\mu} = 0$.

13. En el ejemplo 4.4 se introdujo el concepto de tiempo de avance en el flujo de tráfico y se propuso una distribución particular para X = avance entre dos automóviles consecutivos elegidos al azar (s). Suponga que en un ambiente de tráfico diferente, la distribución del tiempo de avance tiene la forma

$$f(x) = \begin{cases} \dfrac{k}{x^4} & x > 1 \\ 0 & x \le 1 \end{cases}$$

a. Determine el valor de k para el cual $f(x)$ es una fdp legítima.

b. Obtenga la función de distribución acumulada.

c. Utilice la fda del inciso (b) para determinar la probabilidad de que el avance exceda 2 s y también la probabilidad de que el avance esté entre 2 y 3 s.

d. Obtenga el valor medio y la desviación estándar del avance.

e. Cuál es la probabilidad de que el avance esté dentro de una desviación estándar del valor medio?

14. En el artículo "Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants" (*Water Research*, 1984: 1169-1174) se sugiere la distribución uniforme en el intervalo (7.5, 20) como un modelo para la profundidad (cm) de la capa de bioturbación del sedimento de cierta región.

a. ¿Cuáles son la media y la varianza de la profundidad?

b. ¿Cuál es la fda de la profundidad?

c. ¿Cuál es la probabilidad de que la profundidad observada sea a lo sumo 10? ¿Entre 10 y 15?

d. ¿Cuál es la probabilidad de que la profundidad observada esté dentro de una desviación estándar del valor medio? ¿Dentro de dos desviaciones estándar?

15. Sea X la cantidad de espacio que ocupa un artículo colocado en una caja de empaque de 1 pie cúbico. La fdp de X es

$$f(x) = \begin{cases} 90x^8(1 - x) & 0 < x < 1 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Dibuje la fdp. Luego obtenga la fda de X y dibújela.

b. ¿Cuál es $P(X \le 0.5)$ [es decir, $F(0.5)$]?

c. De acuerdo con el inciso (a), ¿cuál es $P(0.25 < X \le 0.5)$? ¿Cuál es $P(0.25 \le X \le 0.5)$?

d. ¿Cuál es el percentil 75 de la distribución?

e. Calcule $E(X)$ y σ_X.

f. ¿Cuál es la probabilidad de que X esté a más de 1 desviación estándar de su valor medio?

16. Conteste los incisos (a) a (f) del ejercicio 15 con X = tiempo de clase después de la hora dada en el ejercicio 5.

17. Considere la fdp de X = peso de arrastre real dada en el ejercicio 6.

a. Obtenga y dibuje la fda de X.

b. De la gráfica de $f(x)$, ¿cuál es $\tilde{\mu}$?

c. Calcule $E(X)$ y $V(X)$.

18. Sea que X tiene una distribución uniforme en el intervalo $[A, B]$.

a. Obtenga una expresión para el percentil $(100p)$.

b. Calcule $E(X)$, $V(X)$ y σ_X.

c. Para n un entero positivo, calcule $E(X^n)$.

19. Sea X una *va* continua con fda

$$F(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x}{4}\left[1 + \ln\left(\dfrac{4}{x}\right)\right] & 0 < x \le 4 \\ 1 & x > 4 \end{cases}$$

[Este tipo de fda se sugiere en el artículo "Variability in Measured Bedload-Transport Rates" (*Water Resources Bull.*, 1985: 39-48) como modelo para cierta variable hidrológica.] ¿Cuál es:

a. $P(X \le 1)$?
b. $P(1 \le X \le 3)$?
c. la fdp de X?

20. Considere la fdp para tiempo de espera total Y de dos autobuses

$$f(y) = \begin{cases} \frac{1}{25}y & 0 \le y < 5 \\ \frac{2}{5} - \frac{1}{25}y & 5 \le y \le 10 \\ 0 & \text{en caso contrario} \end{cases}$$

presentada en el ejercicio 8.

a. Calcule y dibuje la fda de Y. [*Sugerencia*: considere por separado $0 \le y < 5$ y $5 \le y \le 10$ para calcular $F(y)$. Debe ser útil una gráfica de la fdp.]
b. Obtenga una expresión para el percentil $(100p)$. (*Sugerencia*: considere por separado $0 < p < 0.5$ y $0.5 < p < 1$.)
c. Calcule $E(Y)$ y $V(Y)$. ¿Cómo se comparan estos valores con el tiempo de espera promedio y la varianza de un solo autobús cuando el tiempo tiene una distribución uniforme en [0, 5]?

21. Un ecologista desea marcar una región circular de muestreo con radio de 10 m. Sin embargo, el radio de la región resultante es en realidad una variable aleatoria R con fdp

$$f(r) = \begin{cases} \frac{3}{4}[1 - (10 - r)^2] & 9 \le r \le 11 \\ 0 & \text{en caso contrario} \end{cases}$$

¿Cuál es el área esperada de la región circular resultante?

22. La demanda semanal de gas propano (en miles de galones) de una distribuidora es una *va* X con fdp

$$f(x) = \begin{cases} 2\left(1 - \frac{1}{x^2}\right) & 1 \le x \le 2 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Calcule la fda de X.
b. Obtenga una expresión para el percentil $(100p)$. ¿Cuál es el valor de $\tilde{\mu}$?
c. Calcule $E(X)$ y $V(X)$.
d. Si hay en existencia 1500 galones al comienzo de la semana y durante la semana no es conveniente ningún nuevo suministro, ¿cuánto de los 1500 galones se espera que quede al final de la semana? [*Sugerencia*: sea $h(x)$ = cantidad sobrante cuando la demanda es $= x$.]

23. Si la temperatura a la que se funde cierto compuesto es una variable aleatoria con valor medio de 120°C y desviación estándar de 2°C, ¿cuáles son la temperatura media y la desviación estándar medidas en °F? (*Sugerencia*: °F = 1.8°C + 32.)

24. Sea que X tiene la fdp de Pareto

$$f(x; k, \theta) = \begin{cases} \frac{k \cdot \theta^k}{x^{k+1}} & x \ge \theta \\ 0 & x < \theta \end{cases}$$

introducida en el ejercicio 10.

a. Si $k > 1$, calcule $E(X)$.
b. ¿Qué se puede decir acerca de $E(X)$ si $k = 1$?
c. Si $k > 2$, demuestre que $V(X) = k\theta^2(k-1)^{-2}(k-2)^{-1}$.
d. Si $k = 2$, ¿qué se puede decir acerca de $V(X)$?
e. ¿Qué condiciones en k son necesarias para asegurar que $E(X^n)$ es finita?

25. Sea X la temperatura en °C en la cual tiene lugar cierta reacción química, y sea Y la temperatura en °F (así que $Y = 1.8X + 32$).

a. Si la mediana de la distribución de X es $\tilde{\mu}$, demuestre que $1.8\tilde{\mu} + 32$ es la mediana de la distribución de Y.
b. ¿Cómo está relacionado el percentil 90 de la distribución de Y con el percentil 90 de la distribución de X? Verifique su conjetura.
c. En términos generales, si $Y = aX + b$, ¿cómo se relaciona cualquier percentil particular de la distribución de Y con el correspondiente percentil de la distribución de X?

4.3 Distribución normal

La distribución normal es la más importante en toda la probabilidad y la estadística. Muchas poblaciones numéricas tienen distribuciones que se pueden ajustar con mucha precisión mediante una curva normal apropiada. Como ejemplos se tienen alturas, pesos y otras características físicas (en el famoso artículo de *Biometrika*, 1903, "On the Laws of Inheritance in Man" se analizan muchos ejemplos de este tipo), errores de medición en experimentos científicos,

mediciones antropométricas en fósiles, tiempos de reacción en experimentos psicológicos, mediciones de inteligencia y aptitud, puntuaciones en varias pruebas y numerosas medidas económicas e indicadores. Aun cuando la distribución subyacente es discreta, la curva normal suele dar una aproximación excelente. Además, aunque cada una de las variables por sí mismas no tengan una distribución normal, las sumas y promedios de las variables en condiciones adecuadas tendrán una distribución normal aproximada; éste es el contenido del Teorema del límite central analizado en el capítulo siguiente.

DEFINICIÓN

Se dice que una *va* X tiene una **distribución normal** con parámetros μ y σ (o μ y σ^2), donde $-\infty < \mu < \infty$ y $0 < \sigma$, si la fdp de X es

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \qquad -\infty < x < \infty \tag{4.3}$$

De nuevo, e denota la base del sistema de logaritmos naturales y es aproximadamente 2.71828, y π representa la conocida constante matemática con valor aproximado de 3.14159. El enunciado de que X tiene una distribución normal con parámetros μ y σ^2 suele abreviarse como $X \sim N(\mu, \sigma^2)$.

Es evidente que $f(x; \mu, \sigma) \geq 0$ para cualquier número x, pero se deben usar las técnicas del cálculo de varias variables para demostrar que $\int_{-\infty}^{\infty} f(x; \mu, \sigma)\, dx = 1$. Se puede demostrar que $E(X) = \mu$ y $V(X) = \sigma^2$, así que los parámetros son la media y la desviación estándar de X. En la figura 4.13 se presentan gráficas de $f(x; \mu, \sigma)$ para varios pares diferentes (μ, σ). Cada gráfica es simétrica con respecto a μ y en forma de campana, de modo que el centro de la campana (punto de simetría) es tanto la media de la distribución como la mediana. El valor de σ es la distancia desde μ a los puntos de inflexión de la curva (los puntos en los que cambia la concavidad). Los valores grandes de σ producen gráficas que son bastante dispersas con respecto a μ, mientras que los valores pequeños de σ producen gráficas con un máximo arriba de μ y la mayor parte del área bajo la gráfica bastante cerca de μ. Por consiguiente, un valor grande de σ significa que se podría observar un valor de X lejos de μ en tanto que tal valor es muy improbable cuando σ es pequeña.

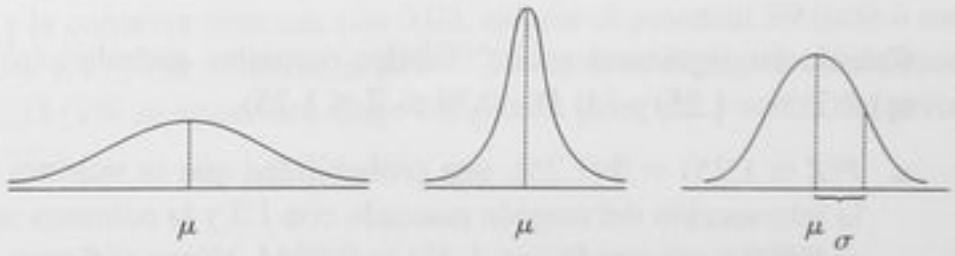


Figura 4.13 Gráficas de fdp normales

Distribución normal estándar

Para calcular $P(a \leq X \leq b)$ cuando X es una *va* normal con parámetros μ y σ, se debe evaluar

$$\int_a^b \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/(2\sigma^2)}\, dx \tag{4.4}$$

Ninguna de las técnicas de integración estándares se pueden usar para evaluar la expresión (4.4). En vez de eso, para $\mu = 0$ y $\sigma = 1$, la expresión (4.4) se ha evaluado numéricamente y tabulado para ciertos valores de a y b. Esta tabla también se puede usar para calcular probabilidades de otros valores de μ y σ en consideración.

DEFINICIÓN

La distribución normal con valores de parámetro $\mu = 0$ y $\sigma = 1$ se llama **distribución normal estándar.** Una variable aleatoria que tiene una distribución normal estándar se llama **variable aleatoria normal estándar** y se denota mediante Z. La fdp de Z es

$$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty$$

La fda de Z es $P(Z \le z) = \int_{-\infty}^{z} f(y; 0, 1)\, dy$, que se denotara mediante $\Phi(z)$.

La distribución normal estándar por lo regular no sirve como modelo para una población que surge de manera natural. En cambio, es una distribución de referencia a partir de la que se puede obtener información acerca de las distribuciones normales. En la tabla A.3 del apéndice se da $\Phi(z) = P(Z \le z)$, el área bajo la gráfica de la fdp normal estándar a la izquierda de z, para $z = -3.49, -3.48, \ldots, 3.48, 3.49$. En la figura 4.14 se ilustra el tipo de área acumulada (probabilidad) tabulada en la tabla A.3. De esta tabla, se pueden calcular otras probabilidades en las que interviene Z.

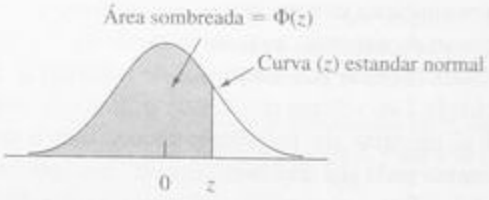


Figura 4.14 Áreas acumuladas de la distribución normal estándar tabuladas en la tabla A.3 del apéndice

Ejemplo 4.12 Calcule las siguientes probabilidades normales estándar: (a) $P(Z \le 1.25)$, (b) $P(Z > 1.25)$, (c) $P(Z \le -1.25)$ y (d) $P(-0.38 \le Z \le 1.25)$.

a. $P(Z \le 1.25) = \Phi(1.25)$, una probabilidad que se muestra en la tabla A.3 del apéndice en la intersección del renglón marcado con 1.2 y la columna marcada con 0.05. El número allí es 0.8944, así que $P(Z \le 1.25) = 0.8944$. Véase al figura 4.15(a).

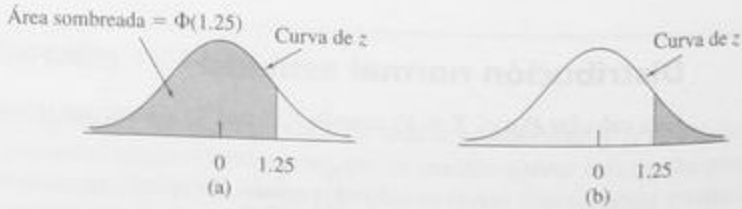


Figura 4.15 Áreas de curvas normales (probabilidades) para el ejemplo 4.12

b. $P(Z > 1.25) = 1 - P(Z \leq 1.25) = 1 - \Phi(1.25)$, el área bajo la curva normal estándar a la derecha de 1.25 (un área del extremo superior). Puesto que $\Phi(1.25) = 0.8944$, se deduce que $P(Z > 1.25) = 0.1056$. Puesto que Z es una *va* continua, $P(Z \geq 1.25)$ también es igual a 0.1056. Véase la figura 4.15(b).

c. $P(Z \leq -1.25) = \Phi(-1.25)$, un área del extremo inferior. Directamente de la tabla A.3 del apéndice, $\Phi(-1.25) = 0.1056$. Por simetría de la curva normal, ésta es la misma respuesta del inciso (b).

d. $P(-0.38 \leq Z \leq 1.25)$ es el área bajo la curva norma estándar arriba del intervalo cuyo punto terminal izquierdo es -0.38 y cuyo punto terminal derecho es 1.25. De la sección 4.2, si X es una *va* continua con fda $F(x)$, entonces $P(a \leq X \leq b) = F(b) - F(a)$. Esto da $P(-0.38 \leq Z \leq 1.25) = \Phi(1.25) - \Phi(-0.38) = 0.8944 - 0.3520 = 0.5424$. (Véase la figura 4.16.)

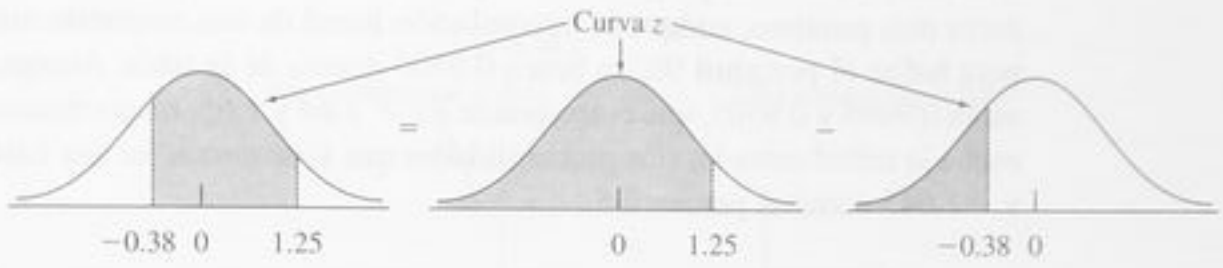


Figura 4.16 $P(-0.38 \leq Z \leq 1.25)$ como la diferencia entre dos áreas acumuladas ■

Percentiles de la distribución normal estándar

Para cualquier p entre 0 y 1, se puede usar la tabla A.3 del apéndice para obtener el percentil $(100p)$ de la distribución normal estándar.

Ejemplo 4.13 El percentil 99 de la distribución normal estándar es el valor del eje horizontal tal, que el área bajo la curva a la izquierda del valor es 0.9900. Ahora la tabla A.3 del apéndice da para z fija el área bajo la curva normal estándar a la izquierda de z, mientras que aquí se tiene el área y se desea el valor de z. Éste es el problema "inverso" a $P(Z \leq z) = ?$ así que se usa la tabla de un modo inverso: encuentre en la mitad de la tabla 0.9900; en el renglón y la columna en los que se ubica este valor, identifique el percentil 99. Aquí 0.9901 se encuentra en el renglón marcado con 2.3 y la columna marcada con 0.03, así que el percentil 99 (más o menos) $z = 2.33$. (Véase la figura 4.17.) Por simetría, el primer percentil es el negativo del percentil 99, así que es igual a -2.33 (1% se encuentra abajo del primer percentil y arriba del percentil 99). (Véase la figura 4.18.)

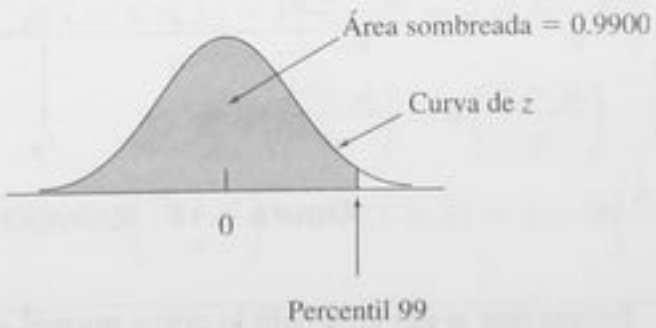


Figura 4.17 Estimación del percentil 99

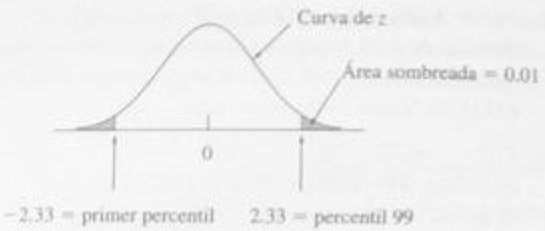


Figura 4.18 Relación entre los percentiles primero y 99

En general, el percentil $(100p)$ se identifica por la columna y el renglón de la tabla A.3 del apéndice en los que se encuentra el elemento p (p. ej., el percentil 67 se obtiene al encontrar 0.6700 en el cuerpo de la tabla, que da $z = 0.44$). Si p no aparece, suele utilizarse el número más próximo, aunque la interpolación lineal da una respuesta más precisa. Por ejemplo, para hallar el percentil 95, se busca 0.9500 dentro de la tabla. Aunque no aparece 0.9500, sí están 0.9495 y 0.9505, que corresponde a $z = 1.64$ y 1.65, respectivamente. Puesto que 0.9500 está a la mitad entre las dos probabilidades que sí aparecen, se usa 1.645 como el percentil 95 y -1.645 como el percentil 5.

Notación z_α

En la inferencia estadística, se necesitan los valores en el eje de medición que captan ciertas áreas extremas pequeñas bajo la curva normal estándar.

> **Notación**
>
> z_α denota el valor en el eje de medición para el cual α del área bajo la curva de z se encuentra a la derecha de z_α. (Véase la figura 4.19.)

Por ejemplo, $z_{0.10}$ captura 0.10 de área del extremo superior y $z_{0.01}$ capta 0.01 de área del extremo superior.

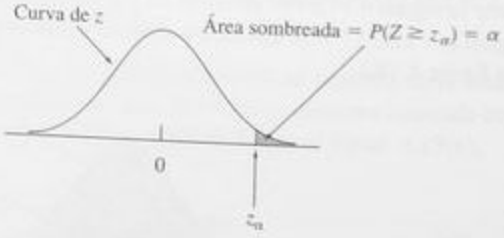


Figura 4.19 Ilustración de la notación z_α

Puesto que α del área bajo la curva normal estándar queda a la derecha de z_α, $1 - \alpha$ del área se encuentra a la izquierda de z_α. Por consiguiente, z_α *es el percentil* $100(1 - \alpha)$ *de la distribución normal estándar*. Por simetría el área bajo la curva normal estándar a la izquierda de

$-z_\alpha$ también es α. Por lo común se hace referencia a las z_α como **valores críticos de z.** En la tabla 4.1 se listan los percentiles más útiles de la distribución normal estándar y los valores de z_α.

Tabla 4.1 Percentiles normales estándar y valores críticos

Percentil	90	95	97.5	99	99.5	99.9	99.95
α (área del extremo)	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
z_α = percentil 100 $(1-\alpha)$	1.28	1.645	1.96	2.33	2.58	3.08	3.27

Ejemplo 4.14 $z_{0.05}$ es el percentil $100(1 - 0.05)$th $= 95$ de la distribución normal estándar, así que $z_{0.05} = 1.645$. El área bajo la curva normal estándar a la izquierda de $-z_{0.05}$ también es 0.05. (Véase la figura 4.20.)

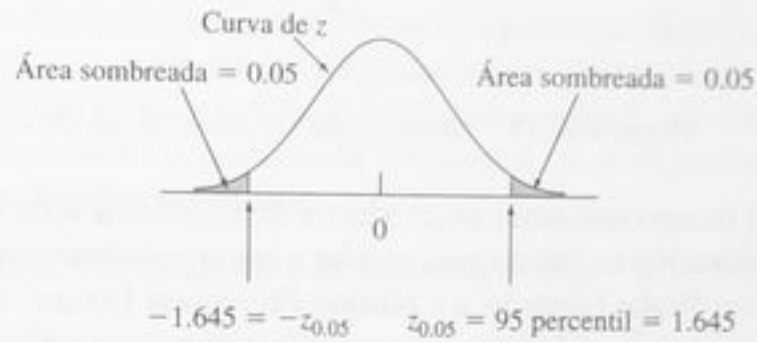


Figura 4.20 Cálculo de $z_{0.05}$ ■

Distribuciones normales no estándar

Cuando $X \sim N(\mu, \sigma^2)$, las probabilidades que tienen que ver con X se calculan mediante "estandarización". La **variable estandarizada** es $(X - \mu)/\sigma$. Al restar μ se desplaza la media de μ a cero, y al dividir entre σ se modifica la escala de la variable de modo que la desviación estándar es 1 y no σ.

PROPOSICIÓN

Si X tiene una distribución normal con media μ y desviación estándar σ, entonces

$$Z = \frac{X - \mu}{\sigma}$$

tiene una distribución normal estándar. Así,

$$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$
$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$
$$P(X \le a) = \Phi\left(\frac{a-\mu}{\sigma}\right) \qquad P(X \ge b) = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right)$$

La idea clave de la proposición es que al estandarizar, cualquier probabilidad en la que interviene X se puede expresar como una probabilidad que tiene que ver con una variable aleatoria

normal estándar Z y, por lo tanto, se puede usar la tabla A.3 del apéndice. Esto se ilustra en la figura 4.21. La proposición se puede demostrar al escribir la fda de $Z = (X - \mu)/\sigma$ como

$$P(Z \le z) = P(X \le \sigma z + \mu) = \int_{-\infty}^{\sigma z + \mu} f(x; \mu, \sigma)\, dx$$

Si se emplea un resultado del cálculo, esta integral se puede derivar con respecto a z para producir la fdp deseada $f(z; 0, 1)$.

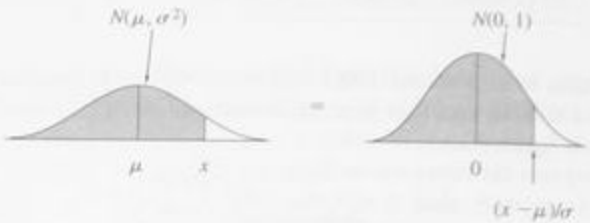


Figura 4.21 Igualdad de las áreas de las curvas normales estándar y no estándar

Ejemplo 4.15 El tiempo que tarda un conductor en reaccionar a las luces del freno de un vehículo en desaceleración es crítico para ayudar a evitar colisiones en la parte posterior. En el artículo "Fast-Rise Brake Lamp as a Collision-Prevention Device" (*Ergonomics*, 1993: 391-395) se indica que el tiempo de reacción para una respuesta en tráfico a una señal de frenado de las luces de freno normales se puede modelar con una distribución normal que tiene valor medio 1.25 s y desviación estándar 0.46 s. ¿Cuál es la probabilidad de que el tiempo de reacción esté entre 1.00 s y 1.75 s? Si X denota el tiempo de reacción, entonces al estandarizar se obtiene

$$1.00 \le X \le 1.75$$

si y sólo si

$$\frac{1.00 - 1.25}{0.46} \le \frac{X - 1.25}{0.46} \le \frac{1.75 - 1.25}{0.46}$$

Así,

$$\begin{aligned} P(1.00 \le X \le 1.75) &= P\left(\frac{1.00 - 1.25}{0.46} \le Z \le \frac{1.75 - 1.25}{0.46}\right) \\ &= P(-0.54 \le Z \le 1.09) = \Phi(1.09) - \Phi(-0.54) \\ &= 0.8621 - 0.2946 = 0.5675 \end{aligned}$$

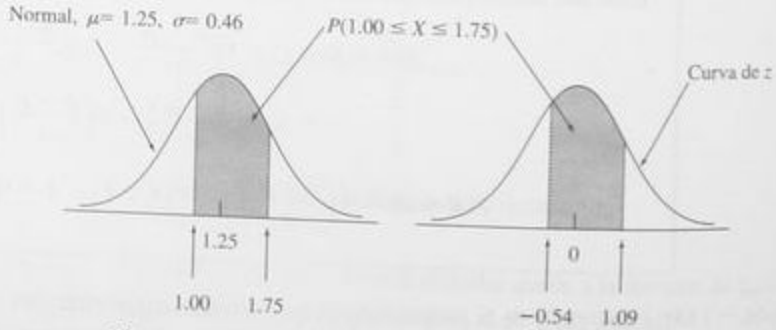


Figura 4.22 Curvas normales para el ejemplo 4.15

Esto se ilustra en la figura 4.22. De manera similar, si se considera que 2 s es mucho tiempo, la probabilidad de que el tiempo de reacción real exceda este valor es

$$P(X > 2) = P\left(Z > \frac{2 - 1.25}{0.46}\right) = P(Z > 1.63) = 1 - \Phi(1.63) = 0.0516$$

■

La estandarización equivale a calcular una distancia a partir de un valor medio y luego volver a expresar la distancia como cierto número de desviaciones estándar. Por ejemplo, si $\mu = 100$ y $\sigma = 15$, entonces $x = 130$ corresponde a $z = (130 - 100)/15 = 30/15 = 2.00$. Es decir, 130 está dos desviaciones estándar arriba del (a la derecha de) del valor medio. De manera similar, al estandarizar 85 se obtiene $(85 - 100)/15 = -1.00$, de modo que 85 está una desviación estándar debajo de la media. La tabla de z se aplica a *cualquier* distribución normal siempre que se piense en términos del número de desviaciones estándar lejos del valor medio.

Ejemplo 4.16 Se sabe que la tensión disruptiva de un diodo de un tipo particular elegido al azar tiene una distribución normal. ¿Cuál es la probabilidad de que la tensión disruptiva del diodo esté dentro de una desviación estándar de su valor medio? Esta pregunta se puede contestar sin conocer el valor de μ o σ, siempre y cuando la distribución sea normal; en otras palabras, la respuesta es la misma para *cualquier* distribución normal:

$$\begin{aligned} P(X \text{ esté dentro de una desviación estándar de su media}) &= P(\mu - \sigma \le X \le \mu + \sigma) \\ &= P\left(\frac{\mu - \sigma - \mu}{\sigma} \le Z \le \frac{\mu + \sigma - \mu}{\sigma}\right) \\ &= P(-1.00 \le Z \le 1.00) \\ &= \Phi(1.00) - \Phi(-1.00) = 0.6826 \end{aligned}$$

La probabilidad de que X esté dentro de 2 desviaciones estándar es $P(-2.00 \le Z \le 2.00) = 0.9544$ y dentro de 3 desviaciones estándar es $P(-3.00 \le Z \le 3.00) = 0.9974$. ■

Los resultados del ejemplo 4.16 suelen presentarse en forma de porcentaje y reciben el nombre de *regla empírica* (debido a que la evidencia empírica ha mostrado que los histogramas de datos reales a menudo se aproximan mediante curvas normales).

> Si la distribución de la población de una variable es (más o menos) normal, entonces:
>
> 1. Alrededor de 68% de los valores están dentro de 1 DE de la media.
> 2. Casi 95% de los valores están dentro de 2 DE de la media.
> 3. Cerca de 99.7% de los valores están dentro de 3 DE de la media.

De hecho, es inusual observar un valor de una distribución normal que esté a mucho más de dos desviaciones estándar de μ. Estos resultados adquieren relevancia en el desarrollo de procedimientos de prueba de hipótesis en capítulos posteriores.

Percentiles de una distribución normal arbitraria

El percentil $(100p)$ de una distribución normal con media μ y desviación estándar σ se relaciona fácilmente con el percentil $(100p)$ de la distribución normal estándar.

PROPOSICIÓN

$$\begin{pmatrix}\text{Percentil } (100p) \\ \text{para } (\mu, \sigma) \text{ normal}\end{pmatrix} = \mu + \begin{bmatrix}(100p) \text{ para} \\ \text{normal estándar}\end{bmatrix} \cdot \sigma$$

Otra forma de decir esto es que si z es el percentil deseado para la distribución normal estándar, entonces el percentil deseado para la distribución normal (μ, σ) está a z desviaciones estándar de μ.

Ejemplo 4.17 La cantidad de agua destilada que entrega cierta máquina tiene una distribución normal con valor medio de 64 oz y una desviación estándar de 0.78 oz. ¿Qué tamaño de recipiente c asegura que el sobreflujo ocurre sólo 5% del tiempo? Si X denota la cantidad entregada, la condición deseada es que $P(X > c) = 0.005$, o bien, lo que es lo mismo, $P(X \le c) = 0.995$. Por consiguiente, c es el percentil 99.5 de la distribución normal con $\mu = 64$ y $\sigma = 0.78$. El percentil 99.5 de la distribución normal estándar es 2.58, de modo que

$$c = \eta(0.995) = 64 + (2.58)(0.78) = 64 + 2.0 = 66 \text{ oz}$$

Esto se ilustra en la figura 4.23.

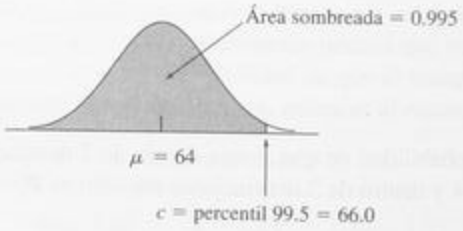


Figura 4.23 Distribución de la cantidad entregada para el ejemplo 4.17 ■

Distribución normal y poblaciones discretas

La distribución normal se usa de ordinario como una aproximación de la distribución de valores en una población discreta. En esas situaciones, se debe tener cuidado de asegurar que las probabilidades se calculan de manera precisa.

Ejemplo 4.18 Se sabe que el CI de cierta población (medido con una prueba estándar) se aproxima a una distribución normal estándar con $\mu = 100$ y $\sigma = 15$. ¿Cuál es la probabilidad de que un individuo elegido al azar tenga un CI de por lo menos 125? Sea que $X =$ CI de una persona elegida al azar, se desea $P(X \ge 125)$. Aquí lo atractivo es estandarizar de inmediato $X \ge 125$ como en los ejemplos anteriores. Sin embargo, la población de CI en realidad es discreta, puesto que los CI son valores enteros, así que la curva normal es una aproximación a un histograma de probabilidad discreta, como se ilustró en la figura 4.24.

Los rectángulos del histograma están *centrados* en enteros, así que los CI de por lo menos 125 corresponden a rectángulos que comienzan en 124.5, la parte sombreada de la figura 4.24. Por lo tanto, en realidad se desea el área bajo la curva normal aproximada a la derecha de 124.5. Al estandarizar este valor se obtiene $P(Z \ge 1.63) = 0.0516$. Si se hubiera estandarizado

$X \geq 125$, se habría obtenido $P(Z \geq 1.67) = 0.0475$. La diferencia no es grande, pero la respuesta 0.0516 es más precisa. De manera similar, $P(X = 125)$ se aproximaría por medio del área entre 124.5 y 125.5, puesto que el área bajo la curva normal arriba del valor único 125 es cero.

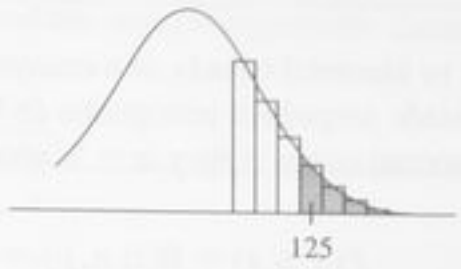


Figura 4.24 Aproximación normal a la distribución discreta ■

La corrección por falta de sucesión de la distribución subyacente del ejemplo 4.18 se denomina **corrección por continuidad**. Es útil en la siguiente aplicación de la distribución normal para el cálculo de probabilidades binomiales.

Aproximación normal a la distribución binomial

Recuerde que el valor medio y la desviación estándar de una variable aleatoria binomial X son $\mu_X = np$ y $\sigma_X = \sqrt{npq}$, respectivamente. En la figura 4.25, se ilustra un histograma de probabilidad binomial para la distribución binomial con $n = 20$, $p = 0.6$ [de modo que $\mu = 20(0.6) = 12$ y $\sigma = \sqrt{20(0.6)(0.4)} = 2.19$]. En el histograma de probabilidad se superpone una curva normal con valor medio y desviación estándar igual a los valores correspondientes para la distribución binomial. Aunque el histograma de probabilidad es un poco sesgado (debido a que $p \neq 0.5$), la curva normal da una buena aproximación, en particular en la parte media de la ilustración. El área del cualquier rectángulo (probabilidad en algún valor particular de X) excepto en los de los extremos, se puede aproximar con precisión mediante el área de la curva normal correspondiente. Por ejemplo, $P(X = 10) = B(10; 20, 0.6) - B(9; 20, 0.6) = 0.117$, mientras que el área bajo la curva normal entre 9.5 y 10.5 es $P(-1.14 \leq Z \leq -0.68) = 0.1212$.

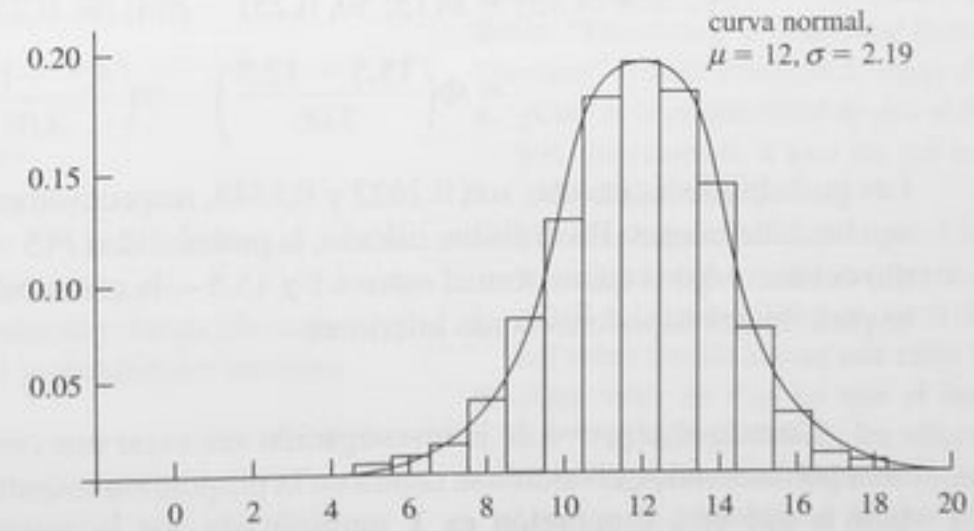


Figura 4.25 Histograma de probabilidad binomial para $n = 20$, $p = 0.6$ con una curva superpuesta de aproximación normal

En términos más generales, siempre y cuando el histograma de la probabilidad binomial no esté muy sesgado, las probabilidades binomiales se pueden aproximar bien mediante áreas de curvas normales. Es común decir que X tiene más o menos una distribución normal.

PROPOSICIÓN

Sea X una *va* binomial basada en n ensayos con probabilidad p de éxito. Entonces, si no está demasiado sesgado el histograma de la probabilidad binomial, X tiene casi una distribución normal con $\mu = np$ y $\sigma = \sqrt{npq}$. En particular, para x = un valor posible de X,

$$P(X \le x) = B(x; n, p) \approx \left(\begin{array}{c}\text{área bajo la curva normal} \\ \text{a la izquierda de } x + 0.5\end{array}\right)$$

$$= \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right)$$

En la práctica, la aproximación es adecuada siempre que $np \ge 10$ y $nq \ge 10$.

Si $np < 10$ o $nq < 10$, la distribución binomial está demasiado sesgada para que la curva normal (simétrica) dé aproximaciones precisas.

Ejemplo 4.19 Suponga que 25% de los conductores con licencia de un estado particular no tienen seguro. Sea X el número de conductores sin seguro en una muestra aleatoria de tamaño 50 (un éxito es un conductor sin seguro), así que $p = 0.25$. Entonces $\mu = 12.5$ y $\sigma = 3.06$. Puesto que $np = 50(0.25) = 12.5 \ge 10$ y $nq = 37.5 \ge 10$, la aproximación se puede aplicar con seguridad:

$$P(X \le 10) = B(10; 50, 0.25) \approx \Phi\left(\frac{10 + 0.5 - 12.5}{3.06}\right)$$

$$= \Phi(-0.65) = 0.2578$$

De manera similar, la probabilidad de que entre 5 y 15 (inclusive) de los conductores seleccionados no estén asegurados es

$$P(5 \le X \le 15) = B(15; 50, 0.25) - B(4; 50, 0.25)$$

$$\approx \Phi\left(\frac{15.5 - 12.5}{3.06}\right) - \Phi\left(\frac{4.5 - 12.5}{3.06}\right) = 0.8320$$

Las probabilidades exactas son 0.2622 y 0.8348, respectivamente, así que las aproximaciones son bastante buenas. En el último cálculo, la probabilidad $P(5 \le X \le 15)$ se aproxima por medio del área bajo la curva normal entre 4.5 y 15.5 —la corrección de continuidad se utiliza tanto para límites superiores como inferiores. ∎

Cuando el objetivo de la investigación sea sacar una conclusión acerca de una proporción poblacional p, el interés se centra en la proporción muestral de éxitos X/n y no en X. Debido a que esta proporción es X multiplicada por la constante $1/n$, también tendrá una distribución normal aproximada (con media $\mu = p$ y desviación estándar $\sigma = \sqrt{pq/n}$) siem-

pre y cuando $np \geq 10$ y $nq \geq 10$. Esta aproximación normal es la base para varios procedimientos inferenciales que se analizan en capítulos posteriores.

Es bastante difícil dar una demostración directa de la validez de esta aproximación normal (la primera se remonta más de 150 años hasta Laplace). En el capítulo siguiente, se verá que es una consecuencia de un resultado general importante llamado Teorema del límite central.

Ejercicios | Sección 4.3 (26 a 54)

26. Sea Z una variable aleatoria normal estándar y calcule las siguientes probabilidades (trace las figuras siempre y cuando sea apropiado).
- **a.** $P(0 \leq Z \leq 2.17)$
- **b.** $P(0 \leq Z \leq 1)$
- **c.** $P(-2.50 \leq Z \leq 0)$
- **d.** $P(-2.50 \leq Z \leq 2.50)$
- **e.** $P(Z \leq 1.37)$
- **f.** $P(-1.75 \leq Z)$
- **g.** $P(-1.50 \leq Z \leq 2.00)$
- **h.** $P(1.37 \leq Z \leq 2.50)$
- **i.** $P(1.50 \leq Z)$
- **j.** $P(|Z| \leq 2.50)$

27. En cada caso, determine el valor de la constante c que hace correcto el enunciado de probabilidad.
- **a.** $\Phi(c) = 0.9838$
- **b.** $P(0 \leq Z \leq c) = 0.291$
- **c.** $P(c \leq Z) = 0.121$
- **d.** $P(-c \leq Z \leq c) = 0.668$
- **e.** $P(c \leq |Z|) = 0.016$

28. Determine los percentiles siguientes para la distribución normal estándar. Interpole cuando sea apropiado.
- **a.** 91
- **b.** 9
- **c.** 75
- **d.** 25
- **e.** 6

29. Determine z_α para lo siguiente:
- **a.** $\alpha = 0.0055$
- **b.** $\alpha = 0.09$
- **c.** $\alpha = 0.663$

30. Si X es una *va* normal con media 80 y desviación estándar 10, calcule las siguientes probabilidades mediante estandarización:
- **a.** $P(X \leq 100)$
- **b.** $P(X \leq 80)$
- **c.** $P(65 \leq X \leq 100)$
- **d.** $P(70 \leq X)$
- **e.** $P(85 \leq X \leq 95)$
- **f.** $P(|X - 80| \leq 10)$

31. Suponga que la fuerza que actúa en una columna que ayuda a sostener un edificio tiene una distribución normal con media de 15.0 kips y desviación estándar de 1.25 kips. ¿Cuál es la probabilidad de que la fuerza
- **a.** sea a lo sumo 18 kips?
- **b.** se encuentre entre 10 y 12 kips?
- **c.** difiera de 15.0 kips en a lo sumo 2 DE?

32. En el artículo "Reliability of Domestic-Waste Biofilm Reactors" (*J. of Envir. Engr.*, 1995: 785-790) se sugiere que la concentración de sustrato (mg/cm^3) de flujo entrante a un reactor tiene una distribución normal con $\mu = 0.30$ y $\sigma = 0.06$.
- **a.** ¿Cuál es la probabilidad de que la concentración sea mayor que 0.25?
- **b.** ¿Cuál es la probabilidad de que la concentración sea a lo sumo 0.10?
- **c.** ¿Cómo caracterizaría el 5% más grande de los valores de concentración?

33. Suponga que el diámetro de árboles de determinado tipo a la altura del pecho (pulg) tiene una distribución normal con $\mu = 8.8$ y $\sigma = 2.8$, como se indica en el artículo "Simulating a Harvester-Forwarder Softwood Thinning" (*Forest Products J.*, mayo de 1997: 36-41).
- **a.** ¿Cuál es la probabilidad de que el diámetro de un árbol, seleccionado al azar sea por lo menos 10 pulg? ¿Mayor de 10 pulgadas?
- **b.** ¿Cuál es la probabilidad de que el diámetro de un árbol seleccionado al azar sea mayor que 20 pulgadas?
- **c.** ¿Cuál es la probabilidad de que el diámetro de un árbol seleccionado al azar esté entre 5 y 10 pulgadas?
- **d.** ¿Qué valor de c es tal que el intervalo $(8.8 - c, 8.8 + c)$ incluya 98% de los valores de diámetro?
- **e.** Si se elige a cuatro árboles de forma independiente, ¿cuál es la probabilidad de que por lo menos uno tenga un diámetro mayor que 10 pulgadas?

34. Se cuenta con dos máquinas para cortar tapones de corcho de botellas de vino. La primera produce corchos con diámetros que tienen una distribución normal con media de 3 cm y desviación estándar de 0.1 cm. La segunda máquina produce corchos con diámetros que tienen una distribución normal con media de 3.04 cm y desviación estándar de 0.02 cm. Los corchos aceptables tienen diámetros entre 2.9 y 3.1 cm. ¿Cuál máquina tiene más probabilidades de producir un corcho aceptable?

35. **a.** Si una distribución normal tiene $\mu = 30$ y $\sigma = 5$, ¿cuál es el percentil 91 de la distribución?
    b. ¿Cual es el sexto percentil de la distribución?
    c. El ancho de una línea grabada en un chip de circuito integrado tiene una distribución normal con media de 3.000 μm y desviación estándar 0.140. ¿Qué valor de amplitud separa el 10% más ancho de tales líneas del otro 90%?

36. En el artículo "Monte Carlo Simulation—Tool for Better Understanding of LRFD" (*J. Structural Engr.*, 1993: 1586-1599) se indica que la resistencia a la deformación permanente (ksi) para acero grado A36 tiene una distribución normal con $\mu = 43$ y $\sigma = 4.5$.
    a. ¿Cuál es la probabilidad de que la resistencia a la deformación permanente sea a lo sumo 40? ¿Mayor que 60?
    b. ¿Cuál valor de resistencia a la deformación permanente separa al 75% más fuerte de los otros?

37. El dispositivo de apertura automática de un paracaídas militar de carga ha sido diseñado para abrirse cuando el paracaídas está a 200 metros sobre el nivel del suelo. Suponga que la altitud de apertura tiene en realidad una distribución normal con valor medio de 200 metros y desviación estándar de 30 metros. El equipo se daña si el paracaídas se abre a una altitud menor que 100 metros. ¿Cuál es la probabilidad de que se dañe la carga de por lo menos uno de cinco paracaídas lanzados de forma independiente?

38. La lectura de temperatura de un termopar colocado en un medio de temperatura constante tiene una distribución normal con media μ, la temperatura real del entorno, y desviación estándar σ. ¿Cuál tendría que ser el valor de σ para asegurar que 95% de las lecturas están dentro de 0.1° de μ?

39. La distribución de resistencia para resistores de cierto tipo es normal, y 10% de los resistores tienen una resistencia mayor que 10.256 ohms y 5% una resistencia menor que 9.671 ohms. ¿Cuáles son la media y la desviación estándar de la distribución de resistencia?

40. Si la longitud de roscado de un perno tiene una distribución normal, ¿cuál es la probabilidad de que la longitud de un perno seleccionado al azar esté
    a. dentro de 1.5 DE de su valor medio?
    b. a más de 2.5 DE de su valor medio?
    c. entre 1 y 2 DE de su valor medio?

41. Una máquina que produce cojinetes se prepara de modo que el diámetro promedio de los cojinetes sea de 0.500 pulg. Un cojinete es aceptable si su diámetro está dentro de 0.004 pulg de su valor objetivo. Sin embargo, suponga que hay variación durante el curso de la producción, de modo que el diámetro de los cojinetes tiene una distribución normal con valor medio de 0.499 pulg y desviación estándar de 0.002 pulg. ¿Qué porcentaje de los cojinetes no es aceptable?

42. La dureza Rockwell de un metal se determina al marcar con un punto acerado la superficie del metal y después medir la profundidad de penetración del punto. Suponga que la dureza Rockwell de cierta aleación tiene una distribución normal con media igual a 70 y desviación estándar 3 (la dureza Rockwell se mide en una escala continua).
    a. Si un espécimen es aceptable sólo si su dureza está entre 67 y 75, ¿cuál es la probabilidad de que un espécimen seleccionado al azar tenga una dureza aceptable?
    b. Si el intervalo aceptable de dureza es $(70 - c, 70 + c)$, ¿para qué valor de c 95% los especímenes tendrían dureza aceptable?
    c. Si el intervalo aceptable es como en el inciso (a) y la dureza de cada diez especímenes elegidos al azar se determina de manera independiente, ¿cuál es el número esperado de especímenes aceptables entre los diez?
    d. ¿Cuál es la probabilidad de que a lo sumo ocho de diez especímenes seleccionados de forma independiente tengan una dureza menor que 73.84? (*Sugerencia:* Y = número entre los diez especímenes con dureza menor que 73.84 es una variable binomial; ¿cuál es p?)

43. La distribución del peso de paquetes enviados de cierto modo es normal con valor medio de 12 libras y desviación estándar de 3.5 libras. El servicio de paquetería desea establecer un valor de peso c más allá del cual habrá sobreprecio. ¿Cuál valor de c es tal que 99% de los paquetes están por lo menos una libra abajo del peso de sobreprecio?

44. Suponga que la tabla A.3 del apéndice tiene $\Phi(z)$ sólo para $z \geq 0$. Explique cómo se podría calcular
    a. $P(-1.72 \leq Z \leq -0.55)$
    b. $P(-1.72 \leq Z \leq 0.55)$

    ¿Es necesario tabular $\Phi(z)$ para z negativa? ¿Qué propiedad de la curva normal estándar justifica su respuesta?

45. Considere a los bebés que nacen en el intervalo "normal" de 37 a 43 semanas de edad de gestación. Gran cantidad de datos apoyan la suposición de que para esos bebés que nacen en Estados Unidos, el peso de nacimiento tiene una distribución normal con media igual a 3432 g y desviación estándar de 482 g. [En el artículo "Are Babies Normal" (*The American Statistician* (1999): 298-302) se analizan datos de un año particular; para una elección sensible de intervalos de clase, la apariencia de un histograma no era nada normal, aunque más tarde se determinó que esto se debía a que en algunos hospitales la medición del peso se hacía en gramos y en otros hasta la onza más próxima y luego se convertía en gramos. Con una elección modificada de intervalos de clase que tomó en cuenta lo anterior se obtuvo un histograma descrito de manera adecuada mediante una distribución normal.

- **a.** ¿Cuál es la probabilidad de que el peso de nacimiento de un bebé de este tipo elegido al azar exceda 4000 gramos?
- **b.** ¿Cuál es la probabilidad de que el peso de nacimiento de un bebé de este tipo elegido al azar sea menos de 2000 o mayor que 5000 gramos?
- **c.** ¿Cuál es la probabilidad de que el peso de nacimiento de un bebé de este tipo elegido al azar sea mayor que 7 libras?
- **d.** ¿Cómo caracterizaría el 1% más extremo de los pesos de nacimiento?
- **e.** Si X es una variable aleatoria con una distribución normal y a es una constante numérica ($a \neq 0$), entonces $Y = aX$ también tiene una distribución normal. Utilice esto para determinar la distribución del peso de nacimiento expresado en libras (forma, media y desviación estándar), y luego calcule de nuevo la probabilidad del inciso (c). ¿De qué manera se compara esto con su respuesta anterior?

46. En respuesta a las preocupaciones acerca del contenido nutrimental de la comida rápida, McDonald's anunció que empleará un nuevo aceite de cocina para sus papas fritas, que disminuirá en forma sustancial la concentración de ácidos grasos, e incrementará la cantidad de grasas poliinsaturadas más benéficas. La compañía afirma que 97 de cada 100 personas no detectan una diferencia de sabor entre los aceites nuevo y antiguo. Suponiendo que esta cifra sea correcta (como proporción de largo plazo), ¿cuál es la probabilidad aproximada de que en una muestra de 100 individuos elegidos al azar que compraron papas en McDonald's,

- **a.** por lo menos 40 noten la diferencia entre los dos aceites?
- **b.** a lo sumo 5% noten la diferencia entre los dos aceites?

47. La desigualdad de Chebyshev, introducida en el ejercicio 43 (capítulo 3) es válida tanto para distribuciones continuas, como discretas. La desigualdad expresa que para cualquier numero k que satisface $k \geq 1$, $P(|X - \mu| \geq k\sigma) \leq 1/k^2$ (véase una interpretación en el ejercicio 43 del capítulo 3). Obtenga esta probabilidad en el caso de una distribución normal para $k = 1$, 2 y 3, y compare con el límite superior..

48. Sea X el número de imperfecciones a lo largo de un carrete de 100 m de cinta magnética (una variable con valores enteros). Suponga que X tiene más o menos una distribución normal con $\mu = 25$ y $\sigma = 5$. Utilice la corrección de continuidad para calcular la probabilidad de que el número de imperfecciones

- **a.** esté entre 20 y 30, inclusive
- **b.** sea a lo sumo 30. Menor que 30.

49. Sea X una distribución binomial con parámetros $n = 25$ y p. Calcule cada una de las probabilidades siguientes por medio de la aproximación normal (con la corrección de continuidad) para los casos $p = 0.5$, 0.6 y 0.8 y compare con las probabilidades exactas calculadas de la tabla A.1 del apéndice.

- **a.** $P(15 \leq X \leq 20)$
- **b.** $P(X \leq 15)$
- **c.** $P(20 \leq X)$

50. Suponga que 10% de los ejes de acero producidos en cierto proceso no cumplen las especificaciones, pero se pueden procesar de nuevo (en lugar de desecharlos). Considere una muestra aleatoria de 200 ejes, y sea X el número entre los que no cumplen las especificaciones pero se pueden reprocesar. ¿Cuál es la probabilidad (aproximada) de que X sea

- **a.** a lo sumo 30?
- **b.** menos de 30?
- **c.** entre 15 y 25 (inclusive)?

51. Suponga que sólo 40% de los conductores de cierto estado usan con regularidad el cinturón de seguridad. Se selecciona al azar una muestra de 500 automovilistas ¿Cuál es la probabilidad de que

- **a.** entre 180 y 230 (inclusive) de los conductores de la muestra utilicen con regularidad el cinturón de seguridad?
- **b.** menos de 175 de los conductores de la muestra utilicen con regularidad el cinturón de seguridad? ¿Menos de 150?

52. Demuestre que la relación entre un percentil normal general y el percentil z correspondiente es como se expresa en esta sección.

53. **a.** Demuestre que si X tiene una distribución normal con parámetros μ y σ, entonces $Y = aX + b$ (una

función lineal de X) también tiene una distribución normal. ¿Cuáles son los parámetros de la distribución de Y [es decir, $E(Y)$ y $V(Y)$]? [*Sugerencia*: escriba la fda de Y, $P(Y \le y)$, como una integral en la que interviene la fdp de X, y luego derive con respecto a y para obtener la fdp de Y.]

b. Si cuando se mide en °C, la temperatura tiene una distribución normal con media 115 y desviación estándar 2, ¿qué se puede decir acerca de la distribución de temperatura medida en °F?

54. No hay una buena fórmula para obtener la fda normal estándar $\Phi(z)$, pero en algunos artículos han aparecido buenas aproximaciones. La siguiente proviene de "Approximations for Hand Calculators Using Small Integer Coefficients" (*Mathematics of Computation*, 1977: 214-222). Para $0 \le z \le 5.5$,

$$P(Z \ge z) = 1 - \Phi(z)$$

$$\approx 0.5 \exp\left\{-\left[\frac{(83z + 351)z + 562}{703/z + 165}\right]\right\}$$

El error relativo de esta aproximación es menor que 0.042%. Utilice esto para calcular aproximaciones a las siguientes probabilidades y compare siempre que sea posible con las probabilidades obtenidas de la tabla A.3 del apéndice.

a. $P(Z \ge 1)$

b. $P(Z < -3)$

c. $P(-4 < Z < 4)$

d. $P(Z > 5)$

4.4 Distribución gamma y sus relativos

La gráfica de cualquier fdp normal tiene forma de campana y, por consiguiente, es simétrica. Hay muchas situaciones prácticas en las que la variable de interés para el experimento podría tener una distribución sesgada. Una familia de fdp que produce una amplia variedad de formas distributivas sesgadas es la gamma. Para definir la familia de distribuciones gamma, primero es necesario introducir una función que desempeña un papel importante en muchas ramas de las matemáticas.

DEFINICIÓN

Para $\alpha > 0$, la **función gamma** $\Gamma(\alpha)$ se define mediante

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx \tag{4.5}$$

Las propiedades más importantes de la función gamma son las siguientes:

1. Para cualquier $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$ [vía integración por partes]
2. Para cualquier entero positivo, n, $\Gamma(n) = (n - 1)!$
3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Mediante la expresión (4.5), si

$$f(x; \alpha) = \begin{cases} \dfrac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} & x \ge 0 \\ 0 & \text{en caso contrario} \end{cases} \tag{4.6}$$

entonces $f(x; \alpha) \ge 0$ y $\int_0^\infty f(x; \alpha)\, dx = \Gamma(\alpha)/\Gamma(\alpha) = 1$, así que $f(x; \alpha)$ satisface las dos propiedades básicas de una fdp.

Familia de distribuciones gamma

DEFINICIÓN

Se dice que una variable continua X tiene una **distribución gamma** si la fdp de X es

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta} & x\geq 0\\ 0 & \text{en caso contrario}\end{cases} \tag{4.7}$$

donde los parámetros α y β satisfacen $\alpha>0$, $\beta>0$. La **distribución gamma estándar** tiene $\beta=1$, de modo que la fdp de una *va* gamma estándar está dada por la ecuación (4.6).

En la figura 4.26a se ilustran las gráficas de la fdp gamma $f(x;\alpha,\beta)$ (4.7) para varios pares (α,β) mientras que la figura 4.26(b) representa gráficas de la fdp gamma estándar. Para la fdp estándar, cuando $\alpha\leq 1$, $f(x;\alpha)$ disminuye en forma estricta cuando x aumenta desde 0; cuando $\alpha>1$, $f(x;\alpha)$ aumenta desde 0 en $x=0$ hasta un máximo y luego disminuye. El parámetro β en (4.7) se llama *parámetro de escala,* porque los valores diferentes de 1 alargan o comprimen la fdp en la dirección x.

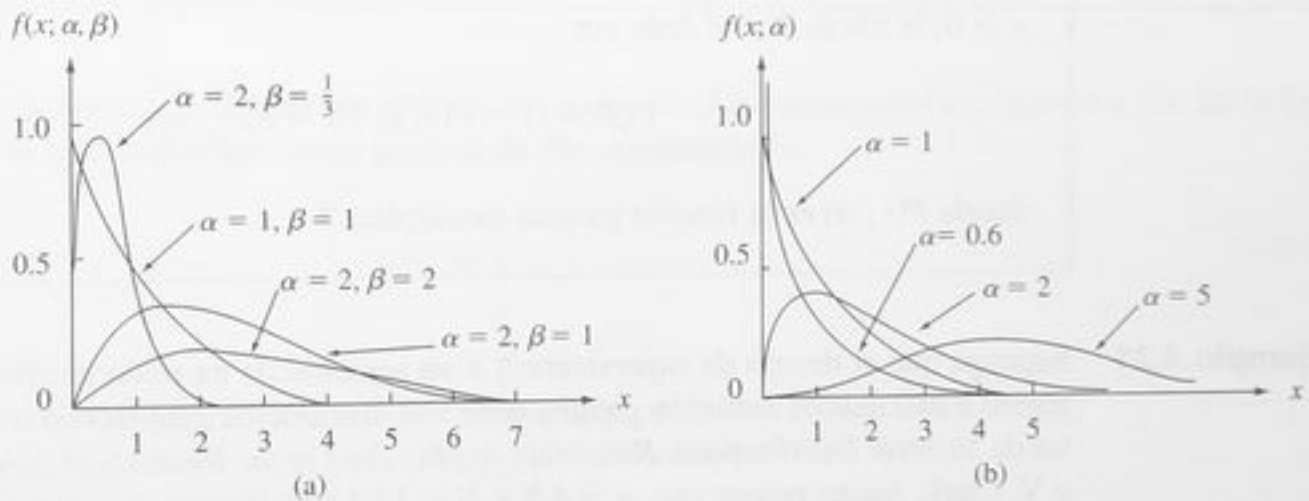


Figura 4.26 (a) Curvas de densidad gamma; (b) curvas de densidad gamma estándar

$E(X)$ y $E(X^2)$ se puede obtener a partir de una integración bastante directa y, por consiguiente, $V(X)=E(X^2)-[E(X)]^2$.

La media y la varianza de un variable aleatoria X que tiene la distribución gamma $f(x;\alpha,\beta)$ son

$$E(X)=\mu=\alpha\beta \qquad V(X)=\sigma^2=\alpha\beta^2$$

Cuando X es una *va* gamma estándar, la fda de X,

$$F(x;\alpha)=\int_0^x \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}\,dy \qquad x>0 \tag{4.8}$$

se llama **función gamma incompleta** [a veces la función gamma incompleta se refiere a la expresión (4.8) sin el denominador $\Gamma(\alpha)$ en el integrando]. Existen tablas extensas de $F(x; \alpha)$ en la tabla A.4 del apéndice, se presenta una tabulación pequeña para $\alpha = 1, 2, \ldots, 10$ y $x = 1, 2, \ldots, 15$.

Ejemplo 4.20 Suponga que el tiempo de reacción X, de un individuo elegido al azar, a cierto estímulo tiene una distribución gamma estándar con $\alpha = 2$ s Puesto que

$$P(a \le X \le b) = F(b) - F(a)$$

cuando X es continua,

$$P(3 \le X \le 5) = F(5; 2) - F(3; 2) = 0.960 - 0.801 = 0.159$$

La probabilidad de que el tiempo de reacción sea más de 4 s es

$$P(X > 4) = 1 - P(X \le 4) = 1 - F(4; 2) = 1 - 0.908 = 0.092$$

■

La función gamma incompleta también se puede usar para calcular probabilidades que tienen que ver con distribuciones gamma no estándar.

PROPOSICIÓN

Sea que X tiene una distribución gamma con parámetros α y β. Entonces, para cualquier $x > 0$, la fda de X está dada por

$$P(X \le x) = F(x; \alpha, \beta) = F\left(\frac{x}{\beta}; \alpha\right)$$

donde $F(\cdot\,; \alpha)$ es la función gamma incompleta.*

Ejemplo 4.21 Suponga que el tiempo de supervivencia X en semanas de un ratón macho elegido al azar expuesto a 240 rads de radiación gamma tiene una distribución gamma con $\alpha = 8$ y $\beta = 15$. (Datos de *Survival Distributions: Reliability Applications in the Biomedical Services*, de A. J. Gross y V. Clark, hacen pensar que $\alpha \approx 8.5$ y $\beta \approx 13.3$.) El tiempo de supervivencia esperado es $E(X) = (8)(15) = 120$ semanas, mientras que $V(X) = (8)(15)^2 = 1800$ y $\sigma_X = \sqrt{1800} = 42.43$ semanas. La probabilidad de que un ratón sobreviva entre 60 y 120 semanas es

$$\begin{aligned} P(60 \le X \le 120) &= P(X \le 120) - P(X \le 60) \\ &= F(120/15; 8) - F(60/15; 8) \\ &= F(8; 8) - F(4; 8) = 0.547 - 0.051 = 0.496 \end{aligned}$$

La probabilidad de que un ratón sobreviva por lo menos 30 semanas es

$$\begin{aligned} P(X \ge 30) &= 1 - P(X < 30) = 1 - P(X \le 30) \\ &= 1 - F(30/15; 8) = 0.999 \end{aligned}$$

■

*MINITAB y otros paquetes de estadística calculan $F(x; \alpha, \beta)$ una vez que se especifican los valores de x, α, y β.

Distribución exponencial

La familia de distribuciones exponenciales proporciona modelos de probabilidad que son muy utilizados en ingeniería y disciplinas científicas.

DEFINICIÓN

Se dice que X tiene una **distribución exponencial** con parámetro λ ($\lambda > 0$) si la fdp de X es

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{en caso contrario} \end{cases} \tag{4.9}$$

La fdp exponencial es un caso especial de la expresión (4.7) de la fdp gamma general en la cual $\alpha = 1$ y β se sustituye con $1/\lambda$ [algunos autores usan la forma $(1/\beta)e^{-x/\beta}$]. Entonces, la media y la varianza de X son

$$\mu = \alpha\beta = \frac{1}{\lambda} \qquad \sigma^2 = \alpha\beta^2 = \frac{1}{\lambda^2}$$

La media y la desviación estándar de la distribución exponencial son iguales a $1/\lambda$. En la figura 4.27 se ilustran varias gráficas de fdp exponenciales.

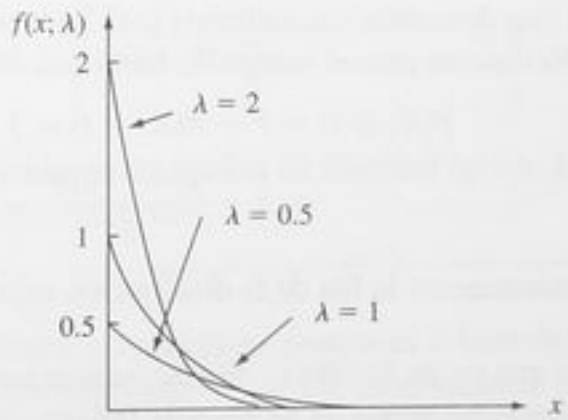


Figura 4.27 Curvas de densidad exponenciales

A diferencia de la fdp gamma general, la fdp exponencial se puede integrar sin dificultad. En particular, la fda de X es

$$F(x; \lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

Ejemplo 4.22 Suponga que el tiempo de respuesta X en cierta terminal de computadora en línea (el tiempo transcurrido entre el final de la petición de un usuario y el comienzo de la respuesta del sistema a esa petición) tiene una distribución exponencial con tiempo de respuesta esperado igual a 5 s. Entonces $E(X) = 1/\lambda = 5$, por lo tanto, $\lambda = 0.2$. La probabilidad de que el tiempo de respuesta sean a lo sumo 10 s es

$$P(X \le 10) = F(10; 0.2) = 1 - e^{-(0.2)(10)} = 1 - e^{-2} = 1 - 0.135 = 0.865$$

La probabilidad de que el tiempo de respuesta esté entre 5 y 10 s es

$$\begin{aligned} P(5 \le X \le 10) &= F(10; 0.2) - F(5; 0.2) \\ &= (1 - e^{-2}) - (1 - e^{-1}) = 0.233 \end{aligned}$$

■

La distribución exponencial se usa muchas veces como modelo para la distribución entre la ocurrencia de eventos sucesivos (p. ej., los clientes que llegan a una instalación de servicio o las llamadas que entran a un conmutador). La razón de esto es que la distribución exponencial tiene una relación estrecha con el proceso de Poisson analizado en el capítulo 3.

PROPOSICIÓN

Suponga que el número de eventos que ocurren en algún intervalo de tiempo de longitud t tiene una distribución de Poisson con parámetro αt (donde α, el índice del proceso del eventos, es el número esperado de eventos que ocurren en una unidad de tiempo) y que el número de ocurrencias en intervalos sin traslape son independientes entre sí. Entonces la distribución del tiempo transcurrido entre la ocurrencia de dos eventos sucesivos es exponencial con parámetro $\lambda = \alpha$.

Aunque una demostración completa está lejos del alcance de este libro, el resultado se comprueba fácilmente para el tiempo X_1 hasta que ocurre el primer evento:

$$\begin{aligned} P(X_1 \le t) &= 1 - P(X_1 > t) = 1 - P[\text{ningún evento en } (0, t)] \\ &= 1 - \frac{e^{-\alpha t} \cdot (\alpha t)^0}{0!} = 1 - e^{-\alpha t} \end{aligned}$$

que es exactamente la fda de la distribución exponencial.

Ejemplo 4.23 Suponga que en un "teléfono de emergencia para suicidios" que da servicio las 24 horas, se reciben llamadas según un proceso de Poisson con tasa $\alpha = 0.5$ llamadas por día. Entonces el número de días X entre llamadas sucesivas tiene una distribución exponencial con valor de parámetro 0.5, de modo que la probabilidad de que transcurran más de dos días entre llamadas es

$$P(X > 2) = 1 - P(X \le 2) = 1 - F(2; 0.5) = e^{-(0.5)(2)} = 0.368$$

El tiempo esperado entre llamadas sucesivas es $1/0.5 = 2$ días. ■

Otra aplicación importante de la distribución exponencial, es modelar la distribución de la vida útil de componentes. Una razón parcial de la popularidad de tales aplicaciones es la **propiedad de "falta de memoria o amnesia"** de la distribución exponencial. Suponga que la vida útil del componente tiene una distribución exponencial con parámetro λ. Después de poner en servicio al componente, se sale durante un periodo de t_0 horas y luego se vuelve para encontrar que aún está trabajando; ahora, ¿cuál es la probabilidad de que dure por lo menos t horas más. En símbolos, se desea $P(X \ge t + t_0 \mid X \ge t_0)$. Por la definición de probabilidad condicional,

$$P(X \geq t + t_0 \mid X \geq t_0) = \frac{P[(X \geq t + t_0) \cap (X \geq t_0)]}{P(X \geq t_0)}$$

Pero el evento $X \geq t_0$ en el numerador es redundante, puesto que ambos eventos ocurren si y sólo si $X \geq t + t_0$. Por lo tanto,

$$P(X \geq t + t_0 \mid X \geq t_0) = \frac{P(X \geq t + t_0)}{P(X \geq t_0)} = \frac{1 - F(t + t_0; \lambda)}{1 - F(t_0; \lambda)} = e^{-\lambda t}$$

Esta probabilidad condicional es idéntica a la probabilidad original $P(X \geq t)$ de que el componente duró t horas. Así, *la distribución de la vida útil adicional es exactamente la misma que la distribución original de la vida útil*, de modo que en cada punto del tiempo el componente no muestre ningún efecto del desgaste. En otras palabras, la distribución de la vida útil restante es independiente de la edad actual.

Aunque la propiedad de falta de memoria se puede justificar por lo menos de forma aproximada en muchos problemas aplicados, en otras situaciones los componentes se deterioran con la edad o, en ocasiones, mejoran (por lo menos hasta cierto punto). Las distribuciones gamma, de Weibull y lognormal proporcionan modelos de vida útil más generales (las dos últimas se analizan en la siguiente sección).

Distribución ji-cuadrada

DEFINICIÓN

Sea ν un entero positivo. Entonces, se dice que una variable aleatoria X tiene una **distribución ji-cuadrada** con parámetro ν si la fdp de X es la densidad gamma con $\alpha = \nu/2$ y $\beta = 2$. Por lo tanto, la fdp de una *va* ji-cuadrada es

$$f(x; \nu) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad (4.10)$$

El parámetro ν se llama **número de grados de libertad** (gl) de X. Por lo común se usa el símbolo χ^2 en lugar de "ji-cuadrada".

La distribución ji-cuadrada es importante, porque es la base de varios procedimientos de inferencia estadística. La razón para esto es que las distribuciones ji-cuadrada están bastante relacionadas con las distribuciones normales (véase el ejercicio 65). En los capítulos acerca de la inferencia, se analiza con más detalle la distribución ji-cuadrada.

Ejercicios | Sección 4.4 (55 a 65)

55. Evalúe lo siguiente:
- **a.** $\Gamma(6)$
- **b.** $\Gamma(5/2)$
- **c.** $F(4; 5)$ (función gamma incompleta)
- **d.** $F(5; 4)$
- **e.** $F(0; 4)$

56. Sea que X tiene una distribución gamma estándar con $\alpha = 7$. Evalúe lo siguiente:
- **a.** $P(X \leq 5)$
- **b.** $P(X < 5)$
- **c.** $P(X > 8)$
- **d.** $P(3 \leq X \leq 8)$
- **e.** $P(3 < X < 8)$
- **f.** $P(X < 4 \text{ o } X > 6)$

57. Suponga que el tiempo que consume un estudiante elegido al azar que utiliza una terminal conectada a una instalación local de computadoras de tiempo compartido tiene una distribución gamma con media 20 min y varianza 80 min^2.

a. ¿Cuáles son los valores de α y β?

b. ¿Cuál es la probabilidad de que el estudiante utilice la terminal a lo sumo 24 minutos?

c. ¿Cuál es la probabilidad de que un estudiante pase entre 20 y 40 minutos usando la terminal?

58. Suponga que cuando un transistor de cierto tipo se somete a una prueba acelerada de duración, la vida útil X (en semanas) tiene una distribución gamma con media de 24 semanas y desviación estándar de 12 semanas.

a. ¿Cuál es la probabilidad de que un transistor dure entre 12 y 24 semanas?

b. ¿Cuál es la probabilidad de que un transistor dure a lo sumo 24 semanas? ¿La mediana de la distribución de la vida útil es menor que 24 semanas? ¿Por qué sí o por qué no?

c. ¿Cuál es el percentil 99 de la distribución de la vida útil?

d. Suponga que la prueba en realidad se termina después de t semanas. ¿Qué valor de t es tal que sólo 0.5% de los transistores aún estará funcionando al término?

59. Sea X = tiempo entre dos llegadas sucesivas en la ventanilla de atención de un banco local. Si X tiene una distribución exponencial con $\lambda = 1$ (que es idéntica a una distribución gamma estándar con $\alpha = 1$), calcule lo siguiente:

a. El tiempo esperado entre dos llegadas sucesivas

b. La desviación estándar del tiempo entre llegadas sucesivas.

c. $P(X \leq 4)$.

d. $P(2 \leq X \leq 5)$.

60. Sea X la distancia (m) que recorre un animal desde el lugar donde nace hasta el primer territorio desocupado que encuentra. Suponga que para las ratas canguro, X tiene una distribución exponencial con parámetro $\lambda = 0.01386$ (como se indica en el artículo "Competition and Dispersal from Multiple Nests", *Ecology*, 1997: 873-883).

a. ¿Cuál es la probabilidad de que la distancia sea a lo sumo 100 m? ¿A lo sumo 200 m? ¿Entre 100 y 200 m?

b. ¿Cuál es la probabilidad de que la distancia sea mayor que la distancia promedio en más de 2 desviaciones estándar?

c. ¿Cuál es el valor de la distancia mediana?

61. La amplia experiencia con ventiladores de cierto tipo utilizados en motores diesel, indica que la distribución exponencial proporciona un buen modelo para el tiempo hasta que se presenta una falla. Suponga que el tiempo medio hasta que se presenta una falla es de 25 000 h. ¿Cuál es la probabilidad de que

a. un ventilador seleccionado al azar dure por lo menos 20 000 h? ¿A lo sumo 30 000 h? ¿Entre 20 000 y 30 000 h?

b. la vida útil de un ventilador exceda el valor medio en más de 2 desviaciones estándar? ¿En más de 3 DE?

62. El caso especial de la distribución gamma en la que α es un entero positivo n se llama distribución de Erlang. Si se sustituye β por $1/\lambda$ en la expresión (4.7), la fdp de Erlang es

$$f(x; \lambda, n) = \begin{cases} \dfrac{\lambda(\lambda x)^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

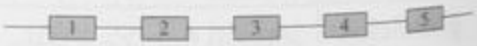
Se puede demostrar que si los tiempos entre eventos sucesivos son independientes, cada uno con una distribución exponencial con parámetro λ, entonces el tiempo total X que transcurre antes que ocurran los siguientes n eventos tiene una fdp $f(x; \lambda, n)$.

a. ¿Cuál es el valor esperado de X? Si el tiempo (en minutos) entre llegadas de clientes sucesivos tiene una distribución exponencial con $\lambda = 0.5$, ¿cuánto tiempo se puede esperar que transcurra antes que llegue el décimo cliente?

b. Si el tiempo entre llegadas de clientes tiene una distribución exponencial con $\lambda = 0.5$, ¿cuál es la probabilidad de que el décimo cliente (después del recién llegado) llegue dentro de los siguientes 30 minutos?

c. El evento $\{X \leq t\}$ ocurre si y sólo si suceden por lo menos n eventos en las siguientes t unidades de tiempo. Utilice el hecho de que el número de eventos que ocurren en un intervalo de longitud t tiene una distribución de Poisson con parámetro λt para escribir una expresión (donde aparezcan probabilidades de Poisson) para la fda de Erlang $F(t; \lambda, n) = P(X \leq t)$.

63. Un sistema consta de cinco componentes idénticos conectados en serie como se muestra:



Al momento que falla un componente, se interrumpe el sistema. Suponga que cada componente tiene una vida útil y una distribución exponencial con $\lambda = 0.01$ y que los componentes fallan sin depender entre sí. Defina los eventos $A_i = \{\text{el } i\text{-ésimo componente dura por lo menos } t \text{ horas}\}$, $i = 1, \ldots, 5$, de modo que las A_i son even-

tos independientes. Sea X = tiempo en que falla el sistema; es decir, la vida útil más corta (mínima) entre los cinco componentes.

a. ¿A qué evento donde intervien $A_1, \ldots, A_5$ es equivalente el evento $\{X \geq t\}$?

b. Por medio de la independencia de las A_i, calcule $P(X \geq t)$. Después, obtenga $F(t) = P(X \leq t)$ y la fdp de X. ¿Qué tipo de distribución tiene X?

c. Suponga que hay n componentes, cada uno con tiempo de vida útil exponencial con parámetro λ. ¿Qué tipo de distribución tiene X?

64. Si X tiene una distribución exponencial con parámetro λ, deduzca una expresión general para el percentil $(100p)$ de la distribución. Después, especifique cómo obtener la mediana.

65. a. ¿A qué evento en el que interviene X es equivalente el evento $\{X^2 \leq y\}$?

b. Si X tiene una distribución normal estándar, utilice el inciso (a) para escribir la integral que es igual a $P(X^2 \leq y)$. Luego, derive con respecto a y para obtener la fdp de X^2 [el cuadrado de una variable $N(0, 1)$ variable]. Por último, demuestre que X^2 tiene una distribución ji-cuadrada con $\nu = 1$ grado de libertad [véase (4.10)]. (*Sugerencia*: utilice la siguiente identidad.)

$$\frac{d}{dy}\left\{\int_{a(y)}^{b(y)} f(x)\,dx\right\} = f[b(y)] \cdot b'(y) - f[a(y)] \cdot a'(y)$$

4.5 Otras distribuciones continuas

Las familias de distribuciones normal, gamma (incluso la exponencial) y uniforme proporcionan una amplia variedad de modelos de probabilidad para variables continuas, pero hay muchas situaciones prácticas en las que ningún miembro de estas familias se ajusta muy bien a un conjunto de datos observados. Los estadísticos y otros investigadores han desarrollado otras familias de distribuciones que suelen ser apropiadas en la práctica.

Distribución de Weibull

El físico sueco Waloddi Weibull introdujo en 1939 la familia de distribuciones de Weibull; en su artículo de 1951 "A Statistical Distribution Function of Wide Applicability" (*J. Applied Mechanics*, vol. 18: 293-297) se analizan varias aplicaciones.

DEFINICIÓN

Se dice que una variable aleatoria X tiene una **distribución Weibull** con parametros α y β $(\alpha > 0, \beta > 0)$ en la fdp X es

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad (4.11)$$

En algunas situaciones, hay justificaciones teóricas para utilizar apropiadamente la distribución de Weibull, pero en muchas aplicaciones $f(x; \alpha, \beta)$ nada más proporciona un buen ajuste para datos observados con valores particulares de α y β. Cuando $\alpha = 1$, la fdp se reduce a la distribución exponencial (con $\lambda = 1/\beta$), así que la distribución exponencial es un caso especial de las distribuciones gamma y de Weibull. Sin embargo, hay distribuciones gamma

que no son distribuciones de Weibull y viceversa, de modo que una familia no es un subconjunto de la otra. Es posible variar a α y β para obtener varias formas de distribución, como se ilustra en la figura 4.28. β es un parámetro de escala, así que diferentes valores alargan o comprimen la gráfica en la dirección x.

La integración para obtener $E(X)$ y $E(X^2)$ produce

$$\mu = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \quad \sigma^2 = \beta^2\left\{\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right\}$$

Por consiguiente, el cálculo de μ y σ^2 requiere usar la función gamma.

La integración $\int_0^x f(y; \alpha, \beta)\, dy$ se realiza sin dificultad para obtener la fda de X.

La fda de una *va* de Weibull con parámetros α y β es

$$F(x; \alpha, \beta) = \begin{cases} 0 & x < 0 \\ 1 - e^{-(x/\beta)^\alpha} & x \geq 0 \end{cases} \tag{4.12}$$

Ejemplo 4.24 En años recientes, la distribución de Weibull se ha empleado para modelar emisiones de varios contaminantes de motores. Sea X la cantidad de NO_x (g/gal) que emite un motor de cuatro tiempos elegido al azar, y suponga que X tiene una distribución de Weibull con $\alpha = 2$ y $\beta = 10$ (según se indica en la información del artículo "Quantification of Variability and Uncertainty en Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors". *J. of the Air and Waste Management Assoc.*, 2002: 435-448). La curva de densidad correspondiente se asemeja a la de la figura 4.28 para $\alpha = 2$, y $\beta = 1$ excepto que ahora los valores 50 y 100 sustituyen a 5 y 10 en el eje horizontal (porque β es un "parámetro de escala"). Entonces,

$$P(X \leq 10) = F(10; 2, 10) = 1 - e^{-(10/10)^2} = 1 - e^{-1} = 0.632$$

De manera similar, $P(X \leq 25) = 0.998$, entonces la distribución está concentrada casi por completo en valores entre 0 y 25. El valor c, que separa 5% de las máquinas que tienen las cantidades más grandes de emisiones de NO_x del 95% restante, satisface

$$0.95 = 1 - e^{-(c/10)^2}$$

Al aislar el término exponencial en un lado, aplicar logaritmos y resolver la ecuación resultante, se obtiene $c = 17.3$ como el percentil 95 de la distribución de vida útil. ∎

Por lo común, en situaciones prácticas un modelo de Weibull podría ser razonable, excepto que el valor más pequeño posible de X puede ser algún valor γ que no se supone sea cero (esto se aplicaría también a un modelo gamma). A la cantidad γ se le puede considerar como un parámetro de la distribución, que es lo que hizo Weibull en su trabajo original. Por ejemplo, para $\gamma = 3$, las curvas de la figura 4.28 se desplazarían tres unidades a la derecha. Esto es equivalente a decir que $X - \gamma$ tiene la fdp (4.11), de modo que la fda de X se obtiene al sustituir x en la ecuación (4.12) por $x - \gamma$.

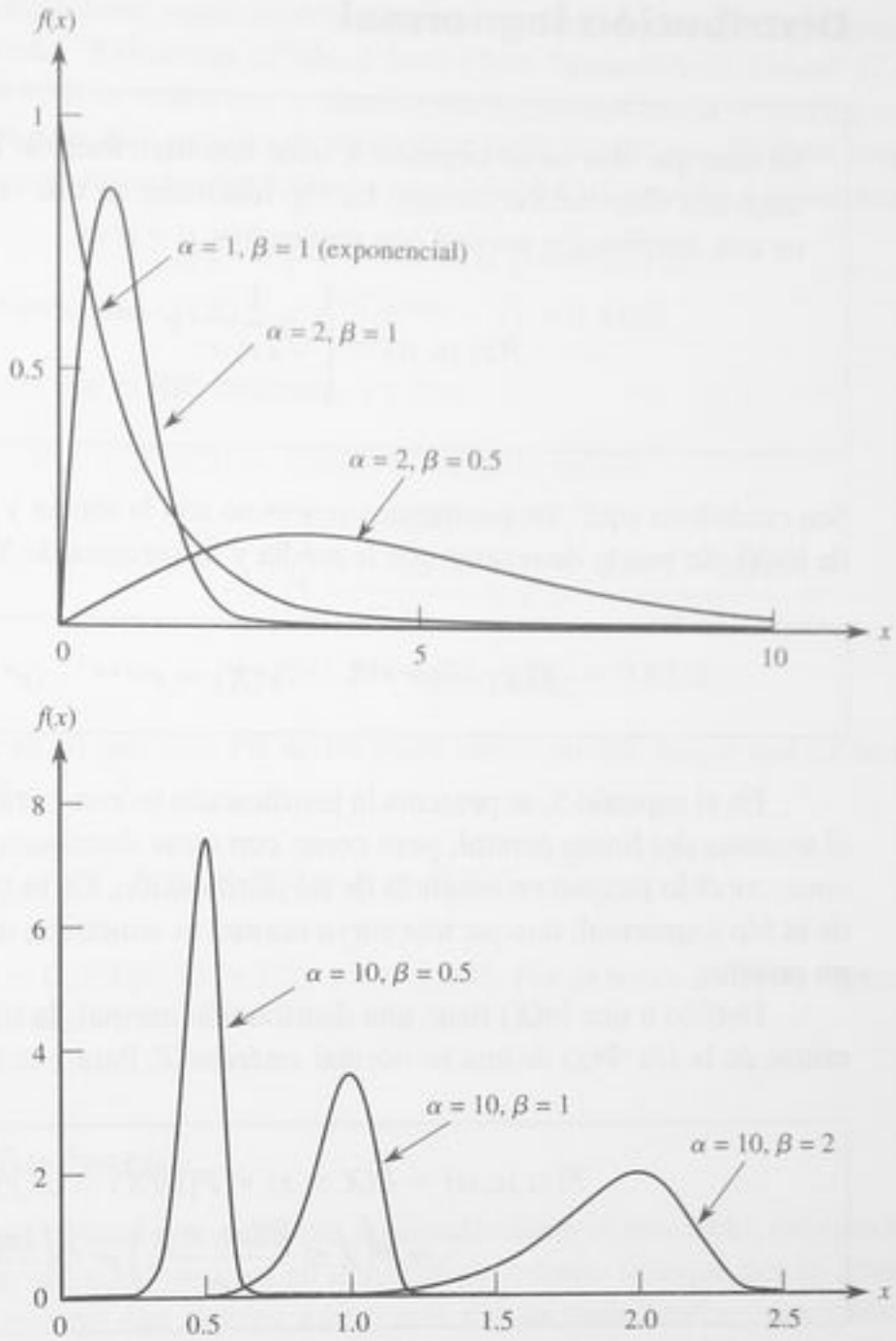


Figura 4.28 Curvas de densidad de Weibull

Ejemplo 4.25 Sea X = la pérdida de peso por corrosión para una pequeña placa cuadrada de aleación de magnesio inmersa durante 7 días en una solución acuosa al 20% de $MgBr_2$. Suponga que la pérdida de peso mínima posible es $\gamma = 3$ y que el exceso $X - 3$ sobre este mínimo tiene una distribución de Weibull con $\alpha = 2$ y $\beta = 4$. (Este ejemplo se consideró en "Practical Applications of the Weibull Distribution". *Industrial Quality Control*, agosto de 1964: 71-78; los valores para α y β se tomaron como 1.8 y 3.67, respectivamente, aunque en el artículo se usó una elección de parámetros un poco distinta.) Entonces, la fda de X es

$$F(x;\alpha,\beta,\gamma) = F(x;2,4,3) = \begin{cases} 0 & x < 3 \\ 1 - e^{-[(x-3)/4]^2} & x \geq 3 \end{cases}$$

Por lo tanto,

$$P(X > 3.5) = 1 - F(3.5;2,4,3) = e^{-0.0156} = 0.985$$

y

$$P(7 \leq X \leq 9) = 1 - e^{-2.25} - (1 - e^{-1}) = 0.895 - 0.632 = 0.263$$

■

Distribución lognormal

DEFINICIÓN

Se dice que una *va* no negativa X tiene una **distribución lognormal** si la *va* $Y = \ln(X)$ tiene una distribución normal. La fdp resultante de una *va* lognormal cuando $\ln(X)$ tiene una distribución normal con parámetros μ y σ es

$$f(x; \mu, \sigma) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,\sigma x} e^{-[\ln(x)-\mu]^2/(2\sigma^2)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Sea cuidadoso aquí; los parámetros μ y σ no son la media y la desviación estándar de X sino de $\ln(X)$. Se puede demostrar que la media y la varianza de X son

$$E(X) = e^{\mu+\sigma^2/2} \qquad V(X) = e^{2\mu+\sigma^2} \cdot (e^{\sigma^2} - 1)$$

En el capítulo 5, se presenta la justificación teórica para esta distribución en relación con el teorema del límite central, pero como con otras distribuciones, la lognormal se puede usar como modelo incluso en ausencia de tal justificación. En la figura 4.29 se ilustran las gráficas de la fdp lognormal; aunque una curva normal es simétrica, una curva lognormal tiene un sesgo positivo.

Debido a que $\ln(X)$ tiene una distribución normal, la fda de X se puede expresar en términos de la fda $\Phi(z)$ de una *va* normal estándar Z. Para $x \geq 0$,

$$\begin{aligned} F(x; \mu, \sigma) &= P(X \leq x) = P[\ln(X) \leq \ln(x)] \\ &= P\left(Z \leq \frac{\ln(x) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right) \end{aligned} \tag{4.13}$$

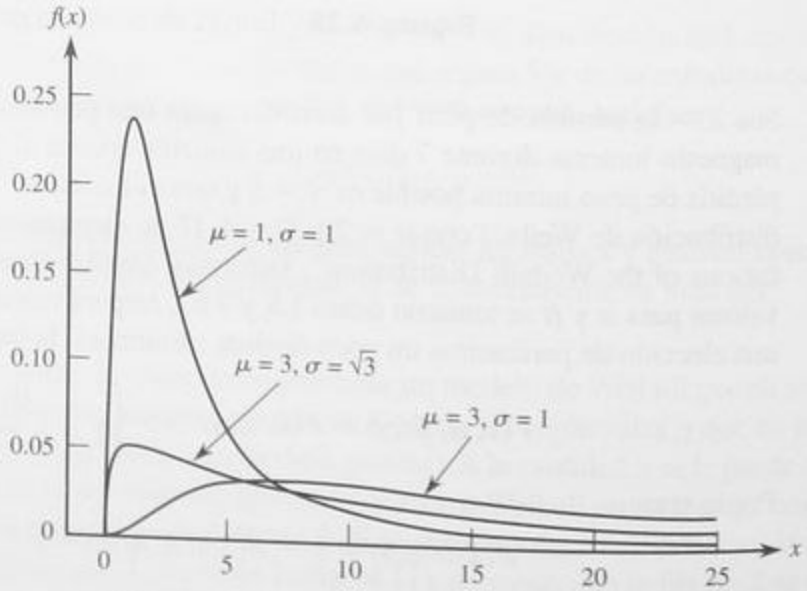


Figura 4.29 Curvas de densidad lognormal

Ejemplo 4.26 La distribución lognormal suele emplearse como un modelo para varias propiedades del material. En el artículo "Reliability of Wood Joist Floor Systems with Creep" (*J. of Structural Engr.*, 1995: 946-954) se indica que la distribución lognormal con $\mu = 0.375$ y $\sigma = 0.25$ es un modelo plausible para X = módulo de elasticidad (ME, en 10^6 psi) de pisos con vigas de madera construidos con abeto hemlock (abeto americano) #2. La media y la varianza del ME son

$$E(X) = e^{0.375 + (0.25)^2/2} = e^{0.40625} = 1.50$$
$$V(X) = e^{0.8125}(e^{0.0625} - 1) = 0.1453$$

La probabilidad de que el ME esté entre 1 y 2 es

$$\begin{aligned} P(1 \le X \le 2) &= P(\ln(1) \le \ln(X) \le \ln(2)) \\ &= P(0 \le \ln(X) \le 0.693) \\ &= P\left(\frac{0 - 0.375}{0.25} \le Z \le \frac{0.693 - 0.375}{0.25}\right) \\ &= \Phi(1.27) - \Phi(-1.50) = 0.8312 \end{aligned}$$

¿Qué valor de c es tal que sólo 1% de los pisos tienen un ME mayor que c? Se desea el valor de c para el cual

$$0.99 = P(X \le c) = P\left(Z \le \frac{\ln(c) - 0.375}{0.25}\right)$$

de donde $(\ln(c) - 0.375)/0.25 = 2.33$ y $c = 2.605$. Por lo tanto, 2.605 es el percentil 99 de la distribución del ME. ■

Distribución beta

Las familias de distribuciones continuas analizadas hasta el momento, excepto la distribución uniforme, tienen densidad positiva en un intervalo infinito (aunque por lo común la función de densidad disminuye con rapidez a cero más allá de unas desviaciones estándar de la media). La distribución beta proporciona densidad positiva sólo para X en un intervalo de longitud finita.

DEFINICIÓN

Se dice que una variable aleatoria X tiene una **distribución beta** con parámetros α, β (ambos positivos), A y B si la fdp de X es

$$f(x;\ \alpha,\ \beta,\ A,\ B) = \begin{cases} \dfrac{1}{B-A} \cdot \dfrac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\dfrac{x-A}{B-A}\right)^{\alpha-1} \left(\dfrac{B-x}{B-A}\right)^{\beta-1} & A \le x \le B \\ 0 & \text{en caso contrario} \end{cases}$$

El caso $A = 0$, $B = 1$ da la **distribución beta estándar.**

En la figura 4.30 se ilustran varias fdp beta estándar. Las gráficas de la fdp general son similares, excepto que están desplazadas y luego se alargan o comprimen para ajustarse en $[A, B]$. A menos que α y β sean enteros, la integración de la fdp para calcular probabilidades es difí-

cil, así que por lo común se emplea una tabla de la función beta incompleta. La media y la varianza de X son

$$\mu = A + (B - A) \cdot \frac{\alpha}{\alpha + \beta} \qquad \sigma^2 = \frac{(B - A)^2 \alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

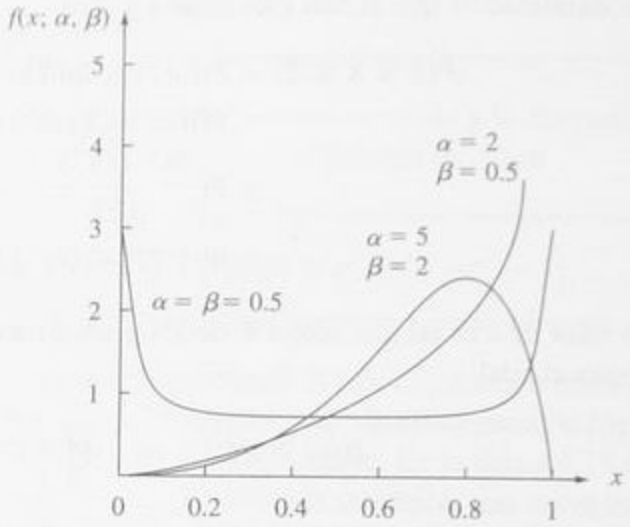


Figura 4.30 Curvas de densidad beta estándar

Ejemplo 4.27 Los administradores de proyecto suelen usar un método denominado PERT —por *program evaluation and review technique*— para coordinar las distintas actividades que conforman un gran proyecto. (una aplicación exitosa se logró en la construcción de la nave espacial *Apolo*). Una suposición normal en el análisis de PERT, es que el tiempo necesario para completar alguna actividad particular, una vez que ha sido iniciada, tiene una distribución beta con A = el tiempo optimista (si todo va bien) y B = el tiempo pesimista (si todo va mal). Suponga que en la construcción de una casa unifamiliar, el tiempo X (en días) necesario para plantar los cimientos tiene una distribución beta con $A = 2$, $B = 5$, $\alpha = 2$ y $\beta = 3$. Entonces $\alpha/(\alpha + \beta) = 0.4$, por lo tanto $E(X) = 2 + (3)(0.4) = 3.2$. Para estos valores de α y β, la fdp de X es una función polinomial simple. La probabilidad de que tome a lo sumo tres días construir la cimentación es

$$P(X \le 3) = \int_2^3 \frac{1}{3} \cdot \frac{4!}{1!2!}\left(\frac{x-2}{3}\right)\left(\frac{5-x}{3}\right)^2 dx$$

$$= \frac{4}{27}\int_2^3 (x-2)(5-x)^2\,dx = \frac{4}{27} \cdot \frac{11}{4} = \frac{11}{27} = 0.407$$

■

La distribución beta estándar se usa por lo regular para modelar la variación en la proporción o porcentaje de una cantidad que ocurre en diferentes muestras como, por ejemplo, la parte de un día de 24 horas que duerme un individuo o la proporción de cierto elemento en un compuesto químico.

Ejercicios | Sección 4.5 (66 a 80)

66. La vida útil X (en cientos de horas) de cierto tipo de tubos al vacío tiene una distribución de Weibull con parámetros $\alpha = 2$ y $\beta = 3$. Calcule lo siguiente:

a. $E(X)$ y $V(X)$

b. $P(X \leq 6)$

c. $P(1.5 \leq X \leq 6)$

(Esta distribución de Weibull se sugiere como modelo para el tiempo de servicio en "On the Assessment of Equipment Reliability: Trading Data Collection Costs for Precision", *J. Engr. Manuf.*, 1991: 105-109.)

67. Los autores del artículo "A Probabilistic Insulation Life Model for Combined Thermal-Electrical Stresses" (*IEEE Trans. on Elect. Insulation*, 1985: 519-522) expresan que "la distribución de Weibull se utiliza ampliamente en problemas de estadística relacionados con la obsolescencia de materiales aislantes sólidos sujetos a envejecimiento y esfuerzo". Proponen el uso de la distribución como modelo para el tiempo (en horas) hasta que fallan especímenes aislantes sólidos sometidos a un voltaje de CA. Los valores de los parámetros dependen del voltaje y temperatura; suponga $\alpha = 2.5$ y $\beta = 200$ (los datos de ese artículo indican estos valores).

a. ¿Cuál es la probabilidad de que la vida útil de un espécimen sea a lo sumo 250? ¿Menos de 250? ¿Más de 300?

b. ¿Cuál es la probabilidad de que la vida útil de un espécimen esté entre 100 y 250?

c. ¿Qué valor es tal que exactamente 50% de los especímenes tienen vidas útiles que exceden ese valor?

68. Sea X = tiempo (en 10^{-1} semanas) desde el envío de un producto defectuoso hasta que el cliente regresa el producto. Suponga que el tiempo mínimo de devolución es $\gamma = 3.5$ y que el exceso $X - 3.5$ sobre el mínimo tiene una distribución de Weibull con parámetros $\alpha = 2$ y $\beta = 1.5$ (véase el artículo *Industrial Quality Control* mencionado en el ejemplo 4.25).

a. ¿Cuál es la fda de X?

b. ¿Cuáles son el tiempo de devolución esperado y la varianza del tiempo de devolución? *[Sugerencia:* Primero obtenga $E(X - 3.5)$ y $V(X - 3.5)$.]

c. Calcule $P(X > 5)$.

d. Calcule $P(5 \leq X \leq 8)$.

69. Considere que X tiene una distribución de Weibull con fdp de la expresión (4.11). Compruebe que $\mu = \beta\Gamma(1 + 1/\alpha)$. (*Sugerencia:* en la integral para $E(X)$, haga el cambio de variable $y = (x/\beta)^{\alpha}$, para que $x = \beta y^{1/\alpha}$.)

70. **a.** En el ejercicio 66, ¿cuál es la mediana de la vida útil de tales tubos? [*Sugerencia:* utilice la expresión (4.12).]

b. En el ejercicio 68, ¿cuál es la mediana del tiempo de devolución?

c. Si X tiene una distribución de Weibull con la fda de la expresión (4.12), obtenga una expresión general para el percentil $(100p)$ de la distribución.

d. En el ejercicio 68, la compañía desea negarse a aceptar devoluciones después de t semanas. ¿Para qué valor de t se rechaza sólo 10% de las devoluciones?

71. Sea X la resistencia última a la tensión (ksi) a $-200°$ de un espécimen de acero de cierto tipo elegido al azar que exhibe "fragilidad en frío" a bajas temperaturas. Suponga que X tiene una distribución de Weibull con $\alpha = 20$ y $\beta = 100$.

a. ¿Cuál es la probabilidad de que X sea a lo sumo 105 ksi?

b. Si se elige espécimen tras espécimen, ¿cuál es la proporción a la larga que tiene valores de resistencia entre 100 y 105 ksi?

c. ¿Cuál es la mediana de la distribución de resistencia?

72. Los autores del artículo del que se tomaron los datos del ejercicio 1.27 sugirieron que un modelo de probabilidad razonable para el tiempo de vida del taladro era una distribución lognormal con $\mu = 4.5$ y $\sigma = 0.8$.

a. ¿Cuáles son el valor medio y la desviación estándar del tiempo de vida?

b. ¿Cuál es la probabilidad de que el tiempo de vida sea a lo sumo 100?

c. ¿Cuál es la probabilidad de que el tiempo de vida sea por lo menos 200? ¿Mayor que 200?

73. Sea X = la potencia mediana por hora (en decibeles) de las señales de radio recibidas que se transmiten entre dos ciudades. Los autores del artículo "Families of Distributions for Hourly Median Power and Instantaneous Power of Received Radio Signals" (*J. Research National Bureau of Standards*, vol. 67D, 1963: 753-762) argumentan que la distribución lognormal proporciona un modelo de probabilidad razonable para X. Si los valores de los parámetros son $\mu = 3.5$ y $\sigma = 1.2$, calcule lo siguiente:

a. El valor medio y la desviación estándar de la potencia recibida.

b. La probabilidad de que la potencia recibida esté entre 50 y 250 dB.

c. La probabilidad de que X sea menor que su valor medio. ¿Por qué esa probabilidad no es 0.5?

74. a. Utilice la ecuación (4.13) para escribir una fórmula para la mediana $\tilde{\mu}$ de la distribución lognormal. ¿Cuál es la mediana para la distribución de potencia del ejercicio 73?
    b. Recuerde que z_α es la notación para el percentil $100(1-\alpha)$ de la distribución normal estándar, escriba una expresión para el percentil $100(1-\alpha)$ de la distribución lognormal. En el ejercicio 73, ¿cuál valor de la potencia recibida es mayor sólo 5% del tiempo?

75. Una justificación teórica basada en un mecanismo de falla de cierto material sustenta la suposición de que la resistencia a la ductilidad X de un material tiene una distribución lognormal. Suponga que los parámetros son $\mu = 5$ y $\sigma = 0.1$.
    a. Calcule $E(X)$ y $V(X)$.
    b. Calcule $P(X > 125)$.
    c. Calcule $P(110 \le X \le 125)$.
    d. ¿Cuál es el valor de la resistencia mediana a la ductilidad?
    e. Si diez muestras diferentes de una aleación de acero de este tipo se someten a una prueba de resistencia, ¿cuántas esperaría que tuvieran resistencia de por lo menos 125?
    f. Si el 5% más pequeño de los valores de resistencia fueran inaceptables, ¿cuál sería la resistencia mínima aceptable?

76. En el artículo "The Statistics of Phytotoxic Air Pollutants" (*J. Royal Stat. Soc.*, 1989: 183-198) se indica la distribución lognormal como un modelo para la concentración de SO_2 en la parte superior de cierto bosque. Suponga que los valores de los parámetros son $\mu = 1.9$ y $\sigma = 0.9$.
    a. ¿Cuáles son el valor medio y la desviación estándar de la concentración?
    b. ¿Cuál es la probabilidad de que la concentración sea a lo sumo 10? ¿De qué esté entre 5 y 10?

77. ¿Qué condición sobre α y β es necesaria para que la fdp beta estándar sea simétrica?

78. Suponga que la proporción X de área superficial de un cuadrado seleccionado al azar cubierto por cierta planta tiene una distribución beta estándar con $\alpha = 5$ y $\beta = 2$.
    a. Calcule $E(X)$ and $V(X)$.
    b. Calcule $P(X \le 0.2)$.
    c. Calcule $P(0.2 \le X \le 0.4)$.
    d. ¿Cuál es la proporción esperada de la región de muestreo que no cubre la planta?

79. Sea que X tiene una densidad beta estándar con parámetros α y β.
    a. Verifique la fórmula para $E(X)$ dada en la sección.
    b. Calcule $E[(1-X)^m]$. Si X representa la proporción de una sustancia que consiste en un ingrediente particular, ¿cuál es la proporción esperada que no consiste en este ingrediente?

80. Se aplica esfuerzo a una barra de acero de 20 pulgadas que está sujeta en una posición fija en cada extremo. Sea Y = distancia desde el extremo izquierdo en el que se rompe la barra. Suponga que $Y/20$ tiene una distribución beta estándar con $E(Y) = 10$ y $V(Y) = \frac{100}{7}$.
    a. ¿Cuáles son los parámetros de la distribución beta estándar de interés?
    b. Calcule $P(8 \le Y \le 12)$.
    c. Calcule la probabilidad de que la barra se rompa a más de 2 pulgadas del lugar previsto.

4.6 Gráficas de probabilidad

DEFINICIÓN

Un investigador con frecuencia obtendrá una muestra numérica $x_1, x_2, \ldots, x_n$ y deseará saber si es posible que provenga de una distribución de población de algún tipo particular (p. ej. de una distribución normal. Entre otras cosas, muchos procedimientos formales de la inferencia estadística se basan en la suposición de que la distribución de la población es de un tipo específico. El uso de un procedimiento de este tipo es inapropiado si la distribución de probabilidad real subyacente difiere mucho del tipo supuesto. Además, entender la distribución subyacente da una idea de los mecanismos físicos que tienen que ver con la generación de datos. Una forma efec-

tiva de comprobar una hipótesis distribucional, es construir lo que se llama **gráfica de probabilidad**. La esencia de tal gráfica es que si la distribución en la que se basa es correcta, los puntos de la gráfica caerán cerca de una línea recta. Si la distribución real es bastante diferente de la utilizada para construir la gráfica, los puntos deben desviarse en forma sustancial del patrón lineal.

Percentiles muestrales

Los detalles relacionados con la construcción de gráficas de probabilidad difieren un poco de una fuente a otra. La base para la construcción es una comparación entre percentiles de los datos muestrales y los percentiles correspondientes de la distribución en consideración. Recuerde que el percentil $(100p)$ de una distribución continua con fda $F(\cdot)$ es el número $\eta(p)$ que satisface $F(\eta(p)) = p$. Es decir, $\eta(p)$ es el número en la escala de medición tal que el área bajo la curva de densidad a la izquierda de $\eta(p)$ es p. Así, el percentil 50 $\eta(0.5)$ satisface $F(\eta(0.5)) = 0.5$, y el percentil 90 satisface $F(\eta(0.9)) = 0.9$. Considere como ejemplo la distribución normal estándar, para la cual se ha denotado la fda mediante $\Phi(\cdot)$. De la tabla A.3 del apéndice, se encuentra el percentil 20 localizando el renglón y la columna donde 0.2000 (o un número lo más cercano posible a éste) aparece dentro de la tabla. Puesto que 0.2005 aparece en la intersección del renglón -0.8 y la columna 0.04, el percentil 20 es más o menos -0.84. De manera similar, el percentil 25 de la distribución normal estándar es (usando la interpolación lineal) casi -0.675.

A grandes rasgos, los percentiles muestrales se definen en la misma forma que los percentiles de una distribución poblacional. El percentil muestral 50 debe separar el 50% más pequeño de la muestra del 50% más grande, el percentil 90 debe ser tal que 90% de la muestra queda debajo de ese valor y 10% queda arriba, etcétera. Infortunadamente, hay problemas cuando en realidad se trata de calcular los percentiles muestrales para una muestra particular de n observaciones. Si, por ejemplo, $n = 10$, se puede dividir 20% de estos valores o 30% de los datos, pero no hay valor que divida en forma exacta 23% de estas diez observaciones. Para avanzar más, se requiere una definición operacional de percentiles muestrales (ésta es una situación donde personas diferentes hacen cosas un poco distintas). Recuerde que cuando n es par, la mediana muestral o el percentil muestral 50 es el valor medio en la lista ordenada, por ejemplo, el sexto valor más grande cuando $n = 11$. Esto implica considerar la observación media como la mitad de la mitad inferior de los datos y la mitad en la mitad superior. De manera similar, suponga que $n = 10$. Entonces si al tercer valor más pequeño se le llama percentil 25, se está considerando al valor como la mitad del grupo inferior (que consiste en las dos observaciones más pequeñas) y la mitad del grupo superior (las siete observaciones más grandes). Esto origina la siguiente definición general de percentiles muestrales.

DEFINICIÓN

> Ordene las n observaciones muestrales de menor a mayor. Luego, la i-ésima observación de la lista se toma como **percentil muestral [$100(i - 0.5)/n$].**

Una vez que han sido calculados los valores de porcentaje $100(i - 0.5)/n$ $(i = 1, 2, \ldots, n)$ los percentiles muestrales correspondientes se pueden obtener mediante interpolación lineal. Por ejemplo, si $n = 10$, los porcentajes que corresponden a las observaciones muestrales son $100(1 - 0.5)/10 = 5\%$, $100(2 - 0.5)/10 = 15\%$, 25%, . . . , y $100(10 - 0.5)/10 = 95\%$.

Entonces el percentil 10 es la mitad entre el percentil 5 (observación muestral más pequeña) y el percentil 15 (segunda observación más pequeña). Para los fines que aquí se persiguen, esta interpolación es innecesaria, porque una gráfica de probabilidad se basará sólo en los porcentajes $100(i - 0.5)/n$ que corresponden a las n observaciones muestrales.

Gráfica de probabilidad

Suponga ahora que para los porcentajes $100(i - 0.5)/n$ $(i = 1, \ldots, n)$ los percentiles se determinan para una distribución específica cuya posibilidad se investiga. Si la muestra se eligiera en realidad de la distribución especificada, los percentiles muestrales (observaciones muestrales ordenadas) deben estar razonablemente cerca de los percentiles correspondientes de la distribución poblacional. Es decir, para $i = 1, 2, \ldots, n$ debe haber un acuerdo razonable entre la observación muestral más pequeña y el percentil $[100(i - 0.5)/n]$ para la distribución específica. Considere los pares (percentil poblacional, percentil muestral); es decir, los pares

$$\begin{pmatrix} \text{percentil } [100(i - 0.5)/n] \\ \text{de la distribución} \end{pmatrix}, \begin{pmatrix} i\text{-ésima observación} \\ \text{muestral más pequeña} \end{pmatrix}$$

para $i = 1, \ldots, n$. Cada par se puede dibujar como un punto en un sistema coordenado bidimensional. Si los percentiles muestrales están cerca de los percentiles correspondientes de la distribución poblacional, el primer número de cada par será más o menos igual al segundo número. Los puntos dibujados caerán entonces cerca de una recta de 45°. Las desviaciones sustanciales de los puntos dibujados con respecto a una recta de 45° arrojan dudas sobre la suposición de que la distribución en consideración es la correcta.

Ejemplo 4.28 Un experimentador conoce el valor de cierta constante física. Él lleva a cabo $n = 10$ mediciones independientes de este valor por medio de un determinado dispositivo de medición y registra los errores de medición resultantes (error = valor observado − valor verdadero). Estas observaciones aparecen en la tabla siguiente.

Porcentaje	5	15	25	35	45
Percentil z	−1.645	−1.037	−0.675	−0.385	−0.126
Observación muestral	−1.91	−1.25	−0.75	−0.53	0.20
Porcentaje	55	65	75	85	95
Percentil z	0.126	0.385	0.675	1.037	1.645
Observación muestral	0.35	0.72	0.87	1.40	1.56

¿Es posible que el *error de medición* de la variable aleatoria tenga una distribución normal estándar? Los percentiles (z) normales estándar necesarios también se muestran en la tabla. Así, los puntos de la gráfica de probabilidad son (−1.645, −1.91), (−1.037, −1.25), . . . , y (1.645, 1.56). En la figura 4.31 se ilustra la gráfica resultante. Aunque los puntos se desvían un poco de la recta de 45°, la impresión predominante es que esta recta se ajusta muy bien a los pun-

tos. En la gráfica se observa que la distribución normal estándar es un modelo de probabilidad razonable para el error de medición.

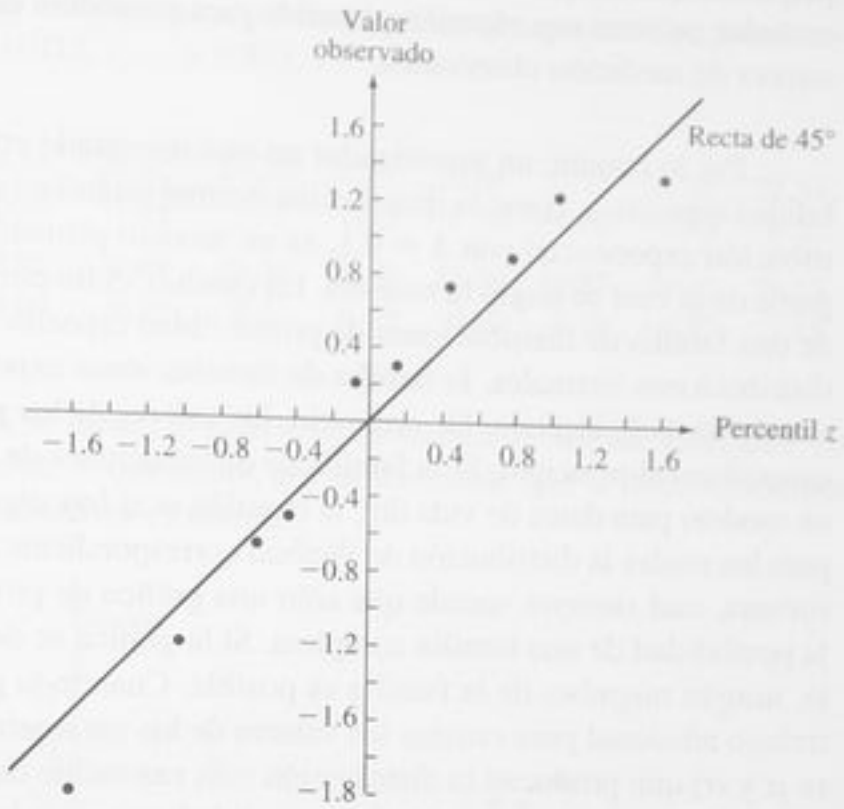


Figura 4.31 Gráficas de pares (percentil z, valor observado) para los datos del ejemplo 4.28: primera muestra

En la figura 4.32 se ilustra una gráfica de pares (percentil z, observación) para una segunda muestra de diez observaciones. La recta de 45° da un buen ajuste para la parte media de la muestra, pero no para los extremos. La gráfica tiene una apariencia de S bien definida. Las dos observaciones muestrales más pequeñas son considerablemente más grandes que los

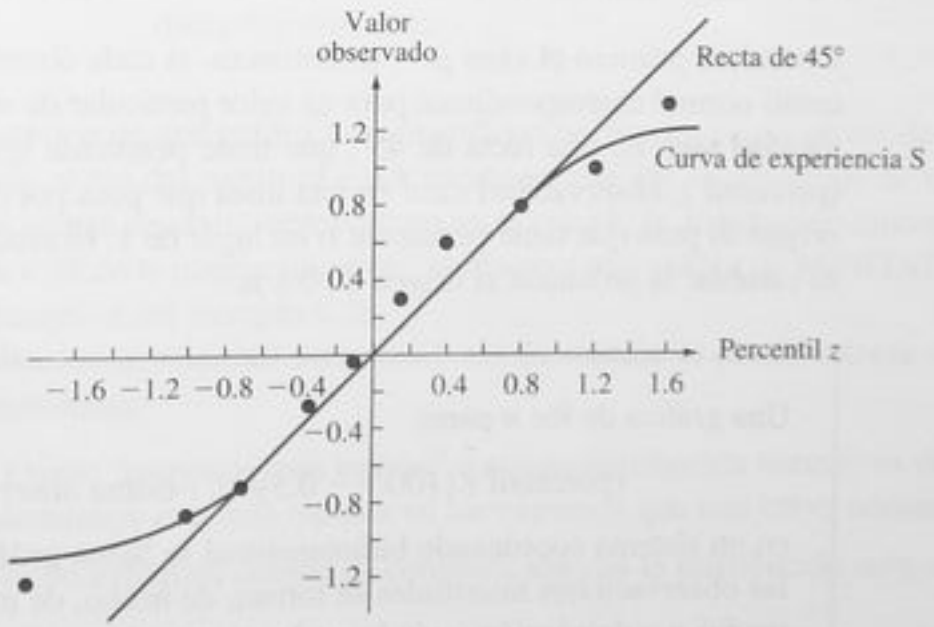


Figura 4.32 Gráficas de pares (percentil z, valor observado) para los datos del ejemplo 4.28: segunda muestra

percentiles z correspondientes (los puntos del extremo izquierdo de la gráfica están bien sobre la línea de 45°. Asimismo las dos observaciones más grandes de la muestra son mucho más pequeños que los percentiles z relacionados). Esta gráfica indica que la distribución normal estándar no sería una elección plausible para el modelo de probabilidad que dio lugar a estos errores de medición observados. ■

Por lo común, un investigador no está interesado en saber si una distribución de probabilidad específica, como la distribución normal estándar (normal con $\mu = 0$ y $\sigma = 1$) o la distribución exponencial con $\lambda = 0.1$, es un modelo plausible para la distribución poblacional a partir de la cual se eligió la muestra. En cambio, el investigador desea saber si *algún* miembro de una familia de distribuciones de probabilidad específica un modelo plausible: la familia de distribuciones normales, la familia de distribuciones exponenciales, la familia de distribuciones de Weibull, etcétera. De ordinario, los valores de los parámetros de una distribución no se especifican al principio. Si la familia de distribuciones de Weibull está en consideración como un modelo para datos de vida útil, la cuestión es si hay *algunos* valores de los parámetros α y β para los cuales la distribución de Weibull correspondiente da un buen ajuste para los datos. Por fortuna, casi siempre sucede que sólo una gráfica de probabilidad es suficiente para evaluar la posibilidad de una familia completa. Si la gráfica se desvía considerablemente de una recta, ningún miembro de la familia es posible. Cuando la gráfica es bastante recta, se requiere trabajo adicional para estimar los valores de los parámetros (p. ej., determinar los valores para μ y σ) que producen la distribución más razonable del tipo especificado.

Con el fin de comprobar la normalidad, se centra la atención en una gráfica. Ésta puede ser muy útil en el trabajo aplicado porque muchos procedimientos estadísticos formales son apropiados (proporcionan inferencias precisas) sólo cuando la distribución poblacional es por lo menos casi normal. Por lo regular, estos procedimientos no se deben usar si la gráfica de probabilidad normal muestra una desviación muy pronunciada con respecto a la linealidad. La clave para construir una gráfica de probabilidad normal general es la relación entre los percentiles (z) normales estándar y los de cualquier otra distribución normal:

$$\begin{pmatrix}\text{percentil para una}\\ \text{distribución normal}(\mu, \sigma)\end{pmatrix} = \mu + \sigma \cdot \begin{pmatrix}\text{percentil } z\\ \text{correspondiente}\end{pmatrix}$$

Considere primero el caso $\mu = 0$. Entonces, si cada observación es exactamente igual al percentil normal correspondiente para un valor particular de σ, los pares ($\sigma \cdot$ [percentil z], observación) caen en una recta de 45°, que tiene pendiente igual a 1. Esto implica que los pares (percentil z, observación) caen en una línea que pasa por (0, 0) (es decir, una con ordenada al origen 0) pero que tiene pendiente σ en lugar de 1. El efecto de un valor de μ distinto de cero es cambiar la ordenada al origen de 0 a μ.

Una gráfica de los n pares

$$(\text{percentil } z\ [100(i - 0.5)/n],\ i\text{-ésima observación más pequeña})$$

en un sistema coordenado bidimensional se llama **gráfica de probabilidad normal.** Si las observaciones muestrales se toman, de hecho, de una distribución normal con valor medio μ y desviación estándar σ, los puntos deben caer cerca de una recta con pendiente σ y ordenada al origen μ. Así, una gráfica para la cual los puntos caen cerca de alguna recta, hace pensar que la suposición de una distribución poblacional normal es razonable.

Ejemplo 4.29 La muestra siguiente consiste en $n = 20$ observaciones en la tensión disruptiva dieléctrica de una pieza de resina epóxica que aparecen en el artículo "Maximum Likelihood Estimation in the 3-Parameter Weibull Distribution (*IEEE Trans. on Dielectrics and Elec. Insul.*, 1996: 43-55). Los valores de $(i - 0.5)/n$ para que se requieren los percentiles z son $(1 - 0.5)/20 = 0.025$, $(2 - 0.5)/20 = 0.075, \ldots,$ y 0.975.

Observación	24.46	25.61	26.25	26.42	26.66	27.15	27.31	27.54	27.74	27.94
Percentil z	−1.96	−1.44	−1.15	−0.93	−0.76	−0.60	−0.45	−0.32	−0.19	−0.06
Observación	27.98	28.04	28.28	28.49	28.50	28.87	29.11	29.13	29.50	30.88
Percentil z	0.06	0.19	0.32	0.45	0.60	0.76	0.93	1.15	1.44	1.96

En la figura 4.33, se ilustra la gráfica de probabilidad normal resultante. El patrón de la gráfica es bastante recto, lo que indica que es razonable que la distribución poblacional de la tensión disruptiva dieléctrica es normal.

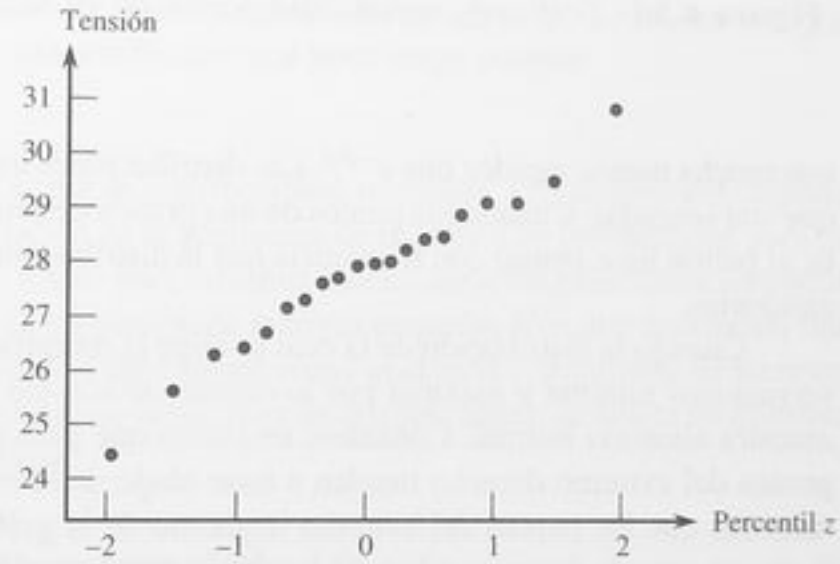


Figura 4.33 Gráfica de probabilidad normal para la muestra de tensión disruptiva dieléctrica ■

Hay otra versión de una gráfica de probabilidad normal en la que un eje de probabilidad no lineal sustituye al eje del percentil z. La escala en este eje se construye de modo que los puntos dibujados caigan de nuevo cerca de una recta cuando la distribución muestreada es normal. En la figura 4.34 de la página siguiente, se muestra una gráfica de MINITAB para los datos de tensión disruptiva del ejemplo 4.29.

Una distribución poblacional no normal con frecuencia se puede colocar en una de las tres categorías siguientes:

1. Es simétrica y tiene "extremos más ligeros" que una distribución normal; es decir, la curva de densidad disminuye con más rapidez en los extremos que una curva normal.
2. Es simétrica y con extremos densos en comparación con la distribución normal.
3. Es sesgada.

Una distribución uniforme tiene extremos ligeros, puesto que su función de densidad disminuye a cero fuera de un intervalo finito. La función de densidad $f(x) = 1/[\pi(1 + x^2)]$ para $-\infty < x < \infty$ es un ejemplo de una distribución con extremos densos, puesto que $1/(1 + x^2)$ disminuye

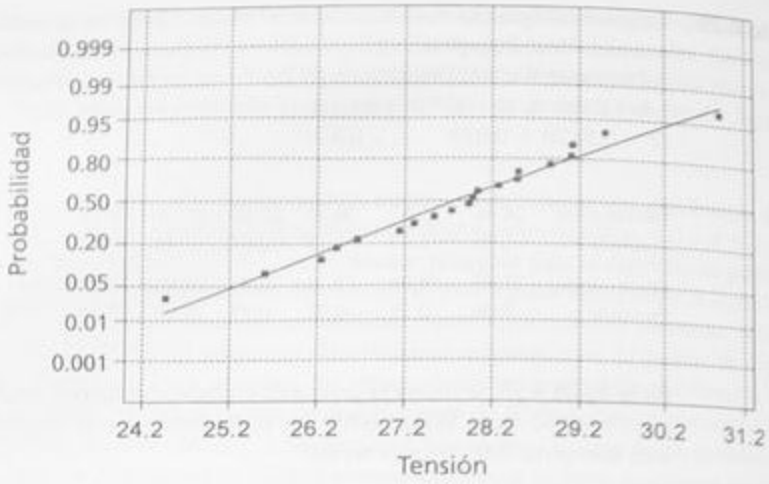


Figura 4.34 Gráfica de probabilidad normal de los datos de tensión disruptiva de MINITAB

con mucha menos rapidez que $e^{-x^2/2}$. Las distribuciones lognormal y de Weibull están entre las que son sesgadas. Cuando los puntos de una gráfica de probabilidad normal no siguen una recta, el patrón hace pensar con frecuencia que la distribución poblacional cae en una de estas tres categorías.

Cuando la distribución de la cual se elige la muestra tiene extremos poco densos, las observaciones mínima y máxima por lo común no son tan extremas como se esperaría de una muestra aleatoria normal. Considere una recta que pasa por la parte media de la gráfica; los puntos del extremo derecho tienden a estar abajo de la recta (valor observado $<$ percentil z), mientras que los puntos del extremo izquierdo de la gráfica tienden a caer arriba de la recta (valor observado $>$ percentil z). El resultado es un patrón en forma de S del tipo ilustrado en la figura 4.32.

Una muestra de una distribución de extremos densos también tiende a producir una gráfica en forma de S. Sin embargo, en contraste con el caso de extremos ligeros, el extremo izquierdo de la gráfica se curva hacia abajo (valor observado $<$ percentil z), como se ilustra en la figura 4.35(a). Si la distribución subyacente tiene sesgo positivo (un extremo izquierdo corto y un extremo derecho largo), las observaciones muestrales más pequeñas serán más grandes que las esperadas de una muestra normal, y lo mismo sucede con las observaciones más grandes. En este caso, los puntos en ambos extremos de la gráfica caen arriba de una recta que pasa por la parte media, produciendo un patrón curvado, como se ilustra en la figura 4.35(b). Una muestra de una distribución lognormal produce por lo común este tipo de patrón. Una gráfica de pares (percentil z, $\ln(x)$) debe parecerse a una recta.

Aun cuando la distribución poblacional es normal, los percentiles muestrales no coincidirán de forma exacta con los percentiles teóricos como resultado de la variabilidad de muestreo. ¿Cuánto se pueden desviar los puntos de la gráfica de probabilidad de un patrón de línea recta antes que ya no sea razonable la suposición de normalidad poblacional? Ésta no es una pregunta fácil de contestar. En términos generales, una muestra pequeña de una distribución normal tiene más probabilidades de producir una gráfica con un patrón no lineal que una muestra grande. En el libro *Fitting Equations to Data* (véase la bibliografía del capítulo 13) presenta los resultados de un estudio de simulación en el que se seleccionaron varias muestras de

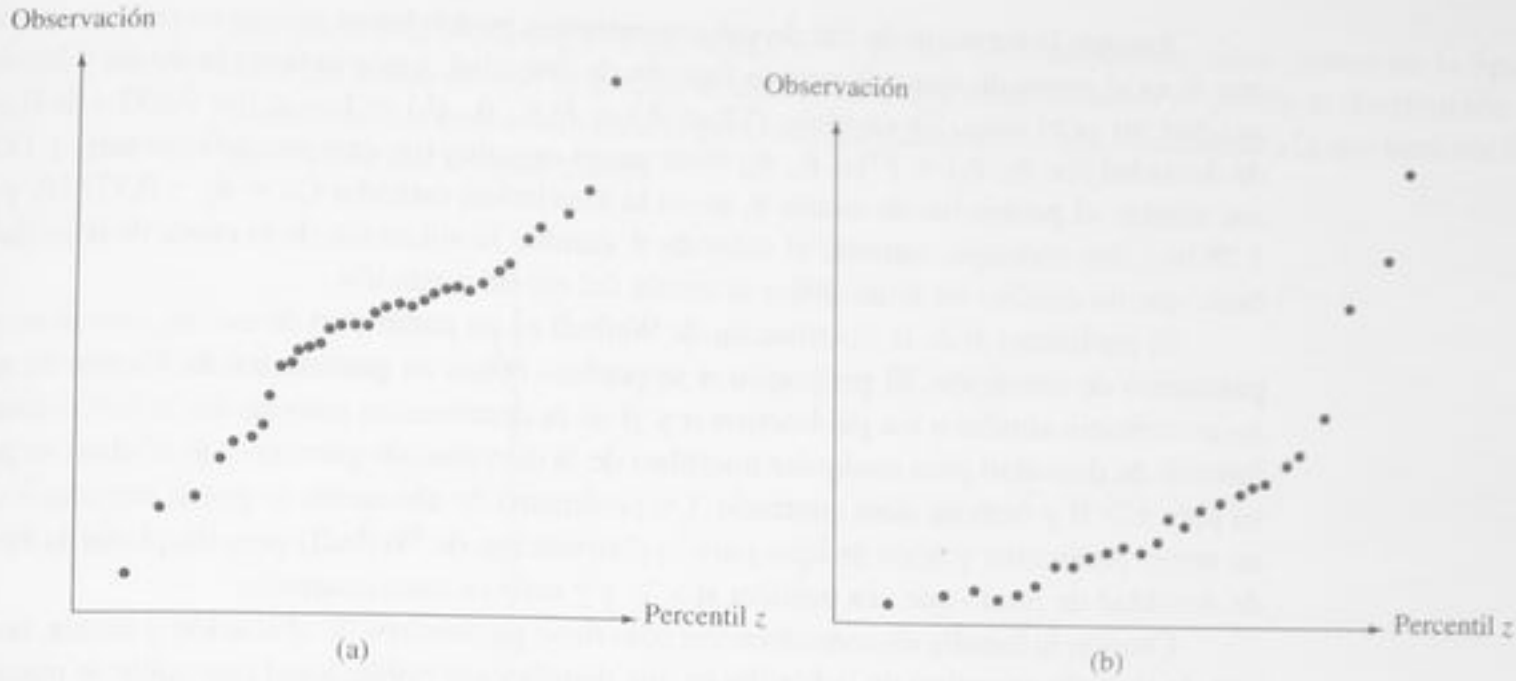


Figura 4.35 Gráficas de probabilidad que indican una distribución no normal: (a) una gráfica consistente con una distribución de extremos densos; (b) una gráfica consistente con una distribución que tiene sesgo positivo

distintos tamaños a partir de distribuciones normales. Los autores concluyeron que por lo común hay mayor variación en la apariencia de la gráfica de probabilidad para tamaños de muestra menores que 30, y sólo para tamaños mucho mayores predomina un patrón lineal. Cuando una gráfica se basa en un tamaño de muestra pequeño, sólo una desviación muy sustancial con respecto a la linealidad se debe tomar como evidencia concluyente de no normalidad. A fin de comprobar la posibilidad de otros tipos de distribuciones, se aplica un comentario similar a las gráficas de probabilidad.

Más allá de la normalidad

Considere una familia de distribuciones de probabilidad que requiere dos parámetros, θ_1 y θ_2, y $F(x; \theta_1, \theta_2)$ representa a las fda correspondientes. La familia de distribuciones normales es una familia de este tipo, con $\theta_1 = \mu$, $\theta_2 = \sigma$, y $F(x; \mu, \sigma) = \Phi[(x - \mu)/\sigma]$. Otro ejemplo es la familia Weibull, con $\theta_1 = \alpha$, $\theta_2 = \beta$, y

$$F(x; \alpha, \beta) = 1 - e^{-(x/\beta)^\alpha}$$

Otra familia más de este tipo es la familia gamma, para la cual la fda es una integral que requiere la función gamma incompleta que no se puede expresar en alguna forma más simple.

Se dice que θ_1 y θ_2 son los **parámetros** de **ubicación** y **escala**, respectivamente, si $F(x; \theta_1, \theta_2)$ es una función de $(x - \theta_1)/\theta_2$. μ y σ de la familia normal son los parámetros de ubicación y escala. Al cambiar μ la ubicación de la curva de densidad en forma de campana se desplaza a la derecha o a la izquierda, y cambiar σ equivale a alargar o comprimir la escala de medición (la escala del eje horizontal cuando se dibuja la función de densidad). La fda es otro ejemplo

$$F(x; \theta_1, \theta_2) = 1 - e^{-e^{(x-\theta_1)/\theta_2}} \qquad -\infty < x < \infty$$

Se dice que una variable aleatoria con esta fda tiene una *distribución de valores extremos*. Se usa en aplicaciones relacionadas con la vida útil de componentes y resistencia de materiales.

Aunque la forma de la fda de valores extremos podría hacer pensar en primera instancia que θ_1 es el punto de simetría para la función de densidad, y por lo tanto la media y la mediana, éste no es el caso. En cambio, $P(X \le \theta_1) = F(\theta_1; \theta_1, \theta_2) = 1 - e^{-1} = 0.632$ y la función de densidad $f(x; \theta_1, \theta_2) = F'(x; \theta_1, \theta_2)$ tiene sesgo negativo (un extremo inferior largo). De forma similar, el parámetro de escala θ_2 no es la desviación estándar ($\mu = \theta_1 - 0.5772\theta_2$ y $\sigma = 1.283\theta_2$). Sin embargo, cambiar el valor de θ_1 cambia la ubicación de la curva de densidad, en tanto que un cambio en θ_2 modifica la escala del eje de medición.

El parámetro β de la distribución de Weibull es un parámetro de escala, pero α no es un parámetro de ubicación. El parámetro α se prefiere como un **parámetro de forma**. Se aplica un comentario similar a los parámetros α y β de la distribución gamma. En la forma usual, la función de densidad para cualquier miembro de la distribución gamma o de Weibull es positiva para $x > 0$ y cero en caso contrario. Un parámetro de ubicación se puede introducir como un tercer parámetro γ (esto se hizo para la distribución de Weibull) para desplazar la función de densidad de modo que sea positiva si $x > \gamma$ y cero en caso contrario.

Cuando la familia en consideración sólo tiene parámetros de ubicación y escala, la cuestión de si algún miembro de la familia es una distribución poblacional razonable se puede tratar vía una sola gráfica de probabilidad construida fácilmente. Primero se obtienen los percentiles de la *distribución estándar*, el que tiene $\theta_1 = 0$ y el que tiene $\theta_2 = 1$, para porcentajes $100(i - 0.5)/n$ $(i = 1, \ldots, n)$. Los n pares (percentil estandarizado, observación) dan los puntos de la gráfica. Por supuesto, esto es exactamente lo que se hizo para obtener una gráfica de probabilidad normal general. Un poco sorprendente es que esta metodología se puede aplicar para obtener una gráfica de probabilidad de Weibull general. El resultado clave es que si X tiene una distribución de Weibull con parámetro de forma α parámetro de escala β, entonces la variable transformada $\ln(X)$ tiene una distribución de valores extremos con parámetro de ubicación $\theta_1 = \ln(\beta)$ y parámetro de escala α. Así, una gráfica de los pares (percentil estandarizado de valor extremo, $\ln(x)$) que muestra un patrón lineal proporciona apoyo para elegir la distribución de Weibull como modelo poblacional.

Ejemplo 4.30 Las observaciones siguientes son acerca de la duración (en horas) del aislamiento de aparatos de potencia cuando la aceleración del esfuerzo térmico y eléctrico se establecieron en valores particulares ("On the Estimation of Life of Power Apparatus Insulation Under Combined Electrical and Thermal Stress". *IEEE Trans. on Electrical Insulation*, 1985: 70-78). Una gráfica de probabilidad de Weibull requiere que se calculen primero los percentiles 5, 15, . . . , y 95 de la distribución estándar de valores extremos. El percentil $(100p)$, $\eta(p)$ satisface

$$p = F(\eta(p)) = 1 - e^{-e^{\eta(p)}}$$

de donde $\eta(p) = \ln[-\ln(1 - p)]$.

Percentil	−2.97	−1.82	−1.25	−0.84	−0.51
x	282	501	741	851	1072
$\ln(x)$	5.64	6.22	6.61	6.75	6.98
Percentil	−0.23	0.05	0.33	0.64	1.10
x	1122	1202	1585	1905	2138
$\ln(x)$	7.02	7.09	7.37	7.55	7.67

Los pares (−2.97, 5.64), (−1.82, 6.22), . . . , (1.10, 7.67) se dibujan como puntos en la figura 4.36. El hecho de que la gráfica se aproxime a una recta favorece el uso de la distribución de Weibull como un modelo para la duración del aislamiento, una conclusión a la que también llegó el autor del artículo citado.

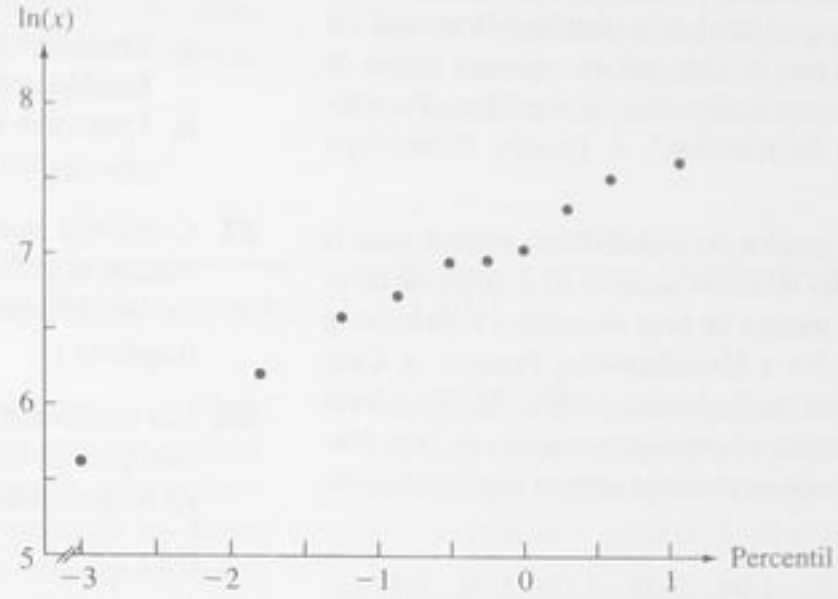
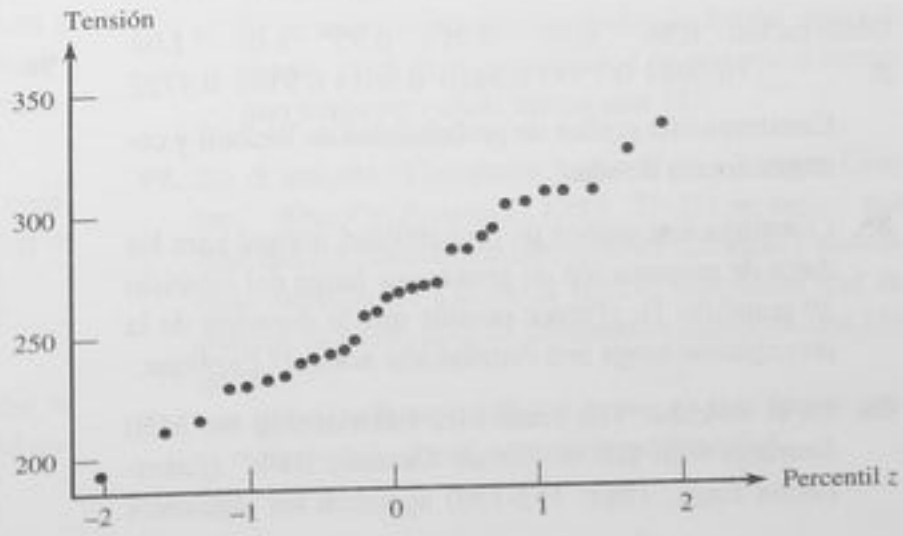


Figura 4.36 Gráfica de probabilidad de Weibull de los datos de vida útil del aislamiento ■

La distribución gamma es un ejemplo de una familia que requiere un parámetro de forma para el cual no hay transformación $h(\cdot)$ tal que $h(X)$ tiene una distribución que depende sólo de los parámetros de ubicación y escala. La construcción de un diagrama de probabilidad requiere estimar primero el parámetro de forma a partir de los datos muestrales (en el capítulo 6 se describen algunos métodos para llevar a cabo lo anterior). En ocasiones un investigador desea saber si la variable transformada X^θ tiene una distribución normal para algún valor de θ (por convención, $\theta = 0$ se identifica con la transformación logarítmica, en cuyo caso X tiene una distribución lognormal). En el libro *Graphical Methods for Data Analysis*, listado en la bibliografía del capítulo 1, se analiza este tipo de problema, así como otros refinamientos de las gráficas de probabilidad.

Ejercicios | Sección 4.6 (81 a 91)

81. La gráfica de probabilidad normal adjunta se construyó a partir de una muestra de 30 lecturas de tensión para mallas detrás de la superficie de cinescopios utilizados en monitores de computadora. ¿Parece factible que la distribución de tensión sea normal?



82. Considere las diez observaciones siguientes acerca de la vida útil de cojinetes (en horas):

152.7	172.0	172.5	173.3	193.0
204.7	216.5	234.9	262.6	422.6

Construya una gráfica de probabilidad normal y comente acerca de la posibilidad de la distribución normal como un modelo para la vida útil de cojinetes (datos de "Modified Moment Estimation for the Three-Parameter Lognormal Distribution", *J. Quality Technology*, 1985: 92-99).

83. Construya una gráfica de probabilidad normal para la siguiente muestra de observaciones de espesor de recubrimiento para pintura de baja viscosidad ("Achieving a Target Value for a Manufacturing Process: A Case Study", *J. of Quality Technology*, 1992: 22-26). ¿Sería conveniente estimar el espesor promedio de la población con un método en el que se supone una distribución normal?

0.83	0.88	0.88	1.04	1.09	1.12	1.29	1.31
1.48	1.49	1.59	1.62	1.65	1.71	1.76	1.83

84. En el artículo "A Probabilistic Model of Fracture in Concrete and Size Effects on Fracture Toughness" (*Magazine of Concrete Res.*, 1996: 311-320) se proporcionan argumentos de por qué la distribución de resistencia a la fractura en especímenes de concreto debe tener una distribución de Weibull y se presentan varios histogramas de los datos que al parecer se ajustan bien mediante curvas de Weibull superpuestas. Considere la muestra siguiente de tamaño $n = 18$ observaciones de resistencia para concreto de alta resistencia (consistente con uno de los histogramas); también se dan los valores de $p_i = (i - 0.5)/18$.

Observación	0.47	0.58	0.65	0.69	0.72	0.74
p_i	0.0278	0.0833	0.1389	0.1944	0.2500	0.3056
Observación	0.77	0.79	0.80	0.81	0.82	0.84
p_i	0.3611	0.4167	0.4722	0.5278	0.5833	0.6389
Observación	0.86	0.89	0.91	0.95	1.01	1.04
p_i	0.6944	0.7500	0.8056	0.8611	0.9167	0.9722

Construya una gráfica de probabilidad de Weibull y comente acerca de ella.

85. Construya una gráfica de probabilidad normal para los datos de propagación de grietas por fatiga del ejercicio 39 (capítulo 1). ¿Parece posible que la duración de la propagación tenga una distribución normal? Explique.

86. En el artículo "The Load-Life Relationship for M50 Bearings with Silicon Nitride Ceramic Balls" (*Lubrication Engr.*, 1984: 153-159) aparecen los siguientes datos sobre la vida útil de cojinetes (millones de revoluciones) probados con una carga de 6.45 kN.

47.1	68.1	68.1	90.8	103.6	106.0	115.0
126.0	146.6	229.0	240.0	240.0	278.0	278.0
289.0	289.0	367.0	385.9	392.0	505.0	

a. Construya una gráfica de probabilidad normal. ¿Es factible la normalidad?

b. Construya una gráfica de probabilidad de Weibull. ¿Es posible la familia de distribución de Weibull?

87. Construya una gráfica de probabilidad que permita evaluar si es factible usar la distribución lognormal como modelo para los datos de lluvia del ejercicio 81 (capítulo 1).

88. Las observaciones siguientes son valores de precipitación pluvial durante marzo, en un periodo de 30 años, en Minneapolis–St. Paul.

0.77	1.20	3.00	1.62	2.81	2.48
1.74	0.47	3.09	1.31	1.87	0.96
0.81	1.43	1.51	0.32	1.18	1.89
1.20	3.37	2.10	0.59	1.35	0.90
1.95	2.20	0.52	0.81	4.75	2.05

a. Construya e interprete una gráfica de probabilidad para este conjunto de datos.

b. Calcule la raíz cuadrada de cada valor y después construya una gráfica de probabilidad normal con base en estos datos transformados. ¿Es factible que la raíz cuadrada de la precipitación tenga una distribución normal?

c. Repita el inciso (b) después de transformar con las raíces cúbicas.

89. Por medio de un paquete de software de estadística, construya una gráfica de probabilidad normal de los datos de resistencia última a la tensión proporcionados en el ejercicio 13 del capítulo 1, y realice comentarios pertinentes.

90. Sean $y_1, y_2, \ldots, y_n$ (observaciones muestrales *ordenadas* (y_1 es la más pequeña y y_n la más grande). A fin de comprobar la normalidad, se dibujan los pares $(\Phi^{-1}((i - 0.5)/n), y_i)$. Suponga que se cree que las observaciones provienen de una distribución con media 0, y sean $w_1, \ldots, w_n$ los valores *absolutos ordenados* de las x_i. Una **gráfica seminormal** es una gráfica de probabilidad de las w_i. De manera más específica, puesto que $P(|Z| \le w) = P(-w \le Z \le w) = 2\Phi(w) - 1$, una gráfica seminormal es un diagrama de los pares $(\Phi^{-1}((i - 0.5)/n) + 1)/2, w_i)$. La ventaja de esta gráfica es que ahora los valores atípicos pequeños o grandes de la muestra

original aparecen sólo en el extremo superior de la gráfica y no en ambos extremos. Construya una gráfica seminormal para la siguiente muestra de errores de medición y comente al respecto: −3.78, −1.27, 1.44, −0.39, 12.38, −43.40, 1.15, −3.96, −2.34, 30.84.

91. Las siguientes observaciones de tiempo de falla (miles de horas) resultaron de la prueba de duración acelerada de 16 circuitos integrados de cierto tipo:

82.8	11.6	359.5	502.5	307.8	179.7
242.0	26.5	244.8	304.3	379.1	212.6
229.9	558.9	366.7	204.6		

Utilice los percentiles correspondientes de la distribución exponencial con $\lambda = 1$ para construir una gráfica de probabilidad. Después, explique por qué la gráfica permite evaluar la factibilidad de que la muestra haya sido generada de *cualquier* distribución exponencial.

Ejercicios complementarios (92 a 120)

92. Sea X = tiempo que una cabeza de lectura y escritura tarda en localizar un registro deseado en un dispositivo de memoria de disco de computadora una vez que la cabeza ha sido colocada en la pista correcta. Si los discos giran una vez cada 25 milisegundos, una suposición razonable es que X tiene una distribución uniforme en el intervalo [0, 25].

a. Calcule $P(10 \le X \le 20)$.
b. Calcule $P(X \ge 10)$.
c. Obtenga la fda $F(X)$.
d. Calcule $E(X)$ y σ_X.

93. Una barra de 12 pulgadas fija en ambos extremos se somete a una creciente cantidad de esfuerzo hasta que se rompe. Sea Y = distancia desde el extremo izquierdo donde ocurre la rotura. Suponga que Y tiene una fdp

$$f(y) = \begin{cases} \left(\frac{1}{24}\right)y\left(1 - \frac{y}{12}\right) & 0 \le y \le 12 \\ 0 & \text{en caso contrario} \end{cases}$$

Calcule lo siguiente:

a. La fda de Y, y dibújela.
b. $P(Y \le 4)$, $P(Y > 6)$ y $P(4 \le Y \le 6)$
c. $E(Y)$, $E(Y^2)$ y $V(Y)$
d. La probabilidad de que el punto de rotura ocurra a más de 2 pulgadas del punto de rotura esperado.
e. La longitud esperada del segmento más corto cuando ocurre la rotura.

94. Sea X el tiempo hasta la falla (en años) de cierto componente hidráulico. Suponga que la fdp de X es $f(x) = 32/(x + 4)^3$ para $x > 0$.

a. Comprueba que $f(x)$ es una fdp legítima.
b. Determine la fda.
c. Utilice el resultado del inciso (b) para calcular la probabilidad de que el tiempo hasta la falla esté entre 2 y 5 años.
d. ¿Cuál es el tiempo esperado hasta la falla?
e. Si el componente tiene un valor de rescate igual a $100/(4 + x)$ cuando su tiempo hasta que se presenta la falla es x, ¿cuál es el valor de rescate esperado?

95. El tiempo de terminación X para cierta tarea tiene una fda $F(x)$ dada por

$$\begin{cases} 0 & x < 0 \\ \frac{x^3}{3} & 0 \le x < 1 \\ 1 - \frac{1}{2}\left(\frac{7}{3} - x\right)\left(\frac{7}{4} - \frac{3}{4}x\right) & 1 \le x \le \frac{7}{3} \\ 1 & x \ge \frac{7}{3} \end{cases}$$

a. Obtenga la fdp $f(x)$ y construya su gráfica.
b. Calcule $P(0.5 \le X \le 2)$.
c. Calcule $E(X)$.

96. Se sabe que la tensión disruptiva de un diodo de cierto tipo seleccionado al azar tiene una distribución normal con valor medio de 40 V y desviación estándar de 1.5 V.

a. ¿Cuál es la probabilidad de que el voltaje de un solo diodo esté entre 39 y 42?
b. ¿Cuál valor es tal que sólo 15% de los diodos tienen voltajes que exceden ese valor?
c. Si se seleccionan cuatro diodos de forma independiente, ¿cuál es la probabilidad de que por lo menos uno tenga un voltaje mayor que 42?

97. En el artículo "Computer Assisted Net Weight Control" (*Quality Progress*, 1983: 22-25) se indica una distribución normal con media de 137.2 onzas y desviación estándar de 1.6 onzas para el contenido real de frascos de cierto tipo. El contenido establecido era 135 onzas.

a. ¿Cuál es la probabilidad de que un solo frasco contenga más que el contenido establecido?

b. Entre diez frascos seleccionados al azar, ¿cuál es la probabilidad de que por lo menos ocho contengan más del contenido indicado?

c. Si se supone que la media permanece en 137.2, ¿a qué valor tendría que cambiar la desviación estándar de modo que 95% de los frascos contengan más de lo establecido?

98. Cuando se prueban tarjetas de circuito utilizadas en la fabricación de reproductores de discos compactos, el porcentaje a largo plazo de tarjetas defectuosas es 5%. Suponga que se recibe un lote de 250 tarjetas y que la condición de cualquier tarjeta es independiente de las demás.

a. ¿Cuál es la probabilidad aproximada de que al menos 10% de las tarjetas del lote estén defectuosas?

b. ¿Cuál es la probabilidad aproximada de que haya exactamente 10 tarjetas defectuosas en el lote?

99. En el artículo "Characterization of Room Temperature Damping in Aluminum-Indium Alloys" (*Metallurgical Trans.*, 1993: 1611-1619) se indica que el tamaño de grano (μm) de la matriz de Al para una aleación que consiste en 2% de Indio se podría modelar con una distribución normal con un valor promedio de 96 y una desviación estándar de 14.

a. ¿Cuál es la probabilidad de que el tamaño de grano sea mayor que 100?

b. ¿Cuál es la probabilidad de que el tamaño de grano esté entre 50 y 80?

c. ¿Cuál intervalo (a, b) incluye al 90% central de los tamaños de grano (de modo que 5% sean menores que a y 5% mayores que b)?

100. El tiempo de reacción (en segundos) a cierto estímulo es una variable aleatoria continua con fdp

$$f(x) = \begin{cases} \frac{3}{2} \cdot \frac{1}{x^2} & 1 \le x \le 3 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Obtenga la fda.

b. ¿Cuál es la probabilidad de que el tiempo de reacción sea a lo sumo 2.5 s? ¿De que esté entre 1.5 y 2.5 s?

c. Calcule el tiempo de reacción esperado.

d. Calcule la desviación estándar del tiempo de reacción.

e. Si un individuo tarda más de 1.5 segundos en reaccionar, se enciende una luz y permanece así hasta que transcurre un segundo más o hasta que reacciona la persona (lo que ocurra primero). Determine el tiempo esperado que permanece encendida la luz. [*Sugerencia:* sea $h(X)$ = tiempo que permanece encendida la luz como función del tiempo de reacción X.]

101. Sea X la temperatura a la que tiene lugar cierta reacción química y suponga que X tiene la siguiente fdp

$$f(x) = \begin{cases} \frac{1}{9}(4 - x^2) & -1 \le x \le 2 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Construya la gráfica de $f(x)$.

b. Determine la fda y trácela.

c. ¿Es 0 la temperatura mediana en la que ocurre la reacción? Si no es así, ¿la temperatura mediana es menor o mayor que 0?

d. Suponga que esta reacción se lleva a cabo de forma independiente una vez en cada uno de diez laboratorios diferentes y que la fdp del tiempo de reacción en cada laboratorio es la que se dio antes. Sea Y = número entre los diez laboratorios en los que la temperatura es mayor que 1. ¿Qué clase de distribución tiene Y? (Dé el nombre y valores de los parámetros.)

102. En el artículo "Determination of the MTF of Positive Photoresists Using the Monte Carlo Method" (*Photographic Sci. and Engr.*, 1983: 254-260) se propone la distribución exponencial con parámetro $\lambda = 0.93$, como modelo para la distribución de una longitud (μm) de trayectoria libre del fotón bajo ciertas circunstancias. Suponga que éste es el modelo correcto.

a. ¿Cuál es la longitud de trayectoria esperada y cuál es la desviación estándar de la longitud de trayectoria?

b. ¿Cuál es la probabilidad de que la longitud de trayectoria sea mayor que 3.0? ¿Cuál es la probabilidad de que la longitud de trayectoria esté entre 1.0 y 3.0?

c. ¿Que valor es excedido por el 10% de todas las longitudes de trayectoria?

103. En el artículo "The Prediction of Corrosion by Statistical Analysis of Corrosion Profiles" (*Corrosion Science*, 1985: 305–315) se indica la siguiente fda para la profundidad X de la picadura más profunda en un experimento que requiere exponer al acero de manganeso al carbono a agua de mar acidulada.

$$F(x; \alpha, \beta) = e^{-e^{-\alpha - \alpha x/\beta}} \qquad -\infty < x < \infty$$

Los autores proponen los valores $\alpha = 150$ y $\beta = 90$. Suponga que éste es el modelo correcto.

a. ¿Cuál es la probabilidad de que la profundidad de la picadura más honda sea a lo sumo 150? ¿A lo sumo 300? ¿De que esté entre 150 y 300?

b. ¿Abajo de qué valor se observa la profundidad de la picadura máxima en 90% de tales experimentos?

c. ¿Cuál es la función de densidad de X?

d. Se puede mostrar que la función de densidad es unimodal (un solo máximo). ¿Arriba de qué valor en el eje de medición ocurre este máximo? (Este valor es la moda.)

e. Se puede demostrar que $E(X) \approx 0.5772\beta + \alpha$. ¿Cuál es la media para los valores de α y β y cómo se compara con la mediana y la moda? Trace la gráfica de la función de densidad. (*Nota*: ésta se llama *distribución del valor extremo más grande*.)

104. La duración X de un componente tiene una distribución exponencial con parámetro λ.

a. Si el costo de operación por unidad de tiempo es c, ¿cuál es el costo esperado de operar este componente en su vida útil?

b. En lugar de un costo constante c como en el inciso (a), suponga que el costo es $c(1 - 0.5e^{ax})$ con $a < 0$, de modo que el costo por unidad de tiempo es menor que c cuando el componente es nuevo y más costoso a medida que envejece el componente. Ahora, calcule el costo esperado de operación durante la vida útil del componente.

105. La *moda* de una distribución continua es el valor x^* que maximiza a $f(x)$.

a. ¿Cuál es la moda de una distribución normal con parámetros μ y σ?

b. ¿La distribución uniforme con parámetros A y B tiene una moda única? ¿Por qué sí o por qué no?

c. ¿Cuál es la moda de una distribución exponencial con parámetro λ? (Dibuje una figura.)

d. Si X tiene una distribución gamma con parámetros α y β, y $\alpha > 1$, encuentre la moda. [*Sugerencia:* $\ln[f(x)]$ se maximiza si y sólo si se maximiza $f(x)$ y podría ser más simple tomar la derivada de $\ln[f(x)]$.]

e. ¿Cuál es la moda de una distribución ji-cuadrada que tiene ν grados de libertad?

106. En el artículo "Error Distribution in Navigation" (*J. Institute of Navigation*, 1971: 429-442) se indica que la distribución de frecuencia de errores positivos (magnitudes de errores) se aproxima bien mediante una distribución exponencial. Sea X = error de posición lateral (millas náuticas), que puede ser negativo o positivo. Suponga que la fdp de X es

$$f(x) = (0.1)e^{-2|x|} \quad -\infty < x < \infty$$

a. Trace una gráfica de $f(x)$ y compruebe que $f(x)$ es una fdp legítima (demuestre que su integral es 1).

b. Obtenga la fda de X y dibújela.

c. Calcule $P(X \le 0)$, $P(X \le 2)$, $P(-1 \le X \le 2)$, y la probabilidad de que se cometa un error de más de dos millas.

107. En algunos sistemas se asigna a un cliente a una de dos instalaciones de servicio. Si el tiempo de servicio para un cliente atendido en la instalación i tiene una distribución exponencial con parámetro λ_i $(i = 1, 2)$ y p es la proporción de los clientes atendidos en la instalación 1, entonces la fdp de X = tiempo de servicio de un cliente seleccionado al azar es

$$f(x; \lambda_1, \lambda_2, p)$$

$$= \begin{cases} p\lambda_1 e^{-\lambda_1 x} + (1-p)\lambda_2 e^{-\lambda_2 x} & x \ge 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Esto recibe a veces el nombre de distribución exponencial mezclada o hiperexponencial. Esta distribución también se propone como modelo para la cantidad de lluvia en "Modeling Monsoon Affected Rainfall of Pakistan by Point Processes" (*J. Water Resources Planning and Mgmnt.*, 1992: 671-688).

a. Compruebe que $f(x; \lambda_1, \lambda_2, p)$ es en realidad fdp.

b. ¿Cuál es la fda $F(x; \lambda_1, \lambda_2, p)$?

c. Si X tiene $f(x; \lambda_1, \lambda_2, p)$ como su fdp, ¿cuál es $E(X)$?

d. Si se utiliza el hecho de que $E(X^2) = 2/\lambda^2$ cuando X tiene una distribución exponencial con parámetro λ, calcule $E(X^2)$ cuando X tiene la fdp $f(x; \lambda_1, \lambda_2, p)$. Después, calcule $V(X)$.

e. El coeficiente de variación de una variable aleatoria (o distribución) es $CV = \sigma/\mu$. ¿Cuál es el CV para una *va* exponencial? ¿Qué puede decir acerca del valor del CV cuando X tiene una distribución hiperexponencial?

f. ¿Cuál es el CV para una distribución de Erlang con parámetros λ y n como se definió en el ejercicio 62? (*Nota:* en el trabajo aplicado, el CV muestral se utiliza para decidir cuál de las tres distribuciones podría ser apropiada.)

108. Suponga que un estado particular permite a las personas presentar declaraciones de renta para desglosar deducciones sólo si el total es por lo menos 5000 dólares. Sea X (en miles de dólares) el total de deducciones desglosadas en una forma seleccionada al azar. Suponga que X tiene la fdp

$$f(x; \alpha) = \begin{cases} k/x^{\alpha} & x \ge 5 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Estime el valor de k. ¿Cuál restricción sobre α es necesaria?

b. ¿Cuál es la fda de X?

c. ¿Cuál es la deducción total esperada en una forma seleccionada al azar? ¿Cuál restricción sobre α es necesaria para que $E(X)$ sea finita?

d. Demuestre que $\ln(X/5)$ tiene una distribución exponencial con parámetro $\alpha = 1$.

109. Sea I_i la corriente de entrada a un transistor e I_o la corriente de salida. Entonces la ganancia de corrientes es proporcional a $\ln(I_o/I_i)$. Suponga que la constante de proporcionalidad es 1 (lo que equivale a seleccionar una unidad particular de medida), así que la ganancia de corriente es $= X = \ln(I_o/I_i)$. Suponga que X tiene una distribución normal con $\mu = 1$ y $\sigma = 0.05$.

a. ¿Qué tipo de distribución tiene la relación I_o/I_i?

b. ¿Cuál es la probabilidad de que la corriente de salida sea más del doble de la corriente de entrada?

c. ¿Cuáles son el valor esperado y la varianza de la relación entre la corriente de salida y la corriente de entrada?

110. En el artículo "Response of SiC_f/Si_3N_4 Composites Under Static and Cyclic Loading—An Experimental and Statistical Analysis" (*J. of Engr. Materials and Technology*, 1997: 186-193) se sugiere que la resistencia a la tensión en (MPa) de materiales compuestos bajo las condiciones especificadas se puede modelar mediante una distribución de Weibull con $\alpha = 9$ y $\beta = 180$.

a. Trace una gráfica de la función de densidad.

b. ¿Cuál es la probabilidad de que la resistencia de un espécimen seleccionado al azar sea mayor de 175? ¿De que esté entre 150 y 175?

c. Si se escogen al azar dos especímenes y sus resistencias son independientes entre sí, ¿cuál es la probabilidad de que al menos uno tenga resistencia entre 150 y 175?

d. ¿Qué valor de resistencia separa al 10% más débil de especímenes del 90% restante?

111. Considere que Z tiene una distribución normal estándar y defina una nueva *va* Y mediante $Y = \sigma Z + \mu$. Demuestre que Y tiene una distribución normal con parámetros μ y σ. (*Sugerencia:* $Y \le y$ si y sólo si $Z \le ?$ Utilice esto para encontrar la fda de Y; luego, derive con respecto a y.)

112. **a.** Suponga que la vida útil X de un componente, cuando se mide en horas, tiene una distribución gamma con parámetros α y β. Sea $Y =$ vida útil medida en minutos. Obtenga la fdp de Y. (*Sugerencia:* $Y \le y$ si y sólo si $X \le y/60$. Utilice esto para obtener la fda de Y; luego, derive para obtener la fdp.)

b. Si X tiene una distribución gamma con parámetros α y β, ¿cuál es la distribución de probabilidad de $Y = cX$?

113. En los ejercicios 111 y 112, así como en otras situaciones, se tiene la fdp $f(x)$ de X y desea conocer la fdp de $Y = h(X)$. Suponga que $h(\cdot)$ es una función invertible, así que $y = h(x)$ se puede resolver para x y obtener $x = k(y)$. Entonces, se puede demostrar que la fdp de Y es

$$g(y) = f[k(y)] \cdot |k'(y)|$$

a. Si X tiene una distribución uniforme con $A = 0$ y $B = 1$, obtenga la fdp de $Y = -\ln(X)$.

b. Resuelva el ejercicio 111 usando este resultado.

c. Resuelva el ejercicio 112(b) con este resultado.

114. Con base en los datos de un experimento de lanzar dardos, en el artículo "Shooting Darts" (*Chance*, Summer 1997, 16-19) se propone que los errores horizontal y vertical de apuntar en un objetivo puntual deben ser independientes entre sí, cada uno con una distribución normal con media 0 y varianza σ^2. Entonces, se puede demostrar que la fdp de la distancia V desde el objetivo al punto de aterrizaje es

$$f(v) = \frac{v}{\sigma^2} \cdot e^{-v^2/2\sigma^2} \quad v > 0$$

a. ¿De qué familia presentada en este capítulo es un miembro esta fdp?

b. Si $\sigma = 20$ mm (cercana al valor sugerido en el artículo), ¿cuál es la probabilidad de que el dardo aterrice dentro de 25 mm (aproximadamente una pulg.) del objetivo?

115. En el artículo "Three Sisters Give Birth on the Same Day" (*Chance*, Spring 2001, 23-25) se utiliza el hecho de que tres hermanas de Utah dieron a luz el 11 de marzo de 1998 como base para plantear algunas preguntas interesantes en relación con coincidencias de nacimiento.

a. Sin tomar en cuenta el año bisiesto, y suponiendo que los otros 365 días tienen las mismas probabilidades, ¿cuál es la probabilidad de que tres nacimientos elegidos al azar ocurran el 11 de marzo? Asegúrese de indicar, si hay alguna, qué suposiciones extra hace.

b. Con las suposiciones utilizadas en el inciso (a), ¿cuál es la probabilidad de que en el mismo día ocurran tres nacimientos elegidos al azar?

c. El autor sugirió que, con base en los datos existentes, la duración de la gestación (tiempo entre la concepción y el nacimiento) se podría modelar como una que tiene una distribución normal con valor medio de 280 días y desviación estándar de 19.88 días. Las fechas límite para las tres hermanas de Utah fueron 15 de marzo, 1 de abril y 4 de abril, respectivamente. Si se supone que las tres fechas límite están en la media de la distribución, ¿cuál es la probabilidad de que los tres nacimientos ocurrieran el 11 de marzo? (*Sugerencia:* la desviación de la fecha de nacimiento con respecto a la fecha límite tiene una distribución normal con media 0.)

d. Explique cómo se usaría la información del inciso (c) para calcular la probabilidad de una fecha de nacimiento común.

116. Sea X la vida útil de un componente, con $f(x)$ y $F(x)$ la fdp y la fda de X. La probabilidad de que el componente falle en el intervalo $(x, x + \Delta x)$ es aproximadamente $f(x) \cdot \Delta x$. La probabilidad condicional de que falle en $(x, x + \Delta x)$ ya que ha durado por lo menos x es $f(x) \cdot \Delta x/[1 - F(x)]$. Al dividir esto entre Δx se obtiene la **función de número de fallas**:

$$r(x) = \frac{f(x)}{1 - F(x)}$$

Una función de número de fallas creciente indica que hay cada vez más probabilidades de que los componentes más viejos se deterioren, en tanto que un número de fallas decreciente es evidencia de confiabilidad creciente con la edad. En la práctica suele suponerse una función de número de fallas en "forma de tina bañera".

a. Si X tiene una distribución exponencial, ¿cuál es $r(x)$?

b. Si X tiene una distribución de Weibull con parámetros α y β, ¿cuál es $r(x)$? ¿Para cuáles valores de los parámetros será creciente $r(x)$? ¿Para cuáles valores de los parámetros disminuirá $r(x)$ con x?

c. Puesto que $r(x) = -(d/dx)\ln[1 - F(x)]$, $\ln[1 - F(x)] = -\int r(x)\,dx$. Suponga que

$$r(x) = \begin{cases} \alpha\left(1 - \dfrac{x}{\beta}\right) & 0 \le x \le \beta \\ 0 & \text{en caso contrario} \end{cases}$$

de modo que si un componente dura β horas, durará por siempre (si bien esto al parecer es ilógico, este modelo se puede emplear para estudiar sólo el "desgaste inicial"). ¿Cuáles son la fda y la fdp de X?

117. Sea que U tiene una distribución uniforme en el intervalo [0, 1]. Entonces, los valores observados que tiene esta distribución se pueden obtener del generador de números aleatorios de una computadora. Sea $X = -(1/\lambda)\ln(1 - U)$.

a. Demuestre que X tiene una distribución exponencial con parámetro λ. [*Sugerencia:* la fda de X es $F(x) = P(X \le x)$; $X \le x$ equivalente a $U \le$?]

b. ¿Cómo se utilizaría el inciso (a) y un generador de números aleatorios para obtener los valores observados de una distribución exponencial con parámetro $\lambda = 10$?

118. Considere una *va* X con media μ y desviación estándar σ, y sea $g(X)$ una función específica de X. La aproximación según la serie de Taylor de primer orden a $g(X)$ en la vecindad de μ es

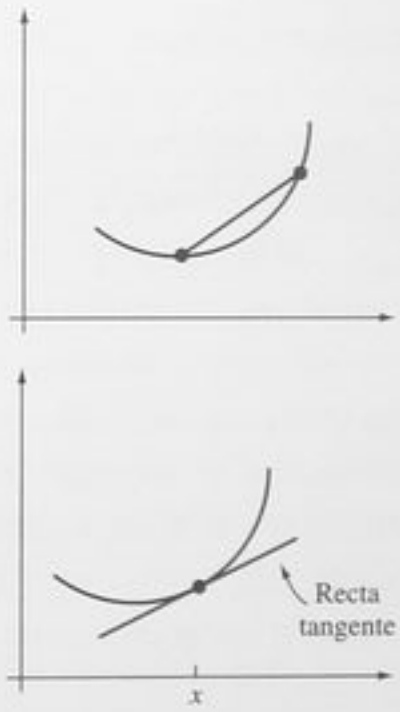
$$g(X) \approx g(\mu) + g'(\mu) \cdot (X - \mu)$$

El lado derecho de esta ecuación es una función lineal de X. Si la distribución de X está concentrada en un intervalo sobre el que $g(\cdot)$ es aproximadamente lineal [p. ej. $\sqrt{x}$ es aproximadamente lineal en (1, 2)], entonces la ecuación produce aproximaciones a $E(g(X))$ y $V(g(X))$.

a. Dé expresiones para estas aproximaciones. (*Sugerencia:* utilice las reglas del valor esperado y la varianza para una función lineal $aX + b$.)

b. Si el voltaje v en un medio es fijo, pero la corriente I es aleatoria, entonces la resistencia también será una variable aleatoria relacionada con I por medio de $R = v/I$. Si $\mu_I = 20$ y $\sigma_I = 0.5$, calcule aproximaciones para μ_R y σ_R.

119. Una función $g(x)$ es *convexa* si la cuerda que une dos puntos cualesquiera en la gráfica de la función queda arriba de la gráfica. Cuando $g(x)$ es diferenciable, una condición equivalente es que para toda x, la recta tangente en x queda por completo sobre o debajo de la gráfica (véanse las figuras siguientes). ¿Cómo se compara $g(\mu) = g(E(X))$ con $E(g(X))$? [*Sugerencia:* la ecuación de la recta tangente en $x = \mu$ es $y = g(\mu) + g'(\mu) \cdot (x - \mu)$. Utilice la condición de convexidad, sustituya a x con X y tome valores esperados. *Nota:* a menos que $g(x)$ sea lineal, la desigualdad resultante (que por lo general es llamada desigualdad de Jensen) es estricta ($<$ en lugar de $\le$); es válida tanto para *va* continuas como para discretas.]



120. Considere que X tiene una distribución de Weibull con parámetros $\alpha = 2$ y β. Demuestre que $Y = 2X^2/\beta^2$ tiene una distribución ji-cuadrada con $\nu = 2$. [*Sugerencia:* la fda de Y es $P(Y \le y)$; exprese esta probabilidad en la forma $P(X \le g(y))$, utilice el hecho de que la fda de X es de la forma de la expresión (4.12) y derive con respecto a y para obtener la fdp de Y.]

Bibliografía

Bury, Karl, *Statistical Distributions in Engineering,* Cambridge Univ. Press, Cambridge, Inglaterra, 1999. Un estudio legible e informativo de las distribuciones y sus propiedades.

Johnson, Norman, Samuel Kotz y N. Balakrishnan *Continuous Univariate Distributions,* vols. 1-2, Wiley, Nueva York, 1994. Estos dos volúmenes presentan un estudio detallado de varias distribuciones continuas.

Nelson, Wayne, *Applied Life Data Analysis,* Wiley, Nueva York, 1982. Proporciona un análisis completo de las distribuciones y métodos que se emplean en el análisis de datos de vida útil.

Olkin, Ingram, Cyrus Derman y Leon Gleser, *Probability Models and Applications* (2a. ed.), Macmillan, Nueva York, 1994. Buena información acerca de propiedades generales y distribuciones específicas.

5

Distribuciones de probabilidad conjunta y muestras aleatorias

Introducción

En los capítulos 3 y 4, se estudiaron modelos de probabilidad para una sola variable aleatoria. Muchos problemas de probabilidad y estadística, dan lugar a modelos en los que intervienen al mismo tiempo varias variables aleatorias. En este capítulo, se analizan primero modelos de probabilidad para el comportamiento conjunto de varias variables aleatorias, poniendo énfasis especial en el caso en el que las variables son independientes entre sí. Luego, se estudian valores esperados de funciones de varias variables aleatorias, entre otras la covarianza y la correlación, como medidas del grado de asociación entre dos variables.

En las últimas tres secciones del capítulo, se consideran funciones de n variables aleatorias $X_1, X_2, \ldots, X_n$, centrando en particular la atención en su promedio $(X_1 + \cdots + X_n)/n$. A cualquier función de esta clase, que por sí misma es una variable aleatoria, se le llama *estadístico*. Los resultados de la probabilidad se utilizan para obtener información acerca de la distribución de un estadístico. El resultado principal de este tipo es el Teorema del Límite Central (TLC), la base para muchos procedimientos de inferencia relacionados con muestras grandes.

5.1 Variables aleatorias con distribución conjunta

Hay muchas situaciones experimentales en las que un investigador se interesa en más de una variable aleatoria (*va*). Primero se consideran las distribuciones de probabilidad conjunta para dos *va* discretas, luego para dos variables continuas y, por último, para más de dos variables.

Función masa de probabilidad conjunta para dos variables aleatorias discretas

La función masa de probabilidad (fmp) de una sola *va* discreta X, especifica cuánta probabilidad se coloca en cada valor posible de X. La fmp conjunta de dos *va* discretas X y Y describe cuánta probabilidad se coloca en cada par posible de valores (x, y).

DEFINICIÓN

Sean X y Y dos *va* discretas definidas en el espacio muestral $\mathcal{S}$ de un experimento. La **función masa de probabilidad conjunta** $p(x, y)$ se define para cada par de números (x, y) mediante

$$p(x, y) = P(X = x \text{ y } Y = y)$$

Sea A cualquier conjunto que consiste en pares de valores (x, y). Entonces la probabilidad $P[(X, Y) \in A]$ se obtiene al sumar la fmp conjunta en los pares de A:

$$P[(X, Y) \in A] = \sum_{(x, y) \in A} \sum p(x, y)$$

Ejemplo 5.1 Una gran agencia de seguros atiende a varios clientes que compraron una póliza para propietario de casa y una póliza de automóvil. Para cada tipo de póliza, se debe especificar una cantidad deducible. Para una póliza de automóvil, las opciones son \$100 y \$250, mientras que para una póliza de propietario de casa, las opciones son 0, \$100 y \$200. Suponga que de los archivos de la agencia se elige un individuo al azar que tiene ambos tipos de póliza. Sea X = cantidad deducible en la póliza del automóvil y Y = cantidad deducible de la póliza del dueño de casa. Los pares posibles (X, Y) son (100, 0), (100, 100), (100, 200), (250, 0), (250, 100), y (250, 200); la fmp especifica la probabilidad relacionada con cada uno de estos pares, con cualquier otro par que tiene probabilidad cero. Supóngase que la fmp conjunta se da en la siguiente **tabla de probabilidad conjunta:**

$p(x, y)$		y = 0	100	200
x	100	0.20	0.10	0.20
	250	0.05	0.15	0.30

Entonces $p(100, 100) = P(X = 100 \text{ y } Y = 100) = P(\$100 \text{ de deducible en ambas pólizas}) = 0.10$. La probabilidad $P(Y \geq 100)$ se calcula al suponer probabilidades de los pares (x, y) para los cuales $y \geq 100$:

$$\begin{aligned} P(Y \geq 100) &= p(100, 100) + p(250, 100) + p(100, 200) + p(250, 200) \\ &= 0.75 \end{aligned}$$

■

Una función $p(x, y)$ se puede usar como una fmp conjunta siempre que $p(x, y) \geq 0$ para toda x y y, y $\sum_x \sum_y p(x, y) = 1$.

La fmp de una de las variables se obtiene al sumar $p(x, y)$ en los valores de la otra variable. El resultado se llama *fmp marginal* porque cuando los valores $p(x, y)$ aparecen en una tabla rectangular, las sumas son totales marginales (renglón o columna).

DEFINICIÓN

Las **funciones masa de probabilidad marginal** de X y de Y, representadas por $p_X(x)$ y $p_Y(y)$, respectivamente, están dadas por

$$p_X(x) = \sum_y p(x, y) \qquad p_Y(y) = \sum_x p(x, y)$$

Así que para obtener la fmp marginal de X evaluada en, por ejemplo, $x = 100$, las probabilidades $p(100, y)$ se suman a los valores posibles de y. Al hacer esto para cada valor posible de X, se obtiene sólo la fmp marginal de X (sin referencia a Y). De las fmp marginales, se pueden calcular las probabilidades de eventos en los que interviene sólo X o sólo Y.

Ejemplo 5.2 (continuación del ejemplo 5.1)

Los valores posibles de X son $x = 100$ y $x = 250$, así que calculando los totales de renglón en la tabla de probabilidad conjunta, se obtiene

$$p_X(100) = p(100, 0) + p(100, 100) + p(100, 200) = 0.50$$

y

$$p_X(250) = p(250, 0) + p(250, 100) + p(250, 200) = 0.50$$

Entonces, la fmp marginal de X es

$$p_X(x) = \begin{cases} 0.5 & x = 100, 250 \\ 0 & \text{en caso contrario} \end{cases}$$

De manera similar, la fmp marginal de Y se obtiene de los totales de columna como

$$p_Y(y) = \begin{cases} 0.25 & y = 0, 100 \\ 0.50 & y = 200 \\ 0 & \text{en caso contrario} \end{cases}$$

de modo que $P(Y \geq 100) = p_Y(100) + p_Y(200) = 0.75$ como antes. ■

Función de densidad de probabilidad conjunta para dos variables aleatorias continuas

La probabilidad de que el valor observado de una *va* continua X esté en un conjunto unidimensional A (p. ej., un intervalo), se obtiene integrando la fdp $f(x)$ sobre el conjunto A. De manera similar, la probabilidad de que el par (X, Y) de *va* continuas se sitúe en un conjunto A bidimensional (p. ej., un rectángulo) se obtiene integrando una función llamada *función de densidad conjunta.*

DEFINICIÓN

Sean X y Y variables aleatorias continuas. Entonces $f(x, y)$ es la **función de densidad de probabilidad conjunta** para X y Y si para cualquier conjunto bidimensional A

$$P[(X, Y) \in A] = \iint_A f(x, y)\, dx\, dy$$

En particular, si A es el rectángulo bidimensional $\{(x, y): a \le x \le b, c \le y \le d\}$, entonces

$$P[(X, Y) \in A] = P(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f(x, y)\, dy\, dx$$

Para que $f(x, y)$ sea una función candidata para una fdp conjunta, se debe satisfacer que $f(x, y) \ge 0$ y $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$. Se puede considerar que $f(x, y)$ especifica una superficie a la altura $f(x, y)$ arriba del punto (x, y) en un sistema coordenado tridimensional. Entonces $P[(X, Y) \in A]$ es el volumen debajo de esta superficie y arriba de la región A, análogo al área bajo la curva en el caso de una dimensión. Esto se ilustra en la figura 5.1.

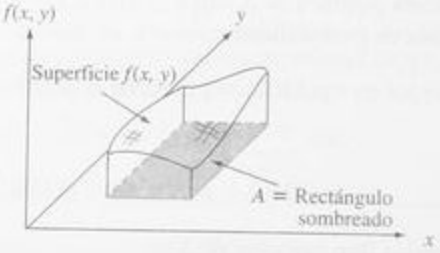


Figura 5.1 $P[(X, Y) \in A]$ = volumen debajo de la superficie de densidad arriba de A

Ejemplo 5.3 Un banco cuenta con una ventanilla para atender a clientes en su automóvil y otra para clientes a pie. En un día elegido al azar, sea X = parte del tiempo que está en uso la ventanilla para atender a los clientes en su automóvil (por lo menos un cliente está siendo atendido o espera ser atendido) y Y = parte del tiempo que está en uso la ventanilla para clientes a pie. Entonces el conjunto de valores posibles para (X, Y) es el rectángulo $D = \{(x, y): 0 \le x \le 1, 0 \le y \le 1\}$. Suponga que la fdp conjunta de (X, Y) está dada por

$$f(x, y) = \begin{cases} \frac{6}{5}(x + y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Para comprobar que ésta es una fdp legítima, observe que $f(x, y) \ge 0$ y

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy &= \int_0^1\int_0^1 \frac{6}{5}(x + y^2)\, dx\, dy \\ &= \int_0^1\int_0^1 \frac{6}{5} x\, dx\, dy + \int_0^1\int_0^1 \frac{6}{5} y^2\, dx\, dy \\ &= \int_0^1 \frac{6}{5} x\, dx + \int_0^1 \frac{6}{5} y^2\, dy = \frac{6}{10} + \frac{6}{15} = 1 \end{aligned}$$

La probabilidad de que ninguna ventanilla está ocupada más de un cuarto del tiempo es

$$P\left(0 \le X \le \frac{1}{4}, 0 \le Y \le \frac{1}{4}\right) = \int_0^{1/4}\int_0^{1/4} \frac{6}{5}(x + y^2)\,dx\,dy$$

$$= \frac{6}{5}\int_0^{1/4}\int_0^{1/4} x\,dx\,dy + \frac{6}{5}\int_0^{1/4}\int_0^{1/4} y^2\,dx\,dy$$

$$= \frac{6}{20}\cdot\frac{x^2}{2}\Bigg|_{x=0}^{x=1/4} + \frac{6}{20}\cdot\frac{y^3}{3}\Bigg|_{y=0}^{y=1/4} = \frac{7}{640}$$

$$= .0109$$

■

Al igual que con las fmp, a partir de la fdp de X y Y, se pueden calcular cada una de las dos funciones de densidad marginal.

DEFINICIÓN

Las **funciones de densidad de probabilidad marginal** de X y Y, denotadas por $f_X(x)$ y $f_Y(y)$, respectivamente, están dadas por

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy \qquad \text{para } -\infty < x < \infty$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx \qquad \text{para } -\infty < y < \infty$$

Ejemplo 5.4 (continuación del ejemplo 5.3)

La fdp marginal de X, que da la distribución de probabilidad del tiempo ocupado para la ventanilla de clientes en automóvil sin referencia a la ventanilla para clientes a pie, es

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy = \int_0^1 \frac{6}{5}(x + y^2)\,dy = \frac{6}{5}x + \frac{2}{5}$$

para $0 \le x \le 1$ y 0 en caso contrario. La fdp marginal de Y es

$$f_Y(y) = \begin{cases} \frac{6}{5}y^2 + \frac{3}{5} & 0 \le y \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Entonces,

$$P\left(\frac{1}{4} \le Y \le \frac{3}{4}\right) = \int_{1/4}^{3/4} f_Y(y)\,dy = \frac{37}{80} = 0.4625$$

■

En el ejemplo 5.3, la región de densidad conjunta positiva fue el rectángulo, que hizo relativamente fácil el cálculo de las fdp marginales. Ahora, considere un ejemplo en el que la región de densidad positiva es una figura más complicada.

Ejemplo 5.5

Una compañía vende latas de nueces mixtas de lujo que contienen almendras, castañas y maní. Suponga que el peso neto de cada lata es exactamente 1 lb, pero la contribución en peso de cada tipo de nuez es aleatoria. Debido a que los tres pesos suman 1, un modelo de probabilidad conjunta para dos tipos de fruto da la información necesaria acerca del peso del tercer tipo. Sea X = peso de las almendras es una lata seleccionada y Y = peso de las castañas. En-

tonces la región de densidad positiva es $D = \{(x, y): 0 \le x \le 1, 0 \le y \le 1, x + y \le 1\}$, la región sombreada ilustrada en la figura 5.2.

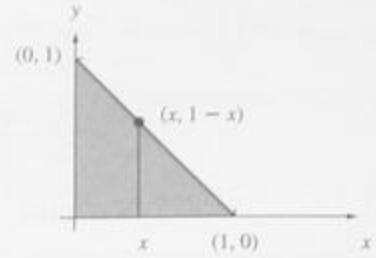


Figura 5.2 Región de densidad positiva para el ejemplo 5.5

Ahora, sea que la fdp conjunta para (X, Y) es

$$f(x, y) = \begin{cases} 24xy & 0 \le x \le 1, 0 \le y \le 1, x + y \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Para cualquier x fija, $f(x, y)$ se incrementa con y; para y fija, $f(x, y)$ se incrementa con x. Esto es apropiado porque la palabra *de lujo* significa que la mayor parte de la lata debe consistir en almendras y castañas en vez de maní, así que la función de densidad debe ser grande cerca del límite superior y pequeña cerca del origen. La superficie determinada por $f(x, y)$ se inclina hacia arriba desde cero cuando (x, y) se aleja desde cualquier eje.

Es evidente que $f(x, y) \ge 0$. Para comprobar la segunda condición de una fdp conjunta, recuerde que una integral doble se calcula como una integral iterada manteniendo fija una variable (p. ej., x en la figura 5.2), integrando sobre valores de la otra variable que yacen a lo largo de la recta que pasa por el valor de la variable fija y, por último, integrando en los valores posibles de la variable fija. Por lo tanto,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dy\, dx = \int_D\int f(x, y)\, dy\, dx = \int_0^1 \left\{\int_0^{1-x} 24xy\, dy\right\} dx$$

$$= \int_0^1 24x \left\{\frac{y^2}{2}\Bigg|_{y=0}^{y=1-x}\right\} dx = \int_0^1 12x(1 - x)^2\, dx = 1$$

Para calcular la probabilidad de que dos tipos de nueces juntos constituyen a lo sumo 50% de la lata, sea $A = \{(x, y): 0 \le x \le 1, 0 \le y \le 1 \text{ y } x + y \le 0.5\}$, como se muestra en la figura 5.3. Entonces,

$$P((X, Y) \in A) = \int_A\int f(x, y)\, dx\, dy = \int_0^{0.5}\int_0^{0.5-x} 24xy\, dy\, dx = 0.0625$$

La fdp para las almendras se obtiene manteniendo fija X en x e integrando $f(x, y)$ a lo largo de la recta vertical que pasa por x:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \begin{cases} \displaystyle\int_0^{1-x} 24xy\, dy = 12x(1 - x)^2 & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

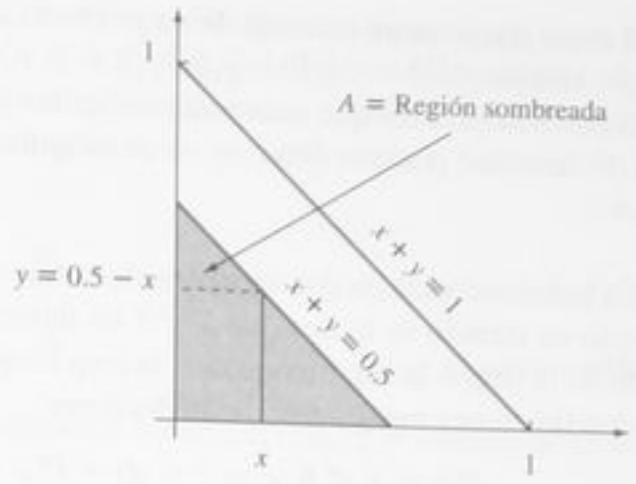


Figura 5.3 Cálculo de $P[(X, Y) \in A]$ para el ejemplo 5.5

Por simetría de $f(x, y)$ y la región D, la fdp marginal de Y se obtiene al sustituir x y X en $f_X(x)$ por y y Y, respectivamente. ■

Variables aleatorias independientes

En muchas situaciones, la información acerca del valor observado de una de las dos variables X y Y proporciona información acerca del valor de la otra variable. En el ejemplo 5.1, la probabilidad marginal de X en $x = 250$ fue 0.5, como lo fue la probabilidad de que $X = 100$. Sin embargo, si se sabe que $Y = 0$ para el individuo elegido, entonces $X = 100$ es cuatro veces tan probable como $X = 250$. Por consiguiente, hay dependencia entre las dos variables.

En el capítulo 2 se señaló que una forma de definir la independencia de dos eventos es señalar que A y B son independientes si $P(A \cap B) = P(A) \cdot P(B)$. A continuación se da una definición análoga para la independencia de dos variables aleatorias.

DEFINICIÓN

Se dice que dos variables aleatorias X y Y son **independientes** si para cada par de valores x y y,

$$p(x, y) = p_X(x) \cdot p_Y(y) \qquad \text{cuando } X \text{ y } Y \text{ son discretas}$$

o bien,

$$f(x, y) = f_X(x) \cdot f_Y(y) \qquad \text{cuando } X \text{ y } Y \text{ son continuas} \tag{5.1}$$

Si no se satisface (5.1) para todo par (x, y), entonces se dice que X y Y son **dependientes.**

La definición establece que dos variables son independientes si su fdp conjunta es el producto de dos fmp o fdp marginales.

Ejemplo 5.6 En la situación del seguro de los ejemplos 5.1 y 5.2,

$$p(100, 100) = 0.10 \neq (0.5)(0.25) = p_X(100) \cdot p_Y(100)$$

así que X y Y no son independientes. La independencia de X y Y requiere que *todo* elemento de la tabla de probabilidad conjunta sea el producto de las probabilidades marginales del renglón y columna correspondientes. ■

Ejemplo 5.7 (continuación del ejemplo 5.5)

Debido a que $f(x, y)$ tiene la forma de un producto, X y Y al parecer serían independientes. Sin embargo, aunque $f_X(\frac{3}{4}) = f_Y(\frac{3}{4}) = \frac{9}{16}, f(\frac{3}{4}, \frac{3}{4}) = 0 \neq \frac{9}{16} \cdot \frac{9}{16}$, entonces las variables de hecho no son independientes. Para que sean independientes, $f(x, y)$ debe tener la forma $g(x) \cdot h(y)$ y la región de densidad positiva debe ser un rectángulo cuyos lados son paralelos a los ejes coordenados. ∎

La independencia de dos variables aleatorias es más útil cuando en la descripción del experimento en estudio se indica que X y Y no tienen efecto entre sí. Entonces, una vez que se especifican la fmp o la fdp marginales, la fmp conjunta, o la fdp, es simplemente el producto de las dos funciones marginales. Se deduce que

$$P(a \le X \le b, c \le Y \le d) = P(a \le X \le b) \cdot P(c \le Y \le d)$$

Ejemplo 5.8

Suponga que los tiempos de vida de dos componentes son independientes entre sí, y que el primer tiempo de vida, X_1, tiene una distribución exponencial con parámetro λ_1 en tanto que el segundo, X_2, tiene una distribución exponencial con parámetro λ_2. Entonces la fdp conjunta es

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

$$= \begin{cases} \lambda_1 e^{-\lambda_1 x_1} \cdot \lambda_2 e^{-\lambda_2 x_2} = \lambda_1 \lambda_2 e^{-\lambda_1 x_1 - \lambda_2 x_2} & x_1 > 0, x_2 > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Sea $\lambda_1 = 1/1000$ y $\lambda_2 = 1/1200$, así que los tiempos de vida esperados son 1000 y 1200 horas, respectivamente. La probabilidad de que ambos tiempos de vida sean por lo menos 1500 horas es

$$\begin{aligned} P(1500 \le X_1, 1500 \le X_2) &= P(1500 \le X_1) \cdot P(1500 \le X_2) \\ &= e^{-\lambda_1(1500)} \cdot e^{-\lambda_2(1500)} \\ &= (0.2231)(0.2865) = 0.0639 \end{aligned}$$ ∎

Más de dos variables aleatorias

Para modelar el comportamiento conjunto de más de dos variables aleatorias, se amplía el concepto de distribución conjunta de dos variables.

DEFINICIÓN

Si $X_1, X_2, \ldots, X_n$ son variables aleatorias discretas, la fmp conjunta de las variables es la función

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

Si las variables son continuas, la fdp conjunta de $X_1, \ldots, X_n$ es la función $f(x_1, x_2, \ldots, x_n)$ tal que para n intervalos cualesquiera $[a_1, b_1], \ldots, [a_n, b_n]$,

$$P(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

En un experimento binomial, cada ensayo podría producir uno de sólo dos resultados posibles. Considere un experimento que consiste en n ensayos independientes e idénticos, en el que cada ensayo produce cualquiera de r resultados posibles. Sea $p_i = P(\text{resultado } i \text{ en cualquier ensayo particular})$, y defina variables aleatorias mediante $X_i =$ número de ensayos que

producen el resultado i ($i = 1, \ldots, r$). Este experimento se llama **experimento multinomial** y la fmp conjunta de $X_1, \ldots, X_r$ se llama **distribución multinomial.** Mediante un argumento de conteo análogo utilizado para obtener la distribución binomial, se puede demostrar que la fmp conjunta de $X_1, \ldots, X_r$ es

$$p(x_1, \ldots, x_r) = \begin{cases} \dfrac{n!}{(x_1!)(x_2!)\cdots(x_r!)} p_1^{x_1} \cdot \cdots \cdot p_r^{x_r} & x_i = 0, 1, 2, \ldots, \text{con } x_1 + \cdots + x_r = n \\ 0 & \text{en caso contrario} \end{cases}$$

El caso $r = 2$ da la distribución binomial, con $X_1 =$ número de éxitos y $X_2 = n - X_1 =$ número de fracasos.

Ejemplo 5.9 Si se determina el alelo de cada una de diez secciones de un chícharo obtenidas de forma independiente y $p_1 = P(\text{AA})$, $p_2 = P(\text{Aa})$, $p_3 = P(\text{aa})$, $X_1 =$ número de pares A, $X_2 =$ número de pares Aa y $X_3 =$ número de pares aa, entonces

$$p(x_1, x_2, x_3) = \frac{10!}{(x_1!)(x_2!)(x_3!)} p_1^{x_1} p_2^{x_2} p_3^{x_3} \qquad x_i = 0, 1, \ldots \quad \text{y } x_1 + x_2 + x_3 = 10$$

Si $p_1 = p_3 = 0.25$, $p_2 = 0.5$, entonces

$$P(X_1 = 2, X_2 = 5, X_3 = 3) = p(2, 5, 3)$$
$$= \frac{10!}{2!\,5!\,3!}(0.25)^2(0.5)^5(0.25)^3 = 0.0769$$ ■

Ejemplo 5.10 Cuando se emplea cierto método para colectar un volumen fijo de muestras de roca en una región, hay cuatro tipos de roca resultantes. Sea que X_1, X_2 y X_3 denotan la proporción en volumen de tipos de roca 1, 2 y 3 en una muestra seleccionada al azar (la proporción de roca tipo 4 es $1 - X_1 - X_2 - X_3$, así que sería redundante una variable X_4). Si la fdp conjunta de X_1, X_2, X_3 es

$$f(x_1, x_2, x_3) = \begin{cases} kx_1x_2(1 - x_3) & 0 \le x_1 \le 1, 0 \le x_2 \le 1, 0 \le x_3 \le 1, x_1 + x_2 + x_3 \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

entonces k se determina mediante

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2, x_3)\, dx_3\, dx_2\, dx_1$$
$$= \int_0^1 \left\{ \int_0^{1-x_1} \left[\int_0^{1-x_1-x_2} kx_1x_2(1 - x_3)\, dx_3 \right] dx_2 \right\} dx_1$$

Esta integral iterada tiene un valor $k/144$, así que $k = 144$. La probabilidad de que las rocas de los tipos 1 y 2 representen a lo sumo 50% de la muestra es

$$P(X_1 + X_2 \le 0.5) = \iiint_{\substack{0 \le x_i \le 1 \text{ para } i = 1, 2, 3 \\ x_1 + x_2 + x_3 \le 1,\ x_1 + x_2 \le 0.5}} f(x_1, x_2, x_3)\, dx_3\, dx_2\, dx_1$$
$$= \int_0^{0.5} \left\{ \int_0^{0.5-x_1} \left[\int_0^{1-x_1-x_2} 144x_1x_2(1 - x_3)\, dx_3 \right] dx_2 \right\} dx_1$$
$$= 0.6066$$ ■

El concepto de independencia de más de dos variables aleatorias, es similar al concepto de independencia de más de dos eventos.

DEFINICIÓN

Las variables aleatorias $X_1, X_2, \ldots, X_n$ son **independientes** si para *todo* subconjunto $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ de las variables (cada par, cada tripleta, etcétera), la fmp o la fdp del subconjunto es igual al producto de la fmp o fdp marginales.

Por lo tanto, si las variables son independientes con $n = 4$, entonces la fmp, o la fdp, conjunta de dos variables cualesquiera es el producto de las dos marginales, y lo mismo para tres variables cualesquiera y las cuatro variables juntas. Lo más importante, una vez que se tiene conocimiento de que las n variables son independientes, entonces la fmp conjunta o fdp es el producto de las n marginales.

Ejemplo 5.11 Si $X_1, \ldots, X_n$ representan los tiempos de vida de n componentes, los componentes operan de forma independiente entre sí y cada tiempo de vida tiene una distribución exponencial con parámetro λ, entonces

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= (\lambda e^{-\lambda x_1}) \cdot (\lambda e^{-\lambda x_2}) \cdot \cdots \cdot (\lambda e^{-\lambda x_n}) \\ &= \begin{cases} \lambda^n e^{-\lambda \Sigma x_i} & x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0 \\ 0 & \text{en caso contrario} \end{cases} \end{aligned}$$

Si estos n componentes constituyen un sistema que falla tan pronto como falla un solo componente, entonces la probabilidad de que el sistema dure pasado el tiempo t es

$$\begin{aligned} P(X_1 > t, \ldots, X_n > t) &= \int_t^\infty \cdots \int_t^\infty f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n \\ &= \left(\int_t^\infty \lambda e^{-\lambda x_1}\, dx_1\right) \cdots \left(\int_t^\infty \lambda e^{-\lambda x_n}\, dx_n\right) \\ &= (e^{-\lambda t})^n = e^{-n\lambda t} \end{aligned}$$

Por lo tanto,

$$P(\text{tiempo de vida del sistema} \leq t) = 1 - e^{-n\lambda t} \qquad \text{para } t \geq 0$$

lo cual muestra que el tiempo de vida del *sistema* tiene una distribución exponencial con parámetro $n\lambda$; el valor esperado del tiempo de vida del sistema es $1/n\lambda$. ∎

En muchas situaciones experimentales por considerar en este libro, la independencia es una suposición razonable, de modo que especificar la distribución conjunta se reduce a decidir acerca de las distribuciones marginales apropiadas.

Distribuciones condicionales

Suponga que X = número de defectos mayores en un automóvil nuevo elegido al azar y Y = número de defectos menores en el mismo auto. Si se sabe que el número de defectos mayores en el automóvil seleccionado es uno, ¿cuál es la probabilidad de que el automóvil tenga a lo sumo tres defectos menores; es decir, ¿cuál es $P(Y \leq 3 \mid X = 1)$? De manera similar, si X y Y son los tiempos de vida de dos componentes de un sistema y sucede que $X = 100$, ¿cuál es la probabilidad de que $Y \geq 200$, y cuál es el tiempo de vida esperado del segundo componente

"condicionado a" este valor de X? Las preguntas de este tipo se pueden contestar mediante el estudio de las distribuciones de probabilidad condicional.

DEFINICIÓN

Sean X y Y dos *va* continuas con fdp conjunta $f(x, y)$ y fdp de X marginal $f_X(x)$. Entonces para cualquier valor x de X para el cual $f_X(x) > 0$, la **función de densidad de probabilidad condicional de Y dado que $X = x$** es

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)} \qquad -\infty < y < \infty$$

Si X y Y son discretas, sustituir las fdp por las fmp en esta definición proporciona la **función masa de probabilidad condicional de Y cuando $X = x$.**

Observe que la definición de $f_{Y|X}(y \mid x)$ se asemeja a la de $P(B \mid A)$, la probabilidad condicional de que ocurra B, dado que ocurrió A. Una vez determinada la fdp o la fmp condicionales, las preguntas del tipo planteado al comienzo de esta subsección se contestan integrando o sumando sobre el conjunto apropiado de valores de Y.

Ejemplo 5.12 Considere de nuevo la situación de los ejemplos 5.3 y 5.4 relacionados con X = parte del tiempo que está ocupada la ventanilla de atención en automóvil de un banco y Y = parte análoga para la ventanilla de atención a clientes a pie. La fdp condicional de Y dado que $X = 0.8$ es

$$f_{Y|X}(y \mid 0.8) = \frac{f(0.8, y)}{f_X(0.8)} = \frac{1.2(0.8 + y^2)}{1.2(0.8 + 0.4)} = \frac{1}{34}(24 + 30y^2) \qquad 0 < y < 1$$

La probabilidad de que la ventanilla de atención a clientes a pie esté ocupada a lo sumo la mitad del tiempo dado que $X = 0.8$ es

$$P(Y \le 0.5 \mid X = 0.8) = \int_{-\infty}^{0.5} f_{Y|X}(y \mid 0.8)\, dy = \int_0^{0.5} \frac{1}{34}(24 + 30y^2)\, dy = 0.390$$

Por medio de la fdp marginal de Y se obtiene $P(Y \le 0.5) = 0.350$. También $E(Y) = 0.6$, mientras que la parte esperada tiempo que está ocupada la ventanilla de clientes a pie dado que $X = 0.8$ (una expectativa *condicional*) es

$$E(Y \mid X = 0.8) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid 0.8)\, dy = \frac{1}{34}\int_0^1 y(24 + 30y^2)\, dy = 0.574$$ ■

Ejercicios | Sección 5.1 (1 a 21)

1. Una gasolinera tiene dos islas, una de autoservicio y otra de servicio completo. En cada isla hay una sola bomba de gasolina sin plomo con dos mangueras. Sea X el número de mangueras que se está usando en una isla de autoservicio en determinado momento, y sea Y el número de mangueras en uso en la isla de servicio completo en ese momento. La fmp conjunta de X y Y se ilustra en la tabla siguiente.

$p(x, y)$		y = 0	1	2
x	0	0.10	0.04	0.02
	1	0.08	0.20	0.06
	2	0.06	0.14	0.30

   a. ¿Cuál es $P(X = 1$ y $Y = 1)$?
   b. Calcule $P(X \le 1$ y $Y \le 1)$.

c. Describa con palabras el evento $\{X \neq 0 \text{ y } Y \neq 0\}$, y calcule la probabilidad de este evento.
d. Calcule la fmp marginal de X y de Y. Por medio de $p_X(x)$, ¿cuál es $P(X \leq 1)$?
e. ¿Son X y Y *va* independientes? Explique su respuesta.

2. Cuando una patrulla de seguridad detiene a un automóvil, se comprueba el desgaste de cada neumático y se verifica si cada luz delantera apunta en la dirección correcta. Sea X el número de luces delanteras que requieren ajuste, y sea Y el número de llantas defectuosas.
a. Si X y Y son independientes con $p_X(0) = 0.5$, $p_X(1) = 0.3$, $p_X(2) = 0.2$ y $p_Y(0) = 0.6$, $p_Y(1) = 0.1$, $p_Y(2) = p_Y(3) = 0.05$, $p_Y(4) = 0.2$, muestre la fmp conjunta de (X, Y) en una tabla de probabilidad conjunta.
b. Calcule $P(X \leq 1 \text{ y } Y \leq 1)$ de la tabla de probabilidad conjunta y compruebe que es igual al producto $P(X \leq 1) \cdot P(Y \leq 1)$.
c. ¿Cuál es $P(X + Y = 0)$ (la probabilidad de que no hay violaciones)?
d. Calcule $P(X + Y \leq 1)$.

3. En cierto supermercado hay dos filas, una para la caja rápida y otra para la superveloz. Sea X_1 el número de clientes en la fila de la caja rápida en determinado momento del día, y sea X_2 el número de clientes en la fila de la caja superveloz en el mismo momento. Suponga que la fmp conjunta de X_1 y X_2 es como se da en la tabla siguiente.

		x_2			
		0	1	2	3
x_1	0	0.08	0.07	0.04	0.00
	1	0.06	0.15	0.05	0.04
	2	0.05	0.04	0.10	0.06
	3	0.00	0.03	0.04	0.07
	4	0.00	0.01	0.05	0.06

a. ¿Cuál es $P(X_1 = 1, X_2 = 1)$, es decir, la probabilidad de que haya exactamente un cliente en cada fila?
b. ¿Cuál es $P(X_1 = X_2)$, es decir, la probabilidad de que el número de clientes en las dos filas sea idéntico?
c. Represente con A el evento de que hay por lo menos dos clientes más en una fila que en la otra. Exprese A en términos de X_1 y X_2, y calcule la probabilidad de este evento.
d. ¿Cuál es la probabilidad de que el número total de clientes en las dos filas sea exactamente cuatro? ¿Por lo menos cuatro?

4. Vuelva a la situación descrita en el ejercicio 3.
a. Determine la fmp marginal de X_1, luego calcule el número esperado de clientes en la fila de la caja rápida.
b. Determine la fmp marginal de X_2.
c. Mediante inspección de las probabilidades $P(X_1 = 4)$, $P(X_2 = 0)$ y $P(X_1 = 4, X_2 = 0)$, ¿X_1 y X_2 son variables aleatorias independientes? Explique.

5. El número de clientes que esperan en el servicio de envoltura de regalos en una tienda de departamentos es una *va* X con valores posibles 0, 1, 2, 3, 4 y probabilidades correspondientes 0.1, 0.2, 0.3, 0.25, 0.15. Un cliente elegido al azar tendrá 1, 2, 3 paquetes para envoltura con probabilidades 0.6, 0.3 y 0.1, respectivamente. Sea Y = número total de paquetes por envolver para los clientes que esperan en la fila (suponga que el número de paquetes de un cliente es independiente de la cantidad de paquetes de otro cliente).
a. Determine $P(X = 3, Y = 3)$, es decir, $p(3, 3)$.
b. Determine $p(4, 11)$.

6. Sea X el número de cámaras digitales Canon vendidas durante una semana particular en cierta tienda. La fmp de X es

x	0	1	2	3	4
$p_X(x)$	0.1	0.2	0.3	0.25	0.15

Sesenta por ciento de los clientes que compran estas cámaras también compran una garantía extendida. Sea Y el número de compras durante esta semana que compran una garantía extendida.
a. ¿Cuál es $P(X = 4, Y = 2)$? [*Sugerencia*: esta probabilidad es igual a $P(Y = 2 \mid X = 4) \cdot P(X = 4)$; ahora considere que las cuatro compras son cuatro ensayos de un experimento binomial, con éxito en un ensayo que corresponde a comprar una garantía extendida.]
b. Calcule $P(X = Y)$.
c. Determine la fmp conjunta de X y Y y luego la fmp marginal de Y.

7. La distribución de probabilidad conjunta del número X de automóviles y el número Y de autobuses por ciclo de señal en un carril propuesto de vuelta a la izquierda se ilustra en la tabla siguiente de probabilidad conjunta.

$p(x, y)$		y		
		0	1	2
x	0	0.025	0.015	0.010
	1	0.050	0.030	0.020
	2	0.125	0.075	0.050
	3	0.150	0.090	0.060
	4	0.100	0.060	0.040
	5	0.050	0.030	0.020

a. ¿Cuál es la probabilidad de que haya exactamente un automóvil y un autobús durante un ciclo?

b. ¿Cuál es la probabilidad de que haya a lo sumo un automóvil y a lo sumo un autobús durante un ciclo?

c. ¿Cuál es la probabilidad de que haya exactamente un automóvil durante un ciclo? ¿Exactamente un autobús?

d. Suponga que el carril de vuelta a la izquierda tendrá una capacidad de cinco automóviles, y que un autobús equivale a tres automóviles. ¿Cuál es la probabilidad de un sobreflujo durante un ciclo?

e. ¿Son *va* independientes X y Y? Explique.

8. Las existencias actuales de un almacén son 30 componentes de cierto tipo, de los cuales ocho son del proveedor 1, 10 del proveedor 2 y 12 del proveedor 3. Seis de éstos se eligen al azar para un determinado ensamble. Sea X = número de componentes elegidos del proveedor 1, Y = número de componentes elegidos del proveedor 2 y $p(x, y)$ representa la fmp conjunta de X y Y.

a. ¿Cuál es $p(3, 2)$? [*Sugerencia*: cada muestra de tamaño seis tiene las mismas probabilidades de ser elegida. Por consiguiente, $p(3, 2)$ = (número de resultados con $X = 3$ y $Y = 2$)/(número total de resultados). Ahora utilice la regla del producto para conteo a fin de obtener el numerador y el denominador.]

b. Por medio del razonamiento del inciso (a), obtenga $p(x, y)$. (Esto se puede considerar como una distribución hipergeométrica de varias variables —muestreo sin reemplazo a partir de una población finita que consiste en más de dos categorías.)

9. Se supone que cada llanta frontal de determinado tipo de vehículo tiene una presión de inflado de 26 psi. Suponga que la presión real en cada llanta es una variable aleatoria, X para la llanta derecha y Y para la llanta izquierda, con fmp conjunta

$$f(x, y) = \begin{cases} K(x^2 + y^2) & 20 \le x \le 30,\ 20 \le y \le 30 \\ 0 & \text{en caso contrario} \end{cases}$$

a. ¿Cuál es el valor de K?

b. ¿Cuál es la probabilidad de que ambas llantas tengan una presión de inflado menor a la supuesta?

c. ¿Cuál es la probabilidad de que la diferencia en la presión de aire entre las dos llantas sea a lo sumo 2 psi?

d. Determine la distribución marginal de la presión de inflado sólo en la llanta derecha.

e. ¿X y Y son variables independientes?

10. Annie y Alvie han acordado reunirse entre las 5:00 P.M. y las 6:00 P.M. para cenar en un restaurante local de alimentos naturales. Sea X = hora de llegada de Annie y Y = hora de llegada de Alvie. Suponga que X y Y son independientes y que cada una tiene una distribución normal en el intervalo [5, 6].

a. ¿Cuál es la fdp conjunta de X y Y?

b. ¿Cuál es la probabilidad de que ambos lleguen entre 5:15 y 5:45?

c. Si el primero en llegar espera sólo 10 minutos antes de salir a comer en alguna otra parte, ¿cuál es la probabilidad de que cenen en el restaurante de alimentos naturales? [*Sugerencia*: el evento de interés es $A = \{(x, y): |x - y| \le \frac{1}{6}\}$.]

11. Dos profesores distintos solicitan la transcripción de exámenes finales. Sea X el número de errores tipográficos en el examen del primer profesor, y Y el número de esta clase de errores en el segundo examen. Suponga que X tiene una distribución de Poisson con parámetro λ, Y tiene una distribución de Poisson con parámetro θ, y X y Y son independientes.

a. ¿Cuál es la fmp conjunta de X y Y?

b. ¿Cuál es la probabilidad de que a lo sumo se cometa un error en ambos exámenes combinados?

c. Obtenga una expresión general para la probabilidad de que el número total de errores en los dos exámenes sea m (donde m es un entero no negativo). [*Sugerencia:* $A = \{(x, y): x + y = m\} = \{(m, 0), (m-1, 1), \ldots, (1, m-1), (0, m)\}$. Ahora sume la fmp conjunta sobre $(x, y) \in A$ y utilice el teorema del binomio que establece que

$$\sum_{k=0}^{m} \binom{m}{k} a^k b^{m-k} = (a + b)^m$$

para a, b cualesquiera.]

12. Dos componentes de una microcomputadora tienen la siguiente fdp conjunta para sus tiempos de vida X y Y:

$$f(x, y) = \begin{cases} xe^{-x(1+y)} & x \ge 0 \text{ y } y \ge 0 \\ 0 & \text{en caso contrario} \end{cases}$$

a. ¿Cuál es la probabilidad de que el tiempo de vida X del primer componente sea mayor que 3?

b. ¿Cuáles son las fdp marginales de X y de Y? ¿Son independientes los dos tiempos de vida? Explique.

c. ¿Cuál es la probabilidad de que el tiempo de vida de por lo menos uno de los componentes sea mayor que 3?

13. Se tienen dos focos de determinada lámpara. Sea X = duración del primer foco y Y = duración del segundo foco (ambas en miles de horas). Suponga que X y Y son independientes y que cada una tiene una distribución exponencial con parámetro $\lambda = 1$.

a. ¿Cuál es la fdp conjunta de X y Y?

b. ¿Cuál es la probabilidad de que cada foco dure a lo sumo 1000 horas (es decir, $X \le 1$ y $Y \le 1$)?

c. ¿Cuál es la probabilidad de que la duración total de los dos focos sea a lo sumo 2? [*Sugerencia:* trace un dibujo de la región $A = \{(x, y): x \geq 0, y \geq 0, x + y \leq 2\}$ antes de integrar.]

d. ¿Cuál es la probabilidad de que la duración total esté entre 1 y 2?

14. Suponga que tiene diez focos, que la duración de cada uno es independiente de las de los demás y que cada duración tiene una distribución exponencial con parámetro λ.

a. ¿Cuál es la probabilidad de que los diez focos fallen antes del tiempo t?

b. ¿Cuál es la probabilidad de que exactamente k de los diez focos fallen antes del tiempo t?

c. Suponga que nueve de los focos tienen vidas útiles distribuidas exponencialmente con parámetro λ y que el foco restante tiene una vida útil distribuida de manera exponencial con parámetro θ (es de otro fabricante). ¿Cuál es la probabilidad de que exactamente cinco de los diez focos fallen antes del tiempo t?

15. Considere un sistema que consiste en tres componentes como se ilustra. El sistema continúa funcionando siempre que funcione el primer componente y también el componente 2 o el 3. Sean X_1, X_2 y X_3 los tiempos de vida útil de los componentes 1, 2 y 3, respectivamente. Suponga que las X_i son independientes entre sí, y que cada X_i tiene una distribución exponencial con parámetro λ.

a. Sea Y la duración del sistema. Obtenga la función de distribución acumulada de Y y derive para obtener la fdp. [*Sugerencia:* $F(y) = P(Y \leq y)$; exprese el evento $\{Y \leq y\}$ en términos de uniones o intersecciones, o ambas, de los tres eventos $\{X_1 \leq y\}$, $\{X_2 \leq y\}$ y $\{X_3 \leq y\}$.]

b. Calcule el tiempo de vida esperado del sistema.

16. a. Para $f(x_1, x_2, x_3)$ como se dio en el ejemplo 5.10, calcule la **función de densidad marginal conjunta** de X_1 y X_3 (integrando en x_2).

b. ¿Cuál es la probabilidad de que las rocas tipos 1 y 3 constituyan a lo sumo 50% de la muestra? [*Sugerencia:* utilice el resultado del inciso (a).]

c. Calcule la fdp de X_1. [*Sugerencia:* utilice el resultado del inciso (a).]

17. Un ecologista desea seleccionar un punto dentro de una región de muestreo circular de acuerdo con una distribución uniforme (en la práctica, esto se podría hacer si se elige primero una dirección y luego una distancia desde el centro en esa dirección). Sea X = coordenada x del punto elegido y Y = coordenada y del punto seleccionado. Si el círculo está centrado en $(0, 0)$ y tiene radio R, entonces la fdp conjunta de X y Y es

$$f(x, y) = \begin{cases} \dfrac{1}{\pi R^2} & x^2 + y^2 \leq R^2 \\ 0 & \text{en caso contrario} \end{cases}$$

a. ¿Cuál es la probabilidad de que el punto seleccionado esté dentro de $R/2$ del centro de la región circular? [*Sugerencia:* trace un esquema de la región de densidad positiva D. Debido a que $f(x, y)$ es constante en D, calcular la probabilidad se reduce al cálculo de un área.]

b. ¿Cuál es la probabilidad de que tanto X como Y difieran de 0 en a lo sumo $R/2$?

c. Conteste el inciso (b) para $R/\sqrt{2}$ en lugar de $R/2$.

d. ¿Cuál es la fdp marginal de X? ¿De Y? ¿X y Y son independientes?

18. Refiérase al ejercicio 1 y conteste las preguntas siguientes:

a. Dado que $X = 1$, determine la fmp condicional de Y; es decir Y por ejemplo, $p_{Y|X}(0 \mid 1)$, $p_{Y|X}(1 \mid 1)$ y $p_{Y|X}(2 \mid 1)$.

b. Dado que dos mangueras están en uso en la isla de autoservicio, ¿cuál es la fmp condicional del número de mangueras en uso en la isla de servicio completo?

c. Utilice el resultado del inciso (b) para calcular la probabilidad condicional $P(Y \leq 1 \mid X = 2)$.

d. Dado que dos mangueras están en uso en la isla de servicio completo, ¿cuál es la fmp condicional del número de mangueras en uso en la isla de autoservicio?

19. La fdp conjunta de las presiones para las llantas frontales izquierda y derecha se dio en el ejercicio 9.

a. Determine la fdp condicional de Y dado que $X = x$ y la fdp condicional de X dado que $Y = y$.

b. Si se encuentra que la presión de la llanta derecha es 22 psi, ¿cuál es la probabilidad de que la llanta izquierda tenga un presión de por lo menos 25 psi? Compare esto con $P(Y \geq 25)$.

c. Si se encuentra que la presión en la llanta derecha es 22 psi, ¿cuál es la presión esperada en la llanta izquierda, y cuál es la desviación estándar de la presión en esta llanta?

20. Sean X_1, X_2 y X_3 los tiempos de vida de los componentes 1, 2 y 3 de un sistema de tres componentes.

a. ¿Cómo definiría la fdp condicional de X_3 dado que $X_1 = x_1$ y $X_2 = x_2$?

b. ¿Cómo definiría la fdp condicional conjunta de X_2 y X_3 dado que $X_1 = x_1$?

21. ¿Qué condición de $f_{Y|X}(y \mid x)$ es equivalente a la independencia de X y Y?

5.2 Valores esperados, covarianza y correlación

Ya se vio que cualquier función $h(X)$ de una sola *va* X es por sí misma una variable aleatoria. Sin embargo, para calcular $E[h(X)]$, no fue necesario obtener la distribución de probabilidad de $h(X)$; en cambio, $E[h(X)]$ se calculó como un promedio ponderado de valores de $h(x)$ donde la función de ponderación fue la fmp $p(x)$ o la fdp $f(x)$ de X. Un resultado similar se cumple para una función $h(X, Y)$ de dos variables aleatorias distribuidas conjuntamente.

PROPOSICIÓN

Sean X y Y dos variables aleatorias distribuidas de manera conjunta con fmp $p(x, y)$ o fdp $f(x, y)$ según si las variables son discretas o continuas. Entonces el valor esperado de una función $h(X, Y)$, denotada por $E[h(X, Y)]$ o $\mu_{h(X,Y)}$, está dada por

$$E[h(X, Y)] = \begin{cases} \sum_x \sum_y h(x, y) \cdot p(x, y) & \text{si } X \text{ y } Y \text{ son discretas} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y)\, dx\, dy & \text{si } X \text{ y } Y \text{ son continuas} \end{cases}$$

Ejemplo 5.13 Cinco amigos compraron boletos para un concierto. Si los boletos son para los asientos 1 a 5 en una fila determinada y los boletos están distribuidos al azar entre los cinco, ¿cuál es el número esperado de asientos que separan a dos amigos cualesquiera de los cinco? Sean X y Y los números de asientos del primer individuo y del segundo, respectivamente. Los pares posibles (X, Y) son $\{(1, 2), (1, 3), \ldots, (5, 4)\}$, y la fmp conjunta de (X, Y) es

$$p(x, y) = \begin{cases} \dfrac{1}{20} & x = 1, \ldots, 5;\ y = 1, \ldots, 5;\ x \neq y \\ 0 & \text{en caso contrario} \end{cases}$$

El número de asientos que separan a dos individuos es $h(X, Y) = |X - Y| - 1$. En la tabla siguiente se da $h(x, y)$ para cada par posible (x, y).

$h(x, y)$		x = 1	2	3	4	5
	1	—	0	1	2	3
	2	0	—	0	1	2
y	3	1	0	—	0	1
	4	2	1	0	—	0
	5	3	2	1	0	—

Así,

$$E[h(X, Y)] = \sum\sum_{(x, y)} h(x, y) \cdot p(x, y) = \sum_{x=1}^{5} \sum_{\substack{y=1 \\ x \neq y}}^{5} (|x - y| - 1) \cdot \frac{1}{20} = 1$$

Ejemplo 5.14 En el ejemplo 5.5, la fdp conjunta de la cantidad X de almendras y la cantidad Y de castañas en una lata de una libra de nueces fue

$$f(x, y) = \begin{cases} 24xy & 0 \le x \le 1, 0 \le y \le 1, x + y \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Si una libra de almendras cuesta a la compañía un dólar, una libra de castañas cuesta \$1.50 y una libra de maní cuesta \$0.50, entonces el costo total del contenido de una lata es

$$h(X, Y) = (1)X + (1.5)Y + (0.5)(1 - X - Y) = 0.5 + 0.5X + Y$$

(puesto que $1 - X - Y$ del peso consiste en maní). El costo esperado total es

$$E[h(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y)\, dx\, dy$$

$$= \int_{0}^{1} \int_{0}^{1-x} (0.5 + 0.5x + y) \cdot 24xy\, dy\, dx = \$1.10$$

La manera de calcular el valor esperado de una función $h(X_1, \ldots, X_n)$ de n variables aleatorias es similar a la utilizada para dos variables aleatorias. Si las X_i son discretas, $E[h(X_1, \ldots, X_n)]$ es una suma de n dimensional; si las X_i son continuas, se tiene una integral n-dimensional.

Covarianza

Cuando dos variables aleatorias X y Y no son independientes, suele ser interesante evaluar cuán estrecha es la relación entre sí.

DEFINICIÓN

La **covarianza** entre dos *va* X y Y es

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$= \begin{cases} \sum\sum_{x \ y} (x - \mu_X)(y - \mu_Y) p(x, y) & X, Y \text{ son discretas} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y)\, dx\, dy & X, Y \text{ son continuas} \end{cases}$$

El fundamento para la definición es el siguiente. Suponga que X y Y tienen una fuerte relación positiva entre sí, mediante la cual se indica que los valores grandes de X tienden a ocurrir con valores grandes de Y y los valores pequeños de X con valores pequeños de Y. Entonces la mayor parte de la masa o densidad de probabilidad estará relacionada con $(x - \mu_X)$ y $(y - \mu_Y)$, ya sea ambas positivas (tanto X como Y arriba de sus medias respectivas) o ambas negativas, así que el producto $(x - \mu_X)(y - \mu_Y)$ tenderá a ser positivo. Por lo tanto, para una fuerte relación positiva, $\text{Cov}(X, Y)$ debe ser bastante positiva. Para una fuerte relación negativa, los signos de $(x - \mu_X)$ y $(y - \mu_Y)$ tenderán a ser opuestos, de modo que el resultado es un

producto negativo. De este modo, para una relación negativa fuerte, $\text{Cov}(X, Y)$ debe ser bastante negativa. Si no hay una relación fuerte entre X y Y, los productos positivos y negativos tienden a cancelarse entre sí, produciendo una covarianza cercana a 0. En la figura 5.4 se ilustran las diferentes posibilidades. La covarianza depende *tanto* del conjunto de pares posible *como* de las probabilidades. En la figura 5.4, las probabilidades se podrían modificar sin alterar el conjunto de pares posibles, y esto podría cambiar de manera drástica el valor de $\text{Cov}(X, Y)$.

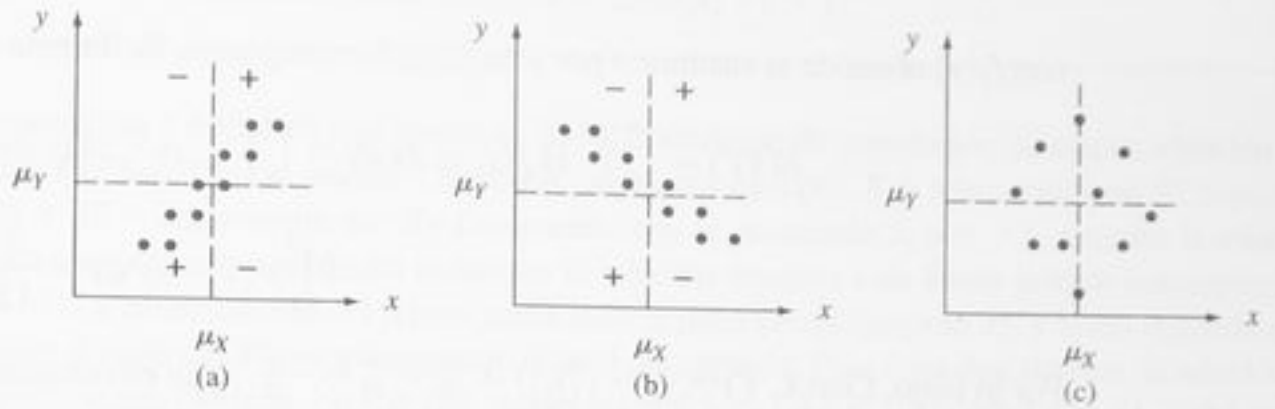


Figura 5.4 $p(x, y) = 1/10$ para cada uno de diez pares que corresponden a puntos indicados; (a) covarianza positiva; (b) covarianza negativa; (c) covarianza cercana a cero

Ejemplo 5.15 Las fmp conjunta y marginal para X = cantidad deducible de la póliza del automóvil y Y = cantidad deducible de la póliza de la casa del ejemplo 5.1 fueron

$p(x, y)$		y = 0	100	200	x	100	250	y	0	100	200
x	100	0.20	0.10	0.20	$p_X(x)$	0.5	0.5	$p_Y(y)$	0.25	0.25	0.5
	250	0.05	0.15	0.30							

de las cuales $\mu_X = \sum x p_X(x) = 175$ y $\mu_Y = 125$. Por lo tanto,

$$\begin{aligned}\text{Cov}(X, Y) &= \sum\sum_{(x, y)} (x - 175)(y - 125)p(x, y) \\ &= (100 - 175)(0 - 125)(0.20) + \cdots \\ &\quad + (250 - 175)(200 - 125)(0.30) \\ &= 1875\end{aligned}$$

■

La siguiente fórmula abreviada para $\text{Cov}(X, Y)$ simplifica los cálculos.

PROPOSICIÓN

$$\text{Cov}(X, Y) = E(XY) - \mu_X \cdot \mu_Y$$

De acuerdo con esta fórmula, ninguna resta intermedia es necesaria; sólo al final del cálculo se resta $\mu_X \cdot \mu_Y$ de $E(XY)$. En la demostración se requiere desarrollar $(X - \mu_X)(Y - \mu_Y)$ y luego tomar el valor esperado de cada término por separado. Observe que $\text{Cov}(X, X) = E(X^2) - \mu_X^2 = V(X)$.

Ejemplo 5.16 (continuación del ejemplo 5.5) Las fdp conjunta y marginal de X = cantidad de almendras y Y = cantidad de castañas fueron

$$f(x, y) = \begin{cases} 24xy & 0 \le x \le 1, 0 \le y \le 1, x + y \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

$$f_X(x) = \begin{cases} 12x(1 - x)^2 & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

con $f_Y(y)$ obtenida al sustituir x por y en $f_X(x)$. Se comprueba fácilmente que $\mu_X = \mu_Y = \frac{2}{5}$, y

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x, y)\, dx\, dy = \int_0^1 \int_0^{1-x} xy \cdot 24xy\, dy\, dx$$

$$= 8\int_0^1 x^2(1 - x)^3\, dx = \frac{2}{15}$$

Por lo tanto, $\text{Cov}(X, Y) = \frac{2}{15} - \left(\frac{2}{5}\right)\left(\frac{2}{5}\right) = \frac{2}{15} - \frac{4}{25} = -\frac{2}{75}$. Aquí es razonable una covarianza negativa, porque más almendras en la lata significan menos castañas. ∎

Parecería que la relación en el ejemplo del seguro es bastante fuerte puesto que $\text{Cov}(X, Y) = 1875$, mientras que $\text{Cov}(X, Y) = -\frac{2}{75}$ en el ejemplo de las nueces parecería implicar una relación bastante débil. Infortunadamente, la covarianza tiene un defecto grave que hace imposible interpretar un valor calculado de la covarianza. En el ejemplo del seguro, suponga que la cantidad deducible se expresó en centavos y no en dólares. Entonces $100X$ remplazaría a X, $100Y$ sustituiría a Y, y la varianza resultante sería $\text{Cov}(100X, 100Y) = (100)(100)\text{Cov}(X, Y) = 18\,750\,000$. Si, por otro lado, la cantidad deducible ha sido expresada en cientos de dólares, la covarianza calculada habría sido $(0.01)(0.01)(1875) = 0.1875$. *El defecto de la covarianza es que su valor calculado depende en gran medida de las unidades de medición.* Desde el punto de vista ideal, la elección de las unidades no debe tener ningún efecto en la medida de la fuerza de la relación. Esto se logra mediante una graduación de la covarianza.

Correlación

DEFINICIÓN

El **coeficiente de correlación** de X y Y, representado por $\text{Corr}(X, Y)$, $\rho_{X,Y}$, o sólo ρ, se define por

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

Ejemplo 5.17 Es fácil comprobar que el problema del seguro del ejemplo 5.15, $E(X^2) = 36\,250$, $\sigma_X^2 = 36\,250 - (175)^2 = 5625$, $\sigma_X = 75$, $E(Y^2) = 22\,500$, $\sigma_Y^2 = 6875$ y $\sigma_Y = 82.92$. Esto da

$$\rho = \frac{1875}{(75)(82.92)} = 0.301$$

∎

La siguiente proposición muestra que ρ corrige el defecto de $\text{Cov}(X, Y)$ y también sugiere cómo reconocer la existencia de una fuerte relación (lineal).

PROPOSICIÓN

1. Si a y c son positivas o negativas,
$$\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$$
2. Para dos *va* cualesquiera X, Y, $-1 \leq \text{Corr}(X, Y) \leq 1$.

El enunciado 1 establece con precisión que el coeficiente de correlación no resulta afectado por un cambio lineal en las unidades de medición (si, por ejemplo, X = temperatura en °C, entonces $9X/5 + 32$ = temperatura en °F). De acuerdo con el enunciado 2, $\rho = +1$, muestra la relación positiva más fuerte posible, en tanto que la relación negativa más fuerte posible corresponde a $\rho = -1$. La demostración del primer enunciado se ilustra en el ejercicio 35, y la del segundo aparece en el ejercicio 87 complementario al final del capítulo. Para fines descriptivos, la relación se describe como fuerte si $|\rho| \geq 0.8$, moderada si $0.5 < |\rho| < 0.8$ y débil si $|\rho| \leq 0.5$.

Si se considera que $p(x, y)$ o $f(x, y)$ dictan un modelo matemático de cómo están distribuidas dos variables numéricas X y Y en alguna población (altura y peso, puntuación del SAT oral o puntuación del SAT cuantitativo, etcétera), entonces ρ es una característica o parámetro de la población cuán estrecha es la relación de X y Y en la población. En el capítulo 12, se considera tomar una muestra de pares $(x_1, y_1), \ldots, (x_n, y_n)$ de la población. Entonces el coeficiente de correlación muestral r se define y se emplea para sacar conclusiones acerca de ρ.

El coeficiente de correlación ρ no es en realidad una medida del todo general de la fuerza de una relación.

PROPOSICIÓN

1. Si X y Y son independientes, entonces $\rho = 0$, pero $\rho = 0$ no implica independencia.
2. $\rho = 1$ o -1 si y sólo si $Y = aX + b$ para algunos números a y b con $a \neq 0$.

Esta proposición establece que ρ es una medida del grado de relación **lineal** entre X y Y, y sólo cuando las dos variables tienen una relación perfecta de una manera lineal ρ será tan positiva o negativa como pueda ser. Una ρ menor que 1 en valor absoluto sólo indica que la relación no es completamente lineal, pero podría haber aún una relación no lineal muy fuerte. También, $\rho = 0$ no significa que X y Y sean independientes, sino que sólo hay una ausencia completa de una relación lineal. Cuando $\rho = 0$, se dice que X y Y **no están correlacionadas**. Dos variables podrían no estar correlacionadas y ser muy dependientes porque hay una fuerte relación no lineal, así que evite concluir demasiado a partir del hecho de que $\rho = 0$.

Ejemplo 5.18 Sean X y Y variables aleatorias discretas con fmp conjunta

$$p(x, y) = \begin{cases} \frac{1}{4} & (x, y) = (-4, 1), (4, -1), (2, 2), (-2, -2) \\ 0 & \text{en caso contrario} \end{cases}$$

Los puntos que reciben masa de probabilidad positiva se identifican en el sistema coordenado (x, y) de la figura 5.5. Es evidente en la figura que el valor de X se determina por completo mediante el valor de Y y viceversa, de modo que las variables son completamente dependientes. Sin embargo, por simetría $\mu_X = \mu_Y = 0$ y $E(XY) = (-4)\frac{1}{4} + (-4)\frac{1}{4} + (4)\frac{1}{4} + (4)\frac{1}{4} = 0$, así que $\text{Cov}(X, Y) = E(XY) - \mu_X \cdot \mu_Y = 0$ y, por consiguiente, $\rho_{X,Y} = 0$. Aunque hay una dependencia perfecta, ¡también hay ausencia completa de relación lineal!

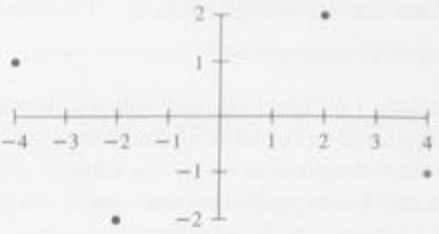


Figura 5.5 Población de pares para el ejemplo 5.18

Un valor de ρ cercano a 1 no necesariamente implica que incrementar el valor de X *ocasione* que aumente Y. Esto sólo implica que valores grandes de X están *relacionados* con valores grandes de Y. Por ejemplo, en la población de niños, el tamaño del vocabulario y el número de cavidades están bastante correlacionados de manera positiva, pero de hecho no es verdad que las cavidades ocasionen que el vocabulario crezca. En cambio, los valores de estas dos variables tienden a incrementarse a medida que aumenta la edad, una tercera variable. Para niños de una edad fija, hay probablemente una correlación muy baja entre el número de cavidades y el tamaño del vocabulario. En resumen, relación (una correlación alta) no es lo mismo que causa.

Ejercicios | Sección 5.2 (22 a 36)

22. Un instructor entrega a sus alumnos un cuestionario corto que consta de dos partes. Para un estudiante seleccionado al azar, sea X = número de puntos obtenidos en la primera parte Y = número de puntos obtenidos en la segunda parte. Suponga que la fmp conjunta de X y Y es la que se da en la tabla siguiente.

$p(x, y)$		y = 0	5	10	15
x	0	0.02	0.06	0.02	0.10
	5	0.04	0.15	0.20	0.10
	10	0.01	0.15	0.14	0.01

a. Si la puntuación registrada en el libro de calificaciones es el número total de puntos obtenidos en las dos partes, ¿cual es la puntuación registrada esperada $E(X + Y)$?

b. Si se registra el máximo de las dos puntuaciones, ¿cuál es la puntuación registrada esperada?

23. La diferencia entre el número de clientes en la fila de la caja rápida y el número en la caja superveloz del ejercicio 3 es $X_1 - X_2$. Calcule la diferencia esperada.

24. Seis individuos, incluso A y B, se sientan por completo al azar alrededor de una mesa redonda. Suponga que los asientos están numerados $1, \ldots, 6$. Sea X = número de asiento de A y Y = número de asiento de B. Si A envía un mensaje escrito a B alrededor de la mesa, en la dirección de mayor cercanía, ¿por cuántos individuos (incluso A y B) se esperaría que pase el mensaje?

25. Un topógrafo desea medir un terreno cuadrado con longitud L en cada lado. Sin embargo, debido a un error de medición, traza un rectángulo donde los lados norte-sur tienen una longitud X y los lados este-oeste tienen longitud Y. Suponga que X y Y son independientes y que cada una tiene una distribución uniforme en el intervalo $[L - A, L + A]$ (donde $0 < A < L$). ¿Cuál es el área esperada del rectángulo resultante?

26. Considere un pequeño transbordador con capacidad para automóviles y autobuses. La cuota para automóviles es tres dólares y la de autobuses 10 dólares. Represente con X y Y el número de automóviles y autobuses, respectivamente, transportados en un solo viaje. Suponga que la distribución conjunta de X y Y es como se da en la tabla del ejercicio 7. Calcule el ingreso esperado de un solo viaje.

27. Annie y Alvie han acordado reunirse para almorzar entre el mediodía (12 horas) y la 1:00 P.M. Represente con X la hora de llegada de Annie y con Y la de Alvie, y suponga que X y Y son independientes con fdp

$$f_X(x) = \begin{cases} 3x^2 & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

$$f_Y(y) = \begin{cases} 2y & 0 \le y \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

¿Cuánto tiempo debe esperar la persona que llega primero a la otra? [*Sugerencia:* $h(X, Y) = |X - Y|$.]

28. Demuestre que si X y Y son *va* independientes, entonces $E(XY) = E(X) \cdot E(Y)$. Después, aplique esto en el ejercicio 25. [*Sugerencia:* considere el caso continuo con $f(x, y) = f_X(x) \cdot f_Y(y)$.]

29. Calcule el coeficiente de correlación ρ para X y Y del ejemplo 5.16 (la covarianza ya ha sido calculada).

30. **a.** Calcule la covarianza para X y Y en el ejercicio 22.
b. Calcule ρ para X y Y en el mismo ejercicio.

31. **a.** Calcule la covarianza entre X y Y en el ejercicio 9.
b. Calcule el coeficiente de correlación ρ para X y Y.

32. Considere de nuevo los tiempos de vida útil X y Y del componente de microcomputadora descrito en el ejercicio 12. Determine $E(XY)$. ¿Qué se puede decir acerca de $\text{Cov}(X, Y)$ y ρ?

33. Utilice el resultado del ejercicio 28 para demostrar que cuando X y Y son independientes, $\text{Cov}(X, Y) = \text{Corr}(X, Y) = 0$.

34. **a.** Recuerde la definición de σ^2 para una sola *va* X, y escriba una fórmula que resulte apropiada para calcular la varianza de una función $h(X, Y)$ de dos variables aleatorias. [*Sugerencia:* recuerde que la varianza es sólo un valor esperado especial.]
b. Utilice esta fórmula para calcular la varianza de la puntuación registrada $h(X, Y)$ [$= \text{máx}(X, Y)$] en el inciso (b) del ejercicio 22.

35. **a.** Utilice las reglas de valor esperado para demostrar que $\text{Cov}(aX + b, cY + d) = ac\,\text{Cov}(X, Y)$.
b. Utilice el inciso (a) junto con las reglas de varianza y desviación estándar para demostrar que $\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$ cuando a y c tienen el mismo signo.
c. ¿Qué sucede si a y c tienen signos opuestos?

36. Demuestre que si $Y = aX + b$ $(a \neq 0)$, entonces $\text{Corr}(X, Y) = +1$ o -1. ¿En qué condiciones se cumple que $\rho = +1$?

5.3 Estadísticos y sus distribuciones

Las observaciones de una sola muestra se denotaron en el capítulo 1 con $x_1, x_2, \ldots, x_n$. Considere elegir dos muestras distintas de tamaño n de la misma distribución poblacional. Las x_i de la segunda muestra casi siempre diferirán por lo menos un poco de las de la primera muestra. Por ejemplo, una primera muestra de $n = 3$ automóviles de un tipo particular podría producir rendimientos de combustible $x_1 = 30.7$, $x_2 = 29.4$, $x_3 = 31.1$, en tanto que una segunda muestra podría dar $x_1 = 28.8$, $x_2 = 30.0$ y $x_3 = 31.1$. Antes de obtener los datos, hay incertidumbre acerca del valor de cada x_i. Debido a esta incertidumbre, *antes* que estén disponibles los datos se ve cada observación como una variable aleatoria y la muestra se representa mediante $X_1, X_2, \ldots, X_n$ (letras mayúsculas para variables aleatorias).

Esta variación de valores observados a su vez implica que el valor de cualquier función de las observaciones muestrales, por ejemplo, la media, la desviación estándar o la cuarta dispersión muestrales, también varía de una muestra a otra. Es decir, antes de obtener $x_1, \ldots, x_n$, hay incertidumbre en cuanto al valor de $\bar{x}$, el valor de s, etcétera.

Ejemplo 5.19 Suponga que la resistencia del material para un espécimen elegido al azar de un tipo particular tiene una distribución de Weibull con valores de parámetro $\alpha = 2$ (forma) y $\beta = 5$ (escala). La curva de densidad correspondiente se muestra en la figura 5.6. Con las fórmulas de la sección 4.5, se obtiene

$$\mu = E(X) = 4.4311 \qquad \tilde{\mu} = 4.1628 \qquad \sigma^2 = V(X) = 5.365 \qquad \sigma = 2.316$$

La media excede la mediana como resultado del sesgo positivo de la distribución.

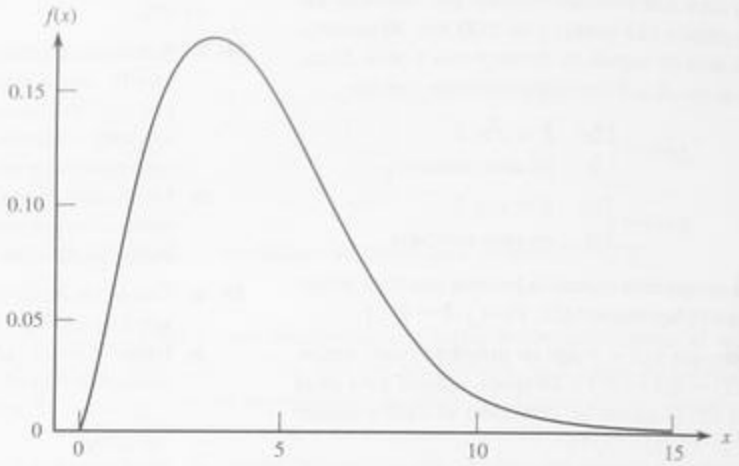


Figura 5.6 Curva de densidad de Weibull para el ejemplo 5.19

Se utilizó MINITAB para generar seis muestras diferentes, cada una con $n = 10$, de esta distribución (resistencias de materiales para seis grupos diferentes de diez especímenes cada uno). El resultado aparece en la tabla 5.1, seguidos de los valores de la media, mediana y desviación estándar para cada muestra. Observe primero que las diez observaciones en cualquier muestra son diferentes de las de cualquier otra. Segundo, los seis valores de la media muestral son diferentes entre sí, al igual que los seis valores de la mediana y los seis valores de la desviación estándar muestrales. Lo mismo se cumple para las medias muestrales recortadas al 10%, las cuartas dispersiones muestrales, etcétera.

Tabla 5.1 Muestras de la distribución de Weibull del ejemplo 5.19

Muestra	1	2	3	4	5	6
1	6.1171	5.07611	3.46710	1.55601	3.12372	8.93795
2	4.1600	6.79279	2.71938	4.56941	6.09685	3.92487
3	3.1950	4.43259	5.88129	4.79870	3.41181	8.76202
4	0.6694	8.55752	5.14915	2.49759	1.65409	7.05569
5	1.8552	6.82487	4.99635	2.33267	2.29512	2.30932
6	5.2316	7.39958	5.86887	4.01295	2.12583	5.94195
7	2.7609	2.14755	6.05918	9.08845	3.20938	6.74166
8	10.2185	8.50628	1.80119	3.25728	3.23209	1.75468
9	5.2438	5.49510	4.21994	3.70132	6.84426	4.91827
10	4.5590	4.04525	2.12934	5.50134	4.20694	7.26081

(continúa)

Muestra	1	2	3	4	5	6
$\bar{x}$	4.401	5.928	4.229	4.132	3.620	5.761
$\tilde{x}$	4.360	6.144	4.608	3.857	3.221	6.342
s	2.642	2.062	1.611	2.124	1.678	2.496

Además, el valor de la media muestral de cualquier muestra se puede considerar como una *estimación puntual* ("puntual", porque se trata de un solo número, que corresponde a un solo punto de la recta numérica) de la media poblacional μ, cuyo valor se sabe que es 4.4311. Ninguna de las estimaciones de estas seis muestras es idéntica a la que está siendo estimada. Las estimaciones a partir de las muestras segunda y sexta son mucho más grandes, en tanto que la quinta muestra proporciona una subestimación sustancial. De manera similar, la desviación estándar muestral da una estimación puntual de la desviación estándar poblacional. Las seis estimaciones resultantes tienen un error de por lo menos una cantidad pequeña.

En resumen, los valores de cada una de las observaciones muestrales varían de una muestra a otra, así que en general el valor de cualquier cantidad calculada a partir de los datos de la muestra, y el valor de una característica muestral utilizado como estimación de la característica poblacional correspondiente, casi nunca coincidirá con lo que está siendo estimado. ■

DEFINICIÓN

Un **estadístico** es cualquier cantidad cuyo valor se calcula a partir de los datos de la muestra. Antes de obtener los datos, hay incertidumbre en cuanto a qué valor se obtendrá para algun estadístico particular. Por lo tanto, un estadístico es una variable aleatoria y se denota con una letra mayúscula; se emplea una letra minúscula para representar el valor calculado u observado del estadístico.

Así, la media muestral, considerada como un estadístico (antes de seleccionar una muestra o realizar un experimento), se denota mediante $\bar{X}$; el valor calculado de este estadístico es $\bar{x}$. De manera similar, S representa la desviación estándar muestral considerada como estadístico, y su valor calculado es s. Si se seleccionan dos tipos de ladrillos y cada una de las resistencias se denotan mediante $X_1, \ldots, X_m$ y $Y_1, \ldots, Y_n$, respectivamente, entonces el estadístico $\bar{X} - \bar{Y}$, la diferencia entre las dos resistencias compresivas medias de la muestra, suele ser de gran interés.

Cualquier estadístico, que es una variable aleatoria, tiene una distribución de probabilidad. En particular, la media muestral $\bar{X}$ tiene una distribución de probabilidad. Suponga, por ejemplo, que se eligen al azar $n = 2$ componentes y se determina para cada uno el número de fallas mientras está vigente la garantía. Los valores posibles para el número promedio muestral de fallas $\bar{X}$ son 0 (si $X_1 = X_2 = 0$), 0.5 (si $X_1 = 0$ y $X_2 = 1$ o $X_1 = 1$ y $X_2 = 0$), 1, 1.5, ... La distribución de probabilidad de $\bar{X}$ especifica $P(\bar{X} = 0)$, $P(\bar{X} = 0.5)$, etcétera, a partir de las cuales se pueden calcular otras probabilidades, como $P(1 \le \bar{X} \le 3)$ y $P(\bar{X} \ge 2.5)$. De forma similar, si para una muestra de tamaño $n = 2$, los únicos valores posibles de la varianza muestral son 0, 12.5 y 50 (que es el caso si X_1 y X_2 toman cada una sólo los valores 40, 45 y 50), entonces la distribución de probabilidad de S^2 produce $P(S^2 = 0)$, $P(S^2 = 12.5)$ y $P(S^2 = 50)$. La distribución de probabilidad de un estadístico se denomina a veces como su **distribución de muestreo** para remarcar que describe cómo varía el valor del estadístico en las muestras que podrían ser seleccionadas.

Muestras aleatorias

La distribución de probabilidad de algún estadístico en particular, depende no sólo de la distribución de la población (normal, uniforme, etcétera) y del tamaño de la muestra n, sino también del método de muestreo. Considere elegir una muestra de tamaño $n = 2$ a partir de una población que consiste en sólo los tres valores 1, 5 y 10, y suponga que el estadístico de interés es la varianza muestral. Si el muestreo se hace "con reemplazo", entonces $S^2 = 0$ resulta si $X_1 = X_2$. Sin embargo, S^2 no puede ser igual a 0 si el muestreo es "sin reemplazo". Por lo tanto $P(S^2 = 0) = 0$ para un método de muestreo, y su probabilidad es positiva para el otro método. En la siguiente definición se describe un método de muestreo que suele encontrarse (por lo menos en forma aproximada) en la práctica.

DEFINICIÓN

Se dice que las *va* $X_1, X_2, \ldots, X_n$ forman una **muestra aleatoria** (simple) de tamaño n si

1. Las X_i son *va* independientes.
2. Toda X_i tiene la misma distribución de probabilidad.

Las condiciones 1 y 2 se pueden parafrasear diciendo que las X_i son *independientes e idénticamente distribuidas* (iid). Si el muestreo es ya sea con reemplazo o a partir de una población infinita (conceptual), las condiciones 1 y 2 se cumplen de forma exacta. Estas condiciones se satisfarán de manera aproximada si el muestreo es sin reemplazo; sin embargo, el tamaño de muestra n es mucho más pequeño que el tamaño de la población N. En la práctica, si $n/N \leq 0.05$ (a lo sumo se muestrea 5% de la población), se puede proceder como si las X_i formaran una muestra aleatoria. La virtud de este método de muestreo es que la distribución de probabilidad de cualquier estadístico se puede obtener con más facilidad que para cualquier otro método de muestreo.

Hay dos métodos generales para obtener información acerca de la distribución de muestreo de un estadístico. En uno de los métodos se requiere hacer cálculos con base en las reglas de la probabilidad y en el otro es necesario llevar a cabo un experimento de simulación.

Cómo obtener la distribución de muestreo de un estadístico

Las reglas de probabilidad se pueden usar para obtener la distribución de un estadístico siempre que sea una función "bastante simple" de las X_i y haya relativamente pocos valores de X diferentes en la población, o bien la distribución poblacional tenga una forma "agradable". En los dos ejemplos siguientes se ilustran estas situaciones.

Ejemplo 5.20 Un gran centro de servicio automotriz cobra 40, 45 y 50 dólares por una afinación de automóviles de cuatro, seis y ocho cilindros, respectivamente. Si 20% de las afinaciones se hacen en automóviles de cuatro cilindros, 30% en autos de seis cilindros y 50% en los de ocho, entonces la distribución de probabilidad del ingreso de una sola afinación elegida al azar está dada por

x	40	45	50
$p(x)$	0.2	0.3	0.5

con $\mu = 46.5$, $\sigma^2 = 15.25$ (5.2)

Suponga que en un determinado día sólo dos servicios requieren afinación. Sea X_1 = ingreso obtenido de la primera afinación y X_2 = ingreso de la segunda. Suponga que X_1 y X_2 son in-

dependientes, cada una con la distribución de probabilidad mostrada en (5.2) [de modo que X_1 y X_2 constituyen una muestra aleatoria de la distribución (5.2)]. En la tabla 5.2 se listan los pares posibles (x_1, x_2) la probabilidad de cada uno [calculada por medio de (5.2) y la suposición de independencia] y los valores de $\bar{x}$ y s^2 resultantes. Ahora, para obtener la distribución de probabilidad de $\bar{X}$, el ingreso promedio muestral por afinación, se debe considerar cada valor posible de $\bar{x}$ y calcular su probabilidad. Por ejemplo, $\bar{x} = 45$ ocurre tres veces en la tabla con probabilidades 0.10, 0.09 y 0.10, así que

$$p_{\bar{x}}(45) = P(\bar{X} = 45) = 0.10 + 0.09 + 0.10 = 0.29$$

Tabla 5.2 Resultados, probabilidades y valores de $\bar{x}$ y s^2 para el ejemplo 5.20

x_1	x_2	$p(x_1, x_2)$	$\bar{x}$	s^2
40	40	0.04	40	0
40	45	0.06	42.5	12.5
40	50	0.10	45	50
45	40	0.06	42.5	12.5
45	45	0.09	45	0
45	50	0.15	47.5	12.5
50	40	0.10	45	50
50	45	0.15	47.5	12.5
50	50	0.25	50	0

De manera similar,

$$p_{S^2}(50) = P(S^2 = 50) = P(X_1 = 40, X_2 = 50 \quad \text{o} \quad X_1 = 50, X_2 = 40)$$
$$= 0.10 + 0.10 = 0.20$$

Las distribuciones de muestreo completas de $\bar{X}$ y S^2 aparecen en (5.3) y (5.4).

$\bar{x}$	40	42.5	45	47.5	50
$p_{\bar{X}}(\bar{x})$	0.04	0.12	0.29	0.30	0.25

(5.3)

s^2	0	12.5	50
$p_{S^2}(s^2)$	0.38	0.42	0.20

(5.4)

En la figura 5.7 se ilustra un histograma de probabilidad para la distribución original (5.2) y la distribución de $\bar{X}$ (5.3). En la figura se indica primero que la media (valor esperado) de la distribución de $\bar{X}$ es igual a la media 46.5 de la distribución original, puesto que ambos histogramas aparecen centrados en el mismo lugar.

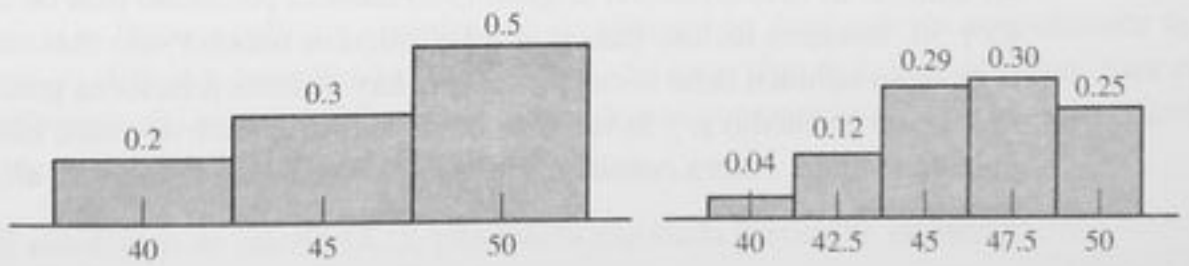


Figura 5.7 Histogramas de probabilidad para la distribución subyacente y la distribución de $\bar{X}$ del ejemplo 5.20

De (5.3),

$$\mu_{\bar{X}} = E(\bar{X}) = \sum \bar{x} p_{\bar{X}}(\bar{x}) = (40)(0.04) + \cdots + (50)(0.25) = 46.5 = \mu$$

Segundo, al parecer la distribución de $\bar{X}$ tiene dispersión más pequeña (variabilidad) que la distribución original, puesto que la masa de probabilidad se movió hacia la media. De nuevo de (5.3),

$$\begin{aligned}\sigma_{\bar{X}}^2 = V(\bar{X}) &= \sum \bar{x}^2 \cdot p_{\bar{X}}(\bar{x}) - \mu_{\bar{X}}^2 \\ &= (40)^2(0.04) + \cdots + (50)^2(0.25) - (46.5)^2 \\ &= 7.625 = \frac{15.25}{2} = \frac{\sigma^2}{2}\end{aligned}$$

La varianza de $\bar{X}$ es precisamente la mitad de la varianza original (debido a que $n = 2$).

El valor medio de S^2 es

$$\begin{aligned}\mu_{S^2} = E(S^2) &= \sum s^2 \cdot p_{S^2}(s^2) \\ &= (0)(0.38) + (12.5)(0.42) + (50)(0.20) = 15.25 = \sigma^2\end{aligned}$$

Es decir, la distribución de muestreo de $\bar{X}$ se centra en la media de la población μ, y la distribución de muestreo de S^2 se centra en la varianza poblacional σ^2.

Si las cuatro afinaciones se hubieran llevado a cabo en el día de interés, el ingreso promedio muestral $\bar{X}$ se basaría en una muestra aleatoria de cuatro X_i, cada una con la distribución (5.2). Con más cálculos se obtiene finalmente la fmp de $\bar{X}$ para $n = 4$ como

$\bar{x}$	40	41.25	42.5	43.75	45	46.25	47.5	48.75	50
$p_{\bar{X}}(\bar{x})$	0.0016	0.0096	0.0376	0.0936	0.1761	0.2340	0.2350	0.1500	0.0625

De esto, para $n = 4$, $\mu_{\bar{X}} = 46.50 = \mu$ y $\sigma_{\bar{X}}^2 = 3.8125 = \sigma^2/4$. La figura 5.8 es un histograma de probabilidad de esta fmp.

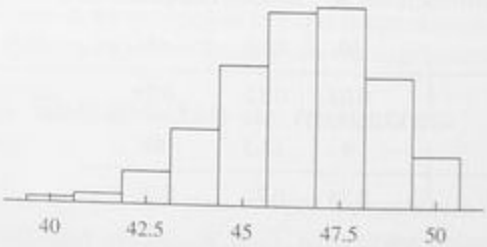


Figura 5.8 Histograma de probabilidad para $\bar{X}$ con base en $n = 4$ en el ejemplo 5.20 ∎

El ejemplo 5.20 debe hacer pensar primero que todo el cálculo de $p_{\bar{X}}(\bar{x})$ y $p_{S^2}(s^2)$ puede ser tedioso. Si la distribución original (5.2) hubiera permitido más de tres valores posibles 40, 45 y 50, entonces incluso para $n = 2$ los cálculos habrían sido más complejos. Sin embargo, el ejemplo también debe hacer pensar que hay algunas relaciones generales entre $E(\bar{X})$, $V(\bar{X})$, $E(S^2)$, y el promedio μ y la varianza σ^2 de la distribución original. Éstas se expresan en la siguiente sección. Ahora considere un ejemplo en el que la muestra aleatoria se extrae de una distribución continua.

Ejemplo 5.21 El tiempo que toma atender a un cliente en la caja de un comercio pequeño es una variable aleatoria que tiene una distribución exponencial con parámetro λ. Suponga que X_1 y X_2 son los

tiempos de servicio para dos clientes distintos, bajo el supuesto de que son independientes entre sí. Considere el tiempo de servicio total $T_o = X_1 + X_2$ para los dos clientes, también un estadístico. La fda de T_o es, para $t \geq 0$,

$$\begin{aligned} F_{T_o}(t) = P(X_1 + X_2 \leq t) &= \iint_{\{(x_1, x_2):\, x_1 + x_2 \leq t\}} f(x_1, x_2)\, dx_1\, dx_2 \\ &= \int_0^t \int_0^{t-x_1} \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda x_2}\, dx_2\, dx_1 = \int_0^t [\lambda e^{-\lambda x_1} - \lambda e^{-\lambda t}]\, dx_1 \\ &= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} \end{aligned}$$

La región de integración se ilustra en la figura 5.9.

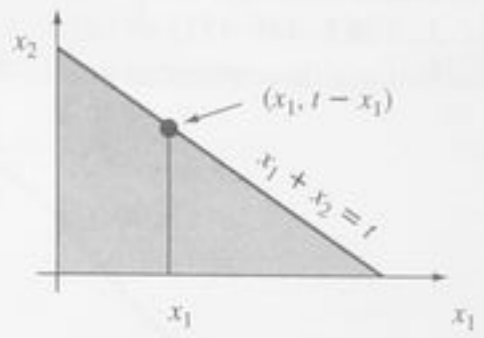


Figura 5.9 Región de integración para obtener la fda de T_o del ejemplo 5.21

La fdp de T_o se obtiene al derivar $F_{T_o}(t)$:

$$f_{T_o}(t) = \begin{cases} \lambda^2 t e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{5.5}$$

Ésta es una fdp gamma ($\alpha = 2$ y $\beta = 1/\lambda$). La fdp de $\overline{X} = T_o/2$ se obtiene de la relación $\{\overline{X} \leq \overline{x}\}$ si y sólo si $\{T_o \leq 2\overline{x}\}$ como

$$f_{\overline{X}}(\overline{x}) = \begin{cases} 4\lambda^2 \overline{x} e^{-2\lambda \overline{x}} & \overline{x} \geq 0 \\ 0 & \overline{x} < 0 \end{cases} \tag{5.6}$$

La media y la varianza de la distribución exponencial subyacente son $\mu = 1/\lambda$ y $\sigma^2 = 1/\lambda^2$. De las expresiones (5.5) y (5.6), se puede comprobar que $E(\overline{X}) = 1/\lambda$, $V(\overline{X}) = 1/(2\lambda^2)$, $E(T_o) = 2/\lambda$ y $V(T_o) = 2/\lambda^2$. Estos resultados de nuevo hacen pensar en algunas relaciones generales entre medias y varianzas de $\overline{X}$, T_o, y la distribución subyacente. ■

Experimentos de simulación

El segundo método para obtener información acerca de la distribución de muestreo de un estadístico es llevar a cabo un experimento de simulación. Este método normalmente se utiliza cuando es difícil o complicado obtener algo vía las reglas de la probabilidad. Esta clase de experimento casi siempre se hace con ayuda de una computadora. Se deben especificar las siguientes características de un experimento:

1. El estadístico de interés ($\overline{X}$, S, una media recortada particular, etcétera).
2. La distribución poblacional (normal con $\mu = 100$ y $\sigma = 15$, uniforme con límite inferior $A = 5$ y límite superior $B = 10$, etcétera).

3. El tamaño de muestra n (p. ej., $n = 10$ o $n = 50$)
4. El número de réplicas k (p. ej., $k = 500$)

Luego, utilice una computadora para obtener k muestras aleatorias distintas, cada una de tamaño n, a partir de la distribución poblacional designada. Para muestra, calcule el valor del estadístico y construya un histograma de los k valores calculados. Este histograma proporciona la distribución de muestreo *aproximada* del estadístico. Mientras más grande sea el valor de k, mejor tiende a ser la aproximación (la distribución de muestreo real surge cuando $k \to \infty$). En la práctica, $k = 500$ o 1000 es suficiente si el estadístico es "bastante simple".

Ejemplo 5.22 La distribución poblacional del primer estudio de simulación es normal con $\mu = 8.25$ y $\sigma = 0.75$, como se ilustra en la figura 5.10. [En el artículo "Platelet Size in Myocardial Infarction" (*British Med. J.*, 1983: 449-451) se indica esta distribución para el volumen de plaquetas en individuos sin historia de problemas graves del corazón.]

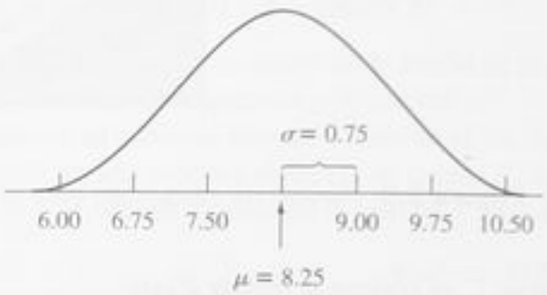


Figura 5.10 Distribución normal, con $\mu = 8.25$ y $\sigma = 0.75$

En realidad se efectuaron cuatro experimentos distintos, con 500 réplicas cada uno. En el primer experimento, se generaron 500 muestras de $n = 5$ observaciones por medio de MINITAB, y los tamaños de muestra para las otras tres fueron $n = 10$, $n = 20$ y $n = 30$, respectivamente. Para cada muestra se calculó la media y los histogramas resultantes de los valores de $\bar{x}$ se ilustran en la figura 5.11.

Lo primero que se debe observar en los histogramas es su forma. Para una aproximación razonable, cada uno de los cuatro se asemeja a una curva normal. La semejanza sería incluso más sorprendente si cada histograma se hubiera basado en muchos más que 500 valores de $\bar{x}$. Segundo, cada histograma está centrado en aproximadamente 8.25, la media de la población en la que se realiza el muestreo. Si los histogramas se hubieran basado en una secuencia interminable de valores de $\bar{x}$ sus centros habrían sido exactamente la media poblacional, 8.25.

El aspecto final que hay que observar en los histogramas es su dispersión relativa entre sí. Mientras más pequeño sea el valor de n, mayor es el grado de dispersión de la distribución de muestreo respecto al valor medio. Por esta razón los histogramas para $n = 20$ y $n = 30$ se basan en intervalos de clase más estrechos que los correspondientes a tamaños de muestra más pequeños. Para tamaños de muestra más grandes, la mayor parte de los valores de $\bar{x}$ están bastante cerca de 8.25. Éste es el efecto de promediar. Cuando n es pequeña, un solo valor inusual de x produce un valor de $\bar{x}$ alejado del centro. Con un tamaño de muestra más grande, cualesquiera valores inusuales de x, cuando se promedian con otros valores de la muestra, aún tienden a producir un valor de $\bar{x}$ cercano a μ. La combinación de estas ideas produce un resultado que debe apelar a la intuición: **la $\bar{X}$ basada en una n grande tiende a estar más cerca de μ que la $\bar{X}$ basada en una n pequeña.**

Ejemplo

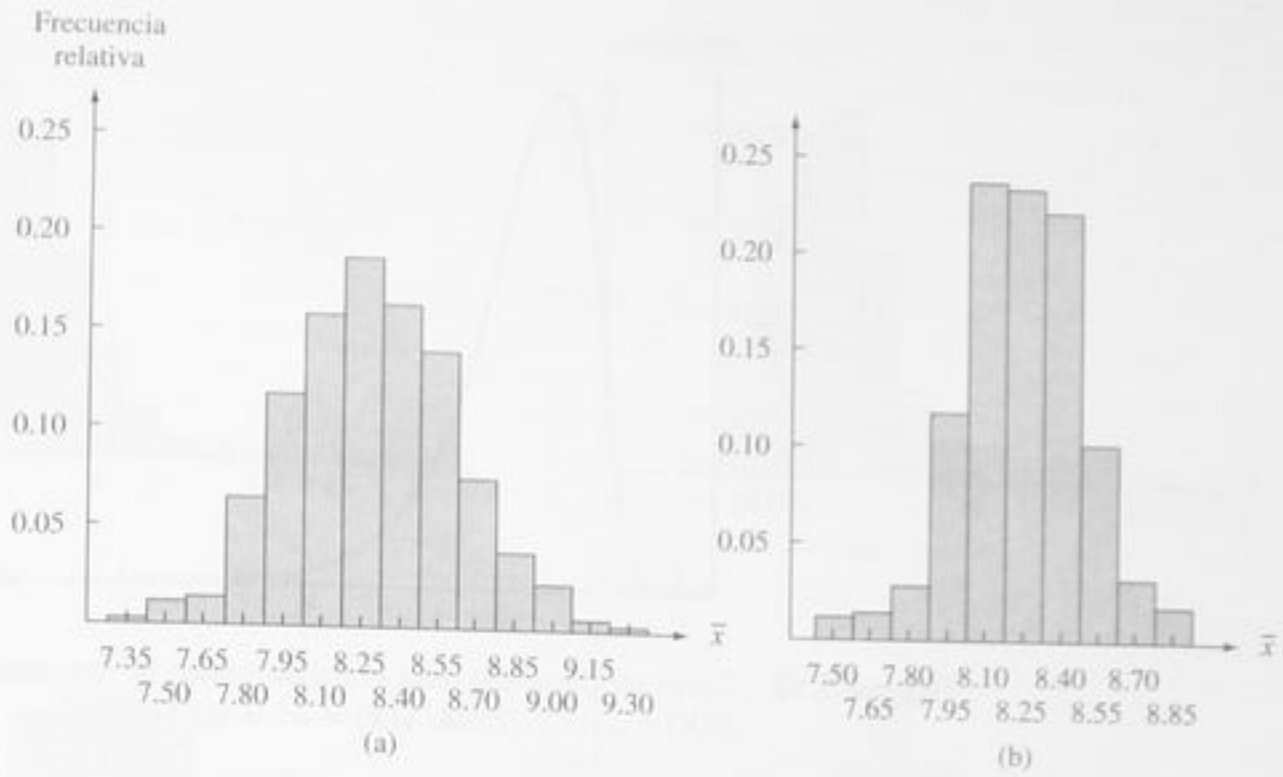


Figura 5.11 Histogramas muestrales para $\bar{x}$ basada en 500 muestras; cada una consta de n observaciones: (a) $n = 5$; (b) $n = 10$; (c) $n = 20$; (d) $n = 30$ ■

Ejemplo 5.23 Considere un experimento de simulación en el que la distribución poblacional está bastante sesgada. En la figura 5.12 se muestra la curva de densidad para los tiempos de vida útil de cierto tipo de control electrónico [ésta es en realidad una distribución lognormal con $E(\ln(X)) = 3$ y $V(\ln(X)) = 0.4$]. De nuevo la estadística de interés es la media muestral $\bar{X}$. En el experimento se utilizan 500 duplicaciones y se consideran los mismos tamaños de muestra que en el ejemplo 5.22. Estos histogramas resultantes junto con una gráfica de probabilidad normal de MINITAB para los 500 valores de $\bar{x}$ con base en $n = 30$ se muestran en la figura 5.13.

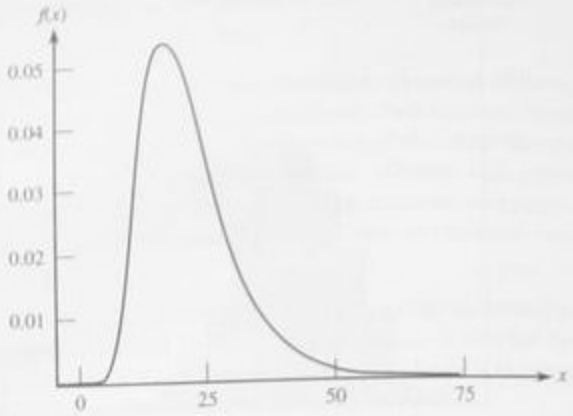


Figura 5.12 Curva de densidad para el experimento de simulación del ejemplo 5.23 [$E(X) = \mu = 21.7584$, $V(X) = \sigma^2 = 82.1449$]

A diferencia del caso normal, estos histogramas difieren en la forma. En particular, se vuelven cada vez menos sesgados a medida que aumenta el tamaño n de la muestra. El promedio de los 500 valores de $\bar{x}$ para los cuatro tamaños de muestra distintos está bastante cerca del valor medio de la distribución poblacional. Si cada histograma se hubiera basado en una secuencia interminable de valores de $\bar{x}$ y no sólo en 500, los cuatro habrían estado centrados en 21.7584. Así, diferentes valores de n cambian la forma, pero no el centro de la distribución de muestreo de $\bar{X}$. La comparación de los cuatro histogramas de la figura 5.13 también muestra que cuando aumenta n, disminuye la dispersión de los histogramas. Si n aumenta se obtiene un mayor grado de concentración respecto a la media poblacional y hace que el histograma se parezca más a una curva normal. El histograma de la figura 5.13(d) y la gráfica de probabilidad normal de la figura 5.13(e) proporcionan pruebas convincentes de que un tamaño de muestra de $n = 30$ es suficiente para vencer el sesgo de la distribución poblacional y da una distribución de muestreo de $\bar{X}$ más o menos normal. ■

Ejercicios | Sección 5.3 (37 a 45)

37. Una marca particular de jabón para lavavajillas se vende en tres tamaños: 25, 40 y 65 onzas. Veinte por ciento de los compradores seleccionan la caja de 25 onzas, 50% la de 40 onzas y el 30% restante la caja de 65 onzas. Sean X_1 y X_2 los tamaños de caja que seleccionan dos compradores elegidos de manera independiente.

a. Determine la distribución de muestreo de $\bar{X}$, calcule $E(\bar{X})$, y compare con μ.

b. Determine la distribución de muestreo de la varianza muestral S^2, calcule $E(S^2)$, y compare con σ^2.

38. En el trayecto de una persona hacia el trabajo hay dos semáforos. Sea X_1 el número de semáforos en los que se debe detener, y suponga que la distribución de X_1 es como sigue:

x_1	0	1	2
$p(x_1)$	0.2	0.5	0.3

$\mu = 1.1$, $\sigma^2 = 0.49$

Sea X_2 el número de semáforos en los que la persona se debe detener camino a casa; X_2 es independiente de X_1. Suponga que X_2 tiene la misma distribución que X_1, de modo que X_1, X_2 es una muestra aleatoria de tamaño $n = 2$.

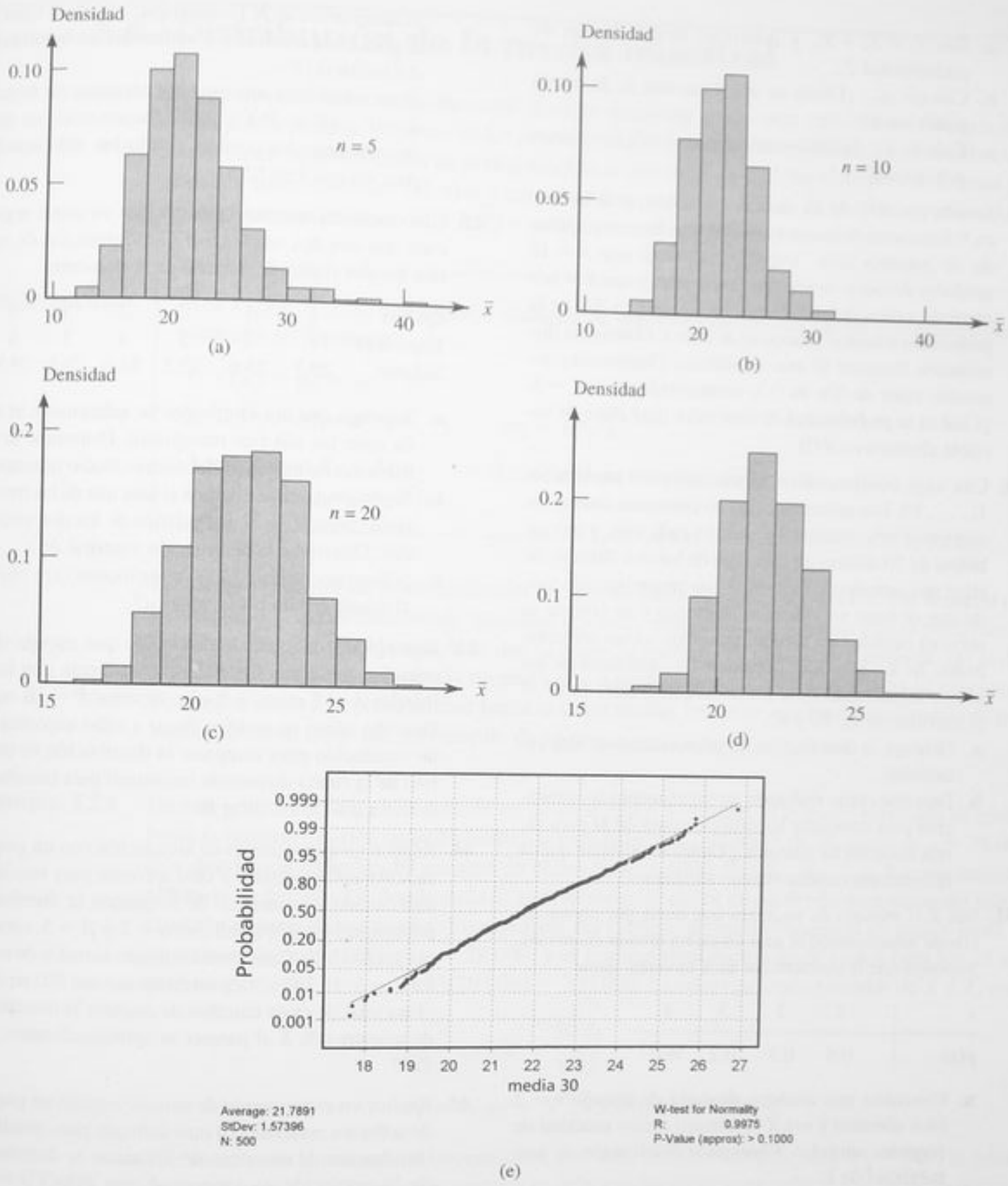


Figura 5.13 Resultados del experimento del ejemplo 5.23: (a) histograma de $\bar{x}$ para $n = 5$; (b) histograma de $\bar{x}$ para $n = 10$; (c) histograma de $\bar{x}$ para $n = 20$; (d) histograma de $\bar{x}$ para $n = 30$; (e) gráfica de probabilidad normal para $n = 30$ (de MINITAB)

a. Sea $T_o = X_1 + X_2$, y determine la distribución de probabilidad T_o.

b. Calcule μ_{T_o}. ¿Cómo se relaciona con μ, la media poblacional?

c. Calcule $\sigma^2_{T_o}$. ¿Cómo se relaciona con σ^2, la varianza poblacional?

39. Se sabe que 80% de las unidades de disco *zip* de la marca A funcionan de manera satisfactoria durante el periodo de garantía (son "éxitos"). Suponga que $n = 10$ unidades de disco se seleccionan al azar y sea X = número de éxitos de la muestra. El estadístico X/n es la proporción muestral (fracción) de éxitos. Obtenga la distribución muestral de este estadístico. [*Sugerencia:* un posible valor de X/n es 0.3, correspondiente a $X = 3$. ¿Cuál es la probabilidad de este valor (qué clase de variable aleatoria es X)?]

40. Una caja contiene diez sobres sellados numerados $1, \ldots, 10$. Los primeros cinco no contienen dinero, los siguientes tres contienen 5 dólares cada uno, y hay un billete de 10 dólares en cada uno de los dos últimos. Se elige una tamaño de muestra 3 *con* reemplazo (de modo que se tiene una muestra aleatoria) y se obtiene la máxima cantidad en cualquiera de los sobres seleccionados. Si X_1, X_2 y X_3 representan las cantidades de los sobres seleccionados, la estadística de interés es M = el máximo de X_1, X_2 y X_3.

a. Obtenga la distribución de probabilidad de este estadístico.

b. Describa cómo realizaría un experimento de simulación para comparar las distribuciones de M para varios tamaños de muestra. ¿Cómo se advierte que la distribución cambia cuando aumenta n?

41. Sea X el número de paquetes que envía por correo un cliente seleccionado al azar en cierta oficina de envíos. Suponga que la distribución de X es como sigue:

x	1	2	3	4
$p(x)$	0.4	0.3	0.2	0.1

a. Considere una muestra aleatoria de tamaño $n = 2$ (dos clientes) y sea $\bar{X}$ el número medio muestral de paquetes enviados. Obtenga la distribución de probabilidad de $\bar{X}$.

b. Refiérase al inciso (a) y calcule $P(\bar{X} \le 2.5)$.

c. Considere de nuevo una muestra aleatoria de tamaño $n = 2$, pero ahora centre su atención en el estadístico R = intervalo muestral (diferencia entre los valores máximo y mínimo de la muestra). Obtenga la distribución de R. [*Sugerencia:* calcule el valor de R para cada resultado, y utilice las probabilidades del inciso (a).]

d. Si se selecciona una muestra aleatoria de tamaño $n = 4$ ¿cuál es $P(\bar{X} \le 1.5)$? (*Sugerencia:* no tiene que enumerar los posibles resultados, sólo aquellos para los que $\bar{x} \le 1.5$.)

42. Una compañía mantiene tres oficinas en cierta región, cada una con dos empleados. La información de salarios anuales (miles de dólares) es la siguiente:

Oficina	1	1	2	2	3	3
Empleado	1	2	3	4	5	6
Salario	29.7	33.6	30.2	33.6	25.8	29.7

a. Suponga que dos empleados se seleccionan al azar de entre los seis (sin reemplazo). Determine la distribución de muestreo del salario medio muestral $\bar{X}$.

b. Suponga que se selecciona al azar una de las tres oficinas. Sean X_1 y X_2 los salarios de los dos empleados. Determine la distribución muestral de $\bar{X}$.

c. ¿Cómo se compara $E(\bar{X})$ de los incisos (a) y (b) con el salario medio poblacional μ?

43. Suponga que la cantidad de líquido que entrega cierta máquina tiene una distribución uniforme con límite inferior $A = 8$ onzas y límite superior $B = 10$ onzas. Describa cómo se podrían llevar a cabo experimentos de simulación para comparar la distribución de muestreo de la cuarta dispersión (muestral) para tamaños de muestra $n = 5$, 10, 20 y 30.

44. Realice un experimento de simulación con un paquete de software estadístico u otro software para estudiar la distribución de muestreo de $\bar{X}$ cuando la distribución poblacional es de Weibull con $\alpha = 2$ y $\beta = 5$, como en el ejemplo 5.19. Considere los cuatro tamaños de muestra $n = 5$, 10, 20 y 30, y en cada caso use 500 réplicas. ¿Para cuál de estos tamaños de muestra la distribución de muestreo de $\bar{X}$ al parecer es aproximadamente normal?

45. Realice un experimento de simulación con un paquete de software estadístico u otro software para estudiar la distribución de muestreo de $\bar{X}$ cuando la distribución de la población es lognormal con $E(\ln(X)) = 3$ y $V(\ln(X)) = 1$. Considere los cuatro tamaños de muestra $n = 10$, 20, 30 y 50, y en cada caso use 500 réplicas. ¿Para cuál de estos tamaños de muestra la distribución de muestreo de $\bar{X}$ al parecer es aproximadamente normal?

5.4 Distribución de la media muestral

La importancia de la media muestral $\overline{X}$ surge de su uso para sacar conclusiones acerca de la media poblacional μ. Algunos de los procedimientos inferenciales utilizados con más frecuencia, se basan en propiedades de la distribución de muestreo de $\overline{X}$. Una vista previa de estas propiedades apareció en los cálculos y experimentos de simulación de la sección anterior, donde se observaron relaciones entre $E(\overline{X})$ y μ también entre $V(\overline{X})$, σ^2 y n.

PROPOSICIÓN

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución con media μ y desviación estándar σ. Entonces

1. $E(\overline{X}) = \mu_{\overline{X}} = \mu$
2. $V(\overline{X}) = \sigma_{\overline{X}}^2 = \sigma^2/n$ y $\sigma_{\overline{X}} = \sigma/\sqrt{n}$

Además, con $T_o = X_1 + \cdots + X_n$ (el total de muestra), $E(T_o) = n\mu$, $V(T_o) = n\sigma^2$ y $\sigma_{T_o} = \sqrt{n}\sigma$.

Las demostraciones de estos resultados se aplazan hasta la siguiente sección. De acuerdo con el resultado 1, la distribución de muestreo (es decir, probabilidad) de $\overline{X}$ se centra precisamente en la media de la población de la que se eligió la muestra. El resultado 2 muestra que la distribución de $\overline{X}$ se concentra más respecto a μ a medida que crece el tamaño de la muestra n. Promediar mueve la probabilidad hacia la parte media, en tanto que totalizar dispersa la probabilidad en un intervalo más amplio de valores.

Ejemplo 5.24 En una prueba de fatiga a la tensión en un espécimen de titanio, el número esperado de ciclos hasta la primera emisión acústica (utilizada para indicar la iniciación de fractura) es $\mu = 28\,000$ y la desviación estándar del número de ciclos es $\sigma = 5000$. Sea $X_1, X_2, \ldots, X_{25}$ una muestra aleatoria de tamaño 25, donde cada X_i es el número de ciclos en un espécimen diferente seleccionado al azar. Entonces el valor esperado del número promedio muestral de ciclos hasta la primera emisión es $E(\overline{X}) = \mu = 28\,000$, y el número total esperado de ciclos para los 25 especímenes es $E(T_o) = n\mu = 25(28\,000) = 700\,000$. Las desviaciones estándar de $\overline{X}$ y T_o son

$$\sigma_{\overline{X}} = \sigma/\sqrt{n} = \frac{5000}{\sqrt{25}} = 1000$$

$$\sigma_{T_o} = \sqrt{n}\sigma = \sqrt{25}(5000) = 25\,000$$

Si el tamaño de la muestra se incrementa a $n = 100$, $E(\overline{X})$ no cambia, pero $\sigma_{\overline{X}} = 500$, la mitad de su valor previo (el tamaño de muestra se debe cuadruplicar para reducir a la mitad la desviación estándar de $\overline{X}$). ■

Caso de una distribución poblacional normal

Considerando de nuevo el experimento de simulación del ejemplo 5.22, se ve que cuando la distribución poblacional es normal, cada histograma de valores de $\bar{x}$ se aproxima bien mediante una curva normal. Se deduce el resultado preciso.

PROPOSICIÓN

Sea $X_1, X_2, \ldots, X_n$ una variable aleatoria de una distribución *normal* con media μ y desviación estándar σ. Entonces para *cualquier* n, $\overline{X}$ tiene una distribución normal (con media μ y desviación estándar $\sigma/\sqrt{n}$), al igual que T_o (con media $n\mu$ y desviación estándar $\sqrt{n}\sigma$).*

Se sabe todo lo que hay que saber acerca de las distribuciones de $\overline{X}$ y T_o cuando la distribución de la población es normal. En particular, probabilidades como $P(a \le \overline{X} \le b)$ y $P(c \le T_o \le d)$ se pueden obtener simplemente por estandarización. En la figura 5.14 se ilustra la proposición.

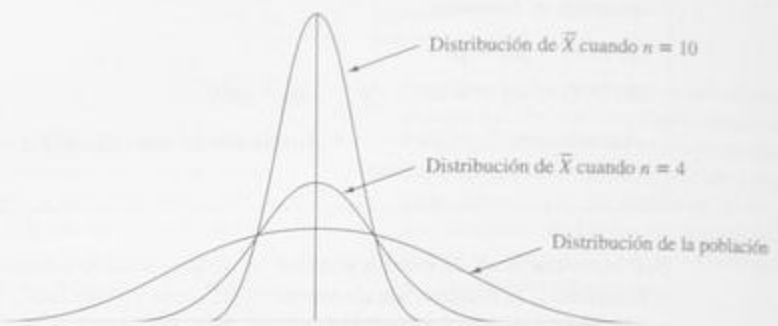


Figura 5.14 Distribución de población normal y distribuciones de muestreo de $\overline{X}$

Ejemplo 5.25 El tiempo que tarda una rata de cierta especie elegida al azar para encontrar su camino por un laberinto es una *va* con distribución normal con $\mu = 1.5$ min y $\sigma = 0.35$ min. Suponga que se eligen cinco ratas. Sean $X_1, \ldots, X_5$ sus tiempos en el laberinto. Suponiendo que las X_i son una variable aleatoria de esta distribución normal, ¿cuál es la probabilidad de que el tiempo total $T_o = X_1 + \cdots + X_5$ para las cinco ratas esté entre 6 y 8 minutos? Por la proposición, T_o tiene una distribución normal con $\mu_{T_o} = n\mu = 5(1.5) = 7.5$ y varianza $\sigma_{T_o}^2 = n\sigma^2 = 5(0.1225) = 0.6125$, por lo tanto $\sigma_{T_o} = 0.783$. Para estandarizar T_o, reste μ_{T_o} y divida entre σ_{T_o}:

$$P(6 \le T_o \le 8) = P\left(\frac{6 - 7.5}{0.783} \le Z \le \frac{8 - 7.5}{0.783}\right)$$
$$= P(-1.92 \le Z \le 0.64) = \Phi(0.64) - \Phi(-1.92) = 0.7115$$

La determinación de la probabilidad de que el tiempo promedio muestral $\overline{X}$ (una variable con distribución normal) es a lo sumo dos min requiere $\mu_{\overline{X}} = \mu = 1.5$ y $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 0.35/\sqrt{5} = 0.1565$. Entonces,

$$P(\overline{X} \le 2.0) = P\left(Z \le \frac{2.0 - 1.5}{0.1565}\right) = P(Z \le 3.19) = \Phi(3.19) = 0.9993$$

■

*Una demostración del resultado para T_o cuando $n = 2$ es posible por medio del método del ejemplo 5.21, pero los detalles son confusos. El resultado general se demuestra por lo común por medio de una herramienta teórica llamada *función generadora de momentos*. Para obtener más información se puede consultar una de las referencias del capítulo.

Teorema del límite central

Si las X_i están distribuidas normalmente, entonces también $\overline{X}$ lo está para cada tamaño de muestra n. El experimento de simulación del ejemplo 5.23, hace pensar que cuando la distribución poblacional es muy poco normal, promediar produce una distribución con más forma de campana que la que está siendo muestreada. Una conjetura razonable es que si n es grande, una curva normal adecuada aproxima la distribución real de $\overline{X}$. El enunciado formal de este resultado es el teorema más importante de la probabilidad.

TEOREMA

Teorema del límite central (TLC)

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2. Entonces si n es suficientemente grande, $\overline{X}$ tiene aproximadamente una distribución normal con $\mu_{\overline{X}} = \mu$ y $\sigma_{\overline{X}}^2 = \sigma^2/n$, y T_o tiene una distribución casi normal con $\mu_{T_o} = n\mu$, $\sigma_{T_o}^2 = n\sigma^2$. Mientras más grande sea el valor de n, mejor es la aproximación.

En la figura 5.15 se ilustra el Teorema del límite central. Según este teorema, cuando n es grande y se desea calcular una probabilidad como $P(a \le \overline{X} \le b)$, sólo se necesita "pretender" que $\overline{X}$ es normal, estandarizarla y usar la tabla normal. La respuesta resultante casi es correcta. La respuesta exacta se podría obtener sólo si se determina primero la distribución de $\overline{X}$, de modo que el TLC proporciona un atajo en verdad impresionante. La demostración del teorema tiene que ver mucho con matemáticas avanzadas.

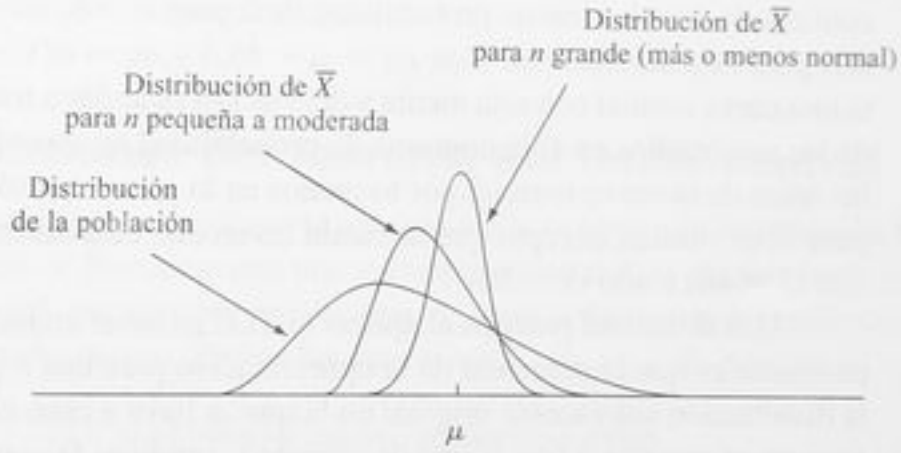


Figura 5.15 Ilustración del Teorema del límite central

Ejemplo 5.26 Cuando se prepara un lote de cierto producto, la cantidad de determinada impureza en el lote es una variable aleatoria con valor medio 4.0 g y una desviación estándar de 1.5 g. Si 50 lotes se preparan de forma independiente, ¿cuál es la probabilidad (aproximada) de que la cantidad promedio muestral de impureza $\overline{X}$ esté entre 3.5 y 3.8 g? De acuerdo con una regla empírica que se enunciará en breve, $n = 50$ es suficientemente grande para que sea aplicable el TLC. Entonces $\overline{X}$ tiene más o menos una distribución normal con valor medio $\mu_{\overline{X}} = 4.0$ y $\sigma_{\overline{X}} = 1.5/\sqrt{50} = 0.2121$, por lo tanto

$$P(3.5 \le \overline{X} \le 3.8) \approx P\left(\frac{3.5 - 4.0}{0.2121} \le Z \le \frac{3.8 - 4.0}{0.2121}\right)$$

$$= \Phi(-0.94) - \Phi(-2.36) = 0.1645 \qquad \blacksquare$$

Ejemplo 5.27 Cierta organización de consumidores informa de manera habitual el número de defectos mayores para cada automóvil nuevo que prueba. Suponga que el número de tales defectos para cierto automóvil es una variable aleatoria con valor medio 3.2 y desviación estándar 2.4. Entre 100 automóviles elegidos al azar de este modelo, ¿qué tan probable es que el número promedio muestral de defectos mayores pase de 4? Sea X_i el número de defectos mayores para el i-ésimo automóvil de la muestra aleatoria. Observe que X_i es una *va* discreta, pero el TLC es aplicable si la variable de interés es discreta o continua. También, aunque el hecho de que la desviación estándar de esta variable no negativa es bastante grande con respecto al valor medio, hace pensar que su distribución tiene un sesgo positivo, el tamaño de muestra grande implica que X tiene más o menos una distribución normal. Con $\mu_{\bar{X}} = 3.2$ y $\sigma_{\bar{X}} = 0.24$,

$$P(\bar{X} > 4) \approx P\left(Z > \frac{4 - 3.2}{0.24}\right) = 1 - \Phi(3.33) = 0.0004$$

■

El TLC da una idea de por qué muchas variables aleatorias tienen distribuciones de probabilidad que son más o menos normales. Por ejemplo, el error de medición en un experimento científico se puede considerar como la suma de varias perturbaciones y errores subyacentes de magnitud pequeña.

Aunque pronto será evidente la utilidad del TLC para la inferencia, el contenido intuitivo del resultado ofrece dificultad a muchos estudiantes principiantes. Volviendo de nuevo a la figura 5.7, el histograma de probabilidad de la izquierda es una representación de la distribución en la que se lleva a cabo el muestreo. Es discreto y bastante sesgado, de modo que no se parece en absoluto a una distribución normal. La distribución de $\bar{X}$ para $n = 2$ comienza a exhibir cierta simetría, y esto es incluso más notable para $n = 4$ en la figura 5.8. La figura 5.16 contiene la distribución de probabilidad de $\bar{X}$ para $n = 8$, así como el histograma de probabilidad para esta distribución. Con $\mu_{\bar{X}} = \mu = 46.5$ y $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 3.905/\sqrt{8} = 1.38$, si se ajusta una curva normal con esta media y desviación estándar a través del histograma de $\bar{X}$, las áreas de los rectángulos en el histograma de probabilidad se aproximan más o menos bien mediante las áreas de la curva normal, por lo menos en la parte central de la distribución. La ilustración para T_o es similar, excepto que la escala horizontal está mucho más dispersa, con T_o que va de 320 ($\bar{x} = 40$) a 400 ($\bar{x} = 50$).

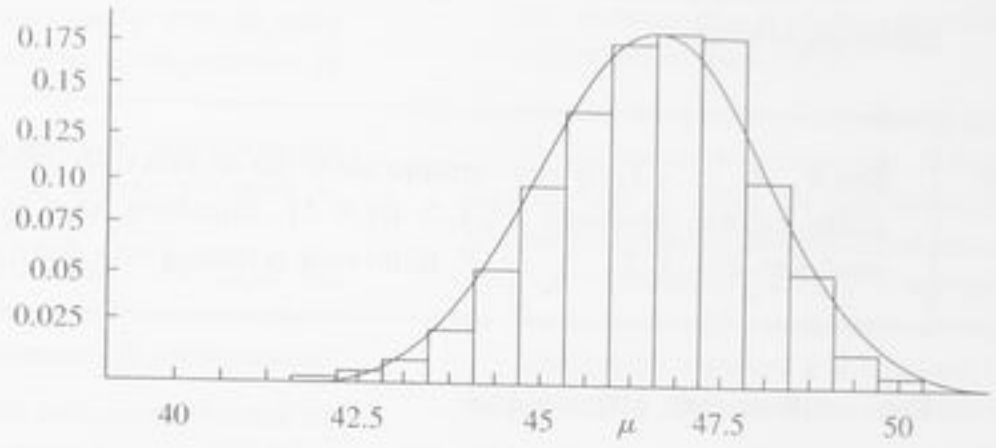
Una dificultad práctica al aplicar el TLC es saber cuándo n es suficientemente grande. El problema es que la precisión de la aproximación para una n particular depende de la forma de la distribución subyacente original en la que se lleva a cabo el muestreo. Si la distribución subyacente es cercana a una forma de campana, entonces la aproximación será buena incluso para una n pequeña, mientras que si está lejos de la forma de campana, entonces se requiere una n grande. Se usará la siguiente regla empírica, que con frecuencia es un poco conservadora.

> **Regla empírica**
>
> Si $n > 30$, se puede usar el Teorema del límite central.

Hay distribuciones poblacionales para las cuales incluso una n de 40 o 50 es insuficiente, pero estas distribuciones rara vez se encuentran en la práctica.

$\bar{x}$	40	40.625	41.25	41.875	42.5	43.125
$p(\bar{x})$	0.0000	0.0000	0.0003	0.0012	0.0038	0.0112
$\bar{x}$	43.75	44.375	45	45.625	46.25	46.875
$p(\bar{x})$	0.0274	0.0556	0.0954	0.1378	0.1704	0.1746
$\bar{x}$	47.5	48.125	48.75	49.375	50	
$p(\bar{x})$	0.1474	0.0998	0.0519	0.0188	0.0039	

(a)



(b)

Figura 5.16 (a) Distribución de probabilidad de $\overline{X}$ para $n = 8$; (b) histograma de probabilidad y distribución normal para la distribución de $\overline{X}$ cuando la distribución original es como en el ejemplo 5.20

Otras aplicaciones del Teorema del límite central

El TLC se puede usar para justificar la aproximación normal a la distribución binomial analizada en el capítulo 4. Recuerde que una variable binomial X es el número de éxitos en un experimento binomial que consiste en ensayos de éxito o fracaso independientes con $p = P(S)$ para un determinado ensayo. Defina las nuevas *va* $X_1, X_2, \ldots, X_n$ por

$$X_i = \begin{cases} 1 & \text{si el } i\text{-ésimo ensayo produce un éxito} \\ 0 & \text{si el } i\text{-ésimo ensayo produce un fracaso} \end{cases} (i = 1, \ldots, n)$$

Debido a que los ensayos son independientes y $P(S)$ es constante de un ensayo a otro, las X_i son independientes y están distribuidas de manera idéntica (una muestra aleatoria de una distribución de Bernoulli). El TLC implica entonces que si n es suficientemente grande, tanto la suma como el promedio de las X_i tienen distribuciones más o menos normales. Cuando se suman las X_i se agrega un 1 a cada S que ocurre y un cero a cada F, así $X_1 + \cdots + X_n = X$. La media muestral de las X_i es X/n, la proporción de éxitos. Es decir, X y X/n son aproximadamente normales cuando n es grande. El tamaño de muestra necesario para esta aproximación depende del valor de p: cuando p es cercana a 0.5, la distribución de cada X_i es simétrica desde un punto de vista razonable (véase la figura 5.17), mientras que la distribución es bastante sesgada cuando p es cercana a 0 o 1. Si se usa la aproximación sólo cuando $np \geq 10$ y $n(1 - p) \geq 10$ se asegura que n es suficientemente grande para superar cualquier sesgo en la distribución de Bernoulli subyacente.

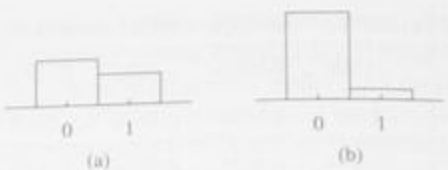


Figura 5.17 Dos distribuciones de Bernoulli: (a) $p = 0.4$ (razonablemente simétrica), (b) $p = 0.1$ (muy sesgada)

Recuerde de la sección 4.5 que X tiene una distribución lognormal si $\ln(X)$ tiene una distribución normal.

PROPOSICIÓN

Sea $X_1, X_2, \ldots, X_n$ es una variable aleatoria de una distribución para la cual sólo son posibles valores positivos [$P(X_i > 0) = 1$]. Entonces, si n es suficientemente grande, el producto $Y = X_1X_2 \cdot \cdots \cdot X_n$ tiene más o menos una distribución lognormal.

Para verificar esto, observe que

$$\ln(Y) = \ln(X_1) + \ln(X_2) + \cdots + \ln(X_n)$$

Puesto que $\ln(Y)$ es una suma de *va* independientes y distribuidas de manera idéntica [el ln de las (X_i)], es más o menos normal cuando n es grande, así que Y tiene más o menos una distribución lognormal. Como ejemplo de la aplicabilidad de este resultado, Bury (*Statistical Models in Applied Science*, Wiley, p. 590) argumenta que el proceso de daño en el flujo plástico y la propagación de fisura es un proceso multiplicativo, de modo que las variables como la elongación porcentual y la resistencia a la ruptura tienen distribuciones lognormales aproximadas.

Ejercicios | Sección 5.4 (46 a 57)

46. El diámetro interior de un anillo de pistón seleccionado al azar es una variable aleatoria con valor medio de 12 cm y desviación estándar de 0.04 cm.

a. Si $\overline{X}$ es el diámetro medio muestral para una muestra aleatoria de $n = 16$ anillos, ¿dónde está centrada la distribución de muestreo de $\overline{X}$ y cuál es la desviación estándar de la distribución de $\overline{X}$?

b. Conteste las preguntas planteadas en el inciso (a) para un tamaño de muestra de $n = 64$ anillos.

c. ¿Para cuál de las dos muestras aleatorias, la del inciso (a) o la del inciso (b), $\overline{X}$ tiene más probabilidades de estar entre 0.01 cm de 12 cm? Explique su razonamiento.

47. Refiérase al ejercicio 46. Suponga que la distribución del diámetro es normal.

a. Calcule $P(11.99 \leq \overline{X} \leq 12.01)$ cuando $n = 16$.

b. ¿Cuán probable es que el diámetro medio muestral pase de 12.01 cuando $n = 25$?

48. Sea que $X_1, X_2, \ldots, X_{100}$ representan los pesos netos reales de 100 bolsas de fertilizante de 50 lb seleccionadas al azar.

a. Si el peso de cada bolsa es 50 y la varianza es 1, calcule $P(49.75 \leq \overline{X} \leq 50.25)$ (aproximadamente) usando el TLC.

b. Si el peso esperado es 49.8 lb en vez de 50 de modo que en promedio las bolsas están incompletas, calcule $P(49.75 \leq \overline{X} \leq 50.25)$.

49. Hay 40 estudiantes en una clase básica de estadística. Con base en años de experiencia, el maestro sabe que el tiempo necesario para evaluar un primer examen elegido

al azar es una variable aleatoria con un valor esperado de seis minutos y una desviación estándar de seis minutos.

a. Si los tiempos de evaluación son independientes y el maestro comienza a calificar a las 6:50 P.M. sin interrupciones, ¿cuál es la probabilidad (aproximada) de que termine de calificar antes de que comiencen las noticias de las 11:00 P.M. en la televisión?

b. Si la sección de deportes comienza a las 11:10, ¿cuál es la probabilidad de que se pierda una parte de la información si espera hasta calificar el último examen antes de encender la televisión?

50. La resistencia a la ruptura de un remache tiene un valor medio de 10 000 lb/pulg^2 y una desviación estándar de 500 lb/pulg^2.

a. ¿Cuál es la probabilidad de que la resistencia media muestral a la ruptura para una muestra aleatoria de 40 remaches esté entre 9900 y 10 200?

b. Si el tamaño muestral hubiera sido 15 en vez de 40, ¿podría haberse calculado la probabilidad pedida en el inciso (a) a partir de la información suministrada?

51. El tiempo que tarda un solicitante seleccionado al azar en llenar cierta forma para una hipoteca tiene una distribución normal con valor medio de 10 minutos y desviación estándar de dos minutos. Si en un día determinado cinco individuos llenan una forma y seis en otro, ¿cuál es la probabilidad de que la cantidad de tiempo promedio de la muestra de cada día sea a lo sumo 11 minutos?

52. La duración de cierto tipo de batería tiene una distribución normal con valor medio de 10 horas y desviación estándar de una hora. Hay cuatro baterías en un paquete. ¿Cuál es el valor del tiempo de vida útil de modo que la duración total de las baterías de un paquete exceda ese valor en sólo 5% de los paquetes?

53. Se sabe que la dureza Rockwell de pernos de cierto tipo tiene un valor medio de 50 y desviación estándar de 1.2.

a. Si la distribución es normal, ¿cuál es la probabilidad de que la dureza muestral media para una muestra aleatoria de nueve pernos sea por lo menos 51?

b. ¿Cuál es la probabilidad (aproximada) de que la dureza muestral media para una muestra aleatoria de 40 pernos sea al menos 51?

54. Suponga que la densidad del sedimento (g/cm) de un espécimen seleccionado al azar, de cierta región, tiene una distribución normal con media 2.65 y desviación estándar 0.85 (sugerida en "Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants", *Water Research*, 1984: 1169-1174).

a. Si se selecciona una muestra aleatoria de 25 especímenes, ¿cuál es la probabilidad de que la densidad promedio de sedimento muestral sea a lo sumo 3.00? ¿De que esté entre 2.65 y 3.00?

b. ¿Qué tan grande se requeriría un tamaño muestral para asegurar que la primera probabilidad del inciso (a) sea por lo menos 0.99?

55. La primera tarea en un curso de cálculo estadístico requiere ejecutar un programa corto. Si la experiencia indica que 40% de todos los estudiantes principiantes no cometerán errores de programación, calcule la probabilidad (aproximada) de que en un grupo de 50 estudiantes

a. Por lo menos 25 no cometan errores. (*Sugerencia:* aproximación normal a la binomial.)

b. Entre 15 y 25 (incluso) no cometan errores.

56. El número de multas por estar mal estacionado emitidas en cierta ciudad en un determinado fin de semana tiene una distribución de Poisson con parámetro $\lambda = 50$. ¿Cuál es la probabilidad aproximada de que:

a. Se emitan entre 35 y 70 infracciones en un día en particular? (*Sugerencia:* cuando λ es grande, una *va* de Poisson tiene una distribución normal aproximada).

b. El número total de infracciones expedidas durante una semana de cinco días esté entre 225 y 275?

57. Suponga que la distribución del tiempo X (en horas) que los estudiantes de cierta universidad emplean en un determinado proyecto es gamma con parámetros $\alpha = 50$ y $\beta = 2$. Debido a que α es grande, se puede demostrar que X tiene más o menos una distribución normal. Utilice este hecho para calcular la probabilidad de que un estudiante seleccionado al azar invierta a lo sumo 125 horas en el proyecto.

5.5 Distribución de una combinación lineal

La media muestral $\overline{X}$ y la muestra total T_o son casos especiales de un tipo de variable aleatoria que surge con mucha frecuencia en aplicaciones estadísticas.

DEFINICIÓN

Dada una colección de n variables aleatorias $X_1, \ldots, X_n$ y n constantes numéricas $a_1, \ldots, a_n$, la *va*

$$Y = a_1X_1 + \cdots + a_nX_n = \sum_{i=1}^{n} a_iX_i \tag{5.7}$$

se llama **combinación lineal** de las X_i.

Con $a_1 = a_2 = \cdots = a_n = 1$ se obtiene $Y = X_1 + \cdots + X_n = T_o$, y $a_1 = a_2 = \cdots = a_n = \frac{1}{n}$ se obtiene $Y = \frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n = \frac{1}{n}(X_1 + \cdots + X_n) = \frac{1}{n}T_o = \bar{X}$. Observe que no se está requiriendo que las X_i sean independientes o que su distribución sea idéntica. Las X_i podrían tener distribuciones diferentes y, por lo tanto, medias y varianzas distintas. Primero se considera el valor esperado y la varianza de una combinación lineal.

PROPOSICIÓN

Sea $X_1, X_2, \ldots, X_n$ tienen valores medios $\mu_1, \ldots, \mu_n$, respectivamente, y varianzas $\sigma_1^2, \ldots, \sigma_n^2$, respectivamente.

1. Si las X_i son independientes o no,

$$\begin{aligned} E(a_1X_1 + a_2X_2 + \cdots + a_nX_n) &= a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n) \\ &= a_1\mu_1 + \cdots + a_n\mu_n \end{aligned} \tag{5.8}$$

2. Si $X_1, \ldots, X_n$ son independientes,

$$\begin{aligned} V(a_1X_1 + a_2X_2 + \cdots + a_nX_n) &= a_1^2V(X_1) + a_2^2V(X_2) + \cdots + a_n^2V(X_n) \\ &= a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2 \end{aligned} \tag{5.9}$$

y

$$\sigma_{a_1X_1 + \cdots + a_nX_n} = \sqrt{a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2} \tag{5.10}$$

3. Para cualquier $X_1, \ldots, X_n$,

$$V(a_1X_1 + \cdots + a_nX_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j\,\mathrm{Cov}(X_i, X_j) \tag{5.11}$$

Al final de la sección se da una idea general de las demostraciones. Una forma diferente de expresar (5.8) es que el valor esperado de una combinación lineal es la misma combinación lineal de los valores esperados; por ejemplo, $E(2X_1 + 5X_2) = 2\mu_1 + 5\mu_2$. El resultado (5.9) del enunciado 2 es un caso especial de (5.11) en el enunciado 3; cuando las X_i son independientes, $\mathrm{Cov}(X_i, X_j) = 0$ para $i \neq j$ e igual a $V(X_i)$ para $i = j$ (esta simplificación ocurre en realidad cuando las X_i no están correlacionadas, una condición más débil que la independencia). Si se particulariza el caso de una muestra aleatoria (las X_i son independientes y están distribuidas de manera idéntica) con $a_i = 1/n$ para toda i, se obtiene $E(\bar{X}) = \mu$ y $V(\bar{X}) = \sigma^2/n$, como se explicó en la sección 5.4. Un componente similar se aplica a las reglas para T_o.

Ejemplo 5.28 Una gasolinera vende tres tipos de gasolina: normal sin plomo, extra sin plomo y súper sin plomo. Éstas se venden a \$1.20, \$1.35 y \$1.50 por galón, respectivamente. Sean X_1, X_2 y X_3 las cantidades de estos tipos de gasolina vendidas (galones) en un día particular. Suponga que las X_i son indepen-

dientes con $\mu_1 = 1000$, $\mu_2 = 500$, $\mu_3 = 300$, $\sigma_1 = 100$, $\sigma_2 = 80$ y $\sigma_3 = 50$. El ingreso obtenido por estas ventas es $Y = 1.2X_1 + 1.35X_2 + 1.5X_3$, y

$$E(Y) = 1.2\mu_1 + 1.35\mu_2 + 1.5\mu_3 = \$2325$$
$$V(Y) = (1.2)^2\sigma_1^2 + (1.35)^2\sigma_2^2 + (1.5)^2\sigma_3^2 = 31\,689$$
$$\sigma_Y = \sqrt{31\,689} = \$178.01$$

■

Diferencia entre dos variables aleatorias

Un caso especial importante de una combinación lineal resulta de tomar $n = 2$, $a_1 = 1$ y $a_2 = -1$:

$$Y = a_1X_1 + a_2X_2 = X_1 - X_2$$

Entonces se tiene el siguiente corolario para la proposición.

COROLARIO

> $E(X_1 - X_2) = E(X_1) - E(X_2)$ y, si X_1 y X_2 son independientes, $V(X_1 - X_2) = V(X_1) + V(X_2)$.

El valor esperado de una diferencia es la diferencia de los dos valores esperados, pero la varianza de una diferencia entre dos variables independientes es la *suma, no* la diferencia, de las dos varianzas. Hay tanta variabilidad en $X_1 - X_2$ como en $X_1 + X_2$ [escribiendo $X_1 - X_2 = X_1 + (-1)X_2$, $(-1)X_2$ tiene la misma cantidad de variabilidad que X_2 por sí misma].

Ejemplo 5.29 Cierto fabricante de automóviles equipa un modelo particular con un motor de seis cilindros o un motor de cuatro cilindros. Sean X_1 y X_2 los rendimientos de combustible para automóviles de seis y cuatro cilindros, respectivamente, seleccionados de manera independiente y aleatoria. Con $\mu_1 = 22$, $\mu_2 = 26$, $\sigma_1 = 1.2$ y $\sigma_2 = 1.5$,

$$E(X_1 - X_2) = \mu_1 - \mu_2 = 22 - 26 = -4$$
$$V(X_1 - X_2) = \sigma_1^2 + \sigma_2^2 = (1.2)^2 + (1.5)^2 = 3.69$$
$$\sigma_{X_1 - X_2} = \sqrt{3.69} = 1.92$$

Si se marca de nuevo de modo que X_1 se refiera al automóvil de cuatro cilindros, entonces $E(X_1 - X_2) = 4$, pero la varianza de la diferencia aún es 3.69. ■

Caso de variables aleatorias normales

Cuando las X_i forman una muestra aleatoria a partir de una distribución normal, $\bar{X}$ y T_o tienen una distribución normal. A continuación se presenta un resultado más general en relación con las combinaciones lineales.

PROPOSICIÓN

> Si $X_1, X_2, \ldots, X_n$ son *va* independientes con distribución normal (posiblemente con medias o varianzas, o ambas, diferentes), entonces cualquier combinación lineal de las X_i también tiene una distribución normal. En particular, la diferencia $X_1 - X_2$ entre dos variables aleatorias distribuidas normalmente por sí misma tiene una distribución normal.

Ejemplo 5.30 (continuación del ejemplo 5.28) El ingreso total de la venta de los tres tipos de gasolina en un día particular fue $Y = 1.2X_1 + 1.35X_2 + 1.5X_3$, y se calculó $\mu_Y = 2325$ y (suponiendo independencia) $\sigma_Y = 178.01$. Si la distribución de las X_i es normal, la probabilidad de que el ingreso pase de 2500 es

$$P(Y > 2500) = P\left(Z > \frac{2500 - 2325}{178.01}\right)$$
$$= P(Z > 0.98) = 1 - \Phi(0.98) = 0.1635$$

El TLC también se puede generalizar de modo que se aplique a ciertas combinaciones lineales. En términos generales, si n es grande y ningún término individual tiene probabilidades de contribuir demasiado al valor global, entonces Y tiene aproximadamente una distribución normal.

Demostraciones para el caso $n = 2$

Para el resultado en relación con los valores esperados, suponga que X_1 y X_2 son continuas con fdp conjunta $f(x_1, x_2)$. Entonces,

$$E(a_1X_1 + a_2X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a_1x_1 + a_2x_2)f(x_1, x_2)\,dx_1\,dx_2$$
$$= a_1\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1 f(x_1, x_2)\,dx_2\,dx_1$$
$$+ a_2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_2 f(x_1, x_2)\,dx_1\,dx_2$$
$$= a_1\int_{-\infty}^{\infty}x_1 f_{X_1}(x_1)\,dx_1 + a_2\int_{-\infty}^{\infty}x_2 f_{X_2}(x_2)\,dx_2$$
$$= a_1E(X_1) + a_2E(X_2)$$

La suma sustituye a la integración en el caso discreto. El argumento para el resultado de la varianza no requiere especificar si la variable es discreta o continua. Recordando que $V(Y) = E[(Y - \mu_Y)^2]$,

$$V(a_1X_1 + a_2X_2) = E\{[a_1X_1 + a_2X_2 - (a_1\mu_1 + a_2\mu_2)]^2\}$$
$$= E\{a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + 2a_1a_2(X_1 - \mu_1)(X_2 - \mu_2)\}$$

La expresión dentro de las llaves es una combinación lineal de las variables $Y_1 = (X_1 - \mu_1)^2$, $Y_2 = (X_2 - \mu_2)^2$ y $Y_3 = (X_1 - \mu_1)(X_2 - \mu_2)$, de modo que al realizar la operación E para los tres términos se obtiene $a_1^2V(X_1) + a_2^2V(X_2) + 2a_1a_2\,\mathrm{Cov}(X_1, X_2)$ como se requería.

Ejercicios | Sección 5.5 (58 a 74)

58. Una compañía de envíos maneja recipientes en tres tamaños distintos: (1) 27 pies^3 ($3 \times 3 \times 3$), (2) 125 pies^3 y (3) 512 pies^3. Sea X_i ($i = 1, 2, 3$) el número de recipientes tipo i enviados durante una semana específica. Con $\mu_i = E(X_i)$ y $\sigma_i^2 = V(X_i)$, suponga que los valores medios y las desviaciones estándar son como sigue:

$\mu_1 = 200$	$\mu_2 = 250$	$\mu_3 = 100$
$\sigma_1 = 10$	$\sigma_2 = 12$	$\sigma_3 = 8$

a. Suponiendo que X_1, X_2, X_3 son independientes, calcule el valor esperado y la varianza del volumen total enviado. [*Sugerencia*: volumen $= 27X_1 + 125X_2 + 512X_3$.]

b. ¿Sus cálculos necesariamente serían correctos si las X_i no fueran independientes? Explique.

59. Sean X_1, X_2 y X_3 los tiempos necesarios para llevar a cabo tres tareas de reparación sucesivas en cierto taller. Suponga que los tiempos son variables aleatorias independientes, normales, con valores esperados μ_1, μ_2 y μ_3 y varianzas σ_1^2, σ_2^2 y σ_3^2, respectivamente.

a. Si $\mu = \mu_2 = \mu_3 = 60$ y $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 15$, calcule $P(X_1 + X_2 + X_3 \le 200)$. ¿Cuál es $P(150 \le X_1 + X_2 + X_3 \le 200)$?

b. Por medio de las μ_i, y las σ_i del inciso (a), calcule $P(55 \le \overline{X})$ y $P(58 \le \overline{X} \le 62)$.

c. Con las μ_i, y las σ_i del inciso (a), calcule $P(-10 \le X_1 - 0.5X_2 - 0.5X_3 \le 5)$.

d. Si $\mu_1 = 40$, $\mu_2 = 50$, $\mu_3 = 60$, $\sigma_1^2 = 10$, $\sigma_2^2 = 12$ y $\sigma_3^2 = 14$, calcule $P(X_1 + X_2 + X_3 \le 160)$ y $P(X_1 + X_2 \ge 2X_3)$.

60. Cinco automóviles del mismo tipo serán conducidos en un viaje de 300 millas. Los primeros dos utilizarán una marca económica de gasolina, y los otros tres automóviles, gasolina de marca registrada. Sean X_1, X_2, X_3, X_4 y X_5 los rendimientos de combustible observados (mpg) para los cinco automóviles. Suponga que estas variables son independientes y tienen una distribución normal con $\mu_1 = \mu_2 = 20$, $\mu_3 = \mu_4 = \mu_5 = 21$ y $\sigma^2 = 4$ para la marca económica y 3.5 para la de marca registrada. Defina una *va* Y mediante

$$Y = \frac{X_1 + X_2}{2} - \frac{X_3 + X_4 + X_5}{3}$$

de modo que Y es una medida de la diferencia de rendimiento entre la gasolina económica y la de marca registrada. Calcule $P(0 \le Y)$ y $P(-1 \le Y \le 1)$. [*Sugerencia:* $Y = a_1X_1 + \cdots + a_5X_5$, con $a_1 = \frac{1}{2}, \ldots, a_5 = -\frac{1}{3}$.]

61. En el ejercicio 26 se introdujeron las variables aleatorias X y Y, el número total de automóviles y autobuses, respectivamente, que lleva un trasbordador en un solo viaje. La fmp conjunta de X y Y se proporciona en el ejercicio 7. Se comprueba con facilidad que X y Y son independientes.

a. Calcule el valor esperado, varianza y desviación estándar del número total de vehículos en un solo viaje.

b. Si cada automóvil paga tres dólares y cada autobús 10 dólares, calcule el valor esperado, varianza y desviación estándar del total obtenido en un solo viaje.

62. La manufactura de cierto componente requiere tres operaciones de maquinado diferentes. El tiempo de maquinado para cada operación tiene una distribución normal y los tres tiempos son independientes entre sí. Los valores medios son 15, 30 y 20 minutos, respectivamente, y las desviaciones estándar son 1, 2 y 1.5 minutos. ¿Cuál es la probabilidad de que tome a lo sumo una hora de maquinado producir un componente seleccionado al azar?

63. Refiérase al ejercicio 3.

a. Calcule la covarianza entre X_1 = número de clientes en la caja rápida y X_2 = número de clientes en la caja superveloz.

b. Calcule $V(X_1 + X_2)$. ¿Cómo se compara esto con $V(X_1) + V(X_2)$?

64. Suponga que el tiempo que una persona espera un autobús en la mañana tiene una distribución uniforme en [0, 8], mientras que el tiempo de espera en la tarde tiene una distribución uniforme en [0, 10], independiente del tiempo de espera en la mañana.

a. Si la persona aborda el autobús cada mañana y tarde durante una semana, ¿cuál es el tiempo de espera total previsto? (*Sugerencia*: defina las *va* $X_1, \ldots, X_{10}$ y utilice una regla de valor esperado.)

b. ¿Cuál es la varianza del tiempo de espera total?

c. ¿Cuáles son el valor esperado y la varianza de la diferencia entre los tiempos de espera matutino y vespertino en un día específico?

d. ¿Cuáles son el valor esperado y la varianza de la diferencia entre el tiempo de espera total matutino y el tiempo de espera total vespertino para una semana en particular?

65. Suponga que cuando el pH de cierto compuesto químico es 5.00, el pH que mide un estudiante de química seleccionado al azar es una variable aleatoria con medida 5.00 y desviación estándar 0.2. Se subdivide un gran lote del compuesto y se da una muestra a cada estudiante del turno matutino y a cada estudiante del turno vespertino. Sea $\overline{X}$ = pH promedio que determinan los estudiantes de la mañana y $\overline{Y}$ = pH promedio que determinan los estudiantes de la tarde.

a. Si el pH es una variable normal y hay 25 estudiantes en cada turno, calcule $P(-0.1 \le \overline{X} - \overline{Y} \le 0.1)$. (*Sugerencia:* $\overline{X} - \overline{Y}$ es una combinación lineal de variables normales, de modo que tiene una distribución normal. Calcule $\mu_{\overline{X}-\overline{Y}}$ y $\sigma_{\overline{X}-\overline{Y}}$.)

b. Si hay 36 estudiantes en cada laboratorio, pero se supone que las determinaciones de pH no son normales, calcule (en forma aproximada) $P(-0.1 \le \overline{X} - \overline{Y} \le 0.1)$.

66. Si se aplican dos cargas a una viga en voladizo como se muestra en la figura siguiente, el momento de flexión en 0 debido a las cargas es $a_1X_1 + a_2X_2$.

- **a.** Suponga que X_1 y X_2 son variables aleatorias independientes, con medias 2 y 4 klb, respectivamente, y desviaciones estándar 0.5 y 1.0 klb, respectivamente. Si $a_1 = 5$ pies y $a_2 = 10$ pies, ¿cuál es el momento de flexión esperado y cuál es la desviación estándar del momento de flexión?
- **b.** Si X_1 y X_2 tienen una distribución normal, ¿cuál es la probabilidad de que el momento de flexión pase de 75 klb-pie?
- **c.** Suponga que las posiciones de las dos cargas son variables aleatorias. Si se representan por A_1 y A_2, suponga que estas variables tienen medias de 5 y 10 pies, respectivamente, que cada una tiene una desviación estándar de 0.5 y que las A_i y X_i son independientes entre sí. ¿Ahora cuál es el momento esperado?
- **d.** Para la situación del inciso (c), ¿cuál es la varianza del momento de flexión?
- **e.** Si la situación es como se describe en el inciso (a), excepto que $\text{Corr}(X_1, X_2) = 0.5$ (de modo que las dos cargas no son independientes), ¿cuál es la varianza del momento de flexión?

67. Una pieza de tubo de PVC se va a insertar dentro de otra pieza. La longitud de la primera tiene una distribución normal con valor medio de 20 pulgadas, y desviación estándar de 0.5 pulgadas. La longitud de la segunda pieza es una variable aleatoria normal con media y desviación estándar de 15 y 4 pulgadas, respectivamente. La cantidad de traslape tiene una distribución normal con valor medio de una pulgada y desviación estándar de 0.1 pulgada. Si se supone que las longitudes y cantidad de traslape son independientes entre sí, ¿cuál es la probabilidad de que la longitud total después de la inserción esté entre 34.5 y 35 pulgadas?

68. Dos aviones vuelan en la misma dirección en corredores paralelos adyacentes. En el tiempo $t = 0$, el primer avión está 10 km adelante del segundo. Suponga que la velocidad del primer avión (km/h) tiene una distribución normal con media de 520 y desviación estándar de 10, y la velocidad del segundo avión también tiene una distribución normal con media y desviación estándar de 500 y 10, respectivamente.

- **a.** ¿Cuál es la probabilidad de que después de dos horas de vuelo el segundo avión no haya alcanzado al primero?
- **b.** Determine la probabilidad de que los aviones estén separados a lo sumo 10 kilómetros después de dos horas.

69. Tres carreteras diferentes convergen en la entrada de una autopista. Suponga que durante un tiempo fijo, el número de automóviles que llegan de cada carretera a la autopista es una variable aleatoria con los valores esperados y desviaciones estándar de la tabla.

	Carretera 1	Carretera 2	Carretera 3
Valor esperado	800	1000	600
Desviación estándar	16	25	18

- **a.** ¿Cuál es el número total esperado de automóviles que entran a la autopista en este punto durante el periodo? (*Sugerencia*: sea X_i = número de la carretera i.)
- **b.** ¿Cuál es la varianza del número total de automóviles entrantes? ¿Hizo algunas suposiciones acerca de la relación entre el número de automóviles en las diferentes carreteras?
- **c.** Con X_i = número de automóviles que llegan de la carretera i durante el periodo, suponga que $\text{Cov}(X_1, X_2) = 80$, $\text{Cov}(X_1, X_3) = 90$ y $\text{Cov}(X_2, X_3) = 100$ (de modo que las tres corrientes de tráfico no son independientes). Calcule el número total esperado de automóviles que entran y la desviación estándar del total.

70. Suponga que se toma una muestra aleatoria de tamaño n de una distribución continua con mediana 0, de modo que la probabilidad de que cualquier observación sea positiva es 0.5. Ahora se hace caso omiso de los signos de las observaciones, se clasifican de menor a mayor en valor absoluto y después se hace W = suma de los rangos de las observaciones que tienen signos positivos. Por ejemplo, si las observaciones son −0.3, +0.7, +2.1 y −2.5, entonces los rangos de observaciones positivas son 2 y 3, así que $W = 5$. En el capítulo 15, a W se le llama *estadístico de rangos con signo de Wilcoxon*. W se puede representar como sigue:

$$W = 1 \cdot Y_1 + 2 \cdot Y_2 + 3 \cdot Y_3 + \cdots + n \cdot Y_n$$

$$= \sum_{i=1}^{n} i \cdot Y_i$$

donde las Y_i son variables aleatorias de Bernoulli independientes, cada una con $p = 0.5$ ($Y_i = 1$ corresponde a la observación con rango i positivo). Calcule lo siguiente:

a. $E(Y_i)$ y después $E(W)$ con la ecuación para W. [*Sugerencia:* los primeros n enteros positivos suman $n(n + 1)/2$.]

b. $V(Y_i)$ y después $V(W)$ [*Sugerencia:* la suma de los cuadrados de los primeros n enteros positivos es $n(n + 1)(2n + 1)/6$.]

71. En el ejercicio 66, el peso de la viga contribuye al momento de flexión. Suponga que la viga es de espesor y densidad uniformes, de modo que la carga resultante tiene una distribución uniforme en la viga. Si el peso de la viga es aleatorio, la carga resultante del peso también es aleatoria; denote esta carga por W (klb/pie).

a. Si la viga mide 12 pies de largo, W tiene media 1.5 y desviación estándar 0.25, y las cargas fijas son como se describe en el inciso (a) del ejercicio 66, ¿cuáles son el valor esperado y la varianza del momento de flexión? (*Sugerencia:* si la carga debida a la viga fuera w klb/pie, la contribución al momento de flexión sería $w \int_0^{12} x\,dx$.)

b. Si las tres variables (X_1, X_2 y W) están distribuidas de manera normal, ¿cuál es la probabilidad de que el momento de flexión sea a lo sumo 200 klb/pie?

72. Una persona tiene que realizar tres encargos en el edificio administrativo. Sea X_i = tiempo que tarda en realizar el i-ésimo encargo ($i = 1, 2, 3$) y sea X_4 = tiempo total en minutos que la persona utiliza en ir de un lado a otro del edificio y entre cada encargo. Suponga que las X_i son independientes, con distribución normal y las siguientes medias y desviaciones estándar: $\mu_1 = 15$, $\sigma_1 = 4$, $\mu_2 = 5$, $\sigma_2 = 1$, $\mu_3 = 8$, $\sigma_3 = 2$, $\mu_4 = 12$, $\sigma_4 = 3$. La persona piensa salir de su oficina justo a las 10:00 A.M. y desea pegar una nota en la puerta de su oficina con el siguiente recado: "regreso a las t A.M." ¿Cuál es el valor de t que debe escribir la persona si quiere que la probabilidad de su regreso después de t sea 0.01?

73. Suponga que la resistencia esperada a la tensión del acero tipo A es 105 klb y que la desviación estándar de la resistencia a la tensión es 8 klb. Para el acero tipo B, suponga que la resistencia esperada a la tensión y la desviación estándar de la resistencia a la tensión son 100 klb y 6 klb, respectivamente. Sea $\overline{X}$ = resistencia promedio a la tensión de una muestra aleatoria de 40 especímenes de tipo A y sea $\overline{Y}$ = resistencia promedio a la tensión de una muestra aleatoria de 35 especímenes tipo B.

a. ¿Cuál es la distribución aproximada de $\overline{X}$? ¿De $\overline{Y}$?

b. ¿Cuál es la distribución aproximada de $\overline{X} - \overline{Y}$? Justifique su respuesta.

c. Calcule (aproximadamente) $P(-1 \le \overline{X} - \overline{Y} \le 1)$.

d. Calcule $P(\overline{X} - \overline{Y} \ge 10)$. Si en realidad observó $\overline{X} - \overline{Y} \ge 10$, ¿dudaría que $\mu_1 - \mu_2 = 5$?

74. En un área que tiene suelo arenoso se plantaron 50 arbolitos de cierto tipo, y otros 50 en un área con terreno arcilloso. Sea X = número de árboles plantados en terreno arenoso que sobreviven un año y Y = número de árboles plantados en terreno arcilloso que sobreviven un año. Si la probabilidad de que un árbol plantado en terreno arenoso sobreviva un año es 0.7, y la probabilidad de que sobreviva un año en terreno arcilloso es 0.6, calcule una aproximación para $P(-5 \le X - Y \le 5)$ (haga caso omiso de la corrección de continuidad).

Ejercicios complementarios (75 a 94)

75. Un restaurante sirve tres cenas de precio fijo que cuestan 12, 15 y 20 dólares. Para una pareja seleccionada al azar que cena en este restaurante, sea X = costo de la cena del hombre y Y = costo de la cena de la mujer. La fmp conjunta de X y Y se proporciona en la tabla siguiente:

$p(x, y)$		y = 12	15	20
	12	0.05	0.05	0.10
x	15	0.05	0.10	0.35
	20	0	0.20	0.10

a. Calcule las fmp marginales de X y Y.

b. ¿Cuál es la probabilidad de que la cena del hombre y la mujer cuesten a lo sumo 15 dólares cada una?

c. ¿Son independientes X y Y? Justifique su respuesta.

d. ¿Cuál es el costo total esperado de la cena para las dos personas?

e. Suponga que cuando una pareja abre galletas de la suerte al terminar de comer, encuentra el mensaje "Recibirá como devolución la diferencia entre el costo de la comida más cara y la menos cara que haya escogido". ¿Cuánto espera devolver el restaurante?

76. En la estimación de costos, el costo total de un proyecto es la suma de los costos de las tareas constitutivas. Cada uno de estos costos es una variable aleatoria con una distribución de probabilidad. Se acostumbra obtener información acerca de la distribución del costo total al sumar características de las distribuciones de cada uno de los costos. Esto se llama procedimiento de "acumulación". Por ejemplo, $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$, así que el procedimiento de acumulación es válido para el costo promedio. Suponga que hay dos tareas constitutivas y que X_1 y X_2 son variables aleatorias independientes con distribución normal. ¿Es válido el procedimiento de acumulación para el percentil 75? Es decir, ¿el percentil 75 de la distribución de $X_1 + X_2$ es el mismo que la suma de los percentiles 75 de cada una las dos distribuciones? Si no, ¿cuál es la relación entre el percentil de la suma y la suma de percentiles? ¿Para cuáles percentiles es válido el procedimiento de acumulación en este caso?

77. Una tienda de alimentos naturales tiene en existencia dos marcas de cierto tipo de grano. Sea X = cantidad (en libras) de la marca A y Y = cantidad de la marca B. Suponga que la fdp conjunta de X y Y es

$$f(x, y) = \begin{cases} kxy & x \geq 0,\ y \geq 0,\ 20 \leq x + y \leq 30 \\ 0 & \text{en caso contrario} \end{cases}$$

- **a.** Trace la región de densidad positiva y determine el valor de k.
- **b.** ¿Son independientes X y Y? Para contestar primero obtenga la fdp marginal de cada variable.
- **c.** Calcule $P(X + Y \leq 25)$.
- **d.** ¿Cuál es la cantidad total esperada de este grano en existencia?
- **e.** Calcule $\text{Cov}(X, Y)$ y $\text{Corr}(X, Y)$.
- **f.** ¿Cuál es la varianza de la cantidad total de grano en existencia?

78. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias que denotan n ofertas independientes para un artículo que está a la venta. Suponga que cada X_i tiene una distribución uniforme en el intervalo [100, 200]. Si el vendedor vende al mejor postor, ¿cuánto se espera que gane en la venta? [*Sugerencia:* sea $Y = \text{máx}(X_1, X_2, \ldots, X_n)$. Primero determine $F_Y(y)$ observando que $Y \leq y$ si y sólo si cada X_i es $\leq y$. Luego obtenga la fdp y $E(Y)$.]

79. Suponga que para cierto individuo, la ingestión de calorías en el desayuno es una variable aleatoria con valor esperado 500 y desviación estándar 50, la del almuerzo también es aleatoria con valor esperado de 900 y desviación estándar de 100, y la de la comida es una variable aleatoria con valor esperado de 2000 y desviación estándar de 180. Si se supone que las ingestiones en las diferentes comidas son independientes entre sí, ¿cuál es la posibilidad de que la ingestión promedio de calorías por día en el siguiente año (365 días) sea a lo sumo 3500? [*Sugerencia:* represente con X_i, Y_i, y Z_i las tres ingestiones de calorías en el día i. Entonces, $\sum(X_i + Y_i + Z_i)$ proporciona la ingestión total.]

80. El peso medio de equipaje que deposita un pasajero de clase turista seleccionado al azar que vuela entre dos ciudades en cierta aerolínea es de 40 libras, y la desviación estándar es 10 lb. La media y desviación estándar para un pasajero de clase ejecutiva son 30 lb y 6 lb, respectivamente.

- **a.** Si hay 12 pasajeros de clase ejecutiva y 50 pasajeros de clase turista en un vuelo, ¿cuáles son el valor esperado y la desviación estándar del peso total del equipaje?
- **b.** Si cada uno de los pesos de equipaje son *no* independientes con distribución normal, ¿cuál es la probabilidad de que el peso total del equipaje sea a lo sumo 2500 lb?

81. Se ha visto que si $E(X_1) = E(X_2) = \cdots = E(X_n) = \mu$, entonces $E(X_1 + \cdots + X_n) = n\mu$. En algunas aplicaciones el número de X_i en consideración no es un número fijo n, sino una variable aleatoria N. Por ejemplo, sea N = número de componentes llevados a un taller de reparación en un día en particular, y sea X_i el tiempo del taller de reparación para el i-ésimo componente. Entonces, el tiempo total de reparación es $X_1 + X_2 + \cdots + X_N$, la suma de un número *aleatorio* de variables aleatorias. Cuando N es independiente de las X_i, se puede demostrar que

$$E(X_1 + \cdots + X_N) = E(N) \cdot \mu$$

- **a.** Si el número esperado de componentes llevados en un día en particular es 10 y el tiempo esperado de reparación para un componente elegido al azar es 40 minutos, ¿cuál es el tiempo esperado de reparación para componentes enviados en un determinado día?
- **b.** Suponga que los componentes de cierto tipo llegan para reparación según un proceso de Poisson con una tasa de 5 por hora. El número esperado de defectos por componente es 3.5 ¿Cuál es el valor esperado del número total de defectos en componentes enviados para reparación durante un periodo de cuatro horas? Asegúrese de indicar cómo deduce su respuesta del resultado general recién proporcionado.

82. Suponga que la proporción de votantes rurales de cierto estado quienes favorecen a cierto candidato a gobernador es 0.45, y que la proporción de votantes urbanos y suburbanos a favor del candidato es 0.60. Si se obtiene una muestra de 200 votantes rurales y 300 votantes urbanos y suburbanos, ¿cuál es la probabilidad aproximada de que por lo menos 250 de estos votantes favorezcan a este candidato?

83. Sea μ el pH real de un compuesto químico. Se llevará a cabo una secuencia de n determinaciones de pH muestrales independientes. Suponga que cada muestra de pH es una variable aleatoria con valor esperado μ y desviación estándar 0.1. ¿Cuántas determinaciones se requieren si se desea que la probabilidad de que el promedio muestral esté dentro de 0.02 del pH real sea por lo menos 0.95? ¿Qué teorema justifica este cálculo de probabilidad?

84. Si la cantidad de bebida gaseosa que consume cierta persona en determinado día es independiente del consumo en cualquier otro día y tiene una distribución normal con $\mu = 13$ onzas y $\sigma = 2$; además, si la persona cuenta con dos paquetes de seis botellas cada uno (16 onzas por envase), ¿cuál es la probabilidad de que todavía tenga algunas botellas al final de dos semanas (14 días)?

85. Refiérase al ejercicio 58 y suponga que las X_i son independientes entre sí, cada una con distribución normal. ¿Cuál es la posibilidad de que el volumen total enviado sea a lo sumo 100 000 pies^3?

86. Un estudiante tiene una clase que se supone termina a las 9:00 A.M. y otra que comienza a las 9:10 A.M. Suponga que el tiempo real de terminación de la clase de las 9:00 es una *va* X_1 que tiene distribución normal con media 9:02 y desviación estándar 1.5 minutos, y que la hora de inicio de la siguiente clase es también una *va* X_2 distribuida de manera normal con media 9:10 y desviación estándar de un minuto. Suponga también que el tiempo necesario para salir de un salón y entrar a otro es una *va* X_3 distribuida normalmente con media de seis minutos, y desviación estándar de un minuto. ¿Cuál es la probabilidad de que el estudiante llegue a la segunda clase antes que ésta comience? (Suponga la independencia de X_1, X_2 y X_3, lo cual es razonable si el estudiante pasa por alto la hora en que termina la primera clase.)

87. **a.** Utilice la fórmula general para la varianza de una combinación lineal a fin de escribir una expresión para $V(aX + Y)$. Entonces, sea $a = \sigma_Y/\sigma_X$ y demuestre que $\rho \geq -1$. [*Sugerencia:* la varianza siempre es ≥ 0 y $\text{Cov}(X, Y) = \sigma_X \cdot \sigma_Y \cdot \rho$.]

b. Considere $V(aX - Y)$ y concluya que $\rho \leq 1$.

c. Utilice el hecho de que $V(W) = 0$ sólo si W es una constante para demostrar que $\rho = 1$ sólo si $Y = aX + b$.

88. Suponga una puntuación verbal X y una puntuación cuantitativa Y de un individuo seleccionado al azar en un examen de aptitud aplicado a nivel nacional que tiene fdp conjunta

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Dé una predicción t de la puntuación total $X + Y$ del individuo. El error de predicción es el error cuadrático medio $E[(X + Y - t)^2]$. ¿Cuál valor de t reduce al mínimo el error de predicción?

89. **a.** Sea que X_1 tiene una distribución chi-cuadrado con parámetro ν_1 (véase la sección 4.4), y sea que X_2 es independiente de X_1 con una distribución chi-cuadrado con parámetro ν_2. Utilice la técnica del ejemplo 5.21 para demostrar que $X_1 + X_2$ tiene una distribución chi-cuadrado con parámetro $\nu_1 + \nu_2$.

b. En el ejercicio 65 del capítulo 4 se pidió demostrar que si Z es una *va* normal estándar, entonces Z^2 tiene una distribución chi-cuadrado con $\nu = 1$. Sea que $Z_1, Z_2, \ldots, Z_n$ son n variables aleatorias normales estándar independientes. ¿Cuál es la distribución de $Z_1^2 + \cdots + Z_n^2$? Justifique su respuesta.

c. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal con media μ y varianza σ^2. ¿Cuál es la distribución de la suma $Y = \sum_{i=1}^{n} [(X_i - \mu)/\sigma]^2$? Justifique su respuesta.

90. **a.** Demuestre que $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$.

b. Sean X_1 y X_2 puntuaciones cuantitativas y verbales en un examen de aptitud, y sean Y_1 y Y_2 las puntuaciones correspondientes a otro examen. Si $\text{Cov}(X_1, Y_1) = 5$, $\text{Cov}(X_1, Y_2) = 1$, $\text{Cov}(X_2, Y_1) = 2$ y $\text{Cov}(X_2, Y_2) = 8$, ¿cuál es la covarianza entre las dos puntuaciones totales $X_1 + X_2$ y $Y_1 + Y_2$?

91. Un espécimen de roca de cierta región se selecciona al azar y se pesa en dos ocasiones diferentes. Represente con W el peso real y con X_1 y X_2 los dos pesos medidos. Entonces, $X_1 = W + E_1$ y $X_2 = W + E_2$, donde E_1 y E_2 son los dos errores de medición. Suponga que las E_i son independientes una de otra y de W y que $V(E_1) = V(E_2) = \sigma_E^2$.

a. Exprese ρ, el coeficiente de correlación entre los dos pesos medidos X_1 y X_2, en términos de σ_W^2, la varianza del peso real, σ_X^2, la varianza del peso medido.

b. Calcule ρ cuando $\sigma_W = 1$ kg y $\sigma_E = 0.01$ kg.

92. Sea A el porcentaje de un elemento en un espécimen de roca seleccionado al azar y sea B el porcentaje de un segundo elemento en ese mismo espécimen. Suponga que D y E son errores de medición al determinar los valores de A y B de modo que los valores medidos son $X = A + D$ y $Y = B + E$, respectivamente. Suponga que los errores de medición son independientes entre sí y de valores reales.

a. Demuestre que

$$\text{Corr}(X, Y) = \text{Corr}(A, B) \cdot \sqrt{\text{Corr}(X_1, X_2)} \cdot \sqrt{\text{Corr}(Y_1, Y_2)}$$

donde X_1 y X_2 son mediciones duplicadas del valor de A, y Y_1 y Y_2 están definidas de manera análoga con respecto a B. ¿Qué efecto tiene la presencia de un error de medición en la correlación?

b. ¿Cuál es el valor máximo de $\text{Corr}(X, Y)$ cuando $\text{Corr}(X_1, X_2) = 0.8100$ y $\text{Corr}(Y_1, Y_2) = 0.9025$? ¿Esto es desconcertante?

93. Sean $X_1, \ldots, X_n$ independientes con valores medios $\mu_1, \ldots, \mu_n$ y varianzas $\sigma_1^2, \ldots, \sigma_n^2$. Considere una función $h(x_1, \ldots, x_n)$, y utilícela para definir una nueva variable aleatoria $Y = h(X_1, \ldots, X_n)$. En condiciones bastante generales en la función h, si las σ_i son pequeñas en relación con las μ_i, correspondientes, se puede demostrar que $E(Y) \approx h(\mu_1, \ldots, \mu_n)$ y

$$V(Y) \approx \left(\frac{\partial h}{\partial x_1}\right)^2 \cdot \sigma_1^2 + \cdots + \left(\frac{\partial h}{\partial x_n}\right)^2 \cdot \sigma_n^2$$

donde cada derivada parcial se evalúa en $(x_1, \ldots, x_n) = (\mu_1, \ldots, \mu_n)$. Suponga que tres resistores con resistencias X_1, X_2, X_3 están conectados en paralelo a una batería con voltaje X_4. Entonces por la ley de Ohm, la corriente es

$$Y = X_4\left[\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3}\right]$$

Sea $\mu_1 = 10$ ohms, $\sigma_1 = 1.0$ ohm, $\mu_2 = 15$ ohms, $\sigma_2 = 1.0$ ohm, $\mu_3 = 20$ ohms, $\sigma_3 = 1.5$ ohms, $\mu_4 = 120$ V, $\sigma_4 = 4.0$ V. Calcule el valor esperado (aproximado) y la desviación estándar de la corriente (sugerido por "Random Samplings", *CHEMTECH*, 1984: 696-697).

94. Una aproximación más precisa a $E[h(X_1, \ldots, X_n)]$ del ejercicio 93 es

$$h(\mu_1, \ldots, \mu_n) + \frac{1}{2}\sigma_1^2\left(\frac{\partial^2 h}{\partial x_1^2}\right) + \cdots + \frac{1}{2}\sigma_n^2\left(\frac{\partial^2 h}{\partial x_n^2}\right)$$

Calcule esto para $Y = h(X_1, X_2, X_3, X_4)$ dada en el ejercicio 93 y compárela con el término inicial $h(\mu_1, \ldots, \mu_n)$.

Bibliografía

Larsen, Richard y Morris Marx, *An Introduction to Mathematical Statistics and Its Applications* (3a. ed.). Prentice Hall, Englewood Cliffs, NJ, 2000. Un estudio más limitado que en el libro de Olkin *et al.*, pero bien escrito y ameno.

Olkin, Ingram, Cyrus Derman y Leon Gleser, *Probability Models and Applications* (2a. ed.), Macmillan, Nueva York, 1994. Contiene una exposición cuidadosa y completa de distribuciones conjuntas, reglas de valores esperados y teoremas límite.

6

Estimación puntual

Introducción

Dado un parámetro de interés, como una media poblacional μ, o una proporción p de la población, el objetivo de la estimación puntual es usar una muestra para calcular un número que representa en cierto sentido una buena estimación del valor real del parámetro. El número resultante se llama *estimación puntual*. En la sección 6.1 se presentan algunos conceptos generales de estimación puntual. En la sección 6.2, se describen e ilustran dos métodos importantes para obtener estimaciones puntuales: el método de los momentos y el método de máxima verosimilitud.

6.1 Algunos conceptos generales de la estimación puntual

La inferencia estadística casi siempre se dirige hacia obtener algún tipo de conclusión acerca de uno o más parámetros (características de la población). Para esto se requiere que un investigador obtenga datos muestrales de cada una de las poblaciones estudiadas. Las conclusiones se pueden basar entonces en los valores calculados de varias cantidades de la muestra. Por ejemplo, sea μ (un parámetro) la resistencia a la ruptura promedio verdadera de conexiones de cables utilizadas en obleas semiconductoras de unión. Se podría tomar una muestra aleatoria de $n = 10$ conexiones y determinar la resistencia a la ruptura de cada una, que producen resistencias observadas $x_1, x_2, \ldots, x_{10}$. La resistencia a la ruptura promedio de la muestra $\bar{x}$ se podría usar entonces para sacar conclusiones acerca del valor de μ. De manera similar, si σ^2 es la varianza de la distribución de resistencia a la ruptura (varianza poblacional, otro parámetro), el valor de la varianza muestral s^2 se puede usar para inferir algo acerca de σ^2.

Al analizar los conceptos generales y métodos de inferencia, es conveniente tener un símbolo genérico para el parámetro de interés. Para este propósito se usa la letra griega θ. El objetivo de la estimación puntual es seleccionar un solo número, con base en los datos de la muestra, que represente un valor sensible para θ. Suponga, por ejemplo, que el parámetro de interés es μ, el tiempo de vida útil promedio real de baterías de cierto tipo. Una muestra aleatoria de $n = 3$ baterías podría producir los tiempos de vida útil observados (horas) $x_1 = 5.0$, $x_2 = 6.4$, $x_3 = 5.9$. El valor calculado del tiempo de vida útil promedio de la muestra es $\bar{x} = 5.77$, y es razonable considerar a 5.77 como un valor muy posible de μ, la "mejor estimación" para el valor de μ con base en la información muestral disponible.

Suponga que se desea estimar un parámetro de una sola población (p. ej., μ o σ) con base en una muestra aleatoria de tamaño n. Recuerde del capítulo anterior que antes que estén disponibles los datos, las observaciones muestrales se deben considerar variables aleatorias (*va*) $X_1, X_2, \ldots, X_n$. Se deduce que cualquier función de las X_i es decir, cualquier estadístico, como la media muestral $\bar{X}$ o la desviación estándar S, también es una variable aleatoria. Lo mismo se cumple si los datos disponibles consisten en más de una muestra. Por ejemplo, las resistencias a la tensión de m especímenes de tipo 1 y n especímenes de tipo 2 se pueden representar mediante $X_1, \ldots, X_m$ y $Y_1, \ldots, Y_n$, respectivamente. La diferencia entre las dos resistencias medias muestrales es $\bar{X} - \bar{Y}$, el estadístico natural para sacar conclusiones acerca de $\mu_1 - \mu_2$, la diferencia entre las resistencias medias poblacionales.

DEFINICIÓN

> Una **estimación puntual** de un parámetro θ es un solo número que se puede considerar como valor sensible para θ. Una estimación puntual se obtiene al seleccionar un estadístico adecuado y calcular su valor a partir de los datos muestrales proporcionados. El estadístico seleccionado se llama **estimador puntual** de θ.

En el ejemplo de la batería recién mencionado, el estimador utilizado para obtener la estimación puntual de μ fue $\bar{X}$, y la estimación puntual de μ fue 5.77. Si en cambio los tres tiempos de vida útil observados hubieran sido $x_1 = 5.6$, $x_2 = 4.5$ y $x_3 = 6.1$, el uso del estimador $\bar{X}$ habría dado como resultado la estimación $\bar{x} = (5.6 + 4.5 + 6.1)/3 = 5.40$. El símbolo $\hat{\theta}$ ("teta con sombrero") por lo común se usa para denotar tanto el estimador de θ como la estimación puntual resultante de una muestra en particular.* Así, $\hat{\mu} = \bar{X}$ se lee como "el estimador

*Siguiendo la notación anterior, se podría usar $\hat{\Theta}$ (teta mayúscula) para el estimador, pero resulta difícil escribirla

puntual de μ es la media muestral $\bar{X}$". El enunciado "el estimado puntual de μ es 5.77" se puede escribir de manera concisa como $\hat{\mu} = 5.77$. Observe que al escribir $\hat{\theta} = 72.5$, no hay indicación de cómo se obtuvo esta estimación puntual (qué estadístico se usó). Se recomienda informar tanto el estimador como la estimación resultante.

Ejemplo 6.1 Un fabricante de automóviles ha desarrollado un nuevo tipo de defensa, que se supone absorbe los impactos con menos daño que las defensas anteriores. El fabricante ha utilizado esta defensa en una secuencia de 25 colisiones controladas contra una pared, cada una a 10 mph, usando uno de sus modelos de automóvil compacto. Sea X = número de colisiones que no producen daño visible al automóvil. El parámetro por estimar es p = proporción de choques que no producen daño [de manera alternativa, $p = P$(ningún daño en una sola colisión)]. Si se observa que X es $x = 15$, el estimador más razonable y la estimación son

$$\text{estimador } \hat{p} = \frac{X}{n} \qquad \text{estimación} = \frac{x}{n} = \frac{15}{25} = 0.60 \qquad \blacksquare$$

Si para cada parámetro de interés sólo hubiera un estimador puntual razonable, no habría mucho para la estimación puntual. Sin embargo, en muchos problemas, habrá más de un estimador razonable.

Ejemplo 6.2 Considere de nuevo las 20 observaciones siguientes acerca de la tensión disruptiva dieléctrica para las piezas de resina epóxica introducidas primero en el ejemplo 4.29 (sección 4.6).

24.46	25.61	26.25	26.42	26.66	27.15	27.31	27.54	27.74	27.94
27.98	28.04	28.28	28.49	28.50	28.87	29.11	29.13	29.50	30.88

El patrón en la gráfica de probabilidad normal proporcionada allí es bastante directo, así que ahora se supone que la distribución de la tensión disruptiva es normal con media μ. Debido a que las distribuciones normales son simétricas, μ también es el tiempo de vida mediano de la distribución. Entonces se supone que las observaciones proporcionadas son el resultado de una muestra aleatoria $X_1, X_2, \ldots, X_{20}$ de esta distribución normal. Considere los siguientes estimadores y las estimaciones resultantes para μ:

a. Estimador = $\bar{X}$, estimación = $\bar{x} = \sum x_i/n = 555.86/20 = 27.793$

b. Estimador = $\tilde{X}$, estimación = $\tilde{x} = (27.94 + 27.98)/2 = 27.960$

c. Estimador = $[\text{mín}(X_i) + \text{máx}(X_i)]/2$ = promedio de los dos tiempos de vida extremos, estimación = $[\text{mín}(x_i) + \text{máx}(x_i)]/2 = (24.46 + 30.88)/2 = 27.670$

d. Estimador = $\bar{X}_{\text{rec}(10)}$, la media recortada al 10% (se descarta el 10% más pequeño y más grande de la muestra y luego se promedia),

$$\begin{aligned}\text{estimación} &= \bar{x}_{\text{rec}(10)} \\ &= \frac{555.86 - 24.46 - 25.61 - 29.50 - 30.88}{16} \\ &= 27.838\end{aligned}$$

Cada uno de los estimadores (a) a (d) emplea una medida diferente del centro de la muestra para estimar μ. ¿Cuál de las estimaciones es la más cercana al valor verdadero? Esto no se puede contestar sin conocer el valor verdadero. Una pregunta que se puede contestar es, "¿Cuál estimador, cuando se emplea en otras muestras de X_i, tenderá a producir estimaciones más cercanas al valor verdadero?". En breve se considerará este tipo de pregunta. ∎

Ejemplo 6.3 En el futuro cercano habrá interés creciente en desarrollar aleaciones de bajo costo basadas en Mg para varios procesos de colado. Por lo tanto, es importante contar con formas prácticas de determinar varias propiedades mecánicas de esta clase de aleaciones. En el artículo "On the Development of a New Approach for the Determination of Yield Strength in Mg-based Alloys" (*Light Metal Age*, Oct. 1998: 50-53) se propone un método ultrasónico para este propósito. Considere la muestra siguiente de observaciones acerca del módulo elástico (GPa) de especímenes de aleación de un proceso de vaciado en molde:

44.2 43.9 44.7 44.2 44.0 43.8 44.6 43.1

Suponga que estas observaciones son el resultado de una muestra aleatoria $X_1, \ldots, X_8$ de la distribución poblacional del módulo elástico en tales circunstancias. Se quiere estimar la varianza poblacional σ^2. Un estimador natural es la varianza muestral:

$$\hat{\sigma}^2 = S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{\sum X_i^2 - (\sum X_i)^2/n}{n-1}$$

La estimación correspondiente es

$$\hat{\sigma}^2 = s^2 = \frac{\sum x_i^2 - (\sum x_i)^2/8}{7} = \frac{15\,533.79 - (352.5)^2/8}{7}$$

$$= 0.25125 \approx 0.251$$

La estimación de σ sería entonces $\hat{\sigma} = s = \sqrt{0.25125} = 0.501$.

Un estimador alterno resultaría de usar el divisor n en vez de $n - 1$ (es decir, la desviación cuadrada promedio):

$$\hat{\sigma}^2 = \frac{\sum(X_i - \bar{X})^2}{n} \qquad \text{estimación} = \frac{1.75875}{8} = 0.220$$

En breve se indica por qué muchos estadísticos prefieren S^2 al estimador con divisor n. ∎

En el mejor de los mundos posibles, se podría encontrar un estimador $\tilde{\theta}$ para el cual siempre se tiene que $\hat{\theta} = \theta$. Sin embargo, $\hat{\theta}$ es una función de las X_i, muestrales, por lo tanto es una variable aleatoria. Para algunas muestras, $\hat{\theta}$ produce un valor más grande que θ, en tanto que para otras muestras $\hat{\theta}$ da una subestimación de θ. Si se escribe

$$\hat{\theta} = \theta + \text{error de estimación}$$

entonces un estimador preciso sería uno que dé como resultado errores de estimación pequeños, de modo que los valores estimados estarán cerca del valor real. Un estimador que tiene las propiedades de insesgamiento y varianza mínima, suele ser preciso en este sentido.

Estimadores insesgados

Suponga que se tienen dos instrumentos de medición, uno ha sido calibrado con precisión, pero el otro da en forma sistemática lecturas más pequeñas que el valor real que está siendo medido. Cuando cada instrumento se usa una y otra vez en el mismo objeto, como resultado del error de medición, las mediciones observadas no serán idénticas. Sin embargo, el primer instrumento produce medidas que tienen una distribución normal con respecto al valor real de tal manera que, en promedio, este instrumento mide lo que pretende medir, de modo que se denomina un instrumento insesgado. El segundo instrumento produce observaciones que tienen un componente de error sistemático o sesgo.

DEFINICIÓN

Se dice que un estimador puntual $\hat{\theta}$ es un **estimador insesgado** de θ si $E(\hat{\theta}) = \theta$ para todo valor posible de θ. Si $\hat{\theta}$ no es insesgado, la diferencia $E(\hat{\theta}) - \theta$ se llama **sesgo** de $\hat{\theta}$.

Es decir, $\hat{\theta}$ es insesgado si su distribución de probabilidad (es decir, muestreo) siempre está "centrada" en el valor verdadero del parámetro. Suponga que $\hat{\theta}$ es un estimador insesgado; entonces si $\theta = 100$, la distribución de muestreo de $\hat{\theta}$ está centrada en 100; si $\theta = 27.5$, entonces la distribución de muestreo está centrada en 27.5, etcétera. En la figura 6.1 se ilustran las distribuciones de varios estimadores sesgados e insesgados. Observe que "centrado" aquí significa que el valor esperado, no la mediana, de la distribución de $\hat{\theta}$ es igual a θ.

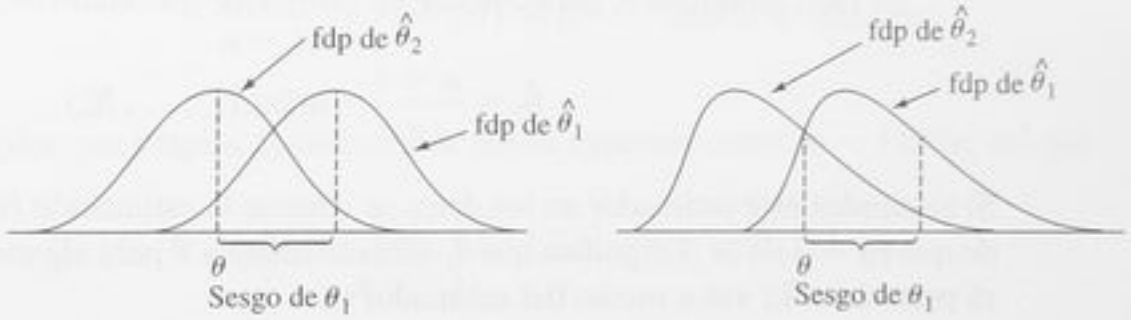


Figura 6.1 Fdp de un estimador sesgado $\hat{\theta}_1$ y un estimador insesgado $\hat{\theta}_2$ para un parámetro θ

Parece como si fuera necesario saber el valor de θ (en cuyo caso es innecesaria la estimación) para ver si $\hat{\theta}$ es insesgada. Esto por lo común no es caso, ya que con frecuencia se puede usar un argumento general de valor esperado para comprobar el insesgamiento.

En el ejemplo 6.1, la proporción muestral X/n se usó como estimador de p, donde X, el número de éxitos en la muestra, tuvo una distribución binomial con parámetros n y p. Así,

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}(np) = p$$

PROPOSICIÓN

Cuando X es una *va* binomial con parámetros n y p, la proporción muestral $\hat{p} = X/n$ es un estimador insesgado de p.

Sin importar cuál sea el valor verdadero de p, la distribución del estimador $\hat{p}$ se centrará en el valor verdadero.

Ejemplo 6.4 Suponga que X, el tiempo de reacción a cierto estímulo, tiene una distribución uniforme en el intervalo de 0 a un límite superior desconocido θ (por lo tanto, la función de densidad de X es de forma rectangular con altura $1/\theta$ para $0 \le x \le \theta$). Se desea estimar θ con base en una muestra aleatoria $X_1, X_2, \ldots, X_n$ de tiempos de reacción. Puesto que θ es el tiempo posible más grande en la población completa de tiempos de reacción, considere como primer estimador el tiempo de reacción más grande de la muestra: $\hat{\theta}_1 = \text{máx}(X_1, \ldots, X_n)$. Si $n = 5$ y $x_1 = 4.2$, $x_2 = 1.7$, $x_3 = 2.4$, $x_4 = 3.9$, $x_5 = 1.3$, la estimación puntual de θ es $\hat{\theta}_1 = \text{máx}(4.2, 1.7, 2.4, 3.9, 1.3) = 4.2$.

El insesgamiento significa que algunas muestras producirán estimaciones que pasan de θ y otras muestras producirán estimaciones más pequeñas que θ, de otro modo θ no podría ser el centro (punto de equilibrio) de la distribución de $\hat{\theta}_1$. El estimador propuesto nunca sobreestimará a θ (el valor muestral más grande no puede pasar del valor más grande de la población) y subestimará a θ a menos que el valor más grande de la muestra sea igual a θ. Este argumento intuitivo muestra que $\hat{\theta}_1$ es un estimador sesgado. Para ser precisos, se puede demostrar (véase el ejercicio 32) que

$$E(\hat{\theta}_1) = \frac{n}{n+1} \cdot \theta < \theta \qquad \left(\text{puesto que } \frac{n}{n+1} < 1\right)$$

El sesgo de $\hat{\theta}_1$ se determina mediante $n\theta/(n+1) - \theta = -\theta/(n+1)$, que tiende a 0 cuando n aumenta.

Es fácil modificar $\hat{\theta}_1$ para obtener un estimador insesgado de θ. Considere el estimador

$$\hat{\theta}_2 = \frac{n+1}{n} \cdot \text{máx}(X_1, \ldots, X_n)$$

Si se emplea este estimador en los datos se obtiene la estimación $(6/5)(4.2) = 5.04$. El hecho de que $(n+1)/n > 1$ significa que $\hat{\theta}_2$ sobreestimara a θ para algunas muestras y la subestimará para otras. El valor medio del estimador es

$$E(\hat{\theta}_2) = E\left[\frac{n+1}{n}\,\text{máx}(X_1, \ldots, X_n)\right] = \frac{n+1}{n} \cdot E[\text{máx}(X_1, \ldots, X_n)]$$

$$= \frac{n+1}{n} \cdot \frac{n}{n+1}\,\theta = \theta$$

Si se emplea $\hat{\theta}_2$ de manera repetida en diferentes muestras para estimar θ, algunas estimaciones serán demasiado grandes y otras demasiado pequeñas, pero a la larga no habrá tendencia sistemática a subestimar o sobreestimar θ. ■

> **Principio de estimación insesgada**
>
> Al elegir entre varios estimadores de θ, seleccione uno que sea insesgado.

De acuerdo con este principio, se debe preferir el estimador insesgado $\hat{\theta}_2$ del ejemplo 6.4 al estimador sesgado $\hat{\theta}_1$. Ahora considere el problema de estimar σ^2.

PROPOSICIÓN

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2. Entonces el estimador

$$\hat{\sigma}^2 = S^2 = \frac{\sum (X_i - \overline{X})^2}{n - 1}$$

es un estimador insesgado de σ^2.

Demostración Para cualquier *va* Y, $V(Y) = E(Y^2) - [E(Y)]^2$, entonces $E(Y^2) = V(Y) + [E(Y)]^2$. Aplicando esto a

$$S^2 = \frac{1}{n-1}\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right]$$

se obtiene

$$\begin{aligned} E(S^2) &= \frac{1}{n-1}\left\{\sum E(X_i^2) - \frac{1}{n} E[(\sum X_i)^2]\right\} \\ &= \frac{1}{n-1}\left\{\sum (\sigma^2 + \mu^2) - \frac{1}{n}\{V(\sum X_i) + [E(\sum X_i)]^2\}\right\} \\ &= \frac{1}{n-1}\left\{n\sigma^2 + n\mu^2 - \frac{1}{n} n\sigma^2 - \frac{1}{n}(n\mu)^2\right\} \\ &= \frac{1}{n-1}\{n\sigma^2 - \sigma^2\} = \sigma^2 \qquad \text{(como se deseaba)} \end{aligned}$$

■

El estimador que emplea el divisor n se puede expresar como $(n-1)S^2/n$, así que

$$E\left[\frac{(n-1)S^2}{n}\right] = \frac{n-1}{n} E(S^2) = \frac{n-1}{n}\sigma^2$$

Este estimador es, por lo tanto, no insesgado. El sesgo es $(n-1)\sigma^2/n - \sigma^2 = -\sigma^2/n$. Debido a que el sesgo es negativo, el estimador con divisor n tiende a subestimar σ^2, y ésta es la razón de que muchos estadísticos prefieran el divisor $n - 1$ (aunque cuando n es grande, el sesgo es pequeño y hay poca diferencia entre los dos).

Aunque S^2 es insesgado para σ^2, S es un estimador sesgado de σ (su sesgo es pequeño, a menos que n sea bastante pequeña). Sin embargo, hay otras buenas razones para usar S como un estimador, en particular cuando la distribución de la población es normal. Éstas serán más evidentes cuando se analicen intervalos de confianza y prueba de hipótesis en los siguientes capítulos.

En el ejemplo 6.2, se proponen varios estimadores para la media μ de una distribución normal. Si hubiera un solo estimador insesgado para μ, el problema de estimación se resolvería por medio de ese estimador. Infortunadamente, éste no es el caso.

PROPOSICIÓN

Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución con media μ, entonces $\overline{X}$ es un estimador insesgado de μ. Si además la distribución es continua y simétrica, entonces $\tilde{X}$ y cualquier media recortada son también estimadores insesgados de μ.

El hecho de que $\overline{X}$ sea insesgado es sólo una reafirmación de una de las reglas de valor esperado: $E(\overline{X}) = \mu$ para todo valor posible de μ (para distribuciones discretas así como continuas). Es más difícil comprobar la falta de sesgo de los otros estimadores.

De acuerdo con esta proposición, el principio de insesgamiento por sí mismo no siempre permite seleccionar un solo estimador. Cuando la población subyacente es normal, incluso el tercer estimador del ejemplo 6.2 es insesgado, y hay muchos otros estimadores insesgados. Lo que se necesita ahora es una forma de seleccionar entre estimadores insesgados.

Estimadores con varianza mínima

Suponga que $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores de θ, ambos insesgados. Entonces, aunque la distribución de cada estimador se centra en el valor verdadero de θ, las dispersiones de las distribuciones con respecto al valor verdadero podrían ser diferentes.

> **Principio de estimación insesgada de varianza mínima**
>
> Entre los estimadores de θ que están insesgados, elija uno que tiene una varianza mínima. La $\hat{\theta}$ resultante se llama **estimador insesgado de varianza mínima (EIVM)** de θ.

En la figura 6.2 se ilustra la fdp de dos estimadores insesgados, con $\hat{\theta}_1$ que tiene varianza más pequeña que $\hat{\theta}_2$. Entonces $\hat{\theta}_1$ tiene más probabilidades que $\hat{\theta}_2$ de producir una estimación cercana al valor verdadero de θ. El EIVM es, en cierto sentido, el más probable entre los estimadores insesgados para producir una estimación cercana al valor verdadero de θ.

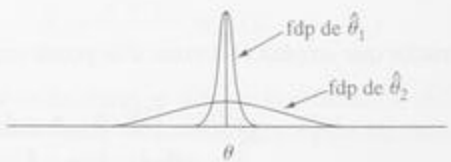


Figura 6.2 Gráficas de las fdp de dos estimadores insesgados diferentes

Example 6.5 En el ejemplo 6.4 se argumentó que cuando $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución uniforme en $[0, \theta]$, el estimador

$$\hat{\theta}_1 = \frac{n+1}{n} \cdot \text{máx}(X_1, \ldots, X_n)$$

es insesgado para θ (antes el estimador se representó con $\hat{\theta}_2$). Éste no es el único estimador insesgado de θ. El valor esperado de una *va* con distribución uniforme es sólo el punto medio del intervalo de densidad positiva, por lo tanto $E(X_i) = \theta/2$. Esto significa que $E(\overline{X}) = \theta/2$, de la cual $E(2\overline{X}) = \theta$. Es decir, el estimador $\hat{\theta}_2 = 2\overline{X}$ es insesgado para θ.

Si X tiene una distribución uniforme en el intervalo $[A, B]$, entonces $V(X) = \sigma^2 = (B - A)^2/12$. Por consiguiente, en la situación que aquí interesa, $V(X_i) = \theta^2/12$, $V(\overline{X}) = \sigma^2/n = \theta^2/(12n)$ y $V(\hat{\theta}_2) = V(2\overline{X}) = 4V(\overline{X}) = \theta^2/(3n)$. Los resultados del ejercicio 32 se pueden usar para demostrar que $V(\hat{\theta}_1) = \theta^2/[n(n+2)]$. El estimador $\hat{\theta}_1$ tiene varianza más pequeña que $\hat{\theta}_2$ si $3n < n(n+2)$ es decir, si $0 < n^2 - n = n(n-1)$. Siempre y cuando $n > 1$, $V(\hat{\theta}_1) < V(\hat{\theta}_2)$, de modo que $\hat{\theta}_1$ es un mejor estimador que $\hat{\theta}_2$. Se pueden usar métodos más avanzados para de-

mostrar que $\hat{\theta}_1$ es el EIVM de θ, cualquier otro estimador insesgado de θ tiene varianza que pasa de $\theta^2/[n(n+2)]$. ■

Uno de los triunfos de la estadística matemática ha sido el desarrollo de metodología para identificar el EIVM en una amplia variedad de situaciones. El resultado más importante de este tipo para los fines de este libro tiene que ver con estimar la media μ de una distribución normal.

TEOREMA

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal con parámetros μ y σ. Entonces el estimador $\hat{\mu} = \overline{X}$ es el EIVM para μ.

Siempre que se esté convencido de que la población en la que se realiza el muestreo es normal, el resultado establece que se debe usar $\overline{X}$ para estimar μ. En el ejemplo 6.2, entonces, la estimación sería $\bar{x} = 27.793$.

En algunas situaciones, es posible obtener un estimador con sesgo pequeño que se preferiría para el mejor estimador insesgado. Esto se ilustra en la figura 6.3. Sin embargo, los EIVM se obtienen con más facilidad que el tipo de estimador sesgado cuya distribución se ilustra.

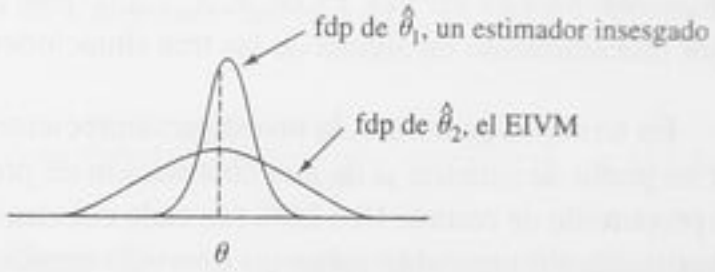


Figura 6.3 Un estimador sesgado preferible al EIVM

Algunas complicaciones

El último teorema no dice que al estimar una media poblacional μ, el estimador $\overline{X}$ se debe usar sin tomar en cuenta la distribución que está siendo muestreada.

Ejemplo 6.6 Suponga que se desea estimar la conductividad térmica μ de cierto material. Por medio de técnicas estándar de medición, se obtiene una muestra aleatoria $X_1, \ldots, X_n$ de n mediciones de conductividad térmica. Supóngase que la distribución poblacional es un miembro de una de las tres familias siguientes:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/(2\sigma^2)} \qquad -\infty < x < \infty \tag{6.1}$$

$$f(x) = \frac{1}{\pi[1 + (x-\mu)^2]} \qquad -\infty < x < \infty \tag{6.2}$$

$$f(x) = \begin{cases} \dfrac{1}{2c} & -c \le x - \mu \le c \\ 0 & \text{en caso contrario} \end{cases} \tag{6.3}$$

La fdp (6.1) es la distribución normal, (6.2) se llama distribución de Cauchy y (6.3) es una distribución uniforme. Las tres distribuciones son simétricas respecto a μ, y de hecho la distribución de Cauchy es con forma de campana, pero con extremos mucho más densos (más probabilidad lejana) que la curva normal. La distribución uniforme no tiene extremos. Los cuatro estimadores para μ considerados antes son $\bar{X}$, $\tilde{X}$, $\bar{X}_e$ (el promedio de las dos observaciones extremas), y $\bar{X}_{\text{rec}(10)}$, una media recortada.

La moraleja es que el mejor estimador para μ depende mucho de cuál distribución esté siendo muestreada. En particular,

1. Si la muestra aleatoria viene de una distribución normal, entonces $\bar{X}$ es el mejor de los cuatro estimadores, puesto que tiene una varianza mínima entre los estimadores insesgados.
2. Si la muestra aleatoria viene de una distribución de Cauchy, entonces $\bar{X}$ y $\bar{X}_e$ son estimadores pésimos para μ, mientras que $\tilde{X}$ es bastante bueno (se desconoce el EIVM); $\bar{X}$ es malo porque es muy sensible a observaciones alejadas, y los extremos densos de la distribución de Cauchy hacen que algunas de estas observaciones aparezcan probablemente en cualquier muestra.
3. Si la distribución subyacente es uniforme, el mejor estimador es $\bar{X}_e$; las observaciones subyacentes afectan en gran medida a este estimador, pero la falta de extremos hace imposibles esta clase de observaciones.
4. *En ninguna de estas tres situaciones la media recortada es la mejor, pero funciona razonablemente bien en las tres.* Es decir, $\bar{X}_{\text{rec}(10)}$ no está en desventaja en comparación con el mejor procedimiento en alguna de las tres situaciones. ∎

En términos generales, la investigación reciente en estadística ha establecido que al estimar un punto de simetría μ de una distribución de probabilidad continua, una media recortada con proporción de recorte 10 o 20% (de cada extremo de la muestra) produce estimaciones con comportamiento razonable sobre un intervalo amplio de modelos posibles. Por esta razón, se dice que una media recortada con porcentaje de recorte pequeño es un **estimador robusto.**

En algunas situaciones, la elección no está entre dos estimadores diferentes construidos a partir de la misma muestra, sino entre estimadores basados en dos experimentos diferentes.

Ejemplo 6.7 Suponga que cierto tipo de componente tiene una distribución de tiempo de vida útil que es exponencial con parámetro λ, de modo que el tiempo de vida útil esperado es $\mu = 1/\lambda$. Se elige una muestra de estos componentes, y se pone en operación cada uno. Si se continúa el experimento hasta haber observado los n tiempos de vida, $X_1, \ldots, X_n$, entonces $\bar{X}$ es un estimador insesgado de μ.

No obstante, en algunos experimentos, los componentes se dejan en operación sólo hasta el momento de la r-ésima falla, donde $r < n$. Este procedimiento se llama **censura**. Sea Y_1 el tiempo de la primera falla (el tiempo de vida mínimo entre los n componentes), Y_2 el momento en que ocurre la segunda falla (el segundo tiempo de vida más pequeño), etcétera. Puesto que el experimento termina en el tiempo Y_r, el tiempo de vida total acumulado en la terminación es

$$T_r = \sum_{i=1}^{r} Y_i + (n - r)Y_r$$

Ahora se demuestra que $\hat{\mu} = T_r/r$ es un estimador insesgado para μ. Para esto se requieren dos propiedades de variables exponenciales:

1. La propiedad de falto de memoria o amnesia (véase la sección 4.4), la cual establece que en cualquier punto del tiempo, el tiempo de vida remanente tiene la misma distribución exponencial que el tiempo de vida original.

2. Si $X_1, \ldots, X_k$ son independientes, cada una exponencial con parámetro λ, entonces $\min(X_1, \ldots, X_k)$ es exponencial con parámetro $k\lambda$ y tiene valor esperado $1/(k\lambda)$.

Puesto que los n componentes duran hasta Y_1, $n-1$ dura una cantidad de tiempo extra $Y_2 - Y_1$, $n-2$ una cantidad extra $Y_3 - Y_2$, etcétera, otra expresión para T_r es

$$T_r = nY_1 + (n-1)(Y_2 - Y_1) + (n-2)(Y_3 - Y_2) + \cdots + (n-r+1)(Y_r - Y_{r-1})$$

Sin embargo Y_1 es el mínimo de n variables exponenciales, así que $E(Y_1) = 1/(n\lambda)$. De manera similar, $Y_2 - Y_1$ es el más pequeño de los $n-1$ tiempos de vida restantes, cada uno exponencial con parámetro λ (por la propiedad de amnesia), así que $E(Y_2 - Y_1) = 1/[(n-1)\lambda]$. Continuando, $E(Y_{i+1} - Y_i) = 1/[(n-i)\lambda]$, así que

$$\begin{aligned} E(T_r) &= nE(Y_1) + (n-1)E(Y_2 - Y_1) + \cdots + (n-r+1)E(Y_r - Y_{r-1}) \\ &= n \cdot \frac{1}{n\lambda} + (n-1) \cdot \frac{1}{(n-1)\lambda} + \cdots + (n-r+1) \cdot \frac{1}{(n-r+1)\lambda} \\ &= \frac{r}{\lambda} \end{aligned}$$

Por lo tanto, $E(T_r/r) = (1/r)E(T_r) = (1/r) \cdot (r/\lambda) = 1/\lambda = \mu$ como se dijo.

Como ejemplo, suponga que se ponen a prueba 20 componentes y que $r = 10$. Entonces de las primeras diez fallas son las veces 11, 15, 29, 33, 35, 40, 47, 55, 58 y 72, la estimación de μ es

$$\hat{\mu} = \frac{11 + 15 + \cdots + 72 + (10)(72)}{10} = 111.5$$

La ventaja del experimento con censura es que termina más rápido que el experimento sin censura. Sin embargo, se puede demostrar que $V(T_r/r) = 1/(\lambda^2 r)$, la cuál es más grande que $1/(\lambda^2 n)$, la varianza de $\overline{X}$ en el experimento sin censura. ■

Informe de una estimación puntual: error estándar

Además de informar el valor de una estimación puntual, se debe proporcionar alguna indicación de su precisión. La medida usual de precisión es el error estándar del estimador empleado.

DEFINICIÓN

El **error estándar** de un estimador $\hat{\theta}$ es su desviación estándar $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$. Si por sí mismo el error estándar tiene que ver con parámetros desconocidos cuyos valores se pueden estimar, la substitución de estas estimaciones en $\sigma_{\hat{\theta}}$ produce el **error estándar estimado** (desviación estándar estimada) del estimador. El error estándar estimado se puede denotar ya sea con $\hat{\sigma}_{\hat{\theta}}$ (el $\hat{}$ sobre σ remarca que se lleva a cabo la estimación de $\sigma_{\hat{\theta}}$) o con $s_{\hat{\theta}}$.

Ejemplo 6.8 (Continuación del ejemplo 6.2)

Suponiendo que la tensión disruptiva tiene una distribución normal, $\hat{\mu} = \overline{X}$ es el mejor estimador de μ. Si se sabe que el valor de $\sigma = 1.5$, el error estándar de $\overline{X}$ es $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 1.5/\sqrt{20} = 0.335$. Si, como por lo común es el caso, se desconoce el valor de σ, la estimación $\hat{\sigma} = s = 1.462$ se sustituye en $\sigma_{\overline{X}}$ para obtener el error estándar estimado $\hat{\sigma}_{\overline{X}} = s_{\overline{X}} = s/\sqrt{n} = 1.462/\sqrt{20} = 0.327$. ■

Ejemplo 6.9 (Continuación del ejemplo 6.1)

El error estándar de $\tilde{p} = X/n$ es

$$\sigma_{\tilde{p}} = \sqrt{V(X/n)} = \sqrt{\frac{V(X)}{n^2}} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}$$

Puesto que se desconocen p y $q = 1 - p$ (¿de otro modo por qué estimar?), se sustituye $\hat{p} = x/n$ y $\hat{q} = 1 - x/n$ en $\sigma_{\tilde{p}}$, y se obtiene el error estándar estimado $\hat{\sigma}_{\tilde{p}} = \sqrt{\hat{p}\hat{q}/n} = \sqrt{(0.6)(0.4)/25} = 0.098$. De manera alternativa, puesto que el valor más grande de pq se obtiene cuando $p = q = 0.5$, un límite superior del error estándar es $\sqrt{1/(4n)} = 0.10$. ■

Cuando el estimador puntual $\hat{\theta}$ tiene más o menos una distribución normal, que a menudo será el caso cuando n es grande, entonces se puede tener cierta confianza de que el valor verdadero de θ se ubica dentro de casi dos errores estándar (desviaciones estándar) de $\hat{\theta}$. Así, si una muestra de $n = 36$ tiempos de vida de los componentes da $\hat{\mu} = \bar{x} = 28.50$ y $s = 3.60$, entonces $s/\sqrt{n} = 0.60$, de modo que dentro de dos errores estándar estimados de $\hat{\mu}$ se traduce al intervalo $28.50 \pm (2)(0.60) = (27.30, 29.70)$.

Si $\hat{\theta}$ no necesariamente es más o menos normal, sino que es insesgada, entonces se puede demostrar que la estimación se desvía de θ por tanto como cuatro errores estándar a lo mucho 6% de las veces. Entonces esperaríamos que el valor verdadero estaría dentro de los 4 errores estándar de $\hat{\theta}$ (y éste es un enunciado muy conservador, puesto que se aplica a *cualquier* $\hat{\theta}$ insesgada). Resumiendo, el error estándar indica más o menos dentro de qué distancia de $\hat{\theta}$ se puede esperar que esté el valor verdadero de θ.

La forma del estimador $\hat{\theta}$ podría ser suficientemente complicada para que no se pueda aplicar la teoría estadística, a fin de obtener una expresión para $\sigma_{\hat{\theta}}$. Esto se cumple, por ejemplo, en el caso $\theta = \sigma$, $\hat{\theta} = S$; en general, no se puede determinar la desviación estándar de la estadística S, σ_S. En años recientes, se introdujo un nuevo método llamado **inicialización (bootstrap)**, que requiere los recursos de muchas computadoras, para atender este problema. Suponga que la fdp es $f(x; \theta)$, un miembro de una familia paramétrica particular, y que con los datos $x_1, x_2, \ldots, x_n$ se obtiene $\hat{\theta} = 21.7$. Ahora, se usa la computadora para obtener "las muestras de inicialización" a partir de la fdp $f(x; 21.7)$, y para cada muestra se calcula una "estimación de inicialización" $\hat{\theta}^*$:

Primera muestra de inicialización: $x_1^*, x_2^*, \ldots, x_n^*$; estimación $= \hat{\theta}_1^*$

Segunda muestra de inicialización: $x_1^*, x_2^*, \ldots, x_n^*$; estimación $= \hat{\theta}_2^*$

$\vdots$

B-ésima muestra de inicialización: $x_1^*, x_2^*, \ldots, x_n^*$; estimación $= \hat{\theta}_B^*$

Con frecuencia se utiliza $B = 100$ o 200. Ahora, sea $\bar{\theta}^* = \sum \hat{\theta}_i^*/B$, la media muestral de las estimaciones de inicialización. La **estimación de inicialización** del error estándar de las $\hat{\theta}$ es ahora la desviación estándar de las $\hat{\theta}_i^*$:

$$S_{\hat{\theta}} = \sqrt{\frac{1}{B-1}\sum(\hat{\theta}_i^* - \bar{\theta}^*)^2}$$

(En las publicaciones acerca de la iniciacilización (*bootstrap*), se utiliza B en lugar de $B - 1$; para valores representativos de B, por lo general hay poca diferencia entre las estimaciones resultantes).

Ejemplo 6.10 Un modelo teórico indica que X, el tiempo hasta que falla un líquido aislante entre electrodos a un voltaje particular, tiene $f(x; \lambda) = \lambda e^{-\lambda x}$, una distribución exponencial de $n = 10$ tiempos de falla (mín) da los datos siguientes:

41.53 18.73 2.99 30.34 12.33 117.52 73.02 223.63 4.00 26.78

Puesto que $E(X) = 1/\lambda$, $E(\bar{X}) = 1/\lambda$, así que una estimación razonable de λ es $\hat{\lambda} = 1/\bar{x} = 1/55.087 = 0.018153$. Entonces se utiliza un paquete de software de estadística para obtener $B = 100$ muestras de inicialización, cada una de tamaño 10, de $f(x; 0.018153)$. La primera muestra fue 41.00, 109.70, 16.78, 6.31, 6.76, 5.62, 60.96, 78.81, 192.25, 27.61, de la cual $\sum x_i^* = 545.8$ y $\hat{\lambda}_1^* = 1/54.58 = 0.01832$. El promedio de las 100 estimaciones de inicialización es $\bar{\lambda}^* = 0.02153$, y la desviación estándar muestral de estas 100 estimaciones es $s_{\hat{\lambda}} = 0.0091$, la estimación de inicialización del error estándar de las $\hat{\lambda}$. Un histograma de las 100 $\hat{\lambda}_i^*$ resultó con un poco de sesgo positivo, lo que hace pensar que la distribución de muestreo de $\hat{\lambda}$ también tiene esta propiedad. ■

En ocasiones, un investigador desea estimar una característica poblacional sin asumir que la distribución de la población pertenece a una familia paramétrica particular. En el ejemplo 6.6 ocurrió un ejemplo de esto, donde se propuso una media recortada al 10% para estimar un centro simétrico de la distribución de la población θ. Con los datos del ejemplo 6.2 se obtuvo $\hat{\theta} = \bar{x}_{\text{rec}(10)} = 27.838$, pero ahora no hay una $f(x; \theta)$, supuesta, así que ¿cómo se obtiene una muestra de inicialización? La respuesta es considerar que la muestra constituye la población (las $n = 10$ observaciones del ejemplo 6.2) y tomar B muestras diferentes, cada una de tamaño n, *con* reemplazo a partir de esta población. En el libro de Bradley Efron y Robert Tibshirani o el de John Rice listados en la bibliografía del capítulo, se proporciona más información.

Ejercicios | Sección 6.1 (1 a 19)

1. Los siguientes datos acerca de la resistencia a la flexión (MPa) de vigas de concreto de cierto tipo se presentaron en el ejemplo 1.2.

   5.9 7.2 7.3 6.3 8.1 6.8 7.0
   7.6 6.8 6.5 7.0 6.3 7.9 9.0
   8.2 8.7 7.8 9.7 7.4 7.7 9.7
   7.8 7.7 11.6 11.3 11.8 10.7

   a. Calcule una estimación puntual del valor medio de resistencia para la población conceptual de todas las vigas fabricadas de este modo, y diga qué estimador usó. (*Sugerencia:* $\sum x_i = 219.8$.)
   b. Calcule una estimación puntual del valor de resistencia que separa al 50% más débil de las vigas del 50% más fuerte, y exprese qué estimador usó.
   c. Calcule e interprete una estimación puntual de la desviación estándar poblacional σ. ¿Qué estimador usó? (*Sugerencia:* $\sum x_i^2 = 1860.94$.)
   d. Calcule una estimación puntual de la proporción de las vigas cuya resistencia a la flexión es mayor que 10 MPa. (*Sugerencia*: imagine que una observación es un "éxito" si pasa de 10.)
   e. Calcule una estimación puntual del coeficiente poblacional de variación σ/μ, e indique qué estimador usó.

2. Una muestra de 20 estudiantes que recientemente tomaron un curso de estadística elemental produjo la siguiente información acerca de la marca de calculadora que tenían (T = Texas Instruments, H = Hewlett Packard, C = Casio, S = Sharp):

   T T H T C T T S C H
   S S T H C T T T H T

a. Estime la proporción real de los estudiantes que poseen una calculadora Texas Instruments.
b. De los 10 estudiantes que poseían una calculadora TI, cuatro tenían calculadoras de graficación. Estime la proporción de estudiantes que no poseen una calculadora de graficación TI.

3. Considere la siguiente muestra de observaciones acerca del espesor de recubrimiento para pintura de baja viscosidad ("Achieving a Target Value for a Manufacturing Process: A Case Study", *J. of Quality Technology*, 1992: 22-26):

0.83 0.88 0.88 1.04 1.09 1.12 1.29 1.31
1.48 1.49 1.59 1.62 1.65 1.71 1.76 1.83

Suponga que la distribución del espesor de recubrimiento es normal (una gráfica de probabilidad normal respalda esta suposición).

a. Calcule una estimación puntual del valor promedio del espesor de recubrimiento y diga qué estimador usó.
b. Calcule una estimación puntual de la mediana de la distribución de espesor de recubrimiento y exprese qué estimador usó.
c. Calcule una estimación puntual del valor que separa el 10% más grande de los valores de la distribución de espesor del 90% restante, e indique qué estimador usó. (*Sugerencia*: exprese lo que trata de estimar en términos de μ y σ.)
d. Estime $P(X < 1.5)$, es decir, la proporción de los valores de espesor menores que 1.5. (*Sugerencia*: Si se conocieran los valores de μ y σ, se podría calcular esta probabilidad. Estos valores no están disponibles, pero se pueden estimar.)
e. ¿Cuál es el error estándar estimado del estimador que usó en el inciso (b)?

4. En el artículo de donde se tomaron los datos del ejercicio 1 también se proporcionaron las siguientes observaciones de resistencia para cilindros:

6.1 5.8 7.8 7.1 7.2 9.2 6.6 8.3 7.0 8.3
7.8 8.1 7.4 8.5 8.9 9.8 9.7 14.1 12.6 11.2

Antes de obtener datos, represente las resistencias de las vigas con $X_1, \ldots, X_m$ y las de los cilindros con $Y_1, \ldots, Y_n$. Suponga que las X_i constituyen una muestra aleatoria de una distribución con media μ_1 y desviación estándar σ_1 y que las Y_i forman una muestra aleatoria (independiente de las X_i) de otra distribución con media μ_2 y desviación estándar σ_2.

a. Utilice las reglas del valor esperado para demostrar que $\overline{X} - \overline{Y}$ es un estimador insesgado de $\mu_1 - \mu_2$. Calcule la estimación para los datos proporcionados.
b. Aplique las reglas de la varianza del capítulo 5 a fin de obtener una ecuación para la varianza y la desviación estándar (error estándar) del estimador del inciso (a); luego, calcule el error estándar estimado.
c. Calcule una estimación puntual de la relación σ_1/σ_2 de las dos desviaciones estándar.
d. Suponga que se elige al azar una sola viga y un solo cilindro. Calcule una estimación puntual de la varianza de la diferencia $X - Y$ entre la resistencia de la viga y la del cilindro.

5. Como ejemplo de una situación en la que se podrían utilizar razonablemente estadísticas diferentes para calcular una estimación puntual, considere una población de N facturas. Relacionado con cada factura está su "valor en libros", la cantidad registrada de esa factura. Sea T el valor total en libros, una cantidad conocida. Algunos de estos valores en libros son erróneos. Se llevará a cabo una auditoría seleccionando al azar n facturas y determinando el valor auditado (correcto) de cada una. Suponga que la muestra proporciona los resultados siguientes (en dólares).

	Factura				
	1	2	3	4	5
Valor en libros	300	720	526	200	127
Valor auditado	300	520	526	200	157
Error	0	200	0	0	−30

Sea

$\overline{Y}$ = valor en libros medio muestral

$\overline{X}$ = valor auditado medio muestral

$\overline{D}$ = error medio muestral

Han sido propuestas varias estadísticas diferentes para estimar el valor auditado total (correcto) (véase "Statistical Models and Analysis in Auditing", *Statistical Science*, 1989: 2-33). Éstas son

Estadístico de media por unidad = $N\overline{X}$

Estadístico de diferencia = $T - N\overline{D}$

Estadístico de relación = $T \cdot (\overline{X}/\overline{Y})$

Si $N = 5000$ y $T = 1\,761\,300$, calcule las tres estimaciones puntuales correspondientes. (En el artículo citado, se analizan las propiedades de estos estimadores.)

6. Considere las siguientes observaciones acerca del caudal (miles de acres-pies) registrado en una estación en Colorado para el periodo del primero de abril al 31 de agosto en un lapso de 31 años (de un artículo del volumen de 1974 de *Water Resources Research*).

127.96	210.07	203.24	108.91	178.21
285.37	100.85	89.59	185.36	126.94
200.19	66.24	247.11	299.87	109.64
125.86	114.79	109.11	330.33	85.54
117.64	302.74	280.55	145.11	95.36
204.91	311.13	150.58	262.09	477.08
94.33				

Una gráfica apropiada de probabilidad apoya el uso de la distribución lognormal (véase la sección 4.5) como un modelo razonable para el caudal.

a. Estime los parámetros de la distribución. [*Sugerencia*: recuerde que X tiene una distribución lognormal con parámetros μ y σ^2 si $\ln(X)$ tiene una distribución normal con media μ y varianza σ^2.]

b. Utilice las estimaciones del inciso (a) para calcular una estimación del valor esperado del caudal [*Sugerencia*: ¿cuál es $E(X)$?]

7. a. Se selecciona una muestra aleatoria de 10 casas en cierta zona, cada una tiene calefacción con gas natural, y para cada casa se determina la cantidad de gas (termias) empleada durante el mes de enero. Las observaciones resultantes son 103, 156, 118, 89, 125, 147, 122, 109, 138, 99. Sea μ el consumo promedio de gas durante enero de las casas de la zona. Calcule una estimación puntual de μ.

b. Suponga que hay 10 mil casas en esta zona que utilizan gas natural para la calefacción. Designe a τ como la cantidad total de gas que emplearon todas las casas durante enero. Estime τ mediante los datos del inciso (a). ¿Qué estimador utilizó para calcular la estimación?

c. Utilice los datos del inciso (a) para estimar p, la proporción de las casas que utilizaron por lo menos 100 termias.

d. Dé una estimación puntual del consumo medio poblacional (el valor medio de la población de todas las casas) con base en la muestra del inciso (a). ¿Cuál estimador utilizó?

8. En una muestra aleatoria de 80 componentes de cierto tipo, 12 resultaron defectuosos.

a. Dé una estimación puntual de la proporción de los componentes que *no* estén defectuosos.

b. Se construirá un sistema seleccionando al azar de dos de estos componentes y conectándolos en serie, como se muestra aquí.

La conexión en serie implica que el sistema funcionará si y sólo si ningún componente está defectuoso (es decir, ambos componentes funcionan de manera correcta). Estime la proporción de los sistemas que funcionan correctamente. [*Sugerencia*: si p representa la probabilidad de que un componente funcione de manera correcta, ¿cómo se puede expresar P(el sistema funciona) en términos de p?]

9. Se examina cada uno de 150 artículos recién fabricados y se registra el número de imperfecciones por pieza (se supone que los artículos están libres de imperfecciones), obteniéndose los datos siguientes:

Número de imperfecciones por artículo	0	1	2	3	4	5	6	7
Frecuencia observada	18	37	42	30	13	7	2	1

Sea X = número de imperfecciones en un artículo elegido al azar, y suponga que X tiene una distribución de Poisson con parámetro λ.

a. Encuentre un estimador insesgado de λ y calcule la estimación para los datos. [*Sugerencia:* $E(X) = \lambda$ para la distribución de Poisson de X, por lo que $E(\bar{X}) = ?$]

b. ¿Cuál es la desviación estándar (error estándar) de su estimador? Calcule el error estándar estimado. (*Sugerencia:* $\sigma_X^2 = \lambda$ para la distribución de Poisson de X.)

10. Mediante una varilla larga de longitud μ se construirá una parcela cuadrada en la que la longitud de cada lado es μ. Por consiguiente, el área de la parcela será μ^2. Sin embargo, se desconoce el valor de μ, por lo que decide hacer n mediciones independientes $X_1, X_2, \ldots X_n$ de la longitud. Suponga que cada X_i tiene media μ (mediciones insesgadas) y varianza σ^2.

a. Demuestre que $\bar{X}^2$ no es un estimador insesgado para μ^2. [*Sugerencia*: para cualquier *va* Y, $E(Y^2) = V(Y) + [E(Y)]^2$. Aplique esto con $Y = \bar{X}$.]

b. ¿Para cuál valor de k el estimador $\bar{X}^2 - kS^2$ es insesgado para μ^2? [*Sugerencia*: calcule $E(\bar{X}^2 - kS^2)$.]

11. De n_1 fumadores varones seleccionados al azar, X_1 fumaron cigarrillos con filtro, mientras que de n_2 fumadoras seleccionadas al azar, X_2 fumaron cigarrillos con filtro.

Represente con p_1 y p_2 las probabilidades de que un hombre y una mujer seleccionados al azar, respectivamente, fumaron cigarrillos con filtro.

a. Demuestre que $(X_1/n_1) - (X_2/n_2)$ es un estimador insesgado para $p_1 - p_2$. [*Sugerencia:* $E(X_i) = n_i p_i$ para $i = 1, 2$.]

b. ¿Cuál es el error estándar del estimador en el inciso (a)?

c. ¿Cómo se utilizarían los valores observados x_1 y x_2 para estimar el error estándar del estimador?

d. Si $n_1 = n_2 = 200$, $x_1 = 127$ y $x_2 = 176$, utilice el estimador del inciso (a) para obtener una estimación de $p_1 - p_2$.

e. Utilice el resultado del inciso (c) y los datos del inciso (d) para estimar el error estándar del estimador.

12. Suponga que cierto tipo de fertilizante tiene un rendimiento esperado por acre de μ_1 con varianza σ^2, mientras que la producción esperada para un segundo tipo de fertilizante es μ_2 con la misma varianza σ^2. Sean S_1^2 y S_2^2 las varianzas muestrales de rendimientos basados en tamaños muestrales n_1 y n_2, respectivamente, de los dos fertilizantes. Demuestre que el estimador agrupado (combinado)

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

es un estimador insesgado de σ^2.

13. Considere una muestra aleatoria $X_1, \ldots, X_n$ de la fdp

$$f(x; \theta) = 0.5(1 + \theta x) \qquad -1 \le x \le 1$$

donde $-1 \le \theta \le 1$ (esta distribución surge en física de partículas). Demuestre que $\hat{\theta} = 3\bar{X}$ es un estimador insesgado de θ. [*Sugerencia*: primero determine $\mu = E(X) = E(\bar{X})$.]

14. Una muestra de n aviones caza a reacción Pandemonium capturados, da como resultado números de serie $x_1, x_2, x_3, \ldots, x_n$. La CIA sabe que las naves fueron numeradas en forma consecutiva en la fábrica, comenzando con α y terminado con β, así que el número total de aviones construidos es $\beta - \alpha + 1$ (p. ej., si $\alpha = 17$ y $\beta = 29$, entonces se fabricaron $29 - 17 + 1 = 13$ aviones con números de serie 17, 18, 19, . . . , 28, 29). Sin embargo, la CIA desconoce los valores de α o β. Un estadístico de la CIA sugiere usar el estimador máx(X_i) − mín(X_i) + 1 para estimar el número total de aviones fabricados.

a. Si $n = 5$, $x_1 = 237$, $x_2 = 375$, $x_3 = 202$, $x_4 = 525$ y $x_5 = 418$, ¿cuál es la estimación correspondiente?

b. ¿Bajo qué condiciones en la muestra el valor de la estimación será exactamente igual al número verdadero total de aviones? ¿La estimación siempre será mayor que el total verdadero? ¿Puede considerarse que el estimador es insesgado para estimar $\beta - \alpha + 1$? Explique con una o dos oraciones.

15. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución de Rayleigh con fdp

$$f(x; \theta) = \frac{x}{\theta} e^{-x^2/(2\theta)} \qquad x > 0$$

a. Se puede demostrar que $E(X^2) = 2\theta$. Utilice este hecho para construir un estimador insesgado de θ con base en $\sum X_i^2$ (y use las reglas de valor esperado para demostrar que es insesgado).

b. Estime θ de las siguientes $n = 10$ observaciones sobre esfuerzo vibratorio de una paleta de turbina bajo condiciones específicas:

16.88	10.23	4.59	6.66	13.68
14.23	19.87	9.40	6.51	10.95

16. Suponga que el crecimiento promedio real μ de un tipo de planta durante un año es idéntico al de un segundo tipo, pero la varianza de crecimiento para el primer tipo es σ^2, mientras que para el segundo tipo la varianza es $4\sigma^2$. Sean $X_1, \ldots, X_m$ m observaciones de crecimiento independientes en el primer tipo [entonces, $E(X_i) = \mu$, $V(X_i) = \sigma^2$], y sean $Y_1, \ldots, Y_n$ observaciones de crecimiento independientes en el segundo tipo [$E(Y_i) = \mu$, $V(Y_i) = 4\sigma^2$].

a. Demuestre que para cualquier δ entre 0 y 1, el estimador $\hat{\mu} = \delta\bar{X} + (1 - \delta)\bar{Y}$ es insesgado para μ.

b. Para m y n, calcule $V(\hat{\mu})$, y luego determine el valor de δ que minimice a $V(\hat{\mu})$. [*Sugerencia:* derive $V(\hat{\mu})$ con respecto a δ.]

17. En el capítulo 3, se define una *va* binomial negativa como el número de fallas que ocurren antes del r-ésimo éxito en una secuencia de ensayos de éxito o fracaso independientes e idénticos. La función de masa de probabilidad (fmp) de X es

$$nb(x; r, p) = \begin{cases} \binom{x + r - 1}{r - 1} p^r (1 - p)^x & x = 0, 1, 2, \ldots \\ 0 & \text{en caso contrario} \end{cases}$$

a. Suponga que $r \ge 2$. Demuestre que

$$\hat{p} = (r - 1)/(X + r - 1)$$

es un estimador insesgado para p. [*Sugerencia:* escriba $E(\hat{p})$ y cancele $x + r - 1$ dentro de la suma].

b. Un reportero que desea entrevistar a cinco personas que apoyan a cierto candidato empieza preguntándoles si apoyan (S) o no apoyan (F) al candidato. Si la secuencia de respuestas es *SFFSFFFSSS*, estime

p = proporción verdadera de quienes apoyan al candidato.

18. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una fdp $f(x)$ simétrica respecto a μ, de modo que $\bar{X}$ es un estimador insesgado de μ. Si n es grande, se puede demostrar que $V(\tilde{X}) \approx 1/(4n[f(\mu)]^2)$.

a. Compare $V(\tilde{X})$ con $V(\bar{X})$ cuando la distribución subyacente es normal.

b. Cuando la fdp subyacente sea de Cauchy (véase el ejemplo 6.6), $V(\bar{X}) = \infty$, entonces $\bar{X}$ es un estimador pésimo. ¿Cuál es $V(\tilde{X})$ en este caso cuando n es grande?

19. Una investigadora desea estimar la proporción de estudiantes de cierta universidad que han violado el Código de Honor. Al haber obtenido una muestra aleatoria de n estudiantes, se percata de que al preguntar a cada uno: ¿has violado el Código de Honor?, probablemente resultarán algunas respuestas falsas. Considere el siguiente esquema, llamado técnica de **respuesta aleatorizada.** La investigadora forma un mazo de 100 cartas, de las cuales 50 son de tipo I y las otras 50 de tipo II.

Tipo I: ¿Has violado el código de honor (sí o no)?

Tipo II: ¿El último dígito de tu número de teléfono es 0, 1 o 2 (sí o no)?

A cada estudiante de la muestra aleatoria se le pide barajar el mazo, sacar una carta y contestar con la verdad la pregunta resultante. Como resultado de la pregunta irrelevante en las cartas tipo II, una respuesta afirmativa ya no deshonra al estudiante, así que se supone que las respuestas son verdaderas. Indique con p la proporción de transgresores del código de honor (es decir, la probabilidad de que un estudiante seleccionado al azar sea un quebrantador) y sea $\lambda = P(\text{respuesta afirmativa})$. Entonces, λ y p están relacionadas mediante $\lambda = 0.5p + (0.5)(0.3)$.

a. Sea Y el número de respuestas afirmativas, así $Y \sim \text{Bin}(n, \lambda)$. Por lo tanto, Y/n es un estimador insesgado de λ. Derive un estimador para p con base en Y. Si $n = 80$ y $y = 20$, ¿cuál es su estimación? (*Sugerencia*: de la ecuación $\lambda = 0.5p + 0.15$ despeje p y luego sustituya λ por Y/n.)

b. Utilice el hecho de que $E(Y/n) = \lambda$ para demostrar que nuestro estimador $\hat{p}$ es insesgado.

c. Si hubiera 70 cartas tipo I y 30 tipo II, ¿cuál sería el estimador para p?

6.2 Métodos de estimación puntual

La definición de insesgamiento no indica en general cómo se pueden obtener estimadores insesgados. Ahora se analizan dos métodos "constructivos" para obtener estimadores puntuales: el método de momentos y el de máxima verosimilitud. Con la palabra constructivo se indica que la definición general de cada tipo de estimador hace pensar de manera explícita en cómo obtener el estimador en algún problema específico. Aunque los estimadores de máxima verosimilitud son en general preferibles a los estimadores de momentos debido a ciertas propiedades de eficiencia, suelen requerir muchos más cálculos que los estimadores de momentos. A veces sucede que estos métodos producen estimadores insesgados.

Método de momentos

La idea básica de este método es igualar ciertas características muestrales, como la media, con los valores esperados poblacionales correspondientes. Entonces, resolver estas ecuaciones para valores de parámetros desconocidos produce los estimadores.

DEFINICIÓN

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una fmp o fdp $f(x)$. Para $k = 1, 2, 3, \ldots$, el **k-ésimo momento poblacional**, o **k-ésimo momento de la distribución $f(x)$**, es $E(X^k)$. El **k-ésimo momento muestral** es $(1/n)\sum_{i=1}^{n} X_i^k$.

Así, el primer momento poblacional es $E(X) = \mu$ y el primer momento muestral es $\sum X_i/n = \bar{X}$. Los segundos momentos poblacionales muestrales son $E(X^2)$ y $\sum X_i^2/n$, respectivamente. Los momentos poblacionales serán funciones de algunos parámetros desconocidos $\theta_1, \theta_2, \ldots$

DEFINICIÓN

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución con fmp o fdp $f(x; \theta_1, \ldots, \theta_m)$, donde $\theta_1, \ldots, \theta_m$ son parámetros cuyos valores se desconocen. Entonces los **estimadores de momento** $\hat{\theta}_1, \ldots, \hat{\theta}_m$ se obtienen igualando los primeros m momentos muestrales con los primeros m momentos poblacionales correspondientes y resolviendo para $\theta_1, \ldots, \theta_m$.

Si, por ejemplo, $m = 2$, $E(X)$ y $E(X^2)$ serán funciones de θ_1 y θ_2. Con $E(X) = (1/n)\sum X_i\ (= \bar{X})$ y $E(X^2) = (1/n)\sum X_i^2$ se obtienen dos ecuaciones en θ_1 y θ_2. Entonces la solución define los estimadores. Para estimar una media poblacional μ, el método da $\mu = \bar{X}$, de modo que el estimador es la media muestral.

Ejemplo 6.11 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tiempos de servicio de n clientes en cierta instalación, donde se supone que la distribución subyacente es exponencial con parámetro λ. Puesto que sólo hay una parámetro por estimar, el estimador se obtiene igualando $E(X)$ con $\bar{X}$. Como $E(X) = 1/\lambda$ para una distribución exponencial, esto da $1/\lambda = \bar{X}$ o $\lambda = 1/\bar{X}$. Entonces el estimador de momento de λ es $\hat{\lambda} = 1/\bar{X}$. ∎

Ejemplo 6.12 Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución gamma con parámetros α y β. De la sección 4.4, $E(X) = \alpha\beta$ y $E(X^2) = \beta^2\Gamma(\alpha + 2)/\Gamma(\alpha) = \beta^2(\alpha + 1)\alpha$. Los estimadores de momento de α y β se obtienen al resolver

$$\bar{X} = \alpha\beta \qquad \frac{1}{n}\sum X_i^2 = \alpha(\alpha + 1)\beta^2$$

Puesto que $\alpha(\alpha + 1)\beta^2 = \alpha^2\beta^2 + \alpha\beta^2$ y la primera ecuación implica que $\alpha^2\beta^2 = \bar{X}^2$, la segunda ecuación se convierte en

$$\frac{1}{n}\sum X_i^2 = \bar{X}^2 + \alpha\beta^2$$

Ahora, dividiendo cada lado de esta segunda ecuación entre el lado correspondiente de la primera ecuación, y sustituyendo hacia atrás, se obtienen los estimadores

$$\hat{\alpha} = \frac{\bar{X}^2}{(1/n)\sum X_i^2 - \bar{X}^2} \qquad \hat{\beta} = \frac{(1/n)\sum X_i^2 - \bar{X}^2}{\bar{X}}$$

Para ilustrar, los datos de tiempo de supervivencia mencionados en el ejemplo 4.21 son

152 115 109 94 88 137 152 77 160 165
125 40 128 123 136 101 62 153 83 69

con $\bar{x} = 113.5$ y $(1/20)\sum x_i^2 = 14\,087.8$. Las estimaciones son

$$\hat{\alpha} = \frac{(113.5)^2}{14\,087.8 - (113.5)^2} = 10.7 \qquad \hat{\beta} = \frac{14\,087.8 - (113.5)^2}{113.5} = 10.6$$

Estas estimaciones de α y β difieren de los valores que sugieren Gross y Clark porque ellos utilizan una técnica de estimación diferente. ■

Ejemplo 6.13 Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución binomial negativa generalizada con parámetros r y p (sección 3.5). Puesto que $E(X) = r(1-p)/p$ y $V(X) = r(1-p)/p^2$, $E(X^2) = V(X) + [E(X)]^2 = r(1-p)(r-rp+1)/p^2$. Al igualar $E(X)$ con $\overline{X}$ y $E(X^2)$ con $(1/n)\sum X_i^2$ finalmente se obtiene

$$\hat{p} = \frac{\overline{X}}{(1/n)\sum X_i^2 - \overline{X}^2} \qquad \hat{r} = \frac{\overline{X}^2}{(1/n)\sum X_i^2 - \overline{X}^2 - \overline{X}}$$

Como ilustración, Reep, Pollard y Benjamin ("Skill and Chance in Ball Games", *J. Royal Stat. Soc.*, 1971: 623-629) consideran la distribución binomial negativa como un modelo para el número de goles por juego que anotan los equipos de la *National Hockey League*. A continuación se dan los datos para la temporada 1966-1967 (420 juegos):

Goles	0	1	2	3	4	5	6	7	8	9	10
Frecuencia	29	71	82	89	65	45	24	7	4	1	3

Entonces,

$$\bar{x} = \sum x_i/420 = [(0)(29) + (1)(71) + \cdots + (10)(3)]/420 = 2.98$$

y

$$\sum x_i^2/420 = [(0)^2(29) + (1)^2(71) + \cdots + (10)^2(3)]/420 = 12.40$$

Por lo tanto,

$$\hat{p} = \frac{2.98}{12.40 - (2.98)^2} = 0.85 \qquad \hat{r} = \frac{(2.98)^2}{12.40 - (2.98)^2 - 2.98} = 16.5$$

Aunque por definición r es positiva, el denominador de $\hat{r}$ podría ser negativo, lo cual indica que la distribución binomial negativa no es apropiada (o bien, que el estimador de momento es deficiente). ■

Estimación de probabilidad máxima

R. A. Fisher, un genetista y estadístico, fue el primero que introdujo en 1920 el método de máxima verosimilitud. La mayor parte de los estadísticos recomiendan este método, por lo menos cuando el tamaño de muestra es grande, puesto que los estimadores resultantes tienen ciertas propiedades de eficiencia deseables (véase la segunda proposición en la página 275).

Ejemplo 6.14 Se obtiene una muestra de diez nuevos cascos para ciclista fabricados por cierta compañía. Al someterlos a una prueba, se encuentra que los cascos primero, tercero y décimo son defectuosos, en tanto que los otros no. Sea $p = P(\text{casco defectuoso})$ y defina $X_1, \ldots, X_{10}$ mediante $X_i = 1$ si el i-ésimo casco es defectuoso y cero en caso contrario. Entonces las x_i observadas son 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, así que la fmp de la muestra es

$$f(x_1, x_2, \ldots, x_{10}; p) = p(1-p)p \cdot \cdots \cdot p = p^3(1-p)^7 \qquad (6.4)$$

Ahora se pregunta, "¿Para qué valor de p la muestra observada tiene más probabilidades de ocurrir?" Es decir, se desea encontrar el valor de p que maximiza la fmp (6.4) o, en forma equivalente, maximiza el logoritmo natural de (6.4).* Puesto que

$$\ln[f(x_1, \ldots, x_{10}; p)] = 3 \ln(p) + 7 \ln(1 - p) \tag{6.5}$$

que es una función derivable de p, igualando a cero la derivada de (6.5) se obtiene el valor de maximización†

$$\frac{d}{dp} \ln[f(x_1, \ldots, x_{10}; p)] = \frac{3}{p} - \frac{7}{1-p} = 0 \Rightarrow p = \frac{3}{10} = \frac{x}{n}$$

donde x es el número observado de éxitos (cascos defectuosos). Ahora la estimación de p es $\hat{p} = \frac{3}{10}$. Se llama estimación de probabilidad máxima o de máxima verosimilitud porque para $x_1, \ldots, x_{10}$ fija, es el valor de parámetro que maximiza la probabilidad (fmp conjunta) de la muestra observada.

Observe que si se hubiera mencionado sólo que entre los diez cascos hubieron tres defectuosos, la fmp binomial $\binom{10}{3} p^3 (1-p)^7$, que también se maximiza para $\hat{p} = \frac{3}{10}$ sustituiría la ecuación (6.4). ∎

DEFINICIÓN

Sea $X_1, X_2, \ldots, X_n$ tiene fmp o fdp conjunta

$$f(x_1, x_2, \ldots, x_n; \theta_1, \ldots, \theta_m) \tag{6.6}$$

donde los parámetros $\theta_1, \ldots, \theta_m$ tienen valores desconocidos. Cuando $x_1, \ldots, x_n$ son los valores muestrales observados y (6.6) es considerada como una función de $\theta_1, \ldots, \theta_m$ se llama **función de verosimilitud.** Las estimaciones de máxima verosimilitud (emv) $\hat{\theta}_1, \ldots, \hat{\theta}_m$ son los valores de las θ_i que maximizan la función de verosimilitud, así que

$$f(x_1, \ldots, x_n; \hat{\theta}_1, \ldots, \hat{\theta}_m) \geq f(x_1, \ldots, x_n; \theta_1, \ldots, \theta_m) \quad \text{para toda } \theta_1, \ldots, \theta_m$$

Cuando las X_i se sustituyen en lugar de las x_i, resultan los **estimadores de máxima verosimilitud**.

La función de verosimilitud señala qué tan probable es la muestra observada como una función de los valores de parámetros posibles. Al maximizar la probabilidad se obtienen los valores de parámetro para los cuales la muestra observada tiene más probabilidades de haber sido generada; es decir, los valores de parámetro que "mejor concuerdan" con los datos observados.

Ejemplo 6.15 Suponga que $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución exponencial con parámetro λ. Debido a la independencia, la función de verosimilitud es un producto de cada una de las fdp:

$$f(x_1, \ldots, x_n; \lambda) = (\lambda e^{-\lambda x_1}) \cdot \cdots \cdot (\lambda e^{-\lambda x_n}) = \lambda^n e^{-\lambda \Sigma x_i}$$

*Puesto que $\ln[g(x)]$ es una función monotónica, encontrar x para maximizar $\ln[g(x)]$es equivalente a maximizar $g(x)$. En estadística, tomar el logaritmo frecuentemente cambia un producto a una suma, con la que es más fácil trabajar.
†Esta conclusión requiere comprobar la segunda derivada, pero se omiten los detalles.

El ln(probabilidad) es

$$\ln[f(x_1, \ldots, x_n; \lambda)] = n \ln(\lambda) - \lambda \sum x_i$$

Al igualar $(d/d\lambda)$[ln(probabilidad)] a cero, se obtiene $n/\lambda - \sum x_i = 0$, o $\lambda = n/\sum x_i = 1/\bar{x}$. Así, el estimador de probabilidad máxima es $\hat{\lambda} = 1/\bar{X}$; es idéntico al del método de momentos [pero no es un estimador insesgado, puesto que $E(1/\bar{X}) \neq 1/E(\bar{X})$]. ■

Ejemplo 6.16 Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal. La función de verosimilitud es

$$f(x_1, \ldots, x_n; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1-\mu)^2/(2\sigma^2)} \cdot \cdots \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_n-\mu)^2/(2\sigma^2)}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\sum(x_i-\mu)^2/(2\sigma^2)}$$

por lo tanto,

$$\ln[f(x_1, \ldots, x_n; \mu, \sigma^2)] = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2$$

Para encontrar los valores de maximización de μ y σ^2, se deben tomar las derivadas parciales de $\ln(f)$ con respecto a μ y σ^2, igualar a cero y resolver las dos ecuaciones resultantes. Omitiendo los detalles, las estimaciones de máxima verosimilitud son

$$\hat{\mu} = \bar{X} \qquad \hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$

La emv de σ^2 no es el estimador insesgado, así que dos principios de estimación diferentes (insesgamiento y máxima verosimilitud) producen dos estimadores distintos. ■

Ejemplo 6.17 En el capítulo 3, se analizó el uso de la distribución de Poisson para modelar el número de "eventos" que ocurren en una región bidimensional. Suponga que cuando la región R en la que se realiza el muestreo tiene área $a(R)$, el número X de eventos que ocurren en R tiene una distribución de Poisson con parámetro $\lambda a(R)$ (donde λ es el número esperado de eventos por unidad de área), y que las regiones que no se traslapan producen X independientes.

Suponga que un ecologista elige n regiones que no se traslapan, $R_1, \ldots, R_n$ y cuenta el número de plantas de cierta especie encontradas en cada región. La fmp conjunta (verosimilitud) es

$$p(x_1, \ldots, x_n; \lambda) = \frac{[\lambda \cdot a(R_1)]^{x_1} e^{-\lambda \cdot a(R_1)}}{x_1!} \cdot \cdots \cdot \frac{[\lambda \cdot a(R_n)]^{x_n} e^{-\lambda \cdot a(R_n)}}{x_n!}$$

$$= \frac{[a(R_1)]^{x_1} \cdot \cdots \cdot [a(R_n)]^{x_n} \cdot \lambda^{\sum x_i} \cdot e^{-\lambda \sum a(R_i)}}{x_1! \cdot \cdots \cdot x_n!}$$

El ln(verosimilitud) es

$$\ln[p(x_1, \ldots, x_n; \lambda)] = \sum x_i \cdot \ln[a(R_i)] + \ln(\lambda) \cdot \sum x_i - \lambda \sum a(R_i) - \sum \ln(x_i!)$$

Tomando $d/d\lambda \ln(p)$ e igualando a cero, se obtiene

$$\frac{\sum x_i}{\lambda} - \sum a(R_i) = 0$$

por lo tanto,

$$\lambda = \frac{\sum x_i}{\sum a(R_i)}$$

Entonces, el estimador de máxima verosimilitud es $\hat{\lambda} = \sum X_i / \sum a(R_i)$. Esto es razonable desde el punto de vista intuitivo, porque λ es la densidad verdadera (plantas por unidad de área), mientras que $\hat{\lambda}$ es la densidad muestral puesto que $\sum a(R_i)$ es el área total muestreada. Debido a que $E(X_i) = \lambda \cdot a(R_i)$, el estimador es insesgado.

Algunas veces se usa un procedimiento de muestreo alternativo. En lugar de fijar regiones para llevar a cabo el muestreo, el ecologista seleccionará n puntos en la región entera de interés y sea y_i = la distancia desde el i-ésimo punto a la planta más cercana. La función de distribución acumulada (fda) de Y = distancia a la planta más cercana es

$$F_Y(y) = P(Y \le y) = 1 - P(Y > y) = 1 - P\begin{pmatrix}\text{ninguna planta en un}\\ \text{círculo de radio } y\end{pmatrix}$$

$$= 1 - \frac{e^{-\lambda \pi y^2}(\lambda \pi y^2)^0}{0!} = 1 - e^{-\lambda \cdot \pi y^2}$$

Tomando la derivada de $F_Y(y)$ con respecto a y, se obtiene

$$f_Y(y; \lambda) = \begin{cases} 2\pi\lambda y e^{-\lambda \pi y^2} & y \ge 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Si ahora se forma la función de verosimilitud $f_Y(y_1; \lambda) \cdot \cdots \cdot f_Y(y_n; \lambda)$, se deriva ln (verosimilitud), etcétera, la estimación de máxima verosimilitud es

$$\hat{\lambda} = \frac{n}{\pi \sum Y_i^2} = \frac{\text{número de plantas observadas}}{\text{área total muestreada}}$$

que también es una densidad muestral. Se puede demostrar que en un medio disperso (λ pequeña), el método de distancia es mejor en cierto sentido, mientras que en un medio denso, el mejor el primer método de muestreo. ∎

Ejemplo 6.18 Sea $X_1, \ldots, X_n$ una muestra aleatoria de una fdp de Weibull

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\alpha}{\beta^\alpha} \cdot x^{\alpha-1} \cdot e^{-(x/\beta)^\alpha} & x \ge 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Al escribir la verosimilitud y ln(verosimilitud), y luego establecer tanto $(\partial/\partial\alpha)[\ln(f)] = 0$ como $(\partial/\partial\beta)[\ln(f)] = 0$ se obtienen las ecuaciones

$$\alpha = \left[\frac{\sum x_i^\alpha \cdot \ln(x_i)}{\sum x_i^\alpha} - \frac{\sum \ln(x_i)}{n}\right]^{-1} \qquad \beta = \left(\frac{\sum x_i^\alpha}{n}\right)^{1/\alpha}$$

Estas dos ecuaciones no se pueden resolver de manera explícita a fin de obtener fórmulas generales para las estimaciones de máxima verosimilitud $\hat{\alpha}$ y $\hat{\beta}$. En cambio, para cada muestra $x_1, \ldots, x_n$, las ecuaciones se deben resolver por medio de un procedimiento numérico iterativo. Los estimadores de momento pares de α y β son un poco complicados (véase el ejercicio 21). ∎

Cómo estimar funciones de parámetros

En el ejemplo 6.16, se obtuvo la emv de σ^2 cuando la distribución subyacente es normal. La emv de $\sigma = \sqrt{\sigma^2}$, así como muchas otras emv, se obtienen sin dificultad por medio de la proposición siguiente.

PROPOSICIÓN

Principio de invarianza

Sean $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$ los emv de los parámetros $\theta_1, \theta_2, \ldots, \theta_m$. Entonces la emv de cualquier función $h(\theta_1, \theta_2, \ldots, \theta_m)$ de estos parámetros es la función $h(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ de las emv.

Ejemplo 6.19 (Continuación del ejemplo 6.16)

En el caso normal, la emv de μ y σ^2 son $\hat{\mu} = \overline{X}$ y $\hat{\sigma}^2 = \sum(X_i - \overline{X})^2/n$. Para obtener la emv de la función $h(\mu, \sigma^2) = \sqrt{\sigma^2} = \sigma$, sustituya la emv en la función:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \left[\frac{1}{n}\sum(X_i - \overline{X})^2\right]^{1/2}$$

La emv de σ no es la desviación estándar muestral S, aunque están cerca a menos que n sea bastante pequeña. ■

Ejemplo 6.20 (Continuación del ejemplo 6.18)

El valor promedio de una *va* X que tiene una distribución de Weibull es

$$\mu = \beta \cdot \Gamma(1 + 1/\alpha)$$

Por lo tanto, la emv de μ es $\hat{\mu} = \hat{\beta}\Gamma(1 + 1/\hat{\alpha})$, donde $\hat{\alpha}$ y $\hat{\beta}$ son las emv de α y β. En particular, $\overline{X}$ no es la emv de μ, aunque es un estimador insesgado. Por lo menos para n grande, $\hat{\mu}$ es un mejor estimador que $\overline{X}$. ■

Comportamiento de las emv con muestra grande

Mientras que un principio de estimación de máxima verosimilitud tiene considerable atractivo intuitivo, la siguiente proposición de esta razón fundamental para el uso de las emv.

PROPOSICIÓN

En condiciones muy generales en la distribución conjunta de la muestra, cuando el tamaño de muestra n es grande, el estimador de máxima verosimilitud de cualquier parámetro θ es aproximadamente insesgado $[E(\hat{\theta}) \approx \theta]$ y tiene varianza que es casi tan pequeña como se pueda lograr mediante cualquier estimador. Expresado de otra manera, la emv de $\hat{\theta}$ es aproximadamente el EIVM de θ.

Debido a este resultado, y al hecho de que por lo común se pueden usar las técnicas basadas en el cálculo para obtener la emv (aunque suelen ser necesarios métodos numéricos, como el de Newton), la estimación de máxima verosimilitud es la técnica de estimación de uso más extendido entre los estadísticos. Muchos de los estimadores utilizados en el resto del libro son emu. Sin embargo, obtener una emv requiere que se especifique la distribución subyacente.

Algunas complicaciones

En ocasiones no se puede usar el cálculo para obtener la emv.

Ejemplo 6.21 Suponga que el tiempo que cierta persona espera el autobús tiene una distribución uniforme en $[0, \theta]$ y se han observado los resultados $x_1, \ldots, x_n$ de una muestra aleatoria a partir de esta distribución. Puesto que $f(x; \theta) = 1/\theta$ para $0 \le x \le \theta$ la probabilidad es 0,

$$f(x_1, \ldots, x_n; \theta) = \begin{cases} \dfrac{1}{\theta^n} & 0 \le x_1 \le \theta, \ldots, 0 \le x_n \le \theta \\ 0 & \text{en caso contrario} \end{cases}$$

Pero tan pronto como $\text{máx}(x_i) \le \theta$, la probabilidad es $1/\theta^n$, que es positiva, pero tan pronto como $\theta < \text{máx}(x_i)$, la probabilidad cae a 0. Esto se ilustra en la figura 6.4. El cálculo no funciona porque el máximo de la probabilidad ocurre en un punto de discontinuidad, pero en la figura se observa que $\hat{\theta} = \text{máx}(X_i)$. Por lo tanto, si los tiempos de espera son 2.3, 3.7, 1.5, 0.4 y 3.2, entonces la emv es $\hat{\theta} = 3.7$.

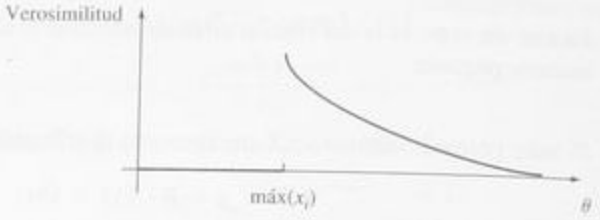


Figura 6.4 Función de probabilidad para el ejemplo 6.21 ■

Ejemplo 6.22 Un método que se utiliza con frecuencia para estimar el tamaño de una población de fauna silvestre tiene que ver con llevar a cabo un experimento de captura y recaptura. En este experimento, se captura una muestra inicial de M animales, se marca cada uno de estos animales y luego se devuelven a la población. Después de un tiempo suficiente para que se mezclen en la población los animales marcados, se captura otra muestra de tamaño n. Con X = número de animales marcados en la segunda muestra, el objetivo es usar la x observada para estimar el tamaño N de la población.

El parámetro de interés es $\theta = N$, que puede tomar sólo valores enteros, así que aun después de determinar la función de probabilidad (en este caso la fmp de X), utilizar el cálculo para obtener N presentaría dificultades. Si se considera que un éxito es capturar de nuevo un animal marcado antes, entonces el muestreo es sin reemplazo a partir de una población que contiene M éxitos y $N - M$ fracasos, de modo que X es una *va* hipergeométrica y la función de verosimilitud es

$$p(x; N) = h(x; n, M, N) = \frac{\binom{M}{x} \cdot \binom{N - M}{n - x}}{\binom{N}{n}}$$

A pesar de la naturaleza de valores enteros de N, sería difícil tomar la derivada de $p(x; N)$. Sin embargo, si se considera la relación de $p(x; N)$ para $p(x; N - 1)$, se tiene

$$\frac{p(x; N)}{p(x; N - 1)} = \frac{(N - M) \cdot (N - n)}{N(N - M - n + x)}$$

Esta relación es más grande que 1 si y sólo si $N < Mn/x$. Por lo tanto, el valor de N para el cual se maximiza $p(x; N)$ el entero más grande menor que Mn/x. Si se emplea notación matemática normal $[r]$ para el entero más grande menor que o igual a r, la emv de N es $\hat{N} = [Mn/x]$.

Como ilustración, si $M = 200$ peces se toman de un lago y se marcan, después se recapturan $n = 100$ peces, y entre los 100 hay $x = 11$ peces marcados, entonces $\hat{N} = [(200)(100)/11] = [1818.18] = 1818$. La estimación es en realidad bastante intuitiva; x/n es la proporción de la muestra recapturada que está marcada, mientras que M/N es la proporción de la población total que está marcada. La estimación se obtiene igualando estas dos proporciones (estimar una proporción poblacional por medio de una proporción muestral). ■

Suponga que $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una fdp $f(x; \theta)$ que es simétrica respecto a θ, pero que el investigador no tiene la seguridad de la forma de la función f. Entonces es deseable usar un estimador $\hat{\theta}$ que es *robusto*; es decir, uno que se desempeña bien para una amplia variedad de fdp subyacentes. Un estimador de este tipo es una media recortada. En años recientes, los estadísticos han propuesto otro tipo de estimador, llamado *estimador M*, con base en una generalización de estimación de máxima verosimilitud. En lugar de maximizar la verosimilitud logarítmica $\sum \ln[f(x; \theta)]$ para una f específica, se maximiza $\sum \rho(x_i; \theta)$. La "función objetivo" ρ se elige para producir un estimador con buenas propiedades de robustez. El libro de David Hoaglin *et al.* (véase la bibliografía) contiene una buena exposición acerca de este tema.

Ejercicios | Sección 6.2 (20 a 30)

20. Se selecciona una muestra aleatoria de n cascos para ciclistas que fabrica cierta compañía. Sea X = número entre los n que son defectuosos y $p = P(\text{defectuosos})$. Suponga que sólo se observa X, en lugar de la secuencia de los éxitos (S) y fracasos (F).

a. Obtenga el estimador de máxima verosimilitud de p. Si $n = 20$ y $x = 3$, ¿cuál es la estimación?

b. ¿Es insesgado el estimador del inciso (a)?

c. Si $n = 20$ y $x = 3$, ¿cuál es la emv de la verosimilitud $(1 - p)^5$ de que ninguno de los siguientes cinco cascos que examinados sean defectuosos?

21. Sea que X tiene una distribución de Weibull con parámetros α y β, de modo que

$$E(X) = \beta \cdot \Gamma(1 + 1/\alpha)$$

$$V(X) = \beta^2\{\Gamma(1 + 2/\alpha) - [\Gamma(1 + 1/\alpha)]^2\}$$

a. Con base en una muestra aleatoria $X_1, \ldots, X_n$, escriba ecuaciones para los estimadores de β y α, del método de momentos. Demuestre que, una vez que se obtuvo la estimación de α la estimación de β se puede encontrar en una tabla de la función gamma, y que la estimación de α es la solución de una ecuación complicada en la cual interviene la función gamma.

b. Si $n = 20$, $\bar{x} = 28.0$ y $\sum x_i^2 = 16\,500$, calcule las estimaciones. [*Sugerencia:* $[\Gamma(1.2)]^2/\Gamma(1.4) = 0.95$.]

22. Sea que X indica la proporción de tiempo asignado que un estudiante seleccionado al azar emplea trabajando en cierta prueba de aptitud. Suponga que la fdp de X es

$$f(x; \theta) = \begin{cases} (\theta + 1)x^\theta & 0 \le x \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

donde $-1 < \theta$. Una muestra aleatoria de diez estudiantes produce los datos $x_1 = 0.92$, $x_2 = 0.79$, $x_3 = 0.90$, $x_4 = 0.65$, $x_5 = 0.86$, $x_6 = 0.47$, $x_7 = 0.73$, $x_8 = 0.97$, $x_9 = 0.94$, $x_{10} = 0.77$.

a. Utilice el método de momentos para obtener un estimador de θ, y después calcule la estimación para estos datos.

b. Obtenga el estimador de máxima verosimilitud de θ, y después calcule la estimación para los datos proporcionados.

23. Se monitorean dos sistemas de computadora diferentes durante un total de n semanas. Represente con X_i el número de interrupciones del primer sistema durante la i-ésima semana, y suponga que las X_i son independientes y extraídas de una distribución de Poisson con parámetro λ_1. De forma similar, represente con Y_i el número de interrupciones del segundo sistema durante la i-ésima semana y suponga independencia en cada Y_i de Poisson con parámetro λ_2. Obtenga las emv de λ_1, λ_2 y $\lambda_1 - \lambda_2$.

[*Sugerencia*: utilice la independencia y escriba la pmf conjunta (probabilidad) de las X_i y las Y_i juntas.]

24. Refiérase al ejercicio 20. En vez de seleccionar $n = 20$ cascos para examinar, suponga se realiza un examen sucesivo de los cascos hasta encontrar $r = 3$ cascos defectuosos. Si el vigésimo casco es el tercero con defectos (de modo que el número de cascos examinados que no tuvieron defectos es $x = 17$), ¿cuál es la emv de p? ¿Ésta es la misma estimación del ejercicio 20? ¿Por qué sí o por qué no? ¿Es la misma que la estimación calculada a partir del estimador insesgado del ejercicio 17?

25. Se determina la resistencia al corte de cada una de diez soldaduras de punto, obteniéndose los datos siguientes (lb/pulg^2):

392 376 401 367 389 362 409 415 358 375

a. Si se supone que la resistencia al corte tiene una distribución normal, estime la resistencia al corte promedio real y su desviación estándar con el método de máxima verosimilitud.

b. Suponga de nuevo una distribución normal y estime el valor de resistencia abajo del cual 95% de las soldaduras tendrán sus resistencias. (*Sugerencia:* ¿cuál es el percentil 95 en términos de μ y σ? Ahora utilice el principio de invarianza).

26. Consulte el ejercicio 25. Suponga que se decide examinar otra soldadura de punto. Sea X = resistencia al corte de la soldadura. Por medio de los datos obtenga la emv de $P(X \le 400)$. [*Sugerencia:* $P(X \le 400) = \Phi((400 - \mu)/\sigma)$.]

27. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución gamma con parámetros α y β.

a. Deduzca las ecuaciones cuya solución produce los estimadores de máxima verosimilitud de α y β. ¿Se pueden resolver de manera explícita?

b. Demuestre que la emv de $\mu = \alpha\beta$ es $\hat{\mu} = \overline{X}$.

28. Sea que $X_1, X_2, \ldots, X_n$ representa una muestra aleatoria de la distribución de Rayleigh con la función de densidad dada en el ejercicio 15. Determine

a. El estimador de máxima verosimilitud de θ y después calcule la estimación para los datos de esfuerzo vibratorio proporcionados en ese ejercicio. ¿Este estimador es el mismo que el insesgado sugerido en el ejercicio 15?

b. La emv de la mediana de la distribución del esfuerzo vibratorio. (*Sugerencia*: primero exprese la mediana en términos de θ.)

29. Considere una muestra aleatoria $X_1, X_2, \ldots, X_n$ de la fdp exponencial desplazada

$$f(x; \lambda, \theta) = \begin{cases} \lambda e^{-\lambda(x-\theta)} & x \ge \theta \\ 0 & \text{en caso contrario} \end{cases}$$

Si se toma $\theta = 0$ se obtiene la fdp de la distribución exponencial considerada antes (con densidad positiva a la derecha de cero). Un ejemplo de la distribución exponencial desplazada apareció en el ejemplo 4.4, en el cual la variable de interés fue el avance del tiempo en el flujo de tránsito y $\theta = 0.5$ el avance del tiempo mínimo posible.

a. Obtenga los estimadores de máxima verosimilitud de θ y λ.

b. Si se hacen $n = 10$ observaciones del avance, que dan como resultado los valores 3.11, 0.64, 2.55, 2.20, 5.44, 3.42, 10.39, 8.93, 17.82 y 1.30, calcule las estimaciones de θ y λ.

30. En el tiempo $t = 0$, 20 se ponen a prueba 20 componentes idénticos. La distribución de tiempo de vida de cada uno es exponencial con parámetro λ. Luego, el experimentador deja sin vigilancia la instalación de prueba. Al regresar, 24 horas después, el experimentador termina de inmediato la prueba después de observar que $y = 15$ de los 20 componentes aún están en operación (por lo tanto, fallaron cinco). Obtenga la emv de λ. [*Sugerencia*: sea Y = número de componentes que resistieron 24 horas. Entonces, $Y \sim \text{Bin}(n, p)$. ¿Cuál es la emv de p? Ahora observe que $p = P(X_i \ge 24)$, donde X_i tiene una distribución exponencial. Esto relaciona λ con p, de modo que la primera se puede estimar una vez estimado la última.]

Ejercicios complementarios (31 a 38)

31. Se dice que un estimador $\hat{\theta}$ es **consistente** si para cualquiera $\epsilon > 0$, $P(|\hat{\theta} - \theta| \ge \epsilon) \to 0$ cuando $n \to \infty$. Es decir, $\hat{\theta}$ es consistente si, a medida que aumenta el tamaño de la muestra, es cada vez menos probable que $\hat{\theta}$ esté más alejada que ϵ del valor verdadero de θ. Demuestre que $\overline{X}$ es un estimador consistente de μ cuando $\sigma^2 < \infty$ por medio de la desigualdad de Chebyshev del ejercicio 43 del capítulo 3. (*Sugerencia*: la desigualdad se puede volver a escribir en la forma

$$P(|Y - \mu_Y| \ge \epsilon) \le \sigma_Y^2/\epsilon$$

Ahora identifique Y con $\overline{X}$.)

32. a. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución uniforme en $[0, \theta]$. Entonces, la emv de θ es $\hat{\theta} = Y = \text{máx}(X_i)$. Utilice el hecho de que $Y \le y$ si y sólo si cada $X_i \le y$ para obtener la fda de Y. Luego, demuestre que la fdp de $Y = \text{máx}(X_i)$ es

$$f_Y(y) = \begin{cases} \dfrac{ny^{n-1}}{\theta^n} & 0 \le y \le \theta \\ 0 & \text{en caso contrario} \end{cases}$$

b. Utilice el resultado del inciso (a) para demostrar que la emv es sesgada, pero que $(n + 1)\text{máx}(X_i)/n$ es insesgada.

33. En el tiempo $t = 0$, hay un individuo vivo en cierta población. Entonces se presenta un **proceso de nacimiento puro**. El tiempo hasta el primer nacimiento tiene una distribución exponencial con parámetro λ. Después del primer nacimiento, hay dos individuos vivos. El tiempo hasta que el primer individuo vuelve a dar a luz es exponencial con parámetro λ, y del mismo modo para el segundo individuo. Por lo tanto, el tiempo hasta el siguiente nacimiento es el mínimo de dos variables exponenciales (λ) el cual es exponencial con parámetro 2λ. De forma similar, una vez que ocurrió el segundo nacimiento hay tres individuos vivos, así que el tiempo hasta el siguiente nacimiento es una *va* exponencial con parámetro 3λ, y así sucesivamente (aquí se emplea la propiedad de amnesia de la distribución exponencial). Suponga que el proceso se observa hasta que ha ocurrido el sexto nacimiento y los tiempos sucesivos de nacimiento son 25.2, 41.7, 51.2, 55.5, 59.5, 61.8 (de los cuales se deben calcular los tiempos entre nacimientos sucesivos). Obtenga la emv de λ. (*Sugerencia*: la verosimilitud es un producto de términos exponenciales.)

34. El **error cuadrático medio** de un estimador $\hat{\theta}$ es $\text{ECM}(\hat{\theta}) = E(\hat{\theta} - \theta)^2$ insesgado, entonces $\text{ECM}(\hat{\theta}) = (\hat{\theta}V)$, pero en general $\text{ECM}(\hat{\theta}) = V(\hat{\theta}) + (\text{sesgo})^2$. Considere el estimador $\hat{\sigma}^2 = KS^2$, donde S^2 = varianza muestral. ¿Qué valor de K reduce al mínimo el error cuadrático medio de este estimador cuando la distribución poblacional es normal? [*Sugerencia*: se puede demostrar que

$$E[(S^2)^2] = (n + 1)\sigma^4/(n - 1)$$

En general, es difícil hallar $\hat{\theta}$ para minimizar $\text{ECM}(\hat{\theta})$, por lo cual sólo se consideran estimadores insesgados y se busca minimizar $V(\hat{\theta})$.]

35. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una fdp simétrica con respecto a μ. Un estimador de μ que funciona bien para varias distribuciones subyacentes es el *estimador de Hodges-Lehmann*. Para definirlo, primero se calcula para cada $i \le j$ y cada $j = 1, 2, \ldots, n$ el promedio por pares $\overline{X}_{i,j} = (X_i + X_j)/2$. Entonces, el estimador es $\hat{\mu}$ = mediana de las $\overline{X}_{i,j}$. Calcule el valor de esta estimación con los datos del ejercicio 44 del capítulo 1. (*Sugerencia:* construya una tabla de cuadrados con las x_i listadas en el margen izquierdo y en la parte superior. Después, calcule promedios en la diagonal y arriba de ésta.)

36. Cuando la distribución poblacional es normal, se puede utilizar la mediana$\{\,|X_1 - \tilde{X}|, \ldots, |X_n - \tilde{X}|\,\}/0.6745$ para estimar σ. Este estimador es más resistente a los efectos de valores atípicos (observaciones lejos de la mayor parte de los datos) que la desviación muestral estándar. Calcule la estimación puntual correspondiente y s para los datos del ejemplo 6.2.

37. Cuando la desviación muestral estándar S se basa en una muestra aleatoria de una distribución poblacional normal, se puede demostrar que

$$E(S) = \sqrt{2/(n-1)}\,\Gamma(n/2)\sigma/\Gamma((n-1)/2)$$

Utilice esta ecuación para obtener un estimador insesgado para σ de la forma cS. ¿Cuánto vale c cuando $n = 20$?

38. Cada uno de n especímenes se va a pesar dos veces en la misma báscula. Sean X_i y Y_i los dos pesos observados para el i-ésimo espécimen. Suponga que X_i y Y_i son independientes entre sí, cada uno distribuido normalmente con valor medio μ_i (el peso verdadero del espécimen i) y varianza σ^2.

a. Demuestre que el estimador de máxima verosimilitud de σ^2 es $\hat{\sigma}^2 = \sum(X_i - Y_i)^2/(4n)$. [*Sugerencia:* si $\overline{z} = (z_1 + z_2)/2$, entonces $\sum(z_i - \overline{z})^2 = (z_1 - z_2)^2/2$.]

b. ¿La emv $\hat{\sigma}^2$ es un estimador insesgado de σ^2? Encuentre un estimador insesgado de σ^2. [*Sugerencia:* para cualquier *va* Z, $E(Z^2) = V(Z) + [E(Z)]^2$. Aplique esto a $Z = X_i - Y_i$.]

Bibliografía

DeGroot, Morris y Mark Schervish, *Probability and Statistics* (3a. ed.), Addison-Wesley, Boston, MA, 2002. Incluye un análisis excelente de propiedades generales y métodos de estimación puntual; de interés especial son

los ejemplos que muestran la forma cómo los principios generales y métodos puedan producir estimadores insatisfactorios en situaciones particulares.

Efron, Bradley y Robert Tibshirani, *An Introduction to the Bootstrap,* Chapman and Hall, Nueva York, 1993. La biblia del bootstrap.

Hoaglin, David, Frederick Mosteller y John Tukey, *Understanding Robust and Exploratory Data Analysis,* Wiley, Nueva York, 1983. Contiene varios buenos capítulos sobre estimación puntual robusta, incluyendo uno sobre estimación *M*.

Hogg, Robert y Allen Craig, *Introduction to Mathematical Statistics* (5a. ed.) Prentice Hall, Englewood Cliffs, NJ, 1995. Un buen estudio sobre el insesgo.

Larsen, Richard y Morris Marx, *Introduction to Mathematical Statistics* (2a. ed.), Prentice Hall, Englewood Cliffs, NJ, 1985. Un estudio muy bueno sobre estimación puntual, desde una perspectiva un poco más matemática que el presente texto.

Rice, John, *Mathematical Statistics and Data Analysis* (2a. ed.), Duxbury Press, Belmont, CA, 1994. Una buena combinación de teoría y datos estadísticos.

7

Intervalos estadísticos basados en una sola muestra

Introducción

Una estimación puntual, debido a que es un solo número, por sí misma no proporciona información acerca de la precisión y confiabilidad de una estimación. Por ejemplo, considere calcular por medio del estadístico $\overline{X}$ una estimación puntual para la resistencia a la ruptura promedio real (g) de toallas de papel de cierta marca, y suponga que $\overline{x} = 9322.7$. Por la variabilidad de muestreo, casi nunca se da el caso de que $\overline{x} = \mu$. La estimación puntual no indica nada en relación con cuán cerca podría estar de μ. Una alternativa para informar un solo valor sensible del parámetro que está siendo estimado es calcular e informar un intervalo completo de valores posibles, una *estimación de intervalo o intervalo de confianza* (IC). Un intervalo de confianza siempre se calcula al seleccionar primero un *nivel de confianza*, que es una medida del grado de confiabilidad del intervalo. Un intervalo de confianza con un nivel de confianza de 95% para la resistencia a la ruptura promedio real podría tener un límite inferior de 9162.5 y un límite superior de 9482.9. Entonces al nivel de confianza de 95%, es posible cualquier valor de μ entre 9162.5 y 9482.9. Un nivel de confianza de 95% significa que 95% de las muestras darían un intervalo que incluye a μ, o cualquier otro parámetro que esté siendo estimado, y sólo 5% de las muestras producirían un intervalo erróneo. Los niveles de confianza que se emplean con más frecuencia son 95%, 99% y 90%. Mientras mayor sea el nivel de confianza, se está más convencido de que el parámetro que está siendo es-

timado se ubica dentro del intervalo (en breve se dará una interpretación de algún nivel de confianza particular).

La información acerca de la precisión de una estimación de intervalo la expresa la extensión del intervalo. Si el nivel de confianza es alto y el intervalo resultante es bastante reducido, el conocimiento del valor del parámetro es razonablemente preciso. Sin embargo, un intervalo de confianza muy amplio, proporciona el mensaje de que hay una gran cantidad de incertidumbre con respecto al valor de lo que se está estimando. En la figura 7.1 se muestran intervalos de confianza de 95% para resistencia a la ruptura promedio real de dos marcas distintas de toallas de papel. Uno de estos intervalos indica un conocimiento preciso acerca de μ, mientras que el otro hace pensar en un intervalo muy amplio de valores posibles.

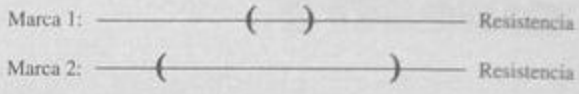


Figura 7.1 Intervalos de confianza que indican información precisa (marca 1) e imprecisa (marca 2) acerca de μ.

7.1 Propiedades básicas de intervalos de confianza

Los conceptos básicos y propiedades de los intervalos de confianza (IC) se introducen con más facilidad si primero se centra la atención en una situación problemática simple, aunque un poco irreal. Suponga que el parámetro de interés es una media poblacional μ y que

1. La distribución de la población es normal.
2. Se conoce el valor de la desviación estándar poblacional σ.

La normalidad de la distribución poblacional suele ser una suposición razonable. Sin embargo, si se ignora el valor de μ no es posible que esté disponible el valor de σ (el conocimiento del centro de una población por lo común precede a la información relacionada con la dispersión). En secciones posteriores se desarrollan métodos basados en suposiciones menos restrictivas.

Ejemplo 7.1 Los ingenieros industriales que se especializan en la ergonomía se interesan en diseñar el espacio de trabajo y dispositivos que operan los trabajadores para lograr alta productividad y confort. En el artículo "Studies on Ergonomically Designed Alphanumeric Keyboards" (*Human Factors*, 1985: 175-187) se proporcionan informes acerca de un estudio de la altura preferida para un teclado experimental con apoyo grande para el antebrazo-muñeca. Se eligió una muestra de $n = 31$ mecanógrafas capacitadas y se determinó la altura de teclado preferida para cada mecanógrafa. La altura preferida promedio resultante de la muestra fue $\bar{x} = 80.0$ cm.

Suponiendo que la altura preferida tiene una distribución normal con $\sigma = 2.0$ cm (un valor que indican los datos del artículo), obtenga un IC para μ, la altura preferida promedio real para la población de mecanógrafas experimentadas. ■

Se supone que las observaciones muestrales reales $x_1, x_2, \ldots, x_n$ son el resultado de una muestra aleatoria $X_1, \ldots, X_n$ de una distribución normal con media μ y desviación estándar σ. Entonces, los resultados del capítulo 5 indican que sin tomar en cuenta el tamaño de muestra n, la media muestral $\overline{X}$ tiene una distribución normal con valor esperado μ y desviación estándar $\sigma/\sqrt{n}$. La estandarización de $\overline{X}$ restando primero su valor esperado y dividiendo después entre su desviación estándar produce la variable

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \tag{7.1}$$

que tiene una distribución normal estándar. Debido a que el área bajo la curva normal estándar entre −1.96 y 1.96 es 0.95,

$$P\left(-1.96 < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < 1.96\right) = 0.95 \tag{7.2}$$

El paso siguiente en el desarrollo es manipular las desigualdades dentro del paréntesis en (7.2) para que aparezcan en la forma equivalente $l < \mu < u$, donde los puntos extremos l y u tienen que ver con $\overline{X}$ y $\sigma/\sqrt{n}$. Esto se logra a través de la siguiente secuencia de operaciones, cada una produce desigualdades equivalentes a las iniciales:

1. Multiplique por $\sigma/\sqrt{n}$ para obtener

$$-1.96 \cdot \frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

2. Reste $\overline{X}$ de cada término para obtener

$$-\overline{X} - 1.96 \cdot \frac{\sigma}{\sqrt{n}} < -\mu < -\overline{X} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

3. Multiplique por −1 para eliminar el signo menos enfrente de μ (lo cual invierte la dirección de cada desigualdad) para obtener

$$\overline{X} + 1.96 \cdot \frac{\sigma}{\sqrt{n}} > \mu > \overline{X} - 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

es decir,

$$\overline{X} - 1.96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

Debido a que cada conjunto de desigualdades de la secuencia es equivalente al original, la probabilidad relacionada con cada uno es 0.95. En particular,

$$P\left(\overline{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95 \tag{7.3}$$

El evento dentro del paréntesis en (7.3) tiene una apariencia algo nueva; antes, la cantidad aleatoria apareció siempre en medio con constantes en ambos extremos, como en $a \le Y \le b$. En (7.3) la cantidad aleatoria apareció en los dos extremos, mientras que la constante desconocida μ aparece a la mitad. Para interpretar (7.3), considere un **intervalo aleatorio** que tiene un punto extremo izquierdo $\bar{X} - 1.96 \cdot \sigma/\sqrt{n}$ y un punto extremo derecho $\bar{X} + 1.96 \cdot \sigma/\sqrt{n}$, que en notación de intervalo es

$$\left(\bar{X} - 1.96 \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right) \tag{7.4}$$

El intervalo (7.4) es aleatorio porque los dos puntos extremos del intervalo tienen que ver con una variable aleatoria (va). Observe que el intervalo está centrado en la media muestral $\bar{X}$ y se extiende $1.96\sigma/\sqrt{n}$ a cada lado de $\bar{X}$. Así, la extensión del intervalo es $2 \cdot (1.96) \cdot \sigma/\sqrt{n}$, que no es aleatorio; sólo la ubicación del intervalo (su punto medio $\bar{X}$) es aleatorio (figura 7.2). Ahora (7.3) se puede interpretar como "*la probabilidad es 0.95 de que el intervalo aleatorio* (7.4) *incluya o cubra al valor verdadero de* μ". Antes de llevar a cabo cualquier experimento y reunir algún dato, es bastante probable (probabilidad 0.95) que μ se encuentre en el intervalo de la expresión (7.4).

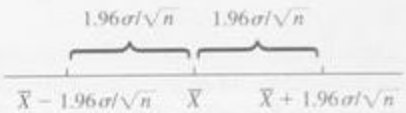


Figura 7.2 Intervalo aleatorio (7.4) centrado en $\bar{X}$

DEFINICIÓN

Si después de observar que $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, se calcula la media muestral observada $\bar{x}$ y luego se sustituye en (7.4) en lugar de $\bar{X}$, el intervalo fijo resultante se llama **intervalo de confianza del 95% para μ.** Este intervalo de confianza de puede expresar como

$$\left(\bar{x} - 1.96 \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right) \quad \text{es un IC de 95\% para } \mu$$

o como

$$\bar{x} - 1.96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96 \cdot \frac{\sigma}{\sqrt{n}} \quad \text{con confianza de 95\%}$$

Una expresion concisa para el intervalo es $\bar{x} \pm 1.96 \cdot \sigma/\sqrt{n}$, donde el signo $-$ proporciona el punto extremo izquierdo (límite inferior) y el signo $+$ da el punto extremo derecho (límite superior).

Ejemplo 7.2 (Continuación del ejemplo 7.1)

Las cantidades necesarias para el cálculo del IC de 95% para la altura preferida promedio real son $\sigma = 2.0$, $n = 31$, y $\bar{x} = 80.0$. El intervalo resultante es

$$\bar{x} \pm 1.96 \cdot \frac{\sigma}{\sqrt{n}} = 80.0 \pm (1.96)\frac{2.0}{\sqrt{31}} = 80.0 \pm 0.7 = (79.3, 80.7)$$

Es decir, se puede estar bastante confiado, al nivel de confianza de 95%, de que $79.3 < \mu < 80.7$. Este intervalo es relativamente estrecho, lo cual indica que se estimó con bastante precisión el valor de μ. ■

Cómo interpretar un intervalo de confianza

El nivel de confianza de 95% para el intervalo recién identificado se heredó de la probabilidad 0.95 para el intervalo aleatorio (7.4). Los intervalos que tienen otros niveles de confianza se introducirán en breve. Por ahora, considere cómo se puede interpretar una confianza de 95%.

Debido a que se comenzó con un evento cuya probabilidad fue 0.95 —de que el intervalo aleatorio (7.4) contendría al valor verdadero de μ— y luego se emplearon los datos del ejemplo 7.1 para calcular el intervalo fijo (79.3, 80.7) es tentador concluir que μ está dentro de este intervalo fijo con probabilidad 0.95. Pero al sustituir $\bar{x} = 80.0$ por $\bar{X}$, desaparece la aleatoriedad; el intervalo (79.3, 80.7) no es un intervalo aleatorio y μ es una constante (que infortunadamente se desconoce), así que es *incorrecto* escribir la proposición $P(\mu$ se ubica en $(79.3, 80.7)) = 0.95$.

Una interpretación correcta de "confianza de 95%" depende de la interpretación de probabilidad de frecuencia relativa a largo plazo: decir que un evento A tiene probabilidad 0.95, equivale a decir que si el experimento en el que está definido A se efectúa una y otra vez, a la larga A ocurrirá 95% de las veces. Suponga que se obtiene otra muestra de las alturas que prefieren las mecanógrafas y se calcula otro intervalo de 95%. Luego, se considera repetir esto para una tercera muestra, una cuarta, etcétera. Sea A el evento en el que $\bar{X} - 1.96 \cdot \sigma/\sqrt{n} < \mu < \bar{X} + 1.96 \cdot \sigma/\sqrt{n}$. Puesto que $P(A) = 0.95$, a la larga 95% de los IC calculados contendrán a μ. Esto se ilustra en la figura 7.3, donde la línea vertical corta el eje de medición en el valor verdadero (pero desconocido) de μ. Observe que de los 11 intervalos ilustrados, sólo los intervalos 3 y 11 no contienen a μ. A la larga, sólo 5% de los intervalos construidos no contendrán a μ.

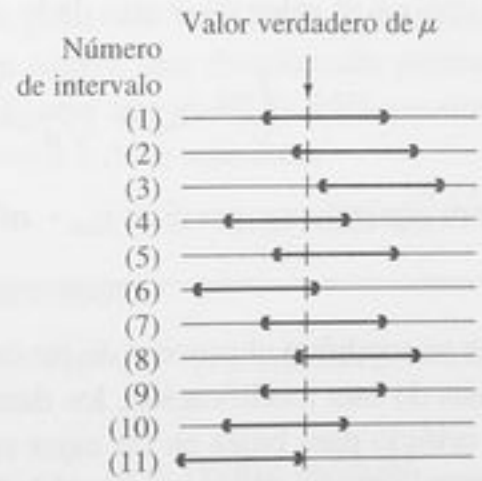


Figura 7.3 Construcción repetida de IC de 95%

De acuerdo con esta interpretación, el nivel de confianza de 95% no es tanto una declaración acerca de algún intervalo particular como (79.3, 80.7), sino que pertenece a lo que sucedería si se construyera un número muy grande de intervalos parecidos. Aunque esto podría parecer insatisfactorio, la raíz del problema radica en la interpretación de probabilidad, se aplica a una larga secuencia de repeticiones de un experimento y no sólo a una repetición. Hay otro método para la construcción e interpretación de IC que utiliza el concepto de probabili-

dad subjetiva y el teorema de Bayes, pero los detalles están más allá del alcance de este texto; el libro de Winkler (véase la bibliografía del capítulo 2) es una buena fuente. El intervalo presentado aquí (así como cada intervalo presentado después) se llama IC "clásico" porque su interpretación se basa en el concepto clásico de probabilidad (aunque las ideas principales se desarrollaron no hace mucho en 1930).

Otros niveles de confianza

El nivel de confianza de 95% se heredó de la probabilidad 0.95 para las desigualdades iniciales en (7.2). Si se desea un nivel de confianza de 99%, se debe remplazar la probabilidad de 0.95 por 0.99, lo cual requiere cambiar el valor crítico z de 1.96 a 2.58. Entonces un IC de 99% resulta de usar 2.58 en lugar de 1.96 en la fórmula para el IC de 95%.

Esto hace pensar que cualquier nivel de confianza deseado se logra remplazando 1.96 o 2.58 con el valor crítico normal estándar apropiado. Como se observa en la figura 7.4, una probabilidad de $1 - \alpha$ se logra si se utiliza $z_{\alpha/2}$ en lugar de 1.96.

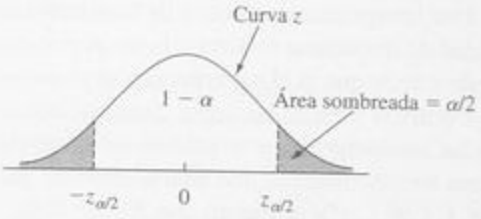


Figura 7.4 $P(-z_{\alpha/2} \le Z < z_{\alpha/2}) = 1 - \alpha$

DEFINICIÓN

Un **intervalo de confianza de $100(1 - \alpha)\%$** para la media μ de una población normal cuando se conoce el valor de σ está dado por

$$\left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) \tag{7.5}$$

o, de manera equivalente, por $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$.

Ejemplo 7.3 Recientemente se modificó el proceso de producción para cajas de control del motor de un tipo particular. Antes de esta modificación, los datos históricos indicaban que la distribución de los diámetros de orificio para bujes en las cajas era normal con una desviación estándar de 0.100 mm. Se cree que la modificación no afectó la forma de la distribución o la desviación estándar, pero que el valor del diámetro medio podría haber cambiado. Se elige una muestra de cajas y se determina el diámetro de orificio para cada una, y el resultado es un diámetro medio muestral de 5.426 mm. Calcúlese un intervalo de confianza para el diámetro de orificio promedio real usando un nivel de confianza de 90%. Esto requiere que $100(1 - \alpha) = 90$, de donde $\alpha = 0.10$ y $z_{\alpha/2} = z_{0.05} = 1.645$ (que corresponde a un área de curva z, acumulada de 0.9500). Entonces el intervalo deseado es

$$5.426 \pm (1.645)\frac{0.100}{\sqrt{40}} = 5.426 \pm 0.026 = (5.400, 5.452)$$

Con un grado de confianza razonablemente alto, se puede decir que $5.400 < \mu < 5.452$. Este intervalo es bastante estrecho por la cantidad pequeña de variabilidad en el diámetro de orificio ($\sigma = 0.100$). ■

Nivel de confianza, precisión y elección del tamaño de muestra

¿Por qué conformarse con un nivel de confianza de 95% cuando se puede lograr un nivel de 99%? Porque el precio que se paga por el nivel de confianza mayor es un intervalo más amplio. Debido a que el intervalo de 95% se extiende $1.96 \cdot \sigma/\sqrt{n}$ a cada lado de $\bar{x}$, la extensión del intervalo es $2(1.96) \cdot \sigma/\sqrt{n} = 3.92 \cdot \sigma/\sqrt{n}$. De manera similar, la extensión del intervalo de 99% es $2(2.58) \cdot \sigma/\sqrt{n} = 5.16 \cdot \sigma/\sqrt{n}$. Es decir, se tiene más confianza en el intervalo de 99% precisamente porque es más amplio. Mientras mayor sea el grado de confianza deseado, más amplio es el intervalo resultante. De hecho, el único IC de 100% para μ es $(-\infty, \infty)$ lo cual no es muy informativo porque, incluso antes del muestreo, se sabe que este intervalo cubre a μ.

Si se considera que la extensión del intervalo especifica su precisión, entonces el nivel de confianza (o confiabilidad) del intervalo tiene una relación inversa con su precisión. Una estimación de intervalo muy confiable podría ser imprecisa en que los puntos finales del intervalo podrían estar muy apartados, en tanto que un intervalo preciso podría conllevar relativamente poca confiabilidad. Por consiguiente, no se puede decir con seguridad que se prefiere un intervalo de 99% a uno de 95%; la ganancia de confiabilidad conlleva una pérdida de precisión.

Una estrategia interesante es especificar tanto el nivel de confianza deseado como la extensión del intervalo y luego determinar el tamaño de muestra necesario.

Ejemplo 7.4 El monitoreo a fondo del sistema de tiempo compartido de computadoras, ha hecho pensar que el tiempo de respuesta a cierto comando de edición tiene una distribución normal con desviación estándar de 25 milisegundos. Se ha instalado un nuevo sistema operativo y se desea estimar el tiempo de respuesta promedio verdadero μ para el nuevo ambiente. Suponiendo que los tiempos de respuesta aún tienen una distribución normal con $\sigma = 25$, ¿qué tamaño de muestra es necesario para asegurar que el IC de 95% resultante tiene una extensión de (a lo sumo) 10? El tamaño de muestra n debe satisfacer

$$10 = 2 \cdot (1.96)(25/\sqrt{n})$$

Al reordenar esta ecuación, se obtiene

$$\sqrt{n} = 2 \cdot (1.96)(25)/10 = 9.80$$

por lo tanto,

$$n = (9.80)^2 = 96.04$$

Puesto que n debe ser un entero, se requiere un tamaño de muestra de 97. ■

La fórmula general para el tamaño de muestra n necesario para asegurar una extensión de intervalo w se obtiene de $w = 2 \cdot z_{\alpha/2} \cdot \sigma/\sqrt{n}$ como

$$n = \left(2z_{\alpha/2} \cdot \frac{\sigma}{w}\right)^2$$

Mientras más pequeña sea la extensión deseada w, el valor de n debe ser más grande. Además, n es una función creciente de σ (una mayor variabilidad poblacional requiere un tamaño de muestra mayor) y del nivel de confianza $100(1 - \alpha)$ (cuando α disminuye, $z_{\alpha/2}$ aumenta).

La extensión media $1.96\sigma/\sqrt{n}$ del IC de 95% a veces se llama **límite sobre el error de estimación** relacionado con un nivel de confianza de 95%; es decir, con confianza de 95%, la estimación puntual $\bar{x}$ no estará más allá de esto con respecto a μ. Antes de obtener los datos, es posible que un investigador desee determinar un tamaño de muestra para el cual se logra un valor particular del límite. Por ejemplo, con μ que representa el rendimiento promedio del combustible (mpg) para los automóviles de cierto tipo, el objetivo de una investigación podría ser estimar μ hasta dentro de 1 mpg con confianza de 95%. En términos generales, si se desea estimar μ hasta dentro de una cantidad B (el límite especificado en el error de estimación) con confianza de $100(1 - \alpha)\%$, el tamaño de muestra necesario resulta de remplazar $2/w$ por $1/B$ en la fórmula de cuadro precedente.

Obtención de un intervalo de confianza

Sea que $X_1, X_2, \ldots, X_n$ representa la muestra en la cual se basará el IC para un parámetro θ. Suponga que se puede hallar una variable aleatoria que satisface las dos propiedades siguientes:

1. La variable depende tanto de $X_1, \ldots, X_n$ como de θ desde el punto de vista de la función.
2. La distribución de probabilidad de la variable no depende de θ o de algún otro parámetro desconocido.

Sea que $h(X_1, X_2, \ldots, X_n; \theta)$ representa esta variable aleatoria. Por ejemplo, si la distribución de probabilidad es normal con σ conocida y $\theta = \mu$, la variable $h(X_1, \ldots, X_n; \mu) = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ satisface ambas propiedades; es claro que depende funcionalmente de μ. Sin embargo, tiene la distribución de probabilidad normal estándar, que no depende de μ. En general, la forma de la función h es evidente al examinar la distribución de un estimador apropiado $\hat{\theta}$.

Para cualquier α entre 0 y 1, se puede hallar constantes a y b que satisfacen

$$P(a < h(X_1, \ldots, X_n; \theta) < b) = 1 - \alpha \tag{7.6}$$

Como resultado de la segunda propiedad, a y b no dependen de θ. En el ejemplo normal, $a = -z_{\alpha/2}$ y $b = z_{\alpha/2}$. Ahora suponga que es posible manipular las desigualdades de (7.6) para aislar θ, de modo que se obtiene la expresión de probabilidad equivalente

$$P(l(X_1, X_2, \ldots, X_n) < \theta < u(X_1, X_2, \ldots, X_n)) = 1 - \alpha$$

Entonces $l(x_1, x_2, \ldots, x_n)$ y $u(x_1, \ldots, x_n)$ son los límites de confianza superior e inferior, respectivamente, para un IC de $100(1 - \alpha)\%$. En el ejemplo con distribucion normal, se vio que $l(X_1, \ldots, X_n) = \bar{X} - z_{\alpha/2} \cdot \sigma/\sqrt{n}$ y $u(X_1, \ldots, X_n) = \bar{X} + z_{\alpha/2} \cdot \sigma/\sqrt{n}$.

Ejemplo 7.5 Un modelo teórico hace pensar que el tiempo hasta que falla un líquido aislante entre electrodos a un determinado voltaje, tiene una distribución exponencial con parámetro λ (véase la sección 4.4). Una muestra de $n = 10$ tiempos de falla produce los siguientes datos muestrales (en minutos): $x_1 = 41.53$, $x_2 = 18.73$, $x_3 = 2.99$, $x_4 = 30.34$, $x_5 = 12.33$, $x_6 = 117.52$, $x_7 = 73.02$, $x_8 = 223.63$, $x_9 = 4.00$, $x_{10} = 26.78$. Se desea un IC de 95% para λ y para el tiempo de falla promedio verdadero.

Sea $h(X_1, X_2, \ldots, X_n; \lambda) = 2\lambda\sum X_i$. Se puede demostrar que esta variable aleatoria tiene una distribución de probabilidad llamada distribución chi-cuadrada con $2n$ grados de libertad (df) ($\nu = 2n$, donde ν es el parámetro de la distribución chi-cuadrada analizada en la sección 4.4). En la tabla A.7 del apéndice se ilustra una curva de densidad chi-cuadrada y se tabulan los valores críticos que capturan áreas extremas específicas. Aquí el número pertinente de grados de libertad es $2(10) = 20$. En el renglón $\nu = 20$ de la tabla se observa que 34.170 capta 0.025 del área extrema superior y 9.591 capta 0.025 del área extrema inferior (0.975 del área extrema superior). Por lo tanto, para $n = 10$,

$$P(9.591 < 2\lambda\sum X_i < 34.170) = 0.95$$

Al dividir entre $2\sum X_i$ se aísla λ, obteniéndose

$$P(9.591/(2\sum X_i) < \lambda < (34.170/(2\sum X_i)) = 0.95$$

El límite inferior del IC de 95% para λ es $9.591/(2\sum x_i)$, y el límite superior es $34.170/(2\sum x_i)$. Para los datos proporcionados, $\sum x_i = 550.87$, así que el intervalo es (0.00871, 0.03101).

El valor esperado de una *va* exponencial es $\mu = 1/\lambda$. Puesto que

$$P(2\sum X_i/34.170 < 1/\lambda < 2\sum X_i/9.591) = 0.95$$

el IC de 95% para el tiempo de falla promedio es $(2\sum x_i/34.170, 2\sum x_i/9.591) = (32.24, 114.87)$. Es evidente que este intervalo es bastante amplio, lo cual refleja variabilidad sustancial en los tiempos de falla y un tamaño de muestra pequeño. ■

En general, los límites de confianza superior e inferior resultan de remplazar cada signo $<$ en (7.6) por $=$ y resolver para θ. En el ejemplo del líquido aislante recién considerado, de $2\lambda\sum x_i = 34.170$ se obtiene que $\lambda = 34.170/(2\sum x_i)$como el límite de confianza superior, y el límite inferior se obtiene de la otra ecuación. Observe que estos dos límites de intervalo no son equidistantes desde la estimación puntual, puesto que el intervalo no es de la forma $\hat{\theta} \pm c$.

Intervalos de confianza de inicialización (Bootstrap)

La técnica de inicialización (bootstrap) se introdujo en el capítulo 6 como una forma de estimar $\sigma_{\hat{\theta}}$. También se puede aplicar para obtener un IC para θ. Considere de nuevo estimar la media μ de una distribución normal cuando se conoce σ. Sustituya a μ por θ y utilice $\hat{\theta} = \overline{X}$ como el estimador puntual. Observe que $1.96\sigma/\sqrt{n}$ es el percentil 97.5 de la distribución de $\hat{\theta} - \theta$ [es decir, $P(\overline{X} - \mu < 1.96\sigma/\sqrt{n}) = P(Z < 1.96) = 0.9750$]. De manera similar, $-1.96\sigma/\sqrt{n}$ es el percentil 2.5, por lo tanto

$$\begin{aligned} 0.95 &= P(\text{percentil 2.5} < \hat{\theta} - \theta < \text{percentil 97.5}) \\ &= P(\hat{\theta} - \text{percentil 2.5} > \theta > \hat{\theta} - \text{percentil 97.5}) \end{aligned}$$

Es decir, con

$$\begin{aligned} l &= \hat{\theta} - \text{percentil 97.5 de } \hat{\theta} - \theta \\ u &= \hat{\theta} - \text{percentil 2.5 de } \hat{\theta} - \theta \end{aligned} \tag{7.7}$$

el IC para θ es (l, u). En muchos casos, no se pueden calcular los percentiles de (7.7), pero se *pueden* estimar a partir de muestras de inicialización. Supóngase que obtiene $B = 1000$ mues-

tras de inicialización y calcule $\hat{\theta}_1^*, \ldots, \hat{\theta}_{1000}^*$ y $\bar{\theta}^*$ y luego las 1000 diferencias $\hat{\theta}_1^* - \bar{\theta}^*, \ldots, \hat{\theta}_{1000}^* - \bar{\theta}^*$. De estas diferencias la 25 más grande y la 25 más pequeña son estimaciones de los percentiles desconocidos en (7.7). Para obtener más información consulte los libros de Rice o Efron citados en el capítulo 6.

Ejercicios | Sección 7.1 (1 a 11)

1. Considere una distribución poblacional normal con el valor de σ conocido.
   a. ¿Cuál es el nivel de confianza para el intervalo $\bar{x} \pm 2.81\sigma/\sqrt{n}$?
   b. ¿Cuál es el nivel de confianza para el intervalo $\bar{x} \pm 1.44\sigma/\sqrt{n}$?
   c. ¿Qué valor de $z_{\alpha/2}$ en la fórmula (7.5) del IC da como resultado un nivel de confianza de 99.7%?
   d. Responda la pregunta del inciso (c) para un nivel de confianza de 75%.

2. Cada uno de los siguientes es un intervalo de confianza de μ = promedio real (es decir, media poblacional) de la frecuencia de resonancia (Hz) para las raquetas de tenis de cierto tipo:

   (114.4, 115.6) (114.1, 115.9)

   a. ¿Cuál es el valor de la frecuencia de resonancia media muestral?
   b. Ambos intervalos se calcularon a partir de los mismos datos de muestra. El nivel de confianza de uno de esos intervalos es 90% y para el otro 99%. ¿Cuál de los intervalos tiene el nivel de confianza de 90% y por qué?

3. Suponga que se selecciona una muestra aleatoria de 50 botellas de una marca de jarabe para la tos y se determina el contenido de alcohol de cada frasco. Sea μ el contenido promedio de alcohol en la población de los frascos de la marca bajo estudio. Suponga que el intervalo de confianza de 95% resultante es (7.8, 9.4).
   a. Un intervalo de confianza de 90% calculado a partir de esta misma muestra, ¿habría sido más reducido o más amplio que el intervalo proporcionado? Explique su razonamiento.
   b. Considere la siguiente afirmación: hay una probabilidad de 95% de que μ esté entre 7.8 y 9.4. ¿Es correcta esa afirmación? ¿Por qué sí o por qué no?
   c. Considere la siguiente afirmación: se puede estar bastante confiado de que 95% de los frascos de esta marca de jarabe para la tos tienen un contenido de alcohol que está entre 7.8 y 9.4. ¿Es correcta esta afirmación? ¿Por qué sí o por qué no?
   d. Tome en cuenta la siguiente afirmación: si el proceso de seleccionar una muestra de tamaño 50 y después calcular el intervalo de confianza de 95% correspondiente se repite 100 veces, 95 de los intervalos resultantes incluirán a μ. ¿Es correcta esta afirmación? ¿Por qué sí o por qué no?

4. Se desea un intervalo de confianza para el promedio verdadero de pérdida suplementaria μ (watts) para cierto tipo de motor de inducción, cuando la corriente de línea se mantiene en 10 amperios para una velocidad de 1500 rpm. Suponga que la pérdida suplementaria tiene una distribución normal con $\sigma = 3.0$.
   a. Calcule un IC de 95% para μ cuando $n = 25$ y $\bar{x} = 58.3$.
   b. Calcule un IC de 95% para μ cuando $n = 100$ y $\bar{x} = 58.3$.
   c. Calcule un IC de 99% para μ cuando $n = 100$ y $\bar{x} = 58.3$.
   d. Calcule un IC de 82% para μ cuando $n = 100$ y $\bar{x} = 58.3$.
   e. ¿Cuál debe ser el valor de n si la longitud del intervalo de 99% para μ debe ser 1.0?

5. Suponga que la porosidad al helio (en porcentaje) de muestras de carbón tomadas de cierta veta tienen una distribución normal con una desviación estándar verdadera de 0.75.
   a. Calcule un IC de 95% para la porosidad promedio real de cierta veta si la porosidad promedio para 20 especímenes de la veta fue 4.85.
   b. Calcule un IC de 98% para la porosidad promedio real de otra veta basada en 16 especímenes con una porosidad promedio muestral de 4.56.
   c. ¿Qué tan grande debe ser el tamaño de la muestra si la extensión del intervalo de 95% va a ser de 0.40?
   d. ¿Qué tamaño de muestra es necesario para estimar un verdadero promedio de porosidad dentro de 0.2 con 99% de confianza?

6. Con base en pruebas detalladas, se sabe que el punto de cedencia de determinado tipo de varilla reforzada con acero dulce tiene una distribución normal con $\sigma = 100$.

Se ha modificado un poco la composición de la barra, pero no se cree que la modificación haya afectado la normalidad o el valor de σ.

a. Suponiendo que éste sea el caso, si una muestra de 25 varillas modificadas produjera un punto de cedencia promedio muestral de 8439 lb, calcule un IC de 90% para el punto de cedencia promedio real de la varilla modificada.

b. ¿Cómo modificaría el intervalo del inciso (a) para obtener un nivel de confianza de 92%?

7. ¿En cuánto se debe aumentar el tamaño muestral n si la extensión del IC (7.5) se reduce a la mitad? Si el tamaño muestral se incrementa en un factor de 25, ¿qué efecto tendrá esto en la extensión del intervalo? Justifique sus afirmaciones.

8. Sea $\alpha_1 > 0$, $\alpha_2 > 0$, con $\alpha_1 + \alpha_2 = \alpha$. Entonces

$$P\left(-z_{\alpha_1} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha_2}\right) = 1 - \alpha$$

a. Utilice esta ecuación a fin de obtener una expresión más general para un IC de $100(1 - \alpha)\%$ para μ del cual el intervalo (7.5) es un caso especial.

b. Sea $\alpha = 0.05$ y $\alpha_1 = \alpha/4$, $\alpha_2 = 3\alpha/4$. ¿Esto origina un intervalo más corto o más amplio que el intervalo (7.5)?

9. a. Bajo las mismas condiciones que originan el intervalo (7.5), resulta que, $P[(\overline{X} - \mu)/(\sigma/\sqrt{n}) < 1.645] = 0.95$. Utilice esto a fin de obtener un intervalo unilateral para μ que tenga extensión infinita y proporcione un límite de confianza inferior sobre μ. ¿Cuál es este intervalo para los datos del ejercicio 5(a)?

b. Generalice el resultado del inciso (a) para obtener un límite inferior con nivel de confianza $100(1 - \alpha)\%$.

c. ¿Cuál es un intervalo análogo al del inciso (b) que proporciona un límite superior sobre μ? Calcule este intervalo al 99% para los datos del ejercicio 4(a).

10. Una muestra aleatoria de $n = 15$ bombas de calor de cierto tipo produjo las siguientes observaciones en relación con el tiempo de vida (en años):

2.0	1.3	6.0	1.9	5.1	0.4	1.0	5.3
15.7	0.7	4.8	0.9	12.2	5.3	0.6	

a. Suponga que la distribución del tiempo de vida es exponencial y utilice un argumento similar al del ejemplo 7.5 a fin de obtener un IC de 95% para el tiempo de vida esperado (promedio verdadero).

b. ¿Cómo se debe modificar el intervalo del inciso (a) para alcanzar un nivel de confianza de 99%?

c. ¿Cuál es un IC de 95% para la desviación estándar de la distribución del tiempo de vida? (*Sugerencia*: ¿Cuál es la desviación estándar de una variable aleatoria exponencial?)

11. Considere los siguientes 1000 intervalos de confianza de 95% para μ que un asesor en estadística obtendrá para varios clientes. Suponga que los conjuntos de datos en los que se basan los intervalos se seleccionan de manera independiente entre sí. ¿Cuántos de estos 1000 intervalos espera que capturen el valor correspondiente de μ? ¿Cuál es la probabilidad de que entre 940 y 960 de esos intervalos contengan el valor correspondiente de μ? (*Sugerencia*: sea Y = número entre los 1000 intervalos que contienen μ. ¿Qué clase de variable aleatoria es Y?)

7.2 Intervalos de confianza de muestras grandes para una media y proporción poblacionales

En el IC para μ dado en la sección anterior se supone que la distribución de la población es normal y que se conoce el valor de σ. Ahora se presenta un IC para muestras grandes cuyos valores no requieren estas suposiciones. Después de mostrar cómo se generaliza el argumento que da lugar a este intervalo para producir otros intervalos de muestras grandes, se centra la atención en un intervalo para una proporción p de la población.

Intervalo de muestra grande para μ

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población que tiene una media μ y desviación estándar σ. Siempre que n sea grande, el Teorema del Límite Central (TLC) implica que $\overline{X}$ tiene una distribución aproximadamente normal sin importar la naturaleza de la distribución

poblacional. Entonces se deduce que $Z = (\overline{X} - \mu)/(\sigma/\sqrt{n})$ tiene aproximadamente una distribución normal estándar, de modo que

$$P\left(-z_{\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

Con un argumento similar al de la sección 7.1 se obtiene $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$ como un IC de muestra grande para μ con un nivel de confianza de *aproximadamente* $100(1 - \alpha)\%$. Es decir, cuando n es grande, el IC para μ dado antes sigue siendo válido cualquiera que sea la distribución poblacional, siempre y cuando se inserte el calificador "aproximadamente" enfrente del nivel de confianza.

Una dificultad práctica con este desarrollo es que el cálculo del intervalo requiere el valor de σ, que casi nunca se conocerá. Considere la variable estandarizada

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

en la cual la desviación estándar muestral S sustituye a σ. Antes sólo había aleatoriedad en el numerador de Z (en virtud de $\overline{X}$). Ahora hay aleatoriedad tanto en el numerador como en el denominador, los valores de $\overline{X}$ y S varían de una muestra a otra. Sin embargo, cuando n es grande, usar S en vez de σ agrega muy poca variabilidad extra a Z. De manera más específica, en este caso la nueva Z tiene aproximadamente una distribución normal estándar. El manejo de las desigualdades en un enunciado de probabilidad en el que interviene esta nueva Z produce un intervalo general de muestra grande para μ.

PROPOSICIÓN

Si n es suficientemente grande, la variable estandarizada

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

tiene una distribución aproximadamente normal estándar. Esto implica que

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \tag{7.8}$$

es un **intervalo de confianza de muestra grande para μ** con nivel de confianza de aproximadamente $100(1 - \alpha)\%$. Esta fórmula es válida sin importar la forma de la distribución de la población.

En términos generales, $n > 40$ será suficiente para justificar el uso de este intervalo. Esto resulta un poco más conservador que la regla empírica para el TLC debido a la variabilidad adicional introducida al usar S en lugar de σ.

Ejemplo 7.6 La tensión disruptiva de corriente alterna (CA) de un líquido aislante indica su resistencia dieléctrica. En el artículo "Testing Practices for the AC Breakdown Voltage Testing of Insulation Liquids" (*IEEE Electrical Insulation Magazine*, 1995; 21-26) se dan las siguientes observa-

Ejemplo 7.

ciones muestrales acerca de la tensión disruptiva (kV) de un circuito particular en ciertas condiciones.

62	50	53	57	41	53	55	61	59	64	50	53	64	62	50	68
54	55	57	50	55	50	56	55	46	55	53	54	52	47	47	55
57	48	63	57	57	55	53	59	53	52	50	55	60	50	56	58

Un diagrama de caja de los datos (figura 7.5) muestra una concentración alta en la mitad media de los datos (extensión del cuadro estrecho). Hay un solo valor atípico en el extremo superior, pero en realidad este valor está un poco más cerca de la mediana (55) de lo que está la observación muestral más pequeña.

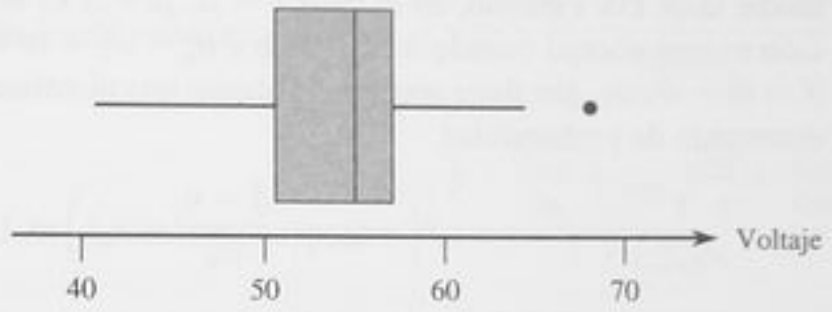


Figura 7.5 Diagrama de caja para los datos de tensión disruptiva del ejemplo 7.6

Las cantidades del resumen son $n = 48, \sum x_i = 2626$, y $\sum x_i^2 = 144\,950$, de las cuales $\bar{x} = 54.7$ y $s = 5.23$. Entonces el intervalo de confianza de 95% es

$$54.7 \pm 1.96\frac{5.23}{\sqrt{48}} = 54.7 \pm 1.5 = (53.2, 56.2)$$

Es decir,

$$53.2 < \mu < 56.2$$

con un nivel de confianza de aproximadamente 95%. El intervalo es razonablemente estrecho, lo cual indica que se estimó con precisión el valor de μ. ■

Infortunadamente, la elección del tamaño de muestra para producir una extensión de intervalo deseada no es tan directa aquí como lo fue para el caso de σ conocida. Esto es porque la extensión de (7.8) es $2z_{\alpha/2}s/\sqrt{n}$. Puesto que el valor de s no está disponible antes de reunir los datos, la extensión del intervalo no se puede determinar sólo mediante la elección de n. La única opción para un investigador que desea especificar una extensión deseada es hacer una suposición con base en la experiencia en cuanto a cuál podría ser el valor de s. Si se es conservador y se supone un valor más grande de s, se elegirá una n más grande de lo necesario. Es posible que el investigador pueda especificar un valor razonablemente preciso del intervalo poblacional (la diferencia entre los valores máximo y mínimo). Entonces, si la distribución de la población no está demasiado sesgada, al dividir el intervalo entre 4 se obtiene un valor aproximado de cuál podría ser s.

Ejemplo 7.7 Refiérase al ejemplo 7.6 acerca de la tensión disruptiva. Suponga que el investigador cree que casi todos los valores de la población están entre 40 y 70. Entonces con $(70 - 40)/4 = 7.5$ se

obtiene un valor razonable para s. El tamaño de muestra apropiado para estimar la tensión disruptiva promedio real hasta dentro de 1 kV con nivel de confianza de 95%, es decir, para que el IC de 95% tenga una extensión de 2 kV, es

$$n = [(1.96)(7.5)/1]^2 \approx 217$$

Intervalo de confianza general de muestra grande

Los intervalos de muestra grandes $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$ y $\bar{x} \pm z_{\alpha/2} \cdot s/\sqrt{n}$ son casos especiales de un IC general de muestra grande para un parámetro θ. Suponga que $\hat{\theta}$ que satisface las propiedades siguientes: (1) tiene aproximadamente una distribución normal; (2) es (por lo menos de forma aproximada) insesgado, y (3) está disponible una expresión para $\sigma_{\hat{\theta}}$, la desviación estándar de $\hat{\theta}$. Por ejemplo, en el caso $\theta = \mu$, $\hat{\mu} = \bar{X}$ es un estimador insesgado cuya distribución es casi normal cuando n es grande y $\sigma_{\hat{\mu}} = \sigma_{\bar{X}} = \sigma/\sqrt{n}$. Al estandarizar $\hat{\theta}$ se obtiene la *va* $Z = (\hat{\theta} - \theta)/\sigma_{\hat{\theta}}$, que tiene aproximadamente una distribución normal estándar. Esto justifica el enunciado de probabilidad

$$P\left(-z_{\alpha/2} < \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}} < z_{\alpha/2}\right) \approx 1 - \alpha \tag{7.9}$$

Suponga, primero, que $\sigma_{\hat{\theta}}$ no requiere parámetros desconocidos (p. ej., σ conocida en el caso $\theta = \mu$). Entonces remplazar cada $<$ en (7.9) por $=$ da como resultado $\theta = \hat{\theta} \pm z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$, así que los límites de confianza inferior y superior son $\hat{\theta} - z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$ y $\hat{\theta} + z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$, respectivamente. Ahora suponga que $\sigma_{\hat{\theta}}$ no tiene relación con θ, pero sí con por lo menos otro parámetro desconocido. Sea $s_{\hat{\theta}}$ la estimación de $\sigma_{\hat{\theta}}$ obtenida al usar estimaciones en lugar de parámetros desconocidos (p. ej., $s/\sqrt{n}$ estima $\sigma/\sqrt{n}$). En condiciones generales (en esencia que $s_{\hat{\theta}}$ esté cerca de $\sigma_{\hat{\theta}}$ para la mayor parte de las muestras), un IC válido es $\hat{\theta} \pm z_{\alpha/2} \cdot s_{\hat{\theta}}$. El intervalo $\bar{x} \pm z_{\alpha/2} \cdot s/\sqrt{n}$ es un ejemplo.

Por último, suponga que $\sigma_{\hat{\theta}}$ requiere el parámetro desconocido θ. Éste es el caso, por ejemplo, cuando $\theta = p$, una proporción poblacional. Entonces puede ser difícil resolver $(\hat{\theta} - \theta)/\sigma_{\hat{\theta}} = z_{\alpha/2}$. Una solución aproximada se obtiene al sustituir θ en $\sigma_{\hat{\theta}}$ por su estimación $\hat{\theta}$. Esto da como resultado una desviación estándar estimada $s_{\hat{\theta}}$, y el intervalo correspondiente es de nuevo $\hat{\theta} \pm z_{\alpha/2} \cdot s_{\hat{\theta}}$.

Intervalo de confianza para una proporción poblacional

Sea p la proporción de "éxitos" en una población, donde *éxito* identifica un individuo u objeto que tiene una propiedad especificada. Se seleccionará una muestra aleatoria de n individuos, y X es el número de éxitos en la muestra. Siempre que n es pequeña en comparación con el tamaño de la población, se puede considerar a X como una *va* binomial con $E(X) = np$ y $\sigma_X = \sqrt{np(1-p)}$. Además, si n es grande ($np \geq 10$ y $nq \geq 10$), X tiene aproximadamente una distribución normal.

El estimador natural de p es $\hat{p} = X/n$, la fracción muestral de éxitos. Puesto que $\hat{p}$ es X multiplicada por la constante $1/n$, $\hat{p}$ también tiene más o menos una distribución normal. Como se muestra en la sección 6.1, $E(\hat{p}) = p$(insesgo) y $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$. La desviación estándar $\sigma_{\hat{p}}$ tiene que ver con el parámetro desconocido p. La estandarización de $\hat{p}$ al restar p y dividir entre $\sigma_{\hat{p}}$ implica que

$$P\left(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

Procediendo como se sugirió en la subsección "Cómo obtener un intervalo de confianza" (sección 7.1), los límites de confianza resultan de remplazar cada $<$ por $=$ y resolver la ecuación cuadrática resultante para p. Con esto se obtienen las dos raíces

$$p = \frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + (z^2_{\alpha/2})/n}$$

PROPOSICIÓN

Un **intervalo de confianza para una proporción poblacional** p con nivel de confianza de aproximadamente $100(1 - \alpha)\%$ tiene

$$\text{límite de confianza inferior} = \frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + (z^2_{\alpha/2})/n}$$

y

$$\text{límite de confianza superior} = \frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + (z^2_{\alpha/2})/n} \qquad (7.10)$$

Si el tamaño de muestra es bastante grande, $z^2/(2n)$ es insignificante en comparación con $\hat{p}$, $z^2/(4n^2)$ bajo la raíz cuadrada es insignificante comparado con $\hat{p}\hat{q}/n$, y z^2/n es insignificante en comparación con 1. Si se descartan estos términos insignificantes, se obtienen límites de confianza aproximados

$$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n} \qquad (7.11)$$

Ésta es la forma general $\hat{\theta} \pm z_{\alpha/2}\,\hat{\sigma}_{\hat{\theta}}$ de un intervalo de muestra grande sugerido en la última subsección. Durante décadas este último intervalo ha sido recomendado siempre y cuando se justifique la aproximación normal para $\hat{p}$. Sin embargo, en investigaciones recientes se ha encontrado que el intervalo un poco más complicado dado en la proposición tiene un nivel de confianza real que tiende a estar más próximo al nivel nominal que el intervalo tradicional (Agresti, Alan y Coull, "Approximate Is Better Than 'Exact' for Interval Estimation of a Binomial Proportion", *The American Statistician*, 1998: 119-126). Es decir, si se utiliza $z_{\alpha/2} = 1.96$ el nivel de confianza para el intervalo "nuevo" tiende a estar más cerca de 95% para casi todos los valores de p que con el intervalo tradicional; esto también se cumple para los niveles de confianza. Además, Agresti y Coull afirman que el intervalo "se puede recomendar para uso con casi todos los tamaños de muestra y valores de parámetro", así que es innecesario comprobar las condiciones $n\hat{p} \geq 10$ y $n\hat{q} \geq 10$.

Ejemplo 7.8 En el artículo "Repeatability and Reproducibility for Pass/Fail Data" (*J. of Testing and Eval.*, 1997: 151-153) se informa que en $n = 48$ ensayos en cierto laboratorio, 16 dieron como resultado la ignición de un tipo particular de sustrato con un cigarro encendido. Sea p la proporción de largo plazo de los ensayos que darían como resultado ignición. Una estimación puntual para p es $\hat{p} = 16/48 = 0.333$. Un intervalo de confianza para p con un nivel de confianza de aproximadamente 95% es

$$\frac{0.333 + (1.96)^2/96 \pm 1.96\sqrt{(0.333)(0.667)/48 + (1.96)^2/9216}}{1 + (1.96)^2/48}$$

$$= \frac{0.373 \pm 0.139}{1.08} = (0.217, 0.474)$$

El intervalo tradicional es

$$0.333 \pm 1.96\sqrt{(0.333)(0.667)/48} = 0.333 \pm 0.133 = (0.200, 0.466)$$

Estos dos intervalos tendrían mucha más coincidencia si el tamaño de muestra fuera sustancialmente más grande. ■

Igualando la extensión del IC para p con una extensión w preespecificada se obtiene una ecuación cuadrática para el tamaño de muestra n necesario para proporcionar un intervalo del grado de precisión deseado. Eliminando el subíndice en $z_{\alpha/2}$, la solución es

$$n = \frac{2z^2\hat{p}\hat{q} - z^2w^2 \pm \sqrt{4z^4\hat{p}\hat{q}(\hat{p}\hat{q} - w^2) + w^2z^4}}{w^2} \tag{7.12}$$

Si se ignoran los términos del numerador en los que interviene w^2, se obtiene

$$n \approx \frac{4z^2\hat{p}\hat{q}}{w^2}$$

Esta última expresión es lo que resulta de igualar la extensión del intervalo tradicional con w.

Infortunadamente, estas fórmulas tienen que ver con la incógnita $\hat{p}$. El método más conservador es aprovechar el hecho de que $\hat{p}\hat{q}$ [$= \hat{p}(1 - \hat{p})$] es un máximo cuando $\hat{p} = 0.5$. Así, si se utiliza $\hat{p} = \hat{q} = 0.5$ en (7.12), la extensión será a lo sumo w sin importar qué valor de $\hat{p}$ resulte de la muestra. En forma alternativa, si el investigador está convencido, como base en información previa, que $p \leq p_0 \leq 0.5$, entonces en lugar de $\hat{p}$ se puede usar p_0. Se aplica un comentario similar cuando $p \geq p_0 \geq 0.5$.

Ejemplo 7.9 La extensión del intervalo de confianza de 95% del ejemplo 7.8 es 0.257. El valor de n necesario para asegurar una extensión de 0.10 sin tomar en cuenta el valor de $\hat{p}$ es

$$n = \frac{2(1.96)^2(.25) - (1.96)^2(0.01) \pm \sqrt{4(1.96)^4(0.25)(0.25 - 0.01) + (0.01)(1.96)^4}}{0.01} = 380.3$$

Así que se debe usar un tamaño de muestra de 381. Con la expresión para n basada en el IC tradicional se obtiene un valor un poco más grande de 385. ■

Intervalos de confianza unilaterales (límites de confianza)

Los intervalos de confianza analizados hasta el momento proporcionan tanto un límite de confianza inferior *como* un límite de confianza superior para el parámetro que está siendo estimado. En algunas circunstancias, un investigador deseará sólo uno de estos dos tipos de límites. Por ejemplo, es posible que un psicólogo desee calcular un límite de confianza superior de 95% para el tiempo de reacción promedio a un estímulo particular, o un ingeniero podría desear sólo un límite de confianza inferior para el tiempo de vida promedio de componentes de cierto tipo. Debido a que el área acumulada bajo la curva normal estándar a la izquierda de 1.645 es 0.95,

$$P\left(\frac{\overline{X} - \mu}{S/\sqrt{n}} < 1.645\right) \approx 0.95$$

Manipulando la desigualdad dentro de los paréntesis para aislar μ en un lado y remplazando las *va* por valores calculados se obtiene la desigualdad $\mu > \bar{x} - 1.645s/\sqrt{n}$; la expresión a la derecha es el límite de confianza inferior deseado. Empezando con $P(-1.645 < Z) \approx 0.95$ y manipulando la desigualdad se obtiene el límite de confianza superior. Un argumento similar proporciona un límite unilateral relacionado con cualquier otro nivel de confianza.

PROPOSICIÓN

Un **límite de confianza superior de muestra grande para μ** es

$$\mu < \bar{x} + z_\alpha \cdot \frac{s}{\sqrt{n}}$$

y un **límite de confianza inferior de muestra grande para μ** es

$$\mu > \bar{x} - z_\alpha \cdot \frac{s}{\sqrt{n}}$$

Un **límite de confianza unilateral para p** resulta de sustituir $z_{\alpha/2}$ por z_α y $\pm$ ya sea por + o − en la fórmula de IC (7.10) para p. En todos los casos el nivel de confianza es aproximadamente $100(1 - \alpha)\%$.

Ejemplo 7.10 La prueba de corte inclinado es el procedimiento más aceptado para evaluar la calidad de una unión entre un material de reparación y su sustrato de concreto. En el artículo "Testing the Bond Between Repair Materials and Concrete Substrate" (*ACI Materials J.*, 1996: 553-558) se informa que en una investigación particular, una muestra de 48 observaciones de resistencia al corte dio una resistencia media muestral de 17.17 N/mm² y una desviación estándar muestral de 3.28 N/mm². Un límite de confianza inferior para la resistencia al corte promedio real μ con nivel de confianza de 95% es

$$17.17 - (1.645)\frac{(3.28)}{\sqrt{48}} = 17.17 - 0.78 = 16.39$$

Es decir, con un nivel de confianza de 95%, el valor de μ se ubica en el intervalo $(16.39, \infty)$. ■

Ejercicios | Sección 7.2 (12 a 27)

12. Una muestra aleatoria de 110 relámpagos en cierta región dio como resultado una duración de eco de radar promedio muestral de 0.81 s y una desviación estándar muestral de 0.34 s ("Lightning Strikes to an Airplane in a Thunderstorm", *J. of Aircraft*, 1984: 607-611). Calcule un intervalo de confianza (bilateral) de 99% para la duración de eco promedio real μ e interprete el intervalo resultante.

13. En el artículo "Extravisual Damage Detection? Defining the Standar Normal Tree" *(Photogrammetric Engr. and Remote Sensing*, 1981: 515-522) se analiza el uso de fotografía infrarroja en color para la identificación de árboles normales en plantaciones de pinos de Oregon. Entre los datos del informe se encuentran estadísticas de resumen para medidas densitométricas ópticas analíticas de filtro verde en muestras de árboles sanos y enfermos. Para una muestra de 69 árboles sanos, la densidad de capa de colorante promedio muestral fue de 1.028 y la desviación estándar muestral fue 0.163.

   a. Calcule un intervalo de confianza de 95% (bilateral) para la densidad de capa de colorante promedio real de todos estos árboles.

   b. Suponga que los investigadores hicieron una suposición aproximada de 0.16 para el valor de s antes de reunir los datos. ¿Qué tamaño de muestra sería necesario a fin de obtener un ancho de intervalo de 0.05 para un nivel de confianza de 95%?

14. En el artículo "Evaluating Tunnel Kiln Performance" *(Amer. Ceramic Soc. Bull.*, agosto de 1997: 59-63) aparece la siguiente información resumida para resistencias a la fractura (MPa) de $n = 169$ barras de cerámica cocidas en determinado horno: $\bar{x} = 89.10$, $s = 3.73$.

a. Calcule un intervalo de confianza (bilateral) para la resistencia a la fractura promedio real usando un nivel de confianza de 95%. ¿Parece que se estimó con precisión la resistencia a la fractura promedio real?

b. Suponga que los investigadores creían, a priori, que la desviación estándar de la población era 4 MPa. Con base en esta suposición, ¿qué tan grande se habría requerido una muestra para estimar μ hasta dentro de 0.5 MPa con confianza de 95%?

15. Determine el nivel de confianza para cada uno de los siguientes límites de confianza unilaterales de muestra grande:

a. Límite superior: $\bar{x} + 0.84s/\sqrt{n}$

b. Límite inferior: $\bar{x} - 2.05s/\sqrt{n}$

c. Límite superior: $\bar{x} + 0.67s/\sqrt{n}$

16. El tiempo entre carga y vaciado (min) para un acero al carbón en un tipo de horno de hogar abierto se determinó para cada hornada en una muestra de tamaño 46. El resultado fue un tiempo promedio de 382.1 y desviación estándar de 31.5. Calcule un límite de confianza superior de 95% para el tiempo promedio real de carga a vaciado.

17. En el ejercicio 1.13 se dio una muestra de observaciones de resistencia a la tensión máxima (klb/pulg^2). Utilice las siguientes estadísticas descriptivas de MINITAB a fin de calcular un límite de confianza inferior de 99% para la resistencia a la tensión máxima promedio real, e interprete el resultado.

N	Mean	Median	TrMean	StDev	SE Mean
153	135.39	135.40	135.41	4.59	0.37

Mínimo	Máximo	Q1	Q3
122.20	147.70	132.95	138.25

18. En el artículo "Ultimate Load Capacities of Expansion Anchor Bolts" (*J. of Energy Engr.*, 1993: 139-158) se dan los siguientes datos de resumen acerca de la resistencia al corte (klb/pulg^2) para una muestra de pernos de anclaje de 3/8 de pulgada: $n = 78$, $\bar{x} = 4.25$, $s = 1.30$. Calcule un límite de confianza inferior utilizando un nivel de confianza de 90% para la resistencia al corte promedio real.

19. En el artículo "Limited Yield Estimation for Visual Defect Sources" (*IEEE Trans. on Semiconductor Manuf.*, 1997: 17-23) se informa que, en un estudio de determinado proceso de inspección de obleas, se examinaron 356 matrices con un probador, de las cuales 201 pasaron el examen. Suponiendo que el proceso es estable, calcule un intervalo de confianza (bilateral) de 95% para la proporción de las matrices que pasan el examen.

20. Associated Press (9 de octubre de 2002) informó que en un estudio de 4722 jóvenes estadounidenses cuyas edades van de los 6 a 19 años, 15% presentaron problemas graves de sobrepeso (un índice de masa corporal de por lo menos 30; este índice es una medida de peso en relación con la estatura). Calcule e interprete un intervalo de confianza para la proporción de los jóvenes estadounidenses que tienen problemas graves de sobrepeso.

21. Se seleccionó una muestra aleatoria de 539 hogares de cierta ciudad de la región central de Estados Unidos, y se determinó que 133 de estos hogares poseían por lo menos un arma de fuego ("The Social Determinants of Gun Ownership: Self-Protection in an Urban Environment", *Criminology*, 1997: 629-640). Con un nivel de confianza de 95%, calcule un límite de confianza inferior para la proporción de casas de esta ciudad que poseen por lo menos un arma de fuego.

22. Se seleccionó una muestra aleatoria de 487 mujeres no fumadoras de peso normal (índice de masa corporal entre 19.8 y 26.0) que dieron a luz en un gran centro médico metropolitano ("The Effects of Cigarette Smoking and Gestational Weight Change on Birth Outcomes in Obese and Normal-Weight Women", *Amer. J. of Public Health*, 1997: 591-596). Se determinó que 7.2% de esos nacimientos dieron como resultado niños de bajo peso al nacer (menos de 2 500 g). Calcule un límite de confianza superior usando un nivel de confianza de 99% para la proporción de nacimientos que dan como resultado niños de bajo peso al nacer.

23. En el artículo "An Evaluation of Football Helmets Under Impact Conditions" *(Amer. J. Sports Medicine*, 1984: 233-237) se informa que cuando se aplicó cierta prueba de impacto a cada casco de fútbol de una muestra aleatoria de 37 cascos del tipo de suspensión, 24 presentaron daño. Sea p la proporción de los cascos de este tipo que presentarían daño al ser probados de la manera prescrita.

a. Calcule un IC de 99% para p.

b. ¿Qué tamaño de muestra se requeriría para que el ancho de un intervalo de confianza de 99% fuera a lo sumo 0.10, sin tomar en cuenta $\hat{p}$?

24. Una muestra de 56 especímenes de algodón para investigación dio como resultado una elongación porcentual promedio muestral de 8.17 y una desviación estándar muestral de 1.42 ("An Apparent Relation Between the Spiral Angle ϕ, the Percent Elongation E_l, and the Dimensions of the Cotton Fiber", *Textile Research J.*, 1978: 407-410). Calcule un intervalo de confianza de muestra grande de 95% para la elongación porcentual promedio real μ. ¿Qué suposiciones se tienen que hacer acerca de la distribución de la elongación porcentual?

25. Una legisladora estatal desea encuestar a los residentes de su distrito para ver qué proporción del electorado está consciente de su postura acerca de usar fondos del estado para pagar abortos.
    a. ¿Qué tamaño de muestra se necesita si el IC de 95% para p debe tener una extensión máxima de 0.10 sin tomar en cuenta p?
    b. Si la legisladora tiene fundamentos para creer que a lo sumo $\frac{2}{3}$ del electorado conoce su posición, ¿qué tamaño de muestra le recomendaría?

26. El superintendente de un gran distrito escolar, que alguna vez cursó probabilidad y estadística, piensa que el número de maestros ausentes en cierto día dado tiene una distribución de Poisson con parámetro λ. Utilice los siguientes datos de ausencias para 50 días, a fin de obtener un intervalo de confianza de muestras grandes para λ. [*Sugerencia*: la media y varianza de una variable de Poisson son iguales a λ, de modo que

$$Z = \frac{\overline{X} - \lambda}{\sqrt{\lambda/n}}$$

tiene aproximadamente una distribución normal estándar. Ahora, proceda como en la obtención del intervalo para p: construya un enunciado de probabilidad (con probabilidad $1 - \alpha$) y despeje una de las desigualdades resultantes (véase el argumento después de (7.10)).]

Número de ausencias	0	1	2	3	4	5	6	7	8	9	10
Frecuencia	1	4	8	10	8	7	5	3	2	1	1

27. Considere de nuevo el IC (7.10) para p y centre su atención en un nivel de confianza de 95%. Demuestre que los límites de confianza concuerdan bastante bien con los del intervalo tradicional (7.11) una vez que se agregan a la muestra dos éxitos y dos fracasos [es decir, la ecuación (7.11), basada en $x + 2$ éxitos (S) en $n + 4$ ensayos]. (*Sugerencia*: $1.96 \approx 2$. *Nota:* Agresti y Coull demostraron que este ajuste del intervalo tradicional también tiene un nivel de confianza real cercano al nivel nominal.)

7.3 Intervalos basados en una distribución poblacional normal

El IC para μ presentado en la sección 7.2 es válido siempre y cuando n sea grande. El intervalo resultante se puede utilizar cualquiera que sea la naturaleza de la distribución poblacional. No obstante, el TLC no se puede invocar cuando n es pequeña. En este caso, una forma de proceder es hacer una suposición específica acerca de la forma de la distribución poblacional y luego obtener un IC adecuado a la suposición. Por ejemplo, se podría desarrollar un IC para μ cuando mediante una distribución gamma se describe la población, otro intervalo para el caso de una población de Weibull, etcétera. De hecho, los estadísticos han llevado a cabo este proceso para distintas familias distributivas. Debido a que la distribución normal es más apropiada como modelo poblacional que cualquier otro tipo de distribución, aquí se pondrá atención en un intervalo de confianza para esta situación.

SUPOSICIÓN

> La población de interés es normal, así que $X_1, \ldots, X_n$ constituye una muestra aleatoria de una distribución normal donde se desconocen μ y σ.

El resultado clave en el que se basa el intervalo de la sección 7.2 fue que para n grande, la *va* $Z = (\overline{X} - \mu)/(S/\sqrt{n})$ tiene aproximadamente una distribución normal estándar. Cuando n es pequeña, S ya no tiene probabilidades de estar cerca de σ, de modo que la variabilidad de la distribución de Z surge de la aleatoriedad tanto en el numerador como en el denominador. Esto significa que la distribución de probabilidad de $(\overline{X} - \mu)/(S/\sqrt{n})$ será más dispersa que la distribución normal estándar. El resultado en el que se basan las inferencias introduce una nueva familia de distribuciones de probabilidad llamada familia de *distribuciones t*.

TEOREMA

Cuando $\overline{X}$ es la media de una muestra aleatoria de tamaño n de una distribución normal con media μ, la *va*

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \tag{7.13}$$

tiene una distribución de probabilidad llamada distribución t con $n - 1$ grados de libertad (gl).

Propiedades de las distribuciones *t*

Antes de aplicar este teorema, es apropiado un análisis de las propiedades de las distribuciones t. Aunque la variable de interés es todavía $(\overline{X} - \mu)/(S/\sqrt{n})$, ahora se denota por T para remarcar que no tiene una distribución normal estándar cuando n es pequeña. Recuerde que dos parámetros gobiernan a la distribución normal, la media μ y la desviación estándar σ. Un solo parámetro gobierna a la distribución t, conocido como el **número de grados de libertad** de la distribución, y se abrevia como gl. Para denotar este parámetro se utiliza la letra griega ν. Los valores posibles de n son los enteros positivos 1, 2, 3, . . . Cada valor de diferente ν corresponde a una distribución de t diferente.

Para cualquier valor fijo del parámetro ν, la función de densidad que especifica la curva t asociada tiene un apariencia aún más complicada que la función de densidad normal. Por fortuna, sólo interesan varias de las características más importantes de estas curvas.

Propiedades de las distribuciónes *t*

Sea que t_ν representa la curva de la función de densidad para ν grados de libertad.

1. Cada curva t_ν es en forma de campana y está centrada en 0.
2. Cada curva t_ν es más dispersa que la curva normal estándar (z).
3. Cuando aumenta ν disminuye la dispersión de la curva t_ν correspondiente.
4. Cuando $\nu \longrightarrow \infty$, la secuencia de curvas t_ν se aproxima a la curva normal estándar (por lo tanto la curva z suele llamarse curva t con grados de libertad $= \infty$).

En la figura 7.6 se ilustran varias de estas propiedades para valores seleccionados de ν.

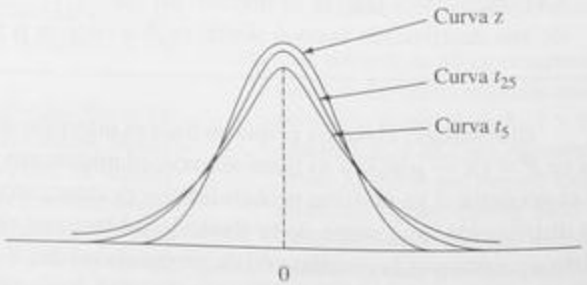


Figura 7.6 Curvas t_ν y z.

El número de gl para T en (7.13) es $n - 1$ porque, aunque S se basa en las n desviaciones $X_1 - \bar{X}, \ldots, \bar{X}_n - \bar{X}, \sum(X_i - \bar{X}) = 0$ implica que sólo $n - 1$ de éstas están "determinadas libremente". El número de gl para una variable t es el número de desviaciones determinadas libremente en las que se basa la desviación estándar estimada del denominador de T.

Puesto que se desea usar T para obtener un IC en el mismo sentido en que se usó antes Z, es necesario establecer notación análoga a z_α para la distribución t.

Notación

Sea $t_{\alpha,\nu}$ = número en el eje de medición para el cual el área bajo la curva t con ν grados de libertad a la derecha de $t_{\alpha,\nu}$ es α; $t_{\alpha,\nu}$ se llama **valor crítico** t.

Esta notación se ilustra en la figura 7.7. En la tabla A.5 del apéndice se da $t_{\alpha,\nu}$ para valores seleccionados de α y ν. Esta tabla también aparece en la tercera de forros. Las columnas de la tabla corresponden a diferentes valores de α. Para obtener $t_{0.05,15}$, vaya a la columna $\alpha = 0.05$, busque el renglón $\nu = 15$ y lea $t_{0.05,15} = 1.735$. De manera similar, $t_{0.05,22} = 1.717$ (columna 0.05, renglón $\nu = 22$), y $t_{0.01,22} = 2.508$.

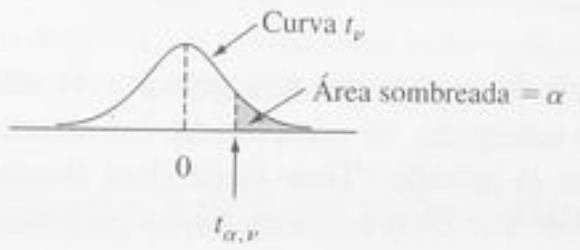


Figura 7.7 Definición gráfica de $t_{\alpha,\nu}$

Los valores de $t_{\alpha,\nu}$ exhiben un comportamiento regular conforme se va por un renglón o columna. Para ν fija, $t_{\alpha,\nu}$ aumenta cuando α disminuye, puesto que es necesario moverse a la derecha de cero para captar el área α del extremo. Para α fija, cuando aumenta ν (es decir, cuando se recorre hacia abajo una determinada columna de la tabla de t) disminuye el valor de $t_{\alpha,\nu}$. Esto es porque un valor más grande de ν implica una distribución t con dispersión más pequeña, así que no es necesario ir tan lejos de cero para captar el área α del extremo. Además, $t_{\alpha,\nu}$ disminuye con más lentitud cuando se incrementa el valor de ν. Por consiguiente, los valores de la tabla se muestran en incrementos de 2 entre 30 gl y 40 gl y luego se salta a $\nu = 50$, 60, 120 y, por último, ∞. Debido a que t_∞ es la curva normal estándar, los valores de z_α comunes aparecen en la última columna de la tabla. La regla empírica sugerida antes para uso del IC de muestra grande (si $n > 40$) viene de la igualdad aproximada de las distribuciones normal estándar y t para $\nu \geq 40$.

Intervalo de confianza t de una sola muestra

La variable estandarizada T tiene una distribución t con $n - 1$ grados de libertad, y el área bajo la curva de densidad t correspondiente entre $-t_{\alpha/2,n-1}$ y $t_{\alpha/2,n-1}$ es $1 - \alpha$ (el área $\alpha/2$ se ubica en cada extremo), por lo tanto

$$P(-t_{\alpha/2,n-1} < T < t_{\alpha/2,n-1}) = 1 - \alpha \tag{7.14}$$

La expresión (7.14) difiere de las expresiones de las secciones anteriores en que T y $t_{\alpha/2,n-1}$ se emplean en lugar de Z y $z_{\alpha/2}$, pero se puede manejar de la misma manera a fin de obtener un intervalo de confianza para μ.

PROPOSICIÓN

Sea $\bar{x}$ y s la media muestral y la desviación estándar muestral calculadas a partir de los resultados de una muestra aleatoria de una población normal con media μ. Entonces un **intervalo de confianza de $100(1-\alpha)\%$ para μ** es

$$\left(\bar{x} - t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}\right) \tag{7.15}$$

o bien, en forma más compacta, $\bar{x} \pm t_{\alpha/2,n-1} \cdot s/\sqrt{n}$.

Un **límite de confianza superior para μ** es

$$\bar{x} + t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}}$$

y al sustituir + por − en esta expresión, se obtiene un **límite de confianza inferior para μ**, con nivel de confianza $100(1-\alpha)\%$.

Ejemplo 7.11 Como parte de un proyecto más grande para estudiar el comportamiento de paneles con recubrimiento esforzado, un componente estructural que se utiliza de manera extensa en Norteamérica, en el artículo "Time-Dependent Bending Properties of Lumber" (*J. of Testing and Eval.*, 1996: 187-193) aparecen varias propiedades mecánicas de especímenes de madera de pino escocés. Considere las observaciones siguientes acerca del módulo de elasticidad (MPa) obtenido un minuto después de cargar en cierta configuración:

10 490	16 620	17 300	15 480	12 970	17 260	13 400	13 900
13 630	13 260	14 370	11 700	15 470	17 840	14 070	14 760

En la figura 7.8 se muestra una gráfica de probabilidad normal obtenida de MINITAB. La rectitud del patrón en la gráfica apoya la suposición de que la distribución poblacional del módulo de elasticidad (MPa) es por lo menos aproximadamente normal.

El cálculo a mano de la media muestral y la desviación estándar se simplificó al restar 10 000 de cada observación: $y_i = x_i - 10\,000$. Se comprueba con facilidad que $\sum y_i = 72\,520$ y $\sum y_i^2 = 392\,083\,800$, de donde $\bar{y} = 4532.5$ y $s_y = 2055.67$. Por consiguiente, $\bar{x} = 14\,532.5$ y $s_x = 2055.67$ (sumar o restar la misma constante de cada observación no afecta la variabilidad). El tamaño de la muestra es 16, así que un intervalo de confianza para la media poblacional del módulo de elasticidad se basa en 15 gl. Un nivel de confianza de 95% para un intervalo bilateral requiere el valor crítico t de 2.131. El intervalo resultante es

$$\bar{x} \pm t_{0.025,15} \cdot \frac{s}{\sqrt{n}} = 14\,532.5 \pm (2.131)\frac{2055.67}{\sqrt{16}}$$

$$= 14\,532.5 \pm 1095.2 = (13\,437.3,\ 15\,627.7)$$

Este intervalo es bastante amplio por el tamaño de muestra pequeño y por la gran cantidad de variabilidad en la muestra. Un límite de confianza inferior de 95% se obtiene al usar − y 1.753 en lugar de ± y 2.131, respectivamente.

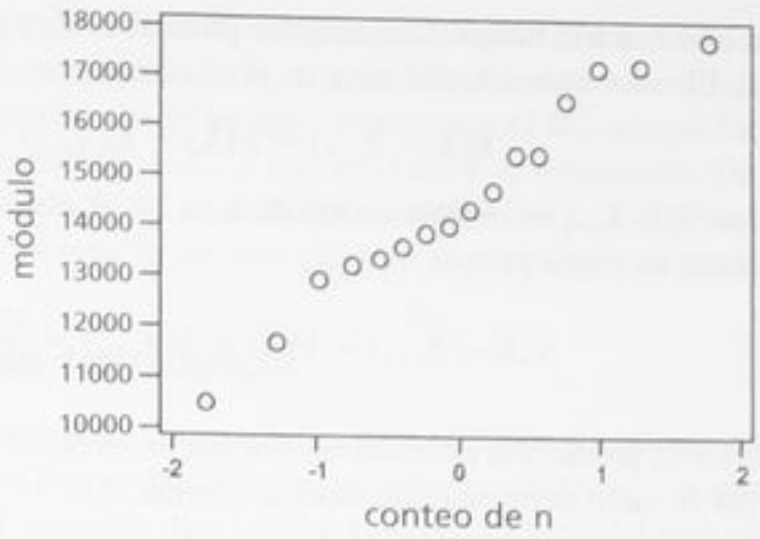


Figura 7.8 Gráfica de probabilidad normal de los datos del módulo de elasticidad ■

Infortunadamente, no es fácil seleccionar n para controlar la extensión del intervalo t. Esto es porque la extensión se relaciona con la s desconocida (antes de reunir los datos) y porque n entra no sólo por $1/\sqrt{n}$, sino también por $t_{\alpha/2,n-1}$. Como resultado, se puede obtener una n apropiada sólo mediante ensayo y error.

En el capítulo 15 se analizará un IC de muestra pequeña para μ que es válido siempre que la distribución poblacional sea simétrica, una suposición más endeble que la de normalidad. Sin embargo, cuando la distribución poblacional es normal, el intervalo t tiende a ser más corto de lo que sería *cualquier* otro intervalo con el mismo nivel de confianza.

Intervalo de predicción para un solo valor futuro

En muchas aplicaciones, un investigador desea *predecir* un solo valor de una variable que será observada en algún tiempo futuro, en vez de *estimar* el valor medio de esa variable.

Ejemplo 7.12 Considere la siguiente muestra de contenido de grasa (en porcentaje) de $n = 10$ *hot dogs* ("Sensory and Mechanical Assessment of the Quality of Frankfurters", *J. Texture Studies*, 1990: 395-409):

25.2 21.3 22.8 17.0 29.8 21.0 25.5 16.0 20.9 19.5

Suponiendo que éstos se seleccionaron de una población con distribución normal, un IC de 95% para (estimación del intervalo de) el contenido de grasa medio poblacional es

$$\bar{x} \pm t_{0.025,9} \cdot \frac{s}{\sqrt{n}} = 21.90 \pm 2.262 \cdot \frac{4.134}{\sqrt{10}} = 21.90 \pm 2.96$$
$$= (18.94, 24.86)$$

No obstante, suponga que una persona se comerá un sólo *hot dog* de este tipo y desea una *predicción* para el contenido de grasa resultante. Una predicción *puntual*, análoga a una estimación *puntual*, es $\bar{x} = 21.90$. Infortunadamente esta predicción no proporciona información acerca de la confiabilidad o precisión. ■

El planteamiento general es como sigue: se tendrá disponible una muestra aleatoria X_1, $X_2, \ldots, X_n$ de una distribución poblacional normal, y se desea predecir el valor de X_{n+1}, a una

sola observación futura. Un predictor puntual es $\overline{X}$, y el error de predicción resultante es $\overline{X} - X_{n+1}$. El valor esperado del error de predicción es

$$E(\overline{X} - X_{n+1}) = E(\overline{X}) - E(X_{n+1}) = \mu - \mu = 0$$

Puesto que X_{n+1} es independiente de $X_1, \ldots, X_n$, es independiente de $\overline{X}$, así que la varianza del error de predicción es

$$V(\overline{X} - X_{n+1}) = V(\overline{X}) + V(X_{n+1}) = \frac{\sigma^2}{n} + \sigma^2 = \sigma^2\left(1 + \frac{1}{n}\right)$$

El error de predicción es una combinación lineal de *va* independientes distribuidas normalmente, por lo tanto tiene una distribución normal. Así,

$$Z = \frac{(\overline{X} - X_{n+1}) - 0}{\sqrt{\sigma^2\left(1 + \frac{1}{n}\right)}} = \frac{\overline{X} - X_{n+1}}{\sqrt{\sigma^2\left(1 + \frac{1}{n}\right)}}$$

tiene una distribución normal estándar. Se puede demostrar que al remplazar σ por la desviación estándar muestral S (de $X_1, \ldots, X_n$), se obtiene

$$T = \frac{\overline{X} - X_{n+1}}{S\sqrt{1 + \frac{1}{n}}} \sim \text{distribución } t \text{ con } n - 1 \text{ grado de libertad}$$

El manejo de esta variable T de la forma como se manejó $T = (\overline{X} - \mu)/(S/\sqrt{n})$ en el desarrollo de un IC proporciona el siguiente resultado.

PROPOSICIÓN

Un **intervalo de predicción** (IP) para una sola observación que se elegirá de una distribución poblacional normal es

$$\bar{x} \pm t_{\alpha/2,n-1} \cdot s\sqrt{1 + \frac{1}{n}} \qquad (7.16)$$

El *nivel de predicción* es $100(1 - \alpha)\%$.

La interpretación de un nivel de predicción de 95% es similar al de un nivel de confianza de 95%; si el intervalo (7.16) se calcula para muestra tras muestra, a la larga 95% de estos intervalos incluirán a los valores futuros correspondientes de X.

Ejemplo 7.13 (Continuación del ejemplo 7.12)

Con $n = 10$, $\bar{x} = 21.90$, $s = 4.134$, y $t_{0.025,9} = 2.262$, un IP de 95% para el contenido de grasa de un solo *hot dog* es

$$21.90 \pm (2.262)(4.134)\sqrt{1 + \frac{1}{10}} = 21.90 \pm 9.81$$
$$= (12.09, 31.71)$$

Este intervalo es bastante ancho, lo cual indica incertidumbre sustancial acerca del contenido de grasa. Observe que el ancho del IP es más de tres veces el del IC. ∎

El error de predicción es $\bar{X} - X_{n+1}$, una diferencia entre dos variables aleatorias, mientras que el error de estimación es $\bar{X} - \mu$, la diferencia entre una variable aleatoria y un valor fijo (pero desconocido). El IP es más amplio que el IC, porque hay más variabilidad en el error de predicción (debido a X_{n+1}) que en el error de estimación. De hecho, cuando n se hace arbitrariamente grande, el IC se reduce a un solo valor μ, y el IP se aproxima a $\mu \pm z_{\alpha/2} \cdot \sigma$. Hay incertidumbre acerca de un solo valor de X aun cuando es innecesario estimar.

Intervalos de tolerancia

Considere una población de automóviles de cierto tipo y suponga que en condiciones específicas el rendimiento de combustible (mpg) tiene una distribución normal con $\mu = 30$ y $\sigma = 2$. Entonces, como el intervalo de -1.645 a 1.645 capta 90% del área bajo la curva z, 90% de estos automóviles tendrán valores de rendimiento de combustible entre $\mu - 1.645\sigma = 26.71$ y $\mu + 1.645\sigma = 33.29$. Pero, ¿qué pasa si se desconocen los valores de μ y σ? Se puede tomar una muestra de tamaño n, determinar los rendimientos de combustible, $\bar{x}$, y s, y formar el intervalo cuyo límite inferior es $\bar{x} - 1.645$ y cuyo límite superior es $\bar{x} + 1.645s$. Sin embargo, debido a la variabilidad del muestreo en las estimaciones de μ y σ, hay una buena probabilidad de que el intervalo resultante incluya menos de 90% de los valores de la población. Desde el punto de vista intuitivo, para tener una posibilidad a priori de 95% de que el intervalo resultante incluya por lo menos 90% de los valores de la población, cuando se usan $\bar{x}$ y s y en lugar de μ y σ, también se debe reemplazar 1.645 por algún número más grande. Por ejemplo, cuando $n = 20$, el valor 2.310 es tal que se puede tener una confianza de 95% de que el intervalo $\bar{x} \pm 2.310s$ incluya por lo menos 90% de los valores de rendimiento de combustible en la población.

Sea k un número entre 0 y 100. Un **intervalo de tolerancia** para captar por lo menos k% de los valores de una distribución poblacional normal con un nivel de confianza de 95% tiene la forma

$$\bar{x} \pm (\text{valor crítico de tolerancia}) \cdot s$$

Los valores críticos de tolerancia para $k = 90$, 95 y 99 en combinación con varios tamaños de muestra se dan en la tabla A.6 del apéndice. Esta tabla también incluye valores críticos para un nivel de confianza de 99% (estos valores son más grandes que los correspondientes valores de 95%. Al reemplazar $\pm$ por $+$ se obtiene un límite de tolerancia superior, y usar $-$ en lugar de $\pm$ origina un límite de tolerancia inferior. Los valores críticos para obtener estos límites unilaterales, aparecen en la tabla A.6 del apéndice.

Ejemplo 7.14 Considere de nuevo los datos del módulo de elasticidad del ejemplo 7.11, donde $n = 16$, $\bar{x} = 14\,532.5$, $s = 2055.67$, y una gráfica de probabilidad normal de los datos, en la cual se indica que fue bastante plausible la normalidad de la población. Para un nivel de confianza de 95%, un intervalo de tolerancia bilateral en el que se capta por lo menos 95% de los valores del módulo de elasticidad para especímenes de madera en la población muestreada utiliza el valor crítico de tolerancia de 2.903. El intervalo resultante es

$$14\,532.5 \pm (2.903)(2055.67) = 14\,532.5 \pm 5967.6 = (8564.9, 20\,500.1)$$

Se puede estar altamente confiado de que por lo menos 95% de los especímenes de madera tienen valores de módulo de elasticidad entre 8 564.9 y 20 500.1.

El IC de 95% para μ fue (13 473.3, 15 627.7), y el intervalo de predicción para el módulo de elasticidad de un solo espécimen de madera es (10 017.0, 19 048.0). Tanto el intervalo de predicción como el intervalo de tolerancia, son mucho más amplios que el intervalo de confianza. ∎

Intervalos basados en distribuciones de poblaciones no normales

El intervalo de confianza de t con una sola muestra para μ es consistente para desviaciones pequeñas o incluso moderadas con respecto a la normalidad, a menos que n sea muy pequeña. Esto significa que si a fin de calcular el intervalo se utiliza un valor crítico para lograr una confianza de 95%, por ejemplo, el nivel de confianza real estará razonablemente cercano al nivel 95% nominal. No obstante, si n es pequeña y la distribución poblacional es bastante anormal, entonces el nivel de confianza real podría ser muy distinto del que se cree se está usando cuando se obtiene un valor crítico particular de la tabla para t. ¡En realidad sería angustiante creer que el nivel de confianza está próximo a 95% cuando de hecho se asemeja más a 88%! La técnica de inicialización (bootstrap), presentada en la sección 7.1, ha resultado bastante exitosa en la estimación de parámetros de una amplia variedad de situaciones no normales.

En contraste con el intervalo de confianza, la validez de los intervalos de predicción y tolerancia descrita en esta sección tiene relación estrecha con la suposición de normalidad. Estos últimos intervalos no se deben usar en ausencia de evidencia convincente de normalidad. En la excelente referencia *Statistical Intervals*, citada en la bibliografía al final de este capítulo, se analizan procedimientos alternativos de este tipo para otras situaciones.

Ejercicios | Sección 7.3 (28 a 41)

28. Determine los valores de las siguientes cantidades:
a. $t_{0.1,15}$ **b.** $t_{0.05,15}$ **c.** $t_{0.05,25}$ **d.** $t_{0.05,40}$ **e.** $t_{0.005,40}$

29. Determine el valor crítico t que contenga el área deseada de la curva t en cada uno de los casos siguientes:
a. Área central = 0.95, gl = 10
b. Área central = 0.95, gl = 20
c. Área central = 0.99, gl = 20
d. Área central = 0.99, gl = 50
e. Área del extremo superior = 0.01, gl = 25
f. Área del extremo inferior = 0.025, gl = 5

30. Determine el valor crítico de t para dos intervalos de confianza bilateral en cada una de las situaciones siguientes.
a. Nivel de confianza = 95%, gl = 10
b. Nivel de confianza = 95%, gl = 15
c. Nivel de confianza = 99%, gl = 15
d. Nivel de confianza = 99%, $n = 5$
e. Nivel de confianza = 98%, gl = 24
f. Nivel de confianza = 99%, $n = 38$

31. Determine el valor crítico de t para un límite de confianza inferior o superior en cada una de las situaciones descritas en el ejercicio 30.

32. Con una muestra aleatoria de $n = 8$ especímenes de prueba de cierto tipo de fibra de vidrio E se obtuvo un límite elástico cortante interfacial promedio muestral de 30.2 y una desviación estándar muestral de 3.1 ("En Interfacial Failure in Notched Unidirectional Glass/Epoxy Composites", *J. of Composite Materials*, 1985: 276-286). Si se supone que el límite elástico cortante interfacial está distribuido normalmente, calcule un intervalo de confianza de 95% para el límite elástico promedio real (como lo hicieron los autores del artículo citado).

33. En el artículo "Measuring and Understanding the Aging of Kraft Insulating Paper in Power Transformers" (*IEEE Electrical Insul. Mag.*, 1996: 28-34) aparecen las observaciones siguientes acerca del grado de polimerización en muestras de papel donde la viscosidad multiplicada por la concentración se ubicó en un cierto intervalo medio.

418 421 421 422 425 427 431
434 437 439 446 447 448 453
454 463 465

a. Trace un diagrama de caja de los datos y comente acerca de algunas características interesantes.

b. ¿Es factible que las observaciones muestrales se eligieran de una distribución normal?

c. Calcule un intervalo de confianza bilateral de 95% para el grado de polimerización promedio real (como lo hicieron los autores del artículo). ¿Al parecer este intervalo indica que 440 es un valor factible del grado de polimerización promedio real? ¿Qué hay acerca de 450?

34. Con una muestra de 14 especímenes de uniones de cierto tipo se obtuvo un esfuerzo límite proporcional medio muestral de 8.48 MPa y desviación estándar muestral de 0.79 MPa ("Characterization of Bearing Strength Factors in Pegged Timber Connections", *J. of Structural Engr.*, 1997: 326-332).

a. Calcule e interprete un límite de confianza inferior de 95% para el esfuerzo límite proporcional promedio real de todas esas uniones. ¿Qué hipótesis, si es que alguna resultó pertinente, planteó acerca de la distribución del esfuerzo límite proporcional?

b. Calcule e interprete un límite inferior de predicción de 95% para el esfuerzo proporcional límite de una sola unión de este tipo.

35. En el ejercicio 46 del capítulo 1 se presentaron las siguientes observaciones muestrales en relación con la viscosidad estabilizada de especímenes de asfalto: 2781, 2900, 3013, 2856, 2888. Una gráfica de probabilidad normal respalda la hipótesis de que la viscosidad tiene por lo menos una distribución normal aproximada.

a. Estime la viscosidad promedio real de una manera que transmita información acerca de la precisión y la confiabilidad.

b. Prediga el valor de la viscosidad para un solo espécimen de asfalto de modo que transmita información acerca de la precisión y la confiabilidad. ¿Cómo se compara la predicción con la estimación calculada en el inciso (a)?

36. Las $n = 26$ observaciones relacionadas con el tiempo de evacuación proporcionadas en el ejercicio 36 del capítulo 1 dan una media y una desviación estándar de 370.69 y 24.36, respectivamente.

a. Calcule un límite de confianza superior para el tiempo promedio de evacuación poblacional, con un nivel de confianza de 95%.

b. Calcule un límite de predicción superior para el tiempo de evacuación de un solo trabajador más, con un nivel de predicción de 95%. ¿Cómo se compara este límite con el límite de confianza del inciso (a)?

c. Suponga que se escogen dos trabajadores más para participar en el ejercicio de evacuación simulada. Represente con X_{27} y X_{28}, los tiempos de evacuación y sea $\overline{X}_{\text{nuevo}}$ el promedio de estos dos valores. Modifique la fórmula del intervalo de predicción para un solo valor de x a fin de obtener el intervalo de predicción para $\overline{X}_{\text{nuevo}}$ y calcule un intervalo bilateral de 95% con base en los datos de tiempos de evacuación.

37. En un estudio de la capacidad de las personas para caminar en línea recta ("Can We Really Walk Straight?" *Amer. J. of Physical Anthro.*, 1992: 19-27) aparecen los datos siguientes acerca de la cadencia (pasos por segundo) para una muestra de $n = 20$ varones saludables elegidos al azar.

0.95 0.85 0.92 0.95 0.93 0.86 1.00 0.92 0.85 0.81
0.78 0.93 0.93 1.05 0.93 1.06 1.06 0.96 0.81 0.96

Una gráfica de probabilidad normal respalda la hipótesis de que la distribución poblacional de la cadencia es aproximadamente normal. A continuación se muestra un resumen descriptivo de los datos obtenidos con MINITAB.

Variable	N	Media	Mediana	Media rec.	DE	EE de la media
cadencia	20	0.9255	0.9300	0.9261	0.0809	0.0181

Variable	Mín	Máx	Q1	Q3
cadencia	0.7800	1.0600	0.8525	0.9600

a. Calcule e interprete un intervalo de confianza de 95% para la cadencia promedio poblacional.

b. Calcule e interprete un intervalo de predicción de 95% para la cadencia de un solo individuo seleccionado al azar de esta población.

c. Calcule un intervalo que comprenda un mínimo de 99% de las cadencias en la distribución poblacional por medio de un nivel de confianza de 95%.

38. Se seleccionó una muestra de 25 piezas de material laminado utilizado en la fabricación de tarjetas de circuito y se determinó la cantidad de alabeo (pulg) para cada pieza bajo ciertas condiciones. El alabeo promedio muestral resultante fue de 0.0635 y la desviación estándar muestral fue de 0.0065.

a. Estime la cantidad de alabeo de una sola pieza de laminado de forma que proporcione información acerca de la precisión y la confiabilidad..

b. Calcule un intervalo con el cual se pueda tener gran confianza de que al menos 95% de las piezas de laminado dan como resultado cantidades de alabeo que están entre los dos límites del intervalo.

39. En el ejercicio 72 del capítulo 1 se proporcionaron las observaciones siguientes acerca de una medida del enlace del receptor (volumen de distribución ajustado) para una muestra de 13 individuos saludables: 23, 39, 40, 41, 43, 47, 51, 58, 63, 66, 67, 69, 72.

a. ¿Es posible que la distribución de la que se eligió esta muestra sea normal?

b. Calcule un intervalo para el cual se pueda tener una confianza de 95% de que por lo menos 95% de los individuos saludables de la población tienen volúme-

nes de distribución ajustados que se ubican entre los límites del intervalo.

c. Dé una predicción del volumen de distribución ajustado de un solo individuo saludable al calcular un intervalo de confianza de 95%. ¿Cómo se compara la extensión de este intervalo con la del intervalo calculado en el inciso (b)?

40. En el ejercicio 13 del capítulo 1 se presentó una muestra de $n = 153$ observaciones en relación con la resistencia a la tensión última o máxima, y en el ejercicio 17 de la sección anterior se proporcionaron cantidades de resumen y se solicitó un intervalo de confianza para muestra grande. Debido a que el tamaño de muestra es grande, son innecesarias las suposiciones acerca de la distribución poblacional para la validez del IC.

a. ¿Se requiere alguna suposición acerca de la distribución de resistencia a la tensión antes de calcular un límite de predicción inferior para la resistencia a la tensión del siguiente espécimen seleccionado mediante el método descrito en esta sección? Explique.

b. Por medio de un programa de estadística investigue la posibilidad de una distribución poblacional normal.

c. Calcule un límite de predicción inferior con un nivel de predicción de 95% para la resistencia a la tensión máxima del siguiente espécimen elegido.

41. Una tabulación más extensa de valores críticos t que la de este libro indica que para la distribución t con 20 gl, las áreas a la derecha de los valores 0.687, 0.860 y 1.064 son 0.25, 0.20 y 0.15, respectivamente. ¿Cuál es el nivel de confianza de cada uno de los siguientes tres intervalos de confianza para la media μ de una población con distribución normal? ¿Cuál de los tres intervalos recomendaría usar y por qué?

a. $(\bar{x} - 0.687s/\sqrt{21}, \bar{x} + 1.725s/\sqrt{21})$

b. $(\bar{x} - 0.860s/\sqrt{21}, \bar{x} + 1.325s/\sqrt{21})$

c. $(\bar{x} - 1.064s/\sqrt{21}, \bar{x} + 1.064s/\sqrt{21})$

7.4 Intervalos de confianza para la varianza y la desviación estándar de una población normal

Aunque las inferencias en relación con una varianza poblacional σ^2 o la desviación estándar son por lo común de menor interés que las de una media o proporción, hay ocasiones en que son necesarios tales procedimientos. En el caso de una distribución poblacional normal, las inferencias se basan en el siguiente resultado que tiene relación con la varianza muestral S^2.

TEOREMA

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con parámetros μ y σ^2. Entonces la *va*

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum(X_i - \bar{X})^2}{\sigma^2}$$

tiene una distribución de probabilidad chi-cuadrada (χ^2) con $n - 1$ grado de libertad.

Como se analizó en las secciones 4.4 y 7.1, la distribución chi-cuadrada es una distribución de probabilidad continua con un solo parámetro ν, llamado número de grados de libertad, con valores posibles 1, 2, 3, ... En la figura 7.9 se ilustran las gráficas de varias distribuciones de probabilidad χ^2 (fdp). Cada fdp $f(x; \nu)$ es positiva sólo para $x > 0$, y cada una tiene sesgo positivo (extremo superior largo), aunque la distribución se mueve a la derecha y se vuelve más simétrica cuando ν aumenta. Para especificar procedimientos inferenciales que utilizan la distribución chi-cuadrada, se necesita una notación análoga a la que se emplea para un valor crítico t, es decir, $t_{\alpha,\nu}$.

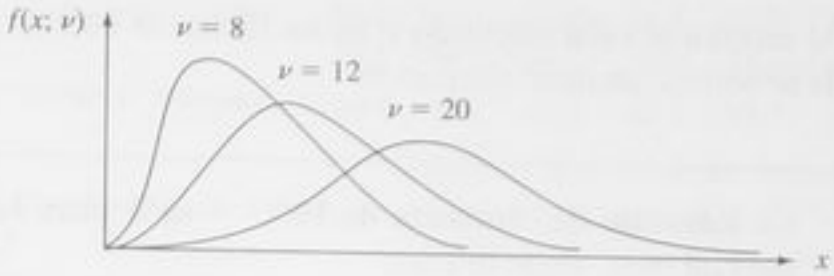


Figura 7.9 Gráficas de funciones de densidad chi-cuadrada

Notación

Sea $\chi^2_{\alpha,\nu}$, llamado **valor crítico chi-cuadrada**, el número en el eje de medición tal que α del área bajo la curva de chi-cuadrada con ν grados de libertad se ubica a la derecha de $\chi^2_{\alpha,\nu}$.

Debido a que la distribución t es simétrica, fue necesario tabular sólo los valores críticos superiores ($t_{\alpha,\nu}$ para valores pequeños de α). La distribución chi-cuadrada no es simétrica, de modo que la tabla A.7 contiene valores de $\chi^2_{\alpha,\nu}$ para α cerca de cero y cerca de 1, como se ilustra en la figura 7.10(b). Por ejemplo, $\chi^2_{0.025,14} = 26.119$ y $\chi^2_{0.95,20}$ (el percentil 5) = 10.851.

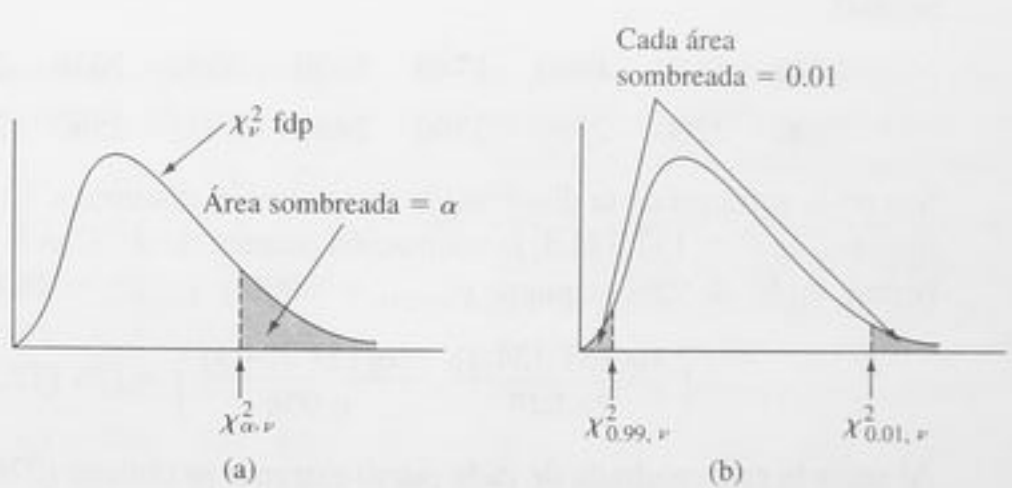


Figura 7.10 Ilustración de la notación $\chi^2_{\alpha,\nu}$

La *va* satisface $(n-1)S^2/\sigma^2$ satisface las dos propiedades en las que se basa el método general para obtener un IC: es una función del parámetro de interés σ^2; sin embargo, su distribución de probabilidad (chi-cuadrada) no depende de este parámetro. El área bajo una curva chi-cuadrada con ν grados de libertad a la derecha de $\chi^2_{\alpha/2,\nu}$ es $\alpha/2$, al igual que el área a la izquierda de $\chi^2_{1-\alpha/2,\nu}$. Así, el área captada entre estos dos valores críticos es $1-\alpha$. Como consecuencia de esto y del teorema antes mencionado,

$$P\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}\right) = 1-\alpha \tag{7.17}$$

Las desigualdades en (7.17) son equivalentes a

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}$$

Al sustituir el valor calculado s^2 en los límites se obtiene un IC para σ^2 y al sacar raíz cuadrada se obtiene un intervalo para σ.

> Un **intervalo de confianza de $100(1-\alpha)\%$ para la varianza σ^2 de una población normal** tiene límite inferior
>
> $$(n-1)s^2/\chi^2_{\alpha/2,n-1}$$
>
> y límite superior
>
> $$(n-1)s^2/\chi^2_{1-\alpha/2,n-1}$$
>
> Un **intervalo de confianza para σ** tiene límites superior e inferior que son raíces cuadradas de los límites correspondientes del intervalo para σ^2.

Ejemplo 7.15 Los datos siguientes acerca de la tensión disruptiva de circuitos forzados eléctricamente se leyeron de una gráfica de probabilidad normal que apareció en el artículo "Damage of Flexible Printed Wiring Boards Associated with Lightning-Induced Voltage Surges" (*IEEE Transactions on Components, Hybrids, and Manuf. Tech.*, 1985: 214-220). La linealidad de la gráfica respaldó la suposición de que la tensión disruptiva tiene una distribución aproximadamente normal.

1470	1510	1690	1740	1900	2000	2030	2100	2190
2200	2290	2380	2390	2480	2500	2580	2700	

Sea σ^2 la varianza de la distribución de la tensión disruptiva. El valor calculado de la varianza muestral es $s^2 = 137\,324.3$, la estimación puntual de σ^2. Con $n - 1 = 16$, a 95% grados de libertad, un IC de 95% requiere $\chi^2_{0.975,16} = 6.908$ y $\chi^2_{0.025,16} = 28.845$. El intervalo es

$$\left(\frac{16(137\,324.3)}{28.845}, \frac{16(137\,324.3)}{6.908}\right) = (76\,172.3, 318\,064.4)$$

Al sacar la raíz cuadrada de cada punto extremo se obtiene (276.0, 564.0) como el IC de 95% para σ. Estos intervalos son bastante amplios, lo que refleja variabilidad sustancial en la tensión disruptiva en combinación con un tamaño de muestra pequeño. ∎

Puede ser difícil obtener los IC para σ^2 y σ cuando la distribución de la población no es normal, incluso cuando el tamaño de muestra es grande. Para esos casos, consulte con un experto en estadística.

Ejercicios | Sección 7.4 (42 a 46)

42. Determine los valores de las siguientes cantidades:

a. $\chi^2_{0.1,15}$ **b.** $\chi^2_{0.1,25}$

c. $\chi^2_{0.01,25}$ **d.** $\chi^2_{0.005,25}$

e. $\chi^2_{0.99,25}$ **f.** $\chi^2_{0.995,25}$

43. Determine lo siguiente:

a. El percentil 95 de la distribución chi-cuadrada $\nu = 10$.

b. El percentil 5 de la distribución chi-cuadrada con $\nu = 10$.

c. $P(10.98 \le \chi^2 \le 36.78)$, donde χ^2 es una *va* chi-cuadrada con $\nu = 22$.

d. $P(\chi^2 < 14.611 \text{ or } \chi^2 > 37.652)$, donde χ^2 a es una *va* chi-cuadrada con $\nu = 25$.

44. Se determinó la cantidad de expansión lateral (milésimas de pulgada) para una muestra de $n = 9$ soldaduras de arco gas-metal activado por impulsos que se utilizan en buques tanque que transportan gas natural licuado. La desviación muestral resultante fue $s = 2.81$ milésimas de pulgada. Suponiendo que se trata de una distribución normal, obtenga un IC de 95% para σ^2 y σ.

45. Las observaciones siguientes se hicieron en relación con la resistencia a la fractura de una placa base de 18% de acero *maragizado* con níquel ["Fracture Testing of Weldments", *ASTM Special Publ. No.* 381, 1965: 328-356 (en klb/pulg^2 $\sqrt{\text{in.}}$, proporcionadas en orden creciente)]:

69.5	71.9	72.6	73.1	73.3	73.5	75.5	75.7
75.8	76.1	76.2	76.2	77.0	77.9	78.1	79.6
79.7	79.9	80.1	82.2	83.7	93.7		

Calcule un IC de 99% para la desviación estándar de la distribución de resistencia a la factura. ¿Es válido este intervalo cualquiera que sea la naturaleza de la distribución? Explique.

46. Los resultados de una prueba de turbiedad de Wagner llevada a cabo en 15 muestras de arena de pruebas Ottawa fueron (en microamperes).

26.7	25.8	24.0	24.9	26.4	25.9	24.4	21.7
24.1	25.9	27.3	26.9	27.3	24.8	23.6	

a. ¿Es posible que esta muestra fuera elegida de una distribución poblacional normal?

b. Calcule un límite de confianza superior con nivel de confianza de 95% para la desviación estándar poblacional de turbiedad.

Ejercicios complementarios (47 a 62)

47. En el ejemplo 1.11 se presentaron las siguientes observaciones en relación con la resistencia de enlace.

11.5	12.1	9.9	9.3	7.8	6.2	6.6	7.0
13.4	17.1	9.3	5.6	5.7	5.4	5.2	5.1
4.9	10.7	15.2	8.5	4.2	4.0	3.9	3.8
3.6	3.4	20.6	25.5	13.8	12.6	13.1	8.9
8.2	10.7	14.2	7.6	5.2	5.5	5.1	5.0
5.2	4.8	4.1	3.8	3.7	3.6	3.6	3.6

a. Estime la resistencia de enlace promedio real de modo que conlleve información acerca de la precisión y la confiabilidad. (*Sugerencia*: $\sum x_i = 387.8$ y $\sum x_i^2 = 4247.08$.)

b. Calcule un intervalo de confianza de 95% para la proporción de enlaces cuyos valores de resistencia serían mayores que 10.

48. Un triatlón que consiste en nadar, andar en bicicleta y correr, es uno de los sucesos deportivos a nivel amateur más extenuantes. En el artículo "Cardiovascular and Thermal Response on Triathlon Performance" (*Medicine and Science in Sports and Exercise*, 1988: 385-389) se informa acerca de un estudio de investigación en el que participaron nueve triatletas varones. Se registró la frecuencia cardiaca máxima (latidos/minuto) durante la actuación en cada uno de los tres sucesos. Para la competencia de nado la media y la desviación estándar muestrales fueron 188.0 y 7.2, respectivamente. Si se supone que la distribución de la frecuencia cardiaca es (aproximadamente) normal, determine un intervalo de confianza de 98% para la frecuencia cardiaca promedio real de los triatletas en la competencia de natación.

49. Para cada uno de 18 núcleos tomados de yacimientos de aceite húmedo de roca carbonatada, se midió la cantidad de saturación de gas residual después de inyectar disolvente e inundar con agua. Las observaciones, en porcentaje de volumen poroso, fueron

23.5	31.5	34.0	46.7	45.6	32.5
41.4	37.2	42.5	46.9	51.5	36.4
44.5	35.7	33.5	39.3	22.0	51.2

(Veáse "Relative Permeability Studies of Gas-Water Flow Following Solvent Injection in Carbonate Rocks", *Soc. Petroleum Engineers J.*, 1976: 23-30).

a. Construya un diagrama de caja para estos datos y comente acerca de algunas características interesantes.

b. ¿Es factible que la muestra haya sido elegida de una distribución poblacional normal?

c. Calcule un intervalo de confianza de 98% para la cantidad promedio real de saturación de gas residual.

50. En un artículo de publicación periódica, se informa que se usó una muestra de tamaño 5 como base para calcular un intervalo de confianza de 95% para la frecuencia natural promedio real (Hz) de vigas deslaminadas de cierto tipo. El intervalo resultante fue (229.764, 233.504). Se decide que un nivel de confianza de 99% es más apropiado que el nivel utilizado de 95%. ¿Cuáles son los límites

del intervalo de 99%? (*Sugerencia*: use el centro del intervalo y su extensión para determinar $\bar{x}$ y s.)

51. El gerente financiero de una gran cadena de tiendas de departamentos, seleccionó una muestra aleatoria de 200 de sus clientes de tarjetas de crédito, y encontró que 136 habían incurrido en cargos por intereses durante el año anterior debido a saldo a pagar.

a. Calcule un intervalo de confianza de 90% para la proporción real de clientes de tarjetas de crédito que incurrieron en cargos por intereses durante el año anterior.

b. Si la extensión deseada del intervalo de 90% es 0.05, ¿qué tamaño de muestra es necesario para asegurar esto?

c. ¿El límite superior del intervalo en el inciso (a) especifica un límite de confianza superior para la proporción que está siendo estimada? Explique.

52. El tiempo de reacción (TR) a un estímulo es el periodo que comienza con la presentación de un estímulo y termina con el primer movimiento discernible de cierto tipo. En el artículo "Relationship of Reaction Time and Movement Time in a Gross Motor Skill" (*Perceptual and Motor Skills*, 1973: 453-454) se informa que el TR promedio muestral de 16 nadadores experimentados a la señal de salida con un disparo fue de 0.214 s y la desviación estándar muestral fue 0.036 s.

a. Establezca las hipótesis necesarias y determine un intervalo de confianza de 90% para el tiempo de reacción promedio real de todos los nadadores experimentados.

b. Calcule un límite de confianza superior de 90% para la desviación estándar de la distribución de tiempos de reacción.

c. Pronostique el tiempo de reacción para otro individuo, de modo que transmita información acerca de la precisión y la confiabilidad.

53. Los pulgones que invaden árboles frutales se controlan mediante aspersión de insecticidas, o al utilizar una gran cantidad de insectos llamados catarinas o mariquitas. En cierta región, se seleccionaron cuatro hileras de árboles frutales para el experimento. Las primeras tres hileras se rociaron con los insecticidas 1, 2, y 3, respectivamente, y la cuarta recibió tratamiento con catarina, con los siguientes resultados:

Tratamiento	n_i = número árboles	$\bar{x}_i$ (grupo/árbol)	s_i
1	100	10.5	1.5
2	90	10.0	1.3
3	100	10.1	1.8
4	120	10.7	1.6

Sea μ_i = la producción promedio real (grupo/árbol) después de recibir el i-ésimo tratamiento. Entonces

$$\theta = \frac{1}{3}(\mu_1 + \mu_2 + \mu_3) - \mu_4$$

mide la diferencia de las producciones promedio reales entre el tratamiento con insecticidas y el tratamiento con mariquitas. Cuando n_1, n_2, n_3 y n_4 son grandes, el estimador $\hat{\theta}$ obtenido al sustituir cada μ_i por $\bar{X}_i$ es aproximadamente normal. Utilice esto para obtener un IC de $100(1 - \alpha)\%$ con muestras grandes para θ y calcule el intervalo de 95% para los datos proporcionados.

54. Es importante que las mascarillas para bomberos resistan altas temperaturas debido a que por lo común trabajan en temperaturas de 200 a 500°F. En una prueba de un tipo de mascarilla, a 11 de 55 se les desprendieron los lentes a 250°. Construya un intervalo de confianza de 90% para proporción real de mascarillas de este tipo cuyos lentes se desprenderían a 250°.

55. Un fabricante de libros de texto para universidades se interesa en estimar la resistencia de la encuadernación llevada a cabo en cierta máquina. La resistencia se puede medir al registrar la fuerza necesaria para arrancar las páginas de la encuadernación. Si esta fuerza se mide en libras, ¿cuántos libros se deben probar para estimar la fuerza promedio requerida para desunir la encuadernación hasta dentro de 0.1 lb con 95% de confianza? Suponga que se sabe que σ es 0.8.

56. Un peligro para la salud bien conocido es la exposición prolongada a fibras de asbesto. En el artículo "The Acute Effects of Chrysotile Asbestos Exposure on Lung Function" (*Environ. Research*, 1978: 360-372) se dan a conocer los resultados de un estudio basado en una muestra de trabajadores de la construcción que estuvieron expuestos al asbesto durante un periodo prolongado. Entre los datos proporcionados en el artículo se encontraron los siguientes valores (ordenados) de distensibilidad pulmonar ($cm^3/cm\ H_2O$) para cada uno de 16 individuos ocho meses después del periodo de exposición (la distensibilidad pulmonar es una medida de la elasticidad de los pulmones, es decir, la eficiencia al inhalar y exhalar):

167.9 180.8 184.8 189.8 194.8 200.2
201.9 206.9 207.2 208.4 226.3 227.7
228.5 232.4 239.8 258.6

a. ¿Es posible que la distribución de la población sea normal?

b. Calcule un IC de 95% para distensibilidad pulmonar promedio real después de la exposición.

c. Calcule un intervalo que, con un nivel de confianza de 95%, incluya por lo menos 95% de los valores de distensibilidad pulmonar en la distribución poblacional.

57. En el ejemplo 6.7 se introdujo el concepto de un experimento censurado donde n componentes se ponen a prueba y el experimento termina tan pronto como fallan r de los componentes. Suponga que los tiempos de vida de los componentes son independientes, cada uno con una distribución exponencial con parámetro λ. Represente con Y_1 el tiempo en el que ocurre la primera falla, Y_2 el tiempo en el que ocurre la segunda falla, y así sucesivamente, de modo que $T_r = Y_1 + \cdots + Y_r + (n - r)Y_r$ es el tiempo de vida acumulado total en la finalización. Entonces se puede demostrar que $2\lambda T_r$ tiene una distribución chi-cuadrada con $2r$ grados de libertad. Utilice este hecho para desarrollar una fórmula del intervalo de confianza de $100(1 - \alpha)\%$ para el tiempo de vida promedio real $1/\lambda$. Calcule un intervalo de confianza de 95% a partir de los datos del ejemplo 6.7.

58. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución de probabilidad continua que tiene mediana $\tilde{\mu}$ (de modo que $P(X_i \le \tilde{\mu}) = P(X_i \ge \tilde{\mu}) = 0.5$).

a. Demuestre que

$$P(\text{mín}(X_i) < \tilde{\mu} < \text{máx}(X_i)) = 1 - \left(\frac{1}{2}\right)^{n-1}$$

de modo (mín(x_i), máx(x_i)) es un intervalo de confianza $100(1 - \alpha)\%$ para $\tilde{\mu}$ con $\alpha = \left(\frac{1}{2}\right)^{n-1}$. [*Sugerencia:* El complemento del suceso $\{\text{mín}(X_i) < \tilde{\mu} < \text{máx}(X_i)\}$ es $\{\text{máx}(X_i) \le \tilde{\mu}\} \cup \{\text{mín}(X_i) \ge \tilde{\mu}\}$. Pero $\text{máx}(X_i) \le \tilde{\mu}$ si y sólo si $X_i \le \tilde{\mu}$ para toda i.]

b. Para cada uno de seis infantes varones normales, se determinó la cantidad del aminoácido alanina (mg/l00 mL) mientras se suministró a los infantes una dieta libre de isoleucina; los datos obtenidos son los siguientes:

2.84 3.54 2.80 1.44 2.94 2.70

Calcule un intervalo de confianza de 97% para la cantidad mediana real de alanina de infantes con esa dieta ("The Essential Amino Acid Requirements of Infants", *Amer. J. Nutrition*, 1964: 322-330).

c. Sea $x_{(2)}$ la segunda x_i más pequeña y $x_{(n-1)}$ la segunda x_i más grande. ¿Cuál es el coeficiente de confianza del intervalo $(x_{(2)}, x_{(n-1)})$ para μ?

59. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución uniforme en el intervalo $[0, \theta]$, de modo que

$$f(x) = \begin{cases} \dfrac{1}{\theta} & 0 \le x \le \theta \\ 0 & \text{en caso contrario} \end{cases}$$

Entonces, si $\lambda = \text{máx}(X_i)$ se puede demostrar que la variable aleatoria $U = Y/\theta$ tiene función de densidad

$$f_U(u) = \begin{cases} nu^{n-1} & 0 \le u \le 1 \\ 0 & \text{en caso contrario} \end{cases}$$

a. Utilice $f_U(u)$ para verificar que

$$P\left((\alpha/2)^{1/n} < \frac{Y}{\theta} \le (1 - \alpha/2)^{1/n}\right) = 1 - \alpha$$

y utilice esta ecuación para obtener un intervalo de confianza de $100(1 - \alpha)\%$ para θ.

b. Compruebe que $P(\alpha^{1/n} \le Y/\theta \le 1) = 1 - \alpha$, y determine un intervalo de confianza de $100(1 - \alpha)\%$ para θ con base en este enunciado de probabilidad.

c. ¿Cuál de los dos intervalos determinados antes es más corto? Si el tiempo que una persona espera el autobús de la mañana tiene una distribución normal y los tiempos de espera observados son $x_1 = 4.2$, $x_2 = 3.5$, $x_3 = 1.7$, $x_4 = 1.2$ y $x_5 = 2.4$, deduzca un intervalo de confianza de 95% para θ usando el más corto de los dos intervalos.

60. Sea $0 \le \gamma \le \alpha$. Entonces un intervalo de confianza de $100(1 - \alpha)\%$ para μ cuando n es grande, es

$$\left(\bar{x} - z_\gamma \cdot \frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha-\gamma} \cdot \frac{s}{\sqrt{n}}\right)$$

La selección $\gamma = \alpha/2$ produce el intervalo usual obtenido en la sección 7.2; si $\gamma \ne \alpha/2$, este intervalo no es simétrico con respecto a $\bar{x}$. La extensión de este intervalo es $w = s(z_\gamma + z_{\alpha-\gamma})/\sqrt{n}$. Demuestre que w se minimiza para cuando se elige $\gamma = \alpha/2$, de modo que el intervalo simétrico es el más corto. [*Sugerencias:* (a) Por definición de z_α, $\Phi(z_\alpha) = 1 - \alpha$, de modo que $z_\alpha = \Phi^{-1}(1 - \alpha)$; (b) la relación entre la derivada de una función $y = f(x)$ y la función inversa $x = f^{-1}(y)$ es (d/dy) $f^{-1}(y) = 1/f'(x)$.]

61. Suponga que $x_1, x_2, \ldots, x_n$ son los valores observados que resultan de una muestra aleatoria de una distribución simétria pero que posiblemente tiene extremos densos. Denótese con $\tilde{x}$ y f_s la mediana muestral y la cuarta dispersión. En el capítulo 11 de *Understanding Robust and Exploratory Data Analysis* (véase la bibliografia del capítulo 6) se indica el siguiente intervalo de confianza robusto al 95% para la media poblacional (punto de simetría):

$$\tilde{x} \pm \left(\frac{\text{valor crítico } t \text{ conservador}}{1.075}\right) \cdot \frac{f_s}{\sqrt{n}}$$

El valor de la cantidad entre paréntesis es 2.10 para $n = 10$, 1.94 para $n = 20$, y 1.91 para $n = 30$. Calcule este intervalo de confianza para los datos del ejercicio 45

y compárelo con el intervalo de confianza t apropiado para una población con distribución normal.

62. **a.** Utilice los resultados del ejemplo 7.5 a fin de obtener un límite de confianza inferior del 95% para el parámetro λ de una distribución exponencial y calcule el límite con base en la información dada en el ejemplo.

b. Si la vida útil X tiene una distribución exponencial, la probabilidad de que pase de t es $P(X > t) = e^{-\lambda t}$. Utilice el resultado del inciso (a) para obtener un límite de confianza inferior de 95% para la probabilidad de que el tiempo de ruptura pase de 100 min.

Bibliografía

DeGroot, Morris y Mark Schervish, *Probability and Statistics* (3a. ed.), Addison-Wesley, Reading, MA, 2002. Una muy buena exposición de los principios generales de inferencia estadística.

Hahn, Gerald y William Meeker, *Statistical Intervals,* Wiley, Nueva York. 1991. Todo lo que el lector siempre ha querido saber acerca de intervalos estadísticos (confianza, predicción, tolerancia y otros).

Larsen, Richard y Morris Marx, *Introduction to Mathematical Statistics* (2a. ed.), Prentice-Hall, Englewood Cliffs, NJ, 1986. Similar a la presentación de DeGroot, pero un poco menos matemático.

8

Pruebas de hipótesis con base en una sola muestra

Introducción

Un parámetro se puede estimar a partir de datos de la muestra, ya sea mediante un solo número (una estimación puntual) o un intervalo completo de valores posibles (un intervalo de confianza). Sin embargo, muchas veces el objetivo de una investigación no es estimar un parámetro, sino decidir cuál de dos afirmaciones contradictorias acerca de un parámetro es correcta. Los métodos para llevar a cabo esto, comprenden la parte de la inferencia estadística llamada *prueba de hipótesis*. En este capítulo, se introducen primero algunos conceptos básicos y terminología de prueba de hipótesis, y luego se desarrollan procedimientos de decisión para la mayor parte de los problemas de prueba comunes con base en una muestra de una sola población.

8.1 Hipótesis y procedimientos de prueba

Una **hipótesis estadística**, o sólo *hipótesis*, es una afirmación acerca del valor de un solo parámetro (característica de la población o característica de una distribución de probabilidad), acerca de los valores de varios parámetros o de la forma de una distribución de probabilidad completa. Un ejemplo de una hipótesis es la afirmación $\mu = 0.75$, donde μ es el diámetro interno promedio real de cierto tipo de tubo de PVC. Otro ejemplo es la expresión $p < 0.10$, donde p es la proporción de tarjetas de circuitos defectuosas entre las tarjetas de circuitos producidas por cierto fabricante. Si μ_1 y μ_2 denotan las resistencias a la ruptura promedio verdaderas de dos tipos de hilo, una hipótesis es la afirmación de que $\mu_1 - \mu_2 = 0$, y otra es la expresión $\mu_1 - \mu_2 > 5$. No obstante, otro ejemplo de una hipótesis es la afirmación de que la distancia de frenado en ciertas condiciones tiene una distribución normal. Las hipótesis de esta última clase se consideran en el capítulo 14. En éste y varios capítulos siguientes, se centra la atención en hipótesis acerca de parámetros.

En cualquier problema de prueba de hipótesis, hay dos hipótesis contradictorias en consideración. Una hipótesis podría ser la afirmación $\mu = 0.75$ y la otra $\mu \neq 0.75$, o bien los dos enunciados contradictorios podrían ser $p \geq 0.10$ y $p < 0.10$. El objetivo es decidir, con base en la información muestral, cuál de las dos hipótesis es correcta. Hay una analogía familiar con esto en un juicio penal. Una afirmación es la aseveración de que el individuo acusado es inocente. En el sistema judicial de Estados Unidos, ésta es la afirmación que al principio se cree cierta. Sólo ante la evidencia contundente de lo contrario, el jurado debe rechazar esta afirmación en favor de la aseveración alternativa de que el acusado es culpable. En este sentido, la afirmación de inocencia es la hipótesis favorecida o protegida, y la carga de demostración se coloca en quienes creen en la afirmación alternativa.

De manera similar, al probar hipótesis estadísticas, el problema se formulará de modo que desde el principio se favorezca una de las afirmaciones. Esta afirmación favorecida al inicio, no se rechazará en favor de la afirmación alternativa a menos que la evidencia muestral la contradiga y proporcione un fuerte respaldo a la aseveración alternativa.

DEFINICIÓN

La **hipótesis nula,** denotada con H_0, es la afirmación que se supone al principio como cierta (la afirmación de "creencia previa"). La **hipótesis alternativa**, denotada por H_a, es la aseveración de que es contradictoria a H_0.

La hipótesis nula se rechaza en favor de la alternativa sólo si la evidencia muestral hace pensar que H_0 es falsa. Si la muestra no contradice de forma contundente a H_0, se continúa con la creencia de que la hipótesis nula es verdadera. Las dos conclusiones posibles de un análisis de prueba de hipótesis son entonces *rechazar* H_0 o *no rechazar* H_0.

Una **prueba de hipótesis** es una forma de usar los datos de la muestra para decidir si se debe rechazar la hipótesis nula. Así, se podría probar H_0: $\mu = 0.75$ contra la alternativa H_a: $\mu \neq 0.75$. Sólo si los datos muestrales hacen pensar sin lugar a dudas que μ es diferente de 0.75, se debe rechazar la hipótesis nula. En ausencia de tal evidencia, no se debe rechazar H_0, puesto que aún es bastante posible.

A veces el investigador no quiere aceptar una aseveración particular, a menos y hasta que los datos respalden con firmeza la aseveración. Por ejemplo, suponga que una compañía está considerando colocar un nuevo tipo de recubrimiento en los cojinetes que produce. La dura-

ción promedio real con el recubrimiento actual es de 1000 horas. Si μ representa la duración promedio real para el nuevo recubrimiento, la compañía se negará a hacer un cambio, a menos que se pruebe que μ pasa de 1000. Una formulación apropiada del problema tendría que ver con probar H_0: $\mu = 1000$ contra H_a: $\mu > 1000$. La conclusión de que se justifica un cambio se identifica con H_a, y llevaría evidencia concluyente para justificar rechazar H_0 y cambiar al nuevo recubrimiento.

En la investigación científica por lo común se requiere tratar de decidir si una explicación plausible y satisfactoria del fenómeno en investigación debe remplazar la teoría actual. Un método conservador es identificar la teoría actual con H_0 y la explicación alternativa del investigador con H_a. Entonces el rechazo de la teoría actual ocurre sólo cuando la evidencia es mucho más consistente con la nueva teoría. En muchas situaciones, H_a se denomina "hipótesis del investigador", puesto que es la afirmación que al investigador le gustaría validar. La palabra *nula* significa "de ningún valor, efecto o consecuencia", lo cual hace pensar que H_0 se debe identificar con la hipótesis de ningún cambio (con respecto a la opinión actual), ninguna diferencia o mejoría, etcétera. Suponga, por ejemplo, que 10% de las tarjetas de circuitos que produjo cierto fabricante durante un periodo reciente resultaron defectuosas. Un ingeniero ha recomendado un cambio en el proceso de producción, con la creencia de que dará como resultado una menor proporción de artículos defectuosos. Sea p la proporción real de tarjetas defectuosas que resultan del proceso modificado. Entonces la hipótesis de investigación, en la que se coloca la carga de la demostración, es la aseveración de que $p < 0.10$. Por lo tanto la hipótesis alternativa es H_a: $p < 0.10$.

En el tratamiento de la prueba de hipótesis, H_0 siempre será expresada como una afirmación de igualdad. Si θ denota el parámetro de interés, la hipótesis nula tendrá la forma H_0: $\theta = \theta_0$, donde θ_0 es un número especificado denominado *valor nulo* del parámetro (valor que se afirma para θ en la hipótesis nula). Por ejemplo, considere la situación de las tarjetas de circuitos recién analizada. La hipótesis alternativa sugerida fue H_a: $p < 0.10$, la afirmación de que la proporción de artículos defectuosos se reduce mediante la modificación del proceso. Una elección natural de H_0 en esta situación es la afirmación de que $p \geq 0.10$, de acuerdo con lo que el nuevo proceso no es capaz de mejorar *o* empeorar en comparación con el actual. En cambio se considerará H_0: $p = 0.10$ contra H_a: $p < 0.10$. El fundamento para usar esta hipótesis nula simplificada, es que cualquier procedimiento de decisión razonable para decidir entre H_0: $p = 0.10$ y H_a: $p < 0.10$ también será razonable para decidir entre la afirmación de que $p \geq 0.10$ y H_a. Se prefiere el uso de una H_0 simplificada porque tiene ciertos beneficios técnicos, que en breve serán evidentes.

La alternativa a la hipótesis nula H_0: $\theta = \theta_0$ se asemejará a una de las tres afirmaciones siguientes: (1) H_a: $\theta > \theta_0$ (en cuyo caso la hipótesis nula implícita es $\theta \leq \theta_0$), (2) H_a: $\theta < \theta_0$ (así que la hipótesis nula implícita establece que $\theta \geq \theta_0$), o bien (3) H_a: $\theta \neq \theta_0$. Por ejemplo, sea que σ representa la desviación estándar de la distribución de los diámetros internos (pulgadas) para cierto tipo de funda metálica. Si se tomó la decisión de usar la funda a menos que la evidencia muestral demuestre de manera concluyente que $\sigma > 0.001$, la hipótesis apropiada sería H_0: $\sigma = 0.001$ contra H_a: $\sigma > 0.001$. El número θ_0 que aparece tanto en H_0 como en H_a (separa a la alternativa de la nula) se llama **valor nulo.**

Procedimientos de prueba

Un procedimiento de prueba es una regla, con base en datos muestrales, para decidir si se rechaza H_0. Una prueba de H_0: $p = 0.10$ contra H_a: $p < 0.10$ en el problema de la tarjeta de circuitos se podría basar en examinar una muestra aleatoria de $n = 200$ tarjetas. Sea X el número de tarjetas defectuosas en la muestra, una variable aleatoria binomial; x representa el valor ob-

servado de X. Si H_0 es cierta, $E(X) = np = 200(0.10) = 20$, mientras que se pueden esperar menos de 20 tarjetas defectuosas si H_a es cierta. Un valor de x un poco abajo de 20 no contradice de manera firme a H_0, así que es razonable rechazar H_0 sólo si x es mucho menor que 20. Un procedimiento de prueba de esta clase es rechazar H_0 si $x \leq 15$ y no rechazar H_0 en caso contrario. Este procedimiento tiene dos constituyentes: 1) una *estadística de prueba* o función de los datos muestrales utilizada para tomar una decisión y 2) una *región de rechazo* que consiste en los valores de x para los cuales H_0 será rechazada en favor de H_a. Para la regla recién sugerida, la región de rechazo consiste en $x = 0, 1, 2, \ldots,$ y 15. H_0 no será rechazada si $x = 16, 17, \ldots, 199,$ o 200.

Un procedimiento de prueba se especifica mediante lo siguiente:

1. Una **estadística de prueba**, una función de los datos muestrales en la que se basará la decisión (rechazar H_0).
2. Una **región de rechazo**, el conjunto de los valores de la estadística de prueba para los que se rechazará H_0.

Entonces, la hipótesis nula será rechazada si y sólo si el valor de la estadística de prueba observada o calculada cae en la región de rechazo.

Como otro ejemplo, suponga que un fabricante de cigarros afirma que el contenido promedio de nicotina μ de los cigarros marca B es (a lo sumo) 1.5 mg. Sería desaconsejable rechazar la afirmación del fabricante sin una prueba contradictoria firme, así que una formulación apropiada del problema es probar H_0: $\mu = 1.5$ contra H_a: $\mu > 1.5$. Considere una regla de decisión basada en analizar una muestra aleatoria de 32 cigarros. Represente con $\overline{X}$ el contenido de nicotina promedio muestral. Si H_0 es verdadera, $E(\overline{X}) = \mu = 1.5$, mientras que si H_0 es falsa, se espera que $\overline{X}$ pase de 1.5. La evidencia firme contra H_0 la proporciona un valor de $\overline{x}$ que sea considerablemente mayor que 1.5. Así, se podría usar $\overline{X}$ como una estadística de prueba junto con la región de rechazo $\overline{x} \geq 1.6$.

Tanto en el ejemplo de las tarjetas de circuitos como en el de la nicotina, la elección de la estadística de prueba y de la región de rechazo tienen sentido desde el punto de vista intuitivo. Sin embargo, la elección del valor de corte utilizado para especificar la región de rechazo es un poco arbitraria. En lugar de rechazar H_0: $p = 0.10$ en favor de H_a: $p < 0.10$ cuando $x \leq 15$, se podría usar la región de rechazo $x \leq 14$. Para esta región, H_0 no sería rechazada si se observan 15 tarjetas defectuosas, en tanto que esta ocurrencia daría lugar al rechazo de H_0 si se emplea la región sugerida al principio. De manera similar, la región de rechazo $\overline{x} \geq 1.55$ se podría usar en el problema de la nicotina en lugar de la región $\overline{x} \geq 1.60$.

Errores en la prueba de hipótesis

La base para elegir una determinada región de rechazo radica en entender los errores que uno podría enfrentar al sacar una conclusión. Considere la región de rechazo $x \leq 15$ en el problema de tarjetas de circuitos. Aun cuando H_0: $p = 0.10$ es cierta podría suceder que una muestra inusual diera como resultado $x = 13$, así que es un error rechazar H_0. Por otro lado, aunque H_a: $p < 0.10$ es verdadero, una muestra inusual podría producir $x = 20$, en cuyo caso H_0 no sería rechazada, de nuevo una conclusión incorrecta. Así, es posible que se pudiera rechazar H_0 cuando es verdadera o que no fuera rechazada cuando es falsa. Los errores posibles no son consecuencias de una región de rechazo elegida sin buen juicio. Cualquiera de estos dos errores podría resultar cuando se emplea la región $x \leq 14$ o de hecho cuando se utiliza cualquier otra región.

DEFINICIÓN

> Un **error de tipo I** consiste en rechazar la hipótesis nula H_0 cuando es verdadera.
> Un **error de tipo II** requiere no rechazar H_0 cuando H_0 es falsa.

En el problema de la nicotina, un error de tipo I consiste en rechazar la afirmación del fabricante de que $\mu = 1.5$ cuando en realidad es cierta. Si se emplea la región de rechazo $\bar{x} \geq 1.6$, podría suceder que $\bar{x} = 1.63$ aun cuando $\mu = 1.5$, que da como resultado un error de tipo I. De otro modo, puede ser que H_0 sea falsa y, sin embargo, se observe $\bar{x} = 1.52$, lo cual da lugar a que H_0 no sea rechazada (un error de tipo II).

En el mejor de los casos, se podrían desarrollar procedimientos de prueba para los cuales ningún tipo de error es posible. Sin embargo, este ideal se puede lograr sólo si una decisión se basa en un examen de la población completa, lo cual casi siempre resulta impráctico. La dificultad con el uso de un procedimiento basado en datos muestrales, es que debido a la variabilidad del muestreo, podría resultar una muestra no representativa. Aunque $E(\bar{X}) = \mu$, el valor observado $\bar{x}$ podría diferir bastante de μ (por lo menos si n es pequeña). Así, cuando $\mu = 1.5$ en el caso de la nicotina, $\bar{x}$ podría ser mucho mayor que 1.5, lo que da como resultado el rechazo erróneo de H_0. Por otro lado, puede ser que $\mu = 1.6$ aunque se observe un valor de $\bar{x}$ mucho más pequeño que éste, lo cual origina un error de tipo II.

El vez de demandar procedimientos sin errores, se deben considerar procedimientos para los que sea poco probable que ocurra cualquier tipo de error. Es decir, un buen procedimiento es aquel para el que la probabilidad de cometer cualquier tipo de error es pequeña. La elección de un valor de corte particular para una región de rechazo fija las probabilidades de tipos I y II. Estas probabilidades de error se denotan por lo común con α y β, respectivamente. Debido a que H_0 especifica un valor único del parámetro, hay un solo valor de α. Sin embargo, hay un valor de β diferente para cada valor del parámetro consistente con H_a.

Ejemplo 8.1

Se sabe que cierto tipo de automóvil permanece sin daño visible 25% del tiempo en pruebas de colisión a 10 mph. En un esfuerzo por incrementar este porcentaje se ha propuesto un diseño de defensa modificada. Sea p la proporción de las colisiones a 10 mph con esta nueva defensa que no producen daño visible. Las hipótesis por probar son H_0: $p = 0.25$ (ninguna mejoría) contra H_a: $p > 0.25$. La prueba se basará en un experimento que requiere $n = 20$ colisiones independientes con prototipos del nuevo diseño. De una manera intuitiva, H_0 se debe rechazar si no muestran daño un número sustancial de colisiones. Considere el procedimiento de prueba siguiente:

Estadística de prueba: X = número de colisiones sin daño visible

Región de rechazo: $R_8 = \{8, 9, 10, \ldots, 19, 20\}$; es decir, rechazar H_0 si $x \geq 8$, donde x es el valor observado de la estadística de prueba.

Esta región de rechazo se llama de *extremo superior*, debido a que consiste sólo en valores grandes de la estadística de prueba.

Cuando H_0 es cierta, X tiene una distribución de probabilidad binomial con $n = 20$ y $p = 0.25$. Entonces,

$$\begin{aligned}\alpha &= P(\text{error de tipo I}) = P(\text{se rechaza } H_0 \text{ cuando es verdadera}) \\ &= P(X \geq 8 \text{ cuando } X \sim \text{Bin}(20, 0.25)) = 1 - B(7; 20, 0.25) \\ &= 1 - 0.898 = 0.102\end{aligned}$$

Es decir, cuando en realidad H_0 es verdadera, alrededor de 10% de los experimentos que consisten en 20 colisiones darían como resultado que se rechazara de forma incorrecta a la hipótesis nula, H_0 (un error de tipo I).

En contraste con α, no hay una sola β. En cambio, hay una β distinta para cada p diferente que pasa de 0.25. Hay un valor de β para $p = 0.3$ (en cuyo caso $X \sim \text{Bin}(20, 0.3)$), otro valor de β para $p = 0.5$, etcétera. Por ejemplo,

$$\begin{aligned}\beta(0.3) &= P(\text{error tipo II cuando } p = 0.3)\\ &= P(\text{no se rechaza } H_0 \text{ cuando es falsa porque } p = 0.3)\\ &= P(X \le 7 \text{ cuando } X \sim \text{Bin}(20, 0.3)) = B(7; 20, 0.3) = 0.772\end{aligned}$$

Cuando p es en realidad 0.3 y no 0.25 (una desviación "pequeña" con respecto a H_0), ¡casi 70% de los experimentos de este tipo darían como resultado el rechazo incorrecto de H_0!

En la tabla siguiente, se muestra β para valores seleccionados de p (cada uno calculado para la región de rechazo R_8). Es obvio que β disminuye cuando el valor de p se aparta a la derecha del valor nulo 0.25. De manera intuitiva, mientras mayor sea la desviación con respecto a H_0, es menos probable que ésta no se detecte.

p	0.3	0.4	0.5	0.6	0.7	0.8
$\beta(p)$	0.772	0.416	0.132	0.021	0.001	0.000

El procedimiento de prueba propuesto aún es razonable para probar la hipótesis nula más real de que $p \le 0.25$. En este caso, ya no hay una sola α, sino una para cada p que es a lo sumo 0.25: $\alpha(0.25)$, $\alpha(0.23)$, $\alpha(0.20)$, $\alpha(0.15)$, etcétera. No obstante, se comprueba fácilmente que $\alpha(p) < \alpha(0.25) = 0.102$ si $p < 0.25$. Es decir, el valor más grande de α ocurre para el valor límite 0.25 entre H_0 y H_a. Por lo tanto, si α es pequeña para la hipótesis nula simplificada, será igual de pequeña o más para la H_0 más real. ∎

Ejemplo 8.2 Se sabe que el tiempo de secado de cierto tipo de pintura en condiciones de prueba específicas, tiene una distribución normal con valor medio de 75 min y desviación estándar de 9 min. Los químicos han propuesto un nuevo aditivo diseñado para disminuir el tiempo promedio de secado. Se cree que los tiempos de secado con este aditivo aún tienen una distribución normal con $\sigma = 9$. Debido al gasto relacionado con el aditivo, la evidencia debe hacer pensar con certeza en una mejoría en el tiempo de secado promedio antes de adoptar tal conclusión. Designe a μ como el tiempo de secado promedio cuando se emplea el aditivo. Las hipótesis apropiadas son H_0: $\mu = 75$ contra H_a: $\mu < 75$. Sólo si se puede rechazar H_0 se declara que el aditivo es exitoso y se usa.

Los datos experimentales consistirán en los tiempos de secado de $n = 25$ especímenes de prueba. Sean $X_1, \ldots, X_{25}$ los 25 tiempos de secado, una muestra aleatoria de tamaño 25 de una distribución normal con valor medio μ y desviación estándar $\sigma = 9$. El tiempo de secado promedio muestral $\overline{X}$ tiene entonces una distribución normal con valor esperado $\mu_{\overline{X}} = \mu$ y desviación estándar $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 9/\sqrt{25} = 1.80$. Cuando H_0 es verdadera, $\mu_{\overline{X}} = 75$, así que un valor de $\overline{x}$ un poco menor que 75 no sería tan contradictoria para H_0. Una región de rechazo razonable tiene la forma $\overline{X} \le c$, donde el valor de corte c se elige de manera adecuada. Considere la elección $c = 70.8$, de modo que el procedimiento de prueba consiste en la estadística de prueba $\overline{X}$ y la región de rechazo $\overline{x} \le 70.8$. Debido a que la región de rechazo consiste sólo en valores pequeños de la estadística de prueba, se dice que la prueba es de *extremo inferior*. El cálculo de α y β ahora requiere una rutina de estandarización de $\overline{X}$ seguida de la referencia a las probabilidades normales estándar de la tabla A.3 del apéndice:

$$\begin{aligned}\alpha &= P(\text{error de tipo I}) = P(H_0 \text{ se rechaza } H_0 \text{ cuando es verdadera})\\ &= P(\overline{X} \le 70.8 \text{ cuando } \overline{X} \sim \text{normal con } \mu_{\overline{X}} = 75,\ \sigma_{\overline{X}} = 1.8)\\ &= \Phi\left(\frac{70.8 - 75}{1.8}\right) = \Phi(-2.33) = 0.01\end{aligned}$$

$$\begin{aligned}\beta(72) &= P(\text{error tipo II cuando } \mu = 72)\\ &= P(\text{no se rechaza } H_0 \text{ cuando es falsa porque } \mu = 72)\\ &= P(\overline{X} > 70.8 \text{ cuando } \overline{X} \sim \text{normal con } \mu_{\overline{X}} = 72 \text{ y } \sigma_{\overline{X}} = 1.8)\\ &= 1 - \Phi\left(\frac{70.8 - 72}{1.8}\right) = 1 - \Phi(-0.67) = 1 - 0.2514 = 0.7486\end{aligned}$$

$$\beta(70) = 1 - \Phi\left(\frac{70.8 - 70}{1.8}\right) = 0.3300 \qquad \beta(67) = 0.0174$$

Para el procedimiento de prueba especificado, sólo 1% de los experimentos realizados como se describió, dan como resultado que se rechace H_0 cuando en realidad es verdadera. Sin embargo, la probabilidad de un error tipo II es muy grande cuando $\mu = 72$ (sólo una pequeña desviación con respecto a H_0), un poco menor cuando $\mu = 70$ y bastante pequeño cuando $\mu = 67$ (una desviación muy sustancial respecto a H_0). Estas probabilidades de error se ilustran en la figura 8.1. Observe que α se calcula usando la distribución de probabilidad de la estadística de prueba cuando H_0 es verdadera, mientras que la determinación de β requiere conocer la distribución de la estadística de prueba cuando H_0 es falsa.

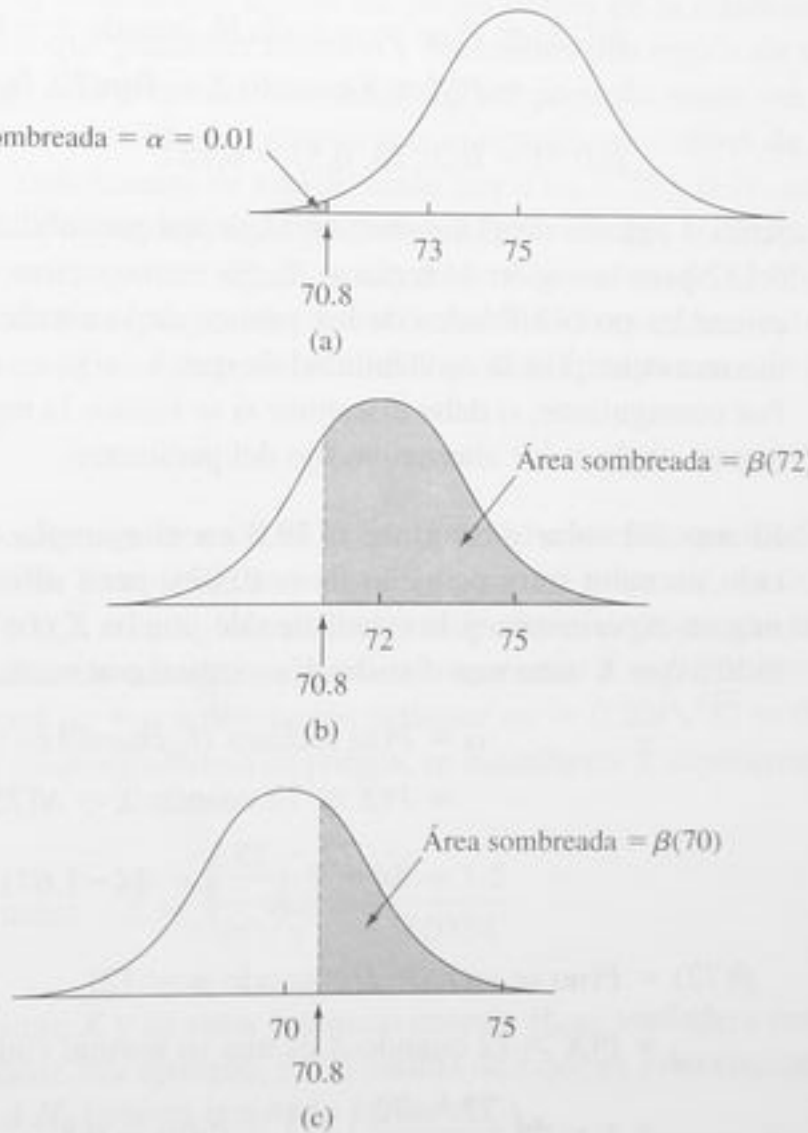


Figura 8.1 Ilustración de α y β para el ejemplo 8.2: (a) distribución de $\overline{X}$ cuando $\mu = 75$ (H_0 verdadera); (b) distribución de $\overline{X}$ cuando $\mu = 72$ (H_0 falsa); (c) distribucion de $\overline{X}$ cuando $\mu = 70$ (H_0 falsa)

Como en el ejemplo 8.1, si se considera la hipótesis nula más real $\mu \geq 75$, hay una α para cada valor de parámetro para el cual H_0 es verdadera: $\alpha(75)$, $\alpha(75.8)$, $\alpha(76.5)$, etcétera. Sin embargo, se comprueba fácilmente que $\alpha(75)$ es la más grande de todas estas probabilidades de error de tipo I. Centrar la atención en el valor límite equivale a trabajar de manera explícita con el "peor caso". ■

La especificación de un valor de corte para la región de rechazo en los ejemplos recién considerados fue un poco arbitraria. El empleo de la región de rechazo $R_8 = \{8, 9, \ldots, 20\}$ en el ejemplo 8.1 dio como resultado $\alpha = 0.102$, $\beta(0.3) = 0.772$, y $\beta(0.5) = 0.132$. Muchos considerarían que estas probabilidades de error son demasiado grandes. Quizá se puedan reducir si se cambia el valor de corte.

Ejemplo 8.3 (Continuación del ejemplo 8.1)

Se emplea el mismo experimento y la estadística de prueba X que se describieron en el problema de la defensa de automóvil, pero ahora considere la región de rechazo $R_9 = \{9, 10, \ldots, 20\}$. Puesto que X aún tiene una distribución binomial con parámetros $n = 20$ y p,

$$\begin{aligned}\alpha &= P(\text{se rechaza } H_0 \text{ cuando } p = 0.25)\\ &= P(X \geq 9 \text{ cuando } X \sim \text{Bin}(20, 0.25)) = 1 - B(8; 20, 0.25) = 0.041\end{aligned}$$

La probabilidad de error de tipo I se redujo al usar la nueva región de rechazo. Sin embargo, se pagó un precio por esta disminución:

$$\begin{aligned}\beta(0.3) &= P(\text{no se rechaza } H_0 \text{ cuando } p = 0.3)\\ &= P(X \leq 8 \text{ cuando } X \sim \text{Bin}(20, 0.3)) = B(8; 20, 0.3) = 0.887\\ \beta(0.5) &= B(8; 20, 0.5) = 0.252\end{aligned}$$

Ambos valores de β son mayores que las probabilidades de error correspondientes 0.772 y 0.132 para la región de rechazo R_8. En retrospectiva, esto no es sorprendente; α se calcula al sumar las probabilidades de los valores de la estadística de prueba *en la región de rechazo*, mientras que β es la probabilidad de que X caiga *en el complemento* de la región de rechazo. Por consiguiente, α debe disminuir si se reduce la región de rechazo al tiempo que β aumenta para algún valor alternativo fijo del parámetro. ■

Ejemplo 8.4 (Continuación del ejemplo 8.2)

El uso del valor de corte $c = 70.8$ en el ejemplo del secado de pintura, dio como resultado un valor muy pequeño de α (0.01), pero valores de β bastante grandes. Considere el mismo experimento y la estadística de prueba $\bar{X}$ con la nueva región de rechazo $\bar{x} \leq 72$. Debido a que $\bar{X}$ tiene una distribución normal con media $\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}} = 1.8$,

$$\begin{aligned}\alpha &= P(\text{se rechaza } H_0 \text{ cuando es verdadera})\\ &= P(\bar{X} \leq 72 \text{ cuando } \bar{X} \sim N(75, 1.8^2))\\ &= \Phi\left(\frac{72 - 75}{1.8}\right) = \Phi(-1.67) = 0.0475 \approx 0.05\end{aligned}$$

$$\begin{aligned}\beta(72) &= P(\text{no se rechaza } H_0 \text{ cuando } \mu = 72)\\ &= P(\bar{X} > 72 \text{ cuando } \bar{X} \text{ es una } va \text{ normal con media 72 y desviación estándar 1.8})\\ &= 1 - \Phi\left(\frac{72 - 72}{1.8}\right) = 1 - \Phi(0) = 0.5\end{aligned}$$

$$\beta(70) = 1 - \Phi\left(\frac{72 - 70}{1.8}\right) = 0.1335 \qquad \beta(67) = 0.0027$$

El cambio en el valor de corte hizo más grande la región de rechazo (incluye más valores de $\bar{x}$, lo que da como resultado una disminución de β para cada valor fijo μ menor que 75. Sin embargo, para esta nueva región, el valor de α se incrementó del valor previo 0.01 a alrededor de 0.05. No obstante, si se puede tolerar una probabilidad de error de tipo I de este tamaño, la segunda región ($c = 72$) es preferible a la primera ($c = 70.8$) debido a las β más pequeñas. ■

Los resultados de estos ejemplos se pueden generalizar de la siguiente manera.

PROPOSICIÓN

> Suponga que permanecen fijos un experimento y un tamaño de muestra, y se elige una estadística de prueba. Entonces, disminuir el tamaño de la región de rechazo para obtener un valor más pequeño de α, da como resultado un valor más grande de β para algún parámetro particular consistente con H_a.

Esta proposición expresa que una vez fijas la estadística de pruebe y n, no hay región de rechazo que al mismo tiempo haga pequeñas a α y β. Se debe elegir una región para llevar a cabo un compromiso entre α y β.

Debido a las normas sugeridas para especificar H_0 y H_a, por lo común un error de tipo I es más grave que un error de tipo II (esto siempre se puede lograr mediante la elección apropiada de la hipótesis). El método que siguen los profesionales de la estadística es especificar el valor más grande de α que puede ser tolerado y determinar una región de rechazo que tiene ese valor de α y no algo más pequeño. Esto hace a β tan pequeña como sea posible sujeta al límite en α. El valor resultante de α se conoce por lo común como **nivel de significación** de la prueba. Los niveles tradicionales de significación son 0.10, 0.05 y 0.01, aunque el nivel en algún problema particular dependerá de la gravedad del error de tipo I, mientras más grave sea este error, menor debe ser el nivel de significación. El procedimiento de prueba correspondiente se llama **prueba de nivel α** (p. ej., una prueba de nivel 0.05 o una de nivel 0.01). Una prueba con nivel de significación α es una para la cual la probabilidad de error de tipo I se controla al nivel especificado.

Ejemplo 8.5 Considere la situación antes mencionada en la que μ fue el contenido promedio de nicotina de los cigarros de la marca B. El objetivo es probar H_0: $\mu = 1.5$ contra H_a: $\mu > 1.5$ con base en la muestra aleatoria $X_1, X_2, \ldots, X_{32}$ de contenidos de nicotina. Suponga que se sabe que la distribución del contenido de nicotina es normal con media $\sigma = 0.20$. Entonces X se distribuye normalmente con $\mu_{\bar{X}} = \mu$ y desviación estándar $\sigma_{\bar{X}} = 0.20/\sqrt{32} = 0.0354$.

En vez de usar $\bar{X}$ como estadística de prueba, se estandariza $\bar{X}$ suponiendo que H_0 es verdadera.

$$\text{Estadística de prueba:} \quad Z = \frac{\bar{X} - 1.5}{\sigma/\sqrt{n}} = \frac{\bar{X} - 1.5}{0.0354}$$

Z expresa la distancia entre $\bar{X}$ y su valor esperado cuando H_0 es verdadera como algún número de desviaciones estándar. Por ejemplo, $z = 3$ resulta de $\bar{x}$ que es 3 desviaciones estándar más grande de lo esperado si H_0 hubiera resultado verdadera.

Rechazar H_0 cuando $\bar{x}$ pasa de 1.5 "en forma considerable" es equivalente a rechazar H_0 cuando z pasa de cero "en forma considerable". Es decir, la forma de la región de rechazo es

$z \geq c$. Ahora se determina c de modo que $\alpha = 0.05$. Cuando H_0 es verdadera, Z tiene una distribución normal estándar. Por consiguiente,

$$\begin{aligned}\alpha &= P(\text{error de tipo I}) = P(\text{rechazar } H_0 \text{ cuando } H_0 \text{ es verdadera})\\ &= P(Z \geq c \text{ cuando } Z \sim N(0, 1))\end{aligned}$$

El valor c debe captar 0.05 del área del extremo superior bajo la curva de z. Ya sea de la sección 4.3 o directamente de la tabla A.3 del apéndice, $c = z_{0.05} = 1.645$.

Observe que $z \geq 1.645$ es equivalente a $\bar{x} - 1.5 \geq (0.0354)(1.645)$, es decir, $\bar{x} \geq 1.56$. Entonces β es la probabilidad de que $\bar{X} < 1.56$ y se puede calcular para cualquier μ mayor que 1.5. ■

Ejercicios | Sección 8.1 (1 a 14)

1. Para cada una de las siguientes afirmaciones, exprese si es una hipótesis estadística legítima y por qué:
   a. H: $\sigma > 100$ b. H: $\bar{x} = 45$
   c. H: $s \leq 0.20$ d. H: $\sigma_1/\sigma_2 < 1$
   e. H: $\bar{X} - \bar{Y} = 5$
   f. H: $\lambda \leq 0.01$, donde λ es el parámetro de una distribución exponencial utilizada para modelar el tiempo de vida de componentes.
2. Para los siguientes pares de afirmaciones, indique cuál no cumple con las reglas para plantear hipótesis y por qué (los subíndices 1 y 2 diferencian entre cantidades para dos poblaciones o muestras diferentes):
   a. H_0: $\mu = 100$, H_a: $\mu > 100$
   b. H_0: $\sigma = 20$, H_a: $\sigma \leq 20$
   c. H_0: $p \neq 0.25$, H_a: $p = 0.25$
   d. H_0: $\mu_1 - \mu_2 = 25$, H_a: $\mu_1 - \mu_2 > 100$
   e. H_0: $S_1^2 = S_2^2$, H_a: $S_1^2 \neq S_2^2$
   f. H_0: $\mu = 120$, H_a: $\mu = 150$
   g. H_0: $\sigma_1/\sigma_2 = 1$, H_a: $\sigma_1/\sigma_2 \neq 1$
   h. H_0: $p_1 - p_2 = -0.1$, H_a: $p_1 - p_2 < -0.1$
3. Para determinar si las soldaduras de tubería en una planta de energía nuclear satisfacen las especificaciones, se selecciona una muestra aleatoria de soldaduras. La resistencia de la soldadura se mide como la fuerza requerida para romper la soldadura. Suponga que las especificaciones establecen que la resistencia promedio de las soldaduras debe pasar de 100 lb/pulg^2; el equipo de inspección decide probar H_0: $\mu = 100$ contra H_a: $\mu > 100$. Explique por qué podría ser preferible usar esta H_a y no $\mu < 100$.
4. Represente con μ el nivel de radiactividad promedio real (picocuries por litro). El valor 5 pCi/L es considerado la línea divisora entre el agua segura e insegura. ¿Recomendaría probar H_0: $\mu = 5$ contra H_a: $\mu > 5$ o H_0: $\mu = 5$ contra H_a: $\mu < 5$? Explique su razonamiento. (*Sugerencia*: considere las consecuencias de un error de tipo I y tipo II para cada posibilidad).
5. Antes de estar de acuerdo en comprar un gran pedido de fundas de polietileno para un tipo particular de cable de energía submarino relleno con aceite a alta presión, una compañía desea ver una evidencia concluyente de que la desviación estándar verdadera del espesor de la funda es menor que 0.05 mm. ¿Qué hipótesis se debe probar y por qué? En este contexto, ¿cuáles son los errores tipo I y tipo II?
6. Muchos hogares antiguos tienen sistemas eléctricos que utilizan fusibles en vez de interruptores de circuito. Un fabricante de fusibles de 40 amperes quiere asegurarse de que el amperaje promedio al que se queman los fusibles es de hecho 40. Si el amperaje promedio es menor que 40, el fabricante podría ser responsable de daño a un sistema eléctrico debido al mal funcionamiento del fusible. Para comprobar el amperaje de los fusibles, se elige e inspecciona una muestra de fusibles. Si se realizara una prueba de hipótesis en los datos resultantes, ¿cuáles hipótesis nula y alternativa serían de interés para el fabricante? Describa los errores tipo I y tipo II en el contexto de esta situación del problema.

7. Se toman muestras del agua utilizada para enfriamiento cuando se descarga de una planta de energía hacia un río. Se ha determinado que siempre y cuando la temperatura promedio del agua descargada sea a lo sumo 150°F, no habrá efectos negativos en el ecosistema del río. Para investigar si la planta cumple con las normas que prohíben una temperatura promedio de agua de descarga arriba de 150°, se tomarán 50 muestras de agua en tiempos elegidos al azar, y se registrará la temperatura de cada muestra. Los datos resultantes se usarán para probar las hipótesis H_0: $\mu = 150°$ contra H_a: $\mu > 150°$. En el contexto de esta situación, describa los errores tipo I y II. ¿Qué tipo de error consideraría más grave? Explique.

8. Un fabricante de tarjetas de circuitos utiliza en la actualidad un tipo regular de laminado. Se ha desarrollado un laminado especial para reducir el alabeo. El laminado regular se usará en una muestra de especímenes y el laminado especial en otra muestra, y luego se determinará la cantidad de alabeo para cada espécimen. El fabricante cambiará al laminado especial sólo si se puede demostrar que la cantidad promedio de alabeo para el laminado es menor que la del laminado regular. Exprese las hipótesis pertinentes y describa los errores de tipo I y II en el contexto de esta situación.

9. Dos compañías diferentes pretenden suministrar el servicio de televisión por cable en una cierta región. Denote con p la proporción de los posibles suscriptores que favorecen a la primera compañía y no a la segunda. Considere probar H_0: $p = 0.5$ contra H_a: $p \neq 0.5$ con base en una muestra aleatoria de 25 individuos. Sea X el número de individuos de la muestra que favorecen a la primera compañía y x el valor observado de X.
   a. ¿Cuál de las siguientes regiones de rechazo es más apropiada y por qué?

      $R_1 = \{x: x \leq 7 \text{ o } x \geq 18\}$, $R_2 = \{x: x \leq 8\}$,

      $R_3 = \{x: x \geq 17\}$

   b. En el contexto de esta situación problema, describa qué errores de tipo I y II se tienen.
   c. ¿Cuál es la distribución de probabilidad de la estadística de prueba X cuando H_0 es verdadera? Utilícela para calcular la probabilidad de un error tipo I.
   d. Calcule la probabilidad de error tipo II para la región seleccionada cuando $p = 0.3$, de nuevo cuando $p = 0.4$ y también para $p = 0.6$ y $p = 0.7$.
   e. Con la región seleccionada, ¿qué concluiría si seis de los 25 interrogados favorecen a la compañía 1?

10. Una mezcla de ceniza de combustible pulverizada y cemento Portland que se usará para lechada debe tener una resistencia compresiva de más de 1300 KN/m². La mezcla no se usará a menos que la evidencia experimental indique que se ha cumplido la especificación de resistencia. Suponga que la resistencia compresiva para especímenes de esta mezcla tiene una distribución normal con $\sigma = 60$. Sea μ la resistencia compresiva promedio real.
    a. ¿Cuáles son las hipótesis nula y alternativa apropiadas?
    b. Represente con $\overline{X}$ la resistencia compresiva promedio muestral para $n = 20$ especímenes seleccionados al azar. Considere el procedimiento de prueba con la estadística de prueba $\overline{X}$ y la región de rechazo $\bar{x} \geq 1331.26$. ¿Cuál es la distribución de la probabilidad de la estadística de la prueba cuando H_0 es verdadera? ¿Cuál es la probabilidad de un error tipo I para este procedimiento de prueba?
    c. ¿Cuál es la distribución de probabilidad de la estadística de prueba cuando $\mu = 1350$? Usando el procedimiento de prueba del inciso (b), ¿cuál es la probabilidad de que se juzgue insatisfactoria la mezcla cuando de hecho $\mu = 1350$ (un error de tipo II)?
    d. ¿Cómo cambiaría el procedimiento de prueba del inciso (b) para obtener una prueba con nivel de significación 0.05? ¿Qué impacto tendría este cambio en la probabilidad de error del inciso (c)?
    e. Considere la estadística de prueba estandarizada $Z = (\overline{X} - 1300)/(\sigma/\sqrt{n}) = (\overline{X} - 1300)/13.42$. ¿Cuáles son los valores de Z que corresponden a la región de rechazo del inciso (b)?

11. La calibración de una balanza se comprobará al pesar 25 veces un espécimen de prueba de 10 kg. Suponga que los resultados de las diferentes pesadas son independientes entre sí y que el peso de cada prueba tiene una distribución normal con $\sigma = 0.200$ kg. Represente con μ la lectura de peso promedio real en la balanza.
    a. ¿Cuáles hipótesis se deben probar?
    b. Suponga que la balanza se calibrará de nuevo si $\bar{x} \geq 10.1032$ o $\bar{x} \leq 9.8968$. ¿Cuál es la probabilidad de que se lleve a cabo la recalibración cuando en realidad es innecesaria?
    c. ¿Cuál es la probabilidad de que se juzgue innecesaria la recalibración cuando de hecho $\mu = 10.1$? ¿Cuando $\mu = 9.8$?
    d. Sea $z = (\bar{x} - 10)/(\sigma/\sqrt{n})$. ¿Para qué valor c la región de rechazo del inciso (b) es equivalente a la región con dos extremos *ya sea* $z \geq c$ *o bien* $z \leq -c$?
    e. Si el tamaño de muestra fuera sólo 10 y no 25, ¿cómo se modificaría el procedimiento del inciso (d) de modo que $\alpha = 0.05$?

f. Si emplea la prueba del inciso (e), ¿qué se concluiría a partir de los siguientes datos muestrales:

9.981 10.006 9.857 10.107 9.888
9.728 10.439 10.214 10.190 9.793

g. Vuelva a expresar el procedimiento de prueba del inciso (b) en términos de la estadística de prueba estandarizada $Z = (\bar{X} - 10)/(\sigma/\sqrt{n})$.

12. Se propuso un nuevo diseño para el sistema de frenado en cierto tipo de automóvil. Para el sistema actual, se sabe que la distancia de frenado promedio real a 40 mph en las condiciones especificadas es 120 pies. Se propone que el nuevo diseño se ponga en práctica sólo si los datos muestrales indican de manera firme una reducción en la distancia promedio de frenado para el nuevo diseño.
   a. Defina el parámetro de interés y exprese las hipótesis pertinentes.
   b. Suponga que la distancia de frenado para el nuevo sistema tiene una distribución normal con $\sigma = 10$. Represente con $\bar{X}$ la distancia de frenado promedio muestral para una muestra aleatoria de 36 observaciones. ¿Cuál de las siguientes regiones de rechazo es apropiada: $R_1 = \{\bar{x}: \bar{x} \geq 124.80\}$, $R_2 = \{\bar{x}: \bar{x} \leq 115.20\}$, $R_3 = \{\bar{x}:$ ya sea $\bar{x} \geq 125.13$ o bien $\bar{x} \leq 114.87\}$?
   c. ¿Cuál es el nivel de significación para la región apropiada del inciso (b)? ¿Cómo cambiaría la región para obtener una prueba con $\alpha = 0.001$?
   d. ¿Cuál es la probabilidad de que el nuevo diseño no se ponga en práctica cuando su distancia de frenado promedio real sea en realidad 115 pies y se utilice la región apropiada del inciso (b)?
   e. Sea $Z = (\bar{X} - 120)/(\sigma/\sqrt{n})$. ¿Cuál es el nivel de significación para la región de rechazo $\{z: z \leq -2.33\}$? ¿Para la región $\{z: z \leq -2.88\}$?

13. Represente con $X_1, \ldots, X_n$ una muestra aleatoria de una distribución poblacional normal con un valor conocido de σ.
   a. Para probar las hipótesis H_0: $\mu = \mu_0$ contra H_a: $\mu > \mu_0$ (donde μ_0 es un número fijo), demuestre que la prueba con la estadística de prueba $\bar{X}$ y región de rechazo $\bar{x} \geq \mu_0 + 2.33\sigma/\sqrt{n}$ tiene nivel de significación 0.01.
   b. Suponga que se utiliza el procedimiento del inciso (a) para probar H_0: $\mu \leq \mu_0$ contra H_a: $\mu > \mu_0$. Si $\mu_0 = 100$, $n = 25$, y $\sigma = 5$, ¿cuál es la probabilidad de cometer un error tipo I cuando $\mu = 99$? ¿Cuando $\mu = 98$? En general, ¿qué se puede decir acerca de la probabilidad de un error de tipo I cuando el valor real de μ es menor que μ_0? Compruebe su aseveración.

14. Considere de nuevo la situación del ejercicio 11 y suponga que la región de rechazo es $\{\bar{x}: \bar{x} \geq 10.1004$ o $\bar{x} \leq 9.8940\} = \{z: z \geq 2.51$ o $z \leq -2.65\}$.
   a. ¿Cuál es el valor de α para este procedimiento?
   b. ¿Cuánto vale β cuando $\mu = 10.1$? ¿Cuándo $\mu = 9.9$? ¿Es deseable esto?

8.2 Pruebas acerca de una media poblacional

La explicación general del capítulo 7 acerca de intervalos de confianza para una media poblacional μ se centró en tres casos diferentes. Ahora se desarrollan procedimientos de prueba para estos mismos tres casos.

Caso I: población normal con σ conocida

Aunque la suposición de que se conoce el valor de σ rara vez se cumple en la práctica, este caso proporciona un buen punto de partida debido a la facilidad con la que se pueden desarrollar procedimientos generales y sus propiedades. La hipótesis nula en los tres casos expresará que μ tiene un valor numérico particular, el *valor nulo*, que se denotará con μ_0. Represente con $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de la población normal. Entonces la media muestral $\bar{X}$ tiene una distribución normal con valor esperado $\mu_{\bar{X}} = \mu$ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Cuando H_0 es verdadera, $\mu_{\bar{X}} = \mu_0$. Ahora considere el estadístico Z obtenido al estandarizar $\bar{X}$ bajo la suposición de que H_0 es verdadera:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

La sustitución de la media muestral calculada $\bar{x}$ produce z, la distancia entre $\bar{x}$ y μ_0 expresada en "unidades de desviación estándar". Por ejemplo, si la hipótesis nula es H_0: $\mu = 100$, $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 10/\sqrt{25} = 2.0$ y $\bar{x} = 103$, entonces el valor del estadístico de prueba es $z = (103 - 100)/2.0 = 1.5$. Es decir, el valor observado de $\bar{x}$ está 1.5 desviaciones estándar (de $\bar{X}$) arriba de lo que se espera que esté cuando H_0 es verdadera. El estadístico Z es una medida natural de la distancia entre $\bar{X}$, el estimado de μ y su valor esperado cuando H_0 es verdadera. Si la distancia es demasiado grande en una dirección consistente con H_a, se debe rechazar la hipótesis nula.

Suponga primero que la hipótesis alternativa tiene la forma H_a: $\mu > \mu_0$. Entonces un valor de $\bar{x}$ menor que μ_0 en realidad no proporciona respaldo para H_a. Esta $\bar{x}$ corresponde a un valor negativo de z (puesto que $\bar{x} - \mu_0$ es negativa y el divisor $\sigma/\sqrt{n}$ es positivo). De manera similar, un valor de $\bar{x}$ que pasa de μ_0 por sólo una pequeña cantidad (correspondiente a z que es positiva pero pequeña) no indica que se deba rechazar H_0 en favor de H_a. El rechazo de H_0 es apropiado sólo cuando $\bar{x}$ pasa de μ_0 en forma considerable; es decir, cuando el valor de z es positivo y grande. En resumen, la región de rechazo apropiada, basada en el estadístico Z y no en $\bar{X}$, tiene la forma $z \geq c$.

Como se explicó en la sección 8.1, el valor de corte c se debe elegir para controlar la probabilidad de un error tipo I al nivel deseado α. Esto se lleva a cabo sin dificultad debido a que la distribución del estadístico de prueba Z cuando H_0 es verdadera es la distribución normal estándar (ésta es la razón de que μ_0 se reste en la estandarización). El valor de corte c requerido es el valor crítico z que capta el área α del extremo superior bajo la curva normal estándar. Como ejemplo, sea $c = 1.645$, el valor que capta el área extrema 0.05 ($z_{0.05} = 1.645$). Entonces,

$$\begin{aligned}\alpha &= P(\text{error de tipo I}) = P(\text{se rechaza } H_0 \text{ cuando es verdadera})\\ &= P(Z \geq 1.645 \text{ cuando } Z \sim N(0, 1)) = 1 - \Phi(1.645) = 0.05\end{aligned}$$

En términos generales, la región de rechazo $z \geq z_\alpha$ tiene probabilidad α de error de tipo I. El procedimiento de prueba es de *extremo superior* porque la región de rechazo consiste sólo en valores grandes del estadístico de prueba.

Un razonamiento análogo para la hipótesis alternativa H_a: $\mu < \mu_0$ indica una región de rechazo de la forma $z \leq c$, donde c es un número negativo elegido de manera adecuada (muy abajo de μ_0 si y sólo si z es bastante negativa). Debido a que Z tiene una distribución normal estándar cuando H_0 es verdadera, al tomar $c = -z_\alpha$ se obtiene $P(\text{error de tipo I}) = \alpha$. Ésta es una prueba de *extremo inferior*. Por ejemplo, $z_{0.10} = 1.28$ implica que la región de rechazo $z \leq -1.28$ especifica un prueba con nivel de significación 0.10.

Por último, cuando la hipótesis alternativa es H_a: $\mu \neq \mu_0$, se debe rechazar H_0 si $\bar{x}$ está demasiado lejos de cualquier lado de μ_0. Esto es equivalente a rechazar H_0 ya sea si $z \geq c$ o si $z \leq -c$. Suponga que se desea $\alpha = 0.05$. Entonces,

$$\begin{aligned}0.05 &= P(Z \geq c \text{ o bien, } Z \leq -c \text{ cuando } Z \text{ tiene una distribución normal estándar})\\ &= \Phi(-c) + 1 - \Phi(c) = 2[1 - \Phi(c)]\end{aligned}$$

Por lo tanto, c es tal que $1 - \Phi(c)$, el área bajo la curva normal estándar a la derecha de c, es 0.025 (¡y no 0.05!). De la sección 4.3 o la tabla A.3 del apéndice, $c = 1.96$, y la región de rechazo es $z \geq 1.96$ *o bien* $z \leq -1.96$. Para cualquier α, la región de rechazo de *dos extremos* $z \geq z_{\alpha/2}$ *o bien* $z \leq -z_{\alpha/2}$ tiene probabilidad α de error tipo I (puesto que el área $\alpha/2$ queda bajo cada uno de los dos extremos de la curva z). De nuevo, la razón clave para usar el estadístico de prueba estandarizado Z es que Z tiene una distribución conocida cuando H_0 es verdadera (normal es-

tándar), una región de rechazo con la probabilidad deseada de error tipo I se obtiene sin dificultad al usar un valor crítico apropiado.

El procedimiento de prueba para el caso I, se resume en el cuadro siguiente, y las regiones de rechazo correspondientes se ilustran en la figura 8.2.

Hipótesis nula: H_0: $\mu = \mu_0$

Valor del estadístico de prueba: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Hipótesis alternativa	**Región de rechazo para la prueba del nivel α.**	
H_a: $\mu > \mu_0$	$z \geq z_\alpha$	(prueba de extremo superior)
H_a: $\mu < \mu_0$	$z \leq -z_\alpha$	(prueba de extremo inferior)
H_a: $\mu \neq \mu_0$	ya sea $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$	(prueba de dos extremos)

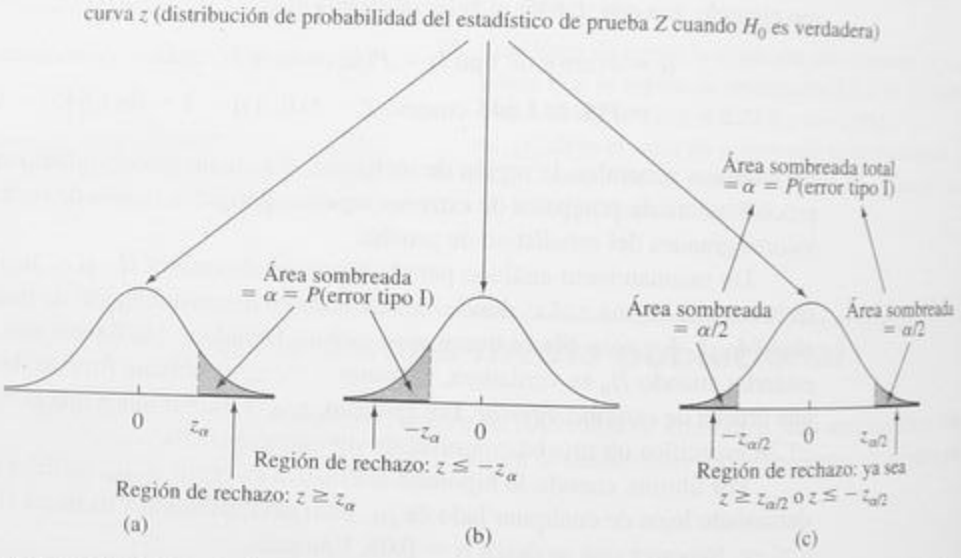


Figura 8.2 Regiones de rechazo para pruebas z (a) prueba de extremo superior; (b) prueba de extremo inferior, (c) prueba de dos extremos

El uso de la siguiente secuencia de pasos se recomienda al probar hipótesis respecto a un parámetro.

1. Identifique el parámetro de interés y descríbalo en el contexto de la situación problema.
2. Determine el valor nulo y exprese la hipótesis nula.
3. Exprese la hipótesis alternativa apropiada.

4. Dé la fórmula para el valor calculado del estadístico de prueba (sustituyendo el valor nulo y los valores conocidos de algunos otros parámetros, pero *no* los de cantidades basadas en la muestra).
5. Exprese la región de rechazo para el nivel de significación seleccionado α.
6. Calcule las cantidades muestrales necesarias, sustituya en la fórmula para el valor del estadístico de prueba y calcule ese valor.
7. Decida si se debe rechazar H_0 y enuncie esta conclusión en el contexto del problema.

La formulación de hipótesis (pasos 2 y 3) se debe hacer antes de examinar los datos.

Ejemplo 8.6 Un fabricante de sistemas rociadores utilizados para protección contra incendios en edificios de oficinas afirma que la temperatura promedio real de activación del sistema es 130°. Una muestra de $n = 9$ sistemas, cuando se prueban, produce una temperatura promedio muestral de activación de 131.08°F. Si la distribución de los tiempos de activación es normal con desviación estándar de 1.5°F, ¿los datos contradicen la afirmación del fabricante al nivel de significación $\alpha = 0.01$?

1. Parámetro de interés: μ = temperatura de activación promedio real.
2. Hipótesis nula: H_0: $\mu = 130$ (valor nulo = $\mu_0 = 130$).
3. Hipótesis alternativa: H_a: $\mu \neq 130$ (una desviación con respecto al valor afirmado en *cualquier* dirección de interés).
4. Valor del estadístico de prueba:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 130}{1.5/\sqrt{n}}$$

5. Región de rechazo: La forma de H_a implica el uso de una prueba de dos extremos con región de rechazo *ya sea* $z \geq z_{0.005}$ *o bien* $z \leq -z_{0.005}$. De la sección 4.3 o de la tabla A.3 del apéndice, $z_{0.005} = 2.58$, así que se rechaza H_0 si $z \geq 2.58$ o $z \leq -2.58$.
6. Sustituyendo $n = 9$ y $\bar{x} = 131.08$,

$$z = \frac{131.08 - 130}{1.5/\sqrt{9}} = \frac{1.08}{0.5} = 2.16$$

Es decir, la media muestral observada está un poco más de dos desviaciones estándar arriba de lo que se esperaría si H_0 fuera verdadera.

7. El valor calculado $z = 2.16$ no cae en la región de rechazo ($-2.58 < 2.16 < 2.58$), así que no se puede rechazar H_0 al nivel de significación 0.01. Los datos no respaldan de manera firme la afirmación de que el promedio real difiere del valor de diseño de 130. ■

Determinación de β y el tamaño de muestra Las pruebas z para el caso I son de las pocas en estadística para las cuales hay fórmulas simples para β, la probabilidad de un error tipo II. Considere primero la prueba de extremo superior con región de rechazo $z \geq z_\alpha$. Esto es equivalente a $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$, por lo tanto no se rechaza H_0 si $\bar{x} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$. Ahora, denote con μ' un valor particular de μ que pasa del valor nulo μ_0. Entonces,

$$\begin{aligned}\beta(\mu') &= P(\text{no se rechaza } H_0 \text{ cuando } \mu = \mu') \\ &= P(\bar{X} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n} \text{ cuando } \mu = \mu') \\ &= P\left(\frac{\bar{X} - \mu'}{\sigma/\sqrt{n}} < z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}} \text{ cuando } \mu = \mu'\right) \\ &= \Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)\end{aligned}$$

A medida que μ' aumenta, $\mu_0 - \mu'$ se vuelve más negativa, así que $\beta(\mu')$ será pequeña cuando μ' excede por mucho a μ_0 (porque el valor al que se evalúa Φ entonces será bastante negativo). Las probabilidades de error para las pruebas de extremo inferior y de dos extremos se derivan de manera análoga.

Si σ es grande, la probabilidad de error de tipo II puede ser grande a un valor alternativo μ' que es de interés particular para un investigador. Suponga que se fija α y también se especifica β para tal valor alternativo. En el ejemplo del rociador, los directivos de la compañía podrían ver a $\mu' = 132$ como una desviación muy sustancial con respecto a H_0: $\mu = 130$ y, por lo tanto, desear que $\beta(132) = 0.10$ además de $\alpha = 0.01$. En términos generales, considere las dos restricciones $P(\text{error tipo I}) = \alpha$ y $\beta(\mu') = \beta$ para valores específicos de α, μ', y β. Entonces para una prueba de extremo superior, se debe elegir el tamaño de muestra n que satisfaga

$$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) = \beta$$

Esto implica que

$$-z_\beta = \begin{array}{l}\text{valor crítico } z \text{ que capta} \\ \text{el área } \beta \text{ del extremo inferior}\end{array} = z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}$$

Es fácil resolver esta ecuación para la n deseada. Con un argumento similar se obtiene el tamaño de muestra necesario para pruebas de extremo inferior y de dos extremos como se resume en el cuadro siguiente.

Hipótesis alternativa	**Probabilidad $\beta(\mu')$ de error de tipo II para una prueba de nivel α**
H_a: $\mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
H_a: $\mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
H_a: $\mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$

donde $\Phi(z) =$ es la fda normal estándar.

El tamaño de muestra n para el cual una prueba de nivel α tiene $\beta(\mu') = \beta$ al valor alternativo μ' es

$$n = \begin{cases} \left[\dfrac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'}\right]^2 & \text{para una prueba de un extremo (superior o inferior)} \\ \left[\dfrac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu'}\right]^2 & \text{para una prueba de dos extremos (una solución aproximada)} \end{cases}$$

Ejemplo 8.7 Represente con μ la duración promedio de cierto tipo de llanta. Considere probar H_0: $\mu = 30\,000$ contra H_a: $\mu > 30\,000$ con base en una muestra de tamaño $n = 16$ de una distribución poblacional normal con $\sigma = 1500$. Una prueba con $\alpha = 0.01$ requiere que $z_\alpha = z_{0.01} = 2.33$. La probabilidad de cometer un error de tipo II cuando $\mu = 31\,000$ es

$$\beta(31\,000) = \Phi\left(2.33 + \frac{30\,000 - 31\,000}{1500/\sqrt{16}}\right) = \Phi(-0.34) = 0.3669$$

Puesto que $z_{0.1} = 1.28$, el requerimiento de que la prueba de nivel 0.01 también tenga $\beta(31\,000) = 0.1$ necesita

$$n = \left[\frac{1500(2.33 + 1.28)}{30\,000 - 31\,000}\right]^2 = (-5.42)^2 = 29.32$$

El tamaño de muestra debe ser un entero, así que se deben usar $n = 30$ llantas. ■

Caso II: Pruebas con muestras grandes

Cuando el tamaño de muestra es grande, las pruebas z para el caso I se modifican fácilmente para producir procedimientos de prueba válidos sin requerir una distribución poblacional normal o σ conocida. El resultado clave se utilizó en el capítulo 7 para justificar intervalos de confianza con muestras grandes: una n grande significa que la desviación estándar muestral s estará cerca de σ para la mayoría de las muestras, de modo que la variable estandarizada

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

tiene *aproximadamente* una distribución normal estándar. La sustitución del valor nulo μ_0 en lugar de μ produce el estadístico de prueba

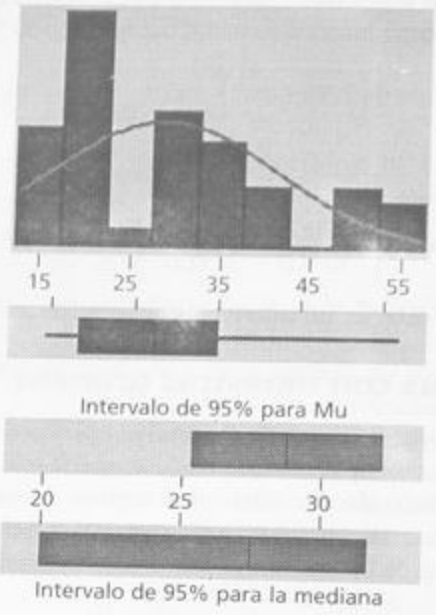
$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

que tiene más o menos una distribución normal estándar cuando H_0 es verdadera. El uso de las regiones de rechazo proporcionadas antes para el caso I (p. ej., $z \geq z_\alpha$ cuando la hipótesis alternativa es H_a: $\mu > \mu_0$) da como resultado procedimientos de prueba para los cuales el nivel de significación es casi (y no exactamente) α. La regla empírica $n > 40$ se empleará de nuevo para caracterizar un tamaño de muestra grande.

Ejemplo 8.8 Para medir la resistencia del material a la penetración cuando se introduce un cono en el pavimento se emplea un penetrómetro de cono dinámico (PCD). Supóngase que para una aplicación particular, se requiere que el valor del PCD promedio real para cierto tipo de pavimento sea menor que 30. No se usará el pavimento a menos que haya evidencia concluyente de que se cumple la especificación. Enúnciese y pruébese la hipótesis apropiada usando los datos siguientes ("Probabilistic Model for the Analysis of Dynamic Cone Penetrometer Test Values in Pavement Structure Evaluation", *J. of Testing and Evaluation*, 1999: 7-14):

14.1	14.5	15.5	16.0	16.0	16.7	16.9	17.1	17.5	17.8
17.8	18.1	18.2	18.3	18.3	19.0	19.2	19.4	20.0	20.0
20.8	20.8	21.0	21.5	23.5	27.5	27.5	28.0	28.3	30.0
30.0	31.6	31.7	31.7	32.5	33.5	33.9	35.0	35.0	35.0
36.7	40.0	40.0	41.3	41.7	47.5	50.0	51.0	51.8	54.4
55.0	57.0								

Estadística descriptiva



Variable: PCD

Prueba de normalidad de Anderson-Darling

A cuadrada	1.902
Valor P	0.000
Media	28.7615
Desv. Est.	12.2647
Varianza	150.423
Sesgo	0.808264
Curtosis	-3.9E-01
N	52
Mínimo	14.1000
Primer cuartil	18.2250
Mediana	27.5000
3er cuartil	35.0000
Máximo	57.0000

Intervalo de 95% para Mu

25.3470	32.1761

Intervalo de confianza de 95% para sigma

10.2784	15.2098

Intervalo de confianza de 95% para la mediana

20.0000	31.7000

Figura 8.3 Resumen descriptivo de MINITAB para los datos del PCD del ejemplo 8.8

En la figura 8.3 se muestra un resumen descriptivo obtenido de MINITAB. El valor medio muestral del PCD es menor que 30. Sin embargo, hay una cantidad sustancial de variación en los datos (coeficiente muestral de variación $= s/\bar{x} = 0.4265$) así que el hecho de que la media sea menor que el corte de especificación de diseño podría ser una consecuencia de la variabilidad de muestreo. Observe que el histograma no se asemeja en absoluto a una curva normal (y una gráfica de probabilidad normal no exhibe un patrón lineal), pero las pruebas z con muestras grandes no requieren una distribución poblacional normal.

1. μ = valor del PCD promedio muestral
2. H_0: $\mu = 30$
3. H_a: $\mu < 30$ (así que el pavimento no se usará a menos que se rechace la hipótesis nula)
4. $z = \dfrac{\bar{x} - 30}{s/\sqrt{n}}$
5. Por medio de una prueba con nivel de significación 0.05, H_0 se rechaza si $z \le -1.645$ (una prueba de extremo inferior).
6. Con $n = 52$, $\bar{x} = 28.76$, y $s = 12.2647$,

$$z = \frac{28.76 - 30}{12.2647/\sqrt{52}} = \frac{-1.24}{1.701} = -0.73$$

7. Puesto que $-0.73 > -1.645$, no se puede rechazar H_0. No se tiene evidencia convincente de que $\mu < 30$, así que no se justifica el uso del pavimento. ■

La determinación de β y el tamaño de muestra necesario para estas pruebas con muestras grandes se puede basar ya sea en especificar un valor posible de σ y usar las fórmulas de caso I (aun cuando se utiliza s en la prueba) o en usar las curvas que se presentarán en breve en relación con el caso III.

Caso III: Distribución poblacional normal

Cuando n es pequeña, ya no se puede utilizar el Teorema del Límite Central (TLC) para justificar el uso de una prueba con muestra grande. Se enfrentó esta dificultad al obtener un intervalo de confianza (IC) de muestra pequeña para μ en el capítulo 7. Aquí se usará el mismo método: se supondrá que la distribución poblacional es por lo menos aproximadamente normal y se describirán procedimientos de prueba cuya validez se base en esta suposición. Si un investigador tiene una buena razón para creer que la distribución poblacional es bastante no normal, se puede usar una prueba sin distribución del capítulo 15. De manera alternativa, se puede consultar a un estadístico en relación con los procedimientos válidos para familias específicas de distribuciones poblacionales que no sean de la familia de distribuciones normales. Se puede elaborar un procedimiento de inicialización (bootstrap).

El resultado clave en el que se basan las pruebas para una media poblacional normal, se usó en el capítulo 7 para obtener el IC de t con una muestra: si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución normal, la variable estandarizada

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

tiene una distribución t con $n - 1$ grados de libertad (gl). Considere probar H_0: $\mu = \mu_0$ contra H_a: $\mu > \mu_0$ mediante el estadístico de prueba $T = (\overline{X} - \mu_0)/(S/\sqrt{n})$. Es decir, el estadístico de prueba resulta de estandarizar $\overline{X}$ bajo la suposición de que H_0 es cierta (usando $S/\sqrt{n}$, la desviación estándar estimada de $\overline{X}$, en vez de $\sigma/\sqrt{n}$). Cuando H_0 es verdadera, el estadístico de prueba tiene una distribución t con $n - 1$ gl. El conocimiento del estadístico de prueba cuando H_0 es verdadera (la "distribución nula") permite construir una región de rechazo, para la cual la probabilidad de error de tipo I se controla al nivel deseado. En particular, el uso del valor crítico t del extremo superior $t_{\alpha,n-1}$ para especificar la región de rechazo $t \geq t_{\alpha,n-1}$ implica que

$$\begin{aligned} P(\text{error de tipo I}) &= P(\text{se rechaza } H_0 \text{ cuando es verdadera}) \\ &= P(T \geq t_{\alpha,n-1} \text{ cuando } T \text{ tiene una distribución } t \text{ con } n - 1 \text{ gl}) \\ &= \alpha \end{aligned}$$

El estadístico de prueba es en realidad el mismo aquí que en el caso de muestras grandes pero se identifica con T para remarcar que su distribución nula es una distribución t con $n - 1$ grados de libertad en vez de la distribución normal estándar (z). La región de rechazo para la prueba t difiere de la de la prueba z sólo en que un valor crítico $t_{\alpha,n-1}$ remplaza al valor crítico $z = z_\alpha$. Comentarios similares se aplican a las alternativas para las cuales es apropiada una prueba de extremo inferior o de dos extremos.

Prueba t con una sola muestra

Hipótesis nula: H_0: $\mu = \mu_0$

Valor del estadístico de prueba: $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Hipótesis alternativa	Región de rechazo para una prueba de nivel α
H_a: $\mu > \mu_0$	$t \geq t_{\alpha,n-1}$ (de extremo superior)
H_a: $\mu < \mu_0$	$t \leq -t_{\alpha,n-1}$ (de extremo inferior)
H_a: $\mu \neq \mu_0$	ya sea $t \geq t_{\alpha/2,n-1}$ o $t \leq -t_{\alpha/2,n-1}$ (de dos extremos)

Ejemplo 8.9 Un lugar de trabajo bien diseñado y seguro puede contribuir en gran medida a incrementar la productividad. Es muy importante que a los trabajadores no se les pida llevar a cabo tareas, como levantar cosas, que excedan sus capacidades. Los datos siguientes acerca del peso máximo de elevación (PME en kg) para una frecuencia de cuatro elevaciones/min se publicó en el artículo "The Effects of Speed, Frequency, and Load on Measured Hand Forces for a Floor-to-Knuckle Lifting Task" (*Ergonomics*, 1992: 883-843); se eligió al azar a los individuos de una población de varones saludables con edades de 18 a 30 años. Suponiendo que el peso máximo de elevación tiene una distribución normal, ¿los datos siguientes indican que el peso máximo de elevación promedio muestral es mayor que 25?

25.8 36.6 26.3 21.8 27.2

Llévese a cabo una prueba con un nivel de significación de 0.05.

1. μ = PME promedio poblacional
2. H_0: $\mu = 25$
3. H_a: $\mu > 25$
4. $t = \dfrac{\bar{x} - 25}{s/\sqrt{n}}$
5. Rechace H_0 si $t \geq t_{\alpha,n-1} = t_{0.05,4} = 2.132$.
6. $\sum x_i = 137.7$ y $\sum x_i^2 = 3911.97$, de donde $\bar{x} = 27.54$, $s = 5.47$, y

$$t = \frac{27.54 - 25}{5.47/\sqrt{5}} = \frac{2.54}{2.45} = 1.04$$

El siguiente resultado de MINITAB de una solicitud para una prueba t con una muestra tiene los mismos valores calculados (el valor P se analizó en la sección 8.4).

```
Prueba de mu = 25.00 contra mu > 25.00
Variable  N   Media   Desv. Est.   EE medio    T    Valor P
PME       5   27.54      5.47        2.45     1.04   0.18
```

7. Puesto que 1.04 no cae en la región de rechazo ($1.04 < 2.132$), H_0 no se puede rechazar al nivel de significación 0.05. Es posible aún que μ sea (a lo sumo) 25. ■

Determinación de β y del tamaño de muestra El cálculo de β al valor alternativo μ' en el caso I se realizó al expresar la región de rechazo en términos de $\bar{x}$ (p. ej., $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$) y luego restar μ' para estandarizar de manera correcta. Un método equivalente requiere observar que cuando $\mu = \mu'$, la estadística de prueba $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ aún tiene una distribución normal con varianza 1, pero ahora el valor medio de Z está dado por $(\mu' - \mu_0)/(\sigma/\sqrt{n})$. Es decir, cuando $\mu = \mu'$, la estadística de prueba aún tiene una distribu-

ción normal aunque no la distribución normal estándar. Debido a esto, $\beta(\mu')$ es un área bajo la curva normal que corresponde al valor medio $(\mu' - \mu_0)/(\sigma/\sqrt{n})$ y varianza 1. Tanto α como β requieren trabajar con variables distribuidas normalmente.

El cálculo de $\beta(\mu')$ para la prueba t es mucho menos directo. Esto se debe a que la distribución de la estadística de prueba $T = (\overline{X} - \mu_0)/(S/\sqrt{n})$ es bastante complicada cuando H_0 es falsa y H_a es verdadera. Así, para una prueba de extremo superior, determinar

$$\beta(\mu') = P(T < t_{\alpha,n-1} \text{ cuando } \mu = \mu' \text{ en vez de } \mu_0)$$

requiere integrar una función de densidad muy desagradable. Esto se debe hacer de forma numérica, pero por suerte los estadísticos se tomaron el trabajo de hacerlo para pruebas t de uno y dos extremos. Los resultados se resumen en las gráficas de β que aparecen en la tabla A.17 del apéndice. Hay cuatro conjuntos de gráficas, que corresponden a pruebas de un extremo al nivel 0.05 y nivel 0.01 y pruebas de dos extremos a los mismos niveles.

Para entender cómo se usan estas gráficas, observe primero que tanto β como el tamaño de muestra necesario n en el caso I no sólo son funciones de la diferencia absoluta $|\mu_0 - \mu'|$ sino de $d = |\mu_0 - \mu'|/\sigma$. Suponga, por ejemplo, que $|\mu_0 - \mu'| = 10$. Esta desviación respecto de H_0 se detecta mucho más fácil (β más pequeña) cuando $\sigma = 2$, en cuyo caso μ_0 y μ' están separadas por cinco desviaciones estándar poblacionales, que cuando $\sigma = 10$. El hecho de que β para la prueba t dependa de d y no sólo de $|\mu_0 - \mu'|$ es infortunado, puesto que para usar las gráficas se debe tener cierta idea del valor verdadero de σ. Una suposición (grande) conservadora para σ produce un valor (grande) conservador de $\beta(\mu')$ y una estimación conservadora del tamaño de muestra necesario para α y $\beta(\mu')$ prescritas.

Una vez que se eligen la μ' alternativa y el valor de σ, se calcula d y se localiza su valor en el eje horizontal del conjunto pertinente de curvas. El valor de β es la altura de la curva con $n - 1$ gl arriba del valor de d (la interpolación visual es necesaria si $n - 1$ no es un valor para el cual aparezca la curva correspondiente), como se ilustra en la figura 8.4.

En vez de fijar n (es decir, $n - 1$ y, por lo tanto, la curva particular a partir de la cual se lee β), se podría prescribir tanto α (0.05 o 0.01 en este caso) como un valor de β para los valores de μ' y σ elegidos. Después de calcular d, se localiza el punto (d, β) en el conjunto pertinente de gráficas. La curva debajo y más cercana a este punto da $n - 1$ y, por lo tanto, n (de nuevo es necesaria la interpolación).

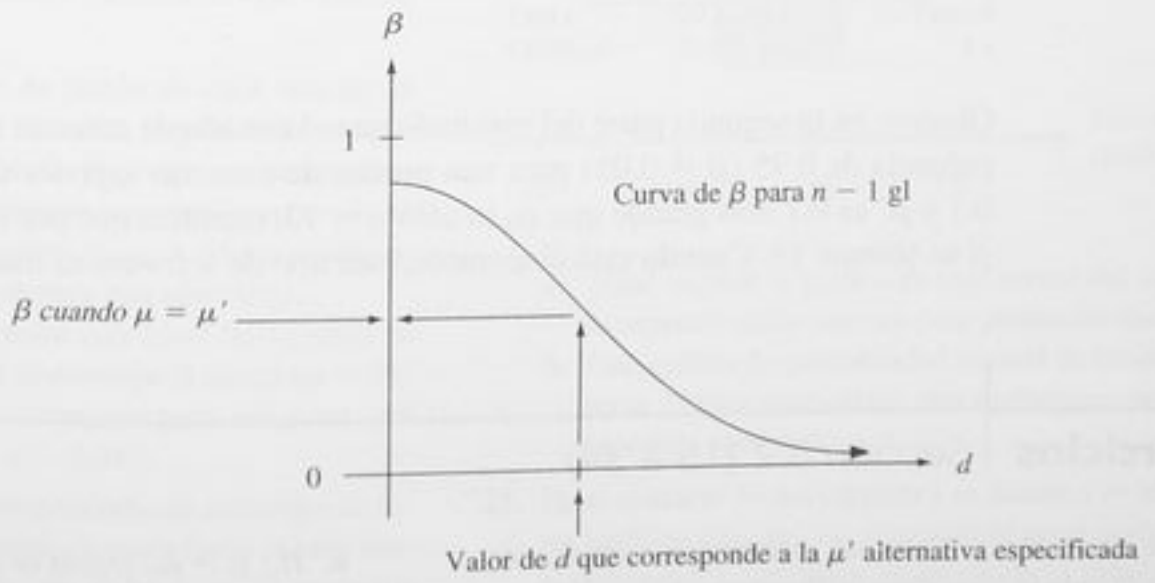


Figura 8.4 Curva de β representativa para la prueba t

Ejemplo 8.10 Se supone que la caída de tensión promedio real del colector al emisor de transistores bipolares de compuerta aislados de cierta clase es a lo sumo 2.5 volts. Un investigador selecciona una muestra de $n = 10$ transistores y usa los voltajes resultantes como base para probar H_0: $\mu = 2.5$ contra H_a: $\mu > 2.5$ por medio de una prueba t con nivel de significación $\alpha = 0.05$. Si la desviación estándar de la distribución de voltaje es $\sigma = 0.100$, ¿qué tan probable es que no se rechace H_0 cuando de hecho $\mu = 2.6$? Con $d = |2.5 - 2.6|/0.100 = 1.0$, el punto en la curva β con 9 gl para una prueba de un extremo con $\alpha = 0.05$ arriba de 1.0 tiene una altura aproximada de 0.1, por lo tanto $\beta \approx 0.1$. El investigador podría pensar que éste es un valor demasiado grande de β para tal desviación sustancial con respecto a H_0 y desea tener $\beta = 0.05$ para este valor alternativo de μ. Puesto que $d = 1.0$, se debe localizar el punto $(d, \beta) = (1.0, 0.05)$. Este punto es muy cercano a la curva de 14 gl, así que con $n = 15$, se obtiene $\alpha = 0.05$ y $\beta = 0.05$ cuando el valor de μ es 2.6 y $\sigma = 0.10$. Con un valor más grande de σ se obtendría una β más grande para esta alternativa, y un valor alternativo de μ más cercano a 2.5 daría como resultado un valor de β mayor. ∎

La mayor parte de los paquetes de software para estadística, también permiten calcular probabilidades de error de tipo II y determinar los tamaños de muestra necesarios. Como un ejemplo, se utiliza MINITAB para llevar a cabo los cálculos del ejemplo 8.10. Sus cálculos se basan en la **potencia**, que es simplemente $1 - \beta$. Se quiere que β sea pequeña, lo que equivale a pedir que la potencia de la prueba sea grande. Por ejemplo, $\beta = 0.05$ corresponde al valor de 0.95 para la potencia. A continuación se presenta el resultado de MINITAB.

```
Potencia y tamaño de muestra

Probar media = nula (contra > nula)
Calcular la potencia para media = nula + 0.1
Alpha = 0.05  Sigma = 0.1

Tamaño de
muestra   Potencia
10          0.8975
```

```
Potencia y tamaño de muestra

Prueba t con 1 muestra

Calcular media = nula (contra > nula)
Calcular la potencia para media = nula + 0.1
Alpha = 0.05  Sigma = 0.1

Tamaño de  Potencia  Potencia
muestra    objetivo     real
13           0.9500   0.9597
```

Observe en la segunda parte del resultado que el tamaño de muestra necesario para obtener una potencia de 0.95 ($\beta = 0.05$) para una prueba de extremo superior con $\alpha = 0.05$ cuando $\sigma = 0.1$ y μ' es 0.1 más grande que μ_0 es sólo $n = 13$, mientras que por inspección de las curvas de β se obtiene 15. Cuando está disponible, este tipo de *software* es más confiable que las curvas.

Ejercicios | Sección 8.2 (15 a 34)

15. Sea el estadístico Z con una distribución normal estándar cuando H_0 es verdadera. Proporcione el nivel de significación para cada una de las situaciones siguientes:
    a. H_a: $\mu > \mu_0$, región de rechazo $z \geq 1.88$
    b. H_a: $\mu < \mu_0$, región de rechazo $z \leq -2.75$
    c. H_a: $\mu \neq \mu_0$, región de rechazo $z \geq 2.88$ o $z \leq -2.88$

16. Sea el estadístico de prueba T con distribución t cuando H_0 es verdadera. Dé el nivel de significación para cada una de las situaciones siguientes:
    a. H_a: $\mu > \mu_0$, gl = 15, región de rechazo $t \geq 3.733$
    b. H_a: $\mu < \mu_0$, $n = 24$, región de rechazo $t \leq -2.500$
    c. H_a: $\mu \neq \mu_0$, $n = 31$, región de rechazo $t \geq 1.697$ o $t \leq -1.697$

17. Conteste las preguntas siguientes para el problema de las llantas del ejemplo 8.7..
    a. Si $\bar{x} = 20\,960$ y se utiliza una prueba con nivel $\alpha = 0.01$, ¿cuál es la decisión?
    b. Si se emplea una prueba con nivel 0.01, ¿cuál es $\beta(20\,500)$?
    c. Si se usa una prueba con nivel 0.01 y también se requiere que $\beta(20\,500) = 0.05$, ¿qué tamaño de muestra n es necesario?
    d. Si $\bar{x} = 20\,960$, ¿cuál es la α más pequeña a la que se puede rechazar H_0 (con base en $n = 16$)?

18. Considere de nuevo la situación del secado de pintura del ejemplo 8.2, en el que el tiempo de secado para un espécimen de prueba tiene una distribución normal con $\sigma = 9$. Las hipótesis H_0: $\mu = 75$ contra H_a: $\mu < 75$ se probarán usando una muestra aleatoria de $n = 25$ observaciones.
    a. ¿A cuántas desviaciones estándar (de $\bar{X}$) abajo del valor nulo está $\bar{x} = 72.3$?
    b. Si $\bar{x} = 72.3$, ¿cuál es la conclusión si se emplea $\alpha = 0.01$?
    c. ¿Cuál es α para el procedimiento de prueba que rechaza H_0 cuando $z \leq -2.88$?
    d. Para el procedimiento de prueba del inciso (c), ¿cuál es $\beta(70)$?
    e. Si se utiliza el procedimiento de prueba del inciso (c), ¿qué valor de n se necesita para asegurar que $\beta(70) = 0.01$?
    f. Si se emplea una prueba de nivel 0.01 con $n = 100$, ¿cuál es la probabilidad de un error de tipo I cuando $\mu = 76$?

19. Se midió la temperatura de fusión de cada una de 16 muestras de cierta marca de aceite vegetal hidrogenado, y se encontró que $\bar{x} = 94.32$. Suponga que la distribución de la temperatura de fusión es normal con $\sigma = 1.20$.
    a. Pruebe H_0: $\mu = 95$ contra H_a: $\mu \neq 95$ por medio de una prueba de dos extremos con nivel 0.01.
    b. Si se emplea una prueba con nivel 0.01, ¿cuál es $\beta(94)$, la probabilidad de error tipo II cuando $\mu = 94$?
    c. ¿Qué valor de n es necesario para asegurar que $\beta(94) = 0.1$ cuando $\alpha = 0.01$?

20. Se advierte que la duración promedio de cierto tipo de focos es de 750 horas. El precio de estos focos es muy asequible, así que un posible cliente ha decidido proseguir con una compra, a menos que se demuestre de manera concluyente que la duración promedio es más pequeña que la anunciada. Se elige una muestra aleatoria de 50 focos, se determina la duración de cada foco y se prueban las hipótesis apropiadas por medio de MINITAB, y se obtiene el resultado siguiente.

Variable	N	Media	Desv. Est.	EE de la media	Z	valor P
Duración	50	738.44	38.20	5.40	-2.14	0.016

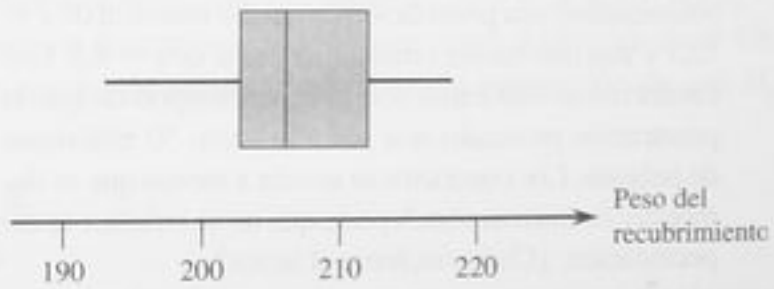
¿Qué conclusión sería apropiada para un nivel de significación de 0.05? ¿Qué nivel de significación y conclusión recomendaría?

21. Se supone que el diámetro promedio real de cojinetes de bolas de cierto tipo es 0.5 pulg. Se llevará a cabo una prueba t con una sola muestra para ver si éste es el caso. ¿Qué conclusión es apropiada en cada una de las situaciones siguientes?
    a. $n = 13$, $t = 1.6$, $\alpha = 0.05$
    b. $n = 13$, $t = -1.6$, $\alpha = 0.05$
    c. $n = 25$, $t = -2.6$, $\alpha = 0.01$
    d. $n = 25$, $t = -3.9$

22. En el artículo "The Foreman's View of Quality Control" (*Quality Engr.*, 1990: 257-280) se describe una investigación en relación con los pesos de recubrimiento para tubos grandes que resultan de un proceso de recubrimiento galvanizado. Los estándares de producción requieren un peso promedio real de 200 lb por tubo. El resumen descriptivo adjunto y la gráfica de caja son de MINITAB.

Variable	N	Media	Mediana	Media rec.	Desv Est.	EE de la media
P. recub	30	206.73	206.00	206.81	6.35	1.16

Variable	Mín	Máx	Q1	Q3
P. recub	193.00	218.00	202.75	212.00



    a. ¿Qué sugiere la gráfica de caja acerca del estado de la especificación para un peso promedio real?
    b. Una gráfica de probabilidad normal de los datos fue recta. Utilice el resultado descriptivo para probar las hipótesis apropiadas.

23. En el ejercicio 36 del capítulo 1 se dieron $n = 26$ observaciones en relación con el tiempo de evacuación (s) para trabajadores petroleros en un ejercicio simulado, del cual la media muestral y la desviación estándar muestral son 370.69 y 24.36, respectivamente. Suponga que los

investigadores tienen la creencia previa de que el tiempo de escape promedio sería a lo sumo seis minutos. ¿Los datos contradicen esta creencia previa? Suponiendo normalidad, pruebe la hipótesis apropiada usando un nivel de significación de 0.05.

24. Considere de nuevo las observaciones muestrales acerca de la viscosidad estabilizada de especímenes de asfalto presentadas en el ejercicio 46 del capítulo 1 (2781, 2900, 3013, 2856 y 2888). Suponga que para una aplicación particular, se requiere que la viscosidad promedio real sea 3000. ¿Al parecer se ha cumplido este requerimiento? Exprese y pruebe la hipótesis apropiada.

25. El porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso es 5.5. Para probar si el porcentaje promedio real es 5.5 para una determinada instalación de producción, se analizan 16 muestras obtenidas de manera independiente. Suponga que el porcentaje de SiO_2 en una muestra tiene una distribución normal con $\sigma = 0.3$ y que $\bar{x} = 5.25$.

- **a.** ¿Esto indica de manera concluyente que el porcentaje promedio real difiere de 5.5? Realice el análisis por medio de la secuencia de pasos sugeridos en el texto.
- **b.** Si el porcentaje promedio real es $\mu = 5.6$ y se emplea una prueba con nivel $\alpha = 0.01$ basada en $n = 16$, ¿cuál es la probabilidad de detectar esta desviación con respecto a H_0?
- **c.** ¿Cuál es el valor de n requerido para satisfacer $\alpha = 0.01$ y $\beta(5.6) = 0.01$?

26. Para obtener información acerca de las propiedades de resistencia a la corrosión de cierto tipo de tubería de acero, se entierran en el suelo 45 especímenes durante un periodo de dos años. Luego, se mide la penetración máxima (en milésimas de pulgada) para cada espécimen, obteniéndose una penetración promedio muestral de $\bar{x} = 52.7$ y una desviación estándar muestral de $s = 4.8$. Los conductos se fabricaron con la especificación de que la penetración promedio real sea a lo sumo 50 milésimas de pulgada. Los conductos se usarán a menos que se demuestre de manera concluyente que no se satisface la especificación. ¿Cuál concluiría el lector?

27. La prueba de impacto de muesca en V de Charpy es la base para estudiar muchos criterios de resistencia del material. Esta prueba se aplicó a 42 especímenes de cierta aleación a 110°F. La cantidad promedio muestral de expansión lateral transversal se calculó como 73.1 milésimas de pulgada, y la desviación estándar muestral fue $s = 5.9$ milésimas de pulgada. A fin de que el material sea adecuado para una aplicación particular, la cantidad promedio real de expansión debe ser menor que 75 milésimas de pulgada. No se usará la aleación a menos que de la muestra se obtenga una evidencia firme de que se satisface este criterio. Pruebe las hipótesis pertinentes con $\alpha = 0.01$ para decidir si es adecuada la aleación.

28. La cirugía menor en caballos en condiciones de campo requiere un anestésico de corto plazo confiable que produzca buena relajación muscular, cambios cardiovasculares y respiratorios mínimos, y recuperación rápida y sin problemas, con efectos secundarios mínimos, de modo que se pueda dejar solos a los caballos. En el artículo "A Field Trial of Ketamine Anesthesia in the Horse" (*Equine Vet. J.* 1984: 176-179) se informa que para una muestra de $n = 73$ caballos a los que se administró ketamina en ciertas condiciones, el tiempo promedio muestral de posición recostada lateral fue 18.86 min y la desviación estándar fue 8.6 min. ¿Estos datos indican que el tiempo promedio real de posición recostada lateral en estas condiciones es menor que 20 min? Pruebe las hipótesis apropiadas al nivel de significación 0.10.

29. Después de una cantidad fija de millas se determinó el desgaste de flecha (0.0001 pulg.) para cada una de $n = 8$ máquinas de combustión interna que tienen cobre y plomo como material de soporte, y se obtiene como resultado que $\bar{x} = 3.72$ y $s = 1.25$.

- **a.** Suponiendo que la distribución del desgaste de flecha es normal con media μ, use la prueba t al nivel 0.05 para probar H_0: $\mu = 3.50$ contra H_a: $\mu > 3.50$.
- **b.** Con $\sigma = 1.25$, ¿cuál es la probabilidad de error de tipo II $\beta(\mu')$ de la prueba para la alternativa $\mu' = 4.00$?

30. La tolerancia dietética diaria recomendada para el cinc entre varones mayores de 50 años es 15 mg/día. En el artículo "Nutrient Intakes and Dietary Patterns of Older Americans: A National Study" (*J. Gerontology*, 1992: M145-150) se proporcionan los siguientes datos de resumen acerca de la ingestión para una muestra de varones de 65 a 74 años de edad: $n = 115$, $\bar{x} = 11.3$, y $s = 6.43$. ¿Estos datos indican que la ingestión promedio diaria de cinc en la población de varones de 65 a 74 años cae por debajo de la tolerancia recomendada?

31. En un experimento diseñado para medir el tiempo necesario para que los ojos de un inspector se acostumbren a la cantidad reducida de luz necesaria para la inspección penetrante, el tiempo promedio muestral para $n = 9$ inspectores fue 6.32 s y la desviación estándar muestral fue 1.65 s. Se supuso que el tiempo de adaptación promedio fue por lo menos 7 s. Suponiendo que el tiempo de adaptación tiene una distribución normal, ¿los datos contradicen la creencia previa? Use la prueba t con $\alpha = 0.1$.

32. Se seleccionó una muestra de 12 detectores de radón de cierto tipo, y cada uno se expuso a 100 pCi/L de radón. Las lecturas resultantes fueron como sigue:

105.6	90.9	91.2	96.9	96.5	91.3
100.1	105.0	99.6	107.7	103.3	92.4

a. ¿Estos datos indican que la lectura media poblacional en estas condiciones difiere de 100? Exprese y pruebe las hipótesis apropiadas con $\alpha = 0.05$.
b. Suponga que antes del experimento, se asumió un valor de $\sigma = 7.5$. ¿Cuántas determinaciones habrían sido apropiadas para obtener $\beta = 0.10$ para la alternativa $\mu = 95$?

33. Demuestre que para cualquier $\Delta > 0$, cuando la distribución poblacional es normal y se conoce el valor de σ, la prueba de dos extremos satisface $\beta(\mu_0 - \Delta) = \beta(\mu_0 + \Delta)$, de modo que $\beta(\mu')$ es simétrica respecto a μ_0.
34. Para un valor alternativo fijo μ', demuestre que $\beta(\mu') \to 0$ cuando $n \to \infty$ ya sea para una prueba z de un extremo o de dos extremos en el caso de una distribución poblacional normal con σ conocida.

8.3 Pruebas relacionadas con una proporción poblacional

Sea p la proporción de individuos u objetos en una población que poseen una propiedad específica (p. ej., automóviles con transmisiones manuales o fumadores que fuman un cigarro con filtro). Si se marca a un individuo u objeto con esa propiedad como un éxito (S), entonces p es la proporción poblacional de éxitos. Las pruebas relacionadas con p se basarán en una muestra aleatoria de tamaño n de la población. Siempre que n sea pequeña en relación con la población, X (el número de éxitos (S) en la muestra) tiene (aproximadamente) una distribución binomial. Además, si n es grande, tanto X como el estimador $\hat{p} = X/n$ tienen aproximadamente una distribución normal. Primero se consideran pruebas con muestras grandes con base en este hecho, y luego se vuelve al caso de muestras pequeñas que utiliza directamente la distribución binomial.

Pruebas con muestras grandes

Las pruebas con muestras grandes en relación con p son un caso especial de los procedimientos más generales con muestras grandes para un parámetro θ. Sea $\hat{\theta}$ un estimador de θ que es (por lo menos en forma aproximada) insesgado y tiene aproximadamente una distribución normal. La hipótesis nula tiene la forma H_0: $\theta = \theta_0$, donde θ_0 denota un número (el valor nulo) apropiado al contexto del problema. Suponga que cuando H_0 es verdadera, la desviación estándar de $\hat{\theta}$, $\sigma_{\hat{\theta}}$ no tiene que ver con parámetros desconocidos. Por ejemplo, si $\theta = \mu$ y $\hat{\theta} = \overline{X}$, $\sigma_{\hat{\theta}} = \sigma_{\overline{X}} = \sigma/\sqrt{n}$, en la que intervienen parámetros desconocidos sólo si se conoce el valor de σ. Un estadístico de prueba de muestra grande resulta de estandarizar $\hat{\theta}$ bajo la suposición de que H_0 es verdadera (de modo que $E(\hat{\theta}) = \theta_0$):

$$\text{Estadístico de prueba: } Z = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}}$$

Si la hipótesis alternativa es H_a: $\theta > \theta_0$, una prueba de extremo superior cuyo nivel de significación es aproximadamente α, se especifica mediante la región de rechazo $z \geq z_\alpha$ Las otras dos alternativas, H_a: $\theta < \theta_0$ y H_a: $\theta \neq \theta_0$, se prueban por medio de una prueba z de extremo superior y una prueba de dos extremos, respectivamente.

En el caso $\theta = p$, $\sigma_{\hat{\theta}}$ no requiere ningún parámetro desconocido cuando H_0 es verdadera, pero esto es poco común. Cuando $\sigma_{\hat{\theta}}$ no tiene que ver con parámetros desconocidos, con frecuencia es posible usar una desviación estándar estimada $S_{\hat{\theta}}$ en lugar de $\sigma_{\hat{\theta}}$ y aún lograr que Z tenga más o menos una distribución normal cuando H_0 es verdadera (porque cuando n es grande, $s_{\hat{\theta}} \approx \sigma_{\hat{\theta}}$ para la mayor parte de las muestras). La prueba con muestras grandes de la sección anterior proporciona un ejemplo de esto: debido a que por lo común se desconoce σ, se utiliza $s_{\hat{\theta}} = s_{\overline{X}} = s/\sqrt{n}$ en lugar de $\sigma/\sqrt{n}$ en el denominador de z.

El estimador $\hat{p} = X/n$ es insesgado ($E(\hat{p}) = p$), tiene una distribución normal aproximada, y su desviación estándar es $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$. Estos hechos se utilizaron en la sección 7.2 para obtener un intervalo de confianza para p. Cuando H_0 es verdadera, $E(\hat{p}) = p_0$ y $\sigma_{\hat{p}} = \sqrt{p_0(1-p_0)/n}$, por lo tanto en $\sigma_{\hat{p}}$ no interviene ningún parámetro desconocido. Entonces, se deduce que cuando n es grande y H_0 es verdadera, el estadístico de prueba

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

tiene aproximadamente una distribución normal estándar. Si la hipótesis alternativa es H_a: $p > p_0$ y se usa la región de rechazo de extremo superior $z \geq z_\alpha$, entonces

$$\begin{aligned} P(\text{error de tipo I}) &= P(\text{se rechaza } H_0 \text{ cuando es verdadera}) \\ &= P(Z \geq z_\alpha \text{ cuando } Z \text{ tiene aproximadamente una distribución} \\ &\qquad \text{normal estándar}) \approx \alpha \end{aligned}$$

Así, el nivel deseado de significación α se obtiene por medio del valor crítico que capta el área α en el extremo superior de la curva z. Las regiones de rechazo para las otras dos hipótesis alternativas, de extremo inferior para H_a: $p < p_0$ y de dos extremos para H_a: $p \neq p_0$, se justifican de manera análoga.

Hipótesis nula: H_0: $p = p_0$

Valor de la estadística de prueba: $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

Hipótesis alternativa	Región de rechazo
H_a: $p > p_0$	$z \geq z_\alpha$ (de extremo superior)
H_a: $p < p_0$	$z \leq -z_\alpha$ (de extremo inferior)
H_a: $p \neq p_0$	ya sea $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$ (de dos extremos)

Estos procedimientos de prueba son válidos siempre y cuando $np_0 \geq 10$ y $n(1-p_0) \geq 10$.

Ejemplo 8.11 Información reciente hace pensar que la obesidad es un problema creciente en América entre todos los grupos de edad. La Associated Press (9 de octubre de 2002) informó que en una muestra de 4115 adultos se encontró que 1276 individuos eran obesos (un índice de masa corporal mayor que 30; este índice es una medida de peso en relación con la estatura). En una encuesta de 1998 basada en la propia evaluación de las personas, se encontró que 20% de los adultos americanos se consideraban obesos. ¿Los datos más recientes indican que la proporción real de adultos que son obesos es más de 1.5 veces el porcentaje de la encuesta de autoevaluación? Para responder esto se lleva a cabo una prueba de hipótesis con un nivel de significación de 0.10.

1. p = proporción de los americanos adultos que son obesos.
2. Decir que el porcentaje actual es 1.5 veces el porcentaje de autoevaluación, equivale a la aseveración de que el porcentaje actual es 30%, de donde se tiene la hipótesis nula H_0: $p = 0.30$.
3. La frase "más de" en la descripción de problema implica que la hipótesis alternativa es H_a: $p > 0.30$

4. Puesto que $np_0 = 4115(0.3) > 10$ y $nq_0 = 4115(0.7) > 10$, se puede usar la prueba z para muestras grandes. El valor del estadístico de prueba es

$$z = (\hat{p} - 0.3)/\sqrt{(0.3)(0.7)/n}$$

5. La forma de H_a implica que es apropiada una prueba de extremo superior: rechazar H_0 si $z > z_{0.10} = 1.28$.

6. $\hat{p} = 1276/4115 = 0.310$, de donde $z = (0.310 - 0.3)/\sqrt{(0.3)(0.7)/4115} = 0.010/0.0071 = 1.40$.

7. Puesto que 1.40 es mayor que el valor crítico 1.28, z se ubica en la región de rechazo. Esto justifica rechazar la hipótesis nula. Con un nivel de significación de 0.10, al parecer más de 30% de los americanos adultos son obesos. ■

Determinación de β y del tamaño de la muestra Cuando H_0 es verdadera, el estadístico de prueba Z tiene aproximadamente una distribución normal estándar. Ahora suponga que H_0 *no* es verdadera y que $p = p'$. Entonces Z aún tiene más o menos una distribución normal (debido a que es una función lineal de $\hat{p}$), pero su valor medio y varianza ya no son 0 y 1, respectivamente. En cambio,

$$E(Z) = \frac{p' - p_0}{\sqrt{p_0(1 - p_0)/n}} \qquad V(Z) = \frac{p'(1 - p')/n}{p_0(1 - p_0)/n}$$

La probabilidad de un error de tipo II para una prueba de extremo superior es $\beta(p') = P(Z < z_\alpha$ cuando $p = p')$. Esto se puede calcular con la media y la varianza proporcionada para estandarizar y luego aplicar la fda normal estándar. Además, si se desea que la prueba de nivel α también tenga $\beta(p') = \beta$ para un valor especificado de β, esta ecuación se puede resolver para la n necesaria como la sección 8.2. Las expresiones generales para $\beta(p')$ y n se dan en el cuadro siguiente.

Hipótesis alternativa	$\beta(p')$
H_a: $p > p_0$	$\Phi\left[\dfrac{p_0 - p' + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$
H_a: $p < p_0$	$1 - \Phi\left[\dfrac{p_0 - p' - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$
H_a: $p \neq p_0$	$\Phi\left[\dfrac{p_0 - p' + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right] - \Phi\left[\dfrac{p_0 - p' - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$

El tamaño de muestra n para el cual la prueba de nivel α también satisface $\beta(p') = \beta$ es

$$n = \begin{cases} \left[\dfrac{z_\alpha\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p'(1-p')}}{p' - p_0}\right]^2 & \text{prueba de un extremo} \\ \left[\dfrac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p'(1-p')}}{p' - p_0}\right]^2 & \text{prueba de dos extremos (una solución aproximada)} \end{cases}$$

Ejemplo 8.12 Un servicio de entrega de paquetes anuncia que por lo menos 90% de los paquetes llevados a su oficina a las 9 a.m. para entrega en la misma ciudad se entregan ese mismo día al mediodía. Sea p la proporción real de los paquetes entregados según el anuncio, y considere la hipótesis H_0: $p = 0.9$ contra H_a: $p < 0.9$. Si sólo 80% de los paquetes se entregan según el anuncio, ¿cuán probable es que una prueba al nivel 0.01 basada en $n = 225$ paquetes detecte tal desviación con respecto a H_0? ¿Cuál debe ser el tamaño de la muestra para asegurar que $\beta(0.8) = 0.01$? Con $\alpha = 0.01$, $p_0 = 0.9$, $p' = 0.8$, y $n = 225$,

$$\beta(0.8) = 1 - \Phi\left(\frac{0.9 - 0.8 - 2.33\sqrt{(0.9)(0.1)/225}}{\sqrt{(0.8)(0.2)/225}}\right)$$
$$= 1 - \Phi(2.00) = 0.0228$$

Así, la probabilidad de que H_0 sea rechazada es 0.9772, usando la prueba cuando $p = 0.8$. Alrededor de 98% de las muestras dan como resultado el rechazo correcto de H_0.

Con $z_\alpha = z_\beta = 2.33$ en la fórmula del tamaño de muestra, se obtiene

$$n = \left[\frac{2.33\sqrt{(0.9)(0.1)} + 2.33\sqrt{(0.8)(0.2)}}{0.8 - 0.9}\right]^2 \approx 266$$

■

Pruebas con muestras pequeñas

Los procedimientos de prueba cuando el tamaño de muestra n es pequeño se basan directamente en la distribución binomial, y no en la aproximación normal. Considere la hipótesis alternativa H_a: $p > p_0$ y de nuevo sea X el número de éxitos en la muestra. Entonces, X es el estadístico de prueba, y la región de rechazo de extremo superior tiene la forma $x \geq c$. Cuando H_0 es verdadera, X tiene una distribución binomial con parámetros n y p_0, por lo tanto

$$\begin{aligned} P(\text{error de tipo I}) &= P(\text{se rechaza } H_0 \text{ cuando es verdadera}) \\ &= P(X \geq c \text{ cuando } X \sim \text{Bin}(n, p_0)) \\ &= 1 - P(X \leq c - 1 \text{ cuando } X \sim \text{Bin}(n, p_0)) \\ &= 1 - B(c - 1; n, p_0) \end{aligned}$$

Cuando disminuye el valor crítico c, se incluyen más valores x en la región de rechazo y se incrementa P(error de tipo I). Debido a que X tiene una distribución de probabilidad discreta, por lo común no es posible hallar un valor de c para el cual P(error de tipo I) sea exactamente el nivel de significación deseado α (p. ej., 0.05 o 0.01). En cambio, se utiliza la región de rechazo más grande de la forma $\{c, c + 1, \ldots, n\}$ que satisface $1 - B(c - 1; n, p_0) \leq \alpha$.

Denótese con p' un valor alternativo de p ($p' > p_0$). Cuando $p = p'$, $X \sim \text{Bin}(n, p')$, por lo tanto

$$\begin{aligned} \beta(p') &= P(\text{error tipo II cuando } p = p') \\ &= P(X < c \text{ cuando } X \sim \text{Bin}(n, p')) = B(c - 1; n, p') \end{aligned}$$

Es decir, $\beta(p')$ es el resultado de un cálculo directo de probabilidad binomial. El tamaño de muestra n necesario para asegurar que una prueba de nivel α también ha especificado β a un valor alternativo particular p' se debe determinar por ensayo y error usando la fda binomial.

Los procedimientos de prueba para H_a: $p < p_0$ y para H_a: $p \neq p_0$ se construyen de manera similar. En el primer caso, la región de rechazo apropiada tiene la forma $x \leq c$ (una prueba de extremo inferior). El valor crítico c es el número más grande que satisface $B(c; n, p_0) \leq \alpha$. La región de rechazo cuando la hipótesis alternativa es H_a: $p \neq p_0$ consiste en valores grandes y pequeños de x.

Ejemplo 8.13 Un fabricante de plásticos desarrolló un nuevo tipo de recipiente de plástico para la basura, y propone vender el producto con una garantía incondicional de seis años. Para ver la factibilidad económica de esto, se someten a una prueba de duración acelerada para simular seis años de uso. La garantía propuesta se modificará sólo si los datos de la muestra indican sin lugar a dudas que menos de 90% de los recipientes durarían seis años. Sea p la proporción de los recipientes que pasan la prueba acelerada. Las hipótesis pertinentes son H_0: $p = 0.9$ contra H_a: $p < 0.9$. Una decisión se basará en la estadística de prueba X, el número entre los 20 recipientes que pasan la prueba. Si el nivel deseado de significación es $\alpha = 0.05$, c debe satisfacer $B(c; 20, 0.9) \leq 0.05$. De la tabla A.1 del apéndice, $B(15; 20, 0.9) = 0.043$, mientras que $B(16; 20, 0.9) = 0.133$. Por lo tanto, la región de rechazo apropiada es $x \leq 15$. Si la prueba acelerada da como resultado $x = 14$, se rechazaría H_0 en favor de H_a, lo que requeriría una modificación de la garantía propuesta. La probabilidad de error de tipo II para el valor alternativo $p' = 0.8$ es

$$\begin{aligned}\beta(0.8) &= P(\text{no se rechaza } H_0 \text{ cuando } X \sim \text{Bin}(20, 0.8)) \\ &= P(X \geq 16 \text{ cuando } X \sim \text{Bin}(20, 0.8)) \\ &= 1 - B(15; 20, 0.8) = 1 - 0.370 = 0.630\end{aligned}$$

Es decir, cuando $p = 0.8$, 63% de las muestras que consisten en $n = 20$ recipientes darían como resultado que no se rechace H_0 lo cual sería incorrecto. Esta probabilidad de error es alta, porque 20 es un tamaño de muestra pequeño y $p' = 0.8$ está cerca del valor nulo $p_0 = 0.9$. ■

Ejercicios | Sección 8.3 (35 a 43)

35. Los registros estatales del Departamento de Vehículos Motorizados indican que de los vehículos sometidos a una prueba de verificación de emisiones el año anterior, 70% pasaron en el primer intento. En una muestra aleatoria de 200 automóviles probados en cierto condado durante el año actual, se encuentra que 124 pasaron en la prueba inicial. ¿Sugiere esto que la proporción real para este condado durante el año actual difiere de la proporción anterior a nivel estatal? Pruebe las hipótesis pertinentes usando $\alpha = 0.05$.

36. Un fabricante de baterías de níquel-hidrógeno selecciona al azar 100 placas de níquel para celdas de prueba, las hace pasar por un ciclo cierto número de veces, y determina que 14 de las placas se han ampollado.

a. ¿Proporciona lo anterior una evidencia convincente para concluir que más de 10% de las placas se ampollan bajo estas circunstancias? Enuncie y pruebe las hipótesis adecuadas con un nivel de significación de 0.05. Para llegar a su conclusión, ¿qué tipo de error podría haber cometido?

b. Si en realidad 15% de las placas se ampollan en estas circunstancias, y si se usa un tamaño de muestra $n = 100$, ¿qué tan probable es que no se rechace la hipótesis nula del inciso (a) con la prueba a nivel 0.05? Conteste esta pregunta para una muestra de tamaño 200.

c. ¿Cuántas placas habría que probar para que $\beta(0.15) = 0.10$ para la prueba del inciso (a)?

37. Una muestra aleatoria de 150 donaciones recientes en cierto banco de sangre, revela que 82 eran sangre tipo A. ¿Sugiere esto que el porcentaje real de donadores tipo A difiere de 40%, el porcentaje de la población con sangre tipo A? Lleve a cabo una prueba de las hipótesis adecuadas con un nivel de significación de 0.01. ¿Habría sido distinta su conclusión si se hubiera usado un nivel de significación de 0.05?

38. Una biblioteca universitaria realiza por lo común una vez al año un inventario completo de los libros. Debido a nuevas reglas de colocación de los libros establecidas el año anterior, el bibliotecario principal considera que es posible ahorrar dinero si se pospone el inventario. El bibliotecario decide seleccionar 1000 libros al azar de la colección de la biblioteca para ver qué pasa con ellos. Si la evidencia indica sin lugar a dudas que la proporción real de libros mal ordenados o no localizables es menor de 0.02, entonces se pospondrá el inventario.

a. Entre los 1000 libros seleccionados para la prueba, 15 estaban mal colocados o no se pudieron localizar. Pruebe las hipótesis pertinentes y asesore al bibliotecario acerca de lo que debe hacer (utilice $\alpha = 0.05$).

b. Si la verdadera proporción de libros mal colocados y perdidos es en realidad 0.01, ¿cuál es la probabi-

lidad de que se realice el inventario (innecesariamente)?

c. Si la proporción real es 0.05, ¿cuál es la probabilidad de que se posponga el inventario?

39. En el artículo "Statistical Evidence of Discrimination" (*J. Amer. Stat. Assoc.*, 1982: 773-783), se analiza el caso judicial *Swain v. Alabama* (1965), en el cual se dijo que había discriminación contra negros en la selección del gran jurado. Según los datos del censo, 25% de los elegibles para prestar servicio como gran jurado era de raza negra, pero en una muestra aleatoria de 1 050 convocados sólo 177 fueron negros. ¿Se puede concluir a partir de los datos que hay discriminación? Utilice una prueba de nivel 0.01.

40. Una línea aérea ha desarrollado un plan para un club del viajero ejecutivo, con base en la suposición de que 5% de sus clientes actuales reunieron las condiciones para obtener una membresía. En una muestra aleatoria de 500 clientes, 40 reunieron las condiciones.

a. Con estos datos, pruebe al nivel 0.01 la hipótesis nula de que la hipótesis de la compañía es correcta contra la alternativa de que no es correcta.

b. ¿Cuál es la probabilidad de que cuando utilice la prueba del inciso (a), se juzgue correcta la premisa de la compañía cuando de hecho 10% de los clientes actuales reúnen las condiciones para obtener la membresía?

41. A cada individuo de un grupo de 20 jugadores de tenis de nivel intermedio se le dan dos raquetas: una con cuerdas de *nylon* y la otra con cuerdas de tripa sintética. Después de jugar varias semanas con las dos raquetas, se le pedirá a cada jugador que opine acerca de cuál de los dos tipos de cuerda prefiere. Sea p la proporción de los jugadores que prefieren la cuerda de tripa en lugar de la de *nylon*, y sea X la cantidad de jugadores en la muestra que prefieren la cuerda de tripa. Debido a que las cuerdas de tripa son más costosas, considere la hipótesis nula de que a lo sumo 50% de los jugadores prefiere la cuerda de tripa. Esto se simplifica a H_0: $p = 0.5$, con la intención de rechazar H_0 sólo si la evidencia muestral favorece sin lugar a dudas las cuerdas de tripa.

a. ¿Cuál de las regiones de rechazo {15, 16, 17, 18, 19, 20}, {0, 1, 2, 3, 4, 5} o {0, 1, 2, 3, 17, 18, 19, 20} es la más apropiada, y por qué las otras dos son inapropiadas?

b. ¿Cuál es la probabilidad de un error tipo I para la región seleccionada del inciso (a)? ¿Especifica la región una prueba de nivel 0.05? ¿Es la mejor prueba de nivel 0.05?

c. Si 60% de los tenistas prefieren cuerdas de tripa, calcule la probabilidad de error de tipo II usando la región apropiada del inciso (a). Repita el cálculo si 80% de los tenistas prefieren las cuerdas de tripa.

d. Si 13 de los 20 jugadores prefieren cuerdas de tripa, ¿se debe rechazar H_0 si se utiliza un nivel de significación de 0.10?

42. Un fabricante de accesorios para plomería elaboró un nuevo tipo de grifo sin arandela. Sea $p = P$(un grifo de este tipo seleccionado al azar presenta una fuga antes de dos años de uso normal). El fabricante ha decidido proceder con la producción a menos que se determine que p es demasiado grande; el valor límite aceptable de p se especifica como 0.10. El fabricante decide someter n de estos grifos a una prueba acelerada (aproximadamente dos años de uso normal). Con X = el número entre los n grifos que presentan fuga antes que concluya la prueba, la producción se iniciará a menos que la X observada sea demasiado grande. Se decide que si $p = 0.10$, la probabilidad de no proseguir debe ser a lo sumo 0.10, mientras que si $p = 0.30$ la probabilidad de proseguir debe ser a lo sumo 0.10. ¿Se puede utilizar $n = 10$? ¿$n = 20$? ¿$n = 25$? ¿Cuál es la región de rechazo adecuada para la n seleccionada, y cuáles son las probabilidades de error reales al utilizar esta región?

43. Según los científicos, los robots jugarán un papel esencial en las fábricas en las décadas siguientes. Suponga que en un experimento para determinar si es factible el uso de robots para trenzar cables de computadora, se empleó un robot para ensamblar 500 cables. Se examinaron los cables y encontraron 15 defectuosos. Si los ensambladores humanos tienen una tasa de cables defectuosos de 0.035 (3.5%), ¿estos datos apoyan la hipótesis de que la proporción de cables defectuosos es menor para robots que para humanos? Utilice un nivel de significación de 0.01.

8.4 Valores *P*

Una manera de informar el resultado de un análisis de prueba de hipótesis, es decir simplemente si se rechazó la hipótesis nula al nivel de significación especificado. Así, un investigador podría expresar que H_0 fue rechazada al nivel de significación 0.05, o bien que el uso de una prueba de nivel 0.01 dio como resultado no rechazar H_0. Este tipo de enunciado es un poco inadecuado, porque no dice nada acerca de si el valor calculado del estadístico de prueba apenas

cae en la región de rechazo, o si excede el valor crítico por una gran cantidad. Una dificultad relacionada es que este tipo de informe, impone el nivel de significación especificado en otros que toman decisiones. En muchas situaciones de decisión, los individuos podrían tener puntos de vista distintos en relación con las consecuencias de un error de tipo I o tipo II. Entonces cada individuo desearía seleccionar su nivel de significación, algunos elegirían $\alpha = 0.05$, otros 0.01, etcétera y, de acuerdo con esto, llegarían a una conclusión. Esto podría dar como resultado que algunos individuos rechacen H_0, mientras que otros concluirían que los datos no muestran una contradicción de H_0 suficientemente firme para justificar su rechazo.

Ejemplo 8.14 Se sabe que el tiempo promedio real en relación con el alivio inicial del dolor para un calmante popular es 10 minutos. Represente con μ el tiempo promedio real respecto al alivio para un calmante recién desarrollado de una compañía. La compañía desea producir y comercializar este calmante sólo si proporciona un alivio más rápido que el calmante más popular, así que desea probar H_0: $\mu = 10$ contra H_a: $\mu < 10$. Sólo si la evidencia experimental da lugar al rechazo de H_0 se introducirá el nuevo calmante. Después de ponderar la gravedad relativa para cada tipo de error, se debe estar de acuerdo en un nivel de significación y tomar una decisión a ese nivel, rechazar H_0 e introducir el calmante o no.

Suponga que se introdujo el nuevo calmante. La compañía respalda su afirmación de alivio más rápido al expresar que, con base en un análisis de datos experimentales, H_0: $\mu = 10$ se rechazó en favor de H_a: $\mu < 10$ con nivel de significación $\alpha = 0.10$. Los individuos que observen un cambio con respecto a este nuevo calmante por supuesto querrán sacar su propia conclusión acerca de la validez de la afirmación. Los individuos satisfechos con el calmante popular verían un error de tipo I (y llegarían a la conclusión de que el nuevo producto proporciona alivio más rápido cuando en realidad no es así), tan grave que desearían usar $\alpha = 0.05$, 0.01 o incluso niveles más pequeños. Infortunadamente, la naturaleza de la afirmación de la compañía evita que quien toma decisiones llegue a una conclusión a ese nivel. La compañía ha impuesto a los demás su propia elección de nivel de significación. El informe podría haberse hecho de manera que hubiera flexibilidad para que cada individuo obtenga una conclusión a un nivel α elegido de manera personal. ■

Un *valor P* lleva mucha información acerca de la fortaleza de la evidencia contra H_0 y permite que quien toma decisiones llegue a una conclusión a algún nivel α especificado. Antes de dar una definición general, considere cómo la conclusión en un problema de prueba de hipótesis depende del nivel α seleccionado.

Ejemplo 8.15 El problema del contenido de nicotina analizado en el ejemplo 8.5 requiere probar H_0: $\mu = 1.5$ contra H_a: $\mu > 1.5$. Como resultado de la desigualdad en H_a, la región de rechazo es de extremo superior, y se rechaza H_0 si $z \geq z_\alpha$. Supóngase que $z = 2.10$. En la tabla siguiente se muestra la región de rechazo para cada uno de cuatro valores de α diferentes junto con la conclusión resultante.

Nivel de significación α	Región de rechazo	Conclusión
0.05	$z \geq 1.645$	Rechazar H_0
0.025	$z \geq 1.96$	Rechazar H_0
0.01	$z \geq 2.33$	No rechazar H_0
0.005	$z \geq 2.58$	No rechazar H_0

Para un valor de α relativamente grande, el valor crítico de z, z_α no está muy alejado en el extremo superior; 2.10 excede el valor crítico y, por lo tanto, se rechaza H_0. Sin embargo, cuando disminuye α, se incrementa el valor crítico. Para α pequeña, el valor crítico z es grande, 2.10 es menor que z_α, y no se rechaza H_0.

Recuerde que para una prueba z de extremo superior, α es el área bajo la curva z a la derecha del valor crítico z_α. Es decir, una vez que se especificó α, se elige el valor crítico para captar el área α del extremo superior. En la tabla A.3 del apéndice se muestra que el área a la derecha de 2.10 es 0.0179. Usar un valor de α mayor que 0.0179 corresponde a $z_\alpha < 2.10$. Una α menor 0.0179 requiere usar un valor crítico z que pase de 2.10. Por consiguiente, la decisión a un nivel α particular depende de cómo se compare la α seleccionada con el área del extremo que capta la z calculada. Esto se ilustra en la figura 8.5. Observe en particular que 0.0179, el área captada del extremo, el menor nivel α al que se rechazaría H_0, porque usar alguna α más pequeña da como resultado un valor crítico z que pasa de 2.10, así que 2.10 no está en la región de rechazo.

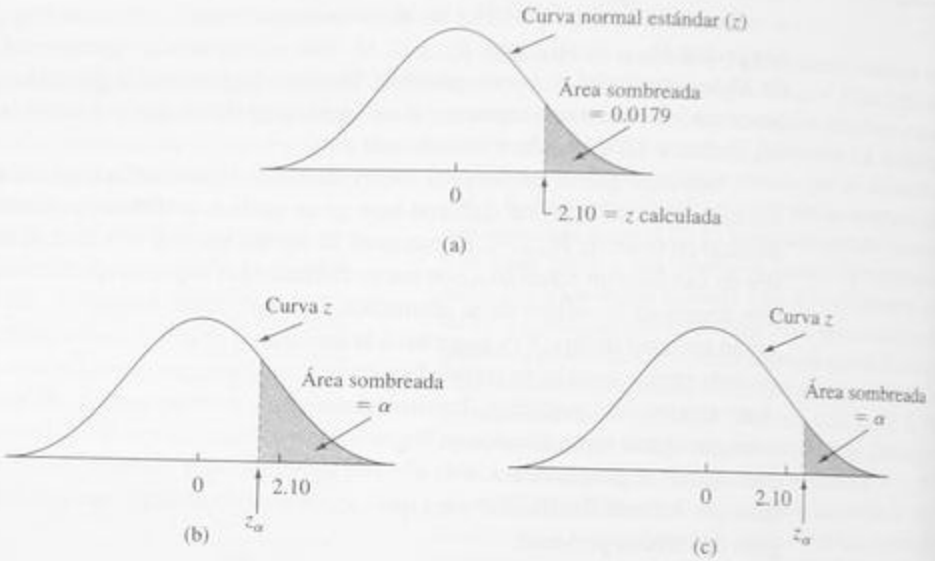


Figura 8.5 Relación entre α y el área del extremo que capta la z calculada: (a) área del extremo que capta la z calculada; (b) cuando $\alpha > 0.0179$, $z_\alpha < 2.10$ y se rechaza H_0; (c) cuando $\alpha < 0.0179$, $z_\alpha > 2.10$ y no se rechaza H_0 ■

En general, suponga que se determinó la distribución de probabilidad de un estadístico de prueba cuando H_0 es verdadera. Luego, para un valor de α especificado, se determina la región de rechazo al hallar un valor crítico o valores que captan el área α del extremo (superior, inferior o de ambos extremos, cualquiera que sea apropiada) bajo la curva de distribución de probabilidad. La α más pequeña para la cual se rechazaría H_0 es el área del extremo que capta el valor calculado de la estadística de prueba. Esta α menor es el valor P.

DEFINICIÓN

El **valor *P*** (o *nivel de significación observado*) es el nivel de significación más pequeño al cual se rechazaría H_0 cuando se emplea un procedimiento de prueba específico en un determinado conjunto de datos. Una vez determinado el valor P, la conclusión a algún nivel α particular resulta de comparar el valor P con α:

1. Valor $P \leq \alpha \Rightarrow$ rechazar H_0 al nivel α.
2. Valor $P > \alpha \Rightarrow$ no rechazar H_0 al nivel α.

De ordinario, se llama *significativos* a los datos cuando se rechaza H_0 y *no significativos* en caso contrario. Entonces el valor P es el nivel menor al cual son significativos los datos. Una forma fácil de visualizar la comparación del valor P con la α elegida es trazar un diagrama como el de la figura 8.6. El cálculo del valor P depende de si la prueba es de extremo superior, inferior o de dos extremos. Sin embargo, una vez calculado, la comparación con α no depende de qué tipo de prueba se utilice.

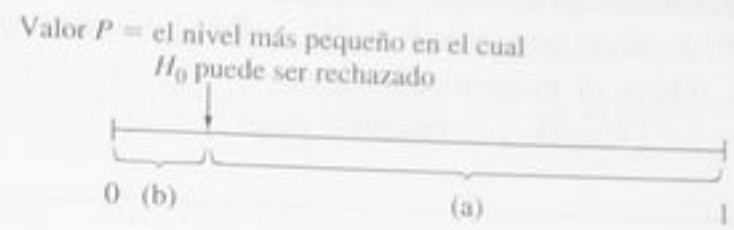


Figura 8.6 Comparación de α con el valor P: (a) rechazar H_0 cuando α se ubica aquí; (b) no rechazar H_0 cuando α se encuentra aquí

Ejemplo 8.16 (Continuación del ejemplo 8.14)

Suponga que cuando se analizan los datos de un experimento relacionados con el nuevo calmante, el valor P para probar H_0: $\mu = 10$ contra H_a: $\mu < 10$ se calculó como 0.0384. Puesto que $\alpha = 0.05$ es mayor que el valor P (0.05 se ubica en el intervalo (a) de la figura 8.6), alguien que realice la prueba al nivel 0.05 rechazaría H_0. Sin embargo, al nivel 0.01, H_0 no sería rechazada porque 0.01 es más pequeño que el nivel menor (0.0384) al cual se puede rechazar H_0. ■

Los paquetes de software para estadística más utilizados incluyen de forma automática un valor P cuando se lleva a cabo un análisis de prueba de hipótesis. Así, se puede sacar una conclusión directamente del resultado, sin referencia a la tabla de valores críticos.

La siguiente es una definición alternativa útil equivalente a la que se proporcionó antes.

DEFINICIÓN

El **valor *P*** es la probabilidad, calculada suponiendo que H_0 es verdadera, de obtener un valor de la estadística de prueba por lo menos tan contradictorio para H_0 como el valor que en realidad resultó. Mientras más pequeño sea el valor P, más contradictorios son los datos para H_0.

Así, si $z = 2.10$ para una prueba z de extremo superior, valor $P = P(Z \geq 2.10$, cuando H_0 es verdadera$) = 1 - \Phi(2.10) = 0.0179$, como antes. Tenga cuidado el valor P no es la probabilidad de que H_0 sea verdadera ¡tampoco es una probabilidad de error!

Valores *P* para pruebas *z*

El valor P para una prueba z (una basada en un estadístico de prueba cuya distribución cuando H_0 es verdadera es por lo menos aproximadamente normal estándar) se determina con facilidad a partir de la información de la tabla A.3 del apéndice. Considere una prueba de extremo superior y denote con z el valor calculado de la estadística de prueba Z. La hipótesis nula se rechaza si $z \geq z_\alpha$, y el valor P es la α más pequeña para la cual éste es el caso. Puesto que z_α se

incrementa cuando disminuye α, el valor P es el valor de α para el cual $z = z_\alpha$. Es decir, el valor P es el área que capta el valor z calculado en el extremo superior de la curva normal estándar. El área acumulada correspondiente es $\Phi(z)$, así que en este caso el valor $P = 1 - \Phi(z)$.

Un argumento análogo para una prueba de extremo inferior, muestra que el valor P es el área que capta el valor z calculado en el extremo inferior de la curva normal estándar. Se debe tener más cuidado en el caso de la prueba de dos extremos. Primero suponga que z es positiva. Entonces el valor P es el valor de α que satisface $z = z_{\alpha/2}$ (es decir, z calculada = valor crítico del extremo superior). Esto dice que el área captada en el extremo superior es un medio del valor P, de modo que valor $P = 2[1 - \Phi(z)]$. Si z es negativo, el valor P es la α para la cual $z = -z_{\alpha/2}$, o bien, de manera equivalente, $-z = z_{\alpha/2}$, por lo tanto valor $P = 2[1 - \Phi(-z)]$. Puesto que $-z = |z|$ cuando z es negativa, valor $P = 2[1 - \Phi(|z|)]$ para z positiva o negativa.

$$\text{Valor } P: \quad P = \begin{cases} 1 - \Phi(z) & \text{para una prueba de extremo superior} \\ \Phi(z) & \text{para una prueba de extremo inferior} \\ 2[1 - \Phi(|z|)] & \text{para una prueba de dos extremos} \end{cases}$$

Cada uno de éstos es la probabilidad de obtener un valor por lo menos tan extremo como el que se obtuvo (suponiendo que H_0 es verdadera). Estos tres casos se ilustran en la figura 8.7.

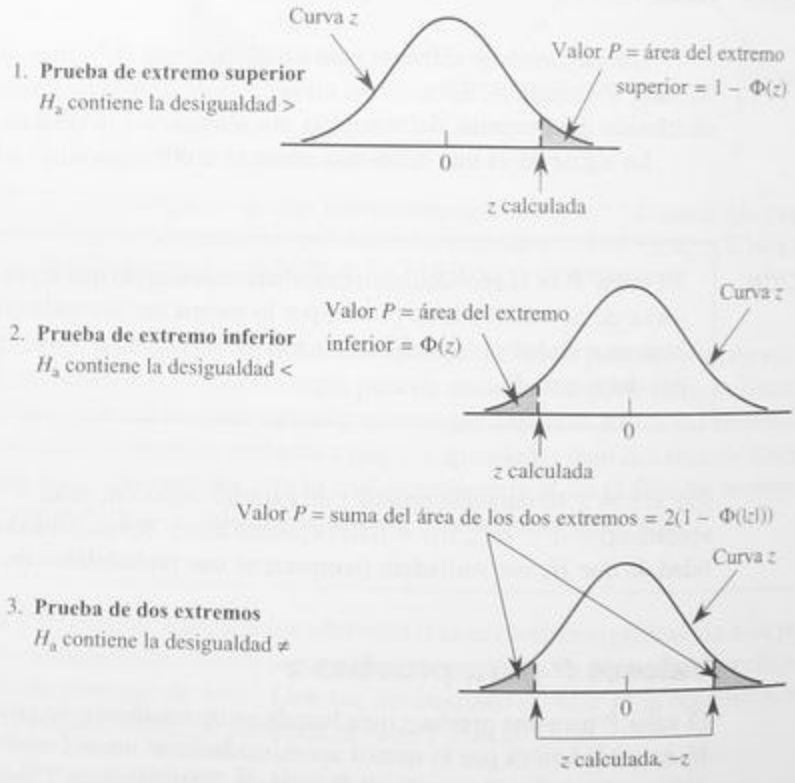


Figura 8.7 Determinación del valor P para una prueba z

En el siguiente ejemplo se ilustra el uso del método del valor P para probar hipótesis por medio de una secuencia de pasos modificados a partir de la secuencia recomendada antes.

Ejemplo 8.17 El espesor objetivo para obleas de silicio utilizado en cierto tipo de circuito integrado es 245 μm. Se obtiene una muestra de 50 obleas y se determina el espesor de cada una, obteniéndose un espesor medio muestral de 246.18 μm y una desviación estándar muestral de 3.60 μm. ¿Indican los datos que el espesor promedio real de la oblea es un poco distinto del valor objetivo?

1. Parámetro de interés: μ = espesor promedio real de la oblea
2. Hipótesis nula H_0: $\mu = 245$
3. Hipótesis alternativa H_a: $\mu \neq 245$
4. Fórmula para el valor del estadístico de prueba: $z = \dfrac{\bar{x} - 245}{s/\sqrt{n}}$
5. Cálculo para el valor del estadístico de prueba: $z = \dfrac{246.18 - 245}{3.60/\sqrt{50}} = 2.32$
6. Determinación del valor P: debido a que la prueba es de dos extremos,
$$\text{Valor } P = 2(1 - \Phi(2.32)) = 0.0204$$
7. Conclusión: usar un nivel de significación de 0.01, H_0 no sería rechazada puesto que $0.0204 > 0.01$. A este nivel de significación, hay evidencia insuficiente para concluir que el espesor promedio real difiere del valor objetivo. ■

Valores *P* para pruebas *t*

Así como el valor P para una prueba z es un área de la curva z, el valor P para una prueba t será un área de la curva t. En la figura 8.8 de la página siguiente se ilustran tres casos. El número de gl para una prueba t con una muestra es $n - 1$.

La tabla de valores críticos t utilizada antes para intervalos de confianza y predicción no contiene información suficiente acerca de alguna distribución t particular que permita determinar con precisión las áreas deseadas. Por lo tanto, se incluye otra tabla para t en la tabla A.8 del apéndice, una que contiene una tabulación de áreas de curva t del extremo superior. Cada columna de la tabla es para un número distinto de grados de libertad, y los renglones son para valores calculados del estadístico de prueba t que van de 0.0 a 4.0 en incrementos de 0.1. Por ejemplo, el número 0.074 aparece en la intersección del renglón 1.6 y la columna de 8 gl, de modo que el área bajo la curva de 8 gl a la derecha de 1.6 (un área de extremo superior) es 0.074. Debido a que las curvas t son simétricas, 0.074 es también el área bajo la curva de 8 gl a la izquierda de −1.6 (un área de extremo inferior).

Suponga, por ejemplo, que una prueba de H_0: $\mu = 100$ contra H_a: $\mu > 100$, se basa en la distribución t de 8 gl. Si el valor calculado de la estadística de prueba es $t = 1.6$, entonces el valor P para esta prueba de extremo superior es 0.074. Debido a que 0.074 es mayor que 0.05, no se podría rechazar H_0 al nivel de significación de 0.05. Si la hipótesis alternativa es H_a: $\mu < 100$ y una prueba basada en 20 gl produce $t = -3.2$, entonces en la tabla A.8 del apéndice se observa que el valor P es el área captada del extremo inferior 0.002. La hipótesis nula se puede rechazar ya sea al nivel 0.05 o 0.01. Considere probar H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 \neq 0$; la hipótesis nula expresa que las medias de las dos poblaciones son idénticas, mientras que la hipótesis alternativa

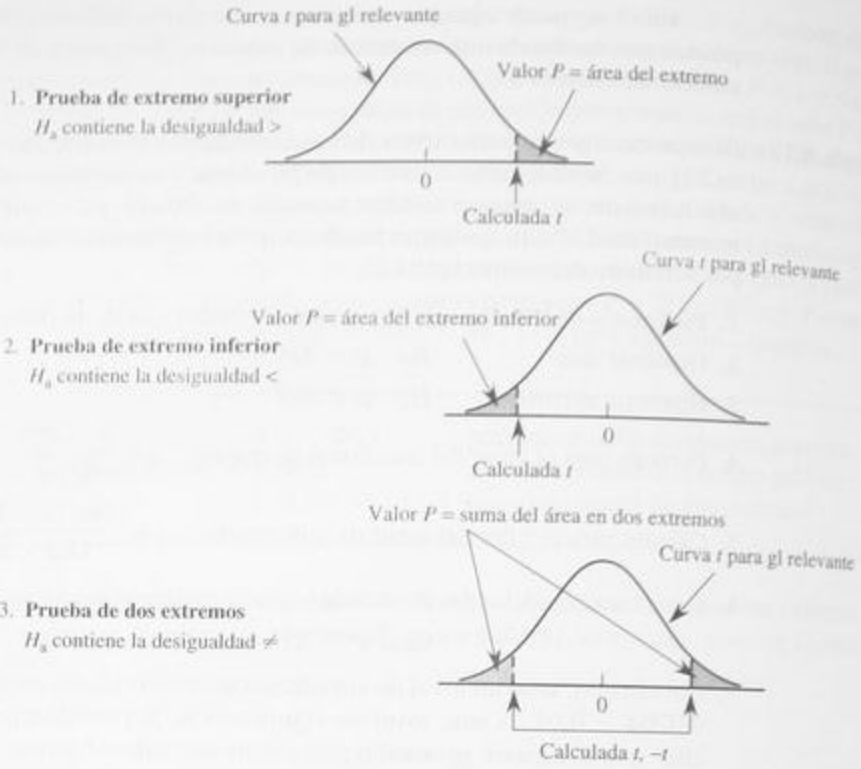


Figura 8.8 Valor *P* para prueba *t*

establece que son diferentes, sin especificar una dirección de desviación con respecto a H_0. Si una prueba *t* se basa en 20 gl y $t = 3.2$, entonces el valor *P* para esta prueba de dos extremos es $2(0.002) = 0.004$. Éste también sería el valor de *P* para $t = -3.2$. El área de extremo es doble porque los valores mayores que 3.2 y menores que −3.2 son más contradictorios para H_0 de lo que se calculó (valores apartados en *cualquier* extremo de la curva *t*).

Ejemplo 8.18 En el ejemplo 8.9, se realizó una prueba de H_0: $\mu = 25$ contra H_a: $\mu > 25$ con base en 4 gl. El valor calculado de *t* fue 1.04. Al observar la columna de 4 gl de la tabla A.8 del apéndice y localizar el renglón 1.0, se ve que el elemento es 0.187, por lo tanto el valor $P \approx 0.187$. Resulta claro que este valor *P* es mayor que cualquier nivel de significación razonable α (0.01, 0.05 e incluso 0.10), así que no hay razón para rechazar la hipótesis nula. El resultado de MINITAB incluido en el ejemplo 8.9 tiene un valor $P = 0.18$. Los valores *P* de los paquetes de *software* serán más precisos que lo que resulte de la tabla A.8 del apéndice puesto que los valores de *t* de la tabla son precisos sólo hasta el dígito de las decenas. ■

Ejercicios | Sección 8.4 (44 a 58)

44. ¿Para cuál de los valores de *P* proporcionados sería rechazada la hipótesis nula al llevar a cabo una prueba de nivel 0.05?

a. 0.001 **b.** 0.021 **c.** 0.078
d. 0.047 **e.** 0.148

45. A continuación se dan pares de valores P y niveles de significación, α. Para cada par, exprese si el valor observado P daría lugar al rechazo de H_0 al nivel de significación determinado.

a. Valor $P = 0.084$, $\alpha = 0.05$
b. Valor $P = 0.003$, $\alpha = 0.001$
c. Valor $P = 0.498$, $\alpha = 0.05$
d. Valor $P = 0.084$, $\alpha = 0.10$
e. Valor $P = 0.039$, $\alpha = 0.01$
f. Valor $P = 0.218$, $\alpha = 0.10$

46. Sea μ el tiempo de reacción medio a cierto estímulo. Para una prueba z de muestra grande de H_0: $\mu = 5$ contra H_a: $\mu > 5$, determine el valor P relacionado con cada uno de los valores proporcionados de la estadística de prueba z.

a. 1.42 **b.** 0.90 **c.** 1.96 **d.** 2.48 **e.** −0.11

47. Se supone que la presión de inflado de llantas recién compradas de un cierto tipo es de 30 lb/pulg^2. Sea μ la presión promedio real. Determine el valor P relacionado con cada estadístico de prueba proporcionada para probar H_0: $\mu = 30$ contra H_a: $\mu \neq 30$.

a. 2.10 **b.** −1.75 **c.** −0.55 **d.** 1.41 **e.** −5.3

48. Proporcione tanta información como sea posible acerca del valor P de una prueba t en cada una de las situaciones siguientes:

a. Prueba de extremo superior, gl = 8, $t = 2.0$
b. Prueba de extremo inferior, gl = 11, $t = -2.4$
c. Prueba de dos extremos, gl = 15, $t = -1.6$
d. Prueba de extremo superior, gl = 19, $t = -0.4$
e. Prueba de extremo superior, gl = 5, $t = 5.0$
f. Prueba de dos extremos, gl = 40, $t = -4.8$

49. La pintura utilizada para hacer líneas en las carreteras debe reflejar luz suficiente para que se vea con claridad en la noche. Sea μ la lectura promedio real del reflectómetro para un nuevo tipo de pintura en consideración. Una prueba de H_0: $\mu = 20$ contra H_a: $\mu > 20$ se basará en una muestra aleatoria de tamaño n de una distribución poblacional normal. ¿Qué conclusión es apropiada en cada una de las situaciones siguientes?

a. $n = 15$, $t = 3.2$, $\alpha = 0.05$
b. $n = 9$, $t = 1.8$, $\alpha = 0.01$
c. $n = 24$, $t = -0.2$

50. Represente con μ la concentración promedio real de receptor sérico para las mujeres embarazadas. Se sabe que el promedio para las mujeres es 5.63. En el artículo "Serum Transferrin Receptor for the Detection of Iron Deficiency in Pregnancy" (*Amer. J. Clinical Nutr.*, 1991: 1077-1081) se informa que el valor $P > 0.10$ para una prueba de H_0: $\mu = 5.63$ contra H_a: $\mu \neq 5.63$ con base en $n = 176$ mujeres embarazadas. Con un nivel de significación de 0.01, ¿cuál sería su conclusión?

51. Un fabricante de aspirinas llena los frascos por peso en vez de por conteo. Puesto que cada frasco debe contener 100 tabletas, el peso promedio por tableta debe ser 5 granos. Se pesa cada una de 100 tabletas tomadas de un lote muy grande, y se obtiene un peso promedio muestral por tableta de 4.87 granos y una desviación estándar muestral de 0.35 granos. ¿Esta información proporciona evidencia inequívoca para concluir que la compañía no está llenando sus frascos como se anuncia? Pruebe las hipótesis apropiadas mediante $\alpha = 0.01$, calculando primero el valor P y comparándolo después al nivel de significación especificado.

52. Como resultado de la variabilidad en el proceso de manufactura, el punto de cedencia real de una muestra de acero dulce sujeto a esfuerzo creciente normalmente diferirá del punto de cedencia teórico. Represente con p la proporción verdadera de muestras que ceden antes de su punto de cedencia teórico. Si con base en una muestra se puede concluir que más de 20% de los especímenes ceden antes del punto teórico, se tendrá que modificar el proceso de producción.

a. Si 15 de 60 especímenes ceden antes del punto teórico, ¿cuál es el valor P cuando se usa la prueba apropiada, y qué recomienda que haga la compañía?
b. Si el porcentaje real de "cedencias iniciales" es en realidad 50% (de modo que el punto teórico es la mediana de la distribución de cedencia) y se utiliza un nivel 0.01, ¿cuál es la probabilidad de que la compañía concluya que es necesaria una modificación del proceso?

53. Muchos consumidores están volviendo a los medicamentos genéricos como una forma de reducir el costo de las medicaciones prescritas. En el artículo "Commercial Information on Drugs: Confusing to the Physician?" (*J. of Drugs Issues*, 1988: 245-257) se proporcionan los resultados de un estudio de 102 médicos. Sólo 47 de los encuestados conocían el nombre genérico para el fármaco metadona. ¿Proporciona lo anterior evidencia suficiente para concluir que menos de la mitad de los médicos conocen el nombre genérico para la metadona? Lleve a cabo una prueba de hipótesis con un nivel de significación de 0.01 por medio del método del valor P.

54. Se obtuvo una muestra aleatoria de especímenes de suelo, y se determinó la cantidad de materia orgánica (%) en el suelo para cada espécimen. Los resultados se muestran a continuación (de "Engineering Properties of Soil", *Soil Science*, 1998: 93-102).

1.10	5.09	0.97	1.59	4.60	0.32	0.55	1.45
0.14	4.47	1.20	3.50	5.02	4.67	5.22	2.69

3.98 3.17 3.03 2.21 0.69 4.47 3.31 1.17
0.76 1.17 1.57 2.62 1.66 2.05

Los valores de la media muestral, desviación estándar muestral y error estándar (estimado) de la media son 2.481, 1.616 y 0.295, respectivamente. ¿Estos datos indican que el porcentaje promedio real de materia orgánica en este tipo de suelo es un poco distinto de 3%? Realice una prueba de hipótesis apropiada al nivel de significación 0.10 determinando primero el valor P. ¿Sería diferente su conclusión se hubiera utilizado $\alpha = 0.05$? (*Nota*: una gráfica de probabilidad normal de los datos muestra un patrón aceptable en vista del tamaño de muestra razonablemente grande.)

55. Los tiempos de activación del primer rociador para una serie de pruebas con sistemas de aspersión para prevención de incendios usando una espuma acuosa formadora de película fueron (en segundos)

27 41 22 27 23 35 30 33 24 27 28 22 24

(véase "Use of AFFF en Sprinkler Systems," *Fire Technology*, 1976: 5). El sistema ha sido diseñado de modo que el tiempo de activación promedio real sea a lo sumo 25 s en estas condiciones. ¿Los datos contradicen sin lugar a dudas la validez de esta especificación de diseño? Pruebe las hipótesis pertinentes al nivel de significación 0.05 con el método del valor P.

56. Cierta pluma ha sido diseñada de modo que la duración promedio real para escritura en condiciones controladas (se requiere una máquina de escritura) es por lo menos 10 horas. Se elige al azar una muestra de 18 plumas, se determina la duración para escritura de cada una y una gráfica de probabilidad normal de los datos resultantes respalda el uso de una prueba t con una sola muestra.
   a. ¿Qué hipótesis se deben probar si el investigador cree a priori que se cumplió la especificación de diseño?
   b. ¿Qué conclusión es apropiada si se prueban las hipótesis del inciso (a), $t = -2.3$, y $\alpha = 0.05$?
   c. ¿Qué conclusión es apropiada si se prueban las hipótesis del inciso (a), $t = -1.8$, y $\alpha = 0.01$?
   d. ¿Qué se debe concluir si se prueban las hipótesis del inciso (a), $t = -3.6$?

57. Se comprueba la precisión de un espectrofotómetro utilizado para medir la concentración de CO [ppm (partes por millón) en volumen] al tomar lecturas en un gas fabricado (conocido como gas de calibración) en el que la concentración de CO se controló con precisión a 70 ppm. Si las lecturas indican que el espectrofotómetro no está funcionando de manera apropiada, éste tendrá que ser recalibrado. Suponga que si está calibrado de manera adecuada, la concentración medida para las muestras de gas de calibración tiene una distribución normal. Con base en las seis lecturas —85, 77, 82, 68, 72, y 69— ¿es necesaria la recalibración? Lleva a cabo una prueba de las hipótesis pertinentes por medio del método del valor P con $\alpha = 0.05$.

58. La conductividad relativa de un dispositivo semiconductor se determina por la cantidad de impureza "dopada" en el dispositivo durante su manufactura. Un diodo de silicio que se utilizará para un propósito específico requiere un voltaje promedio de corte-activación de 0.60 V, y si esto no se logra, se debe ajustar la cantidad de impureza. Se selecciona al azar una muestra de diodos, y se determina el voltaje de corte-activación. El siguiente resultado de SAS se obtuvo de solicitar probar las hipótesis apropiadas.

N	Media	Desv. Est.	T	Prob>|T|
15	0.0453333	0.0899100	1.9527887	0.0711

[*Nota*: SAS prueba de forma explícita H_0: $\mu = 0$, así que para probar H_0: $\mu = 0.60$, el valor nulo 0.60 se debe restar de cada x_i; entonces el valor reportado es el promedio de los valores $(x_i - 0.60)$. Asimismo, el valor P de SAS siempre es para una prueba de dos extremos.] ¿Cuál sería la conclusión para un nivel de significación de 0.01, 0.05, 0.10?

8.5 Algunos comentarios acerca de seleccionar un procedimiento de prueba

Una vez que el experimentador ha decidido en relación con la pregunta de interés y el método para reunir datos (el diseño del experimento), la construcción de un procedimiento apropiado de prueba consiste en tres pasos distintos:

1. Especificar una estadística de prueba (la función de los valores observados a partir de la cual se tomarán decisiones).

2. Decidir acerca de la forma general de la región de rechazo (por lo común rechazar H_0 para valores grandes adecuados del estadístico de prueba, rechazar para valores pequeños apropiados o rechazar para valores grandes o pequeños).
3. Seleccionar el valor o valores críticos numéricos específicos que separan la región de rechazo de la región de aceptación (al obtener la distribución del estadístico de prueba cuando H_0 es verdadera, y seleccionar después un nivel de significación).

En los ejemplos mencionados hasta el momento, los pasos 1 y 2 se llevaron a cabo de una manera *ad hoc* a través de la intuición. Por ejemplo, cuando se supuso que la población subyacente era normal con media μ y σ conocida, se pasó de $\bar{X}$ al estadístico de prueba estandarizado

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Para probar H_0: $\mu = \mu_0$ contra H_a: $\mu > \mu_0$, de manera intuitiva se rechazó H_0 cuando z era grande. Por último, el valor crítico se determinó al especificar el nivel de significación α y usar el hecho de que Z tiene una distribución normal estándar cuando H_0 es verdadera. La confiabilidad de la prueba para llegar a una decisión correcta se puede evaluar mediante el análisis de las probabilidades de error de tipo II.

Las cuestiones por considerar en la realización de los pasos 1-3 comprenden las siguientes preguntas:

1. ¿Cuáles son las implicaciones prácticas y consecuencias de elegir un nivel particular de significación una vez que se determinaron los otros aspectos de un procedimiento de prueba?
2. ¿Existe un principio general, no dependiente sólo de la intuición, que se puede usar para obtener mejores o buenos procedimientos de prueba?
3. Cuando dos o más pruebas son apropiadas en una determinada situación, ¿cómo se pueden comparar las pruebas para decidir cuál se debe usar?
4. Si se deriva una prueba bajo suposiciones específicas acerca de la distribución o población que está siendo muestreada, ¿qué tan bien funciona el procedimiento de prueba cuando se violan las suposiciones?

Significación estadística contra significación práctica

Aunque el proceso de llegar a una decisión mediante la metodología de la prueba de hipótesis clásica tiene que ver con seleccionar un nivel de significación y después rechazar o no rechazar H_0 al nivel α, informar simplemente la α utilizada y la decisión alcanzada transmite poca información contenida en los datos de la muestra. En particular, cuando los resultados de un experimento tienen que ser comunicados a una gran audiencia, el rechazo de H_0 al nivel 0.05 será mucho más convincente si el valor observado del estadístico de prueba excede en gran medida el valor crítico 5% que si apenas excede ese valor. Esto es precisamente lo que origina el concepto de valor P como una manera de informar la significación sin imponer una α particular en otros que quisieran sacar su propia conclusión.

Sin embargo, aun cuando el valor P se incluye en un resumen de resultados, podría haber dificultad para interpretar este valor y tomar una decisión. Esto es porque un valor P pequeño, que por lo común indicaría **significación estadística** en el sentido de que indicaría de manera firme el rechazo de H_0 en favor de H_a, podría ser el resultado de un gran tamaño de muestra con una desviación desde H_0 que tiene poca **significa-**

ción práctica. En muchas situaciones experimentales, sólo valdría la pena detectar las desviaciones de gran magnitud a partir de H_0, mientras que una desviación pequeña de H_0 tendría poca significación práctica.

Considere como ejemplo probar H_0: $\mu = 100$ contra H_a: $\mu > 100$ donde μ es la media de una población normal con $\sigma = 10$. Suponga que un valor verdadero de $\mu = 101$ no representaría una desviación grave de H_0 en el sentido de que no rechazar H_0 cuando $\mu = 101$ sería un error relativamente poco costoso. Para un tamaño de muestra razonablemente grande n, esta μ daría lugar a un valor $\bar{x}$ cercano a 101, así que no desearía que esta evidencia muestral favorezca el rechazo de H_0 cuando se observa que $\bar{x} = 101$. Para varios tamaños de muestra, en la tabla 8.1 se registra tanto el valor P cuando $\bar{x} = 101$ como la probabilidad de no rechazar H_0 al nivel 0.01 cuando $\mu = 101$.

En la segunda columna de la tabla 8.1 se observa que incluso para tamaños de muestra moderadamente grandes, el valor P de $\bar{x} = 101$ es una firme indicación de rechazar H_0, mientras que la $\bar{x}$ observada por sí misma indica que en términos prácticos el valor verdadero de μ difiere poco del valor nulo $\mu_0 = 100$. La tercera columna señala que aun cuando hay poca diferencia práctica entre la μ verdadera y el valor nulo, para un nivel fijo de significación un tamaño de muestra grande casi siempre conducirá al rechazo de la hipótesis nula a ese nivel. En resumen, *se debe tener especial cuidado al interpretar la evidencia cuando el tamaño de muestra es grande, puesto que cualquier desviación pequeña a partir de H_0 casi con certeza será detectada mediante una prueba, aunque esta desviación pudiera tener poca importancia práctica.*

Tabla 8.1 Ilustración del efecto del tamaño de la muestra en el valor P y β

n	Valor P cuando $\bar{x} = 101$	$\beta(101)$ para una prueba de nivel 0.01
25	0.3085	0.9664
100	0.1587	0.9082
400	0.0228	0.6293
900	0.0013	0.2514
1600	0.0000335	0.0475
2500	0.000000297	0.0038
10 000	7.69×10^{-24}	0.0000

Principio de cociente de verosimilitudes

Sean $x_1, x_2, \ldots, x_n$ las observaciones en una muestra aleatoria de tamaño n de una distribución de probabilidad $f(x; \theta)$. La distribución conjunta evaluada en estos valores muestrales es el producto $f(x_1; \theta) \cdot f(x_2; \theta) \cdot \cdots \cdot f(x_n; \theta)$. Al igual que en la explicación de la estimación de máxima verosimilitud, la *función de verosimilitud* es esta distribución conjunta considerada como una función de θ. Considere probar H_0: θ está en Ω_0 contra H_a: θ está en Ω_a, donde Ω_0 y Ω_a son disjuntos (por ejemplo, H_0: $\theta \leq 100$ contra H_a: $\theta > 100$). El **principio de cociente de verosimilitudes** para construcción de la prueba es como sigue:

1. Determine el valor más grande de verosimilitud para cualquier θ en Ω_0 (determine la estimación de máxima verosimilitud dentro de Ω_0 y sustituya en la función de verosimilitud).
2. Calcule el valor más grande de verosimilitud para cualquier θ en Ω_a.
3. Forme la relación

$$\lambda(x_1, \ldots, x_n) = \frac{\text{verosimilitud máxima para } \theta \text{ en } \Omega_0}{\text{verosimilitud máxima para } \theta \text{ en } \Omega_a}$$

La relación $\lambda(x_1, \ldots, x_n)$ se llama *valor estadístico del cociente de verosimilitud*. El procedimiento de prueba consiste en rechazar H_0 cuando su relación es pequeña. Es decir, se elige una constante k y se rechaza H_0 si $\lambda(x_1, \ldots, x_n) \leq k$. Por lo tanto, se rechaza H_0 cuando el denominador de λ excede por mucho al numerador, lo cual indica que los datos son mucho más consistentes con H_a que con H_0.

La constante k se elige para proporcionar la probabilidad de error de tipo I deseada. Con frecuencia la desigualdad $\lambda \leq k$ se puede manejar para producir una condición equivalente más simple. Por ejemplo, para probar H_0: $\mu \leq \mu_0$ contra H_a: $\mu > \mu_0$ en el caso de normalidad, $\lambda \leq k$ es equivalente a $t \geq c$. Por consiguiente, con $c = t_{\alpha,n-1}$, la prueba de cociente de verosimilitudes es una prueba t con una sola muestra.

El principio de cociente de verosimilitudes también se puede aplicar cuando las X_i tienen distribuciones diferentes, e incluso cuando son dependientes, aunque en esos casos, la función de verosimilitud puede ser complicada. Muchos de los procedimientos de prueba que se presentarán en capítulos posteriores se obtienen a partir del principio de cociente de verosimilitudes. Con frecuencia estas pruebas minimizan a β entre las pruebas que tiene la α deseada, así que en realidad son las mejores pruebas. Para más detalles y algunos ejemplos resueltos, refiérase al libro de DeGroot o el de Hogg y Craig mencionados en la bibliografía del capítulo 6.

Una limitación práctica en el uso del principio de relación de probabilidad es que, para construir el estadístico de prueba del cociente de verosimilitudes, se debe especificar la forma de la distribución de probabilidad de la cual proviene la muestra. Para derivar la prueba t del principio de cociente de verosimilitudes, el investigador debe asumir una fdp normal. Si un investigador está dispuesto a asumir que la distribución es simétrica, pero no desea ser específico en cuanto a su forma exacta (como por ejemplo normal, uniforme o de Cauchy), entonces el principio no es válido, porque no hay forma de escribir una fdp conjunta que sea válida al mismo tiempo para todas la distribuciones simétricas. En el capítulo 15, se presentan varios procedimientos de prueba de **distribución libre**, llamados así porque la probabilidad de un error de tipo I se controla de forma simultánea para muchas distribuciones subyacentes distintas. Estos procedimientos son útiles cuando el investigador tiene conocimiento limitado de la distribución subyacente. Se ampliará la explicación acerca de los temas 3 y 4 listados al comienzo de esta sección.

Ejercicios | Sección 8.5 (59 a 60)

59. Reconsidere el problema del secado de pintura analizado en el ejemplo 8.2. Las hipótesis fueron H_0: $\mu = 75$ contra H_a: $\mu < 75$, con un valor supuesto de $\sigma = 9.0$. Considere el valor alternativo $\mu = 74$, que en el contexto del problema es de suponer no sería una desviación importante, desde el punto de vista práctico, respecto a H_0.

- **a.** Para una prueba al nivel 0.01, calcule β en esta alternativa para tamaños de muestra $n = 100$, 900 y 2500.
- **b.** Si el valor observado de $\overline{X}$ es $\bar{x} = 74$, ¿qué se puede decir acerca del valor P resultante cuando $n = 2500$? ¿Tienen importancia estadística los datos a alguno de los valores comunes de α?
- **c.** ¿Desearía usar un tamaño de muestra de 2500 junto con una prueba al nivel 0.01 (sin considerar el costo de tal experimento)? Explique.

60. Considere la prueba de muestra grande al nivel 0.01 de la sección 8.3 para probar H_0: $p = 0.2$ contra H_a: $p > 0.2$.

- **a.** Para el valor alternativo $p = 0.21$, calcule $\beta(0.21)$ para tamaños de muestra $n = 100$, 2500, 10 000, 40 000 y 90 000.
- **b.** Para $\hat{p} = x/n = 0.21$, calcule el valor P cuando $n = 100$, 2500, 10 000 y 40 000.
- **c.** En muchas situaciones, ¿sería razonable usar una prueba al nivel 0.01 junto con un tamaño de muestra de 40 000? ¿Por qué sí o por qué no?

Ejercicios complementarios (61 a 82)

61. Una muestra de 50 lentes utilizadas en anteojos produce un espesor medio muestral de 3.05 mm y una desviación estándar muestral de 0.34 mm. El espesor promedio real deseado de tales lentes es 3.20 mm. ¿Los datos indican firmemente que el espesor promedio muestral de tales lentes es un poco distinto de lo que se desea? Pruebe con $\alpha = 0.05$.

62. En el ejercicio 61, suponga que el experimentador consideró antes de reunir los datos que el valor de σ era de alrededor de 0.30. Si el experimentador deseaba que la probabilidad de un error de tipo II fuera 0.05 cuando $\mu = 3.00$, ¿fue un tamaño de muestra 50 innecesariamente grande?

63. Se especifica que cierto tipo de hierro debe contener 0.85 g de silicio por 10 g de hierro (0.85%). Se determinó el contenido de silicio de cada uno de 25 especímenes seleccionados al azar, y se obtuvo el siguiente resultado de MINITAB de un prueba de las hipótesis apropiadas.

Variable	N	Media	Desv. Est.	EE de la media	T	P
Cont. sil.	25	0.8880	0.1807	0.0361	1.05	0.30

a. ¿Qué hipótesis se probaron?

b. ¿Qué conclusión se debe alcanzar para un nivel de significación de 0.05, y por qué? Conteste la misma pregunta para un nivel de significación de 0.10.

64. Un método para enderezar alambre antes de enfriarlo para hacer un resorte se llama "enderezamiento de rodillo". En el artículo "The Effect of Roller and Spinner Wire Straightening on Coiling Performance and Wire Properties" (*Springs*, 1987: 27-28) se informa acerca de las propiedades de tensión del alambre. Suponga que se selecciona una muestra de 16 alambres y se prueba cada una para determinar la resistencia a la tensión (N/mm^2). La media muestral resultante y la desviación estándar son 2160 y 30, respectivamente.

a. La resistencia promedio a la tensión para resortes fabricados por medio de enderezamiento con hilador es 2150 N/mm^2. ¿Qué hipótesis se deben probar para determinar si la resistencia promedio a la tensión para el método de rodillo pasa de 2150?

b. Si se supone que la distribución de la resistencia a la tensión es aproximadamente normal, ¿qué estadístico de prueba se utilizaría para probar las hipótesis del inciso (a)?

c. ¿Cuál es el valor del estadístico de prueba para estos datos?

d. ¿Cuál es el valor P para el valor del estadístico de prueba calculado en el inciso (c)?

e. Para una prueba al nivel 0.05, ¿a qué conclusión llegaría?

65. En el artículo "A Rapid Method to Determine Total Phosphorus in Soils" (*Soil Sci. Amer. J.*, 1988: 1301-1304) se describe un nuevo método para medir concentraciones de fósforo en el suelo. Suponga que con el nuevo método se analiza una muestra de 11 especímenes de suelo, cada uno con un contenido verdadero de fósforo de 548 mg/kg. La media muestral resultante y la desviación estándar para la concentración de fósforo son 587 y 10, respectivamente.

a. ¿Hay evidencia de que la concentración promedio de fósforo obtenida con el nuevo método difiere en forma importante del valor verdadero de 548 mg/kg? Utilice $\alpha = 0.05$.

b. ¿Cuáles suposiciones se deben hacer para que sea adecuada la prueba del inciso (a)?

66. En el artículo "Orchard Floor Management Utilizing Soil-Applied Coal Dust for Frost Protection" (*Agri. and Forest Meteorology*, 1988: 71-82) aparecen los siguientes valores para flujo térmico del suelo de ocho parcelas cubiertas con polvo de carbón.

34.7 35.4 34.7 37.7 32.5 28.0 18.4 24.9

El flujo térmico promedio para parcelas cubiertas sólo con pasto es 29.0. Si se supone que la distribución de flujo térmico es aproximadamente normal, ¿los datos indican que el polvo de carbón es eficaz para incrementar el flujo térmico promedio por arriba del que se tiene para el pasto? Pruebe las hipótesis adecuadas con $\alpha = 0.05$.

67. En el artículo "Caffeine Knowledge, Attitudes, and Consumption in Adult Women" (*J. of Nutrition Educ.*, 1992: 179-184) aparece el siguiente resumen de datos sobre consumo diario de cafeína para una muestra de mujeres adultas: $n = 47$, $\bar{x} = 215$ mg, $s = 235$ mg, y rango = 5–1176.

a. ¿Al parecer es posible que la distribución poblacional de consumo diario de cafeína sea normal? ¿Es necesario suponer una distribución poblacional normal para probar las hipótesis acerca del valor del consumo medio poblacional? Explique su razonamiento.

b. Suponga que antes se había considerado que el consumo promedio era a lo sumo 200 mg.

¿Los datos proporcionados contradicen la consideración anterior? Pruebe las hipótesis pertinentes al nivel de significación 0.10 e incluya un valor P en su análisis.

68. El resultado siguiente se obtuvo cuando se empleó MINITAB para probar las hipótesis adecuadas acerca del tiempo de activación promedio real con base en los datos del ejercicio 53. Utilice esta información para llegar a una conclusión al nivel de significación 0.05 y también al nivel 0.01.

```
PRUEBA DE MU = 25.000 VS MU G.T. 25.000
     N   MEDIA  DESV.     EE DE     T  VALOR P
                 EST  LA MEDIA
tempo13 27.923  5.619     1.559 1.88    0.043
```

69. Se supone que la resistencia a la ruptura promedio real de aisladores de cerámica de cierto tipo es de por lo menos 10 lb/pulg^2. Los aisladores se utilizarán para una aplicación a menos que los datos de la muestra indiquen de manera concluyente que no se cumple la especificación. Una prueba de hipótesis con $\alpha = 0.01$ se basará en una muestra aleatoria de diez aisladores. Suponga que la distribución de resistencia a la ruptura es normal con desviación estándar desconocida.

a. Si la desviación estándar verdadera es 0.80, ¿cuál es la probabilidad de que los aisladores se consideren satisfactorios cuando la resistencia a la ruptura promedio verdadera sea en realidad de sólo 9.5 o 9.0?

b. ¿Qué tamaño de muestra sería necesario para tener una probabilidad de 75% de detectar que la resistencia a la ruptura promedio real es 9.5 cuando la desviación estándar verdadera es 0.80?

70. Las siguientes observaciones del tiempo de flama residual (segundos) para tiras de ropa de dormir infantil tratada aparecieron en el artículo "An Introduction to Some Precision and Accuracy of Measurement Problems" (*J. of Testing and Eval.*, 1982: 132-140). Suponga que se acordó un tiempo de flama promedio real de a lo sumo 9.75. ¿Los datos indican que no se ha cumplido esta condición? Realice una prueba adecuada después de investigar primero la validez de suposiciones que sustentan el método de inferencia.

9.85	9.93	9.75	9.77	9.67	9.87	9.67
9.94	9.85	9.75	9.83	9.92	9.74	9.99
9.88	9.95	9.95	9.93	9.92	9.89	

71. Se piensa que la incidencia de cierto tipo de defecto cromosómico en la población de varones adultos en Estados Unidos es de uno en 75. Una muestra aleatoria de 800 individuos en instituciones penales norteamericanas, revela que 16 tienen tal defecto. ¿Se puede concluir que la tasa de incidencia de este defecto entre prisioneros difiere de la tasa que se supone para toda la población de varones adultos?

a. Establezca y pruebe las hipótesis pertinentes con $\alpha = 0.05$. ¿Qué tipo de error podría cometer para llegar a una conclusión?

b. ¿Qué valor P se relaciona con esta prueba? Con base en este valor P, ¿se podría rechazar H_0 al nivel de significación 0.20?

72. En una investigación de la toxina producida por cierta serpiente venenosa, un investigador preparó 26 frascos diferentes, cada uno con 1 g de la toxina, y luego determinó la cantidad de antitoxina necesaria para neutralizar la toxina. Se encontró que la cantidad promedio muestral de antitoxina necesaria era de 1.89 mg, y la desviación estándar muestral fue 0.42. En investigaciones anteriores se encontró que la cantidad neutralizante promedio real era de 1.75 mg/g de toxina. ¿Contradice la nueva información el valor encontrado en la investigación anterior? Pruebe las hipótesis pertinentes con el método del valor P. ¿La validez del análisis depende de las suposiciones acerca de la distribución poblacional de la cantidad neutralizante? Explique.

73. Se determinó que la resistencia compresiva promedio muestral no restringida para 45 especímenes de cierto tipo de ladrillo es 3107 lb/pulg^2 y la desviación estándar muestral es 188. La distribución de resistencia compresiva no restringida puede ser un poco sesgada. ¿La información indica que la resistencia compresiva promedio real no restringida es menor que el valor de diseño de 3200? Pruebe con $\alpha = 0.001$.

74. Para probar la capacidad de mecánicos automotrices para identificar problemas simples en motores, se llevó un automóvil con uno de estos problemas a 72 talleres. Sólo 42 de 72 mecánicos que trabajaron en el automóvil identificaron de forma correcta el problema. ¿Esto es evidencia suficiente de que la verdadera proporción de mecánicos que pudieron identificar el problema es menor que 0.75? Calcule el valor P y obtenga una conclusión congruente con lo anterior.

75. Cuando $X_1, X_2, \ldots, X_n$ son variables de Poisson independientes, cada una con parámetro λ, y n es grande, la media muestral $\bar{X}$ tiene aproximadamente una distribución normal con $\mu = E(\bar{X}) = \lambda$ y $\sigma^2 = V(\bar{X}) = \lambda/n$. Esto implica que

$$Z = \frac{\bar{X} - \lambda}{\sqrt{\lambda/n}}$$

tiene aproximadamente una distribución normal estándar. Para probar H_0: $\lambda = \lambda_0$, se puede remplazar λ por λ_0 en la ecuación para Z a fin de obtener un estadístico de prueba. En realidad se prefiere este estadístico al estadístico para muestra grande con denominador $S/\sqrt{n}$ (cuando las X_i son de Poisson) porque se ajusta de forma explícita a la suposición de Poisson. Si el número de solicitudes de consulta que recibe un estadístico durante una semana de trabajo de cinco días tiene una distribución de Poisson y el número total de solicitudes de consulta durante un periodo de 36 semanas es 160, ¿esto indica que el número promedio real de solicitudes es mayor que 4.0? Pruebe con $\alpha = 0.02$.

76. Un fabricante de tinas para baño anuncia que con su equipo de calefacción se puede alcanzar una temperatura de 100°F a lo sumo en 15 minutos. Se selecciona una muestra aleatoria de 32 tinas, y para cada una se determina el tiempo necesario para alcanzar una temperatura de 100°F. El promedio de tiempo y la desviación estándar muestrales son 17.5 min. y 2.2 min., respectivamente. ¿Esta información arroja dudas sobre la afirmación de la compañía? Calcule el valor P y utilícelo para llegar a una conclusión al nivel 0.05 (suponga que la distribución de los tiempos de calentamiento es aproximadamente normal).

77. En el capítulo 7 se presentó un intervalo de confianza para la varianza σ^2 de una distribución poblacional normal. El resultado clave fue que la variable aleatoria $\chi^2 = (n-1)S^2/\sigma^2$ tiene una distribución chi-cuadrada con $n-1$ gl. Considere la hipótesis nula H_0: $\sigma^2 = \sigma_0^2$ (equivalentemente, $\sigma = \sigma_0$). Entonces, cuando H_0 es verdadera, el estadístico de prueba $\chi^2 = (n-1)S^2/\sigma_0^2$ tiene una distribución chi-cuadrada con $n-1$ gl. Si la alternativa pertinente es H_a: $\sigma^2 > \sigma_0^2$, rechazar H_0 si $(n-1)\,s^2/\sigma_0^2 \geq \chi^2_{\alpha,n-1}$ proporciona una prueba con nivel de significación α. Para asegurar características razonablemente uniformes para una aplicación en particular, se desea que la verdadera desviación estándar del punto de ablandamiento de cierto tipo de asfalto sea a lo sumo de 0.50°C. Se determinaron los puntos de ablandamiento de 10 especímenes diferentes, obteniéndose una desviación estándar muestral de 0.58°C ¿Esto contradice la especificación de uniformidad? Pruebe las hipótesis apropiadas con $\alpha = 0.01$.

78. Con referencia al ejercicio 77, suponga que un investigador desea probar H_0: $\sigma^2 = 0.04$ contra H_a: $\sigma^2 < 0.04$, con base en una muestra de 21 observaciones. El valor calculado de $20s^2/0.04$ es 8.58. Asigne límites para el valor P y luego llegue a una conclusión al nivel 0.01.

79. Cuando la distribución poblacional es normal y n es grande, la desviación estándar muestral S tiene aproximadamente una distribución normal con $E(S) \approx \sigma$ y $V(S) \approx \sigma^2/(2n)$. Ya se sabe que en este caso, para cualquier n, $\overline{X}$ es normal con $E(\overline{X}) = \mu$ y $V(\overline{X}) = \sigma^2/n$.

- **a.** Si se supone que la distribución subyacente es normal, ¿cuál es aproximadamente un estimador insesgado del percentil $\theta = \mu + 2.33\sigma$?
- **b.** Cuando las X_i son normales, se puede demostrar que $\overline{X}$ y S son variables aleatorias independientes (una mide la localización, mientras que otra mide la dispersión). Utilice esto para calcular $V(\hat{\theta})$ y $\sigma_{\hat{\theta}}$ para el estimador $\hat{\theta}$ del inciso (a). ¿Cuál es el error estándar estimado $\hat{\sigma}_{\hat{\theta}}$?
- **c.** Escriba una estadística de prueba para probar H_0: $\theta = \theta_0$ que tenga aproximadamente una distribución normal estándar cuando H_0 es verdadera. Si el pH del suelo tiene una distribución normal en cierta región y 64 muestras de suelo proporcionan $\overline{x} = 6.33$, $s = 0.16$, ¿esto proporciona evidencia firme para concluir que a lo sumo 99% de las muestras posibles tendrían un pH menor de 6.75? Pruebe con $\alpha = 0.01$.

80. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución exponencial con parámetro λ. Entonces, se puede demostrar que $2\lambda\sum X_i$ tiene una distribución chi-cuadrada con $\nu = 2n$ (primero demuestre que $2\lambda X_i$ tiene una distribución chi-cuadrada con $\nu = 2$).

- **a.** Utilice este hecho para obtener una estadística de prueba y región de rechazo que juntas especifiquen una prueba de nivel α para H_0: $\mu = \mu_0$ contra cada una de las tres alternativas comunes. [*Sugerencia*: $E(X_i) = \mu = 1/\lambda$, de modo que $\mu = \mu_0$ es equivalente a $\lambda = 1/\mu_0$.]
- **b.** Suponga que se prueban diez componentes idénticos, cada uno con una distribución de tiempo exponencial hasta que falla. Los tiempos de falla resultantes son

  95 16 11 3 42 71 225 64 87 123

  Utilice el procedimiento de prueba del inciso (a) para determinar si los datos indican fuertemente que el tiempo de vida promedio verdadero es menor que el valor afirmado de 75.

81. Suponga que la distribución poblacional es normal con σ conocida. Sea γ tal que $0 < \gamma < \alpha$. Para probar H_0: $\mu = \mu_0$ contra H_a: $\mu \neq \mu_0$, considere la prueba que rechaza H_0 si $z \geq z_\gamma$ o $z \leq -z_{\alpha-\gamma}$, donde la estadística de prueba es $Z = (\overline{X} - \mu_0)/(\sigma/\sqrt{n})$.

- **a.** Demuestre que $P(\text{error de tipo I}) = \alpha$.

b. Deduzca una expresión para $\beta(\mu')$. (*Sugerencia*: Exprese la prueba en la forma "rechazar H_0 si $\bar{x} \geq c_1$ o $\leq c_2$.")

c. Sea $\Delta > 0$. ¿Para qué valores de γ (en relación con α) será $\beta(\mu_0 + \Delta) < \beta(\mu_0 - \Delta)$?

82. Tras un periodo de aprendizaje, una organización aplica un examen que debe ser aprobado a fin de cumplir con los requisitos para ser socio. Sea $p = P$(aprueba el aprendiz seleccionado al azar). La organización desea un examen que la mayoría pero no todos deben aprobar; por lo tanto, decide que es deseable $p = 0.90$. Para un determinado examen las hipótesis pertinentes son H_0: $p = 0.90$ contra la alternativa H_a: $p \neq 0.90$. Suponga que diez personas se presentan al examen y sea $X =$ el número de los que aprueban.

a. ¿La región del extremo inferior $\{0, 1, \ldots, 5\}$ especifica una prueba de nivel 0.01?

b. Demuestre que aunque H_a es bilateral, ninguna prueba de dos extremos es una prueba de nivel 0.01.

c. Trace una gráfica de $\beta(p')$ como una función p' para esta prueba. ¿Esto es deseable?

Bibliografía

Consulte la bibliografía al final de los capítulos 6 y 7.

9

Inferencias basadas en dos muestras

Introducción

En los capítulos 7 y 8 se presentaron intervalos de confianza (IC) y procedimientos de prueba de hipótesis para una sola media μ, una sola proporción p y una sola varianza σ^2. Aquí se amplían estos métodos a situaciones que tienen que ver con medias, proporciones y varianzas de dos distribuciones poblacionales distintas. Por ejemplo, sea μ_1 la dureza Rockwell promedio para especímenes de acero tratados con calor y μ_2 la dureza promedio real para especímenes rolados en frío. Entonces, es posible que un investigador desee usar muestras de observaciones de dureza de cada tipo de acero como base para calcular una estimación del intervalo de $\mu_1 - \mu_2$, la diferencia entre las dos durezas promedio. Otro ejemplo sería que p_1 representa la proporción verdadera de celdas de níquel-cadmio producidas en ciertas condiciones de operación que resultan defectuosas debido a cortos internos, y que p_2 representa la proporción real de celdas con cortos internos producidas en ciertas condiciones de operación modificadas. Si la razón fundamental para estas condiciones modificadas es reducir la proporción de celdas defectuosas, un ingeniero de control de calidad desearía usar información muestral para probar la hipótesis nula H_0: $p_1 - p_2 = 0$ (es decir, $p_1 = p_2$) contra la hipótesis alternativa H_a: $p_1 - p_2 > 0$ (es decir, $p_1 > p_2$).

9.1 Pruebas z e intervalos de confianza para una diferencia entre dos medias poblacionales

Las inferencias analizadas en esta sección, se relacionan con una diferencia $\mu_1 - \mu_2$ entre las medias de dos distribuciones poblacionales distintas. Un investigador podría, por ejemplo, querer probar las hipótesis acerca de la diferencia entre resistencia de ruptura promedio de dos tipos de panel de fibra corrugado. Una hipótesis expresaría que $\mu_1 - \mu_2 = 0$, es decir, que $\mu_1 = \mu_2$. De manera alternativa, podría ser apropiado estimar $\mu_1 - \mu_2$ al calcular un IC de 95%. Este tipo de inferencias se basan en una muestra de observaciones de resistencia para cada tipo de panel de fibra.

Suposiciones básicas

1. $X_1, X_2, \ldots, X_m$ es una muestra aleatoria de una población con media μ_1 y varianza σ_1^2.
2. $Y_1, Y_2, \ldots, Y_n$ es una muestra aleatoria de una población con media μ_2 y varianza σ_2^2.
3. Las muestras X y Y son independientes entre sí.

El estimador natural de $\mu_1 - \mu_2$ es $\overline{X} - \overline{Y}$, la diferencia entre las medias muestrales correspondientes. La estadística de prueba resulta de estandarizar este estimador, así que se requieren expresiones para el valor esperado y la desviación estándar de $\overline{X} - \overline{Y}$.

PROPOSICIÓN

El valor esperado de $\overline{X} - \overline{Y}$ es $\mu_1 - \mu_2$, así que $\overline{X} - \overline{Y}$ es un estimador insesgado de $\mu_1 - \mu_2$. La desviación estándar de $\overline{X} - \overline{Y}$ es

$$\sigma_{\overline{X}-\overline{Y}} = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

Demostración: Ambos resultados dependen de las reglas del valor esperado y la varianza presentadas en el capítulo 5. Puesto que el valor esperado de una diferencia es la diferencia de los valores esperados,

$$E(\overline{X} - \overline{Y}) = E(\overline{X}) - E(\overline{Y}) = \mu_1 - \mu_2$$

Debido a que X y Y son muestras independientes, $\overline{X}$ y $\overline{Y}$ son cantidades independientes, así que la varianza de la diferencia es la *suma* de $V(\overline{X})$ y $V(\overline{Y})$:

$$V(\overline{X} - \overline{Y}) = V(\overline{X}) + V(\overline{Y}) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$$

La desviación estándar de $\overline{X} - \overline{Y}$ es la raíz cuadrada de esta expresión. ■

Si se considera a $\mu_1 - \mu_2$ como parámetro θ, entonces su estimador es $\hat{\theta} = \overline{X} - \overline{Y}$ con desviación estándar $\sigma_{\hat{\theta}}$ dada por la proposición. Cuando σ_1^2 y σ_2^2 tienen valores conocidos, el estadístico de prueba tendrá la forma $(\hat{\theta} - \text{valor nulo})/\sigma_{\hat{\theta}}$; esta forma de un estadístico de prueba se utilizó en varios problemas con una sola muestra del capítulo anterior. Cuando no se conocen los valores de σ_1^2 y σ_2^2, se deben usar las varianzas muestrales para estimar $\sigma_{\hat{\theta}}$.

Procedimientos de prueba para poblaciones normales con varianzas conocidas

En los capítulos 7 y 8, el primer IC y el procedimiento de prueba para una media poblacional μ se basaron en la suposición de la distribución poblacional que era normal y el investigador conocía el valor de la varianza poblacional σ^2. De manera similar, aquí primero se supone que *ambas* distribuciones poblacionales son normales y que se conocen los valores de σ_1^2 y σ_2^2. En breve se presentarán situaciones en las que se puede prescindir de una o ambas suposiciones.

Debido a que las distribuciones poblacionales son normales, tanto $\overline{X}$ como $\overline{Y}$ tienen distribuciones normales. Esto significa que tiene una distribución normal, con valor esperado $\mu_1 - \mu_2$ y desviación estándar $\sigma_{\overline{X}-\overline{Y}}$ dada en la proposición anterior. La estandarización de $\overline{X} - \overline{Y}$ proporciona la variable normal estándar.

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} \tag{9.1}$$

En un problema de prueba de hipótesis, la hipótesis nula expresará que $\mu_1 - \mu_2$ tiene un valor especificado. Si se denota este valor nulo por Δ_0, la hipótesis nula se convierte en H_0: $\mu_1 - \mu_2 = \Delta_0$. A menudo $\Delta_0 = 0$, en cuyo caso H_0 indica que $\mu_1 = \mu_2$. Una estadística de prueba resulta de sustituir $\mu_1 - \mu_2$ en la expresión (9.1) por el valor nulo Δ_0. Debido a que la estadística de prueba Z se obtiene al estandarizar $\overline{X} - \overline{Y}$ bajo el supuesto de que H_0 es verdadera, en este caso tiene una distribución normal estándar. Considere la hipótesis alternativa H_a: $\mu_1 - \mu_2 > \Delta_0$. Un valor $\overline{x} - \overline{y}$ que excede en forma considerable a Δ_0 (el valor esperado de $\overline{X} - \overline{Y}$ cuando H_0 es verdadera) proporciona evidencia contra H_0 y en favor de H_a. Este valor de $\overline{x} - \overline{y}$ corresponde a un valor positivo y grande de z. Por lo tanto, se debe rechazar H_0 en favor de H_a si z es mayor o igual que un valor crítico elegido de manera apropiada. Debido a que el estadístico de prueba Z tiene una distribución normal estándar cuando H_0 es verdadera, la región de rechazo de extremo superior $z \geq z_\alpha$ da una prueba con nivel de significación (probabilidad de error de tipo I) α. Las regiones de rechazo para H_a: $\mu_1 - \mu_2 < \Delta_0$ y H_a: $\mu_1 - \mu_2 \neq \Delta_0$ que producen pruebas con nivel de significación deseado α son de extremo inferior y de dos extremos, respectivamente.

Hipótesis nula: H_0: $\mu_1 - \mu_2 = \Delta_0$

Valor del estadístico de prueba: $z = \dfrac{\overline{x} - \overline{y} - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}}$

Hipótesis alternativa	Región de rechazo para la prueba de nivel α
H_a: $\mu_1 - \mu_2 > \Delta_0$	$z \geq z_\alpha$ (de extremo superior)
H_a: $\mu_1 - \mu_2 < \Delta_0$	$z \leq -z_\alpha$ (de extremo inferior)
H_a: $\mu_1 - \mu_2 \neq \Delta_0$	ya sea $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$ (de dos extremos)

Debido a que éstas son pruebas z, se calcula un valor P como si fuera para las pruebas z del capítulo 8 (p. ej., valor $P = 1 - \Phi(z)$ para una prueba de extremo superior).

Ejemplo 9.1 El análisis de una muestra aleatoria que consiste en $m = 20$ especímenes de acero rolado en frío para determinar las resistencias de cedencia, produjeron una resistencia promedio muestral de $\bar{x} = 29.8$ klb/pulg^2. Una segunda muestra aleatoria de $n = 25$ especímenes de acero galvanizado bilaterales proporcionó una resistencia promedio muestral de $\bar{y} = 34.7$ klb/pulg^2. Suponiendo que las dos distribuciones de resistencia a la cedencia son normales con $\sigma_1 = 4.0$ y $\sigma_2 = 5.0$ (tomados de una gráfica del artículo "Zinc-Coated Sheet Steel: An Overview", *Automotive Engr.*, diciembre de 1984: 39-43), ¿los datos indican que las resistencias a la cedencia promedio reales correspondientes μ_1 y μ_2 son diferentes? Se lleva a cabo una prueba al nivel de significación $\alpha = 0.01$.

1. El parámetro de interés es $\mu_1 - \mu_2$, la diferencia entre las resistencias promedio reales para los dos tipos de acero.
2. La hipótesis nula es H_0: $\mu_1 - \mu_2 = 0$.
3. La hipótesis alternativa es H_a: $\mu_1 - \mu_2 \neq 0$; si H_a es verdadera, entonces μ_1 y μ_2 son diferentes.
4. Con $\Delta_0 = 0$, el valor del estadístico de prueba es
$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}}$$
5. La desigualdad en H_a implica que la prueba es de dos extremos. Para $\alpha = 0.01$, $\alpha/2 = 0.005$ y $z_{\alpha/2} = z_{0.005} = 2.58$. H_0 se rechazará si $z \geq 2.58$ o si $z \leq -2.58$.
6. La sustitución de $m = 20$, $\bar{x} = 29.8$, $\sigma_1^2 = 16.0$, $n = 25$, $\bar{y} = 34.7$ y $\sigma_2^2 = 25.0$ en la fórmula para z produce
$$z = \frac{29.8 - 34.7}{\sqrt{\dfrac{16.0}{20} + \dfrac{25.0}{25}}} = \frac{-4.90}{1.34} = -3.66$$
Es decir, el valor observado de $\bar{x} - \bar{y}$ estría más de tres desviaciones estándar abajo de lo que se esperaría si H_0 fuera cierta.
7. Puesto que $-3.66 < -2.58$, z cae en el extremo inferior de la región de rechazo. Por lo tanto, se rechaza H_0 al nivel 0.01 en favor de la conclusión de que $\mu_1 \neq \mu_2$. Los datos muestrales sugieren fuertemente que la resistencia a la cedencia promedio real para el acero rolado en

frío difiere de la del acero galvanizado. El valor P para esta prueba de dos extremos es $2(1 - \Phi(3.66)) = 2(1 - 1) = 0$, así que se debe rechazar H_0 a *cualquier* nivel de significación razonable. ∎

Cómo usar una comparación para identificar la causalidad

Los investigadores suelen interesarse en comparar ya sea los efectos de dos tratamientos diferentes en una respuesta, o la respuesta después del tratamiento con la respuesta después de ningún tratamiento (tratamiento contra control). Si los investigadores no asignan dos condiciones diferentes a los individuos u objetos que se emplearan en la comparación, se dice que el estudio es **observacional**. La dificultad para sacar conclusiones con base en un estudio observacional es que aunque el análisis estadístico podría indicar una diferencia importante en la respuesta entre los dos grupos, la diferencia podría ser debido a algunos factores subyacentes que no hubieran sido controlados y no a alguna diferencia en los tratamientos.

Ejemplo 9.2 En una carta del *Journal of the American Medical Association* (19 de mayo de 1978) se informa que de 215 médicos varones que se graduaron en Harvard y murieron entre noviembre de 1974 y octubre de 1977, los 125 que ejercieron su profesión tiempo completo, vivieron un promedio de 48.9 años posteriores a su graduación, en tanto que los 90 médicos con afiliaciones académicas vivieron un promedio de 43.2 años contados a partir de su graduación. ¿Los datos indican que el tiempo-de vida promedio después de la graduación para los médicos que ejercieron su profesión tiempo completo es mayor que el tiempo de vida promedio de quienes tuvieron una afiliación académica (en caso afirmativo, los estudiantes de medicina quienes afirman "mueren por obtener una afiliación académica", podrían estar más cerca de la verdad de lo que creen; en otras palabras, la expresión "publica o perece" es en realidad "publica y perece")?

Represente con μ_1 el número promedio de años que viven después de su graduación los médicos que ejercen su profesión tiempo completo, y con μ_2 la misma cantidad para los médicos con afiliaciones académicas. Suponga que los 125 y 90 médicos son muestras aleatorias de las poblaciones 1 y 2, respectivamente (lo cual podría no ser deseable si hay razón para creer que los graduados de Harvard tienen características especiales que los diferencian de los demás médicos; en este caso las inferencias se restringirían sólo a las "poblaciones de Harvard"). La carta de la que se tomaron los datos no proporciona información acerca de las varianzas, así que para ilustración se supone que $\sigma_1 = 14.6$ y $\sigma_2 = 14.4$. Las hipótesis son H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 > 0$, por lo tanto Δ_0 es cero. El valor calculado del estadístico de prueba es

$$z = \frac{48.9 - 43.2}{\sqrt{\dfrac{(14.6)^2}{125} + \dfrac{(14.4)^2}{90}}} = \frac{5.70}{\sqrt{1.70 + 2.30}} = 2.85$$

El valor P para una prueba de extremo superior es $1 - \Phi(2.85) = 0.0022$. Al nivel de significación 0.01, se rechaza H_0 (debido a que $\alpha >$ valor P) en favor de la conclusión de que $\mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$). Esto es consistente con la información registrada en la carta.

Estos datos resultaron de un estudio observacional **retrospectivo**; el investigador no comenzó por elegir una muestra de médicos y asignar algunos al tratamiento "afiliación académica" y los otros al tratamiento "ejercicio de su profesión tiempo completo", sino que identificó a los miembros de los dos grupos al revisar los registros (¡por medio de obituarios!). ¿Se pue-

de distribuir en realidad el resultado estadísticamente importante a una diferencia en el tipo de ejercicio médico después de la graduación, o hay algún otro factor subyacente (p. ej., edad al momento de graduarse, regímenes de ejercicio, etcétera) que podría contribuir también a una explicación posible para la diferencia? Los estudios observacionales han sido utilizados para establecer una relación causal entre el hábito de fumar y el cáncer pulmonar.. ■

Un **experimento controlado aleatorizado** resulta cuando los investigadores asignan sujetos a los dos tratamientos de manera aleatoria. Cuando en este tipo de experimento se observa significación estadística, el investigador y otras personas interesadas tendrán más confianza en la conclusión de que la diferencia en la respuesta se debió a una diferencia en los tratamientos. Un ejemplo muy famoso de este tipo de experimento y conclusión es el experimento de la vacuna contra la polio de Salk descrito en la sección 9.4. Estos temas se analizan con mayor detalle en los libros (no matemáticos) de Moore y de Freedman *et al.*, listados en las referencias del capítulo 1.

β y la elección del tamaño de muestra

La probabilidad de un error de tipo II se calcula con facilidad cuando la distribución poblacional es normal con valores conocidos de σ_1 y σ_2. Considere el caso en el que la hipótesis alternativa es H_a: $\mu_1 - \mu_2 > \Delta_0$. Denote con Δ' un valor de $\mu_1 - \mu_2$ que pase de Δ_0 (un valor para el cual H_0 es falsa). La región de rechazo de extremo superior $z \geq z_\alpha$ se puede expresar de nuevo en la forma $\bar{x} - \bar{y} \geq \Delta_0 + z_\alpha \sigma_{\bar{X}-\bar{Y}}$. Así, la probabilidad de un error de tipo II cuando $\mu_1 - \mu_2 = \Delta'$ es

$$\begin{aligned}\beta(\Delta') &= P(\text{no rechazar } H_0 \text{ cuando } \mu_1 - \mu_2 = \Delta') \\ &= P(\bar{X} - \bar{Y} < \Delta_0 + z_\alpha \sigma_{\bar{X}-\bar{Y}} \text{ cuando } \mu_1 - \mu_2 = \Delta')\end{aligned}$$

Cuando $\mu_1 - \mu_2 = \Delta'$, $\bar{X} - \bar{Y}$ tiene una distribución normal con valor medio Δ' y desviación estándar $\sigma_{\bar{X}-\bar{Y}}$ (la misma desviación estándar que cuando H_0 es verdadera); usando estos valores para estandarizar la desigualdad en el paréntesis se obtiene β.

Hipótesis alternativa	$\beta(\Delta') = P$**(error de tipo II cuando** $\mu_1 - \mu_2 = \Delta'$**)**
H_a: $\mu_1 - \mu_2 > \Delta_0$	$\Phi\left(z_\alpha - \dfrac{\Delta' - \Delta_0}{\sigma}\right)$
H_a: $\mu_1 - \mu_2 < \Delta_0$	$1 - \Phi\left(-z_\alpha - \dfrac{\Delta' - \Delta_0}{\sigma}\right)$
H_a: $\mu_1 - \mu_2 \neq \Delta_0$	$\Phi\left(z_{\alpha/2} - \dfrac{\Delta' - \Delta_0}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \dfrac{\Delta' - \Delta_0}{\sigma}\right)$

donde $\sigma = \sigma_{\bar{X}-\bar{Y}} = \sqrt{(\sigma_1^2/m) + (\sigma_2^2/n)}$

Ejemplo 9.3 (Continuación del ejemplo 9.1) Suponga que cuando μ_1 y μ_2 (las resistencias promedio reales a la cedencia para los dos tipos de acero) difieren por 5, la probabilidad de detectar tal desviación desde H_0 debe ser 0.90. ¿Una prueba de nivel 0.01 con tamaños de muestra $m = 20$ y $n = 25$ satisfacen esta condición? El valor de σ para estos tamaños de muestra (el denominador de z) se calculó antes como 1.34. La probabilidad de un error de tipo II para la prueba de dos extremos al nivel 0.01 cuando $\mu_1 - \mu_2 = \Delta' = 5$ es

$$\begin{aligned}\beta(5) &= \Phi\left(2.58 - \frac{5-0}{1.34}\right) - \Phi\left(-2.58 - \frac{5-0}{1.34}\right)\\ &= \Phi(-1.15) - \Phi(-6.31) = 0.1251\end{aligned}$$

Es fácil comprobar que $\beta(-5) = 0.1251$ también (debido a que la región de rechazo es simétrica). Por consiguiente, la probabilidad de detectar tal desviación es $1 - \beta(5) = 0.8749$. Debido a que este valor es un poco menor que 0.9, se deben usar tamaños de muestra un poco más grandes. ■

Como en el capítulo 8, se determina que los tamaños de muestra m y n satisfarán tanto P(error de tipo I) = una α especificada como P(error de tipo II cuando $\mu_1 - \mu_2 = \Delta'$) = una β especificada. Para una prueba de extremo superior, igualar la expresión previa para $\beta(\Delta')$ a un valor especificado de β, se obtiene

$$\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n} = \frac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2}$$

Cuando los dos tamaños de muestra son iguales, de esta ecuación se obtiene

$$m = n = \frac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\Delta' - \Delta_0)^2}$$

Estas expresiones también son correctas para una prueba de extremo inferior, en tanto que para una prueba de dos extremos $\alpha/2$ sustituye a α.

Pruebas con muestras grandes

Las suposiciones de las distribuciones poblacionales normales y los valores conocidos de σ_1 y σ_2 son innecesarios cuando ambos tamaños de muestra son grandes. En este caso, el Teorema del límite central garantiza que $\overline{X} - \overline{Y}$ tiene aproximadamente una distribución normal, sin considerar las distribuciones poblacionales subyacentes. Además, si se utiliza S_1^2 y S_2^2 en lugar de σ_1^2 y σ_2^2 en la expresión (9.1) se obtiene una variable cuya distribución es aproximadamente normal estándar:

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{m} + \dfrac{S_2^2}{n}}}$$

Un estadístico de prueba de muestra grande resulta de reemplazar $\mu_1 - \mu_2$ por Δ_0, el valor esperado de $\overline{X} - \overline{Y}$ cuando H_0 es verdadera. Entonces este estadístico Z tiene aproximadamente una distribución normal estándar cuando H_0 es verdadera, así que las pruebas de nivel α se obtienen al usar valores críticos z del mismo modo que antes.

El empleo del valor del estadístico de prueba

$$z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

junto con las regiones de rechazo de extremo superior, inferior y de dos extremos mencionadas antes, basadas en valores críticos z, se obtienen pruebas con muestras grandes cuyos niveles de significación son aproximadamente α. Por lo común estas pruebas son apropiadas si $m > 40$ y $n > 40$. Un valor P se calcula del mismo modo que para las pruebas z anteriores.

Ejemplo 9.4 Al seleccionar un concreto azufrado para la construcción de carreteras en regiones que experimentan congelamiento intenso, es importante que el concreto elegido tenga un valor bajo de conductividad térmica, a fin de reducir el daño posterior debido a temperaturas cambiantes. Suponga que para cierta carretera están siendo considerados dos tipos de concreto, uno con agregados graduados y otro sin agregados finos. En la tabla 9.1 se resumen los datos de un experimento realizado para comparar los dos tipos de concreto. ¿Esta información indica que la conductividad promedio real para el concreto graduado es mayor que la del concreto sin agregados finos? Utilice una prueba con $\alpha = 0.01$.

Tabla 9.1 Datos para el ejemplo 9.4

Tipo	Tamaño de muestra	Conductividad promedio muestral	DE muestral
Graduado	42	0.486	0.187
Sin finos	42	0.359	0.158

Represente con μ_1 y μ_2 la conductividad térmica promedio para el concreto con agregados graduados y el concreto sin agregados finos, respectivamente. Las dos hipótesis son H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 > 0$. H_0 será rechazada si $z \geq z_{0.01} = 2.33$. Se calcula

$$z = \frac{0.486 - 0.359}{\sqrt{\frac{(0.187)^2}{42} + \frac{(0.158)^2}{42}}} = \frac{0.127}{0.0378} = 3.36$$

Puesto que $3.36 \geq 2.33$, H_0 se rechaza al nivel de significación 0.01. De manera alternativa, el valor P para una prueba z de extremo superior es

$$\text{Valor } P = 1 - \Phi(z) = 1 - \Phi(3.36) = 0.0004$$

H_0 se debe rechazar no sólo para una prueba con $\alpha = 0.01$ sino también para $\alpha = 0.001$ o cualquier otro valor de α que pase de 0.0004. Los datos favorecen sin lugar a dudas la conclusión de que la conductividad térmica promedio para el concreto graduado no es mayor que la del concreto sin agregados finos. ■

Intervalos de confianza para $\mu_1 - \mu_2$

Cuando las distribuciones poblacionales son normales, la estandarización $\overline{X} - \overline{Y}$ da una variable aleatoria Z con una distribución normal estándar. Puesto que el área bajo la curva z entre $-z_{\alpha/2}$ y $z_{\alpha/2}$ se deduce que

$$P\left(-z_{\alpha/2} < \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} < z_{\alpha/2}\right) = 1 - \alpha$$

El manejo de las desigualdades dentro de los paréntesis para aislar $\mu_1 - \mu_2$ produce el enunciado de probabilidad equivalente

$$P\left(\overline{X} - \overline{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} < \mu_1 - \mu_2 < \overline{X} - \overline{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}\right) = 1 - \alpha$$

Esto significa que un IC de $100(1 - \alpha)\%$ para $\mu_1 - \mu_2$ tiene un límite inferior $\bar{x} - \bar{y} - z_{\alpha/2} \cdot \sigma_{\bar{X}-\bar{Y}}$ y límite superior $\bar{x} - \bar{y} + z_{\alpha/2} \cdot \sigma_{\bar{X}-\bar{Y}}$, donde $\sigma_{\bar{X}-\bar{Y}}$ es la expresión de la raíz cuadrada. Este intervalo es un caso especial de la fórmula general $\hat{\theta} \pm z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$.

Si m y n son grandes, el TLC implica que este intervalo es válido incluso sin la suposición de poblaciones normales; en este caso, el nivel de confianza es *aproximadamente* $100(1 - \alpha)\%$. Además, el uso de las varianzas muestrales S_1^2 y S_2^2 en la variable estandarizada Z produce un intervalo válido en el que s_1^2 y s_2^2 sustituyen a σ_1^2 y σ_2^2.

> Siempre y cuando m y n sean grandes, un IC para $\mu_1 - \mu_2$ con un nivel de confianza de alrededor de $100(1 - \alpha)\%$ es
>
> $$\bar{x} - \bar{y} \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$
>
> donde el signo $-$ da el límite inferior y el $+$ da el límite superior del intervalo. Un límite de confianza superior o inferior también se puede calcular al retener el signo apropiado ($+$ o $-$) y sustituir $z_{\alpha/2}$ por z_α.

La regla empírica estándar para caracterizar tamaños de muestra grandes es $m > 40$ y $n > 40$.

Ejemplo 9.5 Un experimento fue llevado a cabo para estudiar varias características de pernos de anclaje en 78 observaciones de resistencia cortante (klb/pulg^2) de pernos con diámetro de 3/8 de pulgada y 88 observaciones sobre la resistencia de pernos con diámetro de 1/2 pulgada. A continuación se muestran las cantidades de resumen de MINITAB, y las desviaciones estándar muestrales concuerdan con los valores dados en el artículo "Ultimate Load Capacities of Expansion Anchor Bolts" (*J. Energy Engr.*, 1993: 139-158). Los resúmenes indican que la diferencia principal entre las dos muestras está en donde se centran.

```
Variable      N      Media     Mediana      Media      Desv.     EE de
                                            rec.       est.  la media
diám 3/8     78      4.250       4.230     4.238      1.300     0.147

Variable    Mín        Máx          Q1         Q3
diám 3/8  1.634      7.327       3.389      5.075

Variable      N      Media     Mediana      Media      Desv.     EE de
                                            rec.       est.  la media
diám 1/2     88      7.140       7.113     7.150      1.680     0.179

Variable    Mín        Máx          Q1         Q3
diám 1/2  2.450     11.343       5.965      8.447
```

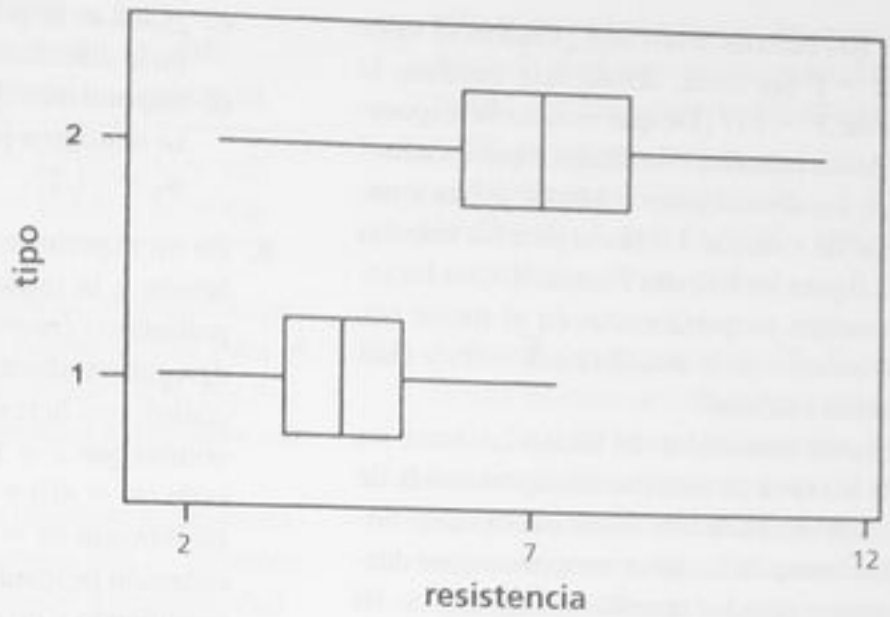


Figura 9.1 Diagrama de caja comparativo de los datos de resistencia al corte

Ahora se calcula un intervalo de confianza para la diferencia entre la resistencia al corte promedio real para los pernos de 3/8 de pulgada (μ_1) y la resistencia al corte promedio real para los pernos de 1/2 pulgada (μ_2) por medio de un nivel de confianza de 95%:

$$4.25 - 7.14 \pm (1.96)\sqrt{\frac{(1.30)^2}{78} + \frac{(1.68)^2}{88}} = -2.89 \pm (1.96)(0.2318)$$

$$= -2.89 \pm 0.45 = (-3.34, -2.44)$$

Es decir, con un nivel de confianza de 95%, $-3.34 < \mu_1 - \mu_2 < -2.44$. Por lo tanto, se puede tener un nivel de confianza alto de que la resistencia al corte promedio real para los pernos de 1/2 pulgada es mayor que la de los pernos de 3/8 de pulgada por entre 2.44 klb/pulg^2 y 3.34 klb/pulg^2. Observe que si se vuelve a marcar de modo que μ_1 se refiera a los pernos de 1/2 pulgada y μ_2 a los pernos de 3/8 pulgada, el intervalo de confianza ahora está centrado en +2.89 y el valor 0.45 aún se resta y se suma para obtener los límites de confianza. El intervalo resultante es (2.44, 3.34), y la interpretación es idéntica a la del intervalo calculado antes. ■

Si las varianzas σ_1^2 y σ_2^2 se conocen por lo menos de forma aproximada y el investigador utiliza tamaños de muestra iguales, entonces el tamaño muestral n para cada muestra que produce un intervalo de $100(1 - \alpha)\%$ de amplitud w es

$$n = \frac{4z_{\alpha/2}^2(\sigma_1^2 + \sigma_2^2)}{w^2}$$

que por lo general tiene que ser redondeado al entero inmediato superior.

Ejercicios | Sección 9.1 (1 a 16)

1. En un artículo de *Consumer Reports* de noviembre de 1983 se comparan varios tipos de baterías. Los tiempos de vida promedio de baterías AA alcalinas Duracell y baterías AA alcalinas Eveready Energizer se dieron como 4.1 horas y 4.5 horas, respectivamente. Suponga que éstos son tiempos de vida promedio poblacionales.
   a. Sea $\overline{X}$ el tiempo de vida promedio muestral de 100 baterías Duracell y $\overline{Y}$ el tiempo de vida promedio

muestral de 100 baterías Eveready. ¿Cuál es el valor medio de $\bar{X} - \bar{Y}$ (es decir, dónde está centrada la distribución de $\bar{X} - \bar{Y}$)? ¿De qué manera la respuesta depende de los tamaños muestrales especificados?

b. Suponga que las desviaciones estándar poblacionales del tiempo de vida son 1.8 horas para las baterías Duracell y 2.0 para las baterías Eveready. Con los tamaños de muestra proporcionados en el inciso (a), ¿cuál es la varianza de la estadística $\bar{X} - \bar{Y}$, y cuál es la desviación estándar?

c. Para los tamaños muestrales del inciso (a), trace un diagrama de la curva de distribución aproximada de $\bar{X} - \bar{Y}$ (incluya una escala de medición en el eje horizontal). ¿La forma de la curva necesariamente describiría lo mismo para los tamaños muestrales de 10 baterías de cada tipo? Explique

2. Sean μ_1 y μ_2 las duraciones promedio verdaderas de la banda de rodamiento de dos marcas contrarias de llantas radiales tamaño P205/65R15. Pruebe H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 \neq 0$ al nivel 0.05 usando los siguientes datos: $m = 45$, $\bar{x} = 42\,500$, $s_1 = 2200$, $n = 45$, $\bar{y} = 40\,400$, y $s_2 = 1900$.

3. Sea μ_1 la duración promedio real de la banda de rodamiento para una marca premium de una llanta radial P205/65R15 y sea μ_2 la duración promedio real de la banda de rodamiento para una marca económica del mismo tamaño. Pruebe H_0: $\mu_1 - \mu_2 = 5000$ contra H_a: $\mu_1 - \mu_2 > 5000$ al nivel 0.01 por medio de los siguientes datos: $m = 45$, $\bar{x} = 42\,500$, $s_1 = 2200$, $n = 45$, $\bar{y} = 36\,800$, y $s_2 = 1500$.

4. **a.** Utilice los datos del ejercicio 2 para calcular un IC de 95% para $\mu_1 - \mu_2$. ¿El intervalo resultante indica que se estimó con precisión $\mu_1 - \mu_2$?

b. Utilice los datos del ejercicio 3 para calcular un límite de confianza superior de 95 para $\mu_1 - \mu_2$.

5. Las personas con síndrome de Reynaud están propensas a sufrir un deterioro repentino de la circulación sanguínea en los dedos de manos y pies. En un experimento para estudiar el grado de este deterioro, cada individuo sumergió un dedo índice en agua y se midió la producción calorífica resultante (cal/cm²/min). Para $m = 10$ individuos con el síndrome la producción calorífica promedio fue $\bar{x} = 0.64$ y para $n = 10$ individuos sin el síndrome, la producción promedio fue 2.05. Sean μ_1 y μ_2 las producciones caloríficas promedio para los dos tipos de individuos. Suponga que las dos distribuciones de producción calorífica son normales con $\sigma_1 = 0.2$ y $\sigma_2 = 0.4$.

a. Considere probar H_0: $\mu_1 - \mu_2 = -1.0$ contra H_a: $\mu_1 - \mu_2 < -1.0$ al nivel 0.01. Describa en palabras lo que dice H_a y luego realice la prueba.

b. Calcule el valor P para el valor de Z obtenido en el inciso (a).

c. ¿Cuál es la probabilidad de un error de tipo II cuando la diferencia real entre μ_1 y μ_2 is $\mu_1 - \mu_2 = -1.2$?

d. Suponiendo que $m = n$, ¿que tamaños de muestra se requieren para asegurar que $\beta = .1$ cuando $\mu_1 - \mu_2 = -1.2$?

6. En un experimento para comparar la resistencia de cohesión a la tensión del mortero modificado con látex polimérico (mortero de cemento Portland al que se le agregan emulsiones de látex polimérico durante el mezclado), con la resistencia del mortero no modificado se obtuvo que $\bar{x} = 18.12$ kgf/cm² para el mortero modificado ($m = 40$) y $\bar{y} = 16.87$ kgf/cm² para el mortero no modificado ($n = 32$). Sean μ_1 y μ_2 las resistencias de cohesión promedio reales a la tensión para los morteros modificado y no modificado, respectivamente. Suponga que las distribuciones de resistencia de cohesión son normales.

a. Suponiendo que $\sigma_1 = 1.6$ y $\sigma_2 = 1.4$, pruebe H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 > 0$ al nivel 0.01.

b. Calcule la probabilidad de un error de tipo II para la prueba del inciso (a) cuando $\mu_1 - \mu_2 = 1$.

c. Suponga que el investigador decidió usar una prueba de nivel 0.05 y deseó una valor de $\beta = 0.10$ cuando $\mu_1 - \mu_2 = 1$. Si $m = 40$, ¿qué valor de n es necesario?

d. ¿Cómo cambiaría el análisis y la conclusión del inciso (a) si no se conocieran σ_1 y σ_2 pero $s_1 = 1.6$ y $s_2 = 1.4$?

7. ¿Se aburren con más facilidad los estudiantes universitarios varones que sus contrapartes mujeres? Esta pregunta se examinó en el artículo "Boredom in Young Adults-Gender and Cultural Comparisons" (*J. of Cross-Cultural Psych.*, 1991: 209-223). Los autores proporcionaron una escala llamada *Boredom Proneness Scale* (Escala de propensión al aburrimiento) para 97 varones y 148 mujeres estudiantes universitarios estadounidenses. ¿Los datos siguientes respaldan la hipótesis de investigación de que el índice medio de propensión al aburrimiento es mayor para varones que para mujeres? Pruebe la hipótesis apropiada con un nivel de significación 0.05.

Género	Tamaño de la muestra	Media muestral	DE de la muestra
Masculino	97	10.40	4.83
Femenino	148	9.26	4.68

8. Se llevaron a cabo pruebas de resistencia a la tensión en dos grados de varilla de alambre ("Fluidized Bed Patenting of Wire Rods", *Wire J.*, junio de 1977: 56-61), y se obtuvieron los datos siguientes.

Grado	Tamaño de la muestra	Media muestral (kg/mm²)	DE muestral
AISI 1064	$m = 129$	$\bar{x} = 107.6$	$s_1 = 1.3$
AISI 1078	$n = 129$	$\bar{y} = 123.6$	$s_2 = 2.0$

a. ¿Los datos proporcionan evidencia convincente para concluir que la resistencia promedio real para el grado 1078 es mayor que la del grado 1064, por más de 10 kg/mm²? Pruebe las hipótesis apropiadas con el método del valor P.
b. Estime la diferencia entre las resistencias promedio reales para los dos grados de modo que proporcione información acerca de la precisión y confiabilidad.

9. En el artículo "Evaluation of a Ventilation Strategy to Prevent Barotrauma in Patients at High Risk for Acute Respiratory Distress Syndrome" (*New Engl. J. of Med.*, 1998: 355-358) se describe un experimento en el que 120 pacientes con características clínicas similares, se dividieron al azar en un grupo de control y un grupo de tratamiento, cada una con 60 pacientes. La estancia promedio muestral en la UCI (días) y la desviación estándar muestral para el grupo de tratamiento, fueron 19.9 y 39.1, respectivamente, mientras que estos valores para el grupo de control fueron 13.7 y 15.8.
   a. Calcule una estimación puntual para la diferencia entre la estancia promedio real en la UCI para los grupos de tratamiento y control. ¿Esta estimación indica que hay una diferencia importante entre las estancias promedio reales en las dos condiciones?
   b. Conteste la pregunta planteada en el inciso (a) mediante la realización de una prueba de hipótesis formal. ¿Es diferente el resultado de lo que se supuso en el inciso (a)?
   c. ¿Al parecer la permanencia en la UCI para los pacientes que recibieron tratamiento ventilatorio tiene una distribución normal? Explique su razonamiento.
   d. Estime la duración promedio real de estancia para pacientes que recibieron tratamiento ventilatorio, de manera que transmita información acerca de la precisión y la confiabilidad.

10. Se llevó a cabo un experimento para comparar la resistencia a la fractura de acero *maraging* (aleación de níquel al 18%) de alta pureza con un acero de pureza comercial del mismo tipo (*Corrosion Science*, 1971: 723-736). Para $m = 32$ especímenes, la resistencia promedio muestral fue $\bar{x} = 65.6$ para el acero de alta pureza, mientras que para $n = 38$ especímenes de acero comercial $\bar{y} = 59.8$. Debido a que el acero de alta pureza es más caro, su uso para cierta aplicación se justifica sólo si su resistencia a la fractura es mayor que la del acero de pureza comercial por más de 0.5. Suponga que ambas distribuciones de resistencia son normales.
   a. Si se supone que $\sigma_1 = 1.2$ y $\sigma_2 = 1.1$, pruebe las hipótesis pertinentes con $\alpha = 0.001$.
   b. Calcule β para la prueba realizada en el inciso (a) cuando $\mu_1 - \mu_2 = 6$.

11. Se determinó la concentración de plomo en la sangre para una muestra de 152 trabajadores varones, con edades entre 20 y 30 años, que manejan desechos peligrosos y también para una muestra de 86 trabajadoras mujeres. Se obtuvo una media ± error estándar de 5.5 ± 0.3 para los varones y 3.8 ± 0.2 para las mujeres ("Temporal Changes in Blood Lead Levels of Hazardous Waste Workers in New Jersey, 1984-1987", *Environ. Monitoring and Assessment*, 1993: 99-107). Calcule una estimación de la diferencia entre las concentraciones promedio verdaderas de plomo en la sangre para varones y mujeres de modo que proporcione información acerca de la confiabilidad y precisión.

12. En la tabla siguiente se dan los datos de resumen en relación con la resistencia a la compresión de cubos (N/mm²) para especímenes de concreto hechos con una mezcla pulverizada de ceniza y combustible ("A Study of Twenty-Five-Year-Old Pulverized Fuel Ash Concrete Used in Foundation Structures", *Proc. Inst. Civ. Engrs.*, marzo de 1985: 149-165):

Edad (días)	Tamaño de la muestra	Media muestral	DE de la muestra
7	68	26.99	4.89
28	74	35.76	6.43

Calcule e interprete un IC de 99% para la diferencia entre la resistencia promedio real de siete días y la de 28 días.

13. Un ingeniero mecánico desea comparar las propiedades de resistencia de vigas de acero con vigas similares hechas con una aleación particular. Se probará el mismo número de vigas, n, de cada tipo. Cada viga se probará en una posición horizontal con un apoyo en cada extremo, se aplicará una fuerza de 2500 lb en el centro y se medirá la deflexión. A partir de experiencias pasadas con este tipo de vigas, el ingeniero está dispuesto a asumir que la desviación estándar verdadera de la deflexión para ambos tipos de viga es 0.05 pulg. Debido a que la aleación es más cara, el ingeniero desea probar al nivel 0.01 si tiene una deflexión promedio menor que la

viga de acero. ¿Qué valor de n es apropiado si la probabilidad deseada de error de tipo II es 0.05 cuando la diferencia en la deflexión promedio real favorece a la aleación por 0.04 pulg.?

14. Al determinar el nivel de actividad de la oxidasa de monoamina (OMA) en las plaquetas de la sangre (nm/mg de proteína/h) para cada individuo de una muestra de 43 esquizofrénicos crónicos, se encontró que $\bar{x} = 2.69$ y $s_1 = 2.30$. Lo mismo se hizo para 45 individuos normales y se obtuvo que $\bar{y} = 6.35$ y $s_2 = 4.03$. ¿Estos datos indican sin lugar a dudas que la actividad promedio real de la OMA para individuos normales es más del doble que el nivel de actividad para esquizofrénicos? Derive un procedimiento de prueba y lleve a cabo la prueba con $\alpha = 0.01$. [*Sugerencia*: aquí H_0 y H_a tienen una forma diferente de la de los tres casos estándar. Asigne μ_1 y μ_2 a la actividad promedio real de la OMA para personas esquizofrénicas y normales, respectivamente, y considere el parámetro $\theta = 2\mu_1 - \mu_2$. Escriba H_0 y H_a en términos de θ, estime θ, y deduzca $\hat{\sigma}_\theta$ ("Reduced Monoamine Oxidase Activity en Blood Platelets from Schizophrenic Patients", *Nature*, 28 de julio, 1972: 225-226).]

15. a. Demuestre para la prueba de extremo superior con σ_1 y σ_2 conocidas que cuando m o n crecen, β disminuye cuando $\mu_1 - \mu_2 > \Delta_0$.
    b. Para el caso de tamaños de muestra iguales ($m = n$) y α fija, ¿qué sucede con el tamaño de muestra necesario n cuando disminuye β, donde β es la probabilidad deseada de error de tipo II a una alternativa fija?

16. Para decidir si dos tipos de acero tienen los mismos valores de resistencia a la fractura promedio reales, se prueban n especímenes de cada tipo, y se obtienen los resultados siguientes:

Tipo	Promedio muestral	DE muestral
1	60.1	1.0
2	59.9	1.0

Calcule el valor P para la prueba z apropiada con dos muestras, suponiendo que los datos se basan en $n = 100$. Luego, repita el cálculo para $n = 400$. ¿El pequeño valor P para $n = 400$ es indicativo de una diferencia que tiene significación práctica? ¿Se sentiría satisfecho con sólo un valor P? Comente de manera breve.

9.2 Prueba *t* de dos muestras e intervalo de confianza

En la práctica, casi siempre sucede que se desconocen los valores de las varianzas poblacionales. En la sección anterior se ilustró, para tamaños de muestras grandes, el uso de un procedimiento de prueba e IC en el que se emplearon las varianzas muestrales en lugar de las varianzas poblacionales. De hecho, para muestras grandes, el TLC permite usar estos métodos aun cuando no son normales las dos poblaciones de interés.

Sin embargo, hay muchos problemas en los que por lo menos un tamaño de muestra es pequeño y las varianzas poblacionales tienen valores desconocidos. En ausencia del TLC, se procede haciendo suposiciones específicas acerca de las distribuciones poblacionales subyacentes. El uso de procedimientos inferenciales que se deducen de estas suposiciones se restringe entonces a situaciones en las que las suposiciones se satisfacen por lo menos de manera aproximada.

SUPOSICIONES

Ambas poblaciones son normales, de modo que $X_1, X_2, \ldots, X_m$ es una muestra aleatoria de una distribución normal, y también lo es $Y_1, \ldots, Y_n$ (con las X y las Y independientes entre sí). La factibilidad de estas suposiciones se juzga al construir una gráfica de probabilidad normal para las x_i y otra para las y_i.

La fórmula del estadístico de prueba y el intervalo de confianza se basa en la misma variable estandarizada de la sección 9.1, pero la distribución pertinente ahora es t y no z.

TEOREMA

Cuando las distribuciones poblacionales son normales, la variable estandarizada

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \tag{9.2}$$

tiene aproximadamente una distribución t con ν gl estimados a partir de los datos mediante la fórmula

$$\nu = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}} = \frac{[(se_1)^2 + (se_2)^2]^2}{\frac{(se_1)^4}{m-1} + \frac{(se_2)^4}{n-1}}$$

donde

$$se_1 = \frac{s_1}{\sqrt{m}}, \; se_2 = \frac{s_2}{\sqrt{n}}$$

(ν se redondea al entero inferior).

Al manipular T en un enunciado de probabilidad para aislar $\mu_1 - \mu_2$ se obtiene un IC, mientras que el estadístico de prueba resulta de reemplazar $\mu_1 - \mu_2$ por el valor nulo Δ_0.

Procedimientos para t con dos muestras

El **intervalo de confianza t con dos muestras para $\mu_1 - \mu_2$** con nivel de confianza $100(1 - \alpha)\%$ es

$$\bar{x} - \bar{y} \pm t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

Un límite de confianza unilateral se puede calcular como se describió antes.

La **prueba t bimuestral** para probar H_0: $\mu_1 - \mu_2 = \Delta_0$ es como sigue:

$$\text{Valor del estadístico de prueba: } t = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

Hipótesis alternativa	**Región de rechazo para una prueba aproximada de nivel α**
H_a: $\mu_1 - \mu_2 > \Delta_0$	$t \geq t_{\alpha,\nu}$ (de extremo superior)
H_a: $\mu_1 - \mu_2 < \Delta_0$	$t \leq -t_{\alpha,\nu}$ (de extrmo inferior)
H_a: $\mu_1 - \mu_2 \neq \Delta_0$	ya sea $t \geq t_{\alpha/2,\nu}$ o $t \leq -t_{\alpha/2,\nu}$ (de dos extremos)

Un valor P se calcula como se describió en la sección 8.4 para la prueba t de una sola muestra.

Ejemplo 9.6 El volumen de huecos dentro de una tela afecta las propiedades de confort, inflamabilidad y aislamiento. La permeabilidad de una tela se refiere a la accesibilidad del espacio de huecos para el flujo de un gas o líquido. En el artículo "The Relationship Between Porosity and Air Permeability of Woven Textile Fabrics" (*J. of Testing and Eval.*, 1997: 108-114) se da la información resumida de la permeabilidad al aire ($cm^3/cm^2/s$) para varios tipos de telas. Considere los siguientes datos acerca de dos tipos de tela de tejido sencillo:

Tipo de tela	Tamaño de la muestra	Media muestral	DE muestral
Algodón	10	51.71	0.79
Triacetato	10	136.14	3.59

Suponiendo que las distribuciones de porosidad para ambos tipos de tela son normales, se calcula un intervalo de confianza para la diferencia entre la porosidad promedio para la tela de algodón y la de la tela de acetato, usando un nivel de confianza de 95%. Antes de poder seleccionar el valor crítico t apropiado, se deben determinar los grados de libertad:

$$\text{gl} = \frac{\left(\dfrac{0.6241}{10} + \dfrac{12.8881}{10}\right)^2}{\dfrac{(0.6241/10)^2}{9} + \dfrac{(12.8881/10)^2}{9}} = \frac{1.8258}{0.1850} = 9.87$$

Por lo tanto, se usa $\nu = 9$; de la tabla A.5 del apéndice se obtiene $t_{.025,9} = 2.262$. El intervalo resultante es

$$51.71 - 136.14 \pm (2.262)\sqrt{\frac{0.6241}{10} + \frac{12.8881}{10}} = -84.43 \pm 2.63$$

$$= (-87.06, -81.80)$$

Con un alto grado de confianza, se puede decir que la porosidad promedio para los especímenes de tela de triacetato sobrepasa la de los especímenes de algodón por entre 81.80 y 87.06 $cm^3/cm^2/$segundo. ∎

Ejemplo 9.7 El deterioro de muchas redes de tubería municipales en todo el país es una preocupación creciente. Una tecnología, propuesta para la rehabilitación de tuberías, utiliza un revestimiento flexible insertado en la tubería existente. En el artículo "Effect of Welding on a High-Density Polyethylene Liner" (*J. of Materials in Civil Engr.*, 1996: 94-100) se proporcionan los siguientes datos acerca de la resistencia a la tensión ($lb/pulg^2$) de especímenes de revestimiento cuando se utilizó cierto tipo de fusión, y cuando no se usó este proceso.

Sin fusión	2748	2700	2655	2822	2511			
	3149	3257	3213	3220	2753			
	$m = 10$	$\bar{x} = 2902.8$		$s_1 = 277.3$				
Fundido	3027	3356	3359	3297	3125	2910	2889	2902
	$n = 8$	$\bar{y} = 3108.1$		$s_2 = 205.9$				

En la figura 9.2 se muestran las gráficas de probabilidad normal de MINITAB. El patrón lineal en cada gráfica, respalda la suposición de que ambas distribuciones de resistencia a la tensión bajo las dos condiciones son normales.

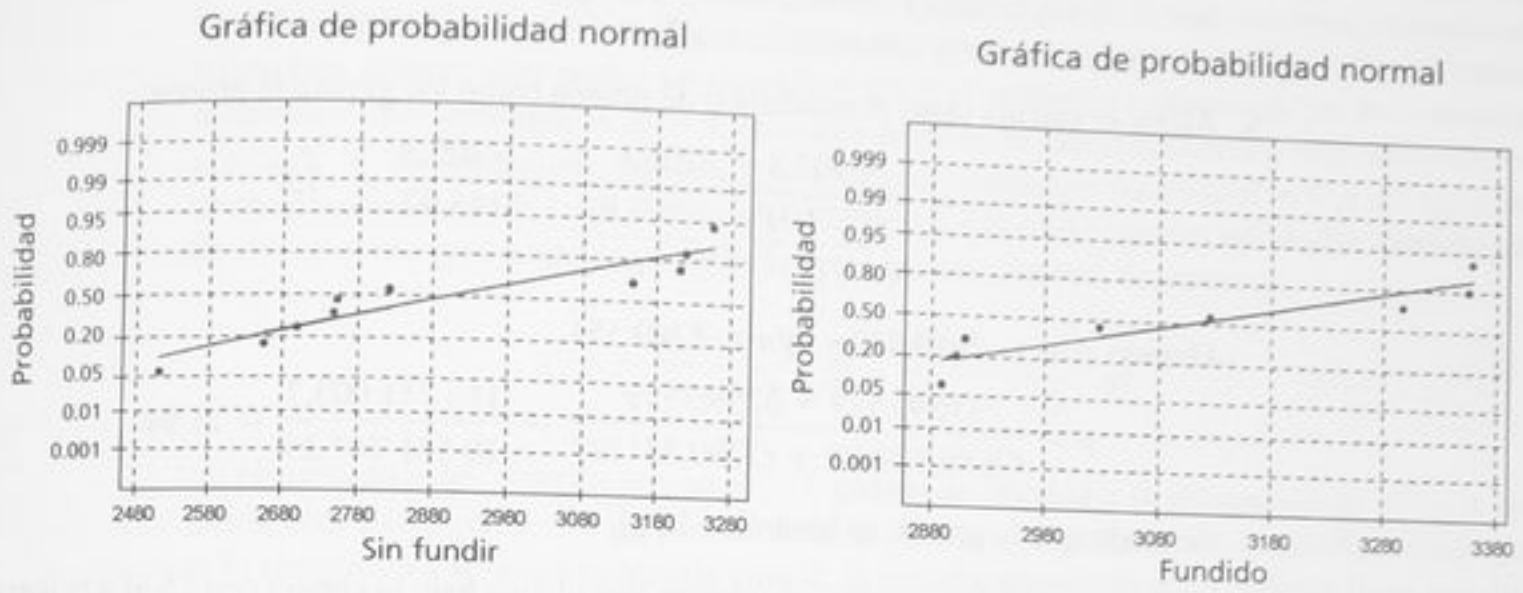


Figura 9.2 Gráficas de probabilidad normal de MINITAB para los datos de resistencia a la tensión

Los autores del artículo expresaron que el proceso de fusión incrementó la resistencia promedio a la tensión. El mensaje del diagrama de caja comparativo de la figura 9.3 no es del todo claro. Se realiza una prueba de hipótesis para ver si los datos apoyan esta conclusión.

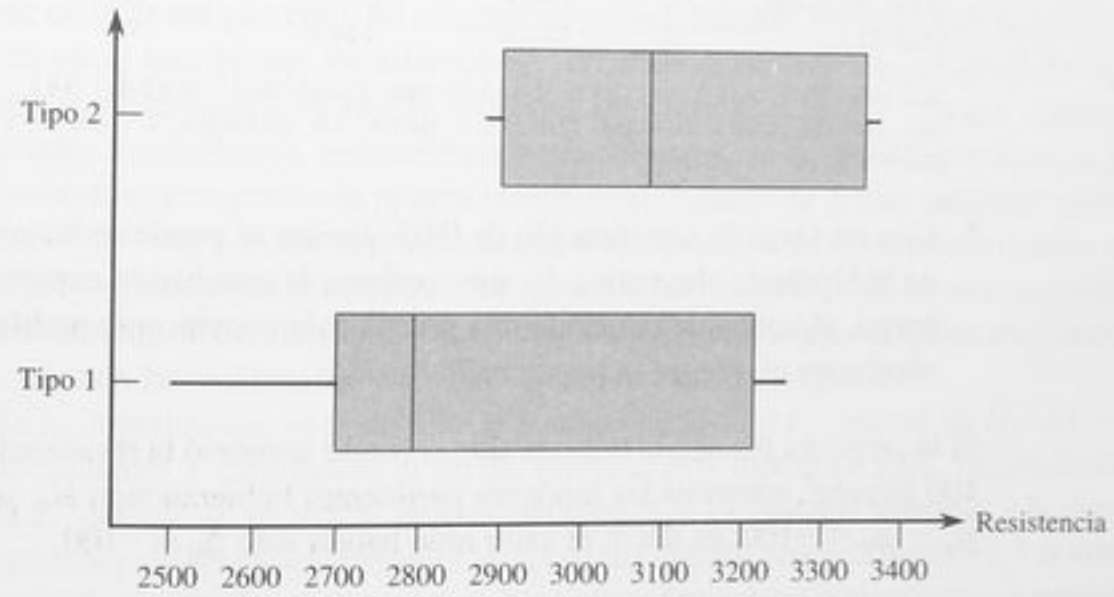


Figura 9.3 Diagrama de caja comparativo de los datos de resistencia a la tensión

1. Sea μ_1 la resistencia promedio real a la tensión de especímenes cuando se emplea tratamiento sin fusión y μ_2 la resistencia promedio real a la tensión cuando se utiliza el tratamiento de fusión.
2. H_0: $\mu_1 - \mu_2 = 0$ (ninguna diferencia en las resistencias a la tensión promedio reales para los dos tratamientos).
3. H_a: $\mu_1 - \mu_2 < 0$ (la resistencia promedio real a la tensión para el tratamiento sin fusión es menor que para el tratamiento con fusión, de modo que es correcta la conclusión de los investigadores).

4. El valor nulo es $\Delta_0 = 0$, por lo tanto el estadístico de prueba es

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

5. Ahora se calcula tanto el estadístico de prueba como los gl para la prueba:

$$t = \frac{2902.8 - 3108.1}{\sqrt{\frac{(277.3)^2}{10} + \frac{(205.9)^2}{8}}} = \frac{-205.3}{113.97} = -1.8$$

Usando $s_1^2/m = 7689.529$ y $s_2^2/n = 5299.351$,

$$\nu = \frac{(7689.529 + 5299.351)^2}{(7689.529)^2/9 + (5299.351)^2/7} = \frac{168\,711\,003.7}{10\,581\,747.35} = 15.94$$

de modo que la prueba se basará en 15 gl.

6. En la tabla A.8 del apéndice se muestra que el área bajo la curva t con 15 gl a la derecha de 1.8 es 0.046, así que el valor P para una prueba de extremo inferior es 0.046. En el siguiente resultado de MINITAB se resumen los cálculos:

```
T con dos muestras para ninguna fusión vs fusión

              N    Media     Desv.      EE de
                           est.    la media
Sin fusión   10     2903      277          88
Fundido       8     3108      206          73

IC de 95% para mu sin fusión-mu fundido: (-488, 38)
mu de no fusión = mu fundido para la prueba T (vs <): T = -1.80 P = 0.046
GL = 15
```

7. Con un nivel de significación de 0.05, apenas se puede rechazar la hipótesis nula en favor de la hipótesis alternativa, lo que confirma la conclusión expresada en el artículo. Sin embargo, alguien que demande una prueba más convincente podría seleccionar $\alpha = 0.01$, un nivel para el que no se puede rechazar H_0.

Si la pregunta planteada hubiera sido si fundir aumentó la resistencia promedio real en más de 100 lb/pulg^2, entonces las hipótesis pertinentes hubieran sido H_0: $\mu_1 - \mu_2 = -100$ contra H_a: $\mu_1 - \mu_2 < -100$; es decir, el valor nulo habría sido $\Delta_0 = -100$. ■

Procedimientos *t* combinados (pooled)

Una alternativa para los procedimientos t con dos muestras recién descritos, resulta de suponer no sólo que las distribuciones poblacionales son normales, sino también que tienen varianzas iguales ($\sigma_1^2 = \sigma_2^2$). Es decir, se supone que las curvas de distribución poblacional son normales con dispersiones iguales, la única diferencia posible entre ellas es dónde están centradas.

Sea que σ^2 representa la varianza poblacional común. Entonces al estandarizar $\bar{X} - \bar{Y}$, se obtiene

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{m} + \frac{\sigma^2}{n}}} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2\left(\frac{1}{m} + \frac{1}{n}\right)}}$$

que tiene una distribución normal estándar. Antes de poder usar esta variable como base para hacer inferencias acerca de $\mu_1 - \mu_2$, se debe estimar la varianza común a partir de los datos muestrales. Un estimador de σ^2 es S_1^2, la varianza de las m observaciones en la primera muestra, y el otro es S_2^2, la varianza de la segunda muestra. Desde el punto de vista intuitivo, un mejor estimador que cualquier varianza muestral individual resulta de combinar las dos varianzas muestrales. Una primera idea podría ser usar $(S_1^2 + S_2^2)/2$, el promedio ordinario de las dos varianzas muestrales. Sin embargo, si $m > n$, entonces la primera muestra contiene más información acerca de σ^2 que la segunda muestra, y se aplica un comentario parecido si $m < n$. El siguiente promedio *ponderado* de las dos varianzas muestrales, conocido como el **estimador combinado de σ^2,** se ajusta para cualquier diferencia entre los dos tamaños de muestra:

$$S_p^2 = \frac{m-1}{m+n-2} \cdot S_1^2 + \frac{n-1}{m+n-2} \cdot S_1^2$$

La primera muestra contribuye con $m - 1$ grados de libertad a la estimación de σ^2, y la segunda contribuye con $n - 1$ gl, para un total de $m + n - 2$ gl. La teoría estadística señala que si S_p^2 reemplaza a σ^2 en la expresión para Z, la variable estandarizada resultante tiene una distribución t basada en $m + n - 2$ gl. En la misma forma que las variables estandarizadas anteriores se usaron como base para obtener intervalos de confianza y procedimientos de prueba, esta variable t da lugar de inmediato al intervalo de confianza t combinado para estimar $\mu_1 - \mu_2$ y la prueba t combinada para probar hipótesis acerca de una diferencia entre medias.

En el pasado, muchos estadísticos recomendaban estos procedimientos t combinados en lugar de los procedimientos t con dos muestras. Por ejemplo, el procedimiento t combinado, se puede deducir del principio de relación de probabilidades, en tanto que la prueba t con dos muestras no es una prueba de relación de probabilidades. Además, el nivel de significación para prueba t combinada es exacto, mientras que para la prueba t con dos muestras sólo es aproximado. Sin embargo, en investigaciones recientes, se demostró que aunque la prueba t combinada supera un poco a la prueba t bimuestral (valores de β más pequeños para la misma α) cuando $\sigma_1^2 = \sigma_2^2$, la primera prueba conduce fácilmente a conclusiones erróneas si se aplica cuando las varianzas son diferentes. Comentarios similares se aplican al comportamiento de los dos intervalos de confianza. Es decir, los procedimientos t combinados no son consistentes en relación con las violaciones de la suposición de varianzas iguales.

Se ha sugerido que se podría llevar a cabo una prueba preliminar de H_0: $\sigma_1^2 = \sigma_2^2$ y usar un procedimiento t combinado si no se rechaza esta hipótesis nula. Infortunadamente, la "prueba F" común de varianzas iguales (sección 9.5) es bastante sensible a la suposición de distribuciones poblacionales normales, mucho más que los procedimientos t. Por consiguiente, se recomienda el método conservador de usar procedimientos t bimuestrales, a menos que haya evidencia realmente convincente para hacer lo contrario, en particular cuando los dos tamaños de muestra son diferentes.

Probabilidades de error de tipo II

Resulta complicado determinar las probabilidades de error de tipo II (o de manera equivalente, potencia $= 1 - \beta$) para la prueba t bimestral es complicado. Al parecer no hay una forma simple de usar las curvas β de la tabla A.17 del apéndice. Con la versión más reciente de MINITAB (Versión 13) se calcula la potencia para la prueba t combinada pero no para la prueba t bimuestral. Sin embargo, la página del Departamento de estadística de la UCLA (http://www.stat.ucla.edu) permite el acceso a una calculadora de potencia que hará esto. Por ejemplo, se especifica $m = 10$, $n = 8$, $\sigma_1 = 300$, $\sigma_2 = 225$ (éstos son los tamaños de muestra

para el ejemplo 9.7, cuyas desviaciones estándar muestrales son un poco más pequeñas que estos valores de σ_1 y σ_2) y se pide la potencia para una prueba de dos extremos al nivel 0.05 de H_0: $\mu_1 - \mu_2 = 0$ cuando $\mu_1 - \mu_2 = 100$, 250, y 500. Los valores resultantes de la potencia fueron 0.1089, 0.4609 y 0.9635 (que corresponden a $\beta = 0.89$, 0.54 y 0.04), respectivamente. En general, β disminuye a medida que aumenta el tamaño de la muestra, ya que α aumenta, y cuando $\mu_1 - \mu_2$ se aparta de 0. El *software* también calcula los tamaños de muestra necesarios para obtener un valor concreto de la potencia para un valor particular de $\mu_1 - \mu_2$.

Ejercicios | Sección 9.2 (17 a 35)

17. Determine el número de grados de libertad para la prueba t de dos muestras o el IC en cada una de las situaciones siguientes:

a. $m = 10$, $n = 10$, $s_1 = 5.0$, $s_2 = 6.0$
b. $m = 10$, $n = 15$, $s_1 = 5.0$, $s_2 = 6.0$
c. $m = 10$, $n = 15$, $s_1 = 2.0$, $s_2 = 6.0$
d. $m = 12$, $n = 24$, $s_1 = 5.0$, $s_2 = 6.0$

18. Sean μ_1 y μ_2 las densidades promedio reales para dos tipos de ladrillo. Suponiendo normalidad de las dos distribuciones de densidad, pruebe H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 \neq 0$ con los siguientes datos: $m = 6$, $\bar{x} = 22.73$, $s_1 = 0.164$, $n = 5$, $\bar{y} = 21.95$ y $s_2 = 0.240$.

19. Suponga que μ_1 y μ_2 son distancias de frenado promedio verdaderas a 50 mph para automóviles de cierto tipo equipados con dos tipos de sistemas de frenado. Utilice la prueba de dos muestras al nivel de significación 0.01 para probar H_0: $\mu_1 - \mu_2 = -10$ contra H_a: $\mu_1 - \mu_2 < -10$ para los datos siguientes: $m = 6$, $\bar{x} = 115.7$, $s_1 = 5.03$, $n = 6$, $\bar{y} = 129.3$ y $s_2 = 5.38$.

20. Utilice los datos del ejercicio 19 a fin de calcular un IC de 95% para la diferencia entre la distancia de frenado promedio real para automóviles equipados con el sistema 1 y automóviles equipados con el sistema 2. ¿El intervalo indica que se tiene información precisa acerca del valor de esta diferencia?

21. Para evaluar en forma rutinaria los síntomas de neuropatías periféricas, como en el síndrome del túnel carpiano (STC), se necesitan técnicas cuantitativas no invasivas. En el artículo "A Gap Detection Tactility Test for Sensory Deficits Associated with Carpal Tunnel Syndrome" (*Ergonomics*, 1995: 2588-2601) se describe una prueba en la que se requirió detectar un pequeño espacio en una superficie por lo demás lisa mediante el sondeo con un dedo; esto, desde el punto de vista de la funcionalidad, se asemeja a muchas actividades táctiles relacionadas con el trabajo, como detectar de arañazos o defectos superficiales. Cuando no se permitió el sondeo con los dedos, el umbral promedio muestral de detección del espacio para $m = 8$ individuos normales fue 1.71 mm, y la desviación estándar muestral fue 0.53; para $n = 10$ personas con STC, el promedio y la desviación estándar muestrales fueron 2.53 y 0.87, respectivamente. ¿Estos datos indican que el umbral promedio real de detección del espacio para personas con STC es mayor que para las personas normales? Enuncie y pruebe las hipótesis pertinentes con un nivel de significación de 0.01.

22. La prueba de corte sesgado es muy aceptada para evaluar la adhesión al concreto de materiales de reparación resinosos; se utilizan especímenes cilíndricos hechos de dos mitades idénticas unidas a 30°. En el artículo "Testing the Bond Between: Repair Materials and Concrete Substrate" (*ACI Materials J.*, 1996: 553-558) se explica que para 12 especímenes preparados por medio de cepillado con alambre, la resistencia promedio al corte (N/mm^2) y la desviación estándar muestrales fueron 19.20 y 1.58, respectivamente, mientras que para 12 especímenes cincelados a mano, los valores correspondientes fueron 23.13 y 4.01. ¿La resistencia promedio al parecer es distinta para los dos métodos de preparación de la superficie? Enuncie y pruebe las hipótesis pertinentes con un nivel de significación de 0.05. ¿Qué supone acerca de las distribuciones de resistencia al corte?

23. Se usan cada vez con más frecuencia entretelas fusibles como soporte de telas exteriores y para mejorar la forma y caída de diversas prendas de vestir. En el artículo "Compatibility of Outer and Fusible Interlining Fabrics in Tailored Garments" (*Textile Res. J.*, 1997: 137-142) aparecen los datos siguientes de extensibilidad (%) en 100 g/cm para especímenes de tela de alta calidad (H) y de mala calidad (P).

H	1.2	0.9	0.7	1.0	1.7	1.7	1.1	0.9	1.7
	1.9	1.3	2.1	1.6	1.8	1.4	1.3	1.9	1.6
	0.8	2.0	1.7	1.6	2.3	2.0			
P	1.6	1.5	1.1	2.1	1.5	1.3	1.0	2.6	

a. Trace gráficas de probabilidad normales para comprobar la posibilidad de que en ambas muestras se seleccionaron distribuciones poblacionales normales.
b. Elabore un diagrama de caja comparativo. ¿Indica que hay una diferencia entre la extensibilidad promedio real para especímenes de tela de alta calidad y la de especímenes de baja calidad?
c. La media y la desviación estándar para la muestra de alta calidad son 1.508 y 0.444, respectivamente, y para la muestra de baja calidad son 1.588 y 0.530. Use la prueba *t* de dos muestras para decidir si la extensibilidad promedio real es distinta para los dos tipos de tela.

24. La firmeza de una fruta es indicador importante de su madurez. La firmeza Magness-Taylor (N) se determinó para una muestra de 20 manzanas golden con vida de anaquel de cero días, obteniéndose como resultado una media muestral de 8.74 y una desviación estándar muestral de 0.66, y otra muestra de 20 manzanas con vida de anaquel de 20 días con media y desviación estándar muestrales de 4.96 y 0.39, respectivamente. Calcule un intervalo de confianza para la diferencia entre la firmeza promedio real de manzanas con cero días y la de manzanas con 20 días por medio de un nivel de confianza de 95%, e interprete el intervalo.

25. El dolor de espalda baja (DEB) es un problema grave de salud en muchos entornos industriales. En el artículo "Isodynamic Evaluation of Trunk Muscles and Low-Back Pain Among Workers in a Steel Factory" (*Ergonomics*, 1995: 2107-2117) se informa el siguiente resumen de datos sobre el alcance lateral de movimiento (grados) para una muestra de trabajadores sin antecedentes de DEB y otra muestra con antecedentes de este padecimiento.

Condición	Tamaño de la muestra	Media muestral	DE de la muestra
Sin DEB	28	91.5	5.5
DEB	31	88.3	7.8

Calcule un intervalo de confianza de 90% para la diferencia entre el alcance medio poblacional del movimiento lateral para las dos condiciones. ¿El intervalo indica que el movimiento promedio lateral poblacional es distinto para las dos condiciones? ¿La conclusión es distinta si se usa un nivel de confianza de 95 por ciento?

26. En el artículo "The Influence of Corrosion Inhibitor and Surface Abrasion on the Failure of Aluminum-Wired Twist-on Connections" (*IEEE Trans. on Components, Hybrids, and Manuf. Tech.*, 1984: 20-25) se dan datos acerca de mediciones de diferencias de potencial para una muestra de conectores cableados con aleación de aluminio y otra muestra cableada con aluminio EC. ¿Los siguientes resultados obtenidos con SAS indican que la caída de potencial promedio real para conexiones de aleación (tipo 1) es mayor que para las conexiones con EC tipo 2 (como se afirma en el artículo)? Lleve a cabo la prueba adecuada con un nivel de significación de 0.01. Para llegar a su conclusión, ¿qué tipo de error podría haber cometido? *Nota:* SAS proporciona el valor *p* para una prueba bilateral).

TIPO	N	Media	Desv. est.	Error est.
1	20	17.49900000	0.55012821	0.12301241
2	20	16.90000000	0.48998389	0.10956373

Varianzas	T	GL	Prob > |T|
Desigual	3.6362	37.5	0.0008
Igual	3.6362	38.0	0.0008

27. Se cree que el codo de tenista se agrava por el impacto experimentado al golpear la pelota. En el artículo "Forces on the Hand in the Tennis One-Handed Backhand" (*Intl. J. of Sport Biomechanics*, 1991: 282-292) se describe la fuerza (N) sobre la mano justo después del impacto en un golpe de revés con una sola mano para seis jugadores expertos y ocho jugadores intermedios.

Tipo de jugador	Tamaño de la muestra	Media muestral	DE de la muestra
1. Avanzado	6	40.3	11.3
2. Intermedio	8	21.4	8.3

En su análisis de los datos, los autores supusieron que en ambos casos las distribuciones de fuerza eran normales. Calcule un intervalo de confianza de 95% para la diferencia entre la fuerza promedio real para jugadores avanzados (μ_1) y la fuerza promedio real de los jugadores intermedios (μ_2). ¿Su intervalo es una prueba convincente para concluir que las dos μ son distintas? ¿Habría llegado a la misma conclusión si hubiera calculado un intervalo de confianza para $\mu_2 - \mu_1$ (es decir, al intercambiar los indicadores 1 y 2 para los dos tipos de jugadores)? Explique.

28. A medida que la población envejece, hay mayor interés acerca de las lesiones de los ancianos relacionadas con accidentes. En el artículo "Age and Gender Differences in Single-Step Recovery from a Forward Fall" (*J. of Gerontology*, 1999-M44 M50) se describe un experimento en el que se determinó el ángulo de inclinación —lo más que se puede inclinar un individuo y volver a su postura vertical en un solo paso— para una muestra

de mujeres jóvenes (21 a 29 años) y una de mujeres mayores (67 a 81 años). Las observaciones siguientes son consistentes con el resumen proporcionado en el artículo:

MJ: 29, 34, 33, 27, 28, 32, 31, 34, 32, 27

MM: 18, 15, 23, 13, 12

¿Los datos indican que el ángulo de inclinación máximo promedio real para mujeres mayores es más de 10 grados más pequeño que el correspondiente para mujeres jóvenes? Exprese y pruebe las hipótesis pertinentes al nivel de significación 0.10 al obtener un valor P.

29. En el artículo "Effect of Internal Gas Pressure on the Compression Strength of Beverage Cans and Plastic Bottles" (*J. Testing and Evaluation*, 1993: 129-131) se proporcionan los datos siguientes acerca de la fuerza de compresión (lb) para una muestra de latas de aluminio de 12 onzas llenas con bebida de fresa y otra muestra de latas que contienen bebida de cola. ¿Los datos indican que el carbonato extra de la bebida de cola da como resultado mayor resistencia promedio a la compresión? Fundamente su respuesta en un valor P. ¿Cuáles hipótesis son necesarias para efectuar el análisis?

Bebida	Tamaño muestral	Media muestral	DE de la muestra
Bebida de fresa	15	540	21
Cola	15	554	15

30. En el artículo "Flexure of Concrete Beams Reinforced with Advanced Composite Orthogrids" (*J. of Aerospace Engr.*, 1997: 7-15) se proporcionan los datos siguientes sobre la carga máxima (kN) para dos tipos de vigas.

Tipo	Tamaño de la muestra	Media muestral	DE de la muestra
Entramado de fibra de vidrio	26	33.4	2.2
Entramado de carbono comercial	26	42.8	4.3

a. Suponiendo que las distribuciones subyacentes son normales, calcule e interprete un IC de 99% para la diferencia entre la carga promedio real para las vigas de fibra de vidrio y la de las vigas de carbono.

b. ¿El límite superior del intervalo calculado en el inciso (a) proporciona un límite de confianza superior para la diferencia entre las dos μ? Si no es así, calcule ese límite. ¿Indica sin lugar a dudas que la carga promedio real para las vigas con carbono es mayor que para las de fibra de vidrio? Explique.

31. Consulte el ejercicio 33 de la sección 7.3. En el artículo citado también se proporcionan las siguientes observaciones sobre el grado de polimerización para especímenes con viscosidad multiplicada por la concentración en un intervalo mayor:

429	430	430	431	436	437
440	441	445	446	447	

a. Trace un diagrama de caja comparativo para las dos muestras y comente acerca de algunas características interesantes.

b. Calcule un intervalo de confianza de 95% para la diferencia entre el grado de polimerización promedio real para el intervalo medio y el del intervalo alto. ¿El intervalo indica que μ_1 y μ_2 de hecho podrían ser diferentes? Explique su razonamiento.

32. En el artículo citado en el ejercicio 34, de la sección 7.3, también se dan los siguientes datos acerca de los límites de esfuerzo proporcional para especímenes construidos con dos tipos de madera.

Tipo de madera	Tamaño de la muestra	Media muestral	DE de la muestra
Roble rojo	14	8.48	0.79
Abeto Douglas	10	6.65	1.28

Suponiendo que ambas muestras se seleccionaron de distribuciones normales, realice una prueba de hipótesis para decidir si el límite de esfuerzo proporcional promedio real para las uniones con roble rojo es mayor que el correspondiente al abeto Douglas en más de un MPa.

33. En la tabla siguiente se resumen los datos acerca de la ganancia de peso corporal (g) tanto para una muestra de animales a los que se les administra una dosis de un mg/pastilla de cierto esteroide blando como para otra muestra de animales de control ("The Soft Drug Approach", *CHEMTECH*, 1984: 28-38).

Tratamiento	Tamaño de la muestra	Media muestral	DE de la muestra
Esteroide	8	32.8	2.6
Control	10	40.5	2.5

¿Los datos indican que la ganancia de peso promedio real en el grupo de control es mayor que para los animales tratados con esteroides en más de 5 g? Enuncie y pruebe las hipótesis adecuadas a un nivel de significación de 0.01 con el método del valor P.

34. Considere la variable t combinada

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

que tiene una distribución t con $m + n - 2$ grados de libertad cuando ambas distribuciones poblacionales son normales con $\sigma_1 = \sigma_2$ (véase la subsección de los procedimientos de t combinada para una descripción de S_p).

a. Utilice esta variable t a fin de obtener una fórmula para el intervalo de confianza de t combinada para $\mu_1 - \mu_2$.

b. Se seleccionó una muestra de humidificadores ultrasónicos de determinada marca para la cual las observaciones de producción máxima de humedad (en onzas) en una cámara controlada fueron 14.0, 14.3, 12.2 y 15.1. Con una muestra de la segunda marca se obtuvieron valores de producción de 12.1, 13.6, 11.9 y 11.2 ("Multiple Comparisons of Means Using Simultaneous Confidence Intervals", *J. of Quality Technology*, 1989: 232-241). Use la fórmula de t combinada del inciso (a) para estimar la diferencia entre las producciones promedio reales para las dos marcas con un intervalo de confianza de 95 por ciento.

c. Estime la diferencia entre las dos μ por medio del intervalo t bimuestral analizado en esta sección y compárelo con el intervalo del inciso (b).

35. Refiérase al ejercicio 34. Describa la prueba t combinada para probar H_0: $\mu_1 - \mu_2 = \Delta_0$ cuando ambas distribuciones poblacionales son normales con $\sigma_1 = \sigma_2$. Luego utilice este procedimiento de prueba para demostrar las hipótesis sugeridas en el ejercicio 33.

9.3 Análisis de pares de datos

En las secciones 9.1 y 9.2, se consideró probar una diferencia entre dos medias μ_1 y μ_2. Esto se hizo utilizando los resultados de una muestra aleatoria $X_1, X_2, \ldots, X_m$ de la distribución con media μ_1 y una muestra independiente por completo $Y_1, \ldots, Y_n$ de la distribución con media μ_2. Es decir, ya sea que se seleccionaran m individuos de la población 1 y n individuos diferentes de la población 2, o que m individuos (u objetos experimentales) recibieran un tratamiento y a otro conjunto de n individuos se les diera el otro tratamiento. En contraste, hay varias situaciones experimentales en las que sólo hay un conjunto de n individuos u objetos experimentales, y se hacen dos observaciones en cada individuo u objeto, lo cual da como resultado pares de valores.

Ejemplo 9.8 Las trazas de metales en el agua potable afectan el sabor, y las concentraciones inusualmente altas pueden poner en riesgo la salud. En el artículo "Trace Metals of South Indian River" (*Envir. Studies*, 1982: 62-66) de describe un estudio en el que se seleccionaron seis lugares de un río (seis objetos experimentales) y se determinó la concentración de cinc (mg/L) tanto para el agua superficial como la del fondo en cada lugar. Los seis pares de observaciones se muestran en la tabla siguiente. ¿Es posible determinar a partir de los datos si la concentración promedio real del agua del fondo es mayor que la del agua superficial?

	Ubicación					
	1	2	3	4	5	6
Concentración de cinc en el agua del fondo (x)	0.430	0.266	0.567	0.531	0.707	0.716
Concentración de cinc en el agua superficial (y)	0.415	0.238	0.390	0.410	0.605	0.609
Diferencia	0.015	0.028	0.177	0.121	0.102	0.107

En la figura 9.4(a) se muestra una gráfica de estos datos. A primera vista, al parecer hay poca diferencia entre las muestras x y y. De un lugar a otro, hay mucha variabilidad en cada muestra, y parece como si las diferencias entre las muestras contribuyeran a esta variabilidad. Sin embargo, cuando se identifican por lugar las observaciones, como en la figura 9.4(b), surge un punto de vista diferente. En cada lugar, la concentración del fondo excede la concentración superficial. Esto se confirma por el hecho de que las diferencias $x - y$ (concentración del agua del fondo-concentración del agua superficial) mostrada en el renglón inferior de la tabla son positivos. Como se verá, un análisis correcto de estos datos se centra en estas diferencias.

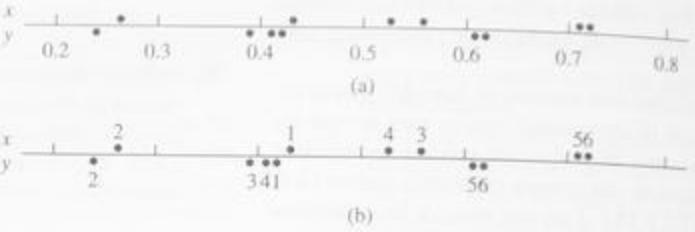


Figura 9.4 Gráfica de pares de datos del ejemplo 9.8: (a) observaciones no identificadas por lugar; (b) observaciones identificadas por lugar. ■

SUPOSICIONES

Los datos consisten en n pares seleccionados de forma independiente $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, con $E(X_i) = \mu_1$ y $E(Y_i) = \mu_2$. Sea $D_1 = X_1 - Y_1$, $D_2 = X_2 - Y_2, \ldots, D_n = X_n - Y_n$, así que las D_i son las diferencias dentro de los pares. Entonces se supone que las D_i están distribuidas normalmente con media μ_D y varianza σ_D^2 (esto por lo común es una consecuencia de que las X_i y las Y_i tienen una distribución normal).

Ejemplo

De nuevo interesa probar hipótesis acerca de la diferencia $\mu_1 - \mu_2$. El denominador de la prueba t bimuestral se obtiene aplicando primero la regla $V(X - Y) = V(X) + V(Y)$. Sin embargo, con los datos en pares, las observaciones X y Y dentro de cada par suelen no ser independientes, así que $\overline{X}$ y $\overline{Y}$ no son independientes entre sí, y la regla no es válida. Por lo tanto, se debe abandonar la prueba t bimuestral y considerar otro método de análisis.

Prueba t de pares de datos

Debido a que pares diferentes son independientes, las D_i son independientes entre sí. Si se hace que $D = X - Y$, donde X y Y son las observaciones primera y segunda, respectivamente, dentro de un par arbitrario, entonces la diferencia esperada es

$$\mu_D = E(X - Y) = E(X) - E(Y) = \mu_1 - \mu_2$$

(la regla de valores esperados utilizada aquí es válida aun cuando X y Y son dependientes.) Por consiguiente, cualquier hipótesis acerca de $\mu_1 - \mu_2$ se puede expresar como una hipótesis acerca de la diferencia media μ_D). Pero puesto que las D_i constituyen una muestra aleatoria normal (de diferencias) con media μ_D, las hipótesis acerca de μ_D se pueden probar por medio de una prueba t con una sola muestra. Es decir, *probar hipótesis acerca de $\mu_1 - \mu_2$ cuando los*

datos están en pares, formar las diferencias $D_1, D_2, \ldots, D_n$ *y realizar una prueba t con una muestra (basada en* $n - 1$ *gl) en las diferencias.*

Prueba *t* de pares de datos

Hipóteses nula: H_0: $\mu_D = \Delta_0$ (donde $D = X - Y$ es la diferencia entre las observaciones primera y segunda dentro de un par, y $\mu_D = \mu_1 - \mu_2$)

Valor del estadístico de prueba $t = \dfrac{\bar{d} - \Delta_0}{s_D/\sqrt{n}}$ (donde $\bar{d}$ y s_D son la media muestral y la desviación estándar, respectivamente de las d_i)

Hipótesis alternativa	**Región de rechazo para la prueba de nivel α**
H_a: $\mu_D > \Delta_0$	$t \geq t_{\alpha,n-1}$
H_a: $\mu_D < \Delta_0$	$t \leq -t_{\alpha,n-1}$
H_a: $\mu_D \neq \Delta_0$	ya sea $t \geq t_{\alpha/2,n-1}$ o $t \leq -t_{\alpha/2,n-1}$

Un valor P se calcula como se hizo para las pruebas t anteriores.

Ejemplo 9.9 Los trastornos musculoesqueléticos de cuello y hombro son muy comunes entre el personal de oficina que realiza tareas repetitivas por medio de unidades de presentación visual. En el artículo "Upper-Arm Elevation During Office Work" (*Ergonomics*, 1996: 1221-1230) se presenta un estudio para determinar si condiciones de trabajo más variadas tendrían un impacto en el movimiento del brazo. Los datos siguientes se obtuvieron de una muestra de $n = 16$ individuos. Cada observación es la cantidad de tiempo, expresada como una proporción del tiempo total observado, durante el cual la elevación del brazo fue por debajo de 30°. Las dos mediciones de cada individuo se obtuvieron con diferencia de 18 meses. Durante este periodo, se modificaron las condiciones de trabajo, y se permitió que los individuos participaran en una variedad más amplia de tareas. ¿Los datos indican que el tiempo promedio real durante el cual la elevación está por debajo de 30° difiere después del cambio de lo que era antes del cambio?

Individuo	1	2	3	4	5	6	7	8
Antes	81	87	86	82	90	86	96	73
Después	78	91	78	78	84	67	92	70
Diferencia	3	−4	8	4	6	19	4	3
Individuo	9	10	11	12	13	14	15	16
Antes	74	75	72	80	66	72	56	82
Después	58	62	70	58	66	60	65	73
Diferencia	16	13	2	22	0	12	−9	9

En la figura 9.5 se muestra una gráfica de probabilidad normal de las 16 diferencias; el patrón de la gráfica es bastante recto, lo cual respalda la suposición de normalidad. Un diagrama de caja de estas diferencias aparece en la figura 9.6; el diagrama de caja se localiza conside-

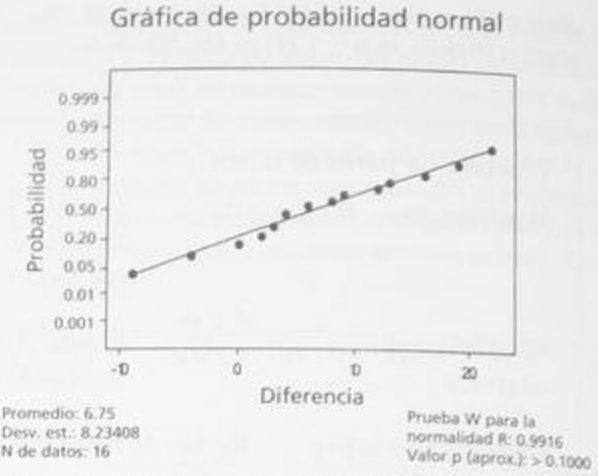


Figura 9.5 Gráfica de probabilidad normal de MINITAB de las diferencias del ejemplo 9.9

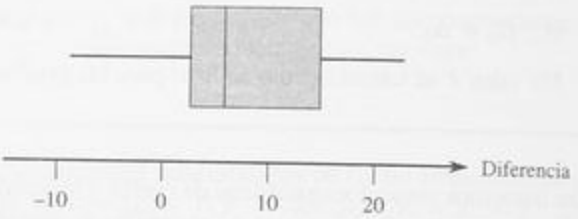


Figura 9.6 Gráfica de las diferencias del ejemplo 9.9

rablemente a la derecha del cero, lo que indica que quizá $\mu_D > 0$ (observe también que 13 de las 16 diferencias son positivas y sólo dos son negativas).

Ahora se utilizará la secuencia recomendada de pasos para probar las hipótesis apropiadas.

1. Denote con μ_D la diferencia promedio real entre el tiempo de elevación antes del cambio en las condiciones de trabajo y el tiempo después del cambio.
2. H_0: $\mu_D = 0$ (no hay diferencia entre el tiempo promedio real antes del cambio y el tiempo promedio real después del cambio)
3. H_a: $\mu_D \neq 0$
4. $t = \dfrac{\bar{d} - 0}{s_D/\sqrt{n}} = \dfrac{\bar{d}}{s_D/\sqrt{n}}$
5. $n = 16, \sum d_i = 108, \sum d_i^2 = 1746$, de lo cual $\bar{d} = 6.75$, $s_D = 8.234$, y

$$t = \frac{6.75}{8.234/\sqrt{16}} = 3.28 \approx 3.3$$

6. En la tabla A.8 del apéndice se muestra que el área a la derecha de 3.3 bajo la curva t con 15 gl es 0.002. La desigualdad en H_a significa que es apropiada una prueba bilateral, de modo que el valor aproximado de P es $2(0.002) = 0.004$ (con MINITAB se obtiene 0.0051).

7. Puesto que $0.004 < 0.01$, la hipótesis nula se rechaza a cualquier nivel de significación 0.05 o 0.01. Al parecer la diferencia promedio real entre tiempos es un poco distinta de cero; es decir, el tiempo promedio real después del cambio es diferente del que se tiene antes del cambio. ■

Cuando el número de pares es grande, resulta innecesaria la suposición de una distribución normal de diferencias. El TLC valida la prueba z resultante.

Intervalo de confianza para μ_D

De la misma forma que el IC de t para una sola media poblacional μ se basa en la variable t, $T = (\overline{X} - \mu)/(S/\sqrt{n})$, un intervalo de confianza t para μ_D $(= \mu_1 - \mu_2)$ se basa en el hecho de que

$$T = \frac{\overline{D} - \mu_D}{S_D/\sqrt{n}}$$

tiene una distribución t con $n - 1$ gl. El manejo de esta variable t, como en las determinaciones anteriores de IC, da como resultado el siguiente IC de $100(1 - \alpha)\%$:

El **IC t de pares para μ_D** es

$$\overline{d} \pm t_{\alpha/2,n-1} \cdot s_D/\sqrt{n}$$

Un límite de confianza unilateral resulta de retener el signo pertinente y reemplazar $t_{\alpha/2}$ por t_α.

Cuando n es pequeña, la validez de este intervalo requiere que la distribución de diferencias sea por lo menos aproximadamente normal. Para n grande, el TLC asegura que el intervalo z resultante es válido sin restricción alguna en la distribución de diferencias.

Ejemplo 9.10 Agregar las imágenes médicas computarizadas a una base de datos promete dar grandes recursos a los médicos. Sin embargo, hay otras formas de obtener este tipo de información, así que se debe investigar la cuestión de la eficiencia de acceso. En el artículo "The Comparative Effectivennes of Conventional and Digital Image Libraries" (*J. of Audiovisual Media in Medicine*, 2001: 8-15) se describe un experimento en el que se tomó el tiempo a 13 médicos expertos en computación cuando recuperaban una imagen de una biblioteca de diapositivas y mientras obtenían la misma imagen de una base de datos computarizada con una interfaz de red.

Individuo:	1	2	3	4	5	6	7	8	9	10	11	12	13
Diapositiva:	30	35	40	25	20	30	35	62	40	51	25	42	33
Digital:	25	16	15	15	10	20	7	16	15	13	11	19	19
Diferencia:	5	19	25	10	10	10	28	46	25	38	14	23	14

Represente con μ_D la diferencia promedio real entre el tiempo de recuperación desde una biblioteca de diapositivas (s) y el tiempo de recuperación digital. Usar el intervalo de confianza t de pares para estimar μ_D requiere que la distribución de diferencias sea por lo menos aproximadamente normal. El patrón lineal de puntos en la gráfica de probabilidad normal de MINITAB (figura 9.7) valida la suposición de normalidad. (Sólo aparecen 9 puntos, debido a que algunas diferencias son iguales.)

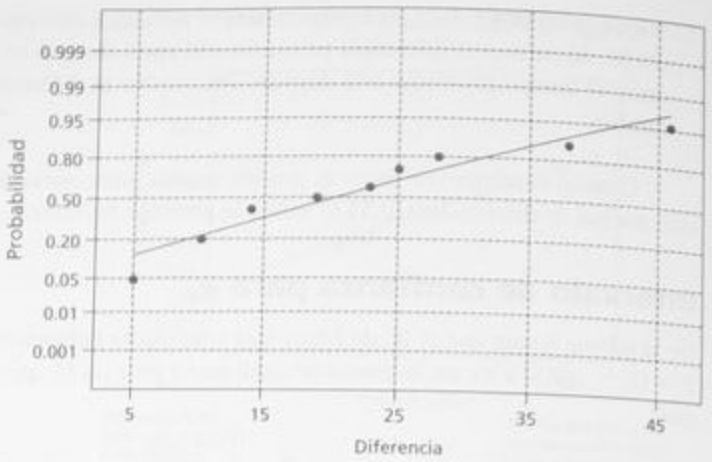


Figura 9.7 Gráfica de probabilidad normal de las diferencias del ejemplo 9.10

Las cantidades de resumen pertinentes son $\sum d_i = 267, \sum d_i^2 = 7201$, de lo cual $\bar{d} = 20.5$, $s_D = 11.96$. El valor crítico t requerido para un nivel de confianza de 95% es $t_{0.025,12} = 2.179$, y el IC de 95% es

$$\bar{d} \pm t_{\alpha/2,n-1} \cdot \frac{s_D}{\sqrt{n}} = 20.5 \pm (2.179) \cdot \frac{11.96}{\sqrt{13}} = 20.5 \pm 7.2 = (13.3, 27.7)$$

Así, se puede tener una alta confianza (al nivel de confianza 95%) de que $13.3 < \mu_D < 27.7$. Este intervalo de valores posibles es bastante amplio, una consecuencia de la desviación estándar muestral que es grande en relación con la media muestral. Se requeriría un tamaño de muestra mucho mayor que 13 para estimar con mucha más precisión. Sin embargo, observe que el 0 queda fuera del intervalo, lo que indica que $\mu_D > 0$; esto se confirma mediante una prueba de hipótesis formal. ∎

Pares de datos y procedimientos *t* con dos muestras

Considere usar la prueba t bimuestral en datos por pares. Los numeradores de los estadísticos de prueba t de pares y la prueba t con dos muestras son idénticas, puesto que $\bar{d} = \sum d/n = [\sum(x_i - y_i)]/n = (\sum x_i)/n - (\sum y_i)/n = \bar{x} - \bar{y}$. La diferencia entre los dos estadísticos se debe por completo a los denominadores. Cada estadístico de prueba se obtiene al estandarizar $\bar{X} - \bar{Y}$ $(= \bar{D})$, pero en presencia de dependencia es incorrecta la estandarización de t bimuestral. Para ver esto, recuerde de la sección 5.5 que

$$V(X \pm Y) = V(X) + V(Y) \pm 2\,\mathrm{Cov}(X, Y)$$

Puesto que la correlación entre X y Y es

$$\rho = \mathrm{Corr}(X, Y) = \mathrm{Cov}(X, Y)/[\sqrt{V(X)} \cdot \sqrt{V(Y)}]$$

se deduce que

$$V(X - Y) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$$

Al aplicar esto a $\overline{X} - \overline{Y}$, se obtiene

$$V(\overline{X} - \overline{Y}) = V(\overline{D}) = V\left(\frac{1}{n}\sum D_i\right) = \frac{V(D_i)}{n} = \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{n}$$

La prueba t bimuestral se basa en la suposición de independencia, en cuyo caso $\rho = 0$. Pero en muchos experimentos con pares de datos, habrá una fuerte dependencia *positiva* entre X y Y (X grande relacionada con Y grande), así que ρ será positiva y la varianza de $\overline{X} - \overline{Y}$ será menor que $\sigma_1^2/n + \sigma_2^2/n$. Por lo tanto, *si hay una dependencia positiva entre pares, el denominador para el estadístico de pares debe ser más pequeño que para t de la prueba con muestras independientes.* Con frecuencia la t bimuestral estará mucho más cercana a cero que la t para pares, entendiendo de forma considerable la significación de los datos.

De manera similar, cuando se tienen pares de datos, el IC de la t por lo común es más estrecho que el IC para la t bimuestral (incorrecto). Esto se debe a que por lo regular hay menos variabilidad en las diferencias que en los valores de x y y.

Experimentos con pares de datos contra experimentos sin pares

En los ejemplos, los pares de datos resultaron de dos observaciones en el mismo individuo (ejemplo 9.9) u objeto experimental (ubicación en el ejemplo 9.8). Incluso cuando esto no se puede hacer, las parejas de datos con dependencia dentro de los pares se obtienen al hacer corresponder individuos u objetos en una o más características que se considera influyen en las respuestas. Por ejemplo, en un experimento médico en el que se compara la eficacia de dos fármacos para disminuir la tensión arterial, el presupuesto del experimentador podría permitir el tratamiento de 20 pacientes. Si se eligen al azar 10 pacientes para el tratamiento con el primer fármaco y otros 10 elegidos de forma independiente para el tratamiento con el segundo fármaco, resulta un experimento con muestras independientes.

Sin embargo, el experimentador, sabiendo que la tensión arterial se ve afectada por la edad y el peso, podría decidir formar parejas de pacientes para que dentro de cada uno de los 10 pares resultantes, el peso y la edad fueran más o menos iguales (aunque podría haber diferencias notables entre pares). Entonces se suministraría cada fármaco a un paciente distinto dentro de cada par a un total de 10 observaciones en cada fármaco.

Sin esta comparación (o "agrupamiento en bloques"), parecería que un fármaco supera al otro sólo porque los pacientes de una muestra fueron de poco peso o más jóvenes y, por lo tanto, más susceptibles a una disminución en la tensión arterial que los pacientes con mayor peso y más edad en la segunda muestra. Sin embargo, hay un precio que pagar por formar parejas, un número más pequeño de grados de libertad para el análisis por pares, así que se debe preguntar cuándo se debe dar preferencia a un tipo de experimento y no a otro.

No existe una respuesta directa y precisa a esta pregunta, pero hay algunas guías útiles. Si se tiene una opción entre dos pruebas t que son válidas (y se llevan a cabo al mismo nivel de significación α), se debe preferir la prueba que tiene el número más grande de grados de libertad. La razón para esto es que un número más grande de grados de libertad significa una β más pequeña para algún valor fijo alternativo del parámetro o parámetros. Es decir, para una probabilidad fija de error de tipo I, se reduce la probabilidad de un error de tipo II al aumentar los grados de libertad.

Sin embargo, si las unidades experimentales son bastante heterogéneas en sus respuestas, será difícil detectar pequeñas diferencias, aunque importantes, entre dos tratamientos. Esto es en esencia lo que sucedió en el conjunto de datos del ejemplo 9.8; para ambos "tratamientos" (agua del fondo y agua superficial), hay una gran variabilidad entre lugares, que tiende a enmascarar diferencias en los tratamientos dentro de los lugares. Si hay una alta correlación positiva dentro de las unidades experimentales o individuos, la varianza de $\overline{D} = \overline{X} - \overline{Y}$ será mucho más pequeña que la varianza sin pares. Como resultado de esta varianza reducida, será más fácil detectar una diferencia con muestras de pares que con muestras independientes. Las ventajas y desventajas de formar parejas se pueden resumir ahora como sigue.

1. Si hay gran heterogeneidad entre unidades experimentales y una gran correlación dentro de unidades experimentales (ρ positiva grande), entonces la pérdida en grados de libertad será compensada por la mayor precisión relacionada con la formación de parejas, de modo que es preferible el experimento con pares a un experimento con muestras independientes.
2. Si las unidades experimentales son relativamente homogéneas y la correlación dentro de los pares no es grande, la ganancia de precisión debida a la formación de parejas se paga con una disminución de grados de libertad, de modo que se debe usar un experimento con muestras independientes.

Por supuesto que no se conocerán con precisión los valores de σ_1^2, σ_2^2 y ρ. Entonces será necesario que un investigador tome una decisión intuitiva en cuanto a si obtiene la situación 1 o 2. En general, si el número de observaciones que se puede obtener es grande, entonces no será grave una pérdida de grados de libertad (p. ej., de 40 a 20); pero si el número es pequeño, entonces la pérdida (p. ej., 16 a 8) debida a la formación de parejas podría ser grave si no se compensa con una mayor precisión. Consideraciones similares se aplican al elegir entre dos tipos de experimentos para estimar $\mu_1 - \mu_2$ con un intervalo de confianza.

Ejercicios | Sección 9.3 (36 a 46)

36. Considere los siguientes datos acerca de la carga de rotura (kg/25 mm de ancho) para varias telas, tanto desgastadas como no desgastadas ("The Effect of Wet Abrasive Wear on the Tensile Properties of Cotton and Polyester-Cotton Fabrics", *J. Testing and Evaluation*, 1993: 84-93). Utilice la prueba t para pares de datos, como lo hicieron los autores del artículo citado, para probar H_0: $\mu_D = 0$ contra H_a: $\mu_D > 0$ al nivel de significación 0.01.

	Tela							
	1	2	3	4	5	6	7	8
U	36.4	55.0	51.5	38.7	43.2	48.8	25.6	49.8
A	28.5	20.0	46.0	34.5	36.5	52.5	26.5	46.5

37. Se ha identificado al cromo hexavalente como cancerígeno por inhalación y una toxina aérea de consideración en varios sitios. En el artículo "Airborne Hexavalent Chromium in Southwestern Ontario" (*J. of Air and Waste Mgmnt. Assoc.*, 1997: 905-910) aparecen los datos siguientes acerca de la concentración en interiores y en exteriores (nanogramos/m^3) para una muestra de casas seleccionadas en cierta región.

Casa	1	2	3	4	5	6	7	8	9
Interior	0.07	0.08	0.09	0.12	0.12	0.12	0.13	0.14	0.15
Exterior	0.29	0.68	0.47	0.54	0.97	0.35	0.49	0.84	0.86

Casa	10	11	12	13	14	15	16	17
Interior	0.15	0.17	0.17	0.18	0.18	0.18	0.18	0.19
Exterior	0.28	0.32	0.32	1.55	0.66	0.29	0.21	1.02

Casa	18	19	20	21	22	23	24	25
Interior	0.20	0.22	0.22	0.23	0.23	0.25	0.26	0.28
Exterior	1.59	0.90	0.52	0.12	0.54	0.88	0.49	1.24
Casa	26	27	28	29	30	31	32	33
Interior	0.28	0.29	0.34	0.39	0.40	0.45	0.54	0.62
Exterior	0.48	0.27	0.37	1.26	0.70	0.76	0.99	0.36

a. Calcule un intervalo de confianza para la diferencia promedio poblacional entre concentraciones en interiores y exteriores con un nivel de confianza de 95%, e interprete el intervalo resultante.

b. Si de la población se seleccionara al azar la casa 34, ¿entre qué valores predeciría que estaría situada la diferencia de concentraciones?

38. De una mezcla de concreto de resistencia normal y de una mezcla de alta resistencia se obtuvieron especímenes de concreto con relaciones variables de altura a diámetro, cortados de diversas posiciones del cilindro original. Se determinó el esfuerzo máximo (MPa) para cada mezcla, y se obtuvieron los datos siguientes ("Effect of Length on Compressive Strain Softening of Concrete", *J. of Engr. Mechanics*, 1997: 25-35):

Condición de la prueba	1	2	3	4	5
Normal	42.8	55.6	49.0	48.7	44.1
Alta	90.9	93.1	86.3	90.3	88.5
Condición de la prueba	6	7	8	9	10
Normal	55.4	50.1	45.7	51.4	43.1
Alta	88.1	93.2	90.8	90.1	92.6
Condición de prueba	11	12	13	14	15
Normal	46.8	46.7	47.7	45.8	45.4
Alta	88.2	88.6	91.0	90.0	90.1

a. Trace un diagrama de caja comparativo de esfuerzos máximos para los dos tipos de concreto y comente algunas características interesantes.

b. Estime la diferencia entre esfuerzos máximos promedio reales para los dos tipos de concreto, de manera que contenga información acerca de la precisión y la confiabilidad. Asegúrese de revisar la factibilidad de las hipótesis necesarias en su análisis. ¿Es posible que las resistencias máximas promedio reales para los dos tipos de concreto sean idénticas? ¿Por qué sí o por qué no?

39. Con frecuencia, científicos e ingenieros desean comparar dos técnicas para medir o determinar el valor de una variable. En este tipo de situaciones, el interés se centra en probar si la diferencia promedio en las mediciones es cero. En el artículo "Evaluation of the Deuterium Dilution Technique Against the Test Weighing Procedure for the Determination of Breast Milk Intake" (*Amer: J. Clinical Nutr.*, 1983: 996-1003) aparecen los datos siguientes sobre la cantidad de leche que ingiere cada uno de 14 infantes seleccionados al azar.

	Infante			
	1	**2**	**3**	**4**
Método isotópico	1509	1418	1561	1556
Método de ponderación de prueba	1498	1254	1336	1565
Diferencia	11	164	225	−9

	Infante			
	5	**6**	**7**	**8**
Método isotópico	2169	1760	1098	1198
Método de ponderación de prueba	2000	1318	1410	1129
Diferencia	169	442	−312	69

	Infante		
	9	**10**	**11**
Método isotópico	1479	1281	1414
Método de ponderación de prueba	1342	1124	1468
Diferencia	137	157	−54

	Infante		
	12	**13**	**14**
Método isotópico	1954	2174	2058
Método de ponderación de prueba	1604	1722	1518
Diferencia	350	452	540

a. ¿Es posible que la distribución poblacional de diferencias sea normal?

b. ¿Parece que la diferencia promedio real entre los valores de ingestión medidos por los dos métodos es distinta de cero? Determine el valor P de la prueba y utilícelo para llegar a una conclusión al nivel de significación 0.05.

40. Se utilizaron dos tipos de señuelos para peces, uno hecho de tubos de arcilla vitrificada y otro de bloques de cemento y maleza, durante 16 periodos diferentes que abarcaron 4 años en el lago Tohopekaliga, Florida ("Two Types of Fish Attractors Compared in Lake Tohopekaliga, Florida", *Trans. Amer. Fisheries Soc.*, 1978: 689-695). Las siguientes observaciones son de captura de peces por día de pesca:

	Periodo							
	1	2	3	4	5	6	7	8
Tubo	6.64	7.89	1.83	0.42	0.85	0.29	0.57	0.63
Maleza	9.73	8.21	2.17	0.75	1.61	0.75	0.83	0.56

	Periodo							
	9	10	11	12	13	14	15	16
Tubo	0.32	0.37	0.00	0.11	4.86	1.80	0.23	0.58
Maleza	0.76	0.32	0.48	0.52	5.38	2.33	0.91	0.79

¿Al parecer un señuelo es más atractivo en promedio que el otro?

a. Utilice la prueba t para pares de datos con $\alpha = 0.01$ para probar H_0: $\mu_D = 0$ contra H_a: $\mu_D \neq 0$.

b. ¿Qué pasa si se utiliza la prueba t bimuestral ($s_1 = 2.48$ y $s_2 = 2.91$)?

41. En un experimento diseñado para estudiar los efectos del nivel de iluminación en el rendimiento para realizar tareas ("Performance of Complex Tasks Under Different Levels of Illumination", *J. Illuminating Eng.*, 1976: 235-242), se pidió a varias personas que insertaran con rapidez y en forma sucesiva una sonda de punta fina en el ojo de cada una de diez agujas, tanto para un nivel bajo de luz con fondo negro como para un nivel mayor de luz con fondo blanco. Cada uno de los datos es el tiempo en segundos necesario para terminar la tarea.

	Individuo				
	1	2	3	4	5
Fondo blanco	25.85	28.84	32.05	25.74	20.89
Fondo negro	18.23	20.84	22.96	19.68	19.50

	Individuo			
	6	7	8	9
Fondo blanco	41.05	25.01	24.96	27.47
Fondo negro	24.98	16.61	16.07	24.59

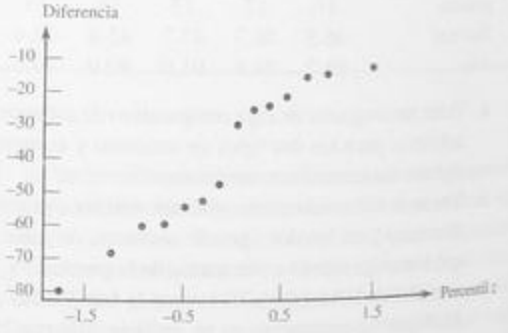
¿Los datos indican que el nivel de iluminación más alto produce una disminución de más de 5 s en el tiempo promedio real para completar la tarea? Pruebe las hipótesis apropiadas con el método del valor P.

42. Se ha estimado que entre 1945 y 1971, 2 millones de niños nacieron de madres tratadas con dietilestilbestrol (DES), un estrógeno no esteroideo recomendado para mantenimiento del embarazo. La FDA prohibió este medicamento en 1971 debido a que en las investigaciones se encontró una relación con la incidencia de cáncer cervical. En el artículo "Effects of Prenatal Exposure to Diethylstilbestrol (DES) on Hemispheric Laterality and Spatial Ability in Human Males" (*Hormones and Behavior*, 1992: 62-75) se describe un estudio en el cual 10 varones expuestos al DES y sus hermanos no expuestos se sometieron a varias pruebas. Los siguientes son los datos de resumen de los resultados de una prueba de habilidad espacial: $\bar{x} = 12.6$ (expuesto), $\bar{y} = 13.7$ y error estándar de la diferencia media = 0.5. Pruebe al nivel 0.05 para ver si la exposición se relaciona con menor capacidad espacial al obtener el valor P.

43. La enfermedad de Cushing se caracteriza por debilidad muscular debido a la disfunción suprarrenal o pituitaria. A fin de proporcionar un tratamiento eficaz, es importante detectar la enfermedad en la infancia tan pronto como sea posible. La edad de inicio de los síntomas y la edad al momento del diagnóstico para 15 niños enfermos se publicaron en el artículo "Treatment of Cushing's Disease in Childhood and Adolescence by Transphenoidal Microadenomectomy" (*New Engl. J. of Med.*, 1984: 889). Los valores de diferencias entre la edad de inicio de los síntomas y la edad al momento del diagnóstico son los siguientes:

−24 −12 −55 −15 −30 −60 −14 −21
−48 −12 −25 −53 −61 −69 −80

a. ¿El siguiente diagrama de probabilidad normal hace dudar acerca de la normalidad aproximada de la distribución de las diferencias?



b. Calcule un límite de confianza de 95% inferior para la diferencia promedio de la población e interprete el resultado.

c. Suponga que se calcularon las diferencias (edad al momento del diagnóstico) − (edad de inicio). ¿Cuál sería límite de confianza de 95% superior para la diferencia promedio poblacional correspondiente?

44. En el ejemplo 7.11 se proporcionaron datos acerca del módulo de elasticidad obtenido un minuto después de cargar en cierta configuración. En el artículo citado también se dan los valores del módulo de elasticidad obtenido cuatro semanas después de colocar cargas en los especímenes. A continuación se presentan los datos.

Observación	1 min	4 semanas	Diferencia
1	10 490	9110	1380
2	16 620	13 250	3370
3	17 300	14 720	2580
4	15 480	12 740	2740
5	12 970	10 120	2850
6	17 260	14 570	2690
7	13 400	11 220	2180
8	13 900	11 100	2800
9	13 630	11 420	2210
10	13 260	10 910	2350
11	14 370	12 110	2260
12	11 700	8620	3080
13	15 470	12 590	2880
14	17 840	15 090	2750
15	14 070	10 550	3520
16	14 760	12 230	2530

(continúa)

Calcule e interprete un límite de confianza superior para la diferencia promedio real entre el módulo de un minuto y el de cuatro semanas; compruebe primero la posibilidad de algunas hipótesis necesarias.

45. Refiérase al ejercicio 41. Calcule una estimación del intervalo para la diferencia entre el tiempo promedio para realizar la tarea con el nivel alto de iluminación y el tiempo promedio real con el nivel bajo.

46. Construya un conjunto de pares de datos para el cual $t = \infty$, de modo que los datos sean muy significativos cuando se utilice el análisis correcto; no obstante, el valor de t es bastante cercano a cero para la prueba t de dos muestras y, por lo tanto, el análisis incorrecto produce un resultado insignificante.

9.4 Inferencias en relación con una diferencia entre proporciones poblacionales

Una vez presentados los métodos para comparar las medias de dos poblaciones distintas, se vuelve a la comparación de dos proporciones poblacionales. La notación para este problema es una extensión de la notación utilizada en el problema correspondiente de una población. Se representan con p_1 y p_2 las proporciones de individuos de las poblaciones 1 y 2, respectivamente, que poseen una característica particular. Otra posibilidad es que si se utiliza la identificación S para un individuo que posee la característica de interés (favorece una proposición particular, ha leído por lo menos un libro en el último mes, etc.), entonces p_1 y p_2 representan las probabilidades de ver la identificación S en un individuo elegido al azar de las poblaciones 1 y 2, respectivamente.

Se supondrá que se tiene una muestra de m individuos de la primera población y n de la segunda. Las variables X y Y representan el número de individuos en cada muestra que poseen la característica que define a p_1 y p_2. Siempre y cuando los tamaños de las poblaciones sean mucho mayores que los tamaños de muestra, la distribución de X se puede tomar como binomial con parámetros m y p_1, y, de forma similar, Y se toma como una variable binomial con parámetros n y p_2. Además, se supone que las muestras son independientes entre sí, de modo que X y Y son variables aleatorias independientes.

El estimador obvio para $p_1 - p_2$, la diferencia de las proporciones poblacionales, es la diferencia correspondiente en las proporciones muestrales $X/m - Y/n$. Con $\hat{p}_1 = X/m$ y $\hat{p}_2 = Y/n$, el estimador de $p_1 - p_2$ se puede expresar como $\hat{p}_1 - \hat{p}_2$.

PROPOSICIÓN

Sea $X \sim \text{Bin}(m, p_1)$ y $Y \sim \text{Bin}(n, p_2)$ con X y Y variables independientes. Entonces

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$$

por lo tanto $\hat{p}_1 - \hat{p}_2$ es un estimador insesgado $p_1 - p_2$, y

$$V(\hat{p}_1 - \hat{p}_2) = \frac{p_1 q_1}{m} + \frac{p_2 q_2}{n} \qquad (\text{donde } q_i = 1 - p_i) \tag{9.3}$$

Demostración Puesto que $E(X) = mp_1$ y $E(Y) = np_2$,

$$E\left(\frac{X}{m} - \frac{Y}{n}\right) = \frac{1}{m}E(X) - \frac{1}{n}E(Y) = \frac{1}{m}mp_1 - \frac{1}{n}np_2 = p_1 - p_2$$

Como $V(X) = mp_1q_1$, $V(Y) = np_2q_2$, y X y Y son independientes,

$$V\left(\frac{X}{m} - \frac{Y}{n}\right) = V\left(\frac{X}{m}\right) + V\left(\frac{Y}{n}\right) = \frac{1}{m^2}V(X) + \frac{1}{n^2}V(Y) = \frac{p_1q_1}{m} + \frac{p_2q_2}{n}$$

■

Primero se atenderán situaciones en las que m y n son grandes. Entonces, como $\hat{p}_1$ y $\hat{p}_2$ tienen cada una distribuciones aproximadamente normales, el estimador $\hat{p}_1 - \hat{p}_2$ también tiene una distribución normal aproximada. Al estandarizar $\hat{p}_1 - \hat{p}_2$ se obtiene una variable Z cuya distribución es aproximadamente normal estándar.

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\dfrac{p_1q_1}{m} + \dfrac{p_2q_2}{n}}}$$

Procedimiento de prueba con muestras grandes

De manera análoga a las hipótesis para $\mu_1 - \mu_2$, la hipótesis nula más general que podría considerar un investigador sería de la forma H_0: $p_1 - p_2 = \Delta_0$, donde Δ_0 es de nuevo un número especificado. Aunque para las medias poblacionales el caso $\Delta_0 \neq 0$ no presenta dificultades, para proporciones poblacionales los casos $\Delta_0 = 0$ y $\Delta_0 \neq 0$ se deben considerar por separado. Puesto que la vasta mayoría de los problemas reales de este tipo se relacionan con $\Delta_0 = 0$ (es decir, la hipótesis nula $p_1 = p_2$), se atenderá este caso. Cuando H_0: $p_1 - p_2 = 0$ es verdadera, represente con p el valor común de p_1 y p_2 (y de manera similar para q). Entonces la variable estandarizada

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{pq\left(\dfrac{1}{m} + \dfrac{1}{n}\right)}} \tag{9.4}$$

tiene aproximadamente una distribución normal estándar cuando H_0 es verdadera. Sin embargo, esta Z no puede servir como un estadístico de prueba, porque se desconoce el valor de p; H_0 sólo asegura que hay un valor común de p, pero no dice cuál es el valor. Para obtener un estadístico de prueba que tiene una distribución normal estándar aproximada cuando H_0 es verdadera (de modo que el uso de un valor crítico z apropiado específica una prueba de nivel α), se debe estimar p a partir de los datos de la muestra.

Suponiendo entonces que $p_1 = p_2 = p$, en lugar de muestras separadas de tamaño m y n de dos poblaciones diferentes (dos distribuciones binomiales distintas), en realidad se tiene una sola muestra de tamaño $m + n$ de una población con proporción p. Puesto que el número to-

tal de individuos en esta muestra combinada que tiene la característica de interés es $X + Y$, el estimador de p es

$$\hat{p} = \frac{X + Y}{m + n} = \frac{m}{m + n}\hat{p}_1 + \frac{n}{m + n}\hat{p}_2 \tag{9.5}$$

La segunda expresión para $\hat{p}$ muestra que en realidad es un promedio ponderado de estimadores $\hat{p}_1$ y $\hat{p}_2$ obtenidos de las dos muestras. Si se toma (9.5) (con $\hat{q} = 1 - \hat{p}$) y se sustituye en la ecuación (9.4), el estadístico resultante tiene una distribución normal estándar aproximada cuando H_0 es verdadera.

Hipótesis nula: H_0: $p_1 - p_2 = 0$

Valor estadístico de prueba (muestras grandes): $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\dfrac{1}{m} + \dfrac{1}{n}\right)}}$

Hipótesis alternativa	**Región de rechazo para la prueba de nivel α aproximada**
H_a: $p_1 - p_2 > 0$	$z \geq z_\alpha$
H_a: $p_1 - p_2 < 0$	$z \leq -z_\alpha$
H_a: $p_1 - p_2 \neq 0$	ya sea $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$

Un valor P se calcula de la misma forma que para las pruebas z previas.

Ejemplo 9.11 Algunas personas acusadas en procesos criminales se declaran culpables y son sentenciadas sin un juicio, mientras que otras que alegan inocencia son encontradas culpables después y, entonces, son sentenciadas. En años recientes, estudiosos de las leyes han especulado en cuanto a si las sentencias de quienes se declaran culpables difieren en severidad de las sentencias para quienes alegan ser inocentes y después son juzgados culpables. Considere los datos siguientes acerca de personas demandadas del condado de San Francisco acusadas de robo, quienes ya han estado en prisión ("Does It Pay to Plead Guilty? Differential Sentencing and the Functioning of Criminal Courts", *Law and Society Rev.*, 1981-1982: 45-69). ¿Estos datos indican que la proporción de acusados en estas circunstancias que se declaran culpables y son enviados a prisión difiere de la proporción de quienes son encarcelados después de alegar inocencia y ser encontrados culpables?

	Alegato	
	Culpable	**Inocente**
Cantidad de personas juzgadas culpables	$m = 191$	$n = 64$
Número de personas sentenciadas	$x = 101$	$y = 56$
Proporción muestral	$\hat{p}_1 = 0.529$	$\hat{p}_2 = 0.875$

Represente con p_1 y p_2 las dos proporciones poblacionales. Las hipótesis de interés son H_0: $p_1 - p_2 = 0$ contra H_a: $p_1 - p_2 \neq 0$. Al nivel 0.01, H_0 se debe rechazar si $z \geq z_{0.005} = 2.58$ o si $z \leq -2.58$. La estimación combinada de la proporción de éxitos comunes es $\hat{p} = (101 + 56)/(191 + 64) = 0.616$. Entonces, el valor del estadístico de prueba es

$$z = \frac{0.529 - 0.875}{\sqrt{(0.616)(0.384)\left(\frac{1}{191} + \frac{1}{64}\right)}} = \frac{-0.346}{0.070} = -4.94$$

Puesto que $-4.94 \leq -2.58$, se debe rechazar H_0.

El valor P para la prueba z bilateral es

$$\text{El valor } P = 2[1 - \Phi(|z|)] = 2[1 - \Phi(4.94)] < 2[1 - \Phi(3.49)] = 0.0004$$

En una tabla normal estándar más extensa se encuentra el valor de $P \approx 0.0000006$. Este valor P es tan pequeño que a cualquier nivel razonable α, se debe rechazar H_0. A partir de los datos se observa que $p_1 \neq p_2$ y, en particular, que declararse culpable al inicio podría ser una buena estrategia para evitar ir a prisión.

En el artículo citado también se publican datos relacionados con acusados en otros condados. Los autores desglosaron los datos por tipo de delito (asalto o robo) y por la naturaleza del registro previo (ninguno, alguno pero sin encarcelamiento y prisión). En cada caso, la conclusión fue la misma: entre los acusados juzgados culpables, los que alegaron ser culpables tuvieron menos probabilidades de recibir sentencias de prisión. ■

Probabilidades de error de tipo II y tamaños de muestra

Aquí la determinación de β es un poco más difícil que para otras pruebas con muestras grandes. La razón es que el denominador de Z es una estimación de la desviación estándar de $\hat{p}_1 - \hat{p}_2$, suponiendo que $p_1 = p_2 = p$. Cuando H_0 es falsa, $\hat{p}_1 - \hat{p}_2$ se debe estandarizar de nuevo por medio de

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1 q_1}{m} + \frac{p_2 q_2}{n}} \tag{9.6}$$

La forma de σ implica que β no es sólo función de $p_1 - p_2$, por consiguiente se denota con $\beta(p_1, p_2)$.

Ejemplo 9

Hipótesis alternativa	$\beta(p_1, p_2)$
H_a: $p_1 - p_2 > 0$	$\Phi\left[\dfrac{z_\alpha\sqrt{\overline{pq}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$
H_a: $p_1 - p_2 < 0$	$1 - \Phi\left[\dfrac{-z_\alpha\sqrt{\overline{pq}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$
H_a: $p_1 - p_2 \neq 0$	$\Phi\left[\dfrac{z_{\alpha/2}\sqrt{\overline{pq}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$
	$-\Phi\left[\dfrac{-z_{\alpha/2}\sqrt{\overline{pq}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$

donde $\overline{p} = (mp_1 + np_2)/(m + n)$, $\overline{q} = (mq_1 + nq_2)/(m + n)$, y σ se determina mediante la ecuación (9.6).

Demostración Para la prueba de extremo superior (H_a: $p_1 - p_2 > 0$),

$$\beta(p_1, p_2) = P\left[\hat{p}_1 - \hat{p}_2 < z_\alpha \sqrt{\hat{p}\hat{q}\left(\frac{1}{m} + \frac{1}{n}\right)}\right]$$

$$= P\left[\frac{(\hat{p}_1 - \hat{p}_2 - (p_1 - p_2))}{\sigma} < \frac{z_\alpha \sqrt{\hat{p}\hat{q}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$$

Cuando m y n son grandes,

$$\hat{p} = (m\hat{p}_1 + n\hat{p}_2)/(m + n) \approx (mp_1 + np_2)/(m + n) = \overline{p}$$

y $\hat{q} \approx \overline{q}$, con lo cual se obtiene la expresión previa (aproximada) para $\beta(p_1, p_2)$. ■

Por otro lado, para p_1 especificada, p_2 con $p_1 - p_2 = d$, se pueden determinar los tamaños de muestra necesarios para lograr que $\beta(p_1, p_2) = \beta$. Por ejemplo, para una prueba de extremo superior, se iguala $-z_\beta$ con el argumento de $\Phi(\cdot)$ (es decir, lo que está dentro del paréntesis) en el cuadro anterior. Si $m = n$, hay una expresión simple para el valor común.

Para el caso $m = n$, la prueba de nivel α tiene una probabilidad β de error de tipo II a los valores alternativos p_1, p_2 con $p_1 - p_2 = d$ cuando

$$n = \frac{\left[z_\alpha \sqrt{(p_1 + p_2)(q_1 + q_2)/2} + z_\beta \sqrt{p_1 q_1 + p_2 q_2}\right]^2}{d^2} \tag{9.7}$$

para una prueba de extremo superior o inferior, con $\alpha/2$ en lugar de α para una prueba bilateral.

Ejemplo 9.12 Una de las aplicaciones realmente impresionantes de la estadística ocurrió en relación con diseño del experimento para la vacuna de la polio de Salk en 1954 y el análisis de los datos resultantes. Parte del experimento se centró en la eficacia de la vacuna para combatir la polio paralítica. Debido a que se consideró que sin un grupo control de niños no habría bases sólidas para la evaluación de la vacuna, se decidió administrar la vacuna a un grupo y una inyección placebo (indistinguible visualmente de la vacuna, pero sin ningún efecto conocido) al grupo de control. Por razones éticas y también porque se consideró que el conocimiento de la administración de la vacuna podría tener un efecto en el tratamiento y el diagnóstico, el experimento se condujo de una manera **doble ciego.** Es decir, ni los individuos que recibieron las inyecciones ni quienes las administraron, sabían en realidad quién estaba recibiendo la vacuna y quién estaba recibiendo el placebo (las muestras se codificaron con números) recuerde, en ese punto no estaba del todo claro si era benéfica la vacuna.

Sean p_1 y p_2 las probabilidades de que un niño contraiga la polio paralítica para las condiciones de control y tratamiento, respectivamente. El objetivo fue probar H_0: $p_1 - p_2 = 0$ contra H_a: $p_1 - p_2 > 0$ (la hipótesis alternativa expresa que un niño vacunado tiene menos probabilidades de contraer la polio que un niño no vacunado). Suponiendo que el valor verdadero de p_1 es 0.0003 (una tasa de incidencia de 30 por 100 000), la vacuna sería una mejora importante si la tasa de incidencia fuese la mitad; es decir, $p_2 = 0.00015$. Usando una prueba al nivel $\alpha = 0.05$, entonces sería razonable pedir tamaños de muestra para los cuales $\beta = 0.1$ cuando $p_1 = 0.0003$ y $p_2 = 0.00015$. Suponiendo que las muestras tienen el mismo tamaño, la n requerida se obtiene de (9.7) como

$$n = \frac{[1.645\sqrt{(0.5)(0.00045)(1.99955)} + 1.28\sqrt{(0.00015)(0.99985) + (0.0003)(0.9997)}]^2}{(0.0003 - 0.00015)^2}$$

$$= [(0.0349 + 0.0271)/0.00015]^2 \approx 171\,000$$

Se deducen los datos reales para este experimento. Se usaron tamaños de muestra de aproximadamente 200 000. El lector puede comprobar sin dificultad que $z = 6.43$, un valor muy significativo. ¡Se juzgó que la vacuna fue un rotundo éxito!

Placebo: $m = 201\,229$, $x =$ número de casos de polio paralítica $= 110$

Vacuna: $n = 200\,745$, $y = 33$ ■

Intervalo de confianza con muestras grandes para $p_1 - p_2$

Del mismo modo que con las medias, muchos problemas con dos muestras se relacionan con el objetivo de la comparación por medio de prueba de hipótesis, pero en ocasiones es apropiada una estimación de intervalo para $p_1 - p_2$. Tanto $\hat{p}_1 = X/m$ como $\hat{p}_2 = Y/n$ tienen distribuciones normales aproximadas cuando m y n son grandes. Si se identifica a θ con $p_1 - p_2$, entonces $\hat{\theta} = \hat{p}_1 - \hat{p}_2$ satisface las condiciones necesarias para obtener un IC de muestras grandes. En particular, la desviación estándar estimada de $\hat{\theta}$ es $\sqrt{(\hat{p}_1\hat{q}_1/m) + (\hat{p}_2\hat{q}_2/n)}$. El intervalo de $100(1 - \alpha)\%$, $\hat{\theta} \pm z_{\alpha/2} \cdot \hat{\sigma}_{\hat{\theta}}$ se convierte en

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{m} + \frac{\hat{p}_2\hat{q}_2}{n}}$$

Observe que la desviación estándar estimada de $\hat{p}_1 - \hat{p}_2$ (la expresión de la raíz cuadrada) es diferente aquí de lo que fue para la prueba de hipótesis cuando $\Delta_0 = 0$.

En investigaciones recientes se ha demostrado que el nivel de confianza real para el IC tradicional recién mencionado a veces desvía sustancialmente del nivel nominal (el nivel que se considere se logra cuando se utiliza un valor crítico z particular (p.ej., 95% cuando $z_{\alpha/2} = 1.96$). La mejoría sugerida es agregar un éxito y un fracaso a cada una de las dos muestras y luego las $\hat{p}$ y las $\hat{q}$ en la fórmula anterior por $\tilde{p}$ y $\tilde{q}$ donde $\tilde{p}_1 = (x + 1)/(m + 2)$, etc. Este intervalo también se puede usar cuando los tamaños de muestra son bastante pequeños.

Ejemplo 9.13 Los autores del artículo "Adjuvant Radiotherapy and Chemotherapy in Node-Positive Premenopausal Women with Breast Cancer" (*New Engl. J. of Med.*, 1997: 956-962) publicaron los resultados de un experimento diseñado para comparar el tratamiento de pacientes enfermos de cáncer sólo con quimioterapia y el tratamiento con una combinación de quimioterapia y radiación. De los 154 individuos que recibieron sólo el tratamiento de quimioterapia, 76 sobrevivieron por lo menos 15 años, mientras que 98 de los 164 pacientes que recibieron el tratamiento híbrido sobrevivieron por lo menos ese tiempo. Si p_1 denota la proporción de las mujeres que, al ser tratadas sólo con quimioterapia, sobreviven por lo menos 15 años y p_2 denota la proporción análoga para el tratamiento híbrido, $\hat{p}_1 = 76/154 = 0.494$ y $98/164 = 0.598$. Un intervalo de confianza para la diferencia entre proporciones basada en la fórmula tradicional con un nivel de confianza de aproximadamente 95% es

$$0.494 - 0.598 \pm (2.58)\sqrt{\frac{(0.494)(0.506)}{154} + \frac{(0.598)(0.402)}{164}} =$$

$$-0.104 \pm 0.143 = (-0.247, 0.039)$$

Al nivel de confianza de 99%, es posible que $-0.247 < p_1 - p_2 < 0.039$. Este intervalo tiene una amplitud razonable, una reflexión del hecho de que los tamaños de muestra no son demasiado grandes para este tipo de intervalo. Observe que 0 es uno de los valores posibles de $p_1 - p_2$, lo cual indica que ningún tratamiento puede ser juzgado superior al otro. Usando $\tilde{p}_1 = 77/156 = 0.494$, $\tilde{q}_1 = 79/156 = 0.506$, $\tilde{p}_2 = 0.596$, $\tilde{q}_2 = 0.404$ con base en tamaños de muestra de 156 y 166, respectivamente, el intervalo "mejorado" aquí es idéntico al intervalo anterior. ■

Inferencias con muestras pequeñas

En ocasiones, una inferencia en relación con $p_1 - p_2$ podría tener que basarse en muestras para las que por lo menos un tamaño de muestra es pequeño. Los métodos apropiados para este tipo de situaciones no son tan directos como los de muestras grandes, y hay más controversia entre los estadísticos en cuanto a procedimientos recomendados. Una prueba utilizada con frecuencia, conocida como prueba de Fisher-Irwin, se basa en la distribución hipergeométrica. Para ampliar su información consulte con un estadístico.

Ejercicios | Sección 9.4 (47 a 56)

47. ¿Alguien que cambia de marcas como resultado de incentivos económicos tiene menos probabilidades de permanecer leal a una marca que alguien que cambia sin incentivo alguno? Represente con p_1 y p_2 las proporciones reales de quienes cambian a una cierta marca con y sin incentivos, respectivamente, y después repiten una compra. Pruebe H_0: $p_1 - p_2 = 0$ contra H_a: $p_1 - p_2 < 0$, usando $\alpha = 0.01$ y los siguientes datos:

$m = 200$ número de éxitos = 30

$n = 600$ número de éxitos = 180

(Se proporcionan datos similares en "Impact of Deals and Deal Retraction on Brand Switching", *J. Marketing*, 1980: 62-70.)

48. En una muestra de 300 residentes adultos de una zona urbana de un estado en particular se encontró que 63 estaban a favor de aumentar el límite de velocidad en carreteras de 55 a 65 millas por hora, mientras que en una muestra de 180 residentes rurales se encontró que 75 favorecían el incremento de velocidad. ¿Estos datos indican que la opinión de incrementar el límite de velocidad es diferente en los dos grupos de residentes?
   a. Pruebe H_0: $p_1 = p_2$ contra H_a: $p_1 \neq p_2$ utilizando $\alpha = 0.05$, donde p_1 se refiere a la población urbana.
   b. Si las proporciones reales en favor del aumento son en realidad $p_1 = 0.20$ (urbana) y $p_2 = 0.40$ (rural), ¿cuál es la probabilidad de que H_0 sea rechazada con una prueba de nivel 0.05 con $m = 300$, y $n = 180$?

49. Se cree que la portada y la naturaleza de la primera pregunta de encuestas por correo influyen en la tasa de respuesta. En el artículo "The Impact of Cover Design and First Questions on Response Rates for a Mail Survey of Skydivers" (*Leisure Sciences*, 1991: 67-76) se probó esta teoría al experimentar con diferentes diseños de portada. Una portada era sencilla y la otra consistía en la fotografía de un paracaidista. Los investigadores especularon que la tasa de devolución sería menor para la portada sencilla.

Portada	Cantidad enviada	Cantidad de vuelta
Sencilla	207	104
Paracaidista	213	109

¿Esta información respalda la hipótesis de los investigadores? Pruebe las hipótesis pertinentes con $\alpha = 0.10$ calculando primero un valor P.

50. ¿Los maestros consideran que su trabajo es remunerador y satisfactorio? En el artículo "Work-Related Attitudes" (*Psychological Reports*, 1991: 443-450) se dan los resultados de un estudio de 395 maestros de escuela primaria y 266 de secundaria. De los primeros, 224 dijeron estar muy satisfechos con su trabajo, mientras que 126 de los maestros de secundaria afirmaron lo mismo. A fin de estimar la diferencia entre la proporción de los maestros de escuela primaria y los de secundaria que están satisfechos calcule un IC.

51. En una muestra aleatoria de 5726 números telefónicos de cierta región, tomada en marzo de 1992, se encontró que 1105 no estaban en el directorio, y un año después, en una muestra de 5384 se determinó que 980 números no estaban listados.

a. Pruebe al nivel 0.10 para ver si hay una diferencia en las proporciones reales de números no listados entre los dos años.

b. Si $p_1 = 0.20$ y $p_2 = 0.18$, ¿qué tamaños de muestra ($m = n$) serían necesarios para detectar tal diferencia con probabilidad 0.90?

52. La radiación ionizante recibe cada vez más atención como método para preservar productos de horticultura. En el artículo "The Influence of Gamma-Irradiation on the Storage Life of Red Variety Garlic" (*J. of Food Processing and Preservation*, 1983: 179-183) se informa que 153, de 180 cabezas de ajo irradiadas eran vendibles (sin brotes, pudrición o reblandecimiento) 240 días después del tratamiento, mientras que sólo 119 de 180 cabezas no tratadas eran vendibles después de ese lapso. ¿Esta información indica que la radiación ionizante es benéfica en cuanto a la comercialización?

53. En investigaciones médicas, la relación $\theta = p_1/p_2$ es con frecuencia más interesante que la diferencia $p_1 - p_2$ (p. ej., ¿por cuánto se debe multiplicar la probabilidad de recuperación de los individuos que reciben el tratamiento 2 a fin de obtener la probabilidad de recuperación de quienes reciben el tratamiento 1?). Sea $\hat{\theta} = \hat{p}_1/\hat{p}_2$. Cuando m y n son grandes, la estadística $\ln(\hat{\theta})$ tiene aproximadamente una distribución normal con valor medio aproximado $\ln(\theta)$ y desviación estándar aproximada $[(m - x)/(mx) + (n - y)/(ny)]^{1/2}$.

a. Utilice la información anterior para obtener una fórmula del intervalo de confianza de 95% con muestras grandes para estimar $\ln(\theta)$ y después un IC para θ.

b. Regrese a los datos de ataques al corazón del ejemplo 1.3 y calcule un intervalo de valores posibles para θ al nivel de confianza de 95%. ¿Qué indica este intervalo acerca de la eficacia del tratamiento con aspirina?

54. A veces los experimentos que requieren respuestas de éxito o fracaso se ejecutan de modo que se obtengan pares de datos o primero una respuesta y después otra. Suponga que antes que un candidato pronuncie su discurso se elige a n personas y se les pregunta si están a favor del candidato (S) o no (F). Luego, tras el discurso, se hace la misma pregunta a las mismas n personas. Las respuestas se anotan en una tabla como sigue:

		Después S	*Después* F
Antes	S	X_1	X_2
	F	X_3	X_4

donde $X_1 + X_2 + X_3 + X_4 = n$. Represente con p_1, p_2, p_3 y p_4 las cuatro probabilidades de celda, de modo que $p_1 = P(S \text{ antes y } S \text{ después})$, etc. Se desea probar la hipótesis de que no se incrementó la proporción real de simpatizantes (S) después del discurso contra la alternativa de que sí aumentó.

a. Exprese las dos hipótesis de interés en términos de p_1, p_2, p_3 y p_4.

b. Construya un estimador para la diferencia antes y después en probabilidades de éxito.

c. Cuando n es grande, se puede demostrar que la variable aleatoria $(X_1 - X_4)/n$ tiene una distribución normal aproximada con varianza dada por $[p_1 + p_4 - (p_1 - p_4)^2]/n$. Utilice esto para construir un estadístico de prueba con una distribución normal estándar aproximada cuando H_0 es verdadera (el resultado se llama prueba de McNemar).

d. Si $x_1 = 350$, $x_2 = 150$, $x_3 = 200$ y $x_4 = 300$, ¿qué se puede concluir?

55. Se utilizaron dos tipos de aleación, A y B, para fabricar especímenes experimentales de un pequeño eslabón de tensión que se utilizará en cierta aplicación de ingeniería. Se determinó la resistencia máxima (en klb/pulg^2) de cada espécimen y los resultados se resumen en la siguiente distribución de frecuencias.

	A	B
$26 - < 30$	6	4
$30 - < 34$	12	9
$34 - < 38$	15	19
$38 - < 42$	7	10
	$m = 40$	$n = 42$

Calcule un intervalo de confianza de 95% para la diferencia entre las proporciones reales de los especímenes de aleaciones A y B que tienen una resistencia máxima de por lo menos 34 klb/pulg^2.

56. Con base en tamaños de muestra iguales se construirá un intervalo de confianza de 95% para $p_1 - p_2$ por medio de la fórmula tradicional. ¿Para qué valor de n ($= m$) el intervalo resultante tendrá una extensión de a lo sumo 0.1 sin importar los resultados del muestreo?

9.5 Inferencias relacionadas con dos varianzas poblacionales

En ocasiones son necesarios métodos para comparar dos varianzas poblacionales (o desviaciones estándar), aunque este tipo de problemas surgen con mucha menos frecuencia que los que tienen que ver con medias o proporciones. Para el caso en el que las poblaciones en investigación son normales, los procedimientos se basan en una nueva familia de distribuciones de probabilidad.

Distribución *F*

La distribución de probabilidad F tiene dos parámetros, que se denotan con ν_1 y ν_2. El parámetro ν_1 se conoce como *número de grados de libertad del numerador* y ν_2 es el *número de grados de libertad del denominador*; aquí ν_1 y ν_2 son enteros positivos. Una variable aleatoria que tiene una distribución F no puede tomar un valor negativo. Puesto que la función de densidad es complicada y no se usará de manera explícita, se omite la fórmula. Hay una relación importante entre una variable F y las variables chi-cuadradas. Si X_1 y X_2 son variables aleatorias chi-cuadradas independientes con ν_1 y ν_2 grados de libertad, respectivamente, entonces se puede demostrar que la *va*

$$F = \frac{X_1/\nu_1}{X_2/\nu_2} \tag{9.8}$$

la relación de las dos variables chi-cuadradas divididas entre sus respectivos grados de libertad, tiene una distribución F.

En la figura 9.8 se ilustra la gráfica de una función de densidad F representativa. Análoga a la notación $t_{\alpha,\nu}$ y $\chi^2_{\alpha,\nu}$ se utiliza F_{α,ν_1,ν_2} para el punto sobre el eje que capta α del área bajo la curva de densidad F con ν_1 y ν_2 gl en el extremo superior. La curva de densidad no es simétrica, así que al parecer se deben tabular los valores críticos de ambos extremos. Sin embargo esto es innecesario debido a la siguiente propiedad.

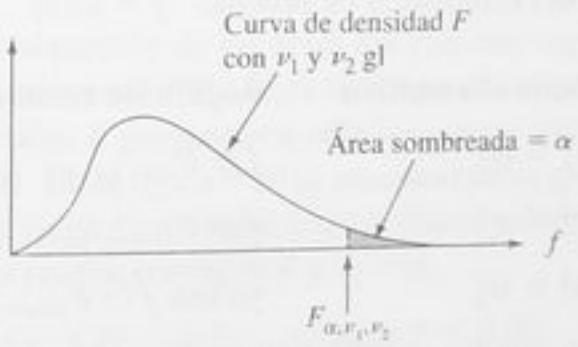


Figura 9.8 Curva de densidad F y valor crítico

$$F_{1-\alpha,\nu_1,\nu_2} = 1/F_{\alpha,\nu_2,\nu_1} \tag{9.9}$$

En la tabla A.9 del apéndice se obtiene F_{α,ν_1,ν_2} para $\alpha = 0.10$, 0.05 y 0.001 y varios valores de ν_1 (en diferentes columnas de la tabla) y ν_2 (en diferentes grupos de renglones de la

tabla). Por ejemplo, $F_{0.05,6,10} = 3.22$ y $F_{0.05,10,6} = 4.06$. Para obtener $F_{0.95,6,10}$, el número que capta 0.95 del área a su derecha (y, por lo tanto, 0.05 a la izquierda) bajo la curva F con $\nu_1 = 6$ y $\nu_2 = 10$, se utiliza (9.9): $F_{0.95,6,10} = 1/F_{0.05,10,6} = 1/4.06 = 0.246$.

Métodos inferenciales

Un procedimiento de prueba para hipótesis que tienen que ver con la relación σ_1^2/σ_2^2, así como con un IC para esta relación, se basan en el siguiente resultado.

TEOREMA

Sea $X_1, \ldots, X_m$ una muestra aleatoria de una distribución normal con varianza σ_1^2, sea $Y_1, \ldots, Y_n$ otra muestra aleatoria (independiente de las X_i) de una distribución normal con varianza σ_2^2, y denote con S_1^2 y S_2^2 las dos varianzas muestrales. Entonces la *va*

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \tag{9.10}$$

tiene una distribución F con $\nu_1 = m - 1$ y $\nu_2 = n - 1$.

Este teorema resulta de combinar (9.8) con el hecho de que cada una de las variables $(m-1)S_1^2/\sigma_1^2$ y $(n-1)S_2^2/\sigma_2^2$ tengan una distribución chi-cuadrada con $m - 1$ y $n - 1$ grados de libertad, respectivamente (véase la sección 7.4). Debido a que F tiene que ver con una relación y no con una diferencia, el estadístico de prueba es la relación de varianzas muestrales; la afirmación de que $\sigma_1^2 = \sigma_2^2$ se rechaza si la relación difiere en gran medida de 1.

Prueba F para igualdad de varianzas

Hipótesis nula: H_0: $\sigma_1^2 = \sigma_2^2$

Valor del estadístico de prueba: $f = s_1^2/s_2^2$

Hipótesis alternativa	**Región de rechazo para la prueba de nivel α**
H_a: $\sigma_1^2 > \sigma_2^2$	$f \geq F_{\alpha,m-1,n-1}$
H_a: $\sigma_1^2 < \sigma_2^2$	$f \leq F_{1-\alpha,m-1,n-1}$
H_a: $\sigma_1^2 \neq \sigma_2^2$	ya sea $f \geq F_{\alpha/2,m-1,n-1}$ o $f \leq F_{1-\alpha/2,m-1,n-1}$

Puesto que los valores críticos se tabulan sólo para $\alpha = 0.10, 0.05, 0.01$ y 0.001, la prueba bilateral se puede llevar a cabo sólo a los niveles 0.20, 0.10, 0.02 y 0.002. En otras partes están disponibles tabulaciones más extensas de los valores críticos F.

Ejemplo 9.14 Con base en los datos publicados en un artículo de 1979 en el *Journal of Gerontology* ("Serum Ferritin in an Elderly Population", pp. 521-254), los autores concluyeron que la distribución

de ferritina en los ancianos tenía una varianza más pequeña que en los adultos jóvenes. (La ferritina sérica se utiliza para diagnosticar deficiencia de hierro.) Para una muestra de 28 hombres ancianos, la desviación estándar muestral de la ferritina sérica (mg/L) fue $s_1 = 52.6$; para 26 varones jóvenes, la desviación estándar de la muestra fue $s_2 = 84.2$. ¿Estos datos respaldan la conclusión aplicada a los varones?

Denote con σ_1^2 y σ_2^2 la varianza de las distribuciones de ferritina sérica para ancianos y jóvenes, respectivamente. Se desea probar H_0: $\sigma_1^2 = \sigma_2^2$ contra H_a: $\sigma_1^2 < \sigma_2^2$. Al nivel 0.01, se rechaza H_0 si $f \le F_{0.99,27,25}$. Para obtener el valor crítico, se necesita $F_{0.01,25,27}$. De la tabla A.9 del apéndice, $F_{0.01,25,27} = 2.54$, por lo tanto $F_{0.99,27,25} = 1/2.54 = 0.394$. El valor calculado de F es $(52.6)^2/(84.2)^2 = 0.390$. Puesto que $0.390 \le 0.394$, se rechaza H_0 al nivel 0.01 en favor de H_a, así que la variabilidad al parecer es mayor en los jóvenes que en los ancianos. ■

Valores *P* para las pruebas *F*

Recuerde que el valor P para una prueba t de extremo superior es el área bajo la curva t pertinente (la que tenga los grados de libertad apropiados) a la derecha de la t calculada. En la misma forma, el valor P para una prueba F de extremo superior es el área bajo la curva F con los grados de libertad apropiados en el numerador y denominador a la derecha de la f calculada. En la figura 9.9 se ilustra esto para una prueba basada en $\nu_1 = 4$ y $\nu_2 = 6$.

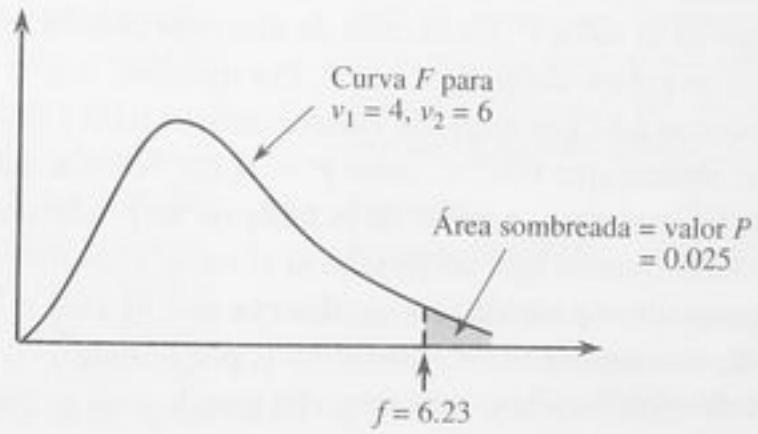


Figura 9.9 Valor *P* para una prueba *F* de extremo superior

Infortunadamente, la tabulación de las áreas del extremo superior de la curva F es mucho más difícil que para las curvas t debido a que intervienen dos grados de libertad. Para cada combinación de ν_1 y ν_2, la tabla F proporciona sólo los cuatro valores críticos que captan las áreas 0.10, 0.05, 0.01 y 0.001. En la figura 9.10 se muestra que lo que se puede decir acerca del valor P en función de donde caiga f con respecto a los cuatro valores críticos.

Por ejemplo, para una prueba con $\nu_1 = 4$ y $\nu_2 = 6$,

$$f = 5.70 \Rightarrow 0.01 < \text{valor } P < 0.05$$
$$f = 2.16 \Rightarrow \text{valor } P > 0.10$$
$$f = 25.03 \Rightarrow \text{valor } P < 0.001$$

Sólo si f es igual a un valor tabulado, se obtiene un valor P exacto (p. ej., si $f = 4.53$, entonces el valor de $P = 0.05$). Una vez que se sabe que $0.01 <$ valor $P < 0.05$, H_0 se rechazaría al nivel de significación de 0.05 pero no a un nivel de 0.01. Cuando el valor $P < 0.001$, H_0 se debe rechazar a cualquier nivel de significación razonable.

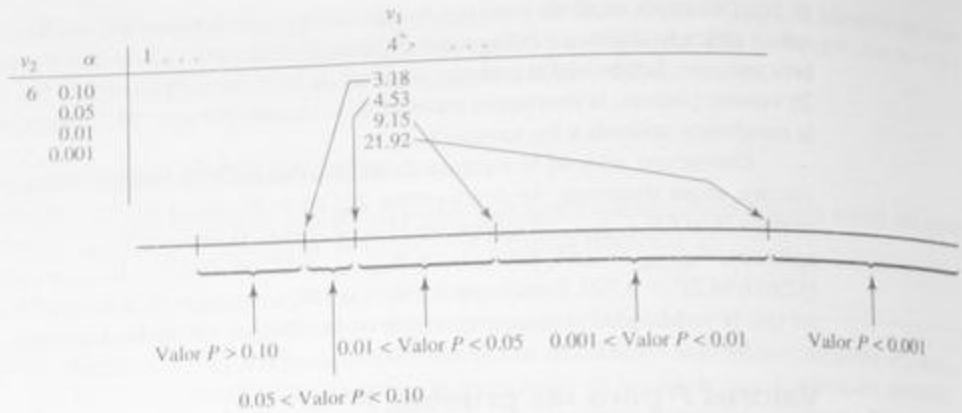


Figura 9.10 Obtención de información del valor P de la tabla F para una prueba F de extremo superior

Las pruebas F analizadas en capítulos posteriores, serán de extremo superior. Sin embargo, si es apropiada la prueba F de extremo inferior, entonces se debe usar la ecuación (9.9) para obtener valores críticos de extremo inferior de modo que se pueda establecer un límite o límites en el valor P. En el caso de una prueba bilateral, el límite o límites de una prueba unilateral se deben multiplicar por 2. Por ejemplo, si $f = 5.82$ cuando $\nu_1 = 4$ y $\nu_2 = 6$, entonces puesto que 5.82 cae entre los valores críticos 0.05 y 0.01, $2(0.01) <$ valor $P < 2(0.05)$, de donde se obtiene que $0.02 <$ valor $P < 0.10$. Se debe rechazar H_0 si $\alpha = 0.10$., pero no si $\alpha = 0.01$. En este caso, a partir de la tabla no se puede decir qué conclusión es apropiada cuando $\alpha = 0.05$ (puesto que no se sabe si el valor P es menor o mayor que esto). Sin embargo, con un programa de estadística se observa que el área a la derecha de 5.82 bajo esta curva F es 0.029, entonces el valor P es 0.058 y, por lo tanto, no se debe rechazar la hipótesis nula al nivel 0.05 (0.058 es la α más pequeña para la cual se puede rechazar H_0 y la α elegida es menor que este valor).

Intervalo de confianza para σ_1/σ_2

El IC para σ_1^2/σ_2^2 se basa en reemplazar F en la expresión de probabilidad

$$P(F_{1-\alpha/2,\nu_1,\nu_2} < F < F_{\alpha/2,\nu_1,\nu_2}) = 1 - \alpha$$

por la variable F (9.10) y hacer operaciones con las desigualdades para aislar σ_1^2/σ_2^2. Un intervalo para σ_1/σ_2 resulta de sacar raíz cuadrada de cada límite. Los detalles se dejan para un ejercicio.

Ejercicios | Sección 9.5 (57 a 64)

57. Obtenga o calcule las siguientes cantidades:
    a. $F_{0.05,5,8}$ b. $F_{0.05,8,5}$ c. $F_{0.95,5,8}$ d. $F_{0.95,8,5}$
    e. El percentil 99 de la distribución F con $\nu_1 = 10$, $\nu_2 = 12$
    f. El primer percentil de la distribución F con $\nu_1 = 10$, $\nu_2 = 12$
    g. $P(F \le 6.16)$ para $\nu_1 = 6$, $\nu_2 = 4$
    h. $P(0.177 \le F \le 4.74)$ para $\nu_1 = 10$, $\nu_2 = 5$

58. Proporcione la mayor información posible acerca del valor P de la prueba F en cada uno de los casos siguientes:
    a. $\nu_1 = 5$, $\nu_2 = 10$, prueba de extremo superior, $f = 4.75$
    b. $\nu_1 = 5$, $\nu_2 = 10$, prueba de extremo superior, $f = 2.00$
    c. $\nu_1 = 5$, $\nu_2 = 10$, prueba bilateral, $f = 5.64$
    d. $\nu_1 = 5$, $\nu_2 = 10$, prueba de extremo inferior, $f = 0.200$
    e. $\nu_1 = 35$, $\nu_2 = 20$, prueba de extremo superior, $f = 3.24$

59. Vuelva a los datos acerca del ángulo de inclinación máxima proporcionados en el ejercicio 28 de este capítulo. Realice una prueba al nivel de significación 0.10 para ver si las desviaciones estándar poblacionales para los dos grupos son diferentes (las gráficas de probabilidad normal respaldan la suposición de normalidad necesaria).

60. Refiérase el ejemplo 9.7. ¿Los datos indican que la desviación estándar de la distribución de resistencias de especímenes fundidos es menor que la de especímenes no fundidos? Realice una prueba al nivel de significación 0.01 al obtener tanta información como sea posible acerca del valor P.

61. El toxafeno es un insecticida que ha sido identificado como contaminante en el ecosistema de los Grandes Lagos. Para investigar el efecto de la exposición al toxafeno en animales, se administró este insecticida a grupos de ratas en sus dietas. En el artículo "Reproduction Study of Toxaphene in the Rat" (*J. of Environ. Sci. Health*, 1988: 101-126) se informa el aumento de peso (en gramos) para ratas a las que se les administró una dosis baja (4 ppm) y para ratas de control cuya dieta no incluyó el insecticida. La desviación estándar muestral para 23 ratas hembras de control fue 32 g y para 20 ratas hembras con dosis baja fue 54 g. ¿Estos datos indican que hay más variabilidad en los aumentos de peso con dosis bajas que en las ganancias de peso de ratas de control? Si se supone normalidad, realice una prueba de hipótesis al nivel de significación 0.05.

62. En un estudio de deficiencia de cobre en ganado se determinaron los valores de cobre (μg Cu/100 mL de sangre) tanto para ganado apacentado en un lugar que se sabe tiene anomalías de molibdeno bien definidas (valores de metales que sobrepasan el intervalo normal de de variación regional), como para ganado pastoreado en un lugar sin anomalías ("An Investigation into Copper Deficiency in Cattle in the Southern Pennines", *J. Agricultural Soc. Cambridge*, 1972: 157-163). El resultado fue $s_1 = 21.5$ ($m = 48$) para la condición anómala y $s_2 = 19.45$ ($n = 45$) para la condición sin anomalías. Pruebe la igualdad contra la desigualdad de varianzas poblacionales al nivel de significación 0.10 por medio del método del valor P.

63. En el artículo "Enhancement of Compressive Properties of Failed Concrete Cylinders with Polymer Impregnation" (*J. Testing and Evaluation*, 1977: 333-337) se dan los siguientes datos acerca del módulo compresivo impregnado (lb/pulg $\times 10^6$) cuando se usaron dos polímeros diferentes para reparar grietas en concreto.

Epóxico	1.75	2.12	2.05	1.97
Prepolímero MMA	1.77	1.59	1.70	1.69

Obtenga un IC de 90% para la relación de varianzas. Utilice primero el método sugerido en el texto al obtener una fórmula del intervalo de confianza general.

64. Reconsidere los datos del ejemplo 9.6 y calcule un límite de confianza superior de 95% para la relación entre la desviación estándar de la distribución de porosidad del triacetato y la que corresponde a la distribución de porosidad del algodón.

Ejercicios complementarios (65 a 93)

65. Los siguientes datos de resumen acerca de la resistencia a la compresión (lb) para cajas de $12 \times 10 \times 8$ pulg se publicaron en el artículo "Compression of Single-Wall Corrugated Shipping Containers Using Fixed and Floating Test Platens" (*J. Testing and Evaluation*, 1992: 318-320). Los autores expresaron que "la diferencia entre la resistencia a la compresión con los métodos de platina fija y flotante fue pequeña en comparación con la variación normal de resistencia a la compresión entre cajas idénticas". ¿Está de acuerdo?

Método	Tamaño de la muestra	Media muestral	DE de la muestra
Fijo	10	807	27
Flotante	10	757	41

66. Los autores del artículo "Dynamics of Canopy Structure and Light Interception in *Pinus elliotti*, North Florida" (*Ecological Monographs*, 1991: 33-51) planearon un experimento para determinar el efecto de fertilizante

en una medida de área de hoja. Para el estudio se disponía de varias parcelas y se seleccionó la mitad al azar para fertilizarlas. A fin de asegurar que las parcelas recibieran el fertilizante y las parcelas de control fueran similares, antes de comenzar el experimento se registró la densidad de árboles (número de árboles por hectárea) para ocho parcelas que recibirían el fertilizante y ocho parcelas de control, obteniéndose los siguientes datos. A continuación se muestra el resultado de MINITAB.

Parcelas con fertilizante	1024	1216	1312	1280
	1216	1312	992	1120
Parcelas de control	1104	1072	1088	1328
	1376	1280	1120	1200

```
T con dos muestras para fertilizante
contra control

                N   Media   Desv.   EE de la
                            est.      media
Fertilizante    8    1184     126         44
Control         8    1196     118         42

IC de 95% para mu de fertilizante - mu de
control: (-144, 120)
```

a. Trace un diagrama de caja comparativo y comente acerca de algunas características interesantes.

b. ¿Concluiría que hay una diferencia importante en la densidad media de árboles para parcelas fertilizadas y de control? Utilice $\alpha = 0.05$.

c. Interprete el intervalo de confianza proporcionado.

67. ¿Se ve afectada la tasa de respuesta para cuestionarios al incluir algún tipo de incentivo que contestar junto con el cuestionario? En un experimento, de 110 cuestionarios sin incentivo regresaron 75, mientras que de 98 cuestionarios que incluían una posibilidad de ganar una lotería se obtuvieron 66 respuestas ("Charities, No; Lotteries, No; Cash, Yes", *Public Opinion Quarterly*, 1996: 542-562). ¿Estos datos indican que incluir un incentivo aumenta la probabilidad de una respuesta? Enuncie y pruebe las hipótesis pertinentes al nivel de significación 0.10 usando el método del valor *P*.

68. Los siguientes datos se obtuvieron de un estudio para evaluar el potencial de licuefacción en una nucleoeléctrica propuesta ("Cyclic Strengths Compared for Two Sampling Techniques", *J. Geotechnical Division, Am. Soc. Civil Engrs. Proceedings*, 1981: 563-576). Antes de probar la resistencia cíclica, se reunieron muestras de suelo con el método de tubo Pitcher y un método de bloques, y se obtuvieron los siguientes valores de densidad en seco (lb/pie^3).

Muestreo con tubo	101.1	111.1	107.6	98.1
	99.5	98.7	103.3	108.9
	109.1	104.1	110.0	98.4
	105.1	104.5	105.7	103.3
	100.3	102.6	101.7	105.4
	99.6	103.3	102.1	104.3
Muestreo de bloques	107.1	105.0	98.0	97.9
	103.3	104.6	100.1	98.2
	97.9	103.2	96.9	

(continúa)

Calcule e interprete un IC de 95% para la diferencia entre las densidades en seco promedio reales para los dos métodos de muestreo.

69. En el artículo "Quantitative MRI and Electrophysiology of Preoperative Carpal Tunnel Syndrome in a Female Population" (*Ergonomics*, 1997: 642-649) se describe que (−473.3, 1961.9) fue un intervalo de confianza de 95% con muestras grandes para la diferencia entre el volumen promedio real del músculo ténar (mm^3) para quienes padecen del síndrome del túnel carpiano y el volumen promedio real para quienes no padecen la enfermedad. Calcule un intervalo de confianza de 90% para esta diferencia.

70. Los siguientes datos de resumen acerca de la resistencia a la flexión (lb-pulg/pulg) de uniones se tomaron del artículo "Bending Strength of Corner Joints Constructed with Injection Molded Splines" (*Forest Products J.*, abril de 1997: 89-92).

Tipo	Tamaño de la muestra	Media muestral	DE de la muestra
Sin cubierta lateral	10	80.95	9.59
Con cubiera lateral	10	63.23	5.96

a. Calcule un límite de confianza inferior de 95% para la resistencia promedio real de uniones con cubierta lateral.

b. Calcule un límite inferior de predicción de 95% para la resistencia de una sola unión con cubierta lateral.

c. Calcule un intervalo que, con 95% de confianza, incluya los valores de resistencia de al menos 95% de la población de todas las uniones con cubierta lateral.

d. Calcule un intervalo de confianza de 95% para la diferencia entre las resistencias promedio reales para los dos tipos de uniones.

71. Se realizó un experimento para comparar diversas propiedades de hilo de algodón y poliéster terminado sólo con suavizador e hilo terminado con suavizador más 5% de resina DP ("Properties of a Fabric Made with Tandem Spun Yarns", *Textile Res. J.*, 1996: 607-611). Una característica particularmente importante de la tela es su durabilidad, es decir, su capacidad de resistir el uso. Para una muestra de 40 especímenes sólo con suavizador, el promedio muestral de la resistencia a la abrasión (ciclos)

en la dirección del hilo fue 3975.0, con una desviación estándar de 245.1. Con otra muestra de especímenes con suavizador y aditivo se obtuvo un promedio y una desviación estándar muestrales de 2795.0 y 293.7, respectivamente. Calcule un intervalo de confianza de 99% para la diferencia entre las resistencias promedio reales a la abrasión en los dos tipos de tela. ¿El intervalo proporciona una prueba convincente de que las resistencias promedio reales difieren para los dos tipos de tela? ¿Por qué sí o por qué no?

72. El descarrilamiento de un tren de carga debido a la falla catastrófica de un cojinete de la armadura de un motor de tracción favoreció un estudio publicado en "Locomotive Traction Motor Armature Bearing Life Study" (*Lubrication Engr.*, agosto de 1997: 12-19). Se seleccionó una muestra de 17 motores de tracción con muchas millas recorridas y se determinó la cantidad de penetración de cono (mm/10) para el cojinete del piñón y para el cojinete de la armadura del conmutador, y se obtuvieron los datos siguientes:

Motor	1	2	3	4	5	6
Conmutador	211	273	305	258	270	209
Piñón	226	278	259	244	273	236
Motor	7	8	9	10	11	12
Conmutador	223	288	296	233	262	291
Piñón	290	287	315	242	288	242
Motor	13	14	15	16	17	
Conmutador	278	275	210	272	264	
Piñón	278	208	281	274	268	

Calcule una estimación de la diferencia promedio poblacional entre la penetración para el cojinete de la armadura del conmutador y la del cojinete del piñón. El cálculo debe proporcionar información acerca de la fiabilidad y la precisión de la estimación. (*Nota:* una gráfica de probabilidad normal valida la hipótesis de normalidad necesaria.) ¿Se podría afirmar que se estimó con precisión la diferencia promedio de la población? ¿Al parecer la penetración promedio poblacional difiere para los dos tipos de cojinete? Explique.

73. La conformabilidad es la capacidad de una pieza cilíndrica de material para darle forma de cabeza de un perno, tornillo u otra parte conformada en frío sin agrietarse. En el artículo "New Methods for Assessing Cold Heading Quality" (*Wire J. Intl.*, octubre de 1996: 66-72) se describe el resultado de una prueba de impacto de conformabilidad aplicada a 30 especímenes de acero desoxidado con aluminio y 30 de acero desoxidado con silicio. El número de clasificación de conformabilidad promedio muestral de los especímenes de acero oxidados con silicio fue 6.43 y el del acero oxidado con aluminio 7.09. Suponga que las desviaciones estándar muestrales fueron 1.08 y 1.19, respectivamente. ¿Concuerda con los autores del artículo en que la diferencia entre conformabilidades es significativa al nivel de 5% (suponiendo que las dos distribuciones de conformabilidad son normales)?

74. En el artículo "Two Parameters Limiting the Sensitivity of Laboratory Tests of Condoms as Viral Barriers" (*J. of Testing and Eval.*, 1996: 279-286) se informa que, en condones de la marca A, entre 16 rasgaduras producidas con una aguja, la longitud promedio muestral de la rasgadura fue 74.0 μm, mientras que para las 14 rasgaduras en la marca B, el promedio muestral fue 61.0 μm (ambas determinadas mediante microscopio de luz y micrográficas de barrido electrónico). Suponga que las desviaciones estándar muestrales son 14.8 y 12.5, respectivamente (consistentes con los intervalos muestrales que aparecen en el artículo). Los autores comentaron que el condón marca B más grueso mostró menor longitud promedio de rasgadura que el de la marca A más delgado. ¿En realidad esta diferencia tiene importancia estadística? Enuncie las hipótesis adecuadas y pruebe con $\alpha = 0.05$.

75. La información acerca de la postura de la mano y las fuerzas que generan los dedos durante el manejo de diversos objetos cotidianos se requiere para diseñar aparatos para prótesis de alta tecnología. En el artículo "Grip Posture and Forces During Holding Cylindrical Objects with Circular Grips" (*Ergonomics*, 1996: 1163-1176) se explica que para una muestra de 11 mujeres, la fuerza promedio muestral de estricción con cuatro dedos (N) fue 98.1 y la desviación estándar muestral fue 14.2. Para una muestra de 15 varones, las cantidades correspondientes fueron 129.2 y 39.1.
    a. Una prueba realizada para comprobar si las fuerzas promedio reales de los dos géneros son diferentes dio como resultado $t = 2.51$ y valor $P = 0.019$. ¿El procedimiento de prueba apropiado descrito en este capítulo produce este valor de t y el valor expresado para P?
    b. ¿Hay una evidencia convincente para concluir que la fuerza promedio real de los varones es mayor que la de las mujeres en más de 25 N? Enuncie y pruebe las hipótesis pertinentes.

76. En el artículo "Pine Needles as Sensors of Atmospheric Pollution" (*Environ. Monitoring*, 1982: 273-286) se describe el uso de análisis por actividad de neutrones para determinar la concentración de contaminantes en agujas de pinos. Según los autores del artículo, "Las observaciones indicaron con certeza que para los elementos que se determinan bien mediante procedimientos analíticos, la distribución de la concentración es lognormal. En conse-

cuencia, en pruebas de significación se usarán los logaritmos de las concentraciones". Los datos se refieren a la concentración de bromo en las agujas tomadas de un sitio cercano a una termoeléctrica alimentada con petróleo y de un lugar relativamente limpio. Los valores de resumen son las medias y desviaciones estándar de las observaciones transformadas a logaritmos.

Sitio	Tamaño de la muestra	Concentración log media	DE de la concentración log
Termoeléctrica	8	18.0	4.9
Sitio limpio	9	11.0	4.6

Sea μ_1^* la concentración *log* promedio real en el primer sitio, y defina μ_2^* de manera análoga para el segundo sitio.

a. Use la prueba t combinada (basada en la hipótesis de normalidad y en las desviaciones estándar iguales) para decidir, al nivel de significación 0.05, si los dos promedios de distribución de concentraciones son iguales.

b. Si σ_1^* y σ_2^*, las desviaciones estándar de las dos distribuciones de concentración logarítmica, no son iguales, ¿serían iguales μ_1 y μ_2, los promedios de las distribuciones de concentración, si $\mu_1^* = \mu_2^*$? Explique su razonamiento.

77. La exposición prolongada de trabajadores textiles al polvo de algodón liberado durante el procesamiento, origina importantes problemas de salud. Por consiguiente, los investigadores textiles han estado estudiando métodos que produzcan riesgos menores a la vez que se conservan propiedades importantes de las telas. Los datos siguientes acerca de la resistencia de cohesión de la mecha (kN•m/kg) para especímenes producidos en cinco múltiples distintos de torcido se tomaron del artículo "Heat Treatment of Cotton: Effect on Endotoxin Content, Fiber and Yarn Properties, and Processability" (*Textile Research J.*, 1996: 727-738):

Múltiple de torsión	1.054	1.141	1.245	1.370	1.481
Resistencia de control	0.45	0.60	0.61	0.73	0.69
Resistencia al calor	0.51	0.59	0.63	0.73	0.74

Los autores del artículo dijeron que, al parecer, la resistencia de los especímenes calentados era ligeramente mayor en promedio que para los especímenes de control. ¿La diferencia es estadísticamente significativa? Enuncie y pruebe las hipótesis pertinentes con $\alpha = 0.05$ al calcular el valor P.

78. Los siguientes datos de resumen acerca de la relación de resistencia con respecto al área de sección transversal para extensores de rodilla se tomaron del artículo "Knee Extensor and Knee Flexor Strength: Cross-Sectional Area Ratios in Young and Elderly Men" (*J. of Gerontology*, 1992: M204-M210).

Grupo	Tamaño de muestra	Media muestral	Error estándar
Jóvenes	13	7.47	0.22
Ancianos	12	6.71	0.28

¿Estos datos indican que la relación promedio real para jóvenes es mayor que la de ancianos? Realice una prueba de hipótesis apropiada con $\alpha = 0.05$. Asegúrese de expresar cualesquiera suposiciones necesarias para su análisis.

79. Los datos siguientes acerca del tiempo de respuesta aparecieron en el artículo "The Extinguishment of Fires Using Low-Flow Water Hose Streams-Part II" (*Fire Technology*, 1991: 291-320).

Buena visibilidad

0.43 1.17 0.37 0.47 0.68 0.58 0.50 2.75

Mala visibilidad

1.47 0.80 1.58 1.53 4.33 4.23 3.25 3.22

Los autores analizaron los datos con la prueba t combinada. ¿Se justifica el uso de esta prueba? (*Sugerencia*: compruebe la normalidad. Los puntuaciones normales para $n = 8$ son −1.53, −0.89, −0.49, −0.15, 0.15, 0.49, 0.89 y 1.53.)

80. En el artículo "The Relationship Between Distress and Delight in Males' and Females' Reactions to Frightening Films" (*Human Communication Research*, 1991: 625-637) se informa que los investigadores midieron respuestas emocionales de 50 hombres y 60 mujeres después de ver un segmento de una película de terror. El artículo incluyó el siguiente enunciado: "La probabilidad de expresar angustia fue mayor para las mujeres que para los varones. Aunque éstos mostraron niveles más altos de disfrute que las mujeres, la diferencia no fue estadísticamente importante". Los siguientes datos también aparecieron en el artículo.

	Índice de angustia		Índice de disfrute	
Género	**Media**	**DE**	**Media**	**DE**
Varones	31.2	10.0	12.02	3.65
Mujeres	40.4	9.1	9.09	5.55
	Valor $P < 0.001$		No significativo (Valor $P > 0.05$)	

Explique de manera breve las hipótesis y conclusiones pertinentes.

81. En un experimento para comparar las resistencias de soporte de espigas insertadas en dos tipos de soportes, una muestra de 14 observaciones acerca del límite de esfuerzo para soportes de roble rojo, dieron como resultado una media muestral y desviación estándar muestral de 8.48 MPa y 0.79 MPa, respectivamente, mientras que una muestra de 12 observaciones cuando se usaron soportes de abeto Douglas dieron una media de 9.36 y una desviación estándar de 1.52 ("Bearing Strength of White Oak Pegs in Red Oak and Douglas Fir Timbers", *J. of Testing and Evaluation*, 1998, 109-114). Considere probar si los límites de esfuerzo promedio reales son idénticos o no para los dos tipos de soporte. Compare los gl y los valores P para las pruebas t no combinada y combinada.

82. ¿Cómo se compara la ingestión de energía con el gasto de energía? Un aspecto de este tema se consideró en el artículo "Measurement of Total Energy Expenditure by the Doubly Labelled Water Method in Professional Soccer Players" (*J. of Sports Sciences*, 2002: 391-397), en el que se aparecen los datos siguientes (MJ/día).

Jugador	1	2	3	4	5	6	7
Gasto	14.4	12.1	14.3	14.2	15.2	15.5	17.8
Ingestión	14.6	9.2	11.8	11.6	12.7	15.0	16.3

Realice una prueba para ver si hay una diferencia importante entre la ingestión y el gasto. ¿La conclusión depende de si se emplea un nivel de significación de 0.05, 0.01 o 0.001?

83. Un experimentador desea obtener un IC para la diferencia entre la resistencia a la ruptura promedio real para los cables que fabrica la compañía I y la II. Suponga que la resistencia a la ruptura tiene una distribución normal para ambos tipos de cable con $\sigma_1 = 30$ lb/pulg^2 y $\sigma_2 = 20$ lb/ pulg^2.

a. Si los costos dictan que el tamaño muestral para el cable tipo I debe ser el triple del tamaño muestral para el cable tipo II, ¿cuántas observaciones se necesitan si el intervalo de confianza de 99% no debe ser mayor de 20 lb/pulg^2?

b. Suponga que se hará un total de 400 observaciones. ¿Cuántas de las observaciones deben hacerse en muestras del cable tipo I si la extensión del intervalo resultante debe ser la mínima?

84. Un experimento para determinar los efectos de la temperatura en la supervivencia de huevos de insectos se describió en el artículo "Development Rates and a Temperature-Dependent Model of Pales Weevil" (*Environ. Entomology*, 1987: 956-962). A 11°C, 73 de 91 huevos sobrevivieron hasta la siguiente etapa de desarrollo. A 30°C, 102 de 110 huevos sobrevivieron. ¿Los resultados de este experimento indican que la tasa de supervivencia (proporción de sobrevivientes) difiere para las dos temperaturas? Calcule el valor p y utilícelo para probar las hipótesis pertinentes.

85. Se midió la capacidad de enlace de insulina (pmol/mg de proteína) para una muestra de ratas diabéticas tratadas con dosis baja de insulina y otra muestra de ratas diabéticas tratadas con una dosis alta. Los resultados son los siguientes:

Dosis baja $m = 8, \bar{x} = 1.98, s_1 = 0.51$

Dosis alta $n = 12, \bar{y} = 1.30, s_2 = 0.35$

(*J. Clinical Investigation*, 1978: 552-560.)

a. ¿Los datos indican que hay alguna diferencia en la capacidad promedio real de enlace de insulina debido al nivel de la dosis? Utilice una prueba con $\alpha = 0.001$.

b. ¿Qué se puede decir acerca del valor P?

86. En el artículo citado en el ejercicio 85, en realidad hubo cuatro tratamientos experimentales: no diabéticas de control, diabéticas no tratadas, diabéticas de dosis baja y diabéticas de dosis alta. Represente con n_i el tamaño muestral para el i-ésimo tratamiento y con S_i^2 ($i = 1, 2, 3, 4$) la varianza muestral. Si se supone que la varianza real para cada tratamiento es σ^2, construya un estimador agrupado de σ^2 que sea insesgado, y verifique por medio de las reglas de valor esperado que en realidad es insesgado. ¿Cuál es su estimación para los siguientes datos reales? (*Sugerencia*: modifique el estimador combinado de la página 377.)

	Tratamiento			
	1	**2**	**3**	**4**
Tamaño de la muestra	16	18	8	12
DE de la muestra	0.64	0.81	0.51	0.35

87. Suponga que se va a realizar una prueba al nivel 0.05 de H_0: $\mu_1 - \mu_2 = 0$ contra H_a: $\mu_1 - \mu_2 > 0$, suponiendo $\sigma_1 = \sigma_2 = 10$ y normalidad de ambas distribuciones con muestras del mismo tamaño ($m = n$). Evalúe la probabilidad de un error tipo II cuando $\mu_1 - \mu_2 = 1$ y $n = 25$, 100, 2500 y 10 000. ¿Puede pensar en problemas reales en los que la diferencia $\mu_1 - \mu_2 = 1$ tenga poca importancia práctica? ¿Serían deseables tamaños muestrales de $n = 10\,000$ en tales problemas?

88. Los siguientes datos se refieren al número de bacterias que flotan en el aire (número de colonias/pie^3) tanto para $m = 8$ salas de hospital alfombradas y $n = 8$ salas no alfombradas ("Microbial Air Sampling in a Carpeted Hospital", *J. Environmental Health*, 1968: 405). ¿Parece haber diferencia en el promedio real del número de bacterias entre las salas alfombradas y las no alfombradas?

Alfombradas	11.8	8.2	7.1	13.0	10.8	10.1	14.6	14.0
No alfombradas	12.1	8.3	3.8	7.2	12.0	11.1	10.1	13.7

Suponga que después se entera de que todas las salas alfombradas eran de hospitales para veteranos de la guerra mientras que las no alfombradas eran de hospitales infantiles. ¿Se podría evaluar el efecto de alfombrar? Comente.

89. Investigadores enviaron 5000 currículos en respuesta a anuncios de empleo que aparecieron en el *Boston Globe* y *Chigago Tribune*. Los currículos fueron idénticos, excepto que 2500 tuvieron nombres que se podían relacionar con personas blancas, como Brett y Emily, mientras que 2500 tuvieron nombres que podían relacionarse con personas negras, como Tamika y Rasheed. Los currículos del primer tipo suscitaron 250 respuestas y los del segundo tipo sólo 167 respuestas (estas cifras son consistentes con la información que apareció en un informe de *Associated Press* el 15 de enero de 2003). ¿Estos datos sugieren fuertemente que un currículum con un nombre de "negro" tiene menos probabilidades de obtener una respuesta que uno con nombre de "blanco"?

90. La prueba de McNemar, desarrollada en el ejercicio 54, también se puede emplear cuando se forman parejas de individuos (similares) para obtener n pares, y después se da el tratamiento 1 a un miembro de cada par y al otro se da el tratamiento 2. Luego, X_1 es el número de parejas en las cuales ambos tratamientos fueron éxitos y, de manera similar, para X_2, X_3 y X_4. El estadístico de prueba para probar igual eficacia de los dos tratamientos está dado por $(X_2 - X_3)/\sqrt{(X_2 + X_3)}$, el cual tiene aproximadamente una distribución normal estándar cuando H_0 es verdadera. Utilice esto para probar si el fármaco ergotamina es eficaz en el tratamiento de la migraña.

		Ergotamina	
		S	***F***
Placebo	***S***	44	34
	F	46	30

Los datos son ficticios, pero la conclusión concuerda con la del artículo "Controlled Clinical Trial of Ergotamine Tartrate" (*British Med. J.*, 1970: 325-327).

91. En el artículo "Evaluating Variability in Filling Operations" (*Food Tech.*, 1984: 51-55) se describen dos operaciones de llenado empleadas en una planta empacadora de carne molida. Ambas operaciones de llenado se fijaron para llenar paquetes con 1400 g de carne molida. En una muestra aleatoria de tamaño 30 tomada de cada operación de llenado, las medias resultantes y desviaciones estándar fueron 1402.24 g y 10.97 g para la operación 1 y 1419.63 g y 9.96 g para la operación 2.

a. Por medio de un nivel de significación 0.05, ¿hay evidencia suficiente para indicar que el peso medio real de los empaques difiere para las dos operaciones?

b. ¿Los datos de la operación 1 indican que el peso medio real de empaques producidos en la operación 1 es mayor que 1400 g? Utilice un nivel de significación de 0.05.

92. Sea $X_1, \ldots, X_m$ una muestra aleatoria de una distribución de Poisson con parámetro λ_1 y sea $Y_1, \ldots, Y_n$ una muestra aleatoria de otra distribución de Poisson con parámetro λ_2. Se desea probar H_0: $\lambda_1 - \lambda_2 = 0$ contra una de las tres alternativas estándar. Puesto que $\mu = \lambda$ para una distribución de Poisson, cuando m y n son grandes, se puede utilizar la prueba z con muestras grandes de la sección 9.1. Sin embargo, el hecho de que $V(\overline{X}) = \lambda/n$ hace pensar que se debe usar un denominador diferente al estandarizar $\overline{X} - \overline{Y}$. Desarrolle un procedimiento de prueba con muestras grandes apropiado para este problema y, luego, aplíquelo a los siguientes datos para probar si las densidades de planta de una especie en particular son iguales en dos regiones diferentes (donde cada observación es el número de plantas encontradas en un cuadro de muestreo localizado al azar que tiene un área de 1 m^2, de modo que para la región 1, habría 40 cuadros donde se observó una planta, etcétera).

	Frecuencia								
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	
Región 1	28	40	28	17	8	2	1	1	$m = 125$
Región 2	14	25	30	18	49	2	1	1	$n = 140$

93. Refiérase al ejercicio 92, obtenga una fórmula del intervalo de confianza con muestras grandes para $\lambda_1 - \lambda_2$. Calcule el intervalo para los datos + proporcionados utilizando un nivel de confianza de 95%.

Bibliografía

Véase la bibliografía al final del capítulo 7.

10

Análisis de varianza

Introducción

Al estudiar métodos para el análisis de datos cuantitativos, primero se atienden los problemas relacionados con una sola muestra de números y luego se vuelve a un análisis comparativo de dos muestras de esta clase distintas. En los problemas con una sola muestra, los datos consistieron en observaciones acerca de las respuestas de individuos u objetos experimentales seleccionados al azar de una sola población. En los problemas con dos muestras, éstas se extrajeron de dos poblaciones diferentes y los parámetros de interés fueron las medias poblacionales, o bien se aplicaron dos tratamientos distintos a las unidades experimentales (individuos u objetos) seleccionados de una sola población; en este último caso, los parámetros de interés reciben el nombre de medias de tratamiento verdadero.

El **análisis de varianza**, o de manera breve **ANOVA**, se refiere en términos generales a una colección de situaciones experimentales y procedimientos estadísticos para el análisis de respuestas cuantitativas de unidades experimentales. El problema ANOVA más simple se conoce como **ANOVA de un solo factor, de una sola clasificación o unidireccional** y tiene que ver con el análisis de datos muestreados de más de dos poblaciones numéricas (distribuciones) o de datos de experimentos en los que se han usado más de dos tratamientos. La característica que distingue a los tratamientos o poblaciones entre sí se conoce como el **factor** bajo estudio, y los diferentes tratamientos o poblaciones se conocen como **niveles** del factor. Ejemplos de este tipo de situaciones son los siguientes:

1. Un experimento para estudiar los efectos de cinco marcas de gasolina en el desempeño de un motor de automóvil (mpg).
2. Un experimento para estudiar los efectos de la presencia de cuatro soluciones de diferentes azúcares (glucosa, sacarosa, fructosa y una mezcla de los tres) en el desarrollo bacteriano.
3. Un experimento para investigar si la concentración de madera dura en la pulpa (%) tiene un efecto en la resistencia a la tensión de bolsas hechas de pulpa
4. Un experimento para decidir si la densidad de color de especímenes de tela depende de la cantidad del colorante utilizado.

En (1) el factor de interés en la marca de gasolina y hay cinco niveles del factor. En (2) el factor es azúcar, con cuatro niveles (o cinco, si se emplea una solución de control libre de azúcar). En (1) y (2), el factor es de naturaleza cuantitativa, y los niveles corresponden a las posibles categorías del factor. En (3) y (4), los factores son la concentración de la madera dura y la cantidad de colorante, respectivamente; ambos factores son de naturaleza cuantitativa, así que los niveles identifican entornos diferentes del factor. Cuando el factor de interés es cuantitativo, las técnicas estadísticas del análisis de regresión (analizadas en los capítulos 12 y 13) también se pueden usar para analizar los datos.

Este capítulo se centra en el ANOVA de un solo factor. En la sección 10.1 se presenta la prueba F para probar la hipótesis nula de que las medias poblacionales o de tratamiento son idénticas. En la sección 10.2 se considera un análisis ulterior de los datos cuando ha sido rechazada H_0. En la sección 10.3 se abarcan otros aspectos del ANOVA de un solo factor. En el capítulo 11 se presentan experimentos de ANOVA que tienen que ver con más de un factor.

10.1 ANOVA de un solo factor

El ANOVA de un solo factor se centra en la comparación de más de dos medias poblacionales o de tratamiento. Sea

I = número de poblaciones o tratamientos que son comparados.

μ_1 = media de la población 1 o la respuesta promedio real cuando se aplica el tratamiento 1.

$\vdots$

μ_I = media de la población I o la respuesta promedio real cuando se aplica el tratamiento I.

Las hipótesis de interés son

$$H_0: \ \mu_1 = \mu_2 = \cdots = \mu_I$$

contra

$$H_a: \ \text{por los menos dos de las } \mu_i \text{ son diferentes}$$

Si $I = 4$, H_0 es verdadera sólo si las cuatro μ_i son idénticas. H_a sería verdadera, por ejemplo, si $\mu_1 = \mu_2 \neq \mu_3 = \mu_4$, si $\mu_1 = \mu_3 = \mu_4 \neq \mu_2$, o si las cuatro μ_i difieren entre sí.

Una prueba de estas hipótesis requiere que se tenga disponible una muestra aleatoria de cada población o tratamiento.

Ejemplo 10.1 En el artículo "Compression of Single –Wall Corrugated Shipping Containers Using Fixed and Floating Test Platens" (*J. Testing and Evaluation*, 1992: 318-320) se describe un experimento en el que se comparan varios tipos de cajas con respecto a la resistencia a la compresión (lb). En la tabla 10.1 se presentan los resultados de un experimento ANOVA de un solo factor relacionado con $I = 4$ tipos de cajas (las medias muestrales y las desviaciones estándar concuerdan bien con las proporcionadas en el artículo).

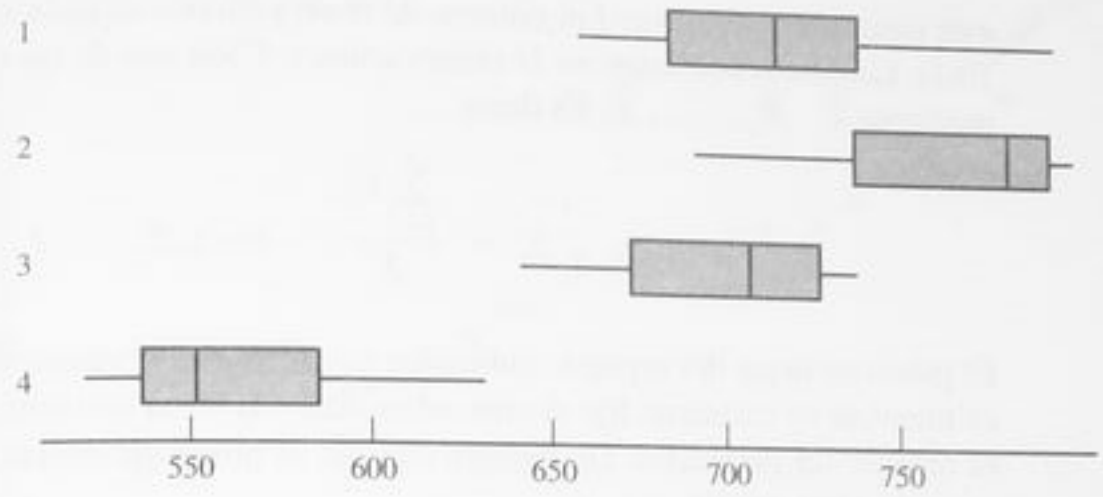
Tabla 10.1 Datos y cantidades de resumen para el ejemplo 10.1

Tipo de caja	Resistencia a la compresión (lb)						Media muestral	DE de la muestra
1	655.5	788.3	734.3	721.4	679.1	699.4	713.00	46.55
2	789.2	772.5	786.9	686.1	732.1	774.8	756.93	40.34
3	737.1	639.0	696.3	671.7	717.2	727.1	698.07	37.20
4	535.1	628.7	542.4	559.0	586.9	520.0	562.02	39.87
						Gran media =	682.50	

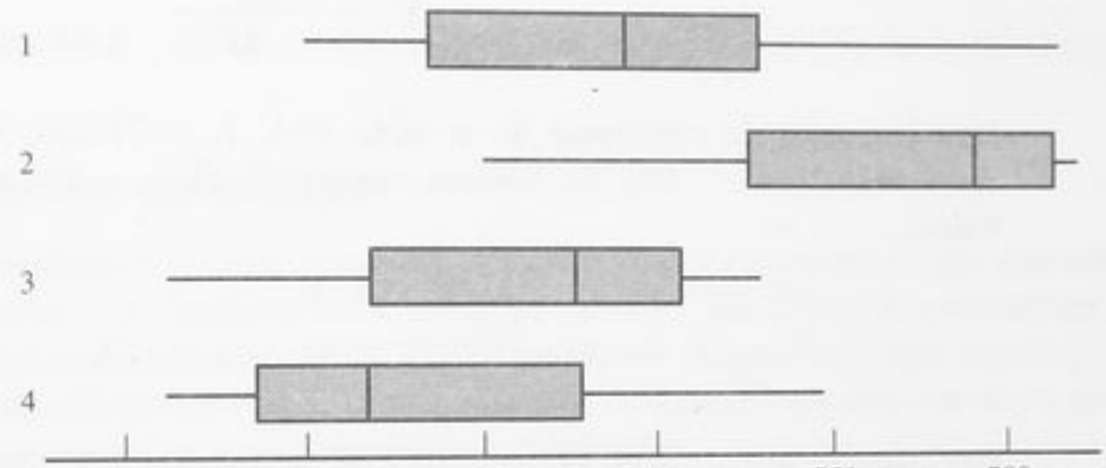
Con μ_i que denota la resistencia promedio real a la compresión para cajas de tipo i ($i = 1, 2, 3, 4$), la hipótesis nula es H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$. En la figura 10.1(a) se muestra un diagrama de caja comparativo para las cuatro muestras. Hay una cantidad sustancial de traslape entre las observaciones en los tres primeros tipos de cajas, pero las resistencias a la compresión para el cuarto tipo parecen considerablemente más pequeñas que para los otros tipos. Esto indica que H_0 no es verdadera. La gráfica comparativa de la figura 10.1(b) se basa en añadir 120 a cada observación en la cuarta muestra (con lo que se obtiene la media 682.02 y la misma desviación estándar) y se dejan sin cambio las otras observaciones. Ya no es obvio si H_0 es verdadera o falsa. En situaciones como ésta, se necesita un procedimiento de prueba formal. ■

Notación y suposiciones

En problemas con dos muestras, se utilizaron las letras X y Y para diferenciar las observaciones de una muestra de las de la otra. Debido a que esto es difícil para tres o más muestras, se acostumbra usar una sola letra con dos subíndices. El primer subíndice identifica el número de muestra, que corresponde a la población o tratamiento que se muestrea, y el segundo



(a)



(b)

Figura 10.1 Diagramas de caja para el ejemplo 10.1: (a) datos originales; (b) datos modificados

subíndice denota la posición de la observación dentro de esa muestra. Sea

X_{ij} = variable aleatoria (*va*) que denota la j-ésima medición tomada de la i-ésima población, o la medición tomada de la j-ésima unidad experimental que recibe el i-ésimo tratamiento

x_{ij} = valor observado de X_{ij} cuando se realiza el experimento

Los datos observados por lo general se muestran en una tabla rectangular, como la tabla 10.1. Tres muestras de diferentes poblaciones aparecen en distintos renglones de la tabla, y x_{ij} es el j-ésimo número en el i-ésimo renglón. Por ejemplo, $x_{2,3} = 786.9$ (la tercera observación a partir de la segunda población), y $x_{4,1} = 535.1$. Cuando no hay ambigüedad, se escribe x_{ij} en vez de x_{ij} (p. ej., si hubiera 15 observaciones en cada uno de 12 tratamientos, x_{112} podría significar $x_{1,12}$ o $x_{11,2}$). Se supone que las X_{ij} dentro de alguna muestra particular son independientes —una muestra aleatoria de la i-ésima distribución de población o tratamiento— y que las diferentes muestras son independientes entre sí.

En algunos experimentos, muestras diferentes contienen distintos números de observaciones. Sin embargo, los conceptos y métodos del ANOVA de un solo factor se desarrollan con más facilidad para el caso de tamaños de muestra iguales. Si por el momento sólo se atiende

este caso, represente con J el número de observaciones en cada muestra ($J = 6$ en el ejemplo 10.1). Los datos consisten en IJ observaciones. Cada una de las medias muestrales se denota mediante $\overline{X}_{1\cdot}, \overline{X}_{2\cdot}, \ldots, \overline{X}_{I\cdot}$. Es decir,

$$\overline{X}_{i\cdot} = \frac{\sum_{j=1}^{J} X_{ij}}{J} \qquad i = 1, 2, \ldots, I$$

El punto en lugar del segundo subíndice significa que se sumaron los valores de ese subíndice mientras se mantiene fijo el otro subíndice, y la barra horizontal indica división entre J para obtener un promedio. De manera similar, el promedio de las IJ observaciones, conocido como la **gran media**, es

$$\overline{X}_{\cdot\cdot} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}}{IJ}$$

Para los datos de resistencia de la tabla 10.1, $\overline{x}_{1\cdot} = 713.00$, $\overline{x}_{2\cdot} = 756.93$, $\overline{x}_{3\cdot} = 698.07$, $\overline{x}_{4\cdot} = 562.02$ y $\overline{x}_{\cdot\cdot} = 682.50$. Además, sea que $S_1^2, S_2^2, \ldots, S_I^2$ represente las varianzas muestrales:

$$S_i^2 = \frac{\sum_{j=1}^{J} (X_{ij} - \overline{X}_{i\cdot})^2}{J - 1} \qquad i = 1, 2, \ldots, I$$

Del ejemplo 10.1, $s_1 = 46.55$, $s_1^2 = 2166.90$, etcétera.

SUPOSICIONES

Las I distribuciones de población o tratamiento son normales con la misma varianza. Es decir, cada X_{ij} tiene una distribución normal con

$$E(X_{ij}) = \mu_i \qquad V(X_{ij}) = \sigma^2$$

Las I desviaciones estándar muestrales por lo general diferirán un poco aun cuando las σ correspondientes sean idénticas. En el ejemplo 10.1, la más grande entre s_1, s_2, s_3 y s_4 es alrededor de 1.25 veces la más pequeña. Una regla empírica es que si la s más grande no es mucho más que dos veces la más pequeña, es razonable suponer que las σ^2 son iguales.

En los capítulos anteriores, se sugirió una gráfica de probabilidad para comprobar la normalidad. Cada uno de los tamaños de muestra en el ANOVA son por lo general demasiado pequeños para que las I gráficas separadas sean informativas. Una sola gráfica se puede construir al restar $\overline{x}_i$ de cada observación en la primera muestra, $\overline{x}_2$ en la segunda, etcétera, y luego graficar estas IJ observaciones contra los percentiles z. En la figura 10.2 se proporciona la gráfica para los datos del ejemplo 10.1. La linealidad del patrón proporciona una prueba convincente de la suposición de normalidad.

Si se juzga imposible la suposición de normalidad o la de varianzas iguales, se debe usar otro método de análisis que no sea la prueba usual F. En este tipo de situaciones consulte a un experto (en la sección 10.3 se sugiere una posibilidad, una transformación de datos).

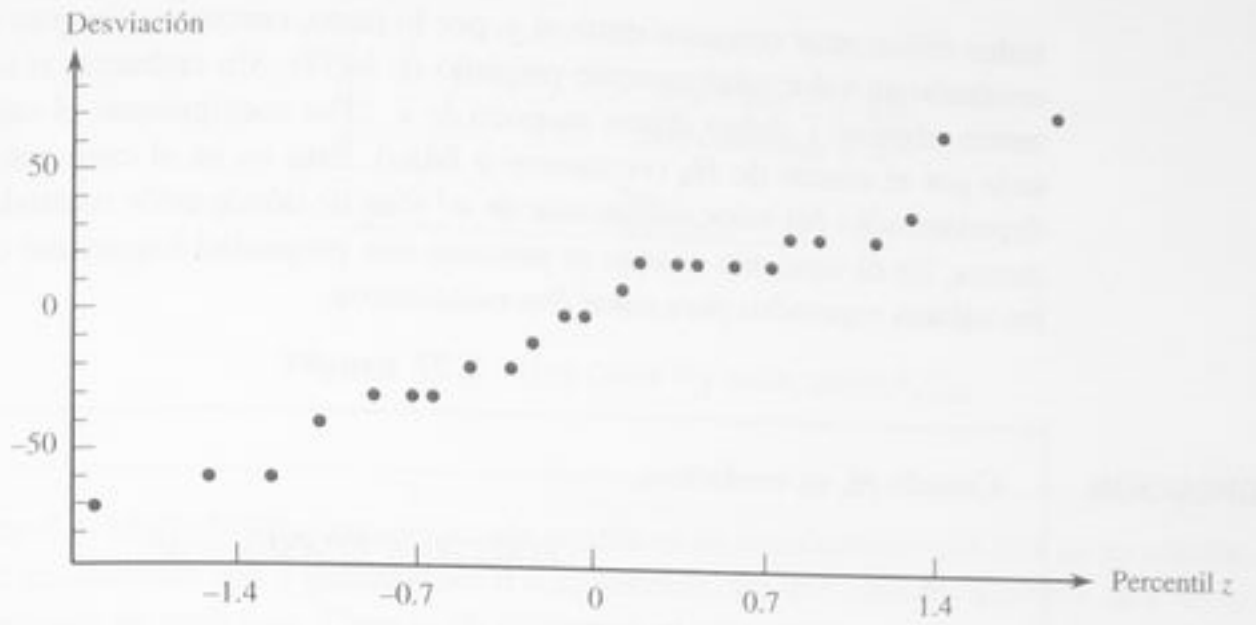


Figura 10.2 Gráfica de probabilidad normal con base en los datos del ejemplo 10.1

Estadístico de prueba

Si H_0 es verdadera, las J observaciones de cada muestra provienen de una distribución poblacional normal con el mismo valor medio μ, en cuyo caso las medias muestrales $\bar{x}_{1\cdot}, \ldots, \bar{x}_{I\cdot}$ deben estar razonablemente cerca. El procedimiento de prueba se basa en comparar una medida de diferencias entre las $\bar{x}_{i\cdot}$ (variación "entre muestras") con respecto a una medida de variación calculada *dentro* de cada una de las muestras.

DEFINICIÓN

El **cuadrado medio para tratamientos** está dado por

$$\text{MSTr} = \frac{J}{I-1}[(\bar{X}_{1\cdot} - \bar{X}_{\cdot\cdot})^2 + (\bar{X}_{2\cdot} - \bar{X}_{\cdot\cdot})^2 + \cdots + (\bar{X}_{I\cdot} - \bar{X}_{\cdot\cdot})^2]$$

$$= \frac{J}{I-1}\sum_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$$

y el **cuadrado medio para el error** es

$$\text{MSE} = \frac{S_1^2 + S_2^2 + \cdots + S_I^2}{I}$$

La estadística de prueba para el ANOVA de un solo factor es $F = \text{MSTr/MSE}$.

En breve se explica la terminología "cuadrado medio". Observe que se utilizan letras mayúsculas $\bar{X}$ y S^2, de modo que MSTr y MSE se definen como estadísticos. Se seguirá esta tendencia y también se usará MSTr y MSE (en lugar de mstr y mse) para denotar los valores calculados de estos estadísticos. Cada S_i^2 evalúa la variación dentro de una muestra particular, de modo que MSE es una medida de variación dentro de las muestras.

¿Qué clase de valor de F proporciona evidencia a favor o en contra de H_0? Observe primero que si H_0 es verdadera (las μ_i son iguales), los valores de cada una de las medias mues-

trales deben estar cercanos entre sí y, por lo tanto, cercanos a la gran media, lo que da como resultado un valor relativamente pequeño de MSTr. Sin embargo, si las μ_i son bastante diferentes, algunas $\overline{x}_i$ deben diferir un poco de $\overline{x}_{\cdot\cdot}$. Por consiguiente, el valor de MSTr se ve afectado por el estado de H_0 (verdadero o falso). Éste no es el caso con el MSE, porque las s_i^2 dependen sólo del valor subyacente de σ^2 y no de dónde estén centradas las distintas distribuciones. En el siguiente cuadro se presenta una propiedad importante de $E(\text{MSTr})$ y $E(\text{MSE})$, los valores esperados para estos dos estadísticos.

PROPOSICIÓN

Cuando H_0 es verdadera,

$$E(\text{MSTr}) = E(\text{MSE}) = \sigma^2$$

mientras que cuando H_0 es falsa,

$$E(\text{MSTr}) > E(\text{MSE}) = \sigma^2$$

Es decir, ambos estadísticos son insesgados para estimar la varianza poblacional común σ^2 cuando H_0 es verdadera, pero MSTr tiende a sobrestimar σ^2 cuando H_0 es falsa.

El insesgo de MSE es una consecuencia de $E(S_i^2) = \sigma^2$ si H_0 es verdadera o falsa. Cuando H_0 es verdadera, cada $\overline{X}_i$ tiene el mismo valor medio μ y varianza σ^2/J, así que $\sum(\overline{X}_i - \overline{X}_{\cdot\cdot})^2/(I-1)$, la "varianza muestral" de las $\overline{X}_i$, estima sin sesgo a σ^2/J; al multiplicar esto por J se obtiene MSTr como estimador insesgado de σ^2. Las $\overline{X}_i$ tienden a dispersarse más cuando H_0 es falsa que cuando es verdadera, lo cual en este caso tiende a inflar el valor de MSTr. Así, un valor de F que sobrepase por mucho a 1, lo que corresponde a un MSTr mucho mayor que MSE, arroja dudas considerables acerca de H_0. La forma de la región de rechazo es $f \geq c$. El corte c se debe elegir para obtener $P(F \geq c$ cuando H_0 es verdadera$) = \alpha$, el nivel de significación deseado. Por lo tanto, cuando H_0 es verdadera se requiere la distribución de F.

Distribuciones *F* y la prueba *F*

En el capítulo 9, se presentó una familia de distribuciones de probabilidad denominadas distribuciones F. Una distribución F surge debido a una relación en la que hay un número de grados de libertad (gl) para el numerador y otro número de grados de libertad para el denominador. Sean ν_1 y ν_2 el número de grados de libertad del numerador y el denominador, respectivamente, para una variable con una distribución F. Tanto ν_1 como ν_2 son enteros positivos. En la figura 10.3 se ilustra una curva de densidad F y el valor crítico de extremo superior F_{α,ν_1,ν_2} correspondiente. En la tabla A.9 del apéndice se dan estos valores críticos para $\alpha = 0.10, 0.05, 0.01$ y 0.001. Los valores de ν_1 se identifican con diferentes columnas de la tabla, y los renglones se marcan con varios valores de ν_2. Por ejemplo, el valor crítico F que capta el área del extremo superior 0.05 bajo la curva F con $\nu_1 = 4$ y $\nu_2 = 6$ es $F_{0.05,4,6} = 4.53$, en tanto que $F_{0.05,6,4} = 6.16$. El resultado teórico es que el estadístico de prueba F tiene una distribución F cuando H_0 es verdadera.

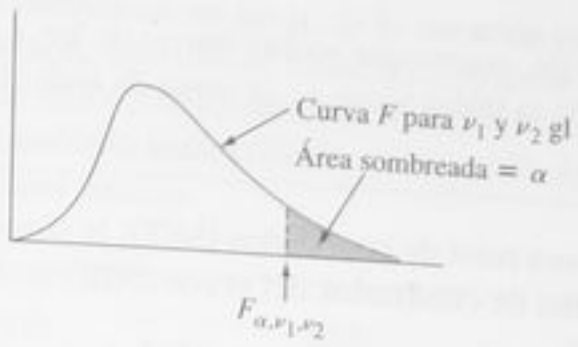


Figura 10.3 Una curva F y valor crítico F_{α,ν_1,ν_2}

TEOREMA

Sea $F = \text{MSTr/MSE}$ el estadístico de prueba en un problema de ANOVA de un solo factor relacionado con I poblaciones o tratamientos con una muestra aleatoria de J observaciones de cada uno. Cuando H_0 es verdadera y se satisfacen las suposiciones básicas de esta sección, F tiene una distribución F con $\nu_1 = I - 1$ y $\nu_2 = I(J - 1)$. Si f denota el valor calculado de F, entonces la región de rechazo $f \geq F_{\alpha,I-1,I(J-1)}$ especifica una prueba con nivel de significación α. Refiérase a la sección 9.5 para ver cómo se obtiene la información del valor P para pruebas F.

El fundamento para $\nu_1 = I - 1$ es que aunque MSTr se basa en las I desviaciones $\overline{X}_{1\cdot} - \overline{X}_{\cdot\cdot}, \ldots, \overline{X}_{I\cdot} - \overline{X}_{\cdot\cdot}, \sum(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot}) = 0$, así que sólo $I - 1$ de éstas se determinan sin restricciones. Debido a que cada muestra contribuye con $J - 1$ gl al MSE y estas muestras son independientes, $\nu_2 = (J - 1) + \cdots + (J - 1) = I(J - 1)$.

Ejemplo 10.2 (Continuación del ejemplo 10.1)

Los valores de I y J para los datos de resistencia son 4 y 6, respectivamente, así que gl del numerador $= I - 1 = 3$ y gl del denominador $= I(J - 1) = 20$. Al nivel de significación 0.05, H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ se rechazará en favor de la conclusión de que por lo menos dos μ_i son diferentes si $f \geq F_{0.05,3,20} = 3.10$. La gran media es $\bar{x}_{\cdot\cdot} = \sum\sum x_{ij}/(IJ) = 682.50$

$$\text{MSTr} = \frac{6}{4-1}[(713.00 - 682.50)^2 + (756.93 - 682.50)^2 + (698.07 - 682.50)^2 + (562.02 - 682.50)^2] = 42\,455.86$$

$$\text{MSE} = \frac{1}{4}[(46.55)^2 + (40.34)^2 + (37.20)^2 + (39.87)^2] = 1691.92$$

$$f = \text{MSTr/MSE} = 42\,455.86/1691.92 = 25.09$$

Puesto que $25.09 \geq 3.10$, se rechaza H_0 al nivel de significación 0.05. Al parecer la resistencia promedio real a la compresión depende del tipo de caja. De hecho, el valor P = área bajo la curva F a la derecha de $25.09 = 0.000$, por lo tanto se debe rechazar H_0 a cualquier nivel de significación razonable. ■

Fórmulas de cálculo para el ANOVA

Los cálculos que dan lugar a f se pueden hacer de manera eficaz por medio de fórmulas similares a la fórmula de cálculo para una sola varianza muestral s^2. Represente con x_i la *suma* (no

el promedio, puesto que no hay barra) de las x_{ij} para i fija (suma de los números del i-ésimo renglón de la tabla) y con $x_{..}$ la suma de *todas* las x_{ij} (el **gran total**).

DEFINICIÓN

La suma total de cuadrados (SST), la suma de cuadrados del tratamiento (SSTr) y la suma de cuadrados del error (SSE) se determinan mediante

$$\text{SST} = \sum_{i=1}^{I}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}x_{ij}^2 - \frac{1}{IJ}x_{..}^2$$

$$\text{SSTr} = \sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{x}_{i.} - \bar{x}_{..})^2 = \frac{1}{J}\sum_{i=1}^{I}x_{i.}^2 - \frac{1}{IJ}x_{..}^2$$

$$\text{SSE} = \sum_{i=1}^{I}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{i.})^2 \qquad \text{donde } x_{i.} = \sum_{j=1}^{J}x_{ij} \quad x_{..} = \sum_{i=1}^{I}\sum_{j=1}^{J}x_{ij}$$

La suma de cuadrados SSTr aparece en el numerador de F y SSE aparece en el denominador de F; en breve será evidente la razón para definir SST.

Las expresiones en el lado derecho extremo de SST y SSTr son las fórmulas de cálculo para estas sumas de cuadrados. Tanto SST como SSTr tienen que ver con $x_{..}^2/(IJ)$ (el cuadrado del gran total dividido entre IJ), que por lo general se llama **factor de corrección para la media** (CF). Después que se calcula el factor de corrección, se obtiene SST al elevar al cuadrado cada número de la tabla de datos, sumar estos cuadrados y restar el factor de corrección. SSTr resulta de elevar al cuadrado cada total de renglón, sumarlos, dividir entre J y restar el factor de corrección.

Una fórmula de cálculo para SSE es una consecuencia de una relación simple entre las tres sumas de cuadrados.

Identidad fundamental

$$\text{SST} = \text{SSTr} + \text{SSE} \tag{10.1}$$

Por consiguiente, si se calculan dos de las sumas de cuadrados, la tercera se puede obtener por medio de la ecuación (10.1); SST y SSTr son fáciles de calcular y entonces SSE = SST − SSTr. La demostración se deduce de elevar al cuadrado ambos lados de la relación

$$x_{ij} - \bar{x}_{..} = (x_{ij} - \bar{x}_{i.}) + (\bar{x}_{i.} - \bar{x}_{..}) \tag{10.2}$$

y sumar las i y las j. Con esto se obtiene SST a la izquierda y SSTr y SSE como los dos términos extremos a la derecha. Se ve con facilidad que el término de producto cruzado es cero.

La interpretación de la identidad fundamental es una ayuda importante para comprender el ANOVA. SST es una medida de la variación total en los datos, la suma de las desviaciones cuadradas respecto a la gran media. La identidad establece que esta variación total se puede dividir en dos partes. SSE mide la variación que se presentaría (dentro de los renglones) incluso si H_0 fuera verdadera y, por lo tanto, es la parte de la variación total que se *explica* mediante la veracidad o falsedad de H_0. SSTr es la cantidad de variación (entre renglones) que se

Ejemplo

puede explicar por posibles diferencias en las μ_i. Si la variación explicada es grande con respecto a la variación no explicada, entonces se rechaza H_0 en favor de H_a.

Una vez que se calcula SSTr y SSE, cada una se divide entre sus grados de libertad relacionados para obtener un cuadrado medio (medio en el sentido de promedio). Entonces F es la relación de los dos cuadrados medios.

$$\text{MSTr} = \frac{\text{SSTr}}{I-1} \qquad \text{MSE} = \frac{\text{SSE}}{I(J-1)} \qquad F = \frac{\text{MSTr}}{\text{MSE}} \qquad (10.3)$$

Los cálculos de resumen en forma tabular, llamada **tabla de ANOVA**, como se muestra en la tabla 10.2. Por lo común las tablas obtenidas por medio de software estadístico incluyen una columna del valor P a la derecha de f.

Tabla 10.2 Tabla de ANOVA

Fuente de variación	gl	Suma de cuadrados	Cuadro medio	f
Tratamientos	$I-1$	SSTr	MSTr = SSTr/$(I-1)$	MSTr/MSE
Error	$I(J-1)$	SSE	MSE = SSE/$[I(J-1)]$	
Total	$IJ-1$	SST		

Ejemplo 10.3 Los siguientes datos se obtuvieron de un experimento en el que se compara el grado de ensuciamiento para tela copolimerizada con tres diferentes mezclas de ácido metacrílico (datos similares aparecen en el artículo "Chemical Factors Affecting Soiling and Soil Release from Cotton DP Fabric", *American Dyestuff Reporter*, 1983: 25-30).

						$x_{i\cdot}$	$\bar{x}_{i\cdot}$
Mezcla 1	0.56	1.12	0.90	1.07	0.94	4.59	0.918
Mezcla 2	0.72	0.69	0.87	0.78	0.91	3.97	0.794
Mezcla 3	0.62	1.08	1.07	0.99	0.93	4.69	0.938
						$x_{\cdot\cdot} = 13.25$	

Represente con μ_i el grado promedio de ensuciamiento cuando se usa la mezcla i ($i = 1, 2, 3$). La hipótesis nula H_0: $\mu_1 = \mu_2 = \mu_3$ expresa que el grado promedio real de ensuciamiento es idéntico para las tres mezclas. Se realizará una prueba al nivel de significación 0.01 para ver si se debe rechazar H_0 en favor de la afirmación del grado promedio de ensuciamiento no es el mismo para las tres mezclas. Puesto que $I - 1 = 2$ y $I(J-1) = 12$, el valor crítico F para la región de rechazo es $F_{0.01,2,12} = 6.93$. Al elevar al cuadrado cada una de las 15 observaciones y sumar se obtiene $\sum\sum x_{ij}^2 = (0.56)^2 + (1.12)^2 + \cdots + (0.93)^2 = 12.1351$. Los valores de las tres sumas de cuadrados son

$$\text{SST} = 12.1351 - (13.25)^2/15 = 12.1351 - 11.7042 = 0.4309$$

$$\text{SSTr} = \frac{1}{5}[(4.59)^2 + (3.97)^2 + (4.69)^2] - 11.7042$$

$$= 11.7650 - 11.7042 = 0.0608$$

$$\text{SSE} = 0.4309 - 0.0608 = 0.3701$$

Los cálculos restantes se resumen en la tabla de ANOVA adjunta. Debido a que $f = 0.99$ no es por lo menos $F_{0.01,2,12} = 6.93$, no se rechaza H_0 al nivel de significación 0.01. Las mezclas al parecer son indistinguibles con respecto al grado de ensuciamiento ($F_{0.10,2,12} = 2.81 \Rightarrow$ valor $P > 0.10$).

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
Tratamientos	2	0.0608	0.0304	0.99
Error	12	0.3701	0.0308	
Total	14	0.4309		

Cuando la prueba F causa que se rechace H_0, con frecuencia el experimentador estará interesado en un análisis más profundo a fin de decidir cuáles μ_i difieren de cuales otras. Los procedimientos para hacer esto se llaman procedimientos de comparación múltiples, y se describen varios en las dos secciones siguientes.

Ejercicios | Sección 10.1 (1 a 10)

1. En un experimento para comparar las resistencias a la tensión de $I = 5$ tipos diferentes de alambre de cobre, se usaron $J = 4$ muestras de cada tipo. Las estimaciones de σ^2 entre muestras y dentro de muestras se calcularon como MSTr = 2673.3 y MSE = 1094.2, respectivamente.
   a. Utilice la prueba F al nivel 0.05 para probar H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ contra H_a: por lo menos dos μ_i no son iguales.
   b. ¿Qué se puede decir acerca del valor P para la prueba?
2. Suponga que las observaciones de resistencia a la compresión en el cuarto tipo de caja del ejemplo 10.1 hubieran sido 655.1, 748.7, 662.4, 679.0, 706.9 y 640.0 (obtenidos al sumar 120 a cada x_{4j} previa). Suponiendo que no hay cambio en las demás observaciones, realice una prueba F con $\alpha = 0.05$.
3. Se determinó la salida de luz para cada una de $I = 3$ marcas diferentes de focos de luz blanca de 60 watts, con $J = 8$ focos probados de cada marca. Las sumas de cuadrados se calcularon como SSE = 4773.3 y SSTr = 591.2. Exprese las hipótesis de interés (incluso las definiciones de los parámetros con palabras), y utilice la prueba F de ANOVA ($\alpha = 0.05$) para decidir si hay algunas diferencias en las salidas de luz promedio reales entre las tres marcas para este tipo de foco al calcular tanta información como sea posible acerca del valor P.
4. En un estudio para evaluar los efectos de la infección por malaria en mosquitos huéspedes ("Plasmodium Cynomolgi: Effects of Malaria Infection on Laboratory Flight Performance of Anopheles Stephensi Mosquitos", *Experimental Parasitology*, 1977: 397-404), se alimentó a los mosquitos en monos rhesus infecciosos y no infecciosos. Después se midió la distancia que volaron durante un periodo de 24 horas por medio de un dispositivo de vuelo. Se dividió a los mosquitos en cuatro grupos de ocho mosquitos cada uno: rhesus infeccioso y esporozoítos presentes (IRS), rhesus infeccioso y oocistos presentes (IRD), rhesus infeccioso y ninguna infección desarrollada (IRN) y no infeccioso (C). Los valores de los datos de resumen son $\bar{x}_{1\cdot} = 4.39$ (IRS), $\bar{x}_{2\cdot} = 4.52$ (IRD), $\bar{x}_{3\cdot} = 5.49$ (IRN), $\bar{x}_{4\cdot} = 6.36$ (C), $\bar{x}_{\cdot\cdot} = 5.19$, y $\sum\sum x_{ij}^2 = 911.91$. Utilice la prueba F de ANOVA al nivel 0.05 para decidir si hay algunas diferencias entre los tiempos de vuelo promedio reales para los cuatro tratamientos.
5. Considere los siguientes datos de resumen acerca del módulo de elasticidad ($\times 10^6$ lb/pulg^2) para madera de tres grados distintos (en estrecha concordancia con los valores del artículo "Bending Strength and Stiffness of Second-Growth Douglas Fir Dimension Lumber" (*Forest Products J.*, 1991: 35-43), excepto que los tamaños de las muestras fueron más grandes):

Grado	J	$\bar{x}_{i\cdot}$	s_i
1	10	1.63	0.27
2	10	1.56	0.24
3	10	1.42	0.26

Utilice estos datos y un nivel de significación de 0.01 para probar la hipótesis nula de ninguna diferencia en el módulo medio de elasticidad para los tres grados.

6. En el artículo "Origin of Precambrian Iron Formations" (*Econ. Geology*, 1964: 1025-1057) aparecen los datos siguientes acerca del Fe total para cuatro tipos de formación de hierro (1 = carbonato, 2 = silicato, 3 = manetita, 4 = hematita).

1: 20.5 28.1 27.8 27.0 28.0
25.2 25.3 27.1 20.5 31.3

2: 26.3 24.0 26.2 20.2 23.7
34.0 17.1 26.8 23.7 24.9

3: 29.5 34.0 27.5 29.4 27.9
26.2 29.9 29.5 30.0 35.6

4: 36.5 44.2 34.1 30.3 31.4
33.1 34.1 32.9 36.3 25.5

Realice un análisis de la prueba de la variación F en el nivel 0.01 de la significación y resuma los resultados en una tabla ANOVA .

7. En un experimento para investigar el desempeño de cuatro marcas de bujías para motocicletas de dos tiempos de 125 cc, se probaron cinco bujías de cada marca y se observó el número de millas (a una velocidad constante). A continuación se proporciona la tabla de ANOVA parcial para los datos. Complete los elementos faltantes, exprese las hipótesis pertinentes y lleve a cabo una prueba al obtener tanta información como sea posible acerca del valor P.

Fuente	gl	Suma de cuadrados	Media cuadrada	f
Marca				
Error			14 713.69	
Total		310 500.76		

8. En un estudio de las propiedades de vigas unidas a placas metálicas utilizadas para soportar techos ("Modeling Joints Made with Light-Gauge Metal Connector Plates", *Forest Products J.*, 1979: 39-44) se hicieron las siguientes observaciones en relación con el índice de rigidez axial (klb/pulg) para longitudes de placa 4, 6, 8, 10 y 12 pulgadas:

4: 309.2 409.5 311.0 326.5 316.8
349.8 309.7

6: 402.1 347.2 361.0 404.5 331.0
348.9 381.7

8: 392.4 366.2 351.0 357.1 409.9
367.3 382.0

10: 346.7 452.9 461.4 433.1 410.6
384.2 362.6

12: 407.4 441.8 419.9 410.7 473.4
441.2 465.8

¿La variación de la longitud de placa tiene algún efecto en la rigidez axial promedio real? Exprese y pruebe las hipótesis pertinentes por medio del análisis de varianza con $\alpha = 0.01$. Muestre sus resultados en una tabla de ANOVA. (*Sugerencia:* $\sum\sum x_{ij}^2 = 5\,241\,420.79$.)

9. Se analizaron seis muestras de cada uno de cuatro tipos de granos de cereal cultivados en cierta región para determinar el contenido de tiamina, y se obtuvieron los siguientes datos (μg/g):

Trigo	5.2	4.5	6.0	6.1	6.7	5.8
Cebada	6.5	8.0	6.1	7.5	5.9	5.6
Maiz	5.8	4.7	6.4	4.9	6.0	5.2
Avena	8.3	6.1	7.8	7.0	5.5	7.2

¿Estos datos indican que por lo menos dos de los granos difieren con respecto al contenido promedio real de tiamina? Utilice una prueba de nivel $\alpha = 0.05$ con base en el método del valor P.

10. En el ANOVA de un solo factor con I tratamientos y J observaciones por tratamiento, sea $\mu = (1/I)\sum\mu_i$.
    a. Exprese $E(\bar{X}_{\cdot\cdot})$ en términos de μ. [*Sugerencia*: $\bar{X}_{\cdot\cdot} = (1/I)\sum\bar{X}_{i\cdot}$]
    b. Calcule $E(\bar{X}_{i\cdot}^2)$. [*Sugerencia*: para cualquier *va* Y, $E(Y^2) = V(Y) + [E(Y)]^2$.]
    c. Calcule $E(\bar{X}_{\cdot\cdot}^2)$.
    d. Calcule $E(\text{SSTr})$ y luego muestre que

$$E(\text{MSTr}) = \sigma^2 + \frac{J}{I-1}\sum(\mu_i - \mu)^2$$

    e. Con el resultado del inciso (d), ¿cuál es $E(\text{MSTr})$ cuando H_0 es verdadera? Cuando H_0 es falsa, ¿cómo se compara $E(\text{MSTr})$ con σ^2?

10.2 Comparaciones múltiples en el ANOVA

Cuando no es importante el valor calculado de la estadística F en el ANOVA de un solo factor, se termina el análisis porque no han sido identificadas diferencias entre las μ_i. Pero cuando se rechaza H_0, el investigador por lo común desea saber cuáles de las μ_i son diferentes entre sí. Un método para llevar a cabo este posible análisis se llama **procedimiento de comparaciones múltiples.**

Varios de los procedimientos de este tipo utilizados con más frecuencia se basan en la siguiente idea central. Primero calcule un intervalo de confianza para cada diferencia por pares $\mu_i - \mu_j$ con $i < j$. Por lo tanto, si $I = 4$ los seis IC requeridos serían para $\mu_1 - \mu_2$ (pero no también para $\mu_2 - \mu_1$), $\mu_1 - \mu_3$, $\mu_1 - \mu_4$, $\mu_2 - \mu_3$, $\mu_2 - \mu_4$ y $\mu_3 - \mu_4$. Entonces si el intervalo para $\mu_1 - \mu_2$ no incluye el 0, se concluye que μ_1 y μ_2 *difieren de forma importante* entre sí; si el intervalo incluye al 0, se considera que las dos μ no son significativamente diferentes. Continuando con la misma línea de razonamiento para cada uno de los otros intervalos, se termina siendo capaz de juzgar para cada par de las μ si difieren o no entre sí de forma importante.

Los procedimientos basados en esta idea difieren en el método utilizado para calcular los diversos intervalos de confianza. A continuación se presenta un método popular que controla el nivel de confianza *simultáneo* para los $I(I-1)/2$ intervalos calculados.

Procedimiento de Tukey (método *T*)

El procedimiento de Tukey tiene que ver con el uso de otra distribución de probabilidad llamada **Distribución de intervalo estudentizado.** La distribución depende de dos parámetros: m y ν. los grados de libertad del numerador y el denominador, respectivamente. Sea $Q_{\alpha,m,\nu}$ el valor crítico α del extremo superior de la distribución de intervalo estudentizado con m grados de libertad en el numerador y ν grados de libertad en el denominador (análogo a F_{α,ν_1,ν_2}). Los valores de $Q_{\alpha,m,\nu}$ se dan en la tabla A.10 del apéndice.

PROPOSICIÓN

Con probabilidad $1 - \alpha$,

$$\bar{X}_{i\cdot} - \bar{X}_{j\cdot} - Q_{\alpha,I,I(J-1)}\sqrt{\text{MSE}/J} \le \mu_i - \mu_j \le \bar{X}_{i\cdot} - \bar{X}_{j\cdot} + Q_{\alpha,I,I(J-1)}\sqrt{\text{MSE}/J} \quad (10.4)$$

para toda i y j ($i = 1, \ldots, I$ y $j = 1, \ldots, I$) con $i < j$.

Ejemplo 1

Observe que los gl del numerador para el valor crítico Q_α apropiado es I, el número de medias poblacionales o de tratamiento que se comparan, y no $I - 1$ como en la prueba F. Cuando las $\bar{x}_{i\cdot}$, $\bar{x}_j$ calculadas y el MSE se sustituyen en la ecuación (10.4), el resultado es una colección de intervalos de confianza con nivel de confianza *simultáneo* $100(1 - \alpha)\%$, para las diferencias por pares de la forma $\mu_i - \mu_j$ con $i < j$. Cada intervalo que no incluye a 0 lleva a la conclusión de que los valores correspondientes de μ_i y μ_j difieren entre sí de manera importante.

Puesto que en realidad no se tiene interés en los límites inferior y superior de los diversos intervalos, sino sólo en cuáles incluyen al 0 y cuáles no, se puede evitar gran parte de la aritmética relacionada con (10.4). En el cuadro siguiente se dan detalles y se describe cómo identificar de forma visual las diferencias por medio de un "patrón de subrayado".

Método *T* para identificar las μ_i significativamente diferentes

Seleccione α, extraiga $Q_{\alpha,I,IJ-I)}$ de la tabla A.10 del apéndice y calcule $w = Q_{\alpha,I,IJ-I)} \cdot \sqrt{\text{MSE}/J}$. Luego, liste las medias muestrales en orden creciente y subraye los pares que difieren en menos que w. Cualquier par de medias muestrales no subrayadas por la misma línea, corresponde a un par de medias poblacionales o de tratamiento considerado significativamente distinto.

Suponga, por ejemplo, que $I = 5$ y que

$$\bar{x}_{2\cdot} < \bar{x}_{5\cdot} < \bar{x}_{4\cdot} < \bar{x}_{1\cdot} < \bar{x}_{3\cdot}$$

Entonces

1. Considere primero la media más pequeña $\bar{x}_{2\cdot}$. Si $\bar{x}_{5\cdot} - \bar{x}_{2\cdot} \geq w$, proceda al paso 2. Sin embargo, si $\bar{x}_{5\cdot} - \bar{x}_{2\cdot} < w$, una estas dos primeras medias con un segmento de línea. Luego, si es posible, prolongue este segmento de recta aún más a la derecha hasta la $\bar{x}_i$ más grande que difiera de $\bar{x}_2$ en menos que w (así, la recta podría conectar dos, tres o incluso más medias).

2. Ahora pase a $\bar{x}_5$ y de nuevo prolongue un segmento de recta hasta la $\bar{x}_i$ más grande a su derecha que difiera de $\bar{x}_{5\cdot}$ en menos que w (es posible que no se pueda trazar esta recta, o de otro modo podría subrayar sólo dos medias o tres o incluso las cuatro medias restantes).

3. Continúe moviéndose a $\bar{x}_4$ y proceda como antes y, por último, muévase a $\bar{x}_1$.

En resumen, comenzando de cada media en la lista ordenada, se prolonga un segmento de recta tanto sea posible a la derecha, siempre y cuando la diferencia entre las medias sea más pequeña que w. Se comprueba con facilidad que un intervalo particular de la forma (10.4) contendrá al 0 si y sólo si el par correspondiente de medias muestrales está subrayado con el mismo segmento de recta.

Ejemplo 10.4 Se lleva a cabo un experimento para comparar cinco marcas de filtros de aceite para automóviles con respecto a su capacidad para captar impurezas. Sea μ_i la cantidad promedio real de material que captan los filtros de la marca i ($i = 1, \ldots, 5$) en condiciones controladas. Se utilizó una muestra de nueve filtros de cada marca, y se obtuvieron las siguientes medias muestrales: $\bar{x}_{1\cdot} = 14.5$, $\bar{x}_{2\cdot} = 13.8$, $\bar{x}_{3\cdot} = 13.3$, $\bar{x}_{4\cdot} = 14.3$, y $\bar{x}_{5\cdot} = 13.1$. La tabla 10.3 es la tabla de ANOVA que resume la primera parte del análisis.

Tabla 10.3 Tabla ANOVA para el ejemplo 10.4

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	*f*
Tratamientos (marcas)	4	13.32	3.33	37.84
Error	40	3.53	0.088	
Total	44	16.85		

Puesto que $F_{0.05,4,40} = 2.61$, H_0 se rechaza (de forma decisiva) al nivel 0.05. Ahora se utiliza el procedimiento de Tukey para buscar diferencias importantes entre las μ_i. De la tabla A.10 del apéndice, $Q_{0.05,5,40} = 4.04$ (el segundo subíndice en Q es I y no $I - 1$ como en F), así que $w = 4.04\sqrt{0.088/9} = 0.4$. Después de disponer las cinco medias muestrales en orden creciente, las dos más pequeñas se pueden unir mediante un segmento de recta porque difieren en menos que 0.4. Sin embargo, este segmento no se puede ampliar más a la derecha, puesto que $13.8 - 13.1 = 0.7 \geq 0.4$. Al moverse una media a la derecha, el par $\bar{x}_{3\cdot}$ y $\bar{x}_{2\cdot}$ no puede ser subrayado, debido a que estas medias difieren en más de 0.4. De nuevo, al moverse a la derecha, la siguiente media, 13.8, no se puede conectar con ninguna otra a la derecha y, por último, se pueden subrayar las dos últimas medias con el mismo segmento de recta.

$\bar{x}_{5\cdot}$	$\bar{x}_{1\cdot}$	$\bar{x}_{2\cdot}$	$\bar{x}_{4\cdot}$	$\bar{x}_{1\cdot}$
13.1	13.3	13.8	14.3	14.5

Las marcas 1 y 4 no son muy diferentes entre sí, pero son significativamente mayores que las otras tres marcas en sus contenidos promedio reales. La marca 2 es mucho mejor que 3 y 5, pero peor que 1 y 4, y las marcas 3 y 5 no difieren de manera importante.

Si $\bar{x}_{2\cdot} = 14.15$ en vez de 13.8 con la misma w calculada, entonces la configuración de medias subrayadas sería

$\bar{x}_{5\cdot}$	$\bar{x}_{3\cdot}$	$\bar{x}_{2\cdot}$	$\bar{x}_{4\cdot}$	$\bar{x}_{1\cdot}$
13.1	13.3	14.15	14.3	14.5

■

Ejemplo 10.5 Un biólogo desea estudiar los efectos del etanol en relación con el tiempo del sueño. Se elige una muestra de 20 ratas que coinciden en edad y otras características, y a cada rata se le suministra una inyección por vía oral que tiene una concentración particular de etanol por peso corporal. Luego se registró el momento del sueño con el movimiento rápido (REM) del ojo para cada rata durante un periodo de 24 horas, con los siguientes resultados:

Tratamiento (concentración de etanol)						$x_{i\cdot}$	$\bar{x}_{i\cdot}$
0 (control)	88.6	73.2	91.4	68.0	75.2	396.4	79.28
1 g/kg	63.0	53.9	69.2	50.1	71.5	307.7	61.54
2 g/kg	44.9	59.5	40.2	56.3	38.7	239.6	47.92
4 g/kg	31.0	39.6	45.3	25.2	22.7	163.8	32.76
						$x_{\cdot\cdot} = 1107.5$	$\bar{x}_{\cdot\cdot} = 55.375$

¿Los datos indican que el momento del sueño REM promedio real depende de la concentración de etanol? (Este ejemplo se basa en un experimento descrito en "Relationship of Ethanol Blood Level to REM and Non-REM Sleep Time and Distribution in the Rat", *Life Sciences*, 1978: 839-846.)

Las $\bar{x}_{i\cdot}$ difieren de manera sustancial entre sí, pero también hay mucha variabilidad dentro de cada muestra, así que para responder la pregunta con precisión, se debe llevar a cabo el ANOVA. Con $\sum\sum x_{ij}^2 = 68\,697.6$ y factor de corrección $x_{\cdot\cdot}^2/(IJ) = (1107.5)^2/20 = 61\,327.8$, las fórmulas de cálculo producen

$$\text{SST} = 68\,697.6 - 61\,327.8 = 7369.8$$

$$\text{SSTr} = \frac{1}{5}[(396.40)^2 + (307.70)^2 + (239.60)^2 + (163.80)^2] - 61\,327.8$$
$$= 67\,210.2 - 61\,327.8 = 5882.4$$

y

$$\text{SSE} = 7369.8 - 5882.4 = 1487.4$$

La tabla 10.4 es una tabla ANOVA de SAS. En la última columna se da el valor P, que es 0.0001. Con un nivel de significación de 0.05, se rechaza la hipótesis nula H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$, puesto que el valor $P = 0.0001 < 0.05 = \alpha$. El tiempo del sueño REM promedio real al parecer depende de la concentración.

Tabla 10.4 Tabla ANOVA de SAS

Procedimiento de análisis de varianza
Variable dependiente: TIEMPO

Fuente	GL	Suma de cuadrados	Cuadrado medio	Valor F	Pr > F
Modelo	3	5882.35750	1960.78583	21.09	0.0001
Error	16	1487.40000	92.96250		
Total Corregido	19	7369.75750			

Hay $I = 4$ tratamientos y 16 gl para el error, así que $Q_{0.05,4,16} = 4.05$ y $w = 4.05\sqrt{93.0/5} = 17.47$. Al ordenar las medias y subrayar, se obtiene

$\bar{x}_{4\cdot}$	$\bar{x}_{3\cdot}$	$\bar{x}_{2\cdot}$	$\bar{x}_{1\cdot}$
32.76	47.92	61.54	79.28

La interpretación de este subrayado se debe hacer con cuidado, puesto que al parecer se concluyó que los tratamientos 2 y 3 no difieren, 3 y 4 no difieren, pero 2 y 4 sí. La manera sugerida de expresar esto es decir que aunque la evidencia permite concluir que los tratamientos 3 y 4 difieren entre sí, no se ha demostrado que alguno sea muy diferente de 3. El tratamiento 1 tiene un tiempo de sueño promedio real significativamente mayor que cualquiera de los otros tratamientos.

En la figura 10.4 se observa el resultado de SAS de la aplicación del procedimiento de Tukey.

Alfa = 0.05 gl = 16 MSE = 92.9625
Valor crítico del intervalo estudentizado = 4.046
Diferencia significativa mínima = 17.446

Las medias con la misma letra no son iguales son significativamente distintas

Agrupamiento de Tukey	Media	N	TRATAMIENTO
A	79.280	5	0(control)
B	61.540	5	1 gm/kg
B C	47.920	5	2 gm/kg
C	32.760	5	4 gm/kg

Figura 10.4 Método de Tukey con SAS

■

Interpretación de α en el método de Tukey

Ya se expresó antes que el nivel de confianza *simultáneo* se controla mediante el método de Tukey. Entonces, ¿qué significa en este caso "simultáneo"? Considere calcular un IC de 95% para una media poblacional μ con base en una muestra de esa población y luego un IC de 95% para una proporción poblacional p con base en otra muestra seleccionada independientemente de la primera. Antes de obtener los datos, la probabilidad de que el primer intervalo incluya a μ es 0.95, y esto también es la probabilidad de que el segundo intervalo incluya a p. Debido a que las dos muestras se seleccionan de forma independiente entre sí, la probabilidad de que *ambos* intervalos incluyan los valores de los parámetros respectivos es $(0.95)(0.95) = (0.95)^2 \approx 0.90$. Así que el nivel de confianza *simultáneo o conjunto* para los dos intervalos es más o menos 90%, si se calculan pares de intervalos una y otra vez a partir de muestras independientes, a la larga alrededor de 90% de las veces el primer intervalo capta a μ y el segundo incluye a p. De forma similar, si se calculan tres IC con base en muestras independientes, el nivel de confianza simultáneo será $100(0.95)^3\% \approx 86\%$. Resulta claro que cuando se incremente el número de intervalos, disminuirá el nivel de confianza simultáneo de que todos los intervalos capten sus parámetros respectivos.

Ahora suponga que se desea mantener el nivel de confianza simultáneo al 95%. Entonces para las dos muestras independientes, el nivel de confianza individual para cada una tendría que ser $100\sqrt{0.95}\% \approx 97.5\%$. Mientras más grande sea el número de intervalos, mayor tendría que ser cada uno de los niveles de confianza para mantener el nivel simultáneo de 95%.

Lo complicado respecto a los intervalos de Tukey es que no se basan en muestras independientes —el MSE aparece en cada uno— y varios intervalos comparten las mismas $\bar{x}_{i\cdot}$ (p. ej., en el caso $I = 4$, tres intervalos diferentes utilizan $\bar{x}_1$). Esto significa que no hay argumento de probabilidad directo para afirmar el nivel de confianza simultáneo a partir de cada uno de los niveles de confianza. Sin embargo, se puede demostrar que si se utiliza $Q_{0.05}$, el nivel de confianza simultáneo se controla al 95%, mientras que con $Q_{0.01}$, se obtiene un nivel simultáneo de 99%. Para obtener un nivel simultáneo de 95%, el nivel de cada intervalo debe ser mucho mayor que 95%. Dicho de una manera un poco distinta, para obtener una tasa de error de 5% *por experimento* o *familia*, la tasa de error individual o por comparación para cada nivel debe ser mucho más pequeña que 0.05. MINITAB pide al usuario especificar la tasa de error por familia (p. ej., 5%) y luego incluye en el resultado la tasa de error individual (véase el ejercicio 16).

Intervalos de confianza para otras funciones paramétricas

En algunas situaciones, se desea un IC para una función de las μ_i más complicada que una diferencia $\mu_i - \mu_j$. Sea $\theta = \sum c_i\mu_i$ donde las c_i son constantes. Una función de este tipo es $\frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5)$, que en el contexto del ejemplo 10.4 mide la diferencia entre el grupo que consiste en las dos primeras marcas y la de las últimas tres marcas. Debido a que las X_{ij} tienen una distribución normal con $E(X_{ij}) = \mu_i$ y $V(X_{ij}) = \sigma^2$, $\hat{\theta} = \sum c_i\bar{X}_{i\cdot}$ está distribuida normalmente, es insesgada para θ y

$$V(\hat{\theta}) = V\left(\sum_i c_i\bar{X}_{i\cdot}\right) = \sum_i c_i^2 V(\bar{X}_{i\cdot}) = \frac{\sigma^2}{J}\sum_i c_i^2$$

Estimar σ^2 mediante MSE y formar $\hat{\sigma}_{\hat{\theta}}$ da como resultado una variable $t\ (\hat{\theta} - \theta)/\hat{\sigma}_{\hat{\theta}}$ que se puede tratar para obtener el siguiente intervalo de confianza de $100(1 - \alpha)\%$ para $\sum c_i\mu_i$:

$$\sum c_i\bar{x}_{i\cdot} \pm t_{\alpha/2,IJ-1)}\sqrt{\frac{\text{MSE}\sum c_i^2}{J}} \qquad (10.5)$$

Ejemplo 10.6 (Continuación del ejemplo 10.4) La función paramétrica para comparar las dos primeras marcas (comerciales) de filtros de aceite con las tres últimas marcas (nacionales) es $\theta = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5)$, de lo cual

$$\sum c_i^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 = \frac{5}{6}$$

Con $\theta = \frac{1}{2}(\bar{x}_{1\cdot} + \bar{x}_{2\cdot}) - \frac{1}{3}(\bar{x}_{3\cdot} + \bar{x}_{4\cdot} + \bar{x}_{5\cdot}) = 0.583$ y MSE = 0.088, un intervalo de 95% es

$$0.583 \pm 2.021\sqrt{5(0.088)/[(6)(9)]} = 0.583 \pm 0.182 = (0.401, 0.765)$$ ■

A veces se lleva a cabo un experimento para comparar cada uno de varios tratamientos "nuevos" con respecto a un tratamiento de control. En tales situaciones, es apropiada una técnica de comparaciones múltiples conocida como método de Dunnett.

Ejercicios | Sección 10.2 (11 a 21)

11. En un experimento para comparar las tasas de dispersión de cinco marcas de pintura látex amarilla para interiores, disponibles en un área particular, se utilizaron cuatro galones ($J = 4$) de cada pintura. Las tasas de dispersión promedio muestrales (ft²/gal) para las cinco marcas fueron $\bar{x}_{1\cdot} = 462.0$, $\bar{x}_{2\cdot} = 512.8$, $\bar{x}_{3\cdot} = 437.5$, $\bar{x}_{4\cdot} = 469.3$ y $\bar{x}_{5\cdot} = 532.1$. El valor calculado de F fue significativo al nivel $\alpha = 0.05$. Con MSE = 272.8, utilice el procedimiento de Tukey para investigar diferencias significativas en las tasas de dispersión promedio entre marcas.

12. En el ejercicio 11, suponga que $\bar{x}_{3\cdot} = 427.5$. Ahora, ¿cuáles tasas de dispersión promedio verdaderas difieren en forma significativa unas de otras? Asegúrese de usar el método de subrayado para ilustrar sus conclusiones, y escriba un párrafo que resuma sus resultados.

13. Repita el ejercicio 12 suponiendo que $\bar{x}_{2\cdot} = 502.8$ además de $\bar{x}_{3\cdot} = 427.5$.

14. Utilice el procedimiento de Tukey en los datos del ejercicio 4 para identificar diferencias en los tiempos de vuelo promedio reales entre los cuatro tipos de mosquito.

15. Aplique el procedimiento de Tukey en los datos del ejercicio 6 para identificar diferencias en el Fe total promedio real entre los cuatro tipos de formaciones (utilice MSE = 15.64).

16. Reconsidere los datos de rigidez axial proporcionados en el ejercicio 8. A continuación se da el resultado para el ANOVA obtenido con MINITAB:

```
Análisis de varianza para la rigidez
Fuente GL      SC      CM      F       P
Longitud
0.000  4       43993   10998   10.48
Error  30      31475   1049
Total  34      75468

Nivel     N    Media  Desv. est
4         7   333.21   36.59
6         7   368.06   28.57
8         7   375.13   20.83
10        7   407.36   44.51
12        7   437.17   26.00
Desv. est. agrupada = 32.39
Comparaciones por pares de Tukey
    Tasa de error por familia = 0.0500
Tasa de error individual = 0.00693
Valor crítico = 4.10
Intervalos para (media de nivel de columna)
- (media de nivel de renglón).
          4          6          8          10
6      -85.0
        15.4
8      -92.1      -57.3
         8.3       43.1
10    -124.3      -89.5      -82.4
        -23.9       10.9       18.0
12    -154.2     -119.3     -112.2      -80.0
        -53.8      -18.9      -11.8       20.4
```

a. ¿Es posible que las varianzas de las cinco distribuciones de índice de rigidez axial sean idénticas? Explique.

b. Utilice el resultado (sin referencia a la tabla de F para probar las hipótesis pertinentes.

c. Utilice los intervalos de Tukey obtenidos en el resultado para determinar cuáles medias difieren y construya el patrón de subrayado correspondiente.

17. Refiérase al ejercicio 5. Calcule un IC t de 95% para $\theta = \frac{1}{2}(\mu_1 + \mu_2) - \mu_3$.

18. Considere los siguientes datos acerca del desarrollo de plantas después de la aplicación de diferentes tipos de hormonas de crecimiento.

	1	13	17	7	14
	2	21	13	20	17
Hormona	3	18	15	20	17
	4	7	11	18	10
	5	6	11	15	8

a. Lleve a cabo un prueba F al nivel $\alpha = 0.05$.

b. ¿Qué pasa cuando se aplica el procedimiento de Tukey?

19. Considere un experimento de ANOVA de un solo factor en el que $I = 3$, $J = 5$, $\bar{x}_{1\cdot} = 10$, $\bar{x}_{2\cdot} = 12$ y $\bar{x}_{3\cdot} = 20$. Determine un valor de SSE para el cual $f > F_{0.05,2,12}$, de modo que se rechaza H_0: $\mu_1 = \mu_2 = \mu_3$ pero cuando se aplica el procedimiento de Tukey no se puede decir que alguna de las μ_i difiere de otra de forma importante.

20. Refiérase al ejercicio 19 y suponga que $\bar{x}_{1\cdot} = 10$, $\bar{x}_{2\cdot} = 15$ y $\bar{x}_{3\cdot} = 20$. ¿Puede encontrar ahora un valor de SSE que produce tal contradicción entre la prueba F y el procedimiento de Tukey?

21. En el artículo "The Effect of Enzyme Inducing Agents on the Survival Times of Rats Exposed to Lethal Levels of Nitrogen Dioxide" (*Toxicology and Applied Pharmacology*, 1978: 169-174) aparecen los siguientes datos acerca de tiempos de supervivencia para ratas expuestas a dióxido de nitrógeno (70 ppm) vía diferentes regímenes de inyección. En cada grupo hubieron $J = 14$ ratas.

Régimen	$\bar{x}_{i\cdot}$ (min)	s_i
1. Control	166	32
2. 3-Metilcolantreno	303	53
3. Alilisopropilacetamida	266	54
4. Fenobarbital	212	35
5. Cloropromacina	202	34
6. Acido *p*-aminobenzoico	184	31

a. Pruebe la hipótesis nula de que el tiempo de supervivencia promedio real no depende del régimen de inyección contra la alternativa de que hay cierta dependencia en el régimen de inyección con $\alpha = 0.01$.

b. Suponga que los intervalos de confianza de $100(1 - \alpha)\%$ para k diferentes funciones paramétricas se calculan a partir del mismo conjunto de datos del ANOVA. Entonces se comprueba fácilmente que el nivel de confianza simultáneo es por lo menos $100(1 - k\alpha)\%$. Calcule IC con nivel de confianza de por lo menos 98% para $\mu_1 - \frac{1}{5}(\mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6)$ y $\frac{1}{4}(\mu_2 + \mu_3 + \mu_4 + \mu_5) - \mu_6$.

10.3 Más acerca de ANOVA de un solo factor

En esta sección se consideran de forma breve algunas cuestiones más en relación con el ANOVA de un solo factor. Entre otras una descripción alternativa de los parámetros del modelo, β para la prueba F, la relación de la prueba para procedimientos considerados antes, transformación de datos, un modelo de efectos aleatorios y fórmulas para el caso de tamaños de muestra distintos.

Descripción alternativa del modelo de ANOVA

Las suposiciones del ANOVA de un solo factor se pueden describir de forma sucinta por medio de la "ecuación modelo"

$$X_{ij} = \mu_i + \epsilon_{ij}$$

donde ϵ_{ij} representa una desviación aleatoria desde la media poblacional o de tratamiento verdadero μ_i. Se supone que las ϵ_{ij} son variables aleatorias independientes, distribuidas normalmente (lo que significa que las X_{ij} también lo son) con $E(\epsilon_{ij}) = 0$ [de modo que $E(X_{ij}) = \mu_i$] y $V(\epsilon_{ij}) = \sigma^2$ [a partir de lo cual $V(X_{ij}) = \sigma^2$ para toda i y j]. Una descripción alternativa del ANOVA de un solo factor proporciona una idea extra y hace pensar en generalizaciones apropiadas a modelos que tienen que ver con más de un factor. Defina un parámetro μ mediante

$$\mu = \frac{1}{I}\sum_{i=1}^{I}\mu_i$$

y los parámetros $\alpha_1, \ldots, \alpha_I$ de acuerdo con

$$\alpha_i = \mu_i - \mu \qquad (i = 1, \ldots, I)$$

Entonces la media de tratamiento μ_i se puede escribir como $\mu + \alpha_i$ donde μ representa la respuesta global promedio verdadera en el experimento, y α_i es el efecto, medido como una desviación desde μ debido al i-ésimo tratamiento. Mientras que al inicio se tenían I parámetros, ahora se tiene $I + 1$ ($\mu, \alpha_1, \ldots, \alpha_I$). Sin embargo, debido a que $\sum \alpha_i = 0$ (la desviación promedio respecto de la respuesta promedio global es cero), sólo I se estos parámetros se determina de forma independiente, así que hay tantos parámetros independientes como había antes. En términos de μ y las α_i, el modelo se convierte en

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij} \qquad (i = 1, \ldots, I, \quad j = 1, \ldots, J)$$

En el capítulo 11 se desarrollarán modelos análogos para el ANOVA de varios factores. La afirmación de que las μ_i son idénticas es equivalente a la igualdad de las α_i y debido a que $\sum \alpha_i = 0$, la hipótesis nula se convierte en

$$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

En la sección 10.1, se expresó que MSTr es un estimador insesgado de σ^2 cuando H_0 es verdadera, pero en caso contrario tiende a sobrestimar σ^2. Para ser precisos,

$$E(\text{MSTr}) = \sigma^2 + \frac{J}{I-1}\sum \alpha_i^2$$

Cuando H_0 es verdadera, $\sum \alpha_i^2 = 0$ de modo que $E(\text{MSTr}) = \sigma^2$ (MSE es insesgado si H_0 es verdadera o no). Si se utiliza a $\sum \alpha_i^2$ como medida de hasta qué grado H_0 es falsa, entonces un valor más grande de $\sum \alpha_i^2$ da como resultado una mayor tendencia para que MSTr sobrestime a σ^2. En el capítulo siguiente las fórmulas para los cuadrados medios esperados de modelos con varios factores se utilizarán para sugerir cómo formar relaciones F a fin de probar varias hipótesis.

Demostración de la fórmula para *E*(MSTr) Para cualquier *va* Y, $E(Y^2) = V(Y) + [E(Y)]^2$, por lo tanto

$$\begin{aligned}
E(\text{SSTr}) &= E\left(\frac{1}{J}\sum_i X_{i\cdot}^2 - \frac{1}{IJ}X_{\cdot\cdot}^2\right) = \frac{1}{J}\sum_i E(X_{i\cdot}^2) - \frac{1}{IJ}E(X_{\cdot\cdot}^2) \\
&= \frac{1}{J}\sum_i \{V(X_{i\cdot}) + [E(X_{i\cdot})]^2\} - \frac{1}{IJ}\{V(X_{\cdot\cdot}) + [E(X_{\cdot\cdot})]^2\} \\
&= \frac{1}{J}\sum_i [J\sigma^2 + [J(\mu + \alpha_i)]^2] - \frac{1}{IJ}[IJ\sigma^2 + (IJ\mu)^2] \\
&= I\sigma^2 + IJ\mu^2 + 2\mu J\sum_i \alpha_i + J\sum_i \alpha_i^2 - \sigma^2 - IJ\mu^2 \\
&= (I-1)\sigma^2 + J\sum_i \alpha_i^2 \qquad \text{(pueto que } \textstyle\sum \alpha_i = 0)
\end{aligned}$$

Entonces el resultado se deduce de la relación $\text{MSTr} = \text{SSTr}/(I-1)$. ■

β para la prueba F

Considere un conjunto de valores de parámetro $\alpha_1, \alpha_2, \ldots, \alpha_I$ para los cuales H_0 no es verdadera. La probabilidad de un error de tipo II, β, es la probabilidad de que H_0 no sea rechazada cuando el conjunto es el de valores verdaderos. Se podría pensar que β tendría que ser determi-

nada por separado para cada configuración diferente de las α_i. Por fortuna, puesto que β para la prueba F depende de las α_i y σ^2 sólo de $\sum\alpha_i^2/\sigma^2$, se puede evaluar de forma simultánea para muchas alternativas distintas. Por ejemplo, $\sum\alpha_i^2 = 4$ para cada uno de los siguientes conjuntos de α_i para las cuales H_0 es falsa, así que β es idéntica para las tres alternativas:

1. $\alpha_1 = -1,\ \alpha_2 = -1,\ \alpha_3 = 1,\ \alpha_4 = 1$
2. $\alpha_1 = -\sqrt{2},\ \alpha_2 = \sqrt{2},\ \alpha_3 = 0,\ \alpha_4 = 0$
3. $\alpha_1 = -\sqrt{3},\ \alpha_2 = \sqrt{1/3},\ \alpha_3 = \sqrt{1/3},\ \alpha_4 = \sqrt{1/3}$

La cantidad $J\sum\alpha_i^2/\sigma^2$ se llama **parámetro de no centralidad** para ANOVA unidireccional (debido a que cuando H_0 es falsa la estadística de prueba tiene una distribución F *no central* con éste como uno de sus parámetros), y β es una función decreciente del valor de este parámetro. Por consiguiente, para valores fijos de σ^2 y J, la hipótesis nula tiene más probabilidades de ser rechazada para alternativas alejadas de H_0 ($\sum\alpha_i^2$ grande) que para alternativas cercanas a H_0. Para un valor fijo de $\sum\alpha_i^2$, β disminuye a medida que crece el tamaño muestral J en cada tratamiento, y se incrementa cuando aumenta la varianza σ^2 (puesto que la mayor variabilidad subyacente hace más difícil detectar alguna desviación por H_0).

Debido a que el cálculo manual de β y la determinación del tamaño de muestra para la prueba F son bastante difíciles (como en el caso de las pruebas t), los estadísticos han construido conjuntos de curvas de las que se puede obtener β. Los conjuntos de curvas para los gl del numerador $\nu_1 = 3$ y $\nu_1 = 4$ se muestran en la figura 10.5* y en la figura 10.6*, respectivamente. Después que se especifican los valores de σ^2 y las α_i para los que se desea β, éstos se utilizan para calcular el valor de ϕ, donde $\phi^2 = (J/I)\sum\alpha_i^2/\sigma^2$. Entonces se ingresa al conjunto apropiado de curvas en el valor de ϕ sobre el eje horizontal, se va hasta la curva relacionada con los ν_2 grados de libertad del error y se pasa al valor de potencia en el eje vertical. Por último, $\beta = 1 -$ potencia.

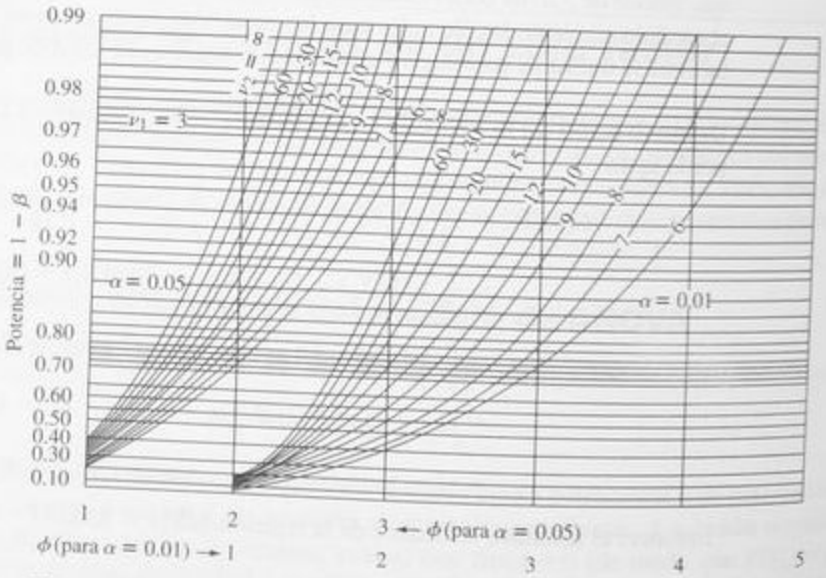


Figura 10.5 Curvas de potencia para la prueba F de ANOVA ($\nu_1 = 3$)

*De E. S. Pearson y H. O. Hartley, "Charts of the Power Function for Analysis of Variance Tests, Derived from the Non-central distribution", *Biometrika*, vol. 38, 1951: 112, con autorización del consejo de administración de Biometrika.

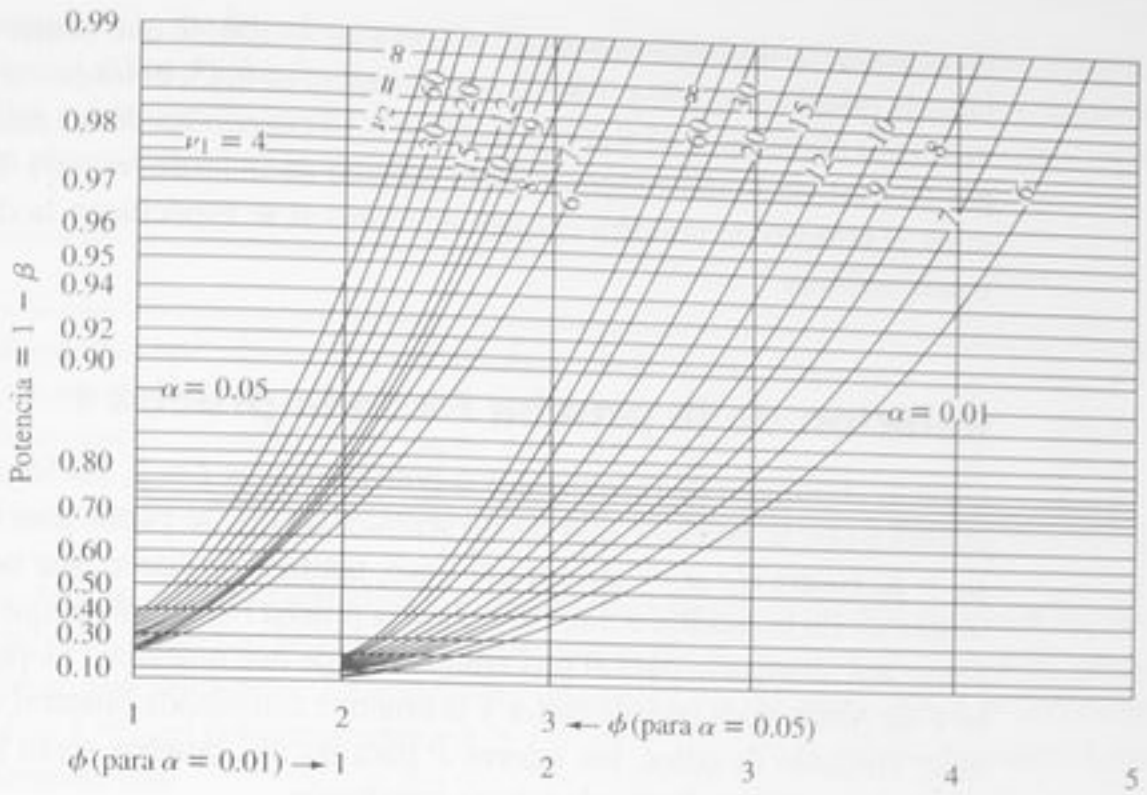


Figura 10.6 Curvas de potencia para la prueba F de ANOVA ($\nu_1 = 4$)

Ejemplo 10.7 Se investigarán los efectos de cuatro tratamientos térmicos en el límite de elasticidad (toneladas/pulg^2) de lingotes de acero. Se van a fundir ocho lingotes usando cada tratamiento. Suponga que la desviación estándar verdadera del límite de elasticidad para cualquiera de los cuatro tratamientos es $\sigma = 1$. ¿Qué tan probable es que no se rechace H_0 al nivel 0.05 si los tres tratamientos tienen el mismo límite de elasticidad esperado y el otro tiene un límite de elasticidad esperado que es una tonelada/pulg^2 mayor que el valor común de los otros tres (es decir, el cuarto límite está en promedio una desviación estándar arriba de los que corresponden a los tres primeros tratamientos)?

Suponga que $\mu_1 = \mu_2 = \mu_3$ y $\mu_4 = \mu_1 + 1$, $\mu = (\sum \mu_i)/4 = \mu_1 + \frac{1}{4}$. Entonces $\alpha_1 = \mu_1 - \mu = -\frac{1}{4}$, $\alpha_2 = -\frac{1}{4}$, $\alpha_3 = -\frac{1}{4}$, $\alpha_4 = \frac{3}{4}$ por lo tanto

$$\phi^2 = \frac{8}{4}\left[\left(\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(\frac{3}{4}\right)^2\right] = \frac{3}{2}$$

y $\phi = 1.22$. Los grados de libertad son $\nu_1 = I - 1 = 3$ y $\nu_2 = I(J - 1) = 28$, así que al interpolar de manera visual entre $\nu_2 = 20$ y $\nu_2 = 30$ se obtiene la potencia ≈ 0.47 y $\beta \approx 0.53$. Esta β es bastante grande, entonces se podría decidir incrementar el valor de J. ¿Cuántos lingotes de cada tipo se requerirían para obtener $\beta \approx 0.05$ para la alternativa en consideración? Al probar diferentes valores de J, se puede comprobar que $J = 24$ satisfará el requerimiento, pero no alguna J más pequeña. ■

Como alternativa para usar las curvas de potencia, el paquete de software de estadística SAS tiene una función que calcula el área acumulada bajo la curva no central F (se introduce F_α, gl del numerador, gl del denominador y ϕ^2), y esta área es β. La versión 13 de MINITAB hace algo bastante diferente. Se le pide al usuario que especifique la diferencia máxima entre las μ_i en vez de cada una de las medias. Por ejemplo, se podría desear calcular la potencia de la prueba cuando $I = 4$, $\mu_1 = 100$, $\mu_2 = 101$, $\mu_3 = 102$, y $\mu_4 = 106$. Entonces la diferencia máxima es $106 - 100 = 6$. Sin embargo, la potencia depende no sólo de esta diferencia máxima, sino de los valores de las μ_i. En esta situación, MINITAB calcula el valor más pequeño posible

de la potencia, dependiendo de $\mu_1 = 100$ y $\mu_4 = 106$, lo que ocurre cuando las otras dos μ están a la mitad entre 100 y 106. Si esta potencia es 0.85, entonces se puede decir que la potencia es por lo menos 0.85 y β es a lo sumo 0.15 cuando las dos μ más extremas están separadas por 6 (también se debe especificar el tamaño de muestra, α y σ). El software también determina el tamaño de muestra común necesario si se especifican la diferencia máxima y la potencia mínima.

Relación de la prueba *F* con la prueba *t*

Cuando el número de tratamientos o poblaciones es $I = 2$, las fórmulas y los resultados relacionados con la prueba F aún tienen sentido, así que se puede usar el ANOVA para probar H_0: $\mu_1 = \mu_2$ contra H_a: $\mu_1 \neq \mu_2$. En este caso, también se puede usar una prueba t bilateral de dos muestras. En la sección 9.3 se mencionó la prueba t combinada, que requiere varianzas iguales, como una alternativa para el procedimiento t de dos muestras. Se puede demostrar que la prueba F de ANOVA de un solo factor y la prueba t combinada bilateral son equivalentes; para cualquier conjunto de datos, los valores P para las dos pruebas serán idénticos, así que mediante cualquier prueba se llega a la misma conclusión.

La prueba t de dos muestras es más flexible que la prueba F cuando $I = 2$ por dos razones. Primero, es válida sin la suposición de que $\sigma_1 = \sigma_2$; segundo, se puede usar para probar H_a: $\mu_1 > \mu_2$ (una prueba de extremo superior) o H_a: $\mu_1 < \mu_2$ así como también H_a: $\mu_1 \neq \mu_2$. En el caso $I \geq 3$, infortunadamente no hay procedimiento de prueba general que se sepa tenga buenas propiedades sin suponer varianzas iguales.

ANOVA de un solo factor cuando los tamaños de muestra no son iguales

Cuando los tamaños de muestra de cada población o tratamiento son distintos, represente con $J_1, J_2, \ldots, J_I$ los I tamaños muestrales y con $n = \sum_i J_i$ el número total de observaciones. En el cuadro siguiente se muestran las fórmulas de ANOVA y el procedimiento de prueba.

$$\text{SST} = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(X_{ij} - \overline{X}_{\cdot\cdot})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J_i}X_{ij}^2 - \frac{1}{n}X_{\cdot\cdot}^2 \qquad \text{gl} = n - 1$$

$$\text{SSTr} = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 = \sum_{i=1}^{I}\frac{1}{J_i}X_{i\cdot}^2 - \frac{1}{n}X_{\cdot\cdot}^2 \qquad \text{gl} = I - 1$$

$$\text{SSE} = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(X_{ij} - \overline{X}_{i\cdot})^2 = \text{SST} - \text{SSTr} \qquad \text{gl} = \sum(J_i - 1) = n - I$$

Valor de la estadística de prueba:

$$f = \frac{\text{MSTr}}{\text{MSE}} \qquad \text{donde MSTr} = \frac{\text{SSTr}}{I - 1} \quad \text{MSE} = \frac{\text{SSE}}{n - I}$$

Región de rechazo: $f \geq F_{\alpha,I-1,n-I}$

Ejemplo 10.8 En el artículo "On the Development of a New Approach for the Determination of Yield Strength in Mg-based Alloys" (*Light Metal Age*, octubre, 1998: 51-53) aparecen los datos siguientes acerca del módulo elástico (GPa) obtenido mediante un método ultrasónico para especímenes de cierta aleación producida con tres procedimientos de colado.

									J_i	$x_{i\cdot}$	$\bar{x}_{i\cdot}$
Moldeado permanente	45.5	45.3	45.4	44.4	44.6	43.9	44.6	44.0	8	357.7	44.71
Colada por presión	44.2	43.9	44.7	44.2	44.0	43.8	44.6	43.1	8	352.5	44.06
Moldeado en yeso	46.0	45.9	44.8	46.2	45.1	45.5			6	273.5	45.58
									22	983.7	

Sean μ_1, μ_2 y μ_3 los módulos elásticos promedio reales para los tres procesos en las circunstancias especificadas. Las hipótesis pertinentes son H_0: $\mu_1 = \mu_2 = \mu_3$ contra H_a: por lo menos dos de las μ_i son diferentes. La estadística de prueba es, por supuesto, $F = \text{MSTr/MSE}$, con base en $I - 1 = 2$ gl del numerador y $n - I = 22 - 3 = 19$ gl del denominador. Las cantidades pertinentes son

$$\sum\sum x_{ij}^2 = 43\,998.73 \qquad \text{CF} = \frac{983.7^2}{22} = 43\,984.80$$

$$\text{SST} = 43\,998.73 - 43\,984.80 = 13.93$$

$$\text{SSTr} = \frac{357.7^2}{8} + \frac{352.5^2}{8} + \frac{273.5^2}{6} - 43\,984.80 = 7.93$$

$$\text{SSE} = 13.93 - 7.93 = 6.00$$

Los demás cálculos se muestran en la siguiente tabla ANOVA. Puesto que $F_{0.001,2,19} = 10.16 < 12.56 = f$, el valor P es más pequeño que 0.001. Por consiguiente, se debe rechazar la hipótesis nula a algún nivel de significación razonable; hay evidencia convincente para concluir que el módulo elástico promedio de alguna manera depende del proceso de colada que se utilice.

Fuente de variación	**gl**	**Suma de cuadrados**	**Cuadrado medio**	f
Tratamientos	2	7.93	3.965	12.56
Error	19	6.00	0.3158	
Total	21	13.93		

■

Comparaciones múltiples cuando los tamaños de muestra son distintos

Hay más controversia entre los estadísticos en relación con el procedimiento de comparaciones múltiples a usar cuando los tamaños de muestra son distintos que en el caso de muestras del mismo tamaño. El procedimiento que se presenta aquí se recomienda en el excelente libro *Beyond ANOVA: Basics of Applied Statistics* (véase la bibliografía del capítulo) para cuando los I tamaños de muestras $J_1, J_2, \ldots, J_I$ están razonablemente cerca entre sí ("desequilibrio leve"). Éste modifica el método de Tukey mediante promedios de pares de los $1/J_i$ en lugar de $1/J$.

Sea

$$w_{ij} = Q_{\alpha,I,n-I} \cdot \sqrt{\frac{\text{MSE}}{2}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)}$$

Entonces la probabilidad es *aproximadamente* $1 - \alpha$ de que

$$\bar{X}_{i\cdot} - \bar{X}_{j\cdot} - w_{ij} \le \mu_i - \mu_j \le \bar{X}_{i\cdot} - \bar{X}_{j\cdot} + w_{ij}$$

Para toda i y j ($i = 1, \ldots, I$ y $j = 1, \ldots, I$) con $i \neq j$.

El nivel de confianza simultáneo $100(1 - \alpha)\%$ es sólo aproximado más que exacto, al igual que con tamaños de muestra iguales. Aún se puede usar el método de subrayado, pero ahora el factor w_{ij} utilizado para decidir si se pueden relacionar $\bar{x}_{i\cdot}$ y $\bar{x}_{j\cdot}$ dependerá de J_i y J_j.

Ejemplo 10.9 (Continuación del ejemplo 10.8)

Los tamaños de muestra para los datos del módulo elástico fueron $J_1 = 8$, $J_2 = 8$, $J_3 = 6$ y $I = 3$, $n - I = 19$, MSE = 0.316. Un nivel de confianza simultáneo de más o menos 95% requiere que $Q_{0.05,3,19} = 3.59$, del cual

$$w_{12} = 3.59\sqrt{\frac{0.316}{2}\left(\frac{1}{8} + \frac{1}{8}\right)} = 0.713, \qquad w_{13} = 0.771 \qquad w_{23} = 0.771$$

Puesto que $\bar{x}_{1\cdot} - \bar{x}_{2\cdot} = 44.71 - 44.06 = 0.65 < w_{12}$, se considera que μ_1 y μ_2 no son significativamente distintas. En el siguiente esquema de subrayado se observa que μ_1 y μ_3 al parecer difieren de manera importante, al igual que μ_2 y μ_3.

2. Por presión	1. Permanente	3. Yeso
44.06	44.71	45.58

■

Transformación de datos

El uso de los métodos de ANOVA se puede invalidar mediante diferencias sustanciales en las varianzas $\sigma_1^2, \ldots, \sigma_I^2$ (que hasta ahora se ha supuesto son iguales con valor común σ^2). A veces sucede que $V(X_{ij}) = \sigma_i^2 = g(\mu_i)$, una función conocida de μ_i (de modo que cuando H_0 es falsa, las varianzas no son iguales). Por ejemplo, si X_{ij} tiene una distribución de Poisson con parámetro λ_i (normalmente aproximado si $\lambda_i \ge 10$), entonces $\mu_i = \lambda_i$ y $\sigma_i^2 = \lambda_i$, por lo tanto $g(\mu_i) = \mu_i$ es la función conocida. En estos casos, es posible transformar las X_{ij} a $h(X_{ij})$ de modo que tendrán más o menos varianzas iguales (mientras se deje a las variables transformadas aproximadamente normales), y entonces la prueba F se puede usar en las observaciones transformadas. La idea clave al elegir una transformación $h(\cdot)$ es que con frecuencia $V[h(X_{ij})] \approx V(X_{ij}) \cdot [h'(\mu_i)]^2 = g(\mu_i) \cdot [h'(\mu_i)]^2$. Se desea encontrar la función $h(\cdot)$ para la cual $g(\mu_i) \cdot [h'(\mu_i)]^2 = c$ (una constante) para toda i.

PROPOSICIÓN

Si $V(X_{ij}) = g(\mu_i)$, una función conocida de μ_i, entonces una transformación $h(X_{ij})$ que "estabiliza la varianza" de modo que $V[h(X_{ij})]$ es aproximadamente la misma para cada i se determina mediante $h(x) \propto \int [g(x)]^{-1/2}\,dx$.

En el caso de Poisson, $g(x) = x$, así que $h(x)$ debe ser proporcional a $\int x^{-1/2}\,dx = 2x^{1/2}$. Los datos de Poisson se deben transformar a $h(x_{ij}) = \sqrt{x_{ij}}$ antes del análisis.

Modelo de efectos aleatorios

Los problemas de un solo factor estudiados hasta el momento han sido considerados ejemplos de un modelo de ANOVA de **efectos fijos**. Esto quiere decir que los niveles elegidos del factor bajo estudio, son los únicos que el experimentador considera pertinentes. El modelo de efectos fijos de un solo factor es

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij} \qquad \sum \alpha_i = 0 \tag{10.6}$$

donde las ϵ_{ij} son aleatorias y tanto μ como las α_i son parámetros fijos cuyos valores se desconocen.

En algunos problemas de un solo factor, se eligen los niveles particulares que estudia el experimentador, ya sea por diseño o muestreo, a partir de una población grande de niveles. Por ejemplo, para estudiar los efectos en el tiempo de ejecución de una tarea al usar operadores distintos en una determinada máquina, se podría elegir una muestra de cinco operadores de un gran grupo. De manera similar, el efecto del pH del suelo en la producción de plantas de maíz se podría estudiar mediante el uso de suelos con cuatro valores específicos de pH elegidos de entre los muchos niveles posibles de pH. Cuando los niveles utilizados se eligen al azar de una población más grande de niveles posibles, se dice que el factor es aleatorio y no fijo, y resulta inapropiado el modelo de efectos fijos (10.6). Un modelo análogo de **efectos aleatorios** se obtiene al reemplazar las α_i fijas en la ecuación (10.6) por variables aleatorias. La descripción del modelo resultante es

$$X_{ij} = \mu + A_i + \epsilon_{ij} \quad \text{con } E(A_i) = E(\epsilon_{ij}) = 0$$
$$V(\epsilon_{ij}) = \sigma^2 \qquad V(A_i) = \sigma_A^2 \tag{10.7}$$

las A_i y las ϵ_{ij} tienen una distribución normal y son independientes entre sí.

La condición $E(A_i) = 0$ en (10.7) es similar a la condición $\sum \alpha_i = 0$ en (10.6); expresa que el efecto esperado o promedio del i-ésimo nivel medido como una desviación desde μ es cero.

Para el modelo de efectos aleatorios (10.7), la hipótesis de ningún efecto debido a diferentes niveles es H_0: $\sigma_A^2 = 0$, que dice que niveles diferentes del factor no contribuyen a la variabilidad de la respuesta. *Aunque las hipótesis en los modelos de efectos fijos y aleatorios de un solo factor son diferentes, se prueban exactamente de la misma forma*, al formar $F =$ MSTr/MSE y rechazar H_0 si $f \geq F_{\alpha,I-1,n-I}$. Esto se puede justificar de manera intuitiva al observar que $E(\text{MSE}) = \sigma^2$ (al igual que para los efectos fijos), mientras que

$$E(\text{MSTr}) = \sigma^2 + \frac{1}{I-1}\left(n - \frac{\sum J_i^2}{n}\right)\sigma_A^2 \tag{10.8}$$

donde $J_1, J_2, \ldots, J_I$ son los tamaños de muestra y $n = \sum J_i$. El factor dentro del paréntesis en el lado derecho de la ecuación (10.8) es no negativo, así que de nuevo $E(\text{MSTr}) = \sigma^2$ si H_0 es verdadera y $E(\text{MSTr}) > \sigma^2$ si H_0 es falsa.

Ejemplo 10.10 El estudio de fuerzas no destructivas y esfuerzos en materiales proporciona información importante para el diseño técnico eficaz. En el artículo "Zero-Force Travel Time Parameters for Ultrasonic Head-Waves in Railroad Rail" (*Materials Evaluation*, 1985: 854-858) se describe un estudio del tiempo de viaje para cierto tipo de onda que resulta del esfuerzo longitudinal de rieles utilizados para vías férreas. Se hicieron tres mediciones en cada uno de seis rieles elegidos al azar de una población de rieles. Los investigadores utilizaron el ANOVA de efectos aleatorios para decidir si alguna variación en el tiempo de viaje se podría atribuir a la "variabilidad entre rieles". Los datos se muestran en la tabla siguiente (cada valor, en nanosegundos, resultó de restar 36.1 μs de la observación original) junto con la tabla ANOVA obtenida. El valor de la relación F es muy significante, así que se rechaza H_0: $\sigma_A^2 = 0$ en favor de la conclusión de las diferencias entre rieles es una fuente de variabilidad del tiempo de viaje.

				$x_{i.}$
1:	55	53	54	162
2:	26	37	32	95
3:	78	91	85	254
4:	92	100	96	288
5:	49	51	50	150
6:	80	85	83	248
				$x_{..} = 1197$

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
Tratamientos	5	9310.5	1862.1	115.2
Error	12	194.0	16.17	
Total	17	9504.5		

■

Ejercicios | Sección 10.3 (22 a 34)

22. Los datos siguientes se refieren a la producción de tomates (kg/parcela) para cuatro niveles de salinidad; aquí el nivel de salinidad se refiere a la conductividad eléctrica (CE), donde los niveles elegidos fueron CE = 1.6, 3.8, 6.0 y 10.2 nmhos/cm:

1.6	59.5	53.3	56.8	63.1	58.7
3.8	55.2	59.1	52.8	54.5	
6.0	51.7	48.8	53.9	49.0	
10.2	44.6	48.5	41.0	47.3	46.1

Utilice la prueba F al nivel $\alpha = 0.05$ para probar algunas diferencias en el rendimiento promedio debido a los diferentes niveles de salinidad.

23. Aplique el método de Tukey modificado a los datos del ejercicio 22 para identificar diferencias significativas entre las μ_i.

24. La siguiente tabla ANOVA parcial se tomó del artículo "Perception of Spatial Incongruity" (*J. Nervous and Mental Disease*, 1961: 222) en el que se evaluaron y compararon las capacidades de tres grupos distintos para identificar una incongruencia perceptiva. Los individuos del experimento habían sido hospitalizados para ser sometidos a tratamiento psiquiátrico. Hubo 21 individuos en el grupo depresivo, 32 en el grupo funcional "otro" y 21 en el grupo con daño cerebral. Complete la tabla de ANOVA y lleve a cabo la prueba F al nivel $\alpha = 0.01$.

Fuente	gl	Suma de cuadrados	Cuadrado medio	f
			76.09	
Grupos				
Error		1123.14		
Total				

25. Los lípidos proporcionan gran parte de la energía de la dieta en los cuerpos de infantes y niños jóvenes. Hay un interés creciente en la calidad del suministro de lípidos de la dieta durante la infancia como determinante principal del crecimiento, desarrollo visual y neural y salud a largo plazo. En el artículo "Essential Fat Requirements of Preterm Infants" (*Amer. J. of Clinical Nutrition*, 2000: 245S-250S) aparecen los datos siguientes acerca de grasas poliinsaturadas totales (%) para infantes que fueron aleatorios para cuatro regímenes de alimentación: leche materna, fórmula basada en aceite de maíz, fórmula con aceite de soya o fórmula a base de aceite de soya y productos del mar:

Régimen	Tamaño de la muestra	Media muestral	DE de la muestra
Leche materna	8	43.0	1.5
AM	13	42.4	1.3
AS	17	43.1	1.2
ASPM	14	43.5	1.2

a. ¿Qué suposiciones se deben hacer acerca de las distribuciones de grasa poliinsaturada total antes de llevar a cabo un ANOVA de un solo factor para decidir si hay algunas diferencias en el contenido de grasa promedio real?

b. Realice la prueba sugerida en el inciso (a). ¿Qué se puede decir acerca del valor P?

26. Se analizaron muestras de seis marcas de margarina de dieta o artificial, para determinar el nivel de ácidos grasos poliinsaturados fisiológicamente activos (PAPFUA, en porcentajes), y se obtuvieron los datos siguientes:

Imperial	14.1	13.6	14.4	14.3	
Parkay	12.8	12.5	13.4	13.0	12.3
Blue Bonnet	13.5	13.4	14.1	14.3	
Chiffon	13.2	12.7	12.6	13.9	
Mazola	16.8	17.2	16.4	17.3	18.0
Fleischmann's	18.1	17.2	18.7	18.4	

(Los números anteriores son ficticios, pero las medias muestrales concuerdan con los datos publicados en el número de enero de 1975 de la revista *Consumer Reports*).

a. Por medio del ANOVA pruebe las diferencias entre los porcentajes de PAPFUA promedio reales para las diferentes marcas.

b. Calcule los IC para las $(\mu_i - \mu_j)$.

c. Mazola y Fleischmann son a base de maíz, mientras que las otras son a base de soya. Calcule un IC para

$$\frac{(\mu_1 + \mu_2 + \mu_3 + \mu_4)}{4} - \frac{(\mu_5 + \mu_6)}{2}$$

[*Sugerencia*: modifique la expresión para $V(\hat{\theta})$ que condujo a la ecuación (10.5) en la sección anterior.]

27. Si bien el te es la bebida de más consumo a nivel mundial después del agua, poco se sabe acerca de su valor nutricional. La folacina es la única vitamina B presente en cantidad insignificante en el te, y avances recientes en los métodos de valoración, han hecho posible la determinación precisa del contenido de folacina. Considere los datos siguientes acerca del contenido de folacina para especímenes elegidos al azar de las cuatro marcas principales de te verde.

Marca	Observaciones						
1	7.9	6.2	6.6	8.6	8.9	10.1	9.6
2	5.7	7.5	9.8	6.1	8.4		
3	6.8	7.5	5.0	7.4	5.3	6.1	
4	6.4	7.1	7.9	4.5	5.0	4.0	

(Los datos se basan en "Folacin Content of Tea", *J. Amer. Dietetic Assoc.*, 1983: 627-632.) ¿Estos datos hacen pensar que el contenido de folacina promedio real es el mismo para las cuatro marcas?

a. Realice una prueba con $\alpha = 0.05$ vía el método del valor P.

b. Evalúe la posibilidad de algunas suposiciones requeridas para su análisis del inciso (a).

c. Realice un análisis de comparaciones múltiples para identificar diferencias importantes entre las marcas.

28. Para un ANOVA de un solo factor con tamaños de muestra J_i $(i = 1, 2, \ldots, I)$, muestre que $\text{SSTr} = \sum J_i(\overline{X}_i - \overline{X}_\cdot)^2 = \sum_i J_i \overline{X}_i^2 - n\overline{X}_\cdot^2$, donde $n = \sum J_i$.

29. Cuando los tamaños de muestra son iguales $(J_i = J)$, los parámetros $\alpha_1, \alpha_2, \ldots, \alpha_I$ de la parametrización alternativa están restringidos por $\sum \alpha_i = 0$. Utilice esto para demostrar que

$$E(\text{MSTr}) = \sigma^2 + \frac{1}{I-1}\sum J_i \alpha_i^2$$

¿Cuál es $E(\text{MSTr})$ cuando H_0 es verdadera? [Esta esperanza es correcta si $\sum J_i \alpha_i = 0$ se sustituye por la restricción $\sum \alpha_i = 0$ (o cualquier otra restricción lineal simple en las α_i utilizada para reducir el modelo a I parámetros independientes), pero $\sum J_i \alpha_i = 0$ simplifica

el álgebra y produce estimaciones naturales para los parámetros del modelo (en particular, $\hat{\alpha}_i = \bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot}$).]

30. Reconsidere el ejemplo 10.7, el cual tiene que ver con una investigación de los efectos de diferentes tratamientos térmicos en el límite de elasticidad de lingotes de acero.
    a. Si $J = 8$ y $\sigma = 1$, ¿cuál es β para la prueba F de nivel 0.05 cuando $\mu_1 = \mu_2$, $\mu_3 = \mu_1 - 1$ y $\mu_4 = \mu_1 + 1$?
    b. Para la alternativa *a*) ¿qué valor de J es necesario para poder obtener $\beta = 0.05$?
    c. Si hay $I = 5$ tratamientos térmicos, $J = 10$ y $\sigma = 1$, ¿cuál es β para la prueba F de nivel 0.05 cuando cuatro de las μ_i son iguales y la quinta difiere en 1 de las otras cuatro?

31. Cuando los tamaños de muestra no son iguales, el parámetro de no centralidad es $\sum J_i\alpha_i^2/\sigma^2$ y $\phi^2 = (1/I)\sum J_i\alpha_i^2/\sigma^2$. En referencia al ejercicio 22, ¿cuál es la potencia de la prueba cuando $\mu_2 = \mu_3$, $\mu_1 = \mu_2 - \sigma$, y $\mu_4 = \mu_2 + \sigma$?

32. En un experimento para comparar la calidad de cuatro marcas de cintas para grabar de bobina a bobina, se seleccionaron cinco bobinas de 2400 pies de cada marca (A-D) y se determinó el número de defectos en cada bobina.

    A: 10 5 12 14 8
    B: 14 12 17 9 8
    C: 13 18 10 15 18
    D: 17 16 12 22 14

    Se cree que el número de defectos tiene más o menos una distribución de Poisson para cada marca. Analice los datos al nivel 0.01 para ver si el número esperado de defectos por bobina es el mismo para cada marca.

33. Suponga que X_{ij} es una variable binomial con parámetros n y p_i (por lo tanto aproximadamente normal cuando $np_i \geq 5$ y $nq_i \geq 5$). Entonces, puesto que $\mu_i = np_i$, $V(X_{ij}) = \sigma_i^2 = np_i(1 - p_i) = \mu_i(1 - \mu_i/n)$. ¿Cómo se deben transformar las X_{ij} de manera que se estabilice la varianza? [*Sugerencia*: $g(\mu_i) = \mu_i(1 - \mu_i/n)$.]

34. Simplifique $E(\text{MSTr})$ para el modelo de efectos aleatorios cuando $J_1 = J_2 = \cdots = J_I = J$.

Ejercicios complementarios (35 a 46)

35. Se llevó a cabo un experimento para comparar los flujos de cuatro tipos de tobera.
    a. Los tamaños de muestra fueron 5, 6, 7 y 6, respectivamente, y a partir de cálculos se obtuvo $f = 3.68$. Exprese y pruebe las hipótesis pertinentes con $\alpha = 0.01$
    b. Un análisis de los datos por medio de un paquete de software, produjo un valor $P = 0.029$. Al nivel 0.01, ¿cuál sería su conclusión y por qué?

36. En el artículo "Computer-Assisted Instruction Augmented with Planned Teacher/Student Contacts" (*J. Exp. Educ.*, invierno, 1980-1981: 120-126) se compararon cinco métodos para enseñar estadística descriptiva. Los cinco métodos clase y discusión (C/D), instrucción con libro de texto programado (R), texto programado con clases (R/L), instrucción con computadora (C) e instrucción con computadora y clases (C/L). Se asignaron al azar 45 estudiantes, nueve a cada método. Después de completar el curso, se aplicó a los estudiantes un examen de una hora. Además, se aplicó un examen de retención de 10 minutos seis semanas después. A continuación se dan las cantidades de resumen.

	Examen		Prueba de retención	
Método	$\bar{x}_{i\cdot}$	s_i	$\bar{x}_{i\cdot}$	s_i
C/D	29.3	4.99	30.20	3.82
R	28.0	5.33	28.80	5.26
R/L	30.2	3.33	26.20	4.66
C	32.4	2.94	31.10	4.91
C/L	34.2	2.74	30.20	3.53

    La gran media para el examen fue 30.82, y la gran media para la prueba de retención fue 29.30.
    a. ¿Los datos indican que hay una diferencia entre los cinco métodos de enseñanza con respecto a la puntuación promedio real del examen? Utilice $\alpha = 0.05$.
    b. Con un nivel de significación 0.05, pruebe la hipótesis nula de ninguna diferencia entre las puntuaciones promedio reales de la prueba de retención para los cinco métodos de enseñanza.

37. Varios factores contribuyen al buen funcionamiento de un motor eléctrico ("Increasing Market Share Through Improved Product and Process Design: An Experimen-

tal Approach", *Quality Engineering*, 1991: 361-369). En particular, es deseable mantener al mínimo el ruido y la vibración del motor. Para estudiar el efecto que tiene la marca del cojinete sobre la vibración del motor, se examinaron cinco marcas de cojinetes al instalar cada uno en muestras aleatorias de seis motores. Se registró la cantidad de vibración (medida en micras) mientras se tenía en funcionamiento cada motor. A continuación se muestran los datos de este estudio. Enuncie y pruebe las hipótesis pertinentes al nivel de significación 0.05 y después efectúe un análisis de comparaciones múltiples si es apropiado.

							Media
Marca 1	13.1	15.0	14.0	14.4	14.0	11.6	13.68
Marca 2	16.3	15.7	17.2	14.9	14.4	17.2	15.95
Marca 3	13.7	13.9	12.4	13.8	14.9	13.3	13.67
Marca 4	15.7	13.7	14.4	16.0	13.9	14.7	14.73
Marca 5	13.5	13.4	13.2	12.7	13.4	12.3	13.08

38. En un artículo de la revista científica británica *Nature* ("Sucrose Induction of Hepatic Hyperplasia in the Rat", 25 de agosto de 1972: 461) aparecen los datos de un estudio en el que se puso a dieta con un carbohidrato diferente a cinco grupos constituidos por seis ratas cada uno. Al finalizar el experimento, se determinó el contenido de DNA en el hígado de cada rata (mg/g de hígado), con los siguientes resultados:

Carbohidrato	$\bar{x}_{i\cdot}$
Almidón	2.58
Sacarosa	2.63
Fructosa	2.13
Glucosa	2.41
Maltosa	2.49

Si se supone también que $\sum\sum x_{ij}^2 = 183.4$, ¿los datos indican que el tipo de carbohidratos de la dieta afecta el contenido promedio verdadero de DNA? Construya una tabla ANOVA y utilice un nivel de significación 0.05.

39. Con base en el ejercicio 38, construyan un intervalo de confianza t para

$$\theta = \mu_1 - (\mu_2 + \mu_3 + \mu_4 + \mu_5)/4$$

que mida la diferencia entre el contenido promedio de DNA para la dieta de almidón, y el promedio combinado para las otras cuatro dietas. ¿El intervalo resultante incluye al cero?

40. Refiérase al ejercicio 38. ¿Cuál es β para la prueba cuando el promedio real del contenido promedio real de DNA es idéntico para tres de las dietas y cae por debajo de este valor común en una desviación estándar (σ) para las otras dos dietas?

41. De una gran población se seleccionan al azar cuatro laboratorios (1 al 4), y se pide a cada uno hacer tres determinaciones del porcentaje de alcohol metílico en especímenes de un compuesto tomado de un solo lote. Con base en los datos siguientes, ¿las diferencias entre laboratorios son una fuente de variación en el porcentaje de alcohol metílico? Establezca y pruebe las hipótesis pertinentes con un nivel de significación de 0.05.

1:	85.06	85.25	84.87
2:	84.99	84.28	84.88
3:	84.48	84.72	85.10
4:	84.10	84.55	84.05

42. La frecuencia crítica de parpadeo (fcp) es la máxima frecuencia (en ciclos por segundo) a la cual una persona puede detectar el parpadeo de una fuente luminosa. A frecuencias mayores arriba de la fcp, la fuente luminosa al parecer es continua, aunque en realidad esté parpadeando. Se realizó un estudio para ver si la fcp promedio real depende del color del iris y se obtuvieron los siguientes datos (basados en el artículo "The Effects of Iris Color on Critical Flicker Frequency", *J. General Psych.*, 1973: 91-95).

	Color del iris		
	1. Café	**2. Verde**	**3. Azul**
	26.8	26.4	25.7
	27.9	24.2	27.2
	23.7	28.0	29.9
	25.0	26.9	28.5
	26.3	29.1	29.4
	24.8		28.3
	25.7		
	24.5		
J_i	8	5	6
$x_{i\cdot}$	204.7	134.6	169.0
$\bar{x}_{i\cdot}$	25.59	26.92	28.17

$n = 19$ $\quad x_{\cdot\cdot} = 508.3$

a. Enuncie y pruebe las hipótesis pertinentes al nivel de significación 0.05 por medio de la tabla F para obtener un límite superior o inferior, o ambos, del valor P. (*Sugerencia:* $\sum\sum x_{ij}^2 = 13\,659.67$ y CF = 13 598.36.)

b. Investigue las diferencias entre colores de iris con respecto al promedio a la fcp promedio.

43. Sean $c_1, c_2, \ldots, c_I$ números que satisfacen $\sum c_i = 0$. Entonces $\sum c_i\mu_i = c_1\mu_1 + \cdots + c_I\mu_I$ se llama *contraste* de las μ_i. Observe que con $c_1 = 1$, $c_2 = -1$, $c_3 = \cdots = c_I = 0$, $\sum c_i\mu_i = \mu_1 - \mu_2$, lo que significa que toda diferencia por pares entre las μ_i es un contraste (también lo es, por ejemplo, $\mu_1 - 0.5\mu_2 - 0.5\mu_3$). Un método atribuido a Scheffé proporciona intervalos de confianza simultáneos con nivel de confianza simultáneo $100(1 - \alpha)\%$ para *todos* los contrastes posibles (¡un número infinito de ellos!). El intervalo para $\sum c_i\mu_i$ es

$$\sum c_i\bar{x}_i \pm (\sum c_i^2/J_i)^{1/2} \cdot [(I-1) \cdot \text{MSE} \cdot F_{\alpha,I-1,n-I}]^{1/2}$$

Por medio de los datos de frecuencia crítica de parpadeo del ejercicio 42, calcule los intervalos de Scheffé para los contrastes $\mu_1 - \mu_2$, $\mu_1 - \mu_3$, $\mu_2 - \mu_3$ y $0.5\mu_1 + 0.5\mu_2 - \mu_3$ (este último contraste compara el azul con el promedio del café y el verde). ¿Cuáles contrastes al parecer difieren significativamente de 0, y por qué?

44. Cuatro tipos de morteros, mortero de cemento ordinario (OCM), mortero impregnado de polímero (PIM), mortero de resina (RM) y mortero de cemento y polímero (PCM) se sometieron a una prueba de compresión para medir resistencia (MPa). Tres observaciones de resistencia para cada tipo de mortero aparecen en el artículo "Polymer Mortar Composite Matrices for Maintenance-Free Highly Durable Ferrocement" (*J. Ferrocement*, 1984: 337-345) y se reproducen aquí. Construya una tabla ANOVA. Con un nivel de significación 0.05 determine si los datos indican que la resistencia promedio real no es la misma para los cuatro tipos de mortero. Si determina que las resistencias promedio reales no son iguales, utilice el método de Tukey para identificar las diferencias importantes.

OCM	32.15	35.53	34.20
PIM	126.32	126.80	134.79
RM	117.91	115.02	114.58
PCM	29.09	30.87	29.80

45. Suponga que las x_{ij} están "codificadas" por $y_{ij} = cx_{ij} + d$. ¿Cómo se compara el valor de la estadística F calculada de las y_{ij} con el valor calculado de las x_{ij}? Justifique su afirmación.

46. En el ejemplo 10.10 reste $\bar{x}_i$ de cada observación en la i-ésima muestra ($i = 1, \ldots, 6$) para obtener un conjunto de 18 residuos. Después construya una gráfica de probabilidad normal y comente acerca de la posibilidad de la suposición de normalidad.

Bibliografía

Miller, Rupert, *Beyond ANOVA: The Basics of Applied Statistics*, Wiley, Nueva York, 1986. Una excelente fuente de información acerca de la comprobación de suposiciones y otros métodos de análisis.

Montgomery, Douglas, *Design and Analysis of Experiments* (5a. ed.), Wiley, Nueva York, 2001. Una presentación actualizada de modelos y metodología ANOVA.

Neter, John, William Wasserman y Michael Kutner, *Applied Linear Statistical Models* (4a. ed.), Irwin, Homewood, IL, 1996. La segunda mitad de este libro contiene un estudio muy bien presentado de ANOVA; el nivel es comparable al del presente texto, pero la explicación es más completa, lo que hace del libro una referencia excelente.

Ott, R. Lyman y Michael Longnecker. *An Introduction to Statistical Methods and Data Analysis* (5a. ed.), Duxbury Press, Belmont, CA, 2001. Incluye varios capítulos sobre metodología del análisis de varianza, que pueden encontrar provechosos los estudiantes que deseen una exposición no matemática; hay un buen capítulo acerca de diversos métodos de comparación múltiple.

11

Análisis de varianza con varios factores

Introducción

En el capítulo anterior, se utilizó el análisis de varianza (ANOVA) para probar la igualdad de I medias poblacionales diferentes o respuestas promedio reales relacionadas con I niveles distintos de un solo factor (que también se conocen como I tratamientos diferentes). En muchas situaciones experimentales, hay dos o más factores que son de interés simultáneo. En este capítulo se amplían los métodos del capítulo 10 para investigar esta clase de situaciones de varios factores.

En las dos primeras secciones, se centra la atención en el caso de dos factores de interés. Se utilizará I para denotar el número de niveles del primer factor (A) y J para denotar el número de niveles del segundo factor (B). Entonces hay IJ combinaciones posibles que consisten en un factor del nivel A y uno del factor B; cada combinación de esta clase se llama tratamiento, así que hay IJ tratamientos distintos. El número de observaciones hechas en el tratamiento (i, j) se denota mediante K_{ij}. En la sección 11.1 se presenta el modelo y análisis cuando $K_{ij} = 1$. Un caso especial importante de este tipo es un diseño de bloques aleatorizado, en el que un solo factor A es de interés primario, pero se crea otro factor "bloques", para controlar la variabilidad extraña en unidades o sujetos experimentales. En la sección 11.2, se atiende el caso $K_{ij} = K > 1$, y se mencionan de manera breve las dificultades relacionadas con K_{ij} distintas.

En la sección 11.3 se consideran experimentos en los que intervienen más de dos factores, entre otros un diseño de cuadrado latino, el cual controla los efectos de dos factores extraños que se considera influyen en la variable de respuesta. Cuando el número de factores es grande, un experimento que consiste en lo menos una observación para cada tratamiento, sería caro y tardado. Un caso especial importante, que se estudia en la sección 11.4, es aquel en el que hay p factores, cada uno de los cuales tiene dos niveles. Hay 2^p tratamientos distintos. Se considera tanto el caso en el que las observaciones se hacen en todos estos tratamientos (un diseño completo) y el caso en el que las observaciones se hacen sólo para un subconjunto seleccionado de tratamientos (un diseño incompleto).

11.1 ANOVA de dos factores con $K_{ij} = 1$

Cuando un factor A consiste en I niveles y un factor B consiste en J niveles, existen IJ combinaciones diferentes (pares) de niveles de los dos factores; a cada combinación se le llama tratamiento. Con K_{ij} = al número de observaciones en el tratamiento del factor A en el nivel i y del factor B en el nivel j, nos enfocaremos, en esta sección, en el caso $K_{ij} = 1$, de tal manera que los datos consistan en IJ observaciones. En primer lugar analizaremos el modelo de efectos fijos, en el cual los únicos niveles de interés para los dos factores son los que en realidad están representados en el experimento. Al final de la sección se estudia el caso en el que uno o los dos factores son aleatorios.

Ejemplo 11.1 ¿En realidad es fácil eliminar las marcas en las telas de plumas borrables como podría implicar la palabra *borrable*? Considere los datos siguientes de un experimento para comparar tres marcas de plumas y cuatro tratamientos de lavado con respecto a su capacidad para eliminar marcas en un tipo particular de tela (basado en "An Assesment of the Effects of Treatment, Time, and Heat on the Removal of Erasable Pen Marks from Cotton and Cotton/Polyester Blend Fabrics," *J. of Testing and Evaluation*, 1991: 394-397). La variable de respuesta es un indicador cuantitativo del cambio de color global del espécimen; mientras más bajo sea este valor, más marcas se eliminaron.

		Tratamiento de lavado				
		1	2	3	4	Total
Marca de la pluma	1	0.97	0.48	0.48	0.46	2.39
	2	0.77	0.14	0.22	0.25	1.38
	3	0.67	0.39	0.57	0.19	1.82
	Total	2.41	1.01	1.27	0.90	5.59

¿Hay alguna diferencia en la cantidad promedio real de cambio de color debida a las marcas de pluma diferentes o a los distintos tratamientos de lavado? ■

Como en el ANOVA de un solo factor, se utilizan subíndices dobles para identificar variables aleatorias y valores observados. Sea

X_{ij} = la variable aleatoria (*va*) que denota la medición cuando se mantiene el factor A al nivel i y el factor B se mantiene al nivel J

x_{ij} = el valor observado de X_{ij}

Las x_{ij} se presentan por lo común en una tabla bidireccional en la que el i-ésimo renglón contiene los valores observados cuando el factor A se mantiene al nivel i y la j-ésima columna contiene los valores observados cuando el factor B se mantiene al nivel j. En el experimento de la pluma borrable del ejemplo 11.1, el número de niveles observados del factor A es $I = 3$, el número de niveles del factor B es $J = 4$, $x_{13} = 0.48$, $x_{22} = 0.14$, etcétera.

Mientras que en el ANOVA de un solo factor se puso atención sólo a las medias de renglón y la gran media, aquí lo que interesa son las medias de columna. Sea

$$\overline{X}_{i\cdot} = \begin{array}{c}\text{el promedio de mediciones obtenidas cuando}\\ \text{el factor } A \text{ se mantiene al nivel } i\end{array} = \frac{\sum_{j=1}^{J} X_{ij}}{J}$$

$$\overline{X}_{\cdot j} = \begin{array}{c}\text{el promedio de mediciones obtenidas cuando}\\ \text{el factor } B \text{ se mantiene al nivel } j\end{array} = \frac{\sum_{i=1}^{I} X_{ij}}{I}$$

$$\overline{X}_{\cdot\cdot} = \text{la gran media} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}}{IJ}$$

con los valores observados $\overline{x}_{i\cdot}$, $\overline{x}_{\cdot j}$ y $\overline{x}_{\cdot\cdot}$. Los totales y no los promedios se denotan omitiendo la barra horizontal (así, $x_{\cdot j} = \sum_i x_{ij}$, etcétera). De manera intuitiva, para ver si hay algún efecto debido a los niveles del factor A, se deben comparar entre sí las $\overline{x}_{i\cdot}$ observadas, y la información acerca de los niveles diferentes del factor B deben provenir de las $\overline{x}_{\cdot j}$.

El modelo

Al proceder por analogía con el ANOVA de un solo factor, la primera inclinación al especificar un modelo es hacer que μ_{ij} = la respuesta promedio cuando el factor A está al nivel i y el factor B al nivel j, con lo que se obtienen IJ parámetros medios. Entonces sea

$$X_{ij} = \mu_{ij} + \epsilon_{ij}$$

donde ϵ_{ij} es la cantidad aleatoria porque el valor observado difiere de su esperanza y se supone que las ϵ_{ij} son normales e independientes con varianza común σ^2. Desafortunadamente, no hay procedimiento de prueba válido para esta elección de parámetros. La razón es que bajo la hipótesis alternativa de interés, las μ_{ij} pueden tomar cualquier valor, mientras que σ^2 puede ser cualquier valor mayor que cero, así que hay $IJ + 1$ parámetros que varían sin restricción. Pero sólo hay IJ observaciones, así que después de usar cada x_{ij} como una estimación de μ_{ij}, no hay manera de estimar σ^2.

Para rectificar este problema de un modelo con más parámetros que valores observados, debemos especificar un modelo realista que involucre relativamente pocos parámetros.

Suponga la existencia de I parámetros $\alpha_1, \alpha_2, \ldots, \alpha_I$ y J parámetros $\beta_1, \beta_2, \ldots, \beta_J$ tal que

$$X_{ij} = \alpha_i + \beta_j + \epsilon_{ij} \qquad (i = 1, \ldots, I, \quad j = 1, \ldots, J) \tag{11.1}$$

de modo que

$$\mu_{ij} = \alpha_i + \beta_j \tag{11.2}$$

Incluyendo a σ^2, hay ahora $I + J + 1$ parámetros de modelo, por lo tanto si $I \geq 3$ y $J \geq 3$, entonces habrá menos parámetros que observaciones (de hecho, en breve se modificará (11.2), así que incluso sean acomodados $I = 2$ o $J = 2$, o ambos).

El modelo especificado (11.1) y (11.2) se llama **modelo aditivo** porque cada respuesta media μ_{ij} es la suma de un efecto debido al factor A al nivel i (α_i) y un efecto debido al factor B al nivel j (β_j). La diferencia entre las respuestas medias para el factor A al nivel i y el nivel i' cuando B se mantiene al nivel j es $\mu_{ij} - \mu_{i'j}$. Cuando el modelo es aditivo,

$$\mu_{ij} - \mu_{i'j} = (\alpha_i + \beta_j) - (\alpha_{i'} + \beta_j) = \alpha_i - \alpha_{i'}$$

que es independiente del nivel j del segundo factor. Un resultado similar se cumple para $\mu_{ij} - \mu_{ij'}$. Por consiguiente, la aditividad significa que la diferencia en las respuestas medias para los dos niveles de uno de los factores es el mismo para todos los niveles del otro factor. En la figura 11.1(a) se muestra un conjunto de respuestas que satisfacen la condición de aditividad, y en la figura 11.1(b) se muestra una configuración no aditiva de respuestas medias.

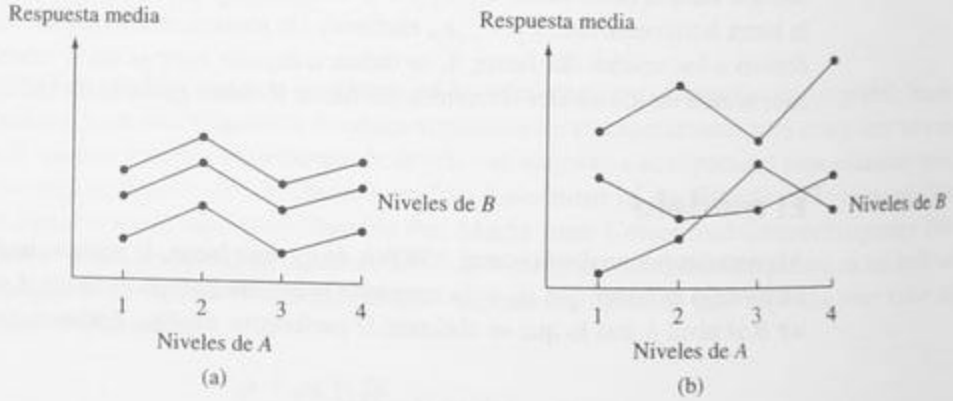


Figura 11.1 Respuestas medias para dos tipos de modelo: (a) aditiva; (b) no aditiva

Ejemplo 11.2 (Continuación del ejemplo 11.1) Cuando se localizan en una gráfica los valores observados de x_{ij} de una manera análoga a la de la figura 11.1, se obtiene el resultado mostrado en la figura 11.2. Aunque hay cierto "cruzamiento" en las x_{ij} observadas, la configuración es razonablemente representativa de lo que se esperaría bajo la aditividad con sólo una observación por tratamiento.

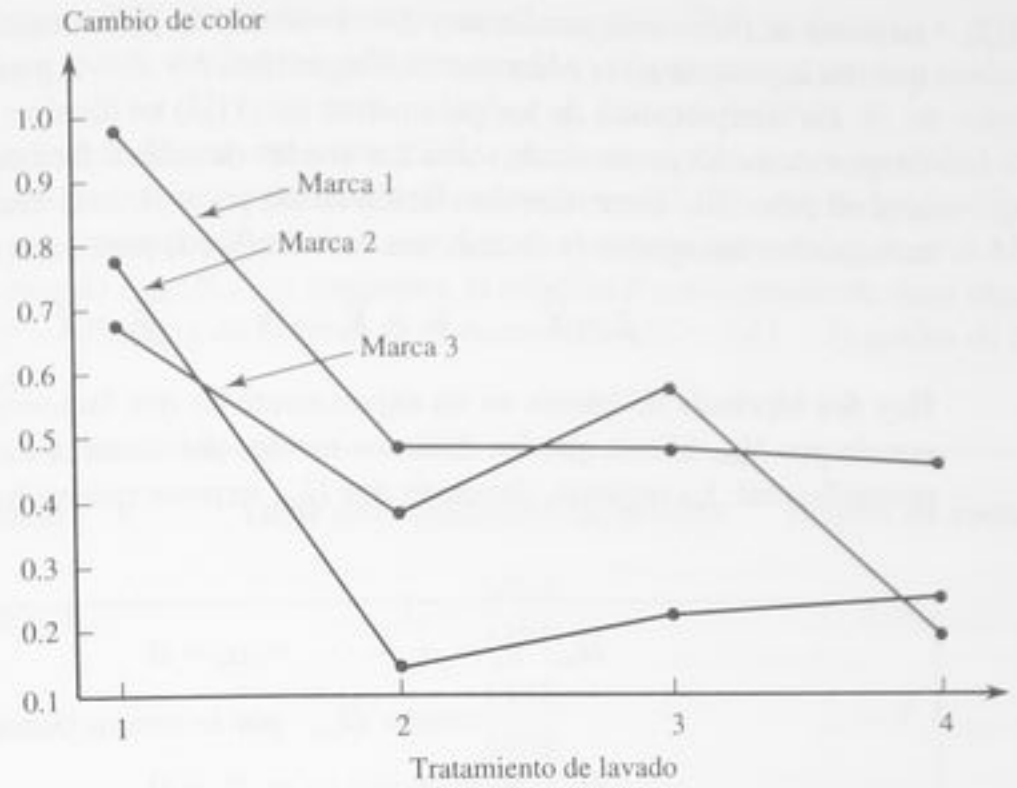


Figura 11.2 Gráfica de los datos del ejemplo 11.1 ■

La expresión (11.2) no necesariamente es la descripción final del modelo porque las α_i y las β_j no se determinan de manera única. A continuación se muestran dos configuraciones diferentes de las α_i y las β_j que producen las mismas μ_{ij} aditivas.

	$\beta_1 = 1$	$\beta_2 = 4$
$\alpha_1 = 1$	$\mu_{11} = 2$	$\mu_{12} = 5$
$\alpha_2 = 2$	$\mu_{21} = 3$	$\mu_{22} = 6$

	$\beta_1 = 2$	$\beta_2 = 5$
$\alpha_1 = 0$	$\mu_{11} = 2$	$\mu_{12} = 5$
$\alpha_2 = 1$	$\mu_{21} = 3$	$\mu_{22} = 6$

Al restar cualquier constante c de las α_i y sumar c a las β_j, se obtienen otras configuraciones correspondientes al mismo modelo aditivo. Esta no unicidad se elimina mediante el uso del modelo siguiente.

$$X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \tag{11.3}$$

donde $\sum_{i=1}^{I} \alpha_i = 0$, $\sum_{j=1}^{J} \beta_j = 0$, y se supone que las ϵ_{ij} son independientes, distribuidas normalmente, con media 0 y varianza común σ^2.

Esto es análogo a la elección alternativa de parámetros para el ANOVA de un solo factor estudiado en la sección 10.3. No es difícil comprobar que (11.3) es un modelo aditivo en el que los parámetros se determinan de manera única (por ejemplo, para las μ_{ij} mencionadas antes, $\mu = 4$, $\alpha_1 = -0.5$, $\alpha_2 = 0.5$, $\beta_1 = -1.5$, y $\beta_2 = 1.5$). Observe que sólo hay $I - 1$ valores de α_i deter-

minados de forma independiente y $J - 1$ valores de β_j determinados de la misma manera, así que (incluyendo a μ) la ecuación 11.3 especifica $I + J - 1$ parámetros medios.

La interpretación de los parámetros de (11.3) es directa: μ es la gran media verdadera (respuesta media promediada sobre los niveles de ambos factores), α_i es el efecto del factor A al nivel i (medido como una desviación desde μ), y β_j es el efecto del factor B al nivel j. Los estimadores insesgados (y de máxima verosimilitud) para estos parámetros son

$$\hat{\mu} = \overline{X}_{..} \qquad \hat{\alpha}_i = \overline{X}_{i.} - \overline{X}_{..} \qquad \hat{\beta}_j = \overline{X}_{.j} - \overline{X}_{..}$$

Hay dos hipótesis de interés en un experimento de dos factores con $K_{ij} = 1$. La primera, denotada por H_{0A}, afirma que los distintos niveles del factor A no tienen efecto en la respuesta promedio real. La segunda, denotada por H_{0B}, expresa que no hay efecto del factor B.

$$\begin{aligned} &H_{0A}\text{: } \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0 \\ &\qquad \text{contra } H_{aA}\text{: por lo menos una } \alpha_i \neq 0 \\ &H_{0B}\text{: } \beta_1 = \beta_2 = \cdots = \beta_J = 0 \\ &\qquad \text{contra } H_{aB}\text{: por lo menos una } \beta_j \neq 0 \end{aligned} \tag{11.4}$$

(Ningún efecto del factor A significa que las α_i sean iguales, así que todas deben ser cero, puesto que su suma es igual a cero, y de manera similar para las β_j.)

Procedimientos de prueba

La descripción y el análisis ahora siguen de cerca al del ANOVA de un solo factor. Las sumas de cuadrados pertinentes y sus formas de cálculo se determinan mediante

$$\begin{aligned} \text{SST} &= \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \overline{X}_{..})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}X_{ij}^2 - \frac{1}{IJ}X_{..}^2 \qquad \text{gl} = IJ - 1 \\ \text{SSA} &= \sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{X}_{i.} - \overline{X}_{..})^2 = \frac{1}{J}\sum_{i=1}^{I}X_{i.}^2 - \frac{1}{IJ}X_{..}^2 \qquad \text{gl} = I - 1 \\ \text{SSB} &= \sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{X}_{.j} - \overline{X}_{..})^2 = \frac{1}{I}\sum_{j=1}^{J}X_{.j}^2 - \frac{1}{IJ}X_{..}^2 \qquad \text{gl} = J - 1 \\ \text{SSE} &= \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}_{..})^2 \qquad \text{gl} = (I-1)(J-1) \end{aligned} \tag{11.5}$$

y la identidad fundamental

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSE} \tag{11.6}$$

permite determinar a SSE mediante sustracción.

Ejemplo 1
Continuación
ejemplo 11.2)

La expresión para SSE resulta de reemplazar μ, α_i y β_j en $\sum[X_{ij} - (\mu + \alpha_i + \beta_j)]^2$ por sus estimadores. Los grados de libertad del error son IJ − número de parámetros medios estimados $= IJ - [1 + (I - 1) + (J - 1)] = (I - 1)(J - 1)$. Con el ANOVA de un solo factor, una variación total se divide en una parte (SSE) que no se explica por la veracidad o falsedad de H_{0A} o H_{0B} y dos partes que se pueden explicar por la posible falsedad de las dos hipótesis nulas.

La teoría estadística ahora dice que si se forman relaciones F como en el ANOVA de un solo factor, cuando H_{0A} (H_{0B}) es verdadera, la relación F correspondiente tiene una distribución con $I - 1$ ($J - 1$) grados de libertad en el numerador e $(I - 1)(J - 1)$ grados de libertad en el denominador.

Hipótesis	Valor del estadístico de prueba	Región de rechazo
H_{0A} contra H_{aA}	$f_A = \dfrac{\text{MSA}}{\text{MSE}}$	$f_A \geq F_{\alpha,I-1,(I-1)(J-1)}$
H_{0B} contra H_{aB}	$f_B = \dfrac{\text{MSB}}{\text{MSE}}$	$f_B \geq F_{\alpha,J-1,(I-1)(J-1)}$

Ejemplo 11.3 (Continuación del ejemplo 11.2)

Las $x_{i\cdot}$ (totales de renglón) y las $x_{\cdot j}$ (totales de columna) para los datos de cambio de color se muestran a lo largo de los márgenes derecho e inferior de la tabla de datos proporcionada antes. Además, $\sum\sum x_{ij}^2 = 3.2987$ y el factor de corrección es $x_{\cdot\cdot}^2/(IJ) = (5.59)^2/12 = 2.6040$. Entonces, las sumas de cuadrados son

$$\text{SST} = 3.2987 - 2.6040 = 0.6947$$

$$\text{SSA} = \frac{1}{4}[(2.39)^2 + (1.38)^2 + (1.82)^2] - 2.6040 = 0.1282$$

$$\text{SSB} = \frac{1}{3}[(2.41)^2 + (1.01)^2 + (1.27)^2 + (0.90)^2] - 2.6040 = 0.4797$$

$$\text{SSE} = 0.6947 - (0.1282 + 0.4797) = 0.0868$$

En la tabla ANOVA siguiente (tabla 11.1) se resumen más cálculos.

Tabla 11.1 Tabla ANOVA para el ejemplo 11.3

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
Factor A (marca)	$I - 1 = 2$	SSA = 0.1282	MSA = 0.0641	$f_A = 4.43$
Factor B (tratamiento lavado)	$J - 1 = 3$	SSB = 0.4797	MSB = 0.1599	$f_B = 11.05$
Error	$(I - 1)(J - 1) = 6$	SSE = 0.0868	MSE = 0.01447	
Total	$IJ - 1 = 11$	SST = 0.6947		

El valor crítico para probar H_{0A} al nivel de significación 0.05 es $F_{0.05,2,6} = 5.14$. Puesto que $4.43 < 5.14$ no se puede rechazar H_{0A} al nivel de significación 0.05. El cambio de color promedio real al parecer no depende de la marca de la pluma. Debido a que $F_{0.05,3,6} = 4.76$ y $11.05 \geq 4.76$, H_{0B} se rechaza al nivel de significación 0.05 en favor de la afirmación de que el cambio de color varía con el tratamiento de lavado. Con un paquete de software para estadística se obtienen valores P de 0.066 y 0.007 para estas dos pruebas. ■

Cuadrados medios esperados

La posibilidad de usar las pruebas F recién descritas, se demuestra al calcular los cuadrados medios esperados. Después de un poco de álgebra,

$$E(\text{MSE}) = \sigma^2 \quad \text{(cuando el modelo es aditivo)}$$

$$E(\text{MSA}) = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I} \alpha_i^2$$

$$E(\text{MSB}) = \sigma^2 + \frac{I}{J-1}\sum_{j=1}^{J} \beta_j^2$$

Cuando H_{0A} es verdadera, MSA es un estimador insesgado de σ^2, así que F es una relación de dos estimadores insesgados de σ^2. Cuando H_{0A} es falsa, MSA tiende a sobrestimar σ^2, entonces se debe rechazar H_{0A} cuando la relación F_A es demasiado grande. Comentarios similares se aplican a MSB y H_{0B}.

Comparaciones múltiples

Cuando han sido rechazadas H_{0A} o H_{0B} se puede usar el procedimiento de Tukey para identificar diferencias importantes entre los niveles del factor bajo investigación. Los pasos del análisis son idénticos a los del ANOVA de un solo factor:

1. Para comparar niveles del factor A, obtenga $Q_{\alpha,IJ-1)(J-1)}$.
   Para comparar niveles del factor B, obtenga $Q_{\alpha,J(I-1)(J-1)}$.

2. Calcule:

$$w = Q \cdot (\text{desviación estándar estimada de las medias muestras que se comparan})$$

$$= \begin{cases} Q_{\alpha,IJ-1)(J-1)} \cdot \sqrt{\text{MSE}/J} & \text{para comparaciones del factor } A \\ Q_{\alpha,J(I-1)(J-1)} \cdot \sqrt{\text{MSE}/I} & \text{para comparaciones del factor } B \end{cases}$$

   (debido a que, p. ej., la desviación estándar de $\overline{X}_{i\cdot}$ es $\sigma/\sqrt{J}$).

3. Disponga en orden creciente las medias muestrales, subraye los pares que difieren por menos de w e identifique los pares no subrayados por la misma línea como correspondientes a niveles significativamente distintos del factor proporcionado.

Ejemplo 11.4 (Continuación del ejemplo 11.3)

La identificación de diferencias importantes entre los cuatro tratamientos de lavado requiere $Q_{0.05,4,6} = 4.90$ y $w = 4.90\sqrt{(0.01447)/3} = 0.340$. Las cuatro medias muestrales del factor B (promedios de columna) ahora se listan en orden creciente, y cualquier par que difiera en menos de 0.340 se subraya con segmento de recta:

$\overline{x}_{4\cdot}$	$\overline{x}_{2\cdot}$	$\overline{x}_{3\cdot}$	$\overline{x}_{1\cdot}$
0.300	0.337	0.423	0.803

El tratamiento de lavado 1 al parecer difiere en forma importante de los otros tres tratamientos, pero no se identifican otras diferencias importantes. En particular, no es evidente entre los tratamientos 2, 3 y 4 cuál es el mejor para eliminar marcas. ∎

Experimentos de bloques aleatorizados

Al usar el ANOVA de un solo factor para probar la presencia de efectos debidos a los I distintos tratamientos en estudio, una vez que han sido elegidos los IJ sujetos o unidades experimentales, los tratamientos se deben asignar en un modo por completo aleatorio. Es decir, se debe elegir al azar J sujetos para el primer tratamiento, luego otra muestra de J elegidos al azar de los $IJ - J$ individuos restantes para el segundo tratamiento, etcétera.

Sin embargo, sucede con frecuencia que los sujetos o unidades experimentales exhiben heterogeneidad con respecto a otras variables que podrían afectas las respuestas observadas. Cuando éste es el caso, la presencia o ausencia de un valor F significativo podría deberse a una variación extraña y no a la presencia o ausencia de efectos del factor. Ésta fue la razón de introducir un experimento por pares en el capítulo 9. La analogía con un experimento por pares cuando $I > 2$ se llama experimento **de bloques aleatorizado.** Se construye un factor extraño, "bloques", al dividir las IJ unidades en J grupos con I unidades en cada grupo. Este agrupamiento se hace de tal forma que dentro de cada bloque, las I unidades son homogéneas con respecto a otros factores que se considera afectan las respuestas. Entonces dentro de cada bloque homogéneo, los I tratamientos se asignan al azar a las I unidades u objetos del bloque.

Ejemplo 11.5 Una organización que prueba productos para el consumidor, desea comparar el consumo de energía anual para cinco marcas de deshumidificadores. Debido a que el consumo de energía depende del nivel de humedad predominante, se decidió monitorear cada marca en cuatro niveles que varían de humedad moderada a extrema (por lo tanto, se construyen bloques del nivel de humedad). Dentro de cada nivel, las marcas se asignaron al azar para los cinco lugares seleccionados. La cantidad resultante de consumo de energía (kWh anuales) aparece en la tabla 11.2.

Tabla 11.2 Datos de consumo de energía para el ejemplo 11.5

Tratamientos (marcas)	Bloques (nivel de humedad) 1	2	3	4	$x_{i\cdot}$	$\bar{x}_{i\cdot}$
1	685	792	838	875	3190	797.50
2	722	806	893	953	3374	843.50
3	733	802	880	941	3356	839.00
4	811	888	952	1005	3656	914.00
5	828	920	978	1023	3749	937.25
$x_{\cdot j}$	3779	4208	4541	4797	17 325	

Puesto que $\sum\sum x_{ij}^2 = 15\ 178\ 901.00$ y $x_{\cdot\cdot}^2/(IJ) = 15\ 007\ 781.25$,

$$\text{SST} = 15\ 178\ 901.00 - 15\ 007\ 781.25 = 171\ 119.75$$

$$\text{SSA} = \frac{1}{4}[60\ 244\ 049] - 15\ 007\ 781.25 = 53\ 231.00$$

$$\text{SSB} = \frac{1}{5}[75\ 619\ 995] - 15\ 007\ 781.25 = 116\ 217.75$$

y

$$\text{SSE} = 171\ 119.75 - 53\ 231.00 - 116\ 217.75 = 1671.00$$

Los cálculos del ANOVA se resumen en la tabla 11.3.

Tabla 11.3 Tabla ANOVA para el ejemplo 11.5

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
Tratamientos (marcas)	4	53 231.00	13 307.75	$f_A = 95.57$
Bloques	3	116 217.75	38 739.25	$f_B = 278.20$
Error	12	1671.00	139.25	
Total	19	171 119.75		

Puesto que $F_{0.05,4,12} = 3.26$ y $f_A = 95.57 \geq 3.26$, se rechaza H_0 en favor de H_a, y se concluye que el consumo de energía eléctrica no depende de la marca del humidificador. Para identificar marcas que difieran de forma importante, se utiliza el procedimiento de Tukey. $Q_{0.05,5,12} = 4.51$ y $w = 4.51\sqrt{139.25/4} = 26.6$.

$\bar{x}_{1.}$	$\bar{x}_{3.}$	$\bar{x}_{2.}$	$\bar{x}_{4.}$	$\bar{x}_{5.}$
797.50	839.00	843.50	914.00	937.25

El subrayado indica que las marcas se pueden dividir en tres grupos con respecto al consumo de energía.

Debido a que el factor de bloque es de interés secundario, $F_{0.05,3,12}$ no es necesario, aunque es evidente que el valor calculado de F_B es muy significativo. En la figura 11.3 se muestra el resultado de SAS para estos datos. Observe que en la primera parte de la tabla de ANOVA, las sumas de cuadrados (SS) para tratamientos (marcas) y bloques (niveles de humedad) se combinan en una sola SS "modelo".

```
                    Análisis del procedimiento de varianza
Variable dependiente: USO DE ENERGÍA
                                  Suma de          Cuadrado
Fuente                    GL      cuadrados          medio          Valor F      Pr > F
Modelo                     7     169448.750      24206.964           173.84      0.0001
Error                     12       1671.000        139.250
Total corregido           19     171119.750

          Cuadrado de R      V.C.          Raíz de MSE      USO DE ENERGÍA     promedio
               0.990235      1.362242          11.8004          866.25000

Fuente                    GL      SS de ANOVA    Cuadrado medio      Valor F      PR > F
MARCA                      4      53231.000      13307.750             95.57      0.0001
HUMEDAD                    3     116217.750      38739.250            278.20      0.0001

                Alfa = 0.05  gl = 12  MSE = 139.25
          Valor crítico del intervalo estudentizado = 4.508
              Diferencia significativa mínima = 26.597

Las medias con la misma letra no son significantes diferentes

Agrupamiento de Tukey        Media      N      Marca
                A            937.250    4      5
                A
                A            914.000    4      4
                B            843.500    4      2
                B
                B            839.000    4      3
                C            797.500    4      1
```

Figura 11.3 Resultado de SAS para los datos de consumo de energía

En muchas situaciones experimentales en las que se aplican tratamientos a los individuos, un solo individuo puede recibir los I tratamientos. La construcción de bloques suele hacerse entonces en los mismos individuos para controlar la variabilidad entre individuos; por lo tanto, se dice que cada individuo actúa como su propio control. A veces los científicos sociales se refieren a esta clase de experimentos como diseños de medidas repetidas. Las "unidades" dentro de un bloque son entonces las "instancias" diferentes de aplicación del tratamiento. De manera similar, los bloques con frecuencia se toman como periodos, lugares u observadores distintos.

Ejemplo 11.6 Los datos de la tabla 11.4 son del artículo "Compounding of Discriminative Stimuli from the Same and Different Sensory Modalities" (*J. Experimental Analysis Behaviour*, 1971:337-342). La respuesta de la rata se mantuvo mediante programas de intervalo fijo de reforzamiento en presencia de un tono o dos luces separadas. Las luces fueron ya sea de intensidad moderada (L1) o baja (L2). Las observaciones se dan como el número medio de respuestas que emite cada sujeto durante las presentaciones de estímulos simples o compuestos en un periodo de cuatro días.

Tabla 11.4 Datos de respuesta para el ejemplo 11.6

	Sujeto					
Estímulo	**1**	**2**	**3**	**4**	$x_{i\cdot}$	$\bar{x}_{i\cdot}$
L1	8.0	17.3	52.0	22.0	99.3	24.8
L2	6.9	19.3	63.7	21.6	111.5	27.9
Tono (T)	9.3	18.8	60.0	28.3	116.4	29.1
L1 + L2	9.2	24.9	82.4	44.9	161.4	40.3
L1 + T	12.0	31.7	83.8	37.4	164.9	41.2
L2 + T	9.4	33.6	96.6	40.6	180.2	45.1
$x_{\cdot j}$	54.8	145.6	438.5	194.8	833.7	

Con $\sum\sum x_{ij}^2 = 44\,614.21$, SST = 15 653.56, SSA = 1428.28, SSB = 13 444.63 y SSE = 780.65. Además los cálculos se resumen en la tabla 11.5.

Tabla 11.5 Tabla ANOVA para el ejemplo 11.6

Fuente de variación	**gl**	**Suma de cuadrados**	**Cuadrado medio**	f
Estímulo (A)	5	1428.28	285.66	$f_A = 5.49$
Sujetos (B)	3	13 444.63	4481.54	$f_B = 86.12$
Error	15	780.65	52.04	
Total	23	15 653.56		

Puesto que $F_{0.05,5,15} = 2.90$ y $5.49 \geq 2.90$, se concluye que hay diferencias en las respuestas promedio reales relacionadas con estímulos diferentes. Para el procedimiento de Tukey, $w = 4.59\sqrt{52.04/4} = 16.56$.

$\bar{x}_{1\cdot}$	$\bar{x}_{2\cdot}$	$\bar{x}_{3\cdot}$	$\bar{x}_{4\cdot}$	$\bar{x}_{5\cdot}$	$\bar{x}_{6\cdot}$
24.8	27.9	29.1	40.3	41.2	45.1

Así que tanto L1 como L2 son significativamente distintas de L2 + T, y no hay otras diferencias significativas entre los estímulos. ■

En la mayor parte de los experimentos de bloques aleatorizados en los que los individuos sirven como bloques, los individuos que en realidad participan en el experimento se eligen de una gran población. Los individuos contribuyen entonces a efectos aleatorios y no a efectos fijos. Esto no afecta al procedimiento para comparar tratamientos cuando $K_{ij} = 1$ (una observación por "celda", como en esta sección), pero el procedimiento se modifica si $K_{ij} = K > 1$. En breve se consideran modelos de dos factores en los que los efectos son aleatorios.

Más acerca del agrupamiento en bloques Cuando $I = 2$, se puede usar la prueba F o la prueba t de diferencias por pares para analizar los datos. La conclusión resultante no dependerá de qué procedimiento se use, puesto que $T^2 = F$ y $t^2_{\alpha/2,\nu} = F_{\alpha,1,\nu}$.

Al igual que con los pares, el agrupamiento en bloques conlleva una posible ganancia como una posible pérdida de precisión. Si en las unidades experimentales hay una gran cantidad de heterogeneidad, será grande el valor del parámetro de varianza σ^2 en un modelo unidireccional. El efecto del agrupamiento en bloques es filtrar la variación representada por σ^2 en el modelo bidireccional apropiado para un experimento de bloques aleatorizado. Si otras cosas permanecen iguales, un valor más pequeño de σ^2 da como resultado una prueba con mayor posibilidad de detectar desviaciones de H_0 (es decir, una prueba con mayor potencia).

Sin embargo, otras cosas no son iguales aquí, puesto que la prueba F de un solo factor se basa en $I(J - 1)$ grados de libertad (gl) para el error, mientras que la prueba F de dos factores se basa en $(I - 1)(J - 1)$ gl para el error. Menos grados de libertad para el error dan como resultado una menor potencia, en esencia debido a que el estimador del denominador de σ^2 no es tan preciso. Esta pérdida de grados de libertad puede ser grave, en particular si el experimentador suministra sólo un pequeño número de observaciones. No obstante, si al parecer el agrupamiento en bloques reduce la variabilidad de manera importante, es probable que valga la pena la pérdida de grados de libertad.

Modelos para efectos aleatorios

En muchos experimentos, los niveles reales de un factor utilizado en el experimento, en vez de ser los únicos de interés para el experimentador, se seleccionan de una población mucho más grande de niveles posibles del factor. En una situación de dos factores, cuando éste es el caso para ambos factores, es apropiado un **modelo de efectos aleatorios.** El caso en el que los niveles de un factor son los únicos de interés y los niveles del otro factor se seleccionan de una población de niveles origina un **modelo de efectos mixtos.** El modelo de efectos aleatorios de dos factores cuando $K_{ij} = 1$ es

$$X_{ij} = \mu + A_i + B_j + \epsilon_{ij} \qquad (i = 1, \ldots, I, \quad j = 1, \ldots, J)$$

donde las A_i, B_j y ϵ_{ij} son *va* independientes, distribuidas normalmente con media 0 y varianzas σ^2_A, σ^2_B y σ^2, respectivamente. Entonces las hipótesis de interés son H_{0A}: $\sigma^2_A = 0$ (el nivel del factor A contribuye a la variación en la respuesta) contra H_{aA}: $\sigma^2_A > 0$ y H_{0B}: $\sigma^2_B = 0$ con-

tra H_{aB}: $\sigma_B^2 > 0$. Mientras que $E(\text{MSE}) = \sigma^2$ como antes, los cuadrados medios esperados para los factores A y B ahora son

$$E(\text{MSA}) = \sigma^2 + J\sigma_A^2 \qquad E(\text{MSB}) = \sigma^2 + I\sigma_B^2$$

Así, cuando H_{0A} (H_{0B}) es verdadera, F_A (F_B) todavía es una relación de dos estimadores insesgados de σ^2. Se puede demostrar que una prueba de nivel α para H_{0A} contra H_{aA} aún rechaza a H_{0A} si $f_A \geq F_{\alpha,I-1,(I-1)(J-1)}$ y, de manera similar, se utiliza el mismo procedimiento que antes para decidir entre H_{0B} y H_{aB}.

Para el caso cuando el factor A es fijo y el factor B es aleatorio, el modelo mixto es

$$X_{ij} = \mu + \alpha_i + B_j + \epsilon_{ij} \qquad (i = 1, \ldots, I, \quad j = 1, \ldots, J)$$

donde $\sum \alpha_i = 0$ y las B_j y las ϵ_{ij} tienen una distribución normal con media 0 y varianzas σ_B^2 y σ^2, respectivamente. Ahora las dos hipótesis nulas son

$$H_{0A}\text{: } \alpha_1 = \cdots = \alpha_I = 0 \qquad \text{y} \qquad H_{0B}\text{: } \sigma_B^2 = 0$$

con cuadrados medios esperados

$$E(\text{MSE}) = \sigma^2 \qquad E(\text{MSA}) = \sigma^2 + \frac{J}{I-1}\sum \alpha_i^2 \qquad E(\text{MSB}) = \sigma^2 + I\sigma_B^2$$

Los procedimientos de prueba para H_{0A} contra H_{aA} y H_{0B} contra H_{aB} son de la misma forma que antes. Por ejemplo, en el análisis de los datos de cambio de color del ejemplo 11.1, si se eligieran al azar los cuatro tratamientos de lavado, entonces debido a que $f_B = 11.05$ y $F_{0.05,3,6} = 4.76$, H_{0B}: $\sigma_B^2 = 0$ se rechaza en favor de H_{aB}: $\sigma_B^2 > 0$. Una estimación del "componente de varianza" σ_B^2 se determina entonces mediante $(\text{MSB} - \text{MSE})/I = 0.0485$.

En resumen, cuando $K_{ij} = 1$, aunque las hipótesis y los cuadrados medios esperados difieran del caso en que ambos efectos son fijos, los procedimientos de prueba son idénticos.

Ejercicios | Sección 11.1 (1 a 15)

1. Se determinó el número de millas (en miles) de vida útil de llantas, midiendo el desgaste de la banda de rodadura para cada una de cinco marcas de automóviles subcompactos (factor A, con $I = 5$) en combinación con cada una de cuatro marcas de llantas radiales (factor B, con $J = 4$), obteniéndose como resultado $IJ = 20$ observaciones. Luego, se calcularon los valores SSA = 30.6, SSB = 44.1 y SSE = 59.2. Suponga que es apropiado un modelo aditivo.

   a. Pruebe H_0: $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$ (ninguna diferencia en la duración promedio de la llanta debida a las marcas de automóviles) contra H_a: por lo menos $\alpha_i \neq 0$ con una prueba de nivel 0.05.

   b. H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (ninguna diferencia en la duración promedio real de las llantas debido a las marcas de llantas) contra H_a: por lo menos una $\beta_j \neq 0$ con una prueba de nivel 0.05.

2. Están en consideración cuatro recubrimientos distintos para proteger tubería metálica contra la corrosión. La tubería será enterrada en cuatro tipos de suelo. Para investigar si la cantidad de corrosión depende del recubrimiento o del tipo de suelo, se seleccionan 12 piezas de tubería. Cada pieza se recubre con uno de los cuatro recubrimientos y se entierra en uno de los tres tipos de suelo durante un tiempo fijo, después del cual se determina la cantidad de corrosión (profundidad de picaduras máximas, en 0.0001 pulg.)

		Tipo de suelo (*B*)		
		1	2	3
Recubrimiento (*A*)	1	64	49	50
	2	53	51	48
	3	47	45	50
	4	51	43	52

a. Suponiendo la validez del modelo aditivo, realice el análisis ANOVA por medio de una tabla ANOVA para ver si la cantidad de corrosión depende del tipo de recubrimiento utilizado o del tipo de suelo. Utilice $\alpha = 0.05$.

b. Calcule $\hat{\mu}$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_3$, $\hat{\alpha}_4$, $\hat{\beta}_1$, $\hat{\beta}_2$ y $\hat{\beta}_3$.

3. En el artículo "Adiabatic Humidification of Air with Water in a Packed Tower" (*Chem. Eng. Prog.*, 1952: 362-370) aparecen los datos acerca del coeficiente de transferencia de calor de película de gas (Btu/hr pie^2 en °F) como función del flujo de gas (factor *A*) y el flujo de líquido (factor *B*).

		B			
		1(190)	2(250)	3(300)	4(400)
A	1(200)	200	226	240	261
	2(400)	278	312	330	381
	3(700)	369	416	462	517
	4(1100)	500	575	645	733

a. Después de construir una tabla de ANOVA, pruebe al nivel 0.01 tanto la hipótesis de ningún efecto del flujo de gas contra la alternativa apropiada, como la hipótesis de ningún efecto del flujo de líquido contra la alternativa apropiada.

b. Utilice el procedimiento de Tukey para investigar diferencias en el coeficiente de transferencia de calor esperado debidas a diferentes flujos de gas.

c. Repita la parte (b) para el flujo del líquido.

4. En un experimento para ver si la cantidad de cobertura de pintura de látex de color azul tenue para interiores depende de la marca de la pintura o la marca del rodillo utilizado, se aplicó un galón de cada una de cuatro marcas de pintura con cada una de tres marcas de rodillo, obteniéndose los datos siguientes (número de pies cuadrados cubiertos).

		Marca del rodillo		
		1	2	3
Pintura Marca	1	454	446	451
	2	446	444	447
	3	439	442	444
	4	444	437	443

a. Construya la tabla ANOVA. [*Sugerencia*: los cálculos se agilizan al restar 400 (o cualquier otro número conveniente) de cada observación. Esto no afecta los resultados finales.]

b. Exprese y pruebe las hipótesis apropiadas para decidir si la marca de pintura tiene algún efecto en la cobertura. Utilice $\alpha = 0.05$.

c. Repita el inciso (b) para la marca del rodillo.

d. Utilice el método de Tukey para identificar diferencias entre las marcas. ¿Hay alguna marca que parezca claramente preferible a las demás?

5. En un experimento para evaluar el efecto del ángulo de tiro en la fuerza requerida para causar la separación en conectores eléctricos, se emplearon cuatro ángulos distintos (factor *A*) y se tiró de cada uno de los cinco conectores de una muestra (factor *B*) ("A Mixed Model Factorial Experiment in Testing Electrical Connectors", *Industrial Quality Control*, 1960: 12-16). Los datos se muestran en la tabla siguiente.

		B				
		1	2	3	4	5
A	0°	45.3	42.2	39.6	36.8	45.8
	2°	44.1	44.1	38.4	38.0	47.2
	4°	42.7	42.7	42.6	42.2	48.9
	6°	43.5	45.8	47.9	37.9	56.4

¿Los datos indican que el ángulo de tiro afecta la fuerza de separación promedio real? Exprese y pruebe las hipótesis apropiadas al nivel 0.01 construyendo primero una tabla ANOVA (SST = 396.13, SSA = 58.16 y SSB = 246.97).

6. Un determinado condado emplea tres asesores quienes deben determinar el valor de la propiedad residencial en el condado. Para ver si estos asesores difieren de forma sistemática en sus estimaciones, se eligen cinco casas, y se pide a cada asesor que determine el valor de mercado de cada casa. Con el factor *A* que denota los asesores ($I = 3$) y el factor *B* que denota las casas ($J = 5$), suponga que SSA = 11.7, SSB = 113.5 y SSE = 25.6.

a. Pruebe H_0: $\alpha_1 = \alpha_2 = \alpha_3 = 0$ al nivel 0.05. (H_0 afirma que no hay diferencias sistemáticas entre los asesores.)

b. Explique por qué se utilizó un experimento de bloques aleatorizado con sólo cinco casas en vez de un experimento ANOVA en el que interviene un total de 15 casas distintas donde se pide a cada asesor que evalúe cinco casas diferentes (un grupo de cinco para cada asesor).

7. En el artículo "Rate of Suttering Adaptation Under Two Electro-Shock Conditions" (*Behavior Research Therapy*, 1967: 49-54) se dan las puntuaciones de adaptación para tres tratamientos distintos: (1) sin choque, (2) choque después de cada palabra tartamudeada y (3) choque durante cada momento de tartamudeo. Estos tratamientos se emplearon en cada uno de 18 tartamudos.
   a. Las estadísticas de resumen son $x_{1\cdot} = 905$, $x_{2\cdot} = 913$, $x_{3\cdot} = 936$, $x_{\cdot\cdot} = 2754$, $\sum_j x_{\cdot j}^2 = 430\,295$ y $\sum\sum x_{ij}^2 = 143\,930$. Construya la tabla ANOVA y pruebe al nivel 0.05 para ver si la puntuación de adaptación promedio real depende del tratamiento suministrado.
   b. A juzgar por la relación F para los individuos (factor B), ¿considera que agrupar en bloques a los individuos resultó efectivo en este experimento? Explique.

8. En la tabla siguiente se da la concentración de epinefrina plasmática para diez individuos experimentales durante la anestesia con (1) isoflurano, (2) halotano y (3) ciclopropano ("Sympathoadrenal and Hemodynamic Effects of Isoflurane, Halothane, and Ciclopropane in Dogs", *Anesthesiology*, 1974: 465-470).
   a. ¿La elección del anestésico afecta la concentración promedio real? Pruebe H_0: $\alpha_1 = \alpha_2 = \alpha_3 = 0$ al nivel 0.05 después de construir la tabla ANOVA.

		Sujeto (B)				
		1	**2**	**3**	**4**	**5**
Anestésico (A)	**1**	0.28	0.51	1.00	0.39	0.29
	2	0.30	0.39	0.63	0.38	0.21
	3	1.07	1.35	0.69	0.28	1.24

		6	**7**	**8**	**9**	**10**
Anestésico (A)	**1**	0.36	0.32	0.69	0.17	0.33
	2	0.88	0.39	0.51	0.32	0.42
	3	1.53	0.49	0.56	1.02	0.30

$\sum\sum x_{ij}^2 = 13.7980$

   b. Utilice el procedimiento de Tukey para investigar diferencias importantes entre los anestésicos.

9. En el artículo "The Effects of a Pneumatic Stool and a One-Legged Stool on Lower Limb Joint Load and Muscular Activity During Sitting and Rising" (*Ergonomics*, 1993: 519-535) aparecen los datos siguientes acerca del esfuerzo requerido de un individuo para levantarse de cuatro tipos de taburete (escala de Borg). Lleve a cabo un análisis de varianza con $\alpha = 0.05$, y a continuación realice un análisis de comparaciones múltiples si es apropiado.

		Individuo									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	$\bar{x}_{i\cdot}$
Tipo de taburete	**1**	12	10	7	7	8	9	8	7	9	8.56
	2	15	14	14	11	11	11	12	11	13	12.44
	3	12	13	13	10	8	11	12	8	10	10.78
	4	10	12	9	9	7	10	11	7	8	9.22

10. La resistencia del concreto utilizado en la construcción comercial tiende a variar de un lote a otro. Por consiguiente, pequeños cilindros de prueba de concreto muestreados de un lote se "curan" por periodos de hasta unos 28 días en condiciones controladas de temperatura y humedad antes de hacer las mediciones de resistencia. Luego, el concreto se "compra y vende con base en los cilindros de prueba de resistencia" (ASTM C 31 Standard Test Method for Making and Curing Concrete Test Specimens in the Field). Los datos siguientes se obtuvieron de un experimento realizado para comparar tres métodos de curación diferentes con respecto a la resistencia compresiva (MPa). Analice estos datos.

Lote	**Método A**	**Método B**	**Método C**
1	30.7	33.7	30.5
2	29.1	30.6	32.6
3	30.0	32.2	30.5
4	31.9	34.6	33.5
5	30.5	33.0	32.4
6	26.9	29.3	27.8
7	28.2	28.4	30.7
8	32.4	32.4	33.6
9	26.6	29.5	29.2
10	28.6	29.4	33.2

11. Los residuos de un ANOVA de dos factores con $K_{ij} = 1$ son las cantidades $x_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j)$. Una gráfica de probabilidad normal de estos residuos se puede usar como comprobación de la suposición de normalidad. Construya este tipo de gráfica para los datos del ejemplo 11.1 y comente.

12. Suponga que en el experimento descrito en el ejercicio 6 las cinco casas en realidad hubieran sido seleccionadas al azar de entre las de cierta edad y tamaño, de modo que el factor B sea aleatorio y no fijo. Pruebe H_0: $\sigma_\beta^2 = 0$ contra H_a: $\sigma_\beta^2 > 0$ por medio de una prueba de nivel 0.01.

13. a. Muestre que se puede agregar una constante d (o restar) de cada x_{ij} sin afectar ninguna de las sumas de cuadrados de ANOVA.
    b. Suponga que cada x_{ij} se multiplica por una constante diferente de cero c. ¿Cómo afecta esto a las sumas de

cuadrados del ANOVA? ¿cómo afecta lo anterior a los valores de los estadísticos F, F_A y F_B? ¿Qué efectos tiene "codificar" los datos mediante $y_{ij} = cx_{ij} + d$ en las conclusiones que resultan de los procedimientos de ANOVA?

14. Utilice el hecho de que $E(X_{ij}) = \mu + \alpha_i + \beta_j$ con $\sum\alpha_i = \sum\beta_j = 0$ para mostrar que $E(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}) = \alpha_i$, de modo que $\hat{\alpha}_i = \bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}$ es un estimador insesgado para α_i.

15. Las curvas de potencia de las figuras 10.5 y 10.6 se pueden usar para obtener $\beta = P$(error de tipo II) para prueba F en el ANOVA de dos factores. Para valores fijos de $\alpha_1, \alpha_2, \ldots, \alpha_I$ se calcula la cantidad $\phi^2 = (J/I)\sum\alpha_i^2/\sigma^2$. Luego, se introduce la cifra correspondiente a $\nu_1 = I - 1$ en el eje horizontal en el valor de ϕ, la potencia se lee en el eje vertical de la curva marcada $\nu_2 = (I - 1)(J - 1)$ y $\beta = 1 -$ potencia.

   a. Para el experimento de corrosión descrito en el ejercicio 2, determine β cuando $\alpha_1 = 4$, $\alpha_2 = 0$, $\alpha_3 = \alpha_4 = -2$, y $\sigma = 4$. Realice los mismo para $\alpha_1 = 6$, $\alpha_2 = 0$, $\alpha_3 = \alpha_4 = -3$, y $\sigma = 4$.

   b. Por simetría, ¿cuál es β para la prueba de H_{oB} contra H_{aB} en el ejemplo 11.1 cuando $\beta_1 = 0.3$, $\beta_2 = \beta_3 = \beta_4 = -0.1$ y $\sigma = 0.3$?

11.2 ANOVA de dos factores con $K_{ij} > 1$

En la sección 11.1, se analizan datos de un experimento de dos factores en el que hubo una observación para cada una de las IJ combinaciones de niveles de los dos factores. Para obtener procedimientos de prueba válidos, se supuso que las μ_{ij} tienen una estructura aditiva con $\mu_{ij} = \mu + \alpha_i + \beta_j$, $\sum\alpha_i = \sum\beta_j = 0$. La aditividad significa que la diferencia en las respuestas promedio reales para dos niveles cualesquiera de los factores es el mismo para cada nivel del otro factor. Por ejemplo, $\mu_{ij} - \mu_{i'j} = (\mu + \alpha_i + \beta_j) - (\mu + \alpha_{i'} + \beta_j) = \alpha_i - \alpha_{i'}$, independiente del nivel j del segundo factor. Esto se ilustra en la figura 11.1(a), en la que las rectas que unen a las respuestas promedio reales son paralelas.

En la figura 11.1(b) se ilustra un conjunto de respuestas promedio reales que no tienen estructura aditiva. Las rectas que unen a estas μ_{ij} no son paralelas, lo cual significa que la diferencia en las respuestas promedio reales para los diferentes niveles de un factor no dependen del nivel del otro factor. Cuando no se cumple la aditividad, se dice que hay una **interacción** entre los distintos niveles de los factores. En la sección 11.1 la suposición de aditividad permitió obtener un estimador de la varianza del error aleatorio σ^2 (MSE) que fue insesgado ya sea si fue verdadera o no la hipótesis nula de interés. Cuando $K_{ij} > 1$ para por lo menos un par (i, j), se puede obtener un estimador válido de σ^2 sin suponer aditividad. Al especificar el modelo apropiado y obtener los procedimientos de prueba, se centrará la atención en el caso $K_{ij} = K > 1$, así que el número de observaciones por "celda" (para cada combinación de niveles) es constante.

Parámetros para el modelo de efectos fijos con interacción

En vez de usar las μ_{ij} como parámetros de modelo, se acostumbra usar un conjunto equivalente que revela de forma más clara el papel de la interacción. Sea

$$\mu = \frac{1}{IJ}\sum_i\sum_j \mu_{ij} \qquad \mu_{i\cdot} = \frac{1}{J}\sum_j \mu_{ij} \qquad \mu_{\cdot j} = \frac{1}{I}\sum_i \mu_{ij} \tag{11.7}$$

Así, μ es la respuesta esperada promediada en los niveles de ambos factores (la gran media verdadera), $\mu_{i\cdot}$ es la respuesta esperada promediada en los niveles del segundo factor cuando el primer factor A se mantiene al nivel i, y de manera similar para $\mu_{\cdot j}$. Ahora defina

$$\alpha_i = \mu_{i\cdot} - \mu = \text{el efecto del factor } A \text{ al nivel } i$$
$$\beta_j = \mu_{\cdot j} - \mu = \text{el efecto del factor } B \text{ al nivel } j \qquad (11.8)$$
$$\gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j)$$

de donde

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \qquad (11.9)$$

El modelo es aditivo si y sólo si las $\gamma_{ij} = 0$. Las γ_{ij} se conocen como **parámetros de interacción**. Las α_i se llaman **efectos principales para el factor *A***, y las β_j son los **efectos principales para el factor *B***. Aunque hay I α_i, J β_j e IJ γ_{ij} además de μ, las condiciones $\sum\alpha_i = 0$, $\sum\beta_j = 0$, $\sum_j\gamma_{ij} = 0$ para cualquier i, y $\sum_i\gamma_{ij} = 0$ para cualquier j [todo en virtud de (11.7) y (11.8)], implica que sólo IJ de estos nuevos parámetros se determinan de modo independiente: μ, $I - 1$ de las α_i, $J - 1$ de las β_j e $(I - 1)(J - 1)$ de las γ_{ij}.

Ahora hay tres conjuntos de hipótesis que serán considerados:

$$H_{0AB}\text{: } \gamma_{ij} = 0 \text{ para toda } i, j \quad \text{contra} \quad H_{aAB}\text{: por lo menos una } \gamma_{ij} \neq 0$$
$$H_{0A}\text{: } \alpha_1 = \cdots = \alpha_I = 0 \quad \text{contra} \quad H_{aA}\text{: por lo menos una } \alpha_i \neq 0$$
$$H_{0B}\text{: } \beta_1 = \cdots = \beta_J = 0 \quad \text{contra} \quad H_{aB}\text{: por lo menos una } \beta_j \neq 0$$

La hipótesis de ninguna interacción H_{0AB} por lo general se prueba primero. Si no se rechaza H_{0AB}, entonces se pueden probar las otras hipótesis para ver si son significativos los efectos principales. Si se rechaza H_{0AB} y luego se prueba H_{0A} y no se rechaza, el modelo resultante $\mu_{ij} = \mu + \beta_j + \gamma_{ij}$ no se presta para una interpretación directa. En tal caso, es mejor construir un cuadro similar al de la figura 11.1(b) para tratar de tener una representación visual de la forma en que interactúan los factores.

Notación, modelo y análisis

Ahora se utilizan subíndices triples tanto para variables aleatorias como valores observados, con X_{ijk} y x_{ijk} que hacen referencia a la k-ésima observación (replicación) cuando el factor A está al nivel i y el factor B está al nivel j. Entonces el modelo es

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \qquad (11.10)$$
$$i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$$

donde las ϵ_{ij} son independientes y tienen una distribución normal, cada una con media 0 y varianza σ^2.

De nuevo un punto en lugar de un subíndice significa que se han sumado los valores de ese subíndice, mientras que una barra horizontal denota promediar. Así, X_{ij} es el total de las K

observaciones hechas para el factor A al nivel i y el factor B al nivel j [las observaciones de la (i, j)-ésima celda], y $\overline{X}_{ij\cdot}$ es el promedio de estas K observaciones.

Ejemplo 11.7 Tres variedades de tomate (Harvester, Pusa Early Dwarf y Ife No. 1) y cuatro densidades de plantas (10, 20, 30 y 40 mil plantas por hectárea) están siendo consideradas para plantar en una determinada región. Para ver si la variedad o la densidad de plantas afectan la producción, se utiliza cada combinación de variedad y densidad de plantas en tres parcelas diferentes, y se obtienen los datos de rendimientos de la tabla 11.6 (con base en el artículo "Effects of Plant Density on Tomato Yields in Western Nigeria", *Experimental Agriculture*, 1976: 43-47).

Tabla 11.6 Datos de producción para el ejemplo 11.7

Variedad	Densidad de plantas 10 000	20 000	30 000	40 000	$x_{i\cdot\cdot}$	$\bar{x}_{i\cdot\cdot}$
H	10.5 9.2 7.9	12.8 11.2 13.3	12.1 12.6 14.0	10.8 9.1 12.5	136.0	11.33
Ife	8.1 8.6 10.1	12.7 13.7 11.5	14.4 15.4 13.7	11.3 12.5 14.5	146.5	12.21
P	16.1 15.3 17.5	16.6 19.2 18.5	20.8 18.0 21.0	18.4 18.9 17.2	217.5	18.13
$x_{\cdot j\cdot}$	103.3	129.5	142.0	125.2	500.00	
$\bar{x}_{\cdot j\cdot}$	11.48	14.39	15.78	13.91		13.89

Aquí, $I = 3$, $J = 4$ y $K = 3$, para un total de $IJK = 36$ observaciones. ∎

Para probar las hipótesis de interés, de nuevo se definen las sumas de cuadrados y se presentan las fórmulas de cálculo:

$$\text{SST} = \sum_i \sum_j \sum_k (X_{ijk} - \overline{X}_{\cdots})^2 = \sum_i \sum_j \sum_k X_{ijk}^2 - \frac{1}{IJK} X_{\cdots}^2 \qquad \text{gl} = IJK - 1$$

$$\text{SSE} = \sum_i \sum_j \sum_k (X_{ijk} - \overline{X}_{ij\cdot})^2$$

$$= \sum_i \sum_j \sum_k X_{ijk}^2 - \frac{1}{K} \sum_i \sum_j X_{ij\cdot}^2 \qquad \text{gl} = IJ(K-1)$$

$$\text{SSA} = \sum_i \sum_j \sum_k (\overline{X}_{i\cdot\cdot} - \overline{X}_{\cdots})^2 = \frac{1}{JK} \sum_i X_{i\cdot\cdot}^2 - \frac{1}{IJK} X_{\cdots}^2 \qquad \text{gl} = I - 1$$

$$\text{SSB} = \sum_i \sum_j \sum_k (\overline{X}_{\cdot j\cdot} - \overline{X}_{\cdots})^2 = \frac{1}{IK} \sum_j X_{\cdot j\cdot}^2 - \frac{1}{IJK} X_{\cdots}^2 \qquad \text{gl} = J - 1$$

$$\text{SSAB} = \sum_i \sum_j \sum_k (\overline{X}_{ij\cdot} - \overline{X}_{i\cdot\cdot} - \overline{X}_{\cdot j\cdot} + \overline{X}_{\cdots})^2 \qquad \text{gl} = (I-1)(J-1)$$

La identidad fundamental

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE}$$

implica que **la suma de cuadrados de la interacción** SSAB se puede obtener por resta.

Por lo tanto, la variación total se divide en cuatro piezas: inexplicada (SSE, que estaría presente si alguna de las tres hipótesis nulas fuera verdadera o no) y tres piezas que se podrían explicar por la verdad o falsedad de las tres H_0. Cada uno de los cuatro cuadrados medios se define mediante MS = SS/gl. Los cuadrados medios esperados indican que se debe probar cada conjunto de hipótesis por medio de la relación apropiada de cuadrados medios con MSE en el denominador:

$$E(\text{MSE}) = \sigma^2$$

$$E(\text{MSA}) = \sigma^2 + \frac{JK}{I-1}\sum_{i=1}^{I}\alpha_i^2$$

$$E(\text{MSB}) = \sigma^2 + \frac{IK}{J-1}\sum_{j=1}^{J}\beta_j^2$$

$$E(\text{MSAB}) = \sigma^2 + \frac{K}{(I-1)(J-1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\gamma_{ij}^2$$

Se puede demostrar que cada una de las tres relaciones de cuadrado medio tiene una distribución F cuando la H_0 relacionada es verdadera, lo cual produce los siguientes procedimientos de prueba al nivel α.

Hipótesis			Valor del estadístico de prueba	Región de rechazo
H_{0A}	contra	H_{aA}	$f_A = \dfrac{\text{MSA}}{\text{MSE}}$	$f_A \geq F_{\alpha,I-1,IJ(K-1)}$
H_{0B}	contra	H_{aB}	$f_B = \dfrac{\text{MSB}}{\text{MSE}}$	$f_B \geq F_{\alpha,J-1,IJ(K-1)}$
H_{0AB}	contra	H_{aAB}	$f_{AB} = \dfrac{\text{MSAB}}{\text{MSE}}$	$f_{AB} \geq F_{\alpha,(I-1)(J-1),IJ(K-1)}$

Como antes, los resultados del análisis se resumen en una tabla de ANOVA.

Ejemplo 11.8 (Continuación del ejemplo 11.7)

A partir de los datos proporcionados, $x_{...}^2 = (500)^2 = 250\,000$,

$$\sum_i\sum_j\sum_k x_{ijk}^2 = (10.5)^2 + (9.2)^2 + \cdots + (18.9)^2 + (17.2)^2 = 7404.80$$

$$\sum_i x_{i..}^2 = (136.0)^2 + (146.5)^2 + (217.5)^2 = 87\,264.50$$

y

$$\sum_j x_{.j.}^2 = 63\,280.18$$

Los totales de celda ($x_{ij.}$) son

	10 000	20 000	30 000	40 000
H	27.6	37.3	38.7	32.4
Ife	26.8	37.9	43.5	38.3
P	48.9	54.3	59.8	54.5

de lo cual $\sum_i \sum_j x_{ij.}^2 = (27.6)^2 + \cdots + (54.5)^2 = 22\,100.28$. Entonces

$$\text{SST} = 7404.80 - \frac{1}{36}(250\,000) = 7404.80 - 6944.44 = 460.36$$

$$\text{SSA} = \frac{1}{12}(87\,264.50) - 6944.44 = 327.60$$

$$\text{SSB} = \frac{1}{9}(63\,280.18) - 6944.44 = 86.69$$

$$\text{SSE} = 7404.80 - \frac{1}{3}(22\,100.28) = 38.04$$

y

$$\text{SSAB} = 460.36 - 327.60 - 86.69 - 38.04 = 8.03$$

En la tabla 11.7 se muestra un resumen de los cálculos.

Tabla 11.7 Tabla de ANOVA para el ejemplo 11.8

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
Variedades	2	327.60	163.8	$f_A = 103.02$
Densidad	3	86.69	28.9	$f_B = 18.18$
Interacción	6	8.03	1.34	$f_{AB} = 0.84$
Error	24	38.04	1.59	
Total	35	460.36		

Puesto que $F_{0.01,6,24} = 3.67$ y $f_{AB} = 0.84$ no es ≥ 3.67, no se puede rechazar H_{0AB} al nivel 0.01, por lo tanto se concluye que los efectos de la interacción son insignificantes. Ahora se puede investigar la presencia o ausencia de efectos principales. Puesto que $F_{0.01,2,24} = 5.61$ y $f_A = 103.02 \geq 5.61$, H_{0A} se rechaza al nivel 0.01 en favor de la conclusión de que variedades diferentes afectan los rendimientos promedio reales. De manera similar, $f_B = 18.18 \geq 4.72 = F_{0.01,3,24}$, por lo tanto se concluye que el rendimiento promedio también depende de la densidad de plantas. ■

Comparaciones múltiples

Cuando no se rechaza la hipótesis de ninguna interacción H_{0AB} y se rechaza por lo menos una de las dos hipótesis nulas de efecto principal, se emplea el método de Tukey para identificar diferencias significativas en los niveles. Para identificar diferencias entre las α_i cuando se rechaza H_{0A},

1. Obtenga $Q_{\alpha,I,IJ(K-1)}$, donde el segundo subíndice I identifica el número de niveles que se comparan y el tercer subíndice se refiere al número de grados de libertad para el error.
2. Calcule $w = Q\sqrt{\text{MSE}/(JK)}$, donde JK es el número de observaciones promediadas para obtener cada una de las $\bar{x}_{i\cdot\cdot}$ comparadas en el paso 3.
3. Ordene las $\bar{x}_{i\cdot\cdot}$ de menor a mayor y, como antes, subraye los pares que difieren por menos de w. Los pares sin subrayar corresponden a niveles significativamente diferentes del factor A.

Para identificar diferentes niveles del factor B cuando se rechaza H_{0B}, sustituya con J el segundo subíndice en Q, reemplace JK por IK en w, y reemplace a $\bar{x}_{i\cdot}$ con $\bar{x}_{\cdot j\cdot}$.

Ejemplo 11.9 (Continuación del ejemplo 11.8)

Para el factor A (variedades), $I = 3$, así que con $\alpha = 0.01$ y $IJ(K-1) = 24$, $Q_{0.01,3,24} = 4.55$. Entonces $w = 4.55\sqrt{1.59/12} = 1.66$, por lo que al ordenar y subrayar se obtiene

$\bar{x}_{1\cdot\cdot}$	$\bar{x}_{2\cdot\cdot}$	$\bar{x}_{3\cdot\cdot}$
11.33	12.21	18.13

Las variedades Harvester y Ife al parecer no difieren de manera importante entre sí en efecto sobre la producción promedio real, pero ambas difieren de la variedad Pusa.

Para el factor B (densidad), $J = 4$, por lo tanto $Q_{0.01,4,24} = 4.91$ y $w = 4.91\sqrt{1.59/9} = 2.06$.

$\bar{x}_{\cdot 1\cdot}$	$\bar{x}_{\cdot 4\cdot}$	$\bar{x}_{\cdot 2\cdot}$	$\bar{x}_{\cdot 3\cdot}$
11.48	13.91	14.39	15.78

Así, con la tasa de error 0.01 por experimento, que es bastante conservadora, sólo la densidad mínima al parecer difiere en forma importante de las demás. Incluso con $\alpha = 0.05$ (de modo que $w = 1.64$), las densidades 2 y 3 no se pueden juzgar significativamente distintas entre sí en su efecto sobre el rendimiento. ■

Modelos con efectos mixtos y aleatorios

En algunos problemas, los niveles de cualquier factor podrían haber sido elegidos de una gran población de niveles posibles, así que los efectos con los que contribuye son aleatorios en vez de fijos. Como en la sección 11.1, si ambos factores contribuyen con efectos aleatorios, el modelo se le denomina de efectos aleatorios, mientras que si un factor está fijo y el otro es aleatorio, resulta un modelo de efectos fijos. Ahora se considera el análisis para un modelo de efectos fijos en el cual el factor A (renglones) es el factor fijo y el factor B (columnas) es el factor aleatorio. El caso en el que ambos factores son aleatorios se trata en el ejercicio 26.

En esta situación, el modelo de efectos mixtos es

$$X_{ijk} = \mu + \alpha_i + B_j + G_{ij} + \epsilon_{ijk}$$
$$i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$$

Aquí μ y α_i son constantes con $\sum \alpha_i = 0$ y las B_j, G_{ij} y ϵ_{ijk} son variables aleatorias independientes, distribuidas normalmente con valor esperado 0 y varianzas σ_B^2, σ_G^2 y σ^2, respectivamente.*

*Éste se conoce como modelo "irrestricto". Un modelo "restringido" alternativo requiere que $\sum_i G_{ij} = 0$ para cada j (de modo que las G_{ij} ya no son independientes). Los cuadrados medios esperados y las relaciones F apropiadas para probar ciertas hipótesis dependen de la elección del modelo. La opción predeterminada de MINITAB da el resultado para el modelo irrestricto.

$$H_{0A}\colon\ \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0 \quad \text{contra} \quad H_{aA}\colon \text{por lo menos una } \alpha_i \neq 0$$
$$H_{0B}\colon\ \sigma_B^2 = 0 \quad \text{contra} \quad H_{aB}\colon\ \sigma_B^2 > 0$$
$$H_{0G}\colon\ \sigma_G^2 = 0 \quad \text{contra} \quad H_{aG}\colon\ \sigma_G^2 > 0$$

Se acostumbra probar H_{0A} y H_{0B} sólo si no se puede rechazar la hipótesis de ninguna interacción H_{0G}.

Las sumas de cuadrados importantes y los cuadrados medios necesarios para los procedimientos de prueba se definen y calculan del mismo modo que para el caso de efectos fijos. Los cuadrados medios esperados para el modelo mixto son

$$E(\text{MSE}) = \sigma^2$$
$$E(\text{MSA}) = \sigma^2 + K\sigma_G^2 + \frac{JK}{I-1}\sum\alpha_i^2$$
$$E(\text{MSB}) = \sigma^2 + K\sigma_G^2 + IK\sigma_B^2$$

y

$$E(\text{MSAB}) = \sigma^2 + K\sigma_G^2$$

Así, para probar la hipótesis de ninguna interacción, de nuevo es apropiada la relación f_{AB} = MSAB/MSE, con H_{0G} rechazada si $f_{AB} \geq F_{\alpha,(I-1)(J-1),IJ(K-1)}$. Sin embargo, para probar H_{0A} contra H_{aA}, los cuadrados medios esperados indican aunque el numerador de la relación F debe ser aún MSA, el denominador debe ser MSAB y no MSE. MSAB también es el denominador de la relación F para probar H_{0B}.

Para probar H_{0A} contra H_{aA} (factor A fijo, B aleatorio), el valor del estadístico de prueba es f_A = MSA/MSAB, y la región de rechazo es $f_A \geq F_{\alpha,I-1,(I-1)(J-1)}$. La prueba de H_{0B} contra H_{aB} utiliza f_B = MSB/MSAB, y la región de rechazo es $f_B \geq F_{\alpha,J-1,(I-1)(J-1)}$.

Ejemplo 11.10 Un ingeniero de proceso ha identificado dos posibles causas de vibración del motor eléctrico: el material usado para la cubierta del motor (factor A) y la fuente de suministro de cojinetes usados en el motor (factor B). Los datos siguientes acerca de la cantidad de vibración (micras) resultaron de un experimento en el que los motores con cubiertas hechas de acero, aluminio y plástico se construyeron con cojinetes provenientes de cinco fuentes seleccionadas al azar.

		Fuente de suministro				
		1	2	3	4	5
Material	Acero	13.1 13.2	16.3 15.8	13.7 14.3	15.7 15.8	13.5 12.5
	Aluminio	15.0 14.8	15.7 16.4	13.9 14.3	13.7 14.2	13.4 13.8
	Plástico	14.0 14.3	17.2 16.7	12.4 12.3	14.4 13.9	13.2 13.1

Sólo los tres materiales de la cubierta utilizados en el experimento se consideran para usarlos en la producción, de modo que el factor A es fijo. Sin embargo, las cinco fuentes de suministro se seleccionaron al azar de una población mucho más grande, por lo tanto el factor B es aleatorio. Las hipótesis nulas pertinentes son

$$H_{0A}\colon\ \alpha_1 = \alpha_2 = \alpha_3 = 0 \qquad H_{0B}\colon\ \sigma_B^2 = 0 \qquad H_{0AB}\colon\ \sigma_G^2 = 0$$

El resultado de MINITAB se ilustra en la figura 11.4. La columna del valor P en la tabla ANOVA indica que las dos últimas hipótesis nulas se deben rechazar al nivel de significación 0.05. Los diferentes materiales de la cubierta al parecer no afectan la vibración, pero la interacción entre el material y el proveedor es una fuente importante de variación en la vibración.

```
Factor
Mat. de la    Tipo   Niveles   Valores
cubierta       fijo         3   1   2   3
Fuente   Aleatorio          5   1   2   3   4   5

Fuente                  GL        SS         MS          F          P
Mat. de la cubierta      2    0.7047     0.3523       0.24      0.790
Fuente                   4   36.6747     9.1687       6.32      0.013
Mat. de la
cubierta*fuente          8   11.6053     1.4507      13.03      0.000
Error                   15    1.6700     0.1113
Total                   29   50.6547

Fuente                          Componente
                                   de la   Término   Cuadrado medio esperado
                                varianza  del error  para cada término (con el
                                                     modelo irrestricto)
1 Mat. de la cubierta                          3     (4) + 2(3) + Q[1]
2 Fuente                         1.2863        3     (4) + 2(3) + 6(2)
3 Mat. de la cubierta*fuente     0.6697        4     (4) + 2(3)
4 Error                          0.1113              (4)
```

Figura 11.4 Resultado de la opción ANOVA equilibrada de MINITAB para los datos del ejemplo 11.10 ■

Cuando por lo menos dos de las K_{ij} son iguales, los cálculos del ANOVA son mucho más complejos que para el caso $K_{ij} = K$, y no hay fórmulas para las estadísticas de prueba apropiadas. Para más información se pueden consultar en las referencias del capítulo.

Ejercicios | Sección 11.2 (16 a 26)

16. En un experimento para evaluar los efectos del tiempo de curación (factor A) y el tipo de mezcla (factor B) en la resistencia compresiva de cubos de cemento endurecido, se emplearon tiempos de curación distintos en combinación con cuatro mezclas distintas, con tres observaciones obtenidas de cada una de 12 combinaciones de tiempos de curación-mezcla. Las sumas de cuadrados resultantes se calcularon como SSA = 30 763.0, SSB = 34 185.6, SSE = 97 436.8 y SST = 205 966.6.
    a. Construya una tabla ANOVA.
    b. Prueba al nivel 0.05 la hipótesis nula H_{0AB}: todas las $\gamma_{ij} = 0$ (ninguna interacción de los factores) contra H_{aAB}: por lo menos una $\gamma_{ij} \neq 0$.
    c. Pruebe al nivel 0.05 la hipótesis nula H_{0A}: $\alpha_1 = \alpha_2 = \alpha_3 = 0$ (están ausentes los efectos principales del factor A) contra H_{aA}: por lo menos una $\alpha_i \neq 0$.
    d. Pruebe H_{0B}: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ contra H_{aB}: por lo menos una $\beta_j \neq 0$ con una prueba de nivel 0.05.
    e. Los valores de las $\bar{x}_{i\cdot\cdot}$ fueron $\bar{x}_{1\cdot\cdot} = 4010.88$, $\bar{x}_{2\cdot\cdot} = 4029.10$ y $\bar{x}_{3\cdot\cdot} = 3960.02$. Utilice el procedimiento de Tukey para investigar diferencias importantes entre los tiempos de curación.
17. En el artículo "Towards Improving the Properties of Plaster Moulds and Castings" (*J. Engr. Manuf.*, 1991: 265-269) se describen varios ANOVA realizados para estudiar cómo la cantidad de adiciones de fibra de car-

bono y arena afectan varias características del proceso de moldeo. A continuación se ilustran los datos de la dureza del vaciado y la resistencia de molde húmedo.

Adición de arena (%)	Adición de fibra de carbono (%)	Dureza del vaciado	Resistencia de molde húmedo
0	0	61.0	34.0
0	0	63.0	16.0
15	0	67.0	36.0
15	0	69.0	19.0
30	0	65.0	28.0
30	0	74.0	17.0
0	0.25	69.0	49.0
0	0.25	69.0	48.0
15	0.25	69.0	43.0
15	0.25	74.0	29.0
30	0.25	74.0	31.0
30	0.25	72.0	24.0
0	0.50	67.0	55.0
0	0.50	69.0	60.0
15	0.50	69.0	45.0
15	0.50	74.0	43.0
30	0.50	74.0	22.0
30	0.50	74.0	48.0

a. Un ANOVA para la resistencia de molde húmedo da SS-arena = 705, SSFibra = 1278, SSE = 843 y SST = 3105. Pruebe la presencia de algunos efectos con $\alpha = 0.05$.
b. Realice un ANOVA en las observaciones de la dureza del vaciado con $\alpha = 0.05$.
c. Trace la gráfica de la dureza promedio muestral contra el porcentaje de arena para los diferentes niveles de fibra de carbono. ¿La gráfica es consistente con el análisis del inciso (b)?

18. Los datos siguientes resultaron de un experimento para investigar si el rendimiento de cierto proceso químico depende de la formulación de un determinado insumo o la velocidad del mezclador.

		Velocidad		
		60	70	80
Formulación	1	189.7	185.1	189.0
		188.6	179.4	193.0
		190.1	177.3	191.1
	2	165.1	161.7	163.3
		165.9	159.8	166.6
		167.6	161.6	170.3

Con un paquete de software para estadística se obtuvo SS(Form) = 2253.44, SS(Velocidad) = 230.81, SS (Form*Velocidad) = 18.58 y SSE = 71.87.
a. ¿Parece haber interacción entre los factores?
b. ¿Al parecer el rendimiento depende de la formulación o la velocidad?
c. Calcule estimaciones de los efectos principales.
d. Los *valores ajustados* son $\hat{x}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}$ y los *residuos* son $x_{ijk} - \hat{x}_{ijk}$. Compruebe que los residuos son 0.23, −0.87, 0.63, 4.50, −1.20, −3.30, −2.03, 1.97, 0.07, −1.10, −0.30, 1.40, 0.67, −1.23, 0.57, −3.43, −0.13 y 3.57.
e. Construya una gráfica de probabilidad normal de los residuos proporcionados en el inciso (d). ¿Al parecer la ϵ_{ijk} tiene una distribución normal?

19. En la siguiente tabla de datos se dan las observaciones de acidez total de muestras de carbón mineral de tres tipos, con determinaciones hechas usando tres concentraciones diferentes de NaOH etanólico ("Chemistry of Brown Coals", *Australian J. Applied Science*, 1958: 375-379).

		Tipo de carbón mineral		
		Morwell	Yallourn	Maddingley
NaOH Conc.	0.404N	8.27, 8.17	8.66, 8.61	8.14, 7.96
	0.626N	8.03, 8.21	8.42, 8.58	8.02, 7.89
	0.786N	8.60, 8.20	8.61, 8.76	8.13, 8.07

Además, $\sum_i \sum_j \sum_k x_{ijk}^2 = 1240.1525$ y $\sum_i \sum_j x_{ij}^2 = 2479.9991$.
a. Suponiendo que ambos efectos son fijos, construya una tabla ANOVA, pruebe la presencia de interacción y luego pruebe la presencia de efectos principales para cada factor (todo con el nivel 0.01).
b. Utilice el procedimiento de Tukey para identificar diferencias significativas entre los tipos de carbón mineral.

20. La corriente (en μA) necesaria para producir un cierto nivel de brillo de un tubo de televisión se midió para dos tipos de vidrio y tres tipos de fósforo, obteniéndose los datos siguientes ("Fundamentals of Analysis of Variance", *Industrial Quality Control*, 1956: 5-8).

		Tipo de fósforo		
		1	2	3
Tipo de vidrio	1	280, 290, 285	300, 310, 295	270, 285, 290
	2	230, 235, 240	260, 240, 235	220, 225, 230

Suponiendo que ambos factores son fijos, pruebe H_{0AB} contra H_{aAB} al nivel 0.01. Entonces, si no se puede rechazar H_{0AB}, pruebe los dos conjuntos de hipótesis de efecto principal.

21. En un experimento para investigar el efecto del "factor cemento" (número de sacos de cemento por yarda cúbica) en la resistencia a la flexión del concreto resultante ("Studies of Flexural Strength of Concrete. Part 3: Effects of Variation in Testing Procedure", *Proceedings ASTM*, 1957: 1127-1139), se usaron $I = 3$ valores de factor distintos, se eligieron $J = 5$ lotes diferentes de cemento y se colaron $K = 2$ vigas de cada combinación factor de cemento/lote. Los valores de resumen son $\sum\sum\sum x_{ijk}^2 = 12\,280\,103$, $\sum\sum x_{ij\cdot}^2 = 24\,529\,699$, $\sum x_{i\cdot\cdot}^2 = 122\,380\,901$, $\sum x_{\cdot j\cdot}^2 = 73\,427\,483$ y $x_{\cdot\cdot\cdot} = 19\,143$.
    a. Construya la tabla ANOVA.
    b. Suponiendo un modelo fijo con el factor cemento (A) fijo y lotes (B) aleatorio, pruebe los tres pares de hipótesis de interés al nivel 0.05.

22. Un estudio realizado para comparar las duraciones de escritura de cuatro marcas de plumas de primera calidad. Se consideró que la superficie de escritura podría afectar la duración, así que se eligieron al azar tres superficies distintas. Se empleó una máquina de escritura para asegurar que las condiciones fueran homogéneas (p. ej., presión constante a un ángulo fijo). En la tabla siguiente se muestran las dos duraciones (min) obtenidas para cada combinación de marca-superficie. Además, $\sum\sum\sum x_{ijk}^2 = 11\,499\,492$ y $\sum\sum x_{ij\cdot}^2 = 22\,982\,552$.

Superficie de escritura

		1	2	3	$x_{i\cdot\cdot}$
Marca de la pluma	1	709, 659	713, 726	660, 645	4112
	2	668, 685	722, 740	692, 720	4227
	3	659, 685	666, 684	678, 750	4122
	4	698, 650	704, 666	686, 733	4137
$x_{\cdot j\cdot}$		5413	5621	5564	16 598

Realice un ANOVA apropiado y enuncie sus conclusiones.

23. La tabla siguiente se obtuvo en un experimento para investigar si la resistencia compresiva de cilindros de concreto depende del tipo de material de sellado empleado o la variabilidad en los diferentes lotes ("The effect of Type of Capping Material on the Compressive Strength of Concrete Cylinders", *Proccedings ASTM*, 1958: 1166-1186). Cada número es un total de celda (x_{ij}) con base en $K = 3$ observaciones.

		Lote				
		1	2	3	4	5
Material de sellado	1	1847	1942	1935	1891	1795
	2	1779	1850	1795	1785	1626
	3	1806	1892	1889	1891	1756

Además, $\sum\sum\sum x_{ijk}^2 = 16\,815\,853$ y $\sum\sum x_{ij}^2 = 50\,443\,409$. Obtenga la tabla ANOVA y luego pruebe al nivel 0.01 las hipótesis H_{0G} contra H_{aG}, H_{0A} contra H_{aA} y H_{0AB} contra H_{aAB}, suponiendo que el sellado es un efecto fijo y los lotes es un efecto aleatorio.

24. a. Muestre que $E(\overline{X}_{i\cdot\cdot} - \overline{X}_{\cdot\cdot\cdot}) = \alpha_i$ de modo que $\overline{X}_{i\cdot\cdot} - \overline{X}_{\cdot\cdot\cdot}$ es un estimador insesgado para α_i (en el modelo de efectos fijos).
    b. Con $\hat{\gamma}_{ij} = \overline{X}_{ij\cdot} - \overline{X}_{i\cdot\cdot} - \overline{X}_{\cdot j\cdot} + \overline{X}_{\cdot\cdot\cdot}$, demuestre que $\hat{\gamma}_{ij}$ es un estimador insesgado para γ_{ij} (en el modelo de efectos fijos).

25. Demuestre cómo se puede obtener un IC de $100(1 - \alpha)\%$ para $\alpha_i - \alpha_{i'}$. Luego, calcule un intervalo de 95% para $\alpha_2 - \alpha_3$ con los datos del ejercicio 19. [*Sugerencia*: con $\theta = \alpha_2 - \alpha_3$, el resultado del ejercicio 24a indica cómo obtener $\hat{\theta}$. Después calcule $V(\hat{\theta})$ y $\sigma_{\hat{\theta}}$, y obtenga una estimación de $\sigma_{\hat{\theta}}$ por medio de $\sqrt{\text{MSE}}$ para estimar σ (lo que identifica el número apropiado de gl).]

26. Cuando ambos factores son aleatorios en un experimento ANOVA bidireccional con K replicaciones por combinación de niveles de factor, los cuadrados medios esperados son $E(\text{MSE}) = \sigma^2$, $E(\text{MSA}) = \sigma^2 + K\sigma_G^2 + JK\sigma_A^2$, $E(\text{MSB}) = \sigma^2 + K\sigma_G^2 + IK\sigma_B^2$ y $E(\text{MSAB}) = \sigma^2 + K\sigma_G^2$.
    a. ¿Qué relación F es apropiada para probar H_{0G}: $\sigma_G^2 = 0$ contra H_{aG}: $\sigma_G^2 > 0$?
    b. Conteste el inciso (a) para probar H_{0A}: $\sigma_A^2 = 0$ contra H_{aA}: $\sigma_A^2 > 0$ y H_{0B}: $\sigma_B^2 = 0$ contra H_{aB}: $\sigma_B^2 > 0$.

11.3 ANOVA de tres factores

Para indicar la naturaleza de modelos y análisis cuando los experimentos de ANOVA tienen que ver con más de dos factores, aquí se atenderá el caso de tres factores fijos, A, B y C. Los números de los niveles de los tres factores se representarán mediante I, J y K, respectivamente, y

L_{ijk} = número de observaciones hechas con el factor A al nivel i, factor B al nivel j y factor C al nivel k. Al igual que con el ANOVA de dos factores, el análisis es bastante complicado cuando las L_{ijk} no son iguales, así que L_{ijk} se restringe a L, es decir, $L_{ijk} = L$. Entonces X_{ijkl} y x_{ijkl} representan el valor observado, antes y después del experimento, de la l-ésima replicación ($l = 1, 2, \ldots, L$) cuando los tres factores son fijos a los niveles i, j y k.

Para entender los parámetros que aparecerán en el modelo ANOVA de tres factores, primero recuerde que en el ANOVA de dos factores con replicaciones, $E(X_{ijk}) = \mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$, donde las restricciones $\sum_i \alpha_i = \sum_j \beta_j = 0$, $\sum_i \gamma_{ij} = 0$ para toda j y $\sum_j \gamma_{ij} = 0$ para toda i fueron necesarias para obtener un conjunto único de parámetros. Si se utilizan subíndices de punto en las μ_{ij} para denotar el promedio (en vez de la suma), entonces

$$\mu_{i\cdot} - \mu_{\cdot\cdot} = \frac{1}{J}\sum_j \mu_{ij} - \frac{1}{IJ}\sum_i \sum_j \mu_{ij} = \alpha_i$$

es el efecto del factor A al nivel i promediado en los niveles del factor B, mientras que

$$\mu_{ij} - \mu_{\cdot j} = \mu_{ij} - \frac{1}{I}\sum_i \mu_{ij} = \alpha_i + \gamma_{ij}$$

es el efecto del factor A al nivel i específico para el factor B al nivel j. Si el efecto de A al nivel i depende del nivel de B, entonces hay interacción entre los factores, y las γ_{ij} no son cero. En particular,

$$\mu_{ij} - \mu_{\cdot j} - \mu_{i\cdot} + \mu_{\cdot\cdot} = \gamma_{ij} \tag{11.11}$$

Modelo de efectos fijos de tres factores

El modelo para el ANOVA de tres factores con $L_{ijk} = L$ es

$$X_{ijkl} = \mu_{ijk} + \epsilon_{ijkl} \qquad \begin{matrix} i = 1, \ldots, I, & j = 1, \ldots, J \\ k = 1, \ldots, K, & l = 1, \ldots, L \end{matrix} \tag{11.12}$$

donde las ϵ_{ijkl} tiene una distribución normal con media 0 y varianza σ^2, y

$$\mu_{ijk} = \mu + \alpha_i + \beta_j + \delta_k + \gamma_{ij}^{AB} + \gamma_{ik}^{AC} + \gamma_{jk}^{BC} + \gamma_{ijk} \tag{11.13}$$

Las restricciones necesarias para obtener parámetros definidos de manera única, son que la suma en cualquier subíndice de cualquier parámetro en el lado derecho de la ecuación (11.13) sea igual a 0.

Los parámetros γ_{ij}^{AB}, γ_{ik}^{AC} y γ_{jk}^{BC} se llaman interacciones de dos factores y γ_{ijk} se llama interacción de tres factores; las α_i, β_j y δ_k son los parámetros de efectos principales. Para cualquier nivel fijo k del tercer factor, análogo a (11.11),

$$\mu_{ijk} - \mu_{i\cdot k} - \mu_{\cdot jk} + \mu_{\cdot\cdot k} = \gamma_{ij}^{AB} + \gamma_{ijk}$$

es la interacción del i-ésimo nivel de A con el j-ésimo nivel de B específica para el k-ésimo nivel de C, mientras que

$$\mu_{ij\cdot} - \mu_{i\cdot\cdot} - \mu_{\cdot j\cdot} + \mu_{\cdot\cdot\cdot} = \gamma_{ij}^{AB}$$

es la interacción entre A al nivel i y B al nivel j promediada en los niveles de C. Si la interacción de A al nivel i y B al nivel j no depende de k, entonces las γ_{ijk} son iguales a 0. Así, las γ_{ijk} distintas de cero representan la no aditividad de las γ_{ij}^{AB} de dos factores sobre los distintos ni-

veles del tercer factor C. Si el experimento incluyó más de tres factores, habría términos de interacción de orden superior correspondientes con interpretaciones análogas. Observe que en el argumento previo, si se hubiera considerado fijar el nivel de A o de B (y no el de C, como se hizo) y examinar las γ_{ijk}, su interpretación sería la misma, si alguna de las interacciones de dos factores dependen del nivel del tercer factor, entonces hay γ_{ijk} que no equivalen a cero.

Análisis de un experimento de tres factores

Cuando $L > 1$, hay una suma de cuadrados para cada efecto principal, cada interacción de dos factores y la interacción de tres factores. Para escribir éstas en una forma que indica cómo se definen las sumas de cuadrados cuando hay más de tres factores, observe que cualquiera de los parámetros del modelo en (11.13) se pueden estimar sin sesgo al promediar X_{ijkl} sobre los subíndices apropiados y tomar diferencias. Por lo tanto,

$$\hat{\mu} = \overline{X}_{\cdots} \qquad \hat{\alpha}_i = \overline{X}_{i\cdots} - \overline{X}_{\cdots} \qquad \hat{\gamma}_{ij}^{AB} = \overline{X}_{ij\cdot\cdot} - \overline{X}_{i\cdots} - \overline{X}_{\cdot j\cdot\cdot} + \overline{X}_{\cdots}$$
$$\hat{\gamma}_{ijk} = \overline{X}_{ijk\cdot} - \overline{X}_{ij\cdot\cdot} - \overline{X}_{i\cdot k\cdot} - \overline{X}_{\cdot jk\cdot} + \overline{X}_{i\cdots} + \overline{X}_{\cdot j\cdot\cdot} + \overline{X}_{\cdot\cdot k\cdot} - \overline{X}_{\cdots}$$

con otros efectos principales y estimadores de interacción obtenidos por simetría. Entonces las sumas de cuadrados son

$$\text{SST} = \sum_i \sum_j \sum_k \sum_l (X_{ijkl} - \overline{X}_{\cdots})^2 \qquad \text{gl} = IJKL - 1$$
$$\text{SSA} = \sum_i \sum_j \sum_k \sum_l \hat{\alpha}_i^2 = JKL \sum_i (\overline{X}_{i\cdots} - \overline{X}_{\cdots})^2 \qquad \text{gl} = I - 1$$
$$\text{SSAB} = \sum_i \sum_j \sum_k \sum_l (\hat{\gamma}_{ij}^{AB})^2 = KL \sum_i \sum_j (\overline{X}_{ij\cdot\cdot} - \overline{X}_{i\cdots} - \overline{X}_{\cdot j\cdot\cdot} + \overline{X}_{\cdots})^2 \qquad \text{gl} = (I-1)(J-1)$$
$$\text{SSABC} = \sum_i \sum_j \sum_k \sum_l \hat{\gamma}_{ijk}^2 = L \sum_i \sum_j \sum_k \hat{\gamma}_{ijk}^2 \qquad \text{gl} = (I-1)(J-1)(K-1)$$
$$\text{SSE} = \sum_i \sum_j \sum_k \sum_l (X_{ijkl} - \overline{X}_{ijk\cdot})^2 \qquad \text{gl} = IJK(L-1)$$

con el otro efecto principal y las sumas de cuadrados de interacción de dos factores obtenidos por simetría. SST es la suma de las otras ocho sumas de cuadrados.

Incluso las fórmulas de cálculo para estas sumas de cuadrados son bastante tediosas para usar, así que se omiten en favor del resultado de un paquete de software de estadística. La versión actual de MINITAB, por ejemplo, se ajusta a un modelo de tres factores con efectos aleatorios fijos, mixtos o aleatorios.

Cada suma de cuadrados (exceptuando SST) cuando se divide entre sus gl da un cuadrado medio, con

$$E(\text{MSE}) = \sigma^2$$
$$E(\text{MSA}) = \sigma^2 + \frac{JKL}{I-1} \sum_i \alpha_i^2$$
$$E(\text{MSAB}) = \sigma^2 + \frac{KL}{(I-1)(J-1)} \sum_i \sum_j (\gamma_{ij}^{AB})^2$$
$$E(\text{MSABC}) = \sigma^2 + \frac{L}{(I-1)(J-1)(K-1)} \sum_i \sum_j \sum_k (\gamma_{ijk})^2$$

y expresiones similares para los otros cuadrados medios esperados. El efecto principal de las hipótesis de interacción se prueban al formar relaciones F con MSE en cada denominador.

Hipótesis nula	Valor estadístico de prueba	Región de rechazo
H_{0A}: todas las $\alpha_i = 0$	$f_A = \dfrac{\text{MSA}}{\text{MSE}}$	$f_A \geq F_{\alpha,I-1,IJK(L-1)}$
H_{0AB}: todas las $\gamma_{ij}^{AB} = 0$	$f_{AB} = \dfrac{\text{MSAB}}{\text{MSE}}$	$f_{AB} \geq F_{\alpha,(I-1)(J-1),IJK(L-1)}$
H_{0ABC}: todas las $\gamma_{ijk} = 0$	$f_{ABC} = \dfrac{\text{MSABC}}{\text{MSE}}$	$f_{ABC} \geq F_{\alpha,(I-1)(J-1)(K-1),IJK(L-1)}$

Por lo común las hipótesis de efectos principales se prueban sólo si se juzga que no son significativas las interacciones.

En este análisis se supone que $L_{ijk} = L > 1$. Si $L = 1$, entonces como en caso de dos factores, se debe suponer que las interacciones de orden superior deben ser igual a cero para obtener un MSE que estime a σ^2. Si $L = 1$ y se descarta la suma del cuarto subíndice sobre l, aún son válidas las fórmulas anteriores para las sumas de cuadrados, y $\text{SSE} = \sum_i \sum_j \sum_k \hat{\gamma}_{ijk}^2$ con $\overline{X}_{ijk\cdot} = X_{ijk}$ en la expresión para $\hat{\gamma}_{ijk\cdot}$.

Ejemplo 11.11 Las siguientes observaciones (temperatura corporal − 100°F) se obtuvieron en un experimento para estudiar la tolerancia del ganado al calor ("The significance of the Coat in Heat Tolerance of Cattle", *Australian J. Agruculture Research*, 1959: 744-748). Las mediciones se hicieron en cuatro periodos distintos (factor A, con $I = 4$) en dos cepas de ganado (factor B, con $J = 2$) con cuatro tipos de tegumento (factor C, con $K = 4$); se hicieron $L = 3$ observaciones para cada una de las $4 \times 2 \times 4 = 32$ combinaciones de niveles de los tres factores.

	B_1				B_2			
	C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4
A_1	3.6	3.4	2.9	2.5	4.2	4.4	3.6	3.0
	3.8	3.7	2.8	2.4	4.0	3.9	3.7	2.8
	3.9	3.9	2.7	2.2	3.9	4.2	3.4	2.9
A_2	3.8	3.8	2.9	2.4	4.4	4.2	3.8	2.0
	3.6	3.9	2.9	2.2	4.4	4.3	3.7	2.9
	4.0	3.9	2.8	2.2	4.6	4.7	3.4	2.8
A_3	3.7	3.8	2.9	2.1	4.2	4.0	4.0	2.0
	3.9	4.0	2.7	2.0	4.4	4.6	3.8	2.4
	4.2	3.9	2.8	1.8	4.5	4.5	3.3	2.0
A_4	3.6	3.6	2.6	2.0	4.0	4.0	3.8	2.0
	3.5	3.7	2.9	2.0	4.1	4.4	3.7	2.2
	3.8	3.9	2.9	1.9	4.2	4.2	3.5	2.3

La tabla de totales de celda (x_{ijk}) para las combinaciones de los tres factores es

	B_1				B_2			
x_{ijk}	C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4
A_1	11.3	11.0	8.4	7.1	12.1	12.5	10.7	8.7
A_2	11.4	11.6	8.6	6.8	13.4	13.2	10.9	7.7
A_3	11.8	11.7	8.4	5.9	13.1	13.1	11.1	6.4
A_4	10.9	11.2	8.4	5.9	12.3	12.6	11.0	6.5

En la figura 11.5 se muestran las gráficas de las medias de celda correspondientes $\bar{x}_{ijk.} = x_{ijk.}/3$. Se volverá a estas gráficas después de considerar pruebas de varias hipótesis. La base para estas pruebas es la tabla ANOVA dada en la tabla 11.8.

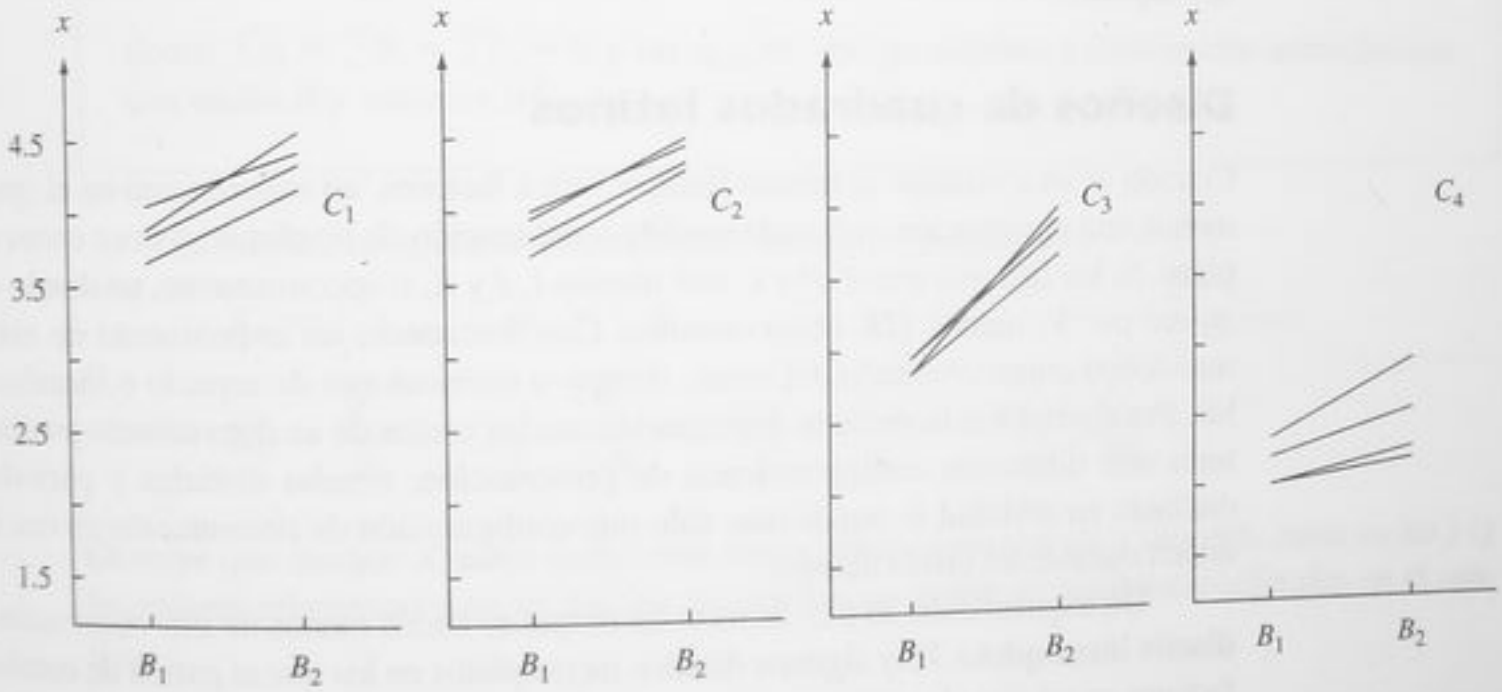


Figura 11.5 Gráficas de $x_{ijk.}$ para el ejemplo 11.11

Tabla 11.8 Tabla de ANOVA para el ejemplo 11.11

Fuente	gl	Suma de cuadrados	Cuadrado medio	f
A	$I - 1 = 3$	0.49	0.163	4.13
B	$J - 1 = 1$	6.45	6.45	163.29
C	$K - 1 = 3$	48.93	16.31	412.91
AB	$(I - 1)(J - 1) = 3$	0.02	0.0067	0.170
AC	$(I - 1)(K - 1) = 9$	1.61	0.179	4.53
BC	$(J - 1)(K - 1) = 3$	0.88	0.293	7.42
ABC	$(I - 1)(J - 1)(K - 1) = 9$	0.25	0.0278	0.704
Error	$IJK(L - 1) = 64$	2.53	0.0395	
Total	$IJKL - 1 = 95$	61.16		

Puesto que $F_{0.01,9,64} \approx 2.70$ y $f_{ABC} = \text{MSABC/MSE} = 0.704$ no es mayor que 2.70, se concluye que no son significativas las interacciones de tres factores. Sin embargo, aunque las interacciones AB también son no insignificantes, las interacciones AC y BC, así como los efectos principales, al parecer son no significativas en el modelo. Cuando no hay interacciones ABC o AB, una gráfica de las $\bar{x}_{ijk\cdot}$ $(= \hat{\mu}_{ijk})$ por separado para cada nivel de C debe revelar interacciones no sustanciales (si sólo las interacciones ABC son cero, es más difícil interpretar las gráficas; véase el artículo "Two-Dimensional Plots for Interpreting Interactions in the Three-Factor Analysis of Variance Model", *Amer. Statistician*, mayo de 1979: 63-69). ■

El procedimiento de Tukey se puede usar en el ANOVA de tres factores (o más). El segundo subíndice en Q es el número de medias muestrales que se comparan y el tercero son los grados del libertad del error.

Los modelos con efectos aleatorios y mixtos también se pueden analizar. Las sumas de cuadrados y los grados de libertad son idénticos para el caso de efectos fijos, pero los cuadrados medios esperados son, por supuesto, diferentes para los efectos aleatorios principales o interacciones. Una buena referencia es el libro de Douglas Montgomery listado en la bibliografía del capítulo.

Diseños de cuadrados latinos

Cuando se va a estudiar al mismo tiempo varios factores, un experimento en el que haya por lo menos una observación para cada posible combinación de niveles se conoce como **diseño completo**. Si los factores son A, B y C con niveles I, J y K, respectivamente, un diseño completo requiere por lo menos IJK observaciones. Con frecuencia un experimento de este tamaño es impráctico como resultado del costo, tiempo o restricciones de espacio o literalmente imposible. Por ejemplo, si la variable de respuesta son las ventas de un determinado producto y los factores son diferentes configuraciones de presentación, tiendas distintas y periodos variables, entonces en realidad se puede usar sólo una configuración de presentación en una determinada tienda durante un cierto tiempo.

Un experimento de tres factores en el que se hacen menos de tres observaciones se llama **diseño incompleto**. Hay algunos diseños incompletos en los que el patrón de combinaciones de factores es tal que el análisis es directo. Un diseño de tres factores de este tipo se llama **cuadrado latino.** Éste es apropiado cuando $I = J = K$ (p. ej., cuatro configuraciones de presentación, cuatro tiendas y cuatro periodos) y se supone que están ausentes los efectos de interacción de dos y tres factores. Si los niveles del factor A se identifican con los renglones de una tabla bidireccional y los niveles de B con las columnas de la tabla, entonces la característica definitoria de un diseño de cuadrado latino es que *todo nivel del factor C aparece una vez en cada renglón y una vez en cada columna*. En la figura 11.6 se ilustran ejemplos de cuadrados latinos de 3×3, 4×4 y 5×5. Hay 12 cuadrados latinos de 3×3 distintos, y el número de cuadrados latinos de $N \times N$ diferentes se incrementa rápido con N (p. ej., cada permutación de renglones de un determinado cuadrado latino produce un cuadrado latino, y de manera similar para las permutaciones de columna). Se recomienda que el cuadrado latino empleado en un experimento particular se elija al azar del conjunto de los cuadrados posibles de la dimensión deseada; para más detalles consulte una de las referencias del capítulo.

Modelo y análisis para cuadrados latinos La letra N se usará para denotar el valor común de I, J y K. Entonces un diseño completo con una observación por combinación requeriría N^3 observaciones, mientras que un cuadrado latino requiere sólo N^2 observaciones. Una vez elegido el cuadrado, el valor de k (el nivel del factor C) se determina por completo

		B		
C	1	2	3	
1	1	2	3	
A 2	2	3	1	
3	3	1	2	

		B			
C	1	2	3	4	
1	3	4	2	1	
A 2	4	2	1	3	
3	2	1	3	4	
4	1	3	4	2	

		B				
C	1	2	3	4	5	
1	4	3	5	2	1	
2	3	1	4	5	2	
A 3	1	5	2	3	4	
4	5	2	1	4	3	
5	2	4	3	1	5	

Figura 11.6 Ejemplos de cuadrados latinos

con los valores de i y j. Para remarcar esto, se usa $x_{ij(k)}$ para denotar el valor observado cuando los tres factores están a los niveles i, j y k, respectivamente, donde k toma sólo un valor para cada i, j. Entonces el modelo es

$$X_{ij(k)} = \mu + \alpha_i + \beta_j + \delta_k + \epsilon_{ij(k)} \qquad i, j, k = 1, \ldots, N$$

donde $\sum\alpha_i = \sum\beta_j = \sum\delta_k = 0$ y las $\epsilon_{ij(k)}$ son independientes y distribuidas normalmente con media 0 y varianza σ^2.

Se emplea la siguiente notación para totales y promedios:

$$X_{i\cdot\cdot} = \sum_j X_{ij(k)} \qquad X_{\cdot j\cdot} = \sum_i X_{ij(k)} \qquad X_{\cdot\cdot k} = \sum_{ij} X_{ij(k)} \qquad X_{\cdots} = \sum_i \sum_j X_{ij(k)}$$

$$\bar{X}_{i\cdot\cdot} = \frac{X_{i\cdot\cdot}}{N} \qquad \bar{X}_{\cdot j\cdot} = \frac{X_{\cdot j\cdot}}{N} \qquad \bar{X}_{\cdot\cdot k} = \frac{X_{\cdot\cdot k}}{N} \qquad \bar{X}_{\cdots} = \frac{X_{\cdots}}{N^2}$$

Observe que aunque $X_{i\cdot\cdot}$ antes indicó una suma, ahora corresponde a una sola suma en las j (y los valores relacionados de k). Las siguientes son las sumas de cuadrados utilizadas en el análisis.

$$\text{SST} = \sum_i \sum_j (X_{ij(k)} - \bar{X}_{\cdots})^2 = \sum_i \sum_j X_{ij(k)}^2 - \frac{X_{\cdots}^2}{N^2} \qquad \text{gl} = N^2 - 1$$

$$\text{SSA} = \sum_i \sum_j (\bar{X}_{i\cdot\cdot} - \bar{X}_{\cdots})^2 = \frac{1}{N}\sum_i X_{i\cdot\cdot}^2 - \frac{X_{\cdots}^2}{N^2} \qquad \text{gl} = N - 1$$

$$\text{SSB} = \sum_i \sum_j (\bar{X}_{\cdot j\cdot} - \bar{X}_{\cdots})^2 = \frac{1}{N}\sum_j X_{\cdot j\cdot}^2 - \frac{X_{\cdots}^2}{N^2} \qquad \text{gl} = N - 1$$

$$\text{SSC} = \sum_i \sum_j (\bar{X}_{\cdot\cdot k} - \bar{X}_{\cdots})^2 = \frac{1}{N}\sum_k X_{\cdot\cdot k}^2 - \frac{X_{\cdots}^2}{N^2} \qquad \text{gl} = N - 1$$

$$\text{SSE} = \sum_i \sum_j [X_{ij(k)} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\delta}_k)]^2$$

$$= \sum_i \sum_j (X_{ij(k)} - \bar{X}_{i\cdot\cdot} - \bar{X}_{\cdot j\cdot} - \bar{X}_{\cdot\cdot k} + 2\bar{X}_{\cdots})^2 \qquad \text{gl} = (N-1)(N-2)$$

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSC} + \text{SSE}$$

Cada cuadrado medio es, por supuesto, la relación SS/gl. Para probar H_{0C}: $\delta_1 = \delta_2 = \cdots = \delta_N = 0$, el valor del estadístico de prueba es $f_C = \text{MSC/MSE}$, con H_{0C} rechazada si $f_C \geq F_{\alpha,N-1,(N-1)(N-2)}$. Las otras dos hipótesis nulas de efecto principal también se rechazan si la relación F correspondiente es mayor que $F_{\alpha,N-1,(N-1)(N-2)}$.

Si se rechaza cualquiera de las hipótesis nulas, las diferencias significativas se identifican mediante el procedimiento de Tukey. Después de calcular $w = Q_{\alpha,N,(N-1)(N-2)} \cdot \sqrt{\text{MSE}/N}$, los pares de medias muestrales (las $\bar{x}_{i\cdot\cdot}$, $\bar{x}_{\cdot j\cdot}$ o $\bar{x}_{\cdot\cdot k}$) que difieren en más de w corresponden a diferencias importantes entre efectos de factores relacionados (las α_i, β_j o δ_k).

La hipótesis H_{0C} es con frecuencia la única de interés central. Se emplea un diseño de cuadrado latino para controlar la variación extraña en los factores A y B, como se hizo mediante el diseño de bloques aleatorizado para el caso de un solo factor extraño. Así, en el ejemplo de ventas de producto mencionado antes, la variación debida a ambas tiendas y los periodos se controla mediante un diseño de cuadrado latino, lo que permite a un investigador probar la presencia de efectos debidos a diferentes configuraciones de presentación de producto.

Ejemplo 11.12 En un experimento para investigar el efecto de la humedad relativa sobre la resistencia a la abrasión de corte de cuero de un patrón rectangular ("The Abrasion of Leather", *J. Inter. Soc. Leather Trades' Chemists*, 1946: 287), se empleó un cuadrado latino de 6×6 para controlar la posible variabilidad debida a la posición del renglón y la columna en el patrón. Los seis niveles de humedad relativa estudiados fueron 1 = 25%, 2 = 37%, 3 = 50%, 4 = 62%, 5 = 75% y 6 = 87%, con los siguientes resultados:

		B (columnas)						
		1	**2**	**3**	**4**	**5**	**6**	$x_{i\cdot\cdot}$
A (renglones)	**1**	$^{3}7.38$	$^{4}5.39$	$^{6}5.03$	$^{2}5.50$	$^{5}5.01$	$^{1}6.79$	35.10
	2	$^{2}7.15$	$^{1}8.16$	$^{5}4.96$	$^{4}5.78$	$^{3}6.24$	$^{6}5.06$	37.35
	3	$^{4}6.75$	$^{6}5.64$	$^{3}6.34$	$^{5}5.31$	$^{1}7.81$	$^{2}8.05$	39.90
	4	$^{1}8.05$	$^{3}6.45$	$^{2}6.31$	$^{6}5.46$	$^{4}6.05$	$^{5}5.51$	37.83
	5	$^{6}5.65$	$^{5}5.44$	$^{1}7.27$	$^{3}6.54$	$^{2}7.03$	$^{4}5.96$	37.89
	6	$^{5}6.00$	$^{2}6.55$	$^{4}5.93$	$^{1}8.02$	$^{6}5.80$	$^{3}6.61$	38.91
	$x_{\cdot j\cdot}$	40.98	37.63	35.84	36.61	37.94	37.98	

También, $x_{\cdot\cdot 1} = 46.10$, $x_{\cdot\cdot 2} = 40.59$, $x_{\cdot\cdot 3} = 39.56$, $x_{\cdot\cdot 4} = 35.86$, $x_{\cdot\cdot 5} = 32.23$, $x_{\cdot\cdot 6} = 32.64$, $x_{\cdots} = 226.98$, y $\sum_i \sum_j x^2_{ij(k)} = 1462.89$. Los demás cálculos se resumen en la tabla 11.9.

Tabla 11.9 Tabla de ANOVA para el ejemplo 11.12

Fuente de variación	**gl**	**Suma de cuadrados**	**Cuadrado medio**	f
A (renglones)	5	2.19	0.438	2.50
B (columnas)	5	2.57	0.514	2.94
C (tratamientos)	5	23.53	4.706	26.89
Error	20	3.49	0.175	
Total	35	31.78		

Puesto que $F_{0.05,5,20} = 2.71$ y $26.89 \geq 2.71$, H_{0C} se rechaza en favor de la hipótesis de que la humedad relativa en promedio afecta la resistencia a la abrasión.

Para aplicar el procedimiento de Tukey, $w = Q_{0.05,6,20} \cdot \sqrt{\text{MSE}/6} = 4.45\sqrt{0.175/6} = 0.76$. Al ordenar las $\bar{x}_{\cdot k}$ y subrayar, se obtiene

75%	87%	62%	50%	37%	25%
5.37	5.44	5.98	6.59	6.77	7.68

En particular, la humedad relativa más baja da como resultado una resistencia a la abrasión promedio real significativamente mayor que para cualquier otra humedad relativa estudiada. ■

Ejercicios | Sección 11.3 (27 a 37)

27. Se estudió la producción de una máquina de extrusión continua que recubre tubo de acero con plástico como función del perfil de la temperatura del termostato (A, a tres niveles), tipo de plástico (B, a tres niveles), y la velocidad del tornillo rotatorio que fuerza el plástico a pasar por una matriz que forma el tubo (C, a tres niveles). Hubo dos replicaciones ($L = 2$) en cada combinación de niveles de los factores, obteniéndose un total de 54 observaciones en la producción. Las sumas de cuadrados fueron SSA = 14 144.44, SSB = 5511.27, SSC = 244 696.39, SSAB = 1069.62, SSAC = 62.67, SSBC = 331.67, SSE = 3127.50 y SST = 270 024.33.
   a. Construya la tabla ANOVA.
   b. Utilice las pruebas F apropiadas para mostrar que ninguna de las relaciones F para interacciones de dos o tres factores es significativa al nivel 0.05.
   c. ¿Qué efectos principales al parecer son significativos?
   d. Con $x_{\cdot 1\cdot} = 8242$, $x_{\cdot 2\cdot} = 9732$ y $x_{\cdot 3\cdot} = 11\,210$, utilice el procedimiento de Tukey para identificar diferencias importantes entre los niveles del factor C.

28. Para ver si la velocidad de taladrado (A), tasa de alimentación (B) o material utilizado (C) afectan la fuerza de empuje al taladrar, se realizó un experimento con cuatro velocidades, tres tasas y dos materiales, con dos muestras ($L = 2$) taladradas en cada combinación de niveles de los tres factores: SSA = 19 149.73, SSB = 2 589 047.62, SSC = 157 437.52, SSAB = 53 238.21, SSAC = 9033.73, SSBC = 91 880.04, SSE = 56 819.50 y SST = 2 983 164.81. Construya la tabla ANOVA e identifique interacciones importantes con $\alpha = 0.01$. ¿Hay algún factor simple que al parecer no tenga efecto en la fuerza de empuje? (En otras palabras, ¿algún factor es no significativo en todo efecto en el que aparece?)

29. En el artículo "An Analysis of Variance Applied to Screw Machines" (*Industrial Quality Control*, 1956: 8-9) se describe un experimento para investigar cómo la hora del día (A), el tratamiento térmico aplicado (B) y la máquina de roscado utilizada (C) afectan la longitud de barras de acero. Los tres tiempos fueron 8:00 A.M., 11:00 A.M. y 3:00 P.M., y hubo dos tratamientos y cuatro máquinas (un experimento factorial de $3 \times 2 \times 4$), y se obtuvieron los datos siguientes [codificados como 1000(longitud − 4.380), que no afecta el análisis].

B_1

	C_1	C_2	C_3	C_4
A_1	6, 9, 1, 3	7, 9, 5, 5	1, 2, 0, 4	6, 6, 7, 3
A_2	6, 3, 1, −1	8, 7, 4, 8	3, 2, 1, 0	7, 9, 11, 6
A_3	5, 4, 9, 6	10, 11, 6, 4	−1, 2, 6, 1	10, 5, 4, 8

B_2

	C_1	C_2	C_3	C_4
A_1	4, 6, 0, 1	6, 5, 3, 4	−1, 0, 0, 1	4, 5, 5, 4
A_2	3, 1, 1, −2	6, 4, 1, 3	2, 0, −1, 1	9, 4, 6, 3
A_3	6, 0, 3, 7	8, 7, 10, 0	0, −2, 4, −4	4, 3, 7, 0

Las sumas de cuadrados son SSAB = 1.646, SSAC = 71.021, SSBC = 1.542, SSE = 447.500 y SST = 1037.833.

a. Construya la tabla ANOVA para estos datos.

b. Pruebe ver si algunos de los efectos de interacción son significativos al nivel 0.05.

c. Realice una prueba para ver si algunos de los efectos principales son significativos al nivel 0.05 (es decir, H_{0A} contra H_{aA}, etcétera).

d. Utilice el procedimiento de Tukey para investigar diferencias importantes entre las cuatro máquinas.

30. Las siguientes cantidades de resumen se calcularon de un experimento relacionado con cuatro niveles de nitrógeno (A), dos tiempos de plantación (B) y dos niveles de potasio (C) ("Use and Misuse of Multiple Comparison Procedures", *Agronomy J.*, 1977: 205-208). Sólo se hizo una observación (contenido de N, en porcentaje, de grano de maíz) para cada una de 16 combinaciones de niveles.

SSA = 0.22625 SSB = 0.000025 SSC = 0.0036
SSAB = 0.004325 SSAC = 0.00065
SSBC = 0.000625 SST = 0.2384.

a. Construya la tabla de ANOVA.

b. Suponga que no hay efectos de interacción de tres vías, de modo que MSABC es una estimación válida de σ^2, y por lo menos al nivel 0.05 para efectos de interacción y principales.

c. Los promedios de nitrógeno son $\bar{x}_{1\cdot\cdot\cdot} = 1.1200$, $\bar{x}_{2\cdot\cdot\cdot} = 1.3025$, $\bar{x}_{3\cdot\cdot\cdot} = 1.3875$ y $\bar{x}_{4\cdot\cdot\cdot} = 1.4300$. Utilice el método de Tukey para examinar diferencias en el porcentaje de N entre los niveles de nitrógeno ($Q_{0.05,4,3} = 6.82$).

31. En el artículo "Kolbe-Schmitt Carbonation of 2-Naphtol" (*Industrial and Eng. Chemistry: Process and Design Development*, 1969: 165-173) se presentan los datos siguientes del rendimiento del ácido BON en porcentaje como función del tiempo de reacción (1, 2 y 3 horas), temperatura (30, 70 y 100°C) y presión (30, 70 y 100 lb/pulg^2). Suponiendo que no hay interacción de tres factores, de modo que SSE = SSABC proporciona una estimación de σ^2, MINITAB dio como resultado la tabla de ANOVA siguiente. Realice las pruebas apropiadas.

	B_1		
	C_1	C_2	C_3
A_1	68.5	73.0	68.7
A_2	74.5	75.0	74.6
A_3	70.5	72.5	74.7

	B_2		
	C_1	C_2	C_3
A_1	72.8	80.1	72.0
A_2	72.0	81.5	76.0
A_3	69.5	84.5	76.0

	B_3		
	C_1	C_2	C_3
A_1	72.5	72.5	73.1
A_2	75.5	70.0	76.0
A_3	65.0	66.5	70.5

Análisis de varianza para el rendimiento

Fuente	GL	SS	MS	F	P
tiempo	2	42.112	21.056	8.76	0.010
temp	2	110.732	55.366	23.04	0.000
pres	2	68.136	34.068	14.18	0.002
tiempo*temp	4	67.761	16.940	7.05	0.010
tiempo*pres	4	35.184	8.796	3.66	0.056
temp*pres	4	136.437	34.109	14.20	0.001
Error	8	19.223	2.403		
Total	26	479.585			

32. Cuando se fijan los factores A y B pero el factor C es aleatorio y se usa el modelo restringido (véase la nota de pie de la página 461; hay una complicación técnica aquí con el modelo irrestricto),

$$E(\text{MSE}) = \sigma^2$$

$$E(\text{MSA}) = \sigma^2 + JL\sigma^2_{AC} + \frac{JKL}{I-1}\sum \alpha_i^2$$

$$E(\text{MSB}) = \sigma^2 + IL\sigma^2_{BC} + \frac{IKL}{J-1}\sum \beta_j^2$$

$$E(\text{MSC}) = \sigma^2 + IJL\sigma^2_C$$

$$E(\text{MSAB}) = \sigma^2 + L\sigma^2_{ABC} + \frac{KL}{(I-1)(J-1)}\sum\sum (\gamma^{AB}_{ij})^2$$

$$E(\text{MSAC}) = \sigma^2 + JL\sigma^2_{AC}$$

$$E(\text{MSBC}) = \sigma^2 + IL\sigma^2_{BC}$$

$$E(\text{MSABC}) = \sigma^2 + L\sigma^2_{ABC}$$

a. Con base en los cuadrados medios esperados, ¿qué relaciones F se usarían para probar H_0: $\sigma^2_{ABC} = 0$; H_0: $\sigma^2_C = 0$; H_0: $\gamma^{AB}_{ij} = 0$ para toda i, j; y H_0: $\alpha_1 = \cdots = \alpha_I = 0$?

b. En un experimento para evaluar los efectos de la edad, tipo de suelo y día de producción en la resistencia compresiva de mezclas de cemento/suelo, se emplearon dos edades (A), cuatro tipos de suelo (B) y 3 días (C, que se supone son aleatorios), con $L = 2$ observaciones hechas para cada combinación de niveles de factor. Las sumas de cuadrados resultantes fueron SSA = 14 318.24, SSB = 9656.40, SSC = 2270.22, SSAB = 3408.93, SSAC = 1442.58, SSBC = 3096.21, SSABC = 2832.72 y SSE = 8655.60. Obtenga la tabla ANOVA y realice las pruebas con el nivel 0.01.

33. Como resultado de la posible variabilidad en el envejecimiento debida a diferentes coladas y segmentos en las coladas, se utilizó un diseño de cuadrado latino con $N = 7$ para investigar el efecto del tratamiento térmico en el envejecimiento. Con A = coladas, B = segmentos, C = tratamientos térmicos, las estadísticas de resumen son $x_{\ldots} = 3815.8$, $\sum x_{i\cdot\cdot}^2 = 297\,216.90$, $\sum x_{\cdot j\cdot}^2 = 297\,200.64$, $\sum x_{\cdot\cdot k}^2 = 297\,155.01$ y $\sum\sum x_{(ijk)}^2 = 297\,317.65$. Obtenga la tabla ANOVA y pruebe al nivel 0.05 la hipótesis de que el tratamiento térmico no tiene efecto en el envejecimiento.

34. En el artículo "The Responsiveness of Food Sales to Shelf Space Requirements" (*J. Marketing Research*, 1964: 63-67) se describe el uso de un diseño de cuadrado latino para investigar el efecto del espacio de anaquel en ventas de alimentos. El experimento se realizó en un periodo de 6 semanas con seis almacenes diferentes, y se obtuvieron los siguientes datos de las ventas de café crema en polvo (con el índice de espacio de anaquel entre paréntesis):

		Semana		
		1	2	3
Almacén	1	27 (5)	14 (4)	18 (3)
	2	34 (6)	31 (5)	34 (4)
	3	39 (2)	67 (6)	31 (5)
	4	40 (3)	57 (1)	39 (2)
	5	15 (4)	15 (3)	11 (1)
	6	16 (1)	15 (2)	14 (6)

		Semana		
		4	5	6
Almacén	1	35 (1)	28 (6)	22 (2)
	2	46 (3)	37 (2)	23 (1)
	3	49 (4)	38 (1)	48 (3)
	4	70 (6)	37 (4)	50 (5)
	5	9 (2)	18 (5)	17 (6)
	6	12 (5)	19 (3)	22 (4)

Construya la tabla ANOVA y enuncie y pruebe al nivel 0.01 la hipótesis de que el espacio de anaquel no afecta las ventas contra la alternativa apropiada.

35. En el artículo "Variation in Moisture and Ascorbic Acid Content From Leaf to Leaf and Plant to Plant in Turnip Greens" (*Southern Cooperative Services Bull.*, 1951: 13-17) se utiliza un diseño de cuadrado latino en el que el factor A es la planta, el factor B es el tamaño de la hoja (de menor a mayor), el factor C (entre paréntesis) es el tiempo de pesaje y la respuesta variable es el contenido de humedad.

		Tamaño de la hoja (B)		
		1	2	3
Planta (A)	1	6.67 (5)	7.15 (4)	8.29 (1)
	2	5.40 (2)	4.77 (5)	5.40 (4)
	3	7.32 (3)	8.53 (2)	8.50 (5)
	4	4.92 (1)	5.00 (3)	7.29 (2)
	5	4.88 (4)	6.16 (1)	7.83 (3)

		Tamaño de la hoja (B)	
		4	5
Planta (A)	1	8.95 (3)	9.62 (2)
	2	7.54 (1)	6.93 (3)
	3	9.99 (4)	9.68 (1)
	4	7.85 (5)	7.08 (4)
	5	5.83 (2)	8.51 (5)

Cuando todos los factores son aleatorios, los cuadrados medios esperados son $E(\text{MSA}) = \sigma^2 + N\sigma_\alpha^2$, $E(\text{MSB}) = \sigma^2 + N\sigma_\beta^2$, $E(\text{MSC}) = \sigma^2 + N\sigma_\gamma^2$ y $E(\text{MSE}) = \sigma^2$. Esto significa que las relaciones F para probar H_{0A}: $\sigma_\alpha^2 = 0$, H_{0B}: $\sigma_\beta^2 = 0$ y H_{0C}: $\sigma_\gamma^2 = 0$ son idénticas a las de efectos fijos. Obtenga la tabla ANOVA y pruebe al nivel 0.05 para ver si hay alguna variación en el contenido de humedad debida a los factores.

36. En el artículo "An Assessment of the Effects of Treatment, Time, and Heat on the Removal of Erasable Pen Marks from Cotton and Cotton/Polyester Blend Fabrics (*J. Testing and Eval.*, 1991: 394-397) se informan las siguientes sumas de cuadrados para la variable de respuesta *grado de eliminación de marcas*: SSA = 39.171, SSB = 0.665, SSC = 21.508, SSAB = 1.432, SSAC = 15.953, SSBC = 1.382, SSABC = 9.016 y SSE = 115.820. En el experimento se usaron cuatro tratamientos de lavandería, tres tipos de pluma y seis telas distintas, y se hicieron tres observaciones para cada combinación de tratamiento-pluma-fábrica. Efectúe un análisis de varianza con $\alpha = 0.01$ para cada prueba, y enuncie sus conclusiones (suponga efectos fijos para los tres factores).

37. Se realizó un experimento ANOVA de cuatro factores para investigar los efectos de la tela (A), tipo de exposición (B), nivel de exposición (C) y dirección de la tela (D) en el grado del cambio de color en la tela expuesta medido con un espectrocolorímetro. Se hicieron dos observaciones para cada una de las tres telas, dos tipos, tres niveles y dos direcciones. Los resultados son MSA = 2207.329, MSB = 47.255, MSC = 491.783, MSD = 0.044, MSAB = 15.303, MSAC = 275.446, MSAD = 0.470, MSBC = 2.141, MSBD = 0.273, MSCD = 0.247, MSABC = 3.714, MSABD = 4.072, MSACD = 0.767, MSBCD = 0.280, MSE = 0.977 y MST = 93.621 ("Accelerated Weathering of Marine Fabrics", *J. Testing and Eval.*, 1992: 139-143). Suponiendo efectos fijos para los factores, realice un análisis de varianza con $\alpha = 0.01$ para las pruebas y resuma sus conclusiones.

11.4 Experimentos factoriales 2^p

Si un experimentador desea estudiar en forma simultánea el efecto de p distintos factores en una respuesta variable y los factores tienen $I_1, I_2, \ldots, I_p$ niveles, respectivamente, entonces un experimento completo requiere al menos $I_1, I_2, \ldots, I_p$ observaciones. En tales situaciones, el experimentador con frecuencia puede efectuar un "experimento de selección" con cada factor en sólo dos niveles para obtener información preliminar acerca de efectos de factor. Un experimento en el que hay p factores, cada uno en dos niveles, se conoce como **experimento factorial 2^p.** El análisis de datos de tal experimento es más simple, desde el punto de vista del cálculo, que para experimentos factoriales más generales. Además, un experimento 2^p proporciona un contexto simple para introducir los conceptos importantes de replicaciones que causan confusión y réplicas fraccionarias.

Experimentos 2^3

Al igual que en la sección 11.3, X_{ijkl} y x_{ijkl} se refieren a la observación de la l-ésima replicación con factores A, B y C a los niveles i, j y k, respectivamente. El modelo para esta situación es

$$X_{ijkl} = \mu + \alpha_i + \beta_j + \delta_k + \gamma_{ij}^{AB} + \gamma_{ik}^{AC} + \gamma_{jk}^{BC} + \gamma_{ijk} + \epsilon_{ijkl} \qquad (11.14)$$

para $i = 1, 2$; $j = 1, 2$; $k = 1, 2$; $l = 1, \ldots, n$. Se supone que las ϵ_{ijkl} son independientes, distribuidas normalmente, con media 0 y varianza σ^2. Debido a que sólo hay dos niveles de cada factor, las condiciones adicionales en los parámetros de (11.14) que especifican unívocamente el modelo se expresan de forma simple: $\alpha_1 + \alpha_2 = 0, \ldots, \gamma_{11}^{AB} + \gamma_{21}^{AB} = 0$, $\gamma_{12}^{AB} + \gamma_{22}^{AB} = 0$, $\gamma_{11}^{AB} + \gamma_{12}^{AB} = 0$, $\gamma_{21}^{AB} + \gamma_{22}^{AB} = 0$, etcétera. Estas condiciones indican que sólo hay un parámetro funcionalmente independiente de cada tipo (para cada efecto principal e interacción). Por ejemplo, $\alpha_2 = -\alpha_1$, mientras que $\gamma_{21}^{AB} = -\gamma_{11}^{AB}$ $\gamma_{12}^{AB} = -\gamma_{11}^{AB}$ y $\gamma_{22}^{AB} = \gamma_{11}^{AB}$. Como resultado de esto, cada suma de cuadrados en el análisis tendrá 1 gl.

Los parámetros del modelo se estiman al tomar promedios en varios subíndices de las X_{ijkl} y luego formar combinaciones lineales apropiadas de los promedios. Por ejemplo,

$$\hat{\alpha}_1 = \overline{X}_{1\ldots} - \overline{X}_{\ldots}$$
$$= \frac{(X_{111\cdot} + X_{121\cdot} + X_{112\cdot} + X_{122\cdot} - X_{211\cdot} - X_{212\cdot} - X_{221\cdot} - X_{222\cdot})}{8n}$$

y

$$\hat{\gamma}_{11}^{AB} = \overline{X}_{11\cdot\cdot} - \overline{X}_{1\ldots} - \overline{X}_{\cdot 1\cdot\cdot} + \overline{X}_{\ldots}$$
$$= \frac{(X_{111\cdot} - X_{121\cdot} - X_{211\cdot} + X_{221\cdot} + X_{112\cdot} - X_{122\cdot} - X_{212\cdot} + X_{222\cdot})}{8n}$$

Cada estimador es, excepto para el factor $1/(8n)$, una función lineal de los totales de celda (X_{ijk}) en la que cada coeficiente es $+1$ o -1, con número igual de cada una; esta clase de funciones se llaman **contrastes** en las X_{ijk}. Además, los estimadores satisfacen las mismas condiciones adicionales que satisfacen los parámetros. Por ejemplo,

$$\hat{\alpha}_1 + \hat{\alpha}_2 = \overline{X}_{1\ldots} - \overline{X}_{\ldots} + \overline{X}_{2\ldots} - \overline{X}_{\ldots} = \overline{X}_{1\ldots} + \overline{X}_{2\ldots} - 2\overline{X}_{\ldots}$$
$$= \frac{1}{4n}X_{1\ldots} + \frac{1}{4n}X_{2\ldots} - \frac{2}{8n}X_{\ldots} = \frac{1}{4n}X_{\ldots} - \frac{1}{4n}X_{\ldots} = 0$$

Ejemplo 11.13 En un experimento para investigar las propiedades de resistencia compresiva de mezclas de cemento-suelo, se emplearon dos periodos de envejecimiento en combinación con dos temperaturas de envejecimiento distintas y dos tipos de suelo. Se hicieron dos replicaciones para cada combinación de niveles de los tres factores, obteniéndose como resultado los siguientes datos:

		Suelo	
Edad	**Temperatura**	**1**	**2**
1	1	471, 413	385, 434
	2	485, 552	530, 593
2	1	712, 637	770, 705
	2	712, 789	741, 806

Los totales de celda calculados son $x_{111\cdot} = 884$, $x_{211\cdot} = 1349$, $x_{121\cdot} = 1037$, $x_{221\cdot} = 1501$, $x_{112\cdot} = 819$, $x_{212\cdot} = 1475$, $x_{122\cdot} = 1123$ y $x_{222\cdot} = 1547$, por lo tanto $x_{\ldots} = 9735$. Entonces

$$\hat{\alpha}_1 = (884 - 1349 + 1037 - 1501 + 819 - 1475 + 1123 - 1547)/16$$
$$= -125.5625 = -\hat{\alpha}_2$$
$$\hat{\gamma}_{11}^{AB} = (884 - 1349 - 1037 + 1501 + 819 - 1475 - 1123 + 1547)/16$$
$$= -14.5625 = -\hat{\gamma}_{12}^{AB} = -\hat{\gamma}_{21}^{AB} = \hat{\gamma}_{22}^{AB}$$

Las otras estimaciones de parámetro se calculan de la misma manera. ■

Análisis de un experimento 2^3 La razón de calcular las estimaciones de parámetros es que las sumas de cuadrados para los distintos efectos se obtienen fácilmente de las estimaciones. Por ejemplo,

$$\text{SSA} = \sum_i \sum_j \sum_k \sum_l \hat{\alpha}_i^2 = 4n \sum_{i=1}^{2} \hat{\alpha}_i^2 = 4n[\hat{\alpha}_1^2 + (-\hat{\alpha}_1)^2] = 8n\hat{\alpha}_1^2$$

y

$$\text{SSAB} = \sum_i \sum_j \sum_k \sum_l (\hat{\gamma}_{ij}^{AB})^2$$

$$= 2n \sum_{i=1}^{2} \sum_{j=1}^{2} (\hat{\gamma}_{ij}^{AB})^2 = 2n[(\hat{\gamma}_{11}^{AB})^2 + (-\hat{\gamma}_{11}^{AB})^2 + (-\hat{\gamma}_{11}^{AB})^2 + (\hat{\gamma}_{11}^{AB})^2]$$

$$= 8n(\hat{\gamma}_{11}^{AB})^2$$

Puesto que cada estimación es un contraste en los totales de celda multiplicados por $1/(8n)$, cada suma de cuadrados tiene la forma (contraste)$^2/8n$. Así que para calcular las distintas sumas de cuadrados, es necesario conocer los coeficientes (+1 o −1) de los contrastes apropiados. La forma más conveniente de mostrar los signos (+ o −) en cada x_{ijk} de cada contraste de efecto es en forma tabular. Se empleará la notación (1) para la condición experimental $i = 1, j = 1, k = 1$, a para $i = 2, j = 1, k = 1$, ab para $i = 2, j = 2, k = 1$, etcétera. Si se considera al nivel 1 como "bajo" y el nivel 2 como "alto", cualquier letra que aparezca denota un nivel alto del factor relacionado. En la tabla 11.10, cada columna proporciona los signos para un contraste de efecto particular en las x_{ijk} relacionadas con las diferentes condiciones experimentales.

Tabla 11.10 Signos para calcular contrastes de efecto

Condición experimental	Total de celda	Efecto factorial						
		A	*B*	*C*	*AB*	*AC*	*BC*	*ABC*
(1)	$x_{111\cdot}$	−	−	−	+	+	+	−
a	$x_{211\cdot}$	+	−	−	−	−	+	+
b	$x_{121\cdot}$	−	+	−	−	+	−	+
ab	$x_{221\cdot}$	+	+	−	+	−	−	−
c	$x_{112\cdot}$	−	−	+	+	−	−	+
ac	$x_{212\cdot}$	+	−	+	−	+	−	−
bc	$x_{122\cdot}$	−	+	+	−	−	+	−
abc	$x_{222\cdot}$	+	+	+	+	+	+	+

En cada una de las tres primeras columnas, el signos es + si el factor correspondiente está en el nivel alto y − si está en el nivel bajo. Entonces todo signo de la columna *AB* es el "producto" de los signos de las columnas *A* y *B* con (+)(+) = (−)(−) = + y (+)(−) = (−)(+) = −, y de manera similar para las columnas *AC* y *BC*. Por último, los signos de la columna *ABC* son los productos de *AB* con *C* (o *B* con *AC* o *A* con *BC*). Así, por ejemplo,

$$AC \text{ contraste} = + x_{111\cdot} - x_{211\cdot} + x_{121\cdot} - x_{221\cdot} - x_{112\cdot} + x_{212\cdot} - x_{122\cdot} + x_{222\cdot}$$

Una vez que se calculan los siete contrastes de efecto,

$$\text{SS(efecto)} = \frac{(\text{contraste efecto})^2}{8n}$$

Incluso con una tabla de signos, el cálculo de los contrastes es tedioso. Una técnica de cálculo eficaz, debida a Yates, es como sigue. Escriba en una columna los ocho totales en el **orden estándar** de la misma forma que en la tabla de signos, y establezca tres columnas adicionales. En cada una de estas tres columnas, las primeras cuatro entradas son las sumas de las

entradas 1 y 2, 3 y 4 y 6, 7 y 8 de las columnas previas. Los últimos cuatro elementos son las diferencias entre los elementos 2 y 1, 4 y 3, 6 y 5 y 8 y 7 de la columna previa. Entonces la última columna contiene $x_{\ldots}$ y los siete contrastes de efecto en orden estándar. Al elevar al cuadrado cada contraste y dividir entre $8n$ se obtienen las siete sumas de cuadrados.

Ejemplo 11.14 (Continuación del ejemplo 11.13) Puesto que $n = 2$, $8n = 16$. El método de Yates se ilustra en la tabla 11.11.

Tabla 11.11 Método de cálculo de Yates

Condición de tratamiento	$x_{ijk\cdot}$	1	2	Contraste de efecto	SS = (contraste)²/16
$(1) = x_{111\cdot}$	884	2233	4771	9735	
$a = x_{211\cdot}$	1349	2538	4964	2009	252 255.06
$b = x_{121\cdot}$	1037	2294	929	681	28 985.06
$ab = x_{221\cdot}$	1501	2670	1080	−233	3393.06
$c = x_{112\cdot}$	819	465	305	193	2328.06
$ac = x_{212\cdot}$	1475	464	376	151	1425.06
$bc = x_{122\cdot}$	1123	656	−1	71	315.06
$abc = x_{222\cdot}$	1547	424	−232	−231	3335.06
					292 036.42

De los datos originales, $\sum_i \sum_j \sum_k \sum_l x_{ijkl}^2 = 6\,232\,289$ y

$$\frac{x_{\ldots}^2}{16} = 5\,923\,139.06$$

así,

$$\text{SST} = 6\,232\,289 - 5\,923\,139.06 = 309\,149.94$$

$$\text{SSE} = \text{SST} - [\text{SSA} + \cdots + \text{SSABC}] = 309\,149.94 - 292\,036.42$$

$$= 17\,113.52$$

Los cálculos del ANOVA se resumen en la tabla 11.12.

Tabla 11.12 Tabla de ANOVA para el ejemplo 11.14

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
A	1	252 255.06	252 255.06	117.92
B	1	28 985.06	28 985.06	13.55
C	1	2328.06	2328.06	1.09
AB	1	3393.06	3393.06	1.59
AC	1	1425.06	1425.06	0.67
BC	1	315.06	315.06	0.15
ABC	1	3335.06	3335.06	1.56
Error	8	17 113.52	2139.19	
Total	15	309 149.94		

En la figura 11.7 se muestra el resultado de SAS para este ejemplo. Sólo los valores P para edad (A) y temperatura (B) son menores que 0.01, así que se considera que sólo estos efectos son importantes.

Análisis del procedimiento de varianza
Variable dependiente: RESISTENCIA

Fuente	GL	Suma de cuadrados	Cuadrado medio	Valor F	Pr > F
Modelo	7	292036.4375	41719.4911	19.50	0.0002
Error	8	17113.5000	2139.1875		
Total corregido	15	309149.9375			

Cuadrado de R	V.C.	Raíz de MSE	USO DE ENERGÍA promedio
0.944643	7.601660	46.25135	608.437500

Fuente	GL	SS de Anova	Cuadrado medio	Valor F	Pr > F
EDAD	1	252255.0625	252255.0625	117.92	0.0001
TEMPO	1	28985.0625	28985.0625	13.55	0.0062
EDAD*TIEMPO	1	3393.0625	3393.0625	1.59	0.2434
SUELO	1	2328.0625	2328.0625	1.09	0.3273
EDAD*SUELO	1	1425.0625	1425.0625	0.67	0.4380
TEMP*SUELO	1	315.0625	315.0625	0.15	0.7111
EDAD*TEMP*SUELO	1	3335.0625	3335.0625	1.56	0.2471

Figura 11.7 Resultado de SAS para los datos de resistencia del ejemplo 11.14 ∎

Experimentos 2^p para $p > 3$

Aunque los cálculos cuando $p > 3$ son bastante tediosos, el análisis se asemeja al del caso de tres factores. Por ejemplo, si hay cuatro factores A, B, C y D, hay 16 condiciones experimentales diferentes. Los primeros 8 en el orden estándar son exactamente los listados para un experimento de tres factores. Los segundos 8 se obtienen al remplazar la letra d además de cada condición en el primer grupo. El método de Yates se inicia entonces calculando los totales en las replicaciones, listando estos totales en orden estándar y procediendo como antes; con p factores, la p-ésima columna a la derecha de los totales de tratamiento producirán los contrastes de efecto.

Para $p > 3$, no habrá replicaciones del experimento (así que sólo se tiene disponible un duplicado completo). Una posible forma de probar hipótesis, es suponer que están ausentes ciertos efectos de orden superior y luego sumar las sumas de cuadrado correspondientes para obtener una SSE. Sin embargo, una suposición de este tipo puede ser confusa en ausencia de conocimiento previo (véase el libro de Montgomery listado en la bibliografía del capítulo). Un método alternativo tiene que ver con trabajar de manera directa con los contrastes de efecto. Cada contraste tiene una distribución normal con la misma varianza. Cuando está ausente un efecto particular, el valor esperado del contraste correspondiente es 0, pero esto no es así cuando está presente el efecto. El método sugerido de análisis es construir una gráfica de probabilidad normal de los contrastes de efecto (o, de manera equivalente, las estimaciones del parámetro de efecto, puesto que estimación = contraste/2^p cuando $n = 1$). Los puntos correspondientes a efectos ausentes tenderán a caer cerca de una recta, mientras que los puntos relacionados con efectos sustanciales por lo común quedarán lejos de esta recta.

Ejemplo 11.15 Los datos siguientes se tomaron del artículo "Quick and Easy Analysis of Unreplicated Factorials" (*Technometrics*, 1989: 469-473). Los cuatro factores son A = resistencia al ácido, B = tiempo, C = cantidad de ácido y D = temperatura, y la variable de respuesta es el rendimiento de isatina. Las observaciones, en orden estándar son 0.08, 0.04, 0.53, 0.43, 0.31, 0.09, 0.12, 0.36, 0.79, 0.68, 0.73, 0.08, 0.77, 0.38, 0.49 y 0.23. En la tabla 11.13 se muestran las estimaciones de efecto como aparecen el artículo (en el que se utiliza constraste/8 en vez de contraste/16).

Tabla 11.13 Estimaciones de efecto para el ejemplo 11.15

Estimación del efecto	A	B	AB	C	AC	BC	ABC	D
	−0.191	−0.021	−0.001	−0.076	0.034	−0.066	0.149	0.274
Estimación del efecto	AD	BD	ABD	CD	ACD	BCD	$ABCD$	
	−0.161	−0.251	−0.101	−0.026	−0.066	0.124	0.019	

La figura 11.8 es una gráfica de probabilidad normal de las estimaciones de efecto. Los puntos de la gráfica caen cerca de la misma recta, lo que hace pensar en la ausencia completa de cualquier efecto (en breve se da un ejemplo en el que éste no es el caso).

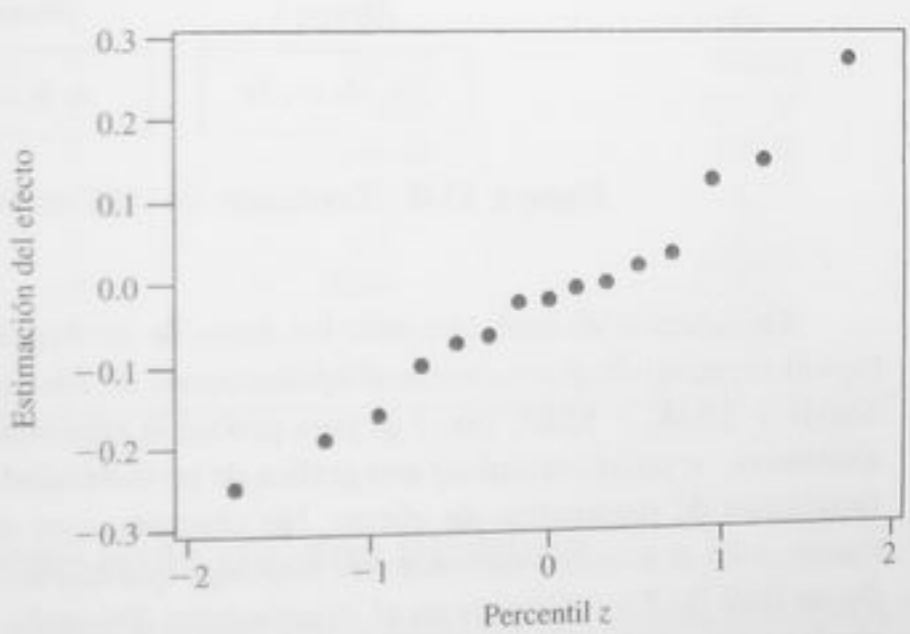


Figura 11.8 Gráfica de probabilidad normal de las estimaciones de efecto del ejemplo 11.15 ■

Los juicios visuales de desviación de la linealidad en una gráfica de probabilidad normal son bastante subjetivos. En el artículo citado en el ejemplo 11.15 se describe una técnica más objetiva para identificar efectos significativos en un experimento sin replicación.

Confusión

Con frecuencia se imposibilita llevar a cabo 2^p condiciones experimentales de un experimento factorial 2^p en un ambiente homogéneo. En esta clase de situaciones, podría ser posible separar las condiciones experimentales en 2^r bloques homogéneos ($r < p$), de modo que hay 2^{p-r}

condiciones experimentales en cada bloque. Los bloques podrían, por ejemplo, corresponder a diferentes laboratorios, periodos diferentes o distintos operadores o equipos de trabajo. En el caso más simple, $p = 3$ y $r = 1$, así que hay dos bloques constituidos cada uno por ocho condiciones experimentales.

Como siempre, el bloqueo es eficaz para reducir la variación relacionada con fuentes extrañas. Sin embargo, cuando las 2^p condiciones experimentales se colocan en dos bloques, el precio que se paga por este bloqueo es que no se pueden estimar $2^r - 1$ de los efectos de factor. Esto es porque $2^r - 1$ efectos de factor (efectos principales o interacciones, o ambos) se mezclan o **confunden** con los efectos de bloque. La asignación de condiciones experimentales a los bloques por lo general se hace de modo que sólo se confundan las interacciones de nivel superior, mientras que los efectos principales y las interacciones de orden bajo sean estimables y se puedan probar las hipótesis.

Para ver cómo se lleva a cabo la asignación de bloques, considere primero un experimento 2^3 con dos bloques ($r = 1$) y cuatro tratamientos por bloque. Suponga que se selecciona ABC como el efecto por confundir con bloques. Entonces cualquier condición experimental que tiene un número impar de letras en común con ABC, como por ejemplo b (una letra) o abc (tres letras), se coloca en un bloque, mientras que cualquier condición que tiene un número par de letras en común con ABC (donde 0 es par) va en el otro bloque. En la figura 11.9 se muestra esta asignación de tratamiento de los dos bloques.

Bloque 1	*Bloque 2*
(1), *ab*, *ac*, *bc*	*a*, *b*, *c*, *abc*

Figura 11.9 Confusión de ABC en un experimento 2^3

En ausencia de replicaciones, los datos de un experimento normalmente se analizarían bajo el supuesto de que no hubiera interacciones de dos factores (aditividad) y usando SSE = SSAB + SSAC + SSBC con 3 gl para probar la presencia de efectos principales. De manera alternativa, se podría examinar una gráfica de probabilidad normal de contrastes de efecto o estimaciones de parámetros de efecto. No obstante, con mucha frecuencia hay replicaciones cuando sólo se estudian tres factores. Suponga que hay u replicaciones, que dan como resultado un total de $2^r \cdot u$ bloques en el experimento. Después de restar de SST las sumas de cuadrados relacionadas con los efectos no confundidos con los bloques (calculados por medio del método de Yates), la suma de cuadrados de bloque se calcula por medio de los $2^r \cdot u$ totales de bloque y luego se resta para obtener SSE (de modo que hay $2^r \cdot u - 1$ gl para los bloques).

Ejemplo 11.16 En el artículo "Factorial Experiments in Pilot Plant Studies" (*Industrial and Eng. Chemistry*, 1951:1300-1306) aparecen los resultados de un experimento para evaluar los efectos de la temperatura del reactor (A), producción de gas (B) y concentración del constituyente activo (C) en la resistencia de la solución del producto (medida en unidades arbitrarias) en una unidad de recirculación. Se emplearon dos bloques, donde el efecto ABC se confundió con los bloques, y hubo dos replicaciones, obteniéndose los datos de la figura 11.10. Los cuatro totales de bloque × replicación son 288, 212, 88 y 220, con un gran total de 808, así que

$$\text{SSBl} = \frac{(288)^2 + (212)^2 + (88)^2 + (220)^2}{4} - \frac{(808)^2}{16} = 5204.00$$

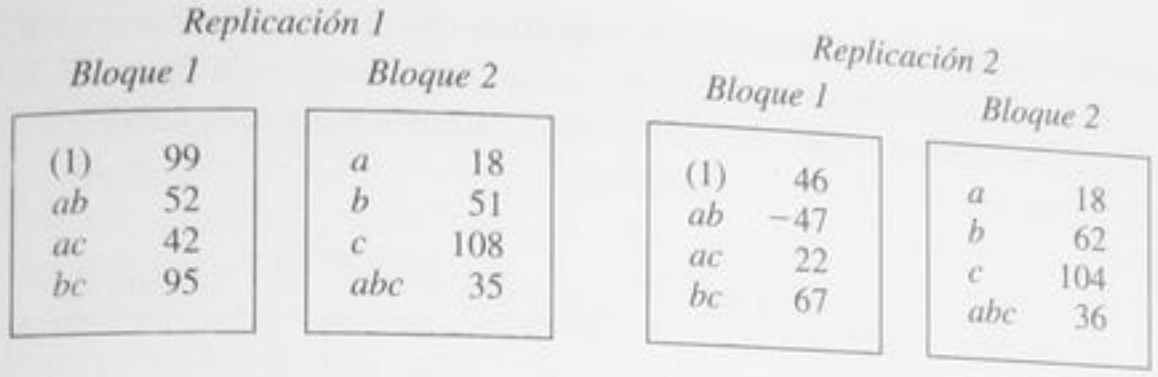


Figura 11.10 Datos para el ejemplo 11.16

Las otras sumas de cuadrados se calculan con el método de Yates usando los ocho totales de condición experimental, y se obtiene la tabla ANOVA ilustrada en la tabla 11.14. Por comparación con $F_{0.05,1,6} = 5.99$, se concluye que sólo los efectos principales para A y C difieren de manera importante de cero.

Tabla 11.14 Tabla de ANOVA para el ejemplo 11.16

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
A	1	12 996	12 996	39.82
B	1	702.25	702.25	2.15
C	1	2 756.25	2756.25	8.45
AB	1	210.25	210.25	0.64
AC	1	30.25	30.25	0.093
BC	1	25	25	0.077
Bloques	3	5204	1734.67	5.32
Error	6	1958	326.33	
Total	15	23 882		

■

Confusión con más de dos bloques

En el caso $r = 2$ (cuatro bloques), se confundieron con bloques tres efectos. El experimentador elige primero dos efectos definitorios por confundir. Por ejemplo, en un experimento de cinco factores (A, B, C, D y E), se podría elegir que las dos interacciones de tres factores por confundir fueran BCD y CDE. Entonces, el tercer efecto confundido es la **generalización** de los dos **interacciones**, obtenida al escribir los dos efectos elegidos lado con lado y cancelar después letras en común para ambos: $(BCD)(CDE) = BE$. Observe que si se eligen ABC y CDE para confundir, su interacción generalizada es $(ABC)(CDE) = ABDE$, así que no se confunden efectos principales o interacciones de dos factores.

Una vez seleccionados los dos efectos por confundir, un bloque consiste en las condiciones de tratamiento que tiene un número par de letras en común con ambos defectos definitorios. El segundo bloque consiste en las condiciones que tiene un número par de letras en común con el primer definitorio y un número impar de letras en común con el segundo contraste, y el tercero y cuarto bloques consisten en los contrastes "impar/par" e "impar/impar".

En un experimento de cinco factores con efectos definitorios ABC y CDE, esto da como resultado la asignación a bloques como se muestra en la figura 11.11 (con el número de letras en común con cada definitorio que aparece junto a cada condición experimental).

Bloque 1		*Bloque 2*		*Bloque 3*		*Bloque 4*	
(1)	(0, 0)	*d*	(0, 1)	*a*	(1, 0)	*c*	(1, 1)
ab	(2, 0)	*e*	(0, 1)	*b*	(1, 0)	*ad*	(1, 1)
de	(0, 2)	*ac*	(2, 1)	*cd*	(1, 2)	*ae*	(1, 1)
acd	(2, 2)	*bc*	(2, 1)	*ce*	(1, 2)	*bd*	(1, 1)
ace	(2, 2)	*abd*	(2, 1)	*ade*	(1, 2)	*be*	(1, 1)
bcd	(2, 2)	*abe*	(2, 1)	*bde*	(1, 2)	*abc*	(3, 1)
bce	(2, 2)	*acde*	(2, 3)	*abcd*	(3, 2)	*cde*	(1, 3)
abde	(2, 2)	*bcde*	(2, 3)	*abce*	(3, 2)	*abcde*	(3, 3)

Figura 11.11 Cuatro bloques de un experimento factorial 2^5 con efectos definitorios ABC y CDE

El bloque que contiene al (1) se llama **bloque principal**. Una vez construido, un segundo bloque se obtiene al seleccionar cualquier condición experimental que no sea del bloque principal y obtener su interacción generalizada con cada condición en el bloque principal. Los otros bloques se construyen de la misma forma al seleccionar primero una condición en un bloque que ya se construyó y hallar condiciones generalizadas con el bloque principal.

Para las situaciones experimentales con $p > 3$, con frecuencia no hay replicación, así que las sumas de cuadrados con interacciones no confusas de orden superior por lo común se agrupan para obtener una suma de cuadrados de errores que se pueda emplear en los denominadores de las distintas estadísticas. Los cálculos de nuevo se realizan mediante el método de Yates, con SSBl como la suma de sumas de cuadrados relacionada con los efectos confundidos.

Cuando $r > 2$, primero se eligen r efectos definitorios por confundir con los bloques, asegurándose de que ninguno de los efectos elegidos es la interacción generalizada de otros dos seleccionados. Los $2^r - r - 1$ efectos confundidos con los bloques son las interacciones generalizadas de los efectos en el conjunto definitorio (incluso no sólo las interacciones generalizadas de los pares de efectos, sino también de conjuntos de tres, cuatro, etcétera). Consulte los detalles en el libro de Montgomery.

Replicación fraccional

Cuando el número p de factores es grande, incluso una sola replicación de un experimento 2^p puede ser cara y tardada. Por ejemplo, una replicación de un experimento fraccional 2^6 requiere una observación de cada una de 64 condiciones experimentales diferentes. Una estrategia interesante en estas situaciones es hacer observaciones para sólo una fracción de las 2^p condiciones. Siempre y cuando se ponga cuidado en la elección de las condiciones a observar, se puede obtener mucha información acerca de los efectos de factor.

Suponga que se decide incluir sólo 2^{p-1} (la mitad) de las 2^p condiciones posibles en el experimento; esto por lo general se llama **semirreplicación**. El precio que se paga por economizar el doble. Primero, la información acerca de un solo efecto (determinado por las 2^{p-1} con-

diciones seleccionadas para observación) se pierde por completo para el experimentador en el sentido de que no es posible ninguna estimación razonable del efecto. Segundo, con los $2^n - 2$ efectos principales e interacciones restantes se forman pares hasta que algún efecto de un par determinado se confunda con el otro efecto del mismo par. Por ejemplo, un par podría ser $\{A, BCD\}$, de modo que las estimaciones separadas del efecto principal A y la interacción BCD no sean posibles. Entonces es deseable seleccionar una semirreplicación para la cual los efectos principales y las interacciones de orden bajo formen pares (confundidos) sólo con las interacciones de orden superior y con ninguna otra.

El primer paso para seleccionar una semirreplicación es seleccionar un efecto definitorio como el efecto no estimable. Suponga que en un experimento de cinco factores se elija a $ABCDE$ como el efecto definitorio. Ahora las $2^5 = 32$ condiciones de tratamiento posibles se dividen en dos grupos con 16 condiciones cada uno, uno que consiste en las condiciones que tienen un número impar de letras en común con $ABCDE$ y el otro que contiene un número par de letras en común con el contraste definitorio. Entonces cada grupo de 16 condiciones se usa como la semirreplicación. El grupo "impar" es

$$a, b, c, d, e, abc, abd, abe, acd, ace, ade, bcd, bce, bde, cde, abcde$$

Cada efecto principal e interacción distinta de $ABCDE$ se confunden con (tiene como **alias** a) su interacción generalizada con $ABCDE$. Por lo tanto, $(AB)(ABCDE) = CDE$, de modo que la interacción AB y la interacción CDE se confunden entre sí. Los **pares de alias** resultantes son

$\{A, BCDE\}$	$\{B, ACDE\}$	$\{C, ABDE\}$	$\{D, ABCE\}$	$\{E, ABCD\}$
$\{AB, CDE\}$	$\{AC, BDE\}$	$\{AD, BCE\}$	$\{AE, BCD\}$	$\{BC, ADE\}$
$\{BD, ACE\}$	$\{BE, ACD\}$	$\{CD, ABE\}$	$\{CE, ABD\}$	$\{DE, ABC\}$

Observe en particular que todo efecto principal tiene como alias una interacción de cuatro factores. Suponer que estas interacciones son insignificantes permite probar la presencia de efectos principales.

Para seleccionar un cuarto de replicación de un experimento factorial 2^p (2^{p-2} de las 2^p condiciones de tratamientos posibles), se seleccionan dos efectos definitorios. Estas dos y su interacción generalizada se convierten de efectos no estimables. En lugar de los pares de alias como en la semirreplicación, ahora cada efecto restante se confunde con otros tres factores; cada uno es su interacción generalizada con uno de los tres efectos no estimables.

Ejemplo 11.17 En el artículo "More on Planning Experiments to Increase Research Efficiency" (*Industrial and Eng. Chemistry*, 1970: 60-65) se encuentran los resultados de un cuarto de replicación de un experimento 2^5 en el cual los cinco factores fueron A = temperatura de condensación, B = cantidad de material B, C = volumen de disolvente, D = tiempo de condensación y E = cantidad de material E. La variable de respuesta fue el rendimiento del proceso químico. Los contrastes definitorios elegidos fueron ACE y BDE, con interacción generalizada $(ACE)(BDE) = ABCD$. Los 28 efectos principales restantes e interacciones se dividen en siete grupos de cuatro efectos cada uno tal que los efectos dentro de un grupo no se puedan evaluar por separado. Por ejemplo, las interacciones generalizadas de A con los efectos no estimables son $(A)(ACE) = CE$, $(A)(BDE) = ABDE$, y $(A)(ABCD) = BCD$, de manera que un grupo de alias es $\{A, CE, ABDE, BCD\}$. El conjunto completo de grupos de alias es

$\{A, CE, ABDE, BCD\}$ $\{B, ABCE, DE, ACD\}$ $\{C, AE, BCDE, ABD\}$
$\{D, ACDE, BE, ABC\}$ $\{E, AC, BD, ABCDE\}$ $\{AB, BCE, ADE, CD\}$
$\{AD, CDE, ABE, BC\}$

■

Análisis de una replicación fraccional Una vez elegidos los contrastes definitorios para un cuarto de replicación, se emplean como en el análisis de la confusión para dividir las 2^p condiciones de tratamiento en cuatro grupos de 2^{p-2} condiciones cada uno. Entonces se selecciona cualquiera de los cuatro grupos como el conjunto de condiciones para el que se reunirán los datos. Se aplican comentarios similares a una replicación $1/2^r$ de un experimento factorial 2^p.

Una vez realizadas las observaciones para las combinaciones de tratamiento seleccionadas, se construye una tabla de signos similar a la tabla 11.10. La tabla contiene un solo renglón para cada una de las combinaciones de tratamiento observadas en realidad en vez de los 2^p renglones, y hay una sola columna para cada grupo de alias (puesto que cada efecto en el grupo tendría el mismo conjunto de signos para las condiciones de tratamiento seleccionadas para observación). Como siempre, los signos en cada columna indican cómo se calculan los contrastes para las distintas sumas de cuadrados. También se puede usar el método de Yates, pero la regla se debe modificar para disponer las condiciones observadas en el orden estándar.

La parte difícil de un análisis de replicación fraccional por lo común tiene que ver con decidir qué usar para la suma de cuadrados de errores. Puesto que normalmente no habrá replicación (aunque uno podría observar, p. ej., dos réplicas de un cuarto de réplica), algunas sumas de cuadrados de efectos se deben combinar para obtener una suma de cuadrados de errores. En una media réplica de un experimento 2^8, por ejemplo, se puede elegir una estructura de alias para que los ocho efectos principales y las 28 interacciones de dos factores se confundan sólo con interacciones de orden superior y haya otros 27 grupos de alias relacionados sólo con interacciones de orden superior. Suponiendo la ausencia de efectos de orden superior, las 27 sumas de cuadrados resultantes se suman para obtener una suma de cuadrados de errores, lo que permite pruebas de 1 gl para los efectos principales e interacciones de dos factores. Sin embargo, en muchos casos las pruebas para efectos principales sólo se obtienen combinando algunas de las sumas de cuadrados relacionadas con grupos de alias relacionados con interacciones de dos factores, y no se pueden investigar las interacciones correspondientes de dos factores.

Ejemplo 11.18 (Continuación del ejemplo 11.17) El conjunto de condiciones de tratamiento elegido y los rendimientos resultantes para el cuarto de réplica del experimento 2^5 fueron

e	ab	ad	bc	cd	ace	bde	$abcde$
23.2	15.5	16.9	16.2	23.8	23.4	16.8	18.1

La tabla de signos abreviada se muestra en la tabla 11.15.

Con SSA que denota la suma de cuadrados para los efectos en el grupo de alias $\{A, CE, ABDE, BCD\}$,

$$\text{SSA} = \frac{(-23.2 + 15.5 + 16.9 - 16.2 - 23.8 + 23.4 - 16.8 + 18.1)^2}{8} = 4.65$$

De manera similar, SSB = 53.56, SSC = 10.35, SSD = 0.91, SSE′ = 10.35 (la ′ diferencia esta cantidad de la suma de cuadrados de errores SSE), SSAB = 6.66 y SSAD = 3.25 obteniéndose SST = 4.65 + 53.56 + · · · + 3.25 = 89.73. Para probar los efectos principales, se utiliza SSE = SSAB + SSAD = 9.91 con 2 gl. La tabla de ANOVA se ilustra en la tabla 11.16.

Tabla 11.15 Tabla de signos para el ejemplo 11.18

	A	*B*	*C*	*D*	*E*	*AB*	*AD*
e	−	−	−	−	+	+	+
ab	+	+	−	−	−	+	−
ad	+	−	−	+	−	−	+
bc	−	+	+	−	−	−	+
cd	−	−	+	+	−	+	−
ace	+	−	+	−	+	−	−
bde	−	+	−	+	+	−	−
abcde	+	+	+	+	+	+	+

Tabla 11.16 Tabla ANOVA para el ejemplo 11.18

Fuente	gl	Suma de cuadrados	Cuadrado medio	*f*
A	1	4.65	4.65	0.94
B	1	53.56	53.56	10.80
C	1	10.35	10.35	2.09
D	1	0.91	0.91	0.18
E	1	10.35	10.35	2.09
Error	2	9.91	4.96	
Total	7	89.73		

Puesto que $F_{0.05,1,2} = 18.51$, ninguno de los cinco efectos principales se puede considerar significativo. Por supuesto, con sólo dos gl para el error, la prueba no es muy poderosa (es decir, es bastante probable que no detecte la presencia de efectos). El artículo de *Industrial and Engineering Chemistry* del que en realidad provinieron los datos tuvo una estimación independiente del error estándar de los efectos de tratamiento con base en la experiencia previa, así que se empleó un análisis un poco diferente. El análisis realizado aquí se hizo sólo para propósitos ilustrativos, puesto que por lo común se desearía mucho más que dos grados de libertad para el error. ■

Como una alternativa para pruebas *F* basadas en sumas de cuadrados de agrupamiento para obtener SSE, se puede examinar una gráfica de probabilidad normal de contrastes de efecto.

Ejemplo 11.19 Se realizó un experimento para investigar la contracción de la cubierta de plástico utilizada para cables de velocímetro ("An Explanation and Critique of Taguchi's Contribution to Quality Engineering," *Quality and Reliability Engr. Intl.*, 1988: 123-131). Los ingenieros comenzaron con 15 factores: diámetro exterior del forro, matriz del forro, material de forro, velocidad de línea de forro, tipo de trenzado del alambre, tensión del trenzado, diámetro del alambre, tensión lineal, temperatura del forro, material de cobertura, tipo de matriz para la cobertura, temperatura de fusión, paquete protector, método de enfriamiento y velocidad de línea. Se sospechó que sólo algunos de estos factores fueron importantes, así que se realizó un experimento de discriminación en la forma de un factorial 2^{15-11} (una fracción $1/2^{11}$ de un experimento fraccional 2^{15}). La estructura de alias resultante es bastante complicada; en particular, todo efecto principal se confunde con interacciones de dos factores. La variable de respuesta fue el porcentaje de contracción para una muestra de cable producida a niveles diseñados de los factores.

En la figura 11.12 se muestra una gráfica de probabilidad normal de los contrastes de efecto. Todos, excepto dos de los puntos caen bastante cerca de una recta. Los puntos discrepantes corresponden a los efectos E = tipo de trenzado del alambre y G = diámetro de alambre, lo que hace pensar que estos dos factores son los únicos que afectan la cantidad de contracción.

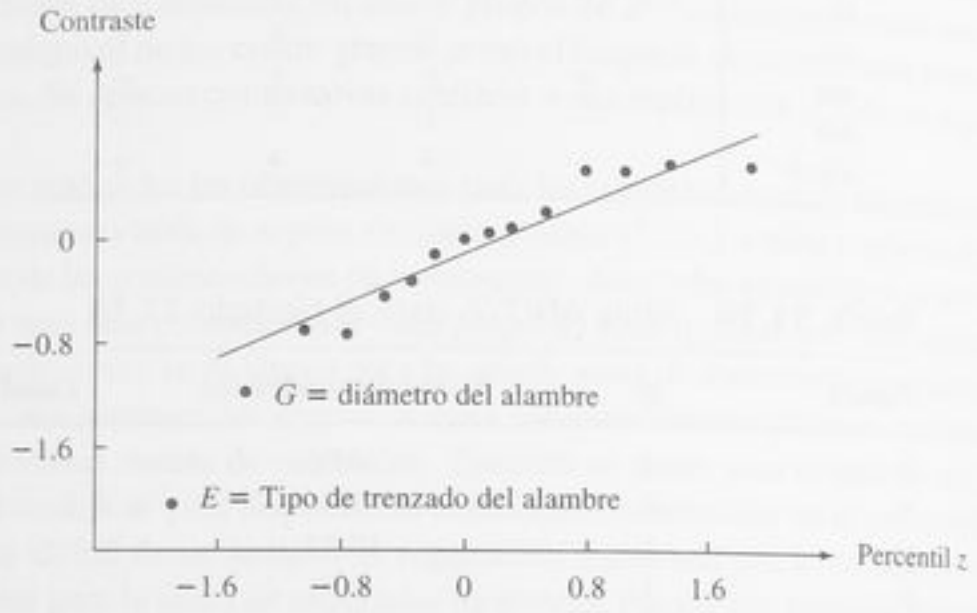


Figura 11.12 Gráfica de probabilidad normal de contrastes del ejemplo 11.19 ∎

Los temas de la experimentación factorial, confusión y replicación factorial abarcan muchos modelos y técnicas que se han analizado aquí. Para más información consulte las referencias del capítulo.

Ejercicios | Sección 11.4 (38 a 49)

38. Los datos siguientes se obtuvieron de un experimento para estudiar la naturaleza de la dependencia de la corriente de soldadura en tres factores: voltaje de soldadura, velocidad de alimentación del alambre y distancia de la punta a la pieza de trabajo. Hubo dos niveles de cada factor (un experimento 2^3) con dos replicaciones por combinación de niveles (los promedios en las replicaciones concuerdan con los valores proporcionados en el artículo "A Study on Prediction of Welding Current in Gas Metal Arc Welding", *J. Engr. Manuf.*, 1991: 64-69). Los dos primeros números son para el tratamiento (1), los dos siguientes para a, y así sucesivamente en orden estándar: 200.0, 204.2, 215.5, 219.5, 272.7, 276.9, 299.5, 302.7, 166.6, 172.6, 186.4, 192.0, 232.6, 240.8, 253.4, 261.6.

a. Compruebe que las sumas de cuadrados son como las que se proporcionan en la siguiente tabla ANOVA de MINITAB.

b. ¿Qué efectos al parecer son importantes y por qué?

```
Análisis de varianza para la corriente
Fuente            GL      SC      CM       F      P
Volt               1  1685.1  1685.1  102.38  0.000
Velocidad          1 21272.2 21272.2 1292.37  0.000
Dist               1  5076.6  5076.6  308.42  0.000
Volt*velocidad     1    36.6    36.6    2.22  0.174
Volt*dist          1     0.4     0.4    0.03  0.877
Velocidad*dist     1   109.2   109.2    6.63  0.033
Volt*velocidad*
  dist             1    23.5    23.5    1.43  0.266
Error              8   131.7    16.5
Total             15 28335.3
```

39. Los datos siguientes resultaron de un experimento 2^3 con tres replicaciones por combinación de tratamientos diseñados para estudiar los efectos de la concentración de detergente (A), concentración de carbonato de sodio (B) y concentración de carboximetil celulosa (C) en la capacidad de limpieza de una solución en pruebas de lavado (un número más grande indica mejor capacidad de limpieza que un número pequeño).

Niveles de factor				
A	*B*	*C*	Condición	Observaciones
1	1	1	(1)	106, 93, 116
2	1	1	*a*	198, 200, 214
1	2	1	*b*	197, 202, 185
2	2	1	*ab*	329, 331, 307
1	1	2	*c*	149, 169, 135
2	1	2	*ac*	243, 247, 220
1	2	2	*bc*	255, 230, 252
2	2	2	*abc*	383, 360, 364

a. Después de obtener los totales de celda x_{ijk}, calcule las estimaciones de β_1, γ_{11}^{AC} y γ_{21}^{AC}.

b. Utilice los totales de celda junto con el método de Yates para calcular los contrastes de efecto y las sumas de cuadrados. Luego construya una tabla ANOVA y pruebe las hipótesis apropiadas con $\alpha = 0.05$.

40. En un estudio de procesos utilizados para eliminar impurezas de artículos de celulosa ("Optimization of Rope-Range Bleaching of Cellulosic Fabrics", *Textile Research J.*, 1976: 493-496), se obtuvieron los datos siguientes de un experimento 2^4 relacionado con un proceso de desencolado. Los cuatro factores fueron concentración de la enzima (*A*), pH (*B*), temperatura (*C*) y tiempo (*D*).

Tratamiento	Enzima (g/L)	pH	Temp. (°C)	Tiempo (h)	Almidón % en peso 1a réplica	Almidón % en peso 2a réplica
(1)	0.50	6.0	60.0	6	9.72	13.50
a	0.75	6.0	60.0	6	9.80	14.04
b	0.50	7.0	60.0	6	10.13	11.27
ab	0.75	7.0	60.0	6	11.80	11.30
c	0.50	6.0	70.0	6	12.70	11.37
ac	0.75	6.0	70.0	6	11.96	12.05
bc	0.50	7.0	70.0	6	11.38	9.92
abc	0.75	7.0	70.0	6	11.80	11.10
d	0.50	6.0	60.0	8	13.15	13.00
ad	0.75	6.0	60.0	8	10.60	12.37
bd	0.50	7.0	60.0	8	10.37	12.00
abd	0.75	7.0	60.0	8	11.30	11.64
cd	0.50	6.0	70.0	8	13.05	14.55
acd	0.75	6.0	70.0	8	11.15	15.00
bcd	0.50	7.0	70.0	8	12.70	14.10
abcd	0.75	7.0	70.0	8	13.20	16.12

a. Utilice el algoritmo de Yates para obtener sumas de cuadrados y la tabla ANOVA.

b. ¿Al parecer están presentes efectos de interacción de segundo, tercer o cuarto orden? Explique su razonamiento. ¿Cuáles efectos principales al parecer son significativos?

41. En el ejercicio 39, suponga que se utilizó una temperatura baja del agua para obtener los datos. El experimento completo se repite con una mayor temperatura del agua para obtener los datos siguientes. Utilice el algoritmo de Yates en el conjunto de 48 observaciones para obtener las sumas de cuadrados y la tabla ANOVA, y luego pruebe las hipótesis apropiadas al nivel 0.05.

Condición	Observaciones
d	144, 154, 158
ad	239, 227, 244
bd	232, 242, 246
abd	364, 362, 346
cd	194, 162, 203
acd	284, 295, 291
bcd	291, 287, 297
abcd	411, 406, 395

42. Los datos siguientes acerca del consumo de energía en hornos eléctricos (kW consumidos por tonelada de producto fundido) se obtuvieron de un experimento factorial 2^4 con tres réplicas ("Studies on a 10-cwt Arc Furnace", *J. Iron and Steel Institute*, 1956: 22). Los factores fueron la naturaleza del techo *A* (baja, alta), selección de potencia *B* (alta, baja), chatarra empleada *C* (tubo, placa) y carga *D* (700 lb, 1000 lb).

Tratamiento	x_{ijklm}	Tratamiento	x_{ijklm}
(1)	866, 862, 800	*d*	988, 808, 650
a	946, 800, 840	*ad*	966, 976, 876
b	774, 834, 746	*bd*	702, 658, 650
ab	709, 789, 646	*abd*	784, 700, 596
c	1017, 990, 954	*cd*	922, 808, 868
ac	1028, 906, 977	*acd*	1056, 870, 908
bc	817, 783, 771	*bcd*	798, 726, 700
abc	829, 806, 691	*abcd*	752, 714, 714

Construya la tabla ANOVA y pruebe las hipótesis de interés con $\alpha = 0.01$.

43. En el artículo "Statistical Design and Analysis of Qualification Test Program for a Small Rocket Engine" (*Industrial Quality Control*, 1964: 14-18) aparecen los datos de un experimento para evaluar los efectos de la vibración (*A*), ciclo de temperatura (*B*), ciclo de altitud (*C*) y temperatura para el ciclo de altitud y encendido (*D*) en la duración del empuje. A continuación se proporciona un subconjunto de los datos. (En el artículo, hubo cuatro niveles de *D* en vez de sólo dos.) Utilice el método de Ya-

tes para obtener las sumas de cuadrados y la tabla ANOVA. Luego, suponga que están ausentes las interacciones de tres y cuatro factores, combine las sumas de cuadrados correspondientes para obtener una estimación de σ^2 y pruebe las hipótesis apropiadas al nivel 0.05.

		D_1		D_2	
		C_1	C_2	C_1	C_2
A_1	B_1	21.60	21.60	11.54	11.50
	B_2	21.09	22.17	11.14	11.32
A_2	B_1	21.60	21.86	11.75	9.82
	B_2	19.57	21.85	11.69	11.18

44. a. En un experimento 2^4, suponga que se emplean dos bloques y se decide confundir la interacción *ABCD* con el efecto de bloque. ¿Qué tratamientos se deben realizar en el primer bloque [el que contiene el tratamiento (1)], y qué tratamientos se asignan al segundo bloque?

b. En un experimento para investigar la retención de niacina en verduras como función de la temperatura de cocción (*A*), tamaño del colador (*B*), tipo de procesamiento (*C*) y tiempo de cocción (*D*), cada factor se mantuvo a dos niveles. Se emplearon dos bloques, con la asignación de bloques del inciso (a) para confundir sólo la interacción *ABCD* con los bloques. Utilice el procedimiento de Yates para obtener la tabla ANOVA para los datos siguientes.

Tratamiento	x_{ijkl}	Tratamiento	x_{ijkl}
(1)	91	*d*	72
a	85	*ad*	78
b	92	*bd*	68
ab	94	*abd*	79
c	86	*cd*	69
ac	83	*acd*	75
bc	85	*bcd*	72
abc	90	*abcd*	71

c. Suponga que están ausentes los efectos de interacción de tres vías, de modo que las sumas de cuadrados se combinan para producir una estimación de σ^2, y realice las pruebas apropiadas al nivel 0.05.

45. a. Se llevó a cabo un experimento para investigar los efectos en la sensibilidad de audio de resistencia variable (*A*), dos capacitancias (*B*, *C*) e inductancia de una bobina (*D*) en parte de un circuito de televisión. Si se emplearon cuatro bloques con cuatro tratamientos por bloque, y los efectos definitorios para la confusión fueron *AB* y *CD*, ¿qué tratamientos aparecieron en cada bloque?

b. Suponga que se llevaron a cabo dos replicaciones del experimento en el inciso (a) y se obtuvieron los siguientes datos. Obtenga la tabla ANOVA y pruebe las hipótesis relevantes al nivel 0.01.

Tratamiento	x_{ijkl1}	x_{ijkl2}	Tratamiento	x_{ijkl1}	x_{ijkl2}
(1)	618	598	*d*	598	585
a	583	560	*ad*	587	541
b	477	525	*bd*	480	508
ab	421	462	*abd*	462	449
c	601	595	*cd*	603	577
ac	550	589	*acd*	571	552
bc	505	484	*bcd*	502	508
abc	452	451	*abcd*	449	455

46. En un experimento relacionado con cuatro factores (*A*, *B*, *C* y *D*) y cuatro bloques, muestre que por lo menos un efecto principal o el efecto de interacción de dos factores se debe confundir con el efecto de bloque.

47. a. En un experimento de siete factores (*A*, . . . , *G*), suponga que en realidad se realiza un cuarto de réplica. Si los efectos definitorios son *ABCDE* y *CDEFG*, ¿cuál es el tercer efecto no estimable y qué tratamientos del grupo contienen (1)? ¿Cuáles son los grupos de alias de los siete efectos principales?

b. Si el cuarto de réplica se va a realizar usando cuatro bloques (con ocho tratamientos por bloque), ¿cuáles son los bloques si los efectos de confusión elegidos son *ACF* y *BDG*?

48. Suponga que en el problema del empuje del cohete del ejercicio 43, hubiera habido recursos suficientes para sólo una mitad de réplica del experimento 2^4.

a. Si se elige el efecto *ABCD* como el efecto definitorio para la réplica y el grupo de ocho tratamientos para los que se obtienen los datos incluye el tratamiento (1), ¿qué otros tratamientos están en el grupo observado y cuáles son los pares de alias?

b. Suponga que los resultados de llevar a cabo el experimento descrito en el inciso (a) son los que se dan a continuación (proporcionados en orden después de borrar la mitad no observada). Suponiendo que las interacciones de dos y tres factores son insignificantes, pruebe al nivel 0.05 la presencia de efectos principales. También construya una gráfica de probabilidad normal.

19.09 20.11 21.66 20.44
13.72 11.26 11.72 12.29

49. Una mitad de réplica de un experimento 2^5 para investigar los efectos del tiempo del calentamiento (A), tiempo de temple (B), tiempo de estirado (C), posición de las espiras de calentamiento (D) y posición de medición (E) en la dureza de coladas de acero produjo los siguientes datos. Construya la tabla de ANOVA, y (suponiendo que las interacciones de segundo y tercer orden son insignificantes) pruebe al nivel 0.01 la presencia de efectos principales. También construya una gráfica de probabilidad normal.

Tratamiento	Observación	Tratamiento	Observación
a	70.4	acd	66.6
b	72.1	ace	67.5
c	70.4	ade	64.0
d	67.4	bcd	66.8
e	68.0	bce	70.3
abc	73.8	bde	67.9
abd	67.0	cde	65.9
abe	67.8	$abcde$	68.0

Ejercicios complementarios (50 a 61)

50. Los resultados de un estudio acerca de la eficacia de secado en tendedero sobre la suavidad de telas se resumen en el artículo "Line-Dried vs. Machine-Dried Fabrics: Comparison of Appearance, Hand, and Consumer Acceptance" (*Home Econ. Research J.*, 1984: 27-35). Se asignaron puntuaciones de suavidad para nueve tipos de tela y cinco métodos de secado: (1) secado a máquina, (2) secado en tendedero, (3) secado en tendedero seguido de 15 minutos de secado en tambor giratorio, (4) secado en tendedero con suavizante y (5) secado en tendedero con movimiento de aire. Considerando los diferentes tipos de tela como bloques, construya una tabla ANOVA. Por medio de un nivel de significación 0.05, pruebe si hay una diferencia en la puntuación de suavidad promedio real para los métodos de secado.

		Método de secado				
		1	2	3	4	5
Tela	Crepé	3.3	2.5	2.8	2.5	1.9
	Doble punto	3.6	2.0	3.6	2.4	2.3
	Tela cruzada	4.2	3.4	3.8	3.1	3.1
	Mezcla de tela cruzada	3.4	2.4	2.9	1.6	1.7
	Felpa	3.8	1.3	2.8	2.0	1.6
	Paño fino	2.2	1.5	2.7	1.5	1.9
	Tela de sábanas	3.5	2.1	2.8	2.1	2.2
	Pana	3.6	1.3	2.8	1.7	1.8
	Mezclilla	2.6	1.4	2.4	1.3	1.6

51. La absorción de agua de dos tipos de mortero empleados para reparar cemento dañado, se estudió en el artículo "Polymer Mortar Composite Matrices for Maintenance-Free, Highly Durable Ferrocement" (*J. Ferrocement*, 1964: 337-345). Se sumergieron especímenes de mortero de cemento ordinario (MCO) y mortero de cemento polimérico (MCP) durante tiempos variables (5, 9, 24 o 48 horas) y se registró la absorción de agua (% en peso). Con el tipo de mortero como factor A (con dos niveles) y el periodo de inmersión como factor B (con cuatro niveles), se hicieron tres observaciones por cada combinación de nivel de factores. Se utilizaron los datos incluidos en el artículo para calcular las sumas de cuadrados, que fueron SSA = 322.667, SSB = 35.623, SSAB = 8.557 y SST = 372.113. Utilice esta información para construir una tabla ANOVA. Pruebe las hipótesis apropiadas en un nivel de significación 0.05.

52. En cuatro parcelas se realizó un experimento para comparar la acumulación de tréboles para cuatro proporciones de siembra ("Performance of Overdrilled Red Clover with Different Sowing Rates and Initial Grazing Managements", *N. Zeal J. Exp. Ag.*, 1984: 71-81). Puesto que antes del experimento se apacentó ganado de distinta manera en las cuatro parcelas y se pensó que esto podría afectar la acumulación de tréboles, se utilizó un experimento de bloques aleatorizado con las cuatro proporciones de siembra en una sección de cada parcela. Utilice los datos para probar la hipótesis nula de ninguna diferencia en la acumulación promedio real de tréboles (kg DM/ha) para las diferentes proporciones de siembra.

		Proporción de siembra (kg/ha)			
		3.6	6.6	10.2	13.5
Parcela	1	1155	2255	3505	4632
	2	123	406	564	416
	3	68	416	662	379
	4	62	75	362	564

53. En un proceso automatizado de recubrimiento químico, la velocidad a la que los objetos en una banda transportadora pasan por una dispersión química (velocidad de

la banda), la cantidad de sustancia atomizada (volumen de aspersión) y la marca (marca) de la sustancia son factores que podrían afectar la uniformidad del recubrimiento aplicado. En un esfuerzo por incrementar la uniformidad del recubrimiento se efectuó un experimento 2^3 con réplica. En la siguiente tabla, los valores mayores de la variable de respuesta se relacionan con mayor uniformidad en la superficie:

				Uniformidad de la superficie	
Corrida	Volumen de dispersión	Velocidad de la banda	Marca	Replicación 1	Replicación 2
1	−	−	−	40	36
2	+	−	−	25	28
3	−	+	−	30	32
4	+	+	−	50	48
5	−	−	+	45	43
6	+	−	+	25	30
7	−	+	+	30	29
8	+	+	+	52	49

Analice estos datos y escriba sus conclusiones.

54. Las plantas de energía alimentadas con carbón mineral utilizadas en la industria eléctrica, han llamado cada vez más la atención pública por los problemas ambientales relacionados con los desechos sólidos generados en la combustión a gran escala ("Fly Ash Binders in Stabilization of FGD Wastes", *J. of Environmental Engineering*, 1998: 43-49). Se efectuó un estudio para analizar la influencia de tres factores: tipo de aglomerante (*A*), cantidad de agua (*B*) y condiciones de eliminación en terreno (*C*), que afectan ciertas características de lixiviación de desechos sólidos de la combustión. Cada factor se estudió en dos niveles. Se realizó un experimento 2^3 sin réplica y se midió un valor de respuesta EC50 (la concentración efectiva, en mg/L, que disminuye la luz en 50% en un bioensayo de luminiscencia) para cada combinación de niveles de factor. Los datos experimentales son los siguientes:

	Factor			Respuesta
Corrida	*A*	*B*	*C*	EC50
1	−1	−1	−1	23 100
2	1	−1	−1	43 000
3	−1	1	−1	71 400
4	1	1	−1	76 000
5	−1	−1	1	37 000
6	1	−1	1	33 200
7	−1	1	1	17 000
8	1	1	1	16 500

Realice un análisis de varianza adecuado y escriba sus conclusiones.

55. Las impurezas en forma de óxidos de hierro disminuyen el valor económico y la utilidad de minerales industriales, como el caolín, utilizados en las industrias cerámica y de procesamiento de papel. Se efectuó un experimento 2^4 para evaluar los efectos de cuatro factores sobre el porcentaje de hierro eliminado de muestras de caolín ("Factorial Experiments in the Development of a Kaolin Bleaching Process Using Thiourea in Sulphuric Acid Solutions", *Hydrometallurgy*, 1997: 181-197). Los factores y sus niveles se listan en la siguiente tabla:

Factor	Descripción	Unidades	Nivel bajo	Nivel alto
A	H_2SO_4	M	0.10	0.25
B	Toiurea	g/L	0.0	5.0
C	Temperatura	°C	70	90
D	Tiempo	min.	30	150

En la tabla siguiente se enumeran los datos de un experimento 2^4 sin réplica.

Corrida de prueba	Extracción de hierro (%)	Corrida de prueba	Extraccion de hierro (%)
(1)	7	*d*	28
a	11	*ad*	51
b	7	*bd*	33
ab	12	*abd*	57
c	21	*cd*	70
ac	41	*acd*	95
bc	27	*bcd*	77
abc	48	*abcd*	99

a. Calcule estimaciones de los efectos principales y efectos de interacción de dos factores para este experimento.

b. Trace una gráfica de probabilidad de los efectos. ¿Cuáles efectos al parecer son importantes?

56. En silvicultura se utilizan diseños factoriales para evaluar los efectos de varios factores en el crecimiento de árboles. En un experimento, los investigadores consideraron que los almácigos de abeto saludables debían germinar más pronto que los de abeto enfermos ("Practical Analysis of Factorial Experiments in Forestry", *Canadian J. of Forestry*, 1995: 446-461). Además, antes de la siembra, las semillas se expusieron a tres niveles de pH para ver si este factor tiene algún efecto sobre la captación de virus en la raíz. La siguiente tabla muestra los datos de un experimento 2×3 para estudiar ambos factores:

		pH		
		3	5.5	7
Condición	Enferma	1.2, 1.4, 1.0, 1.2, 1.4	0.8, 0.6, 0.8, 1.0, 0.8	1.0, 1.0, 1.2, 1.4, 1.2
	Saludable	1.4, 1.6, 1.6, 1.6, 1.4	1.0, 1.2, 1.2, 1.4, 1.4	1.2, 1.4, 1.2, 1.2, 1.4

La variable de respuesta es una evaluación promedio de cinco brotes de un almácigo. Las calificaciones son 0 (sin germinar), 1 (expandido en parte) y 2 (expandido por completo). Analice estos datos.

57. Una propiedad de las bolsas de aire de automóviles que contribuye a su capacidad para absorber energía es la permeabilidad (pies^3/pies^2/min) del tejido con el que se fabrican. Es importante entender cómo diversos factores influyen en la permeabilidad para aumentar la eficacia de las bolsas de aire. En un estudio se analizaron los efectos de tres factores, cada uno a tres niveles ("Analysis of Fabrics used in Passive Restraint Systems-Airbags", *J. of the Textile Institute*, 1996: 554-571):

A (Temperatura): 8°C, 50°C, 75°C

B (Denier de la tela): 420-D, 630-D, 840-D

C (Presión de aire): 17.2 kPa, 34.4 kPa, 103.4 kPa

Temperatura 8°

Denier	Presión 17.2	34.4	103.4
420-D	73	157	332
	80	155	322
630-D	35	91	288
	433	98	271
840-D	125	234	477
	111	233	464

Temperatura 50°

Denier	Presión 17.2	34.4	103.4
420-D	52	125	281
	51	118	264
630-D	16	72	169
	12	78	173
840-D	96	149	338
	100	155	350

Temperatura 75°

Denier	Presión 17.2	34.4	103.4
420-D	37	95	276
	31	106	281
630-D	30	91	213
	41	100	211
840-D	102	170	307
	98	160	311

Analice estos datos y enuncie sus conclusiones (suponga que los factores son fijos).

58. Un ingeniero químico realizó un experimento para estudiar los efectos de los factores fijos presión del tanque (*A*), tiempo de cocción de la pulpa (*B*) y concentración de madera dura (*C*) en la resistencia del papel. En el experimento intervienen dos presiones, cuatro tiempos de cocción, tres concentraciones y dos observaciones en cada combinación de estos niveles. Las sumas de cuadrados calculadas son SSA = 6.94, SSB = 5.61, SSC = 12.33, SSAB = 4.05, SSAC = 7.32, SSBC = 15.80, SSE = 14.40 y SST = 70.82. Construya la tabla ANOVA y realice las pruebas apropiadas al nivel de significación 0.05.

59. Se estudió la resistencia de unión al montar un circuito integrado en un substrato de vidrio metalizado como función del factor *A* = tipo de adhesivo, factor *B* = tiempo de curado y factor *C* = material conductor (cobre y níquel). A continuación se proporcionan los datos junto con una tabla ANOVA de MINITAB. ¿Qué conclusiones se obtienen de los datos?

Cobre		Tiempo cuadrado 1	2	3
Adhesivo	1	72.7	74.6	80.0
		80.0	77.5	82.7
	2	77.8	78.5	84.6
		75.3	81.1	78.3
	3	77.3	80.9	83.9
		76.5	82.6	85.0
Níquel		**1**	**2**	**3**
Adhesivo	1	74.7	75.7	77.2
		77.4	78.2	74.6
	2	79.3	78.8	83.0
		77.8	75.4	83.9
	3	77.2	84.5	89.4
		78.4	77.5	81.2

```
Análisis de varianza para la resistencia
Fuente              GL      SC      SM      F      P
Adhesivo             2  101.317  50.659  6.54  0.007
Tiempo curado        2  151.317  75.659  9.76  0.001
Mat. cond.           1    0.722   0.722  0.09  0.764
Adhes*tcura          4   30.526   7.632  0.98  0.441
Adhes*mcond          2    8.015   4.008  0.52  0.605
Tcura*mcond          2    5.952   2.976  0.38  0.687
Adhes*tcura*mcond    4   33.298   8.325  1.07  0.398
Error               18  139.515   7.751
Total               35  470.663
```

60. En el artículo "Food Consumption and Energy Requirements of Captive Bald Eagles" (*J. Wildlife Mgmt.*, 1982: 646-654) se investigó la ingestión diaria de energía promedio (la variable de respuesta) para diferentes dietas (factor *A*, con tres niveles) y la temperatura (factor *B*, con tres niveles). Las cantidades de resumen del artículo se utilizaron para generar datos, y los resultados fueron SSA = 18 138, SSB = 5182, SSAB = 1737, SST = 36 348 y grados de libertad del error = 36. Construya una tabla ANOVA y pruebe las hipótesis pertinentes.

61. Análogo a un cuadrado latino, un diseño de cuadrado grecolatino se emplea cuando se sospecha que tres factores extraños podrían afectar la variable de respuesta y los cuatro factores (los tres extraños y el de interés) tienen el mismo número de niveles. En un cuadrado latino, cada nivel del factor de interés (*C*) aparece una vez en cada renglón (con cada nivel de *A*) y una vez en cada columna (con cada nivel de *B*). En un cuadrado grecolatino, cada nivel del factor *D* aparece una vez en cada renglón, en cada columna y también con cada nivel del tercer factor extraño *C*. En otro caso, el diseño se puede utilizar cuando los cuatro factores sean de igual interés, el número de niveles de cada uno sea *N* y se disponga de recursos sólo para N^2 observaciones. En (a) se ilustra un cuadrado de 5×5, con (k, l) en cada celda que denota el *k*-ésimo nivel de *C* y el *l*-ésimo nivel de *D*. En (b) se muestran los datos de la pérdida de peso en barras de silicio empleadas para material semiconductor como función del volumen de grabado (*A*), color del ácido nítrico de la solución de grabado (*B*), tamaño de las barras (C) y tiempo en la solución de grabado (*D*) (de "Applications of Analytic Techniques to the Semiconductor Industry", Fourteenth Midwest Quality Control Conference, 1959).

Sea $X_{(ijkl)}$ la pérdida de peso observada cuando el factor *A* está a nivel *i*, *B* al nivel *j*, *C* al nivel *k* y *D* al nivel *l*. Si se supone que no hay interacción entre factores, la suma total de cuadrados SST (con $N^2 - 1$ grados de libertad) se divide en SSA, SSB, SSC, SSD y SSE. Determine las expresiones para estas sumas de cuadrados, incluso las fórmulas de cálculo, obtenga la tabla ANOVA para los datos y pruebe cada una de las cuatro hipótesis de efecto principal mediante $\alpha = 0.05$.

(C, D)		*B* 1	2	3	4	5
	1	(1, 1)	(2, 3)	(3, 5)	(4, 2)	(5, 4)
	2	(2, 2)	(3, 4)	(4, 1)	(5, 3)	(1, 5)
A	3	(3, 3)	(4, 5)	(5, 2)	(1, 4)	(2, 1)
	4	(4, 4)	(5, 1)	(1, 3)	(2, 5)	(3, 2)
	5	(5, 5)	(1, 2)	(2, 4)	(3, 1)	(4, 3)

(a)

65	82	108	101	126
84	109	73	97	83
105	129	89	89	52
119	72	76	117	84
97	59	94	78	106

(b)

Bibliografía

Box, George, William Hunter y Stuart Hunter, *Statistics for Experimenters*, Wiley, Nueva York, 1978. Contiene una vasta cantidad de sugerencias e ideas acerca del análisis de datos con base en la amplia experiencia de los autores como consultores.

DeVor, R., T. Chang y J. W. Sutherland, *Statistical Quality Design and Control*, Macmillan, Nueva York, 1992. Contiene un moderno estudio de la experimentación factorial y factorial fraccional con un mínimo de matemáticas.

Hocking, Ronald, *The Analysis of Linear Models*, Brooks/Cole, Pacific Grove, CA, 1985. Un tratamiento muy general del análisis de varianza, escrito por una de las autoridades más importantes en este campo.

Kleinbaum, David, Lawrence Kupper, Keith Muller y Azhar Nizam, *Applied Regression Analysis and Other Multivariable Methods* (3a. ed.), Duxbury Press, Pacific Grove, 1998. Contiene un análisis particularmente bueno de problemas relacionados con el análisis de "datos no equilibrados", es decir, K_{ij} distintas.

Kuehl, Robert O., *Statistical Principles of Research Design and Analysis*, Wadsworth, Belmont, CA. 1994. Un tratamiento actualizado y completo de diseño de experimentos y análisis de los datos resultantes.

Montgomery, Douglas. *Design and Analysis of Experiments* (5a. ed.), Wiley, Nueva York, 2001. Véase la bibliografía del capítulo 10.

Neter, John, William Wasserman y Michael Kutner, *Applied Linear Statistical Models* (4a. ed.), Irwin, Homewood, IL, 1996. Véase la bibliografía del capítulo 10.

Vardeman, Stephen, *Statistics for Engineering Problem Solving*, PWS, Boston, 1994. Una introducción general para ingenieros, con mucha metodología descriptiva e inferencial para datos de experimentos diseñados.

12

Regresión lineal simple y correlación

Introducción

En los problemas de dos muestras analizados en el capítulo 9, se trató con la comparación de valores de parámetros para la distribución x y la distribución y. Aun cuando las observaciones se combinaron en pares, no se intentó usar información acerca de una variable para estudiar la otra variable. Éste es precisamente el objetivo del análisis de regresión: explotar la relación entre dos (o más) variables, de modo que se obtenga información acerca de una de ellas al conocer valores de la(s) otra(s).

Gran parte de las matemáticas se dedica a estudiar variables que están relacionadas determinísticamente. Decir que x y y están relacionadas de esta manera, significa que una vez que se sabe el valor de x, el valor de y se especifica por completo. Por ejemplo, suponga que se decide rentar una camioneta durante un día, y que el costo de la renta es de 25 dólares más \$30 dólares por milla conducida. Si x = número de millas conducidas y y = cobro por la renta, entonces $y = 25 + 0.3x$. Si se conduce la camioneta 100 millas ($x = 100$), entonces $y = 25 + 0.3(100) = 55$. Como otro ejemplo, si la velocidad inicial de una partícula es v_0 y experimenta una aceleración constante a, entonces la distancia recorrida $= y = v_0x + \frac{1}{2}ax^2$, donde x = tiempo.

Hay muchas variables x y y que parecerían estar relacionadas entre sí, pero no en un modo determinístico. Un ejemplo familiar a muchos estudiantes lo proporcionan las variables x = calificación promedio de la secundaria (CP) y CP de la

universidad. El valor de y no se puede determinar sólo por conocer x, y dos estudiantes podrían tener el mismo valor de x, pero valores de y distintos. Sin embargo, hay una tendencia porque los estudiantes que tienen CP altas (bajas) en la preparatoria también tienen CP altas (bajas) en la universidad. Conocer la CP de un estudiante de preparatoria debe ser bastante útil para poder predecir cuál será el desempeño de esa persona en la universidad.

Otros ejemplos de variables relacionadas de un modo determinístico son x = edad de un niño y y = tamaño del vocabulario de ese niño, x = tamaño de un motor en centímetros cúbicos y y = rendimiento de combustible para un automóvil equipado con ese motor, y x = fuerza de tensión aplicada y y = cantidad de elongación de una tira de metal.

El **análisis de regresión** es la parte de la estadística que trata con la investigación de la relación entre dos o más variables relacionadas de un modo no determinístico. En este capítulo, se generaliza la relación lineal determinística $y = \beta_0 + \beta_1 x$ a una relación probabilística lineal, se desarrollan procedimientos para hacer inferencias acerca de los parámetros del modelo y se obtiene una media cuantitativa (el coeficiente de correlación) del grado de relación de las dos variables. En el capítulo siguiente, se consideran técnicas para validar un modelo particular e investigar relaciones no lineales y relaciones en las que intervienen más de dos variables.

12.1 Modelo de regresión lineal simple

La relación matemática determinística más simple entre dos variables x y y es una relación lineal $y = \beta_0 + \beta_1 x$. El conjunto de pares (x, y) para los cuales $y = \beta_0 + \beta_1 x$ determina una recta con pendiente β_1 y ordenada al origen β_0.* El objetivo de esta sección es desarrollar un modelo probabilístico lineal.

Si las dos variables no se relacionan de modo determinístico, entonces para un valor fijo de x, el valor de la segunda variable es aleatorio. Por ejemplo, si se está investigando la relación entre la edad de un niño y el tamaño del vocabulario, y se decide seleccionar un niño de edad $x = 5.0$ años, entonces antes de hacer la selección, el tamaño del vocabulario es una variable aleatoria Y. Después de seleccionar y probar un determinado niño de cinco años, podría resultar un vocabulario de 2000 palabras. Se diría entonces que el valor observado de Y relacionado con fijar $x = 5.0$ fue $y = 2000$.

*La pendiente de una recta es el cambio en y para un incremento de una unidad en x. Por ejemplo, si $y = -3x + 10$, entonces y disminuye en tres cuando x aumenta en uno, de modo que la pendiente es -3. La ordenada al origen es la altura a la cual la recta cruza el eje vertical y se obtiene al hacer $x = 0$ en la ecuación.

En términos generales, la variable cuyo valor lo fija el experimentador se denota con x y se llama **variable independiente, predictora** o **explicativa.** Para x fija, la segunda variable será aleatoria; esta variable aleatoria y su valor observado, se denotan mediante Y y y, respectivamente, y se hace referencia a ella como la **variable dependiente** o de **respuesta.**

Por lo común, las observaciones se hacen para diversos valores de la variable independiente. Sean $x_1, x_2, \ldots, x_n$ los valores de la variable independiente para los que se hacen las observaciones, y sean Y_i y y_i la variable aleatoria y el valor observado en relación con x_i, respectivamente. Los datos bivariados disponibles consisten entonces en los n pares $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Un primer paso en el análisis de regresión en el que intervienen dos variables es construir una **gráfica de dispersión** de los datos observados. En tal gráfica, cada (x_i, y_i) se representa como un punto dibujado en un sistema coordenado bidimensional.

Ejemplo 12.1 Los problemas visuales y musculoesqueléticos relacionados con el uso de terminales de visualización (VDT) se han vuelto bastante comunes en años recientes. Algunos investigadores se han interesado en la dirección vertical como causa de vista cansada e irritación ocular. Se sabe que esta dirección tiene estrecha relación con el área de superficie ocular (ASO), de modo que es necesario un método para medir el ASO. Los siguientes datos representativos acerca de y = ASO (cm^2) y x = amplitud de la fisura palpebral (es decir, la amplitud horizontal de la apertura del ojo, en cm) se tomaron del artículo "Analysis of Ocular Surface Area for Comfortable VDT Workstation Layout" (*Ergonomics*, 1996: 877-884). No se da el orden en el que se obtuvieron las observaciones, así que por conveniencia se listan en orden creciente de valores de x.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	0.40	0.42	0.48	0.51	0.57	0.60	0.70	0.75	0.75	0.78	0.84	0.95	0.99	1.03	1.12
y_i	1.02	1.21	.88	.98	1.52	1.83	1.50	1.80	1.74	1.63	2.00	2.80	2.48	2.47	3.05

i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
x_i	1.15	1.20	1.25	1.25	1.28	1.30	1.34	1.37	1.40	1.43	1.46	1.49	1.55	1.58	1.60
y_i	3.18	3.76	3.68	3.82	3.21	4.27	3.12	3.99	3.75	4.10	4.18	3.77	4.34	4.21	4.92

Así $(x_1, y_1) = (0.40, 1.02)$, $(x_5, y_5) = (0.57, 1.52)$, etcétera. Una gráfica de dispersión de MINITAB se ilustra en la figura 12.1; se empleó una opción que produjo un diagrama de puntos de los valores de x y de y a lo largo de los márgenes derecho y superior de la gráfica, que facilita visualizar las distribuciones de cada una de las variables (los histogramas o diagramas de puntos son opciones alternativas). A continuación se dan algunos aspectos a observar acerca de los datos y la gráfica:

- Varias observaciones tienen valores de x idénticos pero diferentes valores de y (p. ej., $x_8 = x_9 = 0.75$, pero $y_8 = 1.80$ y $y_9 = 1.74$). Así que el valor de y *no* se determina sólo mediante x, sino también mediante otros factores.
- Hay una fuerte tendencia al incremento de y cuando aumenta x. Es decir, los valores más grandes de ASO tienden a relacionarse con valores más grandes de amplitud de fisura, una relación positiva entre las variables.

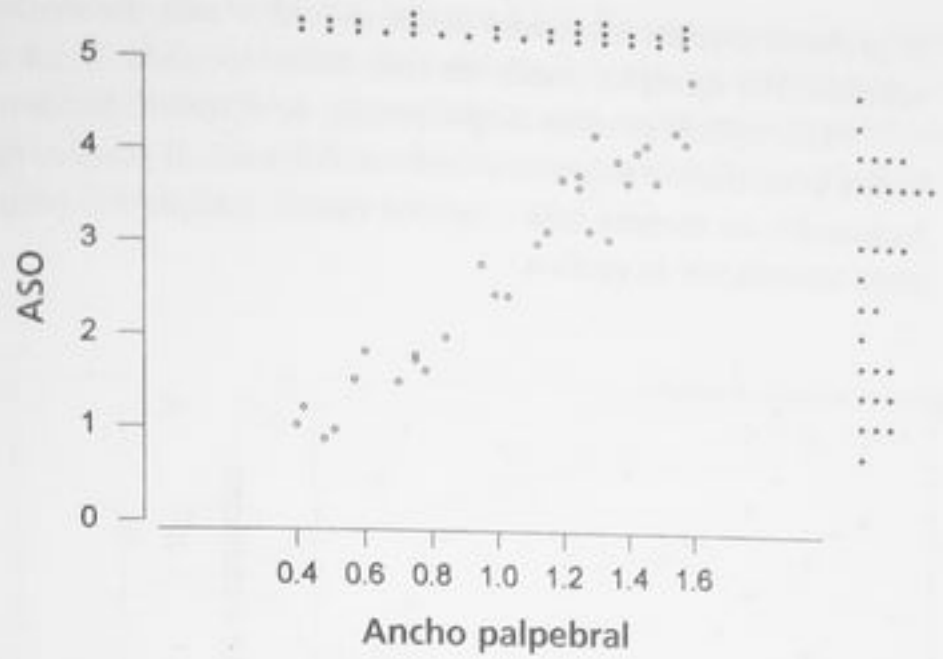


Figura 12.1 Gráfica de dispersión de MINITAB para los datos del ejemplo 12.1, junto con los diagramas de puntos de los valores de x y y

- Al parecer, el valor de y se podría predecir de x al hallar una recta que esté razonablemente de los puntos de la gráfica (los autores del citado artículo superpusieron una recta en su gráfica). En otras palabras, hay evidencia de una relación lineal sustancial (aunque no perfecta) entre las dos variables. ■

Los ejes horizontal y vertical del diagrama de dispersión de la figura 12.1 se intersecan en el punto (0, 0). En muchos conjuntos de datos, los valores de x o y, o los valores de ambas variables, difieren en forma considerable de cero en relación con el intervalo o intervalos de los valores. Por ejemplo, un estudio de cómo se relaciona la eficiencia del acondicionador de aire con la temperatura máxima exterior diaria, podría relacionarse con observaciones para temperaturas que van de 80°F a 100°F. Cuando éste es el caso, una gráfica más importante mostraría los ejes marcados de manera apropiada que se cruzan el algún punto distinto de (0, 0).

Ejemplo 12.2 El crecimiento del bosque y los fenómenos de deterioro de todo el mundo, han despertado interés público y científico considerable. En el artículo "Relationships Among Crown Condition, Growth, and Stand Nutrition in Seven Northern Vermont Sugarbushes" (*Canad. J. of Forest Res.*, 1995: 386-397) incluyeron una gráfica de dispersión de y = deterioro promedio de la corona (%), un indicador de retardo del desarrollo, y x = pH del suelo (el pH mayor corresponde a suelo más ácido), de lo cual se tomaron las observaciones siguientes:

x	3.3	3.4	3.4	3.5	3.6	3.6	3.7	3.7	3.8	3.8
y	7.3	10.8	13.1	10.4	5.8	9.3	12.4	14.9	11.2	8.0
x	3.9	4.0	4.1	4.2	4.3	4.4	4.5	5.0	5.1	
y	6.6	10.0	9.2	12.4	2.3	4.3	3.0	1.6	1.0	

En la figura 12.2 de la página siguiente, se muestran dos gráficas de dispersión de MINITAB de estos datos. En la figura 12.2(a), MINITAB seleccionó la escala para ambos ejes. Se obtuvo la figura 12.2(b) al especificar los valores mínimo y máximo para x y y de modo que los ejes se intersecarían más o menos en el punto (0, 0). La segunda gráfica está más hacinada que

la primera; este hacinamiento puede dificultar más discernir la naturaleza general de cualquier relación. Por ejemplo, puede ser más difícil localizar la curvatura en una gráfica hacinada.

Los valores grandes de porcentaje de deterioro tienden a relacionarse con un pH bajo del suelo, una relación negativa o inversa. Además, al parecer las dos variables están al menos relacionadas de manera más o menos lineal, aunque los puntos se dispersarían en torno a una recta trazada por la gráfica.

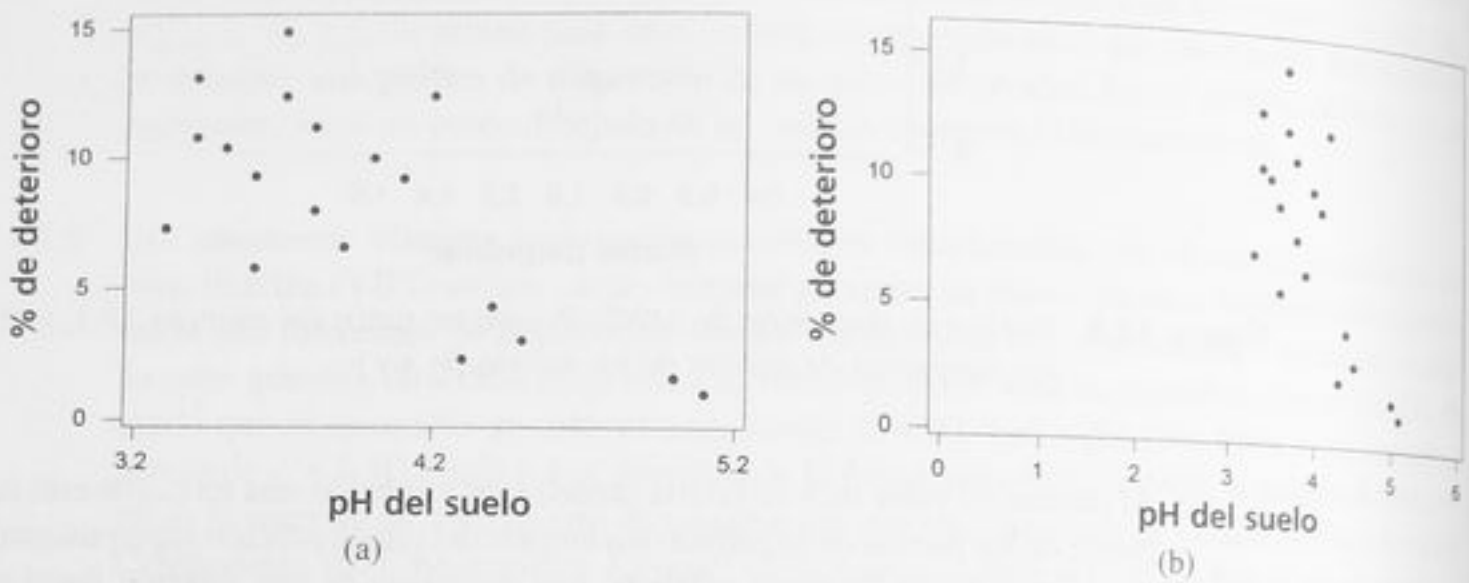


Figura 12.2 Gráficas de dispersión de MINITAB de los datos del ejemplo 12.2 ■

Modelo probabilístico lineal

Para el modelo determinístico $y = \beta_0 + \beta_1 x$, el valor observado real de y es una función lineal de x. La generalización apropiada de esto a un modelo probabilístico, da por hecho que el *valor esperado de Y es una función lineal de x*, pero que para x fija, la variable Y difiere de su valor esperado por una cantidad aleatoria.

Modelo de regresión lineal simple

Hay parámetros β_0, β_1 y σ^2 tal que para cualquier valor fijo de la variable independiente x, la variable dependiente se relaciona con x por medio de la **ecuación modelo**

$$Y = \beta_0 + \beta_1 x + \epsilon \qquad (12.1)$$

La cantidad ϵ en la ecuación modelo es una variable aleatoria, que se supone tiene una distribución normal con $E(\epsilon) = 0$ y $V(\epsilon) = \sigma^2$.

La variable ϵ se conoce como la **desviación aleatoria** o **término de error aleatorio** en el modelo. Sin ϵ, cualquier par observado (x, y) correspondería a un punto que cae exactamente en la recta $y = \beta_0 + \beta_1 x$, llamada **recta de regresión verdadera** (o **poblacional**). La inclusión de un término de error aleatorio permite que (x, y) caiga ya sea arriba de la recta de regresión verdadera (cuando $\epsilon > 0$) o abajo de la recta (cuando $\epsilon < 0$). Los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ que resultan de n observaciones independientes se dispersarán respecto a la

recta de regresión, como se ilustra en la figura 12.3. En ocasiones, las consideraciones teóricas podrían hacer pensar en la pertinencia del modelo de regresión lineal simple (p. ej., hay una relación lineal exacta entre las dos variables, con ϵ que representa el error de medición). Con mucha más frecuencia, una gráfica de dispersión que exhibe un patrón lineal sustancial, indica lo razonable del modelo (como en la figura 12.1).

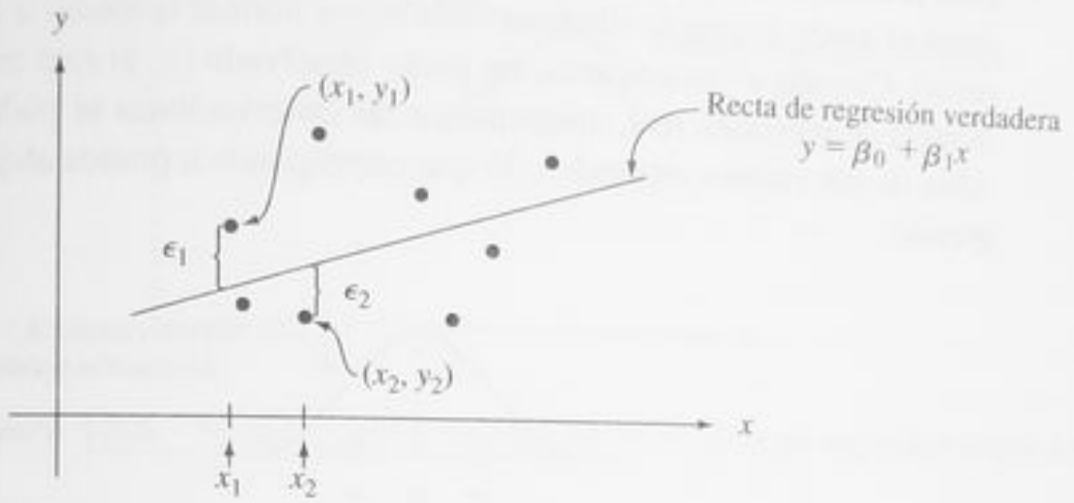


Figura 12.3 Puntos que corresponden a observaciones del modelo de regresión lineal simple

Las implicaciones de la ecuación modelo (12.1) se entienden mejor con la ayuda de la siguiente notación. Sea x^* un valor particular de la variable independiente x y

$$\mu_{Y \cdot x^*} = \text{valor esperado (o medio) de } Y \text{ cuando } x = x^*$$

$$\sigma^2_{Y \cdot x^*} = \text{varianza de } Y \text{ cuando } x = x^*$$

La notación alternativa es $E(Y \mid x^*)$ y $V(Y \mid x^*)$. Por ejemplo, si x = esfuerzo aplicado (kg/mm^2) y y = tiempo hasta la fractura (h), entonces $\mu_{Y \cdot 20}$ denotaría el valor esperado del tiempo hasta la fractura cuando el esfuerzo aplicado es 20 kg/mm^2. Si se considera una población completa de pares (x, y), entonces $\mu_{Y \cdot x}$ es la media de los valores y para los cuales $x = x^*$, y $\sigma^2_{Y \cdot x}$ es una medida de cuántos de estos valores de y se dispersan respecto a la media. Si, por ejemplo, x = edad de un niño y y = tamaño del vocabulario, entonces $\mu_{Y \cdot 5}$ es el tamaño de vocabulario promedio para los niños de cinco años en la población, y $\sigma^2_{Y \cdot 5}$ describe la cantidad de variabilidad en el tamaño de vocabulario para esta parte de la población. Una vez que x está fija, la única aleatoriedad en el lado derecho de la ecuación modelo (12.1) es el error aleatorio ϵ, y su valor medio y varianza son 0 y σ^2, respectivamente, cualquiera que sea el valor de x. Esto implica que

$$\mu_{Y \cdot x^*} = E(\beta_0 + \beta_1 x^* + \epsilon) = \beta_0 + \beta_1 x^* + E(\epsilon) = \beta_0 + \beta_1 x^*$$

$$\sigma^2_{Y \cdot x^*} = V(\beta_0 + \beta_1 x^* + \epsilon) = V(\beta_0 + \beta_1 x^*) + V(\epsilon) = 0 + \sigma^2 = \sigma^2$$

Al sustituir x^* en $\mu_{Y \cdot x^*}$ por x, se obtiene la relación $\mu_{Y \cdot x} = \beta_0 + \beta_1 x$, la cual establece que el *valor medio* de Y, y no Y, es una función lineal de x. Por lo tanto, la recta de regresión real $y = \beta_0 + \beta_1 x$ es la *línea de valores medios*; su altura arriba de cualquier valor x particular, es el valor esperado de Y para ese valor de x. La pendiente β_1 de la recta de regresión verdadera se interpreta como el cambio *esperado* en Y relacionado con un incremento de una unidad en el valor de x. La segunda relación expresa que la cantidad de variabilidad en la distribución de valores de Y es la misma en cada valor diferente de x (homogeneidad de varianza). En el ejem-

plo relacionado con la edad de un niño y el tamaño de su vocabulario, el modelo implica que el tamaño del vocabulario promedio cambia linealmente con la edad (es de esperar que β_1 sea positiva) y que la cantidad de variabilidad en el tamaño del vocabulario a una determinada edad sea la misma a cualquier otra edad. Por último, para x fija, Y es la suma de una constante $\beta_0 + \beta_1 x$ y una variable aleatoria ϵ con distribución normal, así que por sí misma tiene una distribución normal. Estas propiedades se ilustran en la figura 12.4. El parámetro de varianza σ^2 determina el grado al que se dispersa cada curva normal respecto a su valor medio (la altura de la recta). Cuando σ^2 es pequeña, un punto observado (x, y) casi siempre caerá bastante cerca de la recta de regresión real, mientras que las observaciones se podrían desviar en forma considerable de sus valores esperados (lo que corresponde a puntos alejados de la recta) cuando σ^2 es grande.

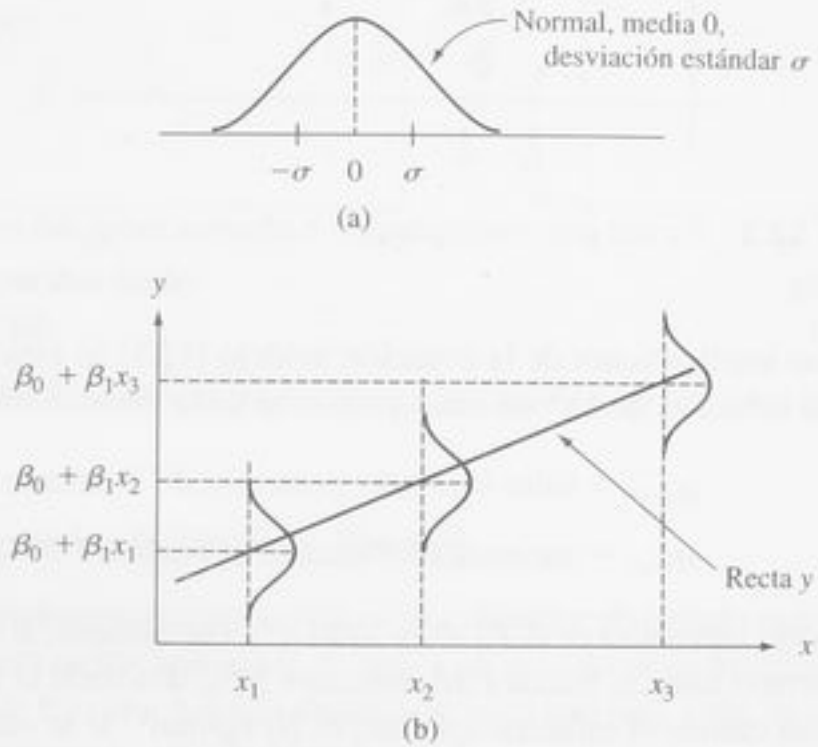


Figura 12.4 a) Distribución de ϵ; b) distribución de Y para diferentes valores x

Ejemplo 12.3 Suponga que la relación entre el esfuerzo aplicado x y el tiempo hasta la falla y se describe mediante el modelo de regresión lineal simple con recta de regresión verdadera $y = 65 - 1.2x$ y $\sigma = 8$. Entonces para cualquier valor fijo x^* del esfuerzo, el tiempo hasta la falla tiene una distribución normal con valor medio $65 - 1.2x^*$ y desviación estándar 8. En términos generales, en la población que consiste en los puntos (x, y), la magnitud de una desviación típica a partir de la recta de regresión verdadera es alrededor de 8. Para $x = 20$, Y tiene el valor medio $\mu_{Y \cdot 20} = 65 - 1.2(20) = 41$, así que

$$P(Y > 50 \text{ cuando } x = 20) = P\left(Z > \frac{50 - 41}{8}\right) = 1 - \Phi(1.13) = 0.1292$$

La probabilidad de que el tiempo hasta la falla pase de 50 cuando el esfuerzo aplicado es 25 es, debido a que $\mu_{Y \cdot 25} = 35$,

$$P(Y > 50 \text{ cuando } x = 25) = P\left(Z > \frac{50 - 35}{8}\right) = 1 - \Phi(1.88) = 0.0301$$

Estas probabilidades se ilustran como las áreas sombreadas en la figura 12.5.

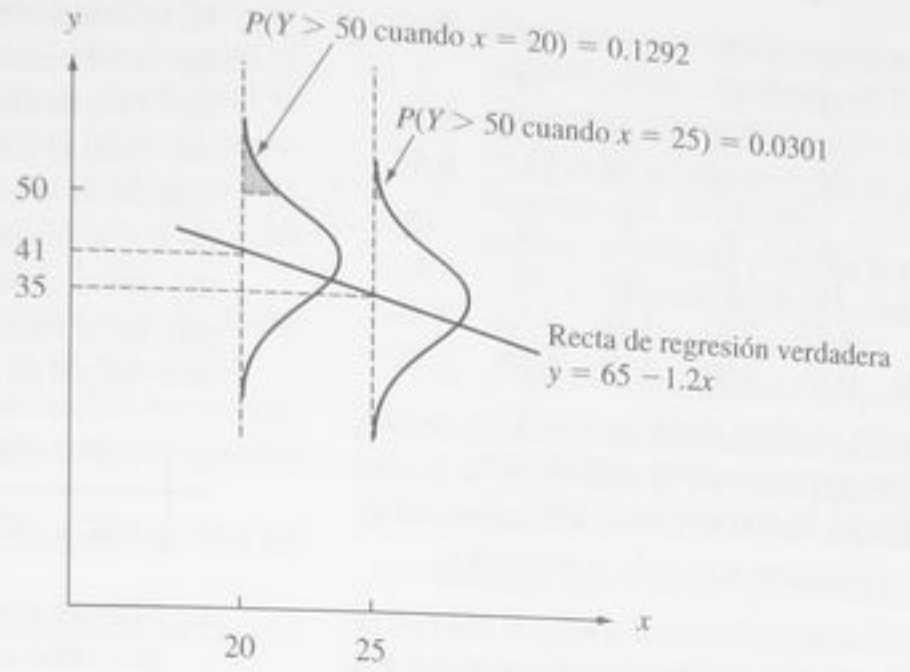


Figura 12.5 Probabilidades basadas en el modelo de regresión lineal simple

Suponga que Y_1 denota una observación en el tiempo hasta la falla hecho con $x = 25$ y Y_2 denota una observación independiente hecha con $x = 24$. Entonces $Y_1 - Y_2$ tiene una distribución normal con valor medio $E(Y_1 - Y_2) = \beta_1 = -1.2$, varianza $V(Y_1 - Y_2) = \sigma^2 + \sigma^2 = 128$ y desviación estándar $\sqrt{128} = 11.314$. La probabilidad de que Y_1 sea mayor que Y_2 es

$$P(Y_1 - Y_2 > 0) = P\left(Z > \frac{0 - (-1.2)}{11.314}\right) = P(Z > 0.11) = 0.4562$$

Es decir, aun cuando se espera que Y disminuya cuando x aumenta en una unidad, no es improbable que el valor observado Y en $x + 1$ sea mayor que la Y observada en x. ■

Ejercicios | Sección 12.1 (1 a 11)

1. La relación de eficiencia para un espécimen de acero sumergido en un tanque fosfatizado es el peso del recubrimiento de fosfato dividido entre la pérdida de metal (ambos en mg/pie²). En el artículo "Statistical Process Control of a Phosphate Coating Line" (*Wire J. Intl.*, mayo de 1997: 78-81) aparecen los siguientes datos sobre la temperatura del tanque (x) y relación de eficiencia (y).

Temp.	170	172	173	174	174	175	176
Relación	0.84	1.31	1.42	1.03	1.07	1.08	1.04
Temp.	177	180	180	180	180	180	181
Relación	1.80	1.45	1.60	1.61	2.13	2.15	0.84
Temp.	181	182	182	182	182	184	184
Relación	1.43	0.90	1.81	1.94	2.68	1.49	2.52
Temp.	185	186	188				
Relación	3.00	1.87	3.08				

   a. Trace diagramas de tallo y hojas de la temperatura y la relación de eficiencia, y comente acerca de características interesantes.
   b. ¿El valor de la relación de eficiencia está determinado por completo y de manera única por la temperatura del tanque? Explique su razonamiento.
   c. Construya un diagrama de dispersión de los datos. ¿Al parecer la relación de eficiencia se podría predecir por el valor de la temperatura? Explique su razonamiento.

2. En el artículo "Exhaust Emissions from Four-Stroke Lawn Mower Engines" (*J. of the Air and Water Mgmnt. Assoc.*, 1997: 945-952) aparecen los datos de un estudio donde se usaron una mezcla base de gasolina y una gasolina reformulada. Considere las siguientes observaciones acerca de la edad (años) y emisiones de NO, (g/kWh):

Motor	1	2	3	4	5
Edad	0	0	2	11	7
Base	1.72	4.38	4.06	1.26	5.31
Reformulada	1.88	5.93	5.54	2.67	6.53
Motor	6	7	8	9	10
Edad	16	9	0	12	4
Base	0.57	3.37	3.44	0.74	1.24
Reformulada	0.74	4.94	4.89	0.69	1.42

Construya gráficas de dispersión de emisiones de NO_x contra edad. ¿Cuál parece ser la naturaleza de la relación entre estas dos variables? (*Nota:* los autores del artículo citado comentaron acerca de la relación.)

3. Los datos bivariados con frecuencia surgen del uso de dos técnicas distintas para medir la misma cantidad. Por ejemplo, las observaciones siguientes en x = concentración de hidrógeno (ppm) con un método de cromatografía de gases y y = concentración con un nuevo método de sensor, se leyeron de una gráfica del artículo "A New Method to Measure the Diffusible Hydrogen Content in Steel Weldments Using a Polymer Electrolyte-Based Hydrogen Sensor" (*Welding Res.*, julio de 1997: 251s-256s).

x	47	62	65	70	70	78	95	100	114	118
y	38	62	53	67	84	79	93	106	117	116
x	124	127	140	140	140	150	152	164	198	221
y	127	114	134	139	142	170	149	154	200	215

Trace una gráfica de dispersión. ¿Al parecer hay una relación muy fuerte entre los dos tipos de mediciones de concentración? ¿Los dos métodos al parecer miden de forma aproximada la misma cantidad? Explique su razonamiento.

4. Un estudio realizado para evaluar la capacidad de sistemas pantanosos de flujo subterráneo para eliminar la demanda bioquímica de oxígeno (DBO) y otros componentes químicos, dio como resultado los datos siguientes en relación con x = carga másica de DBO (kg/ha/d) y y = eliminación de masa de DBO (kg/ha/d) ("Subsurface Flow Wetlands: A Performance Evaluation", *Water Envir: Res.*, 1995: 244-247).

x	3	8	10	11	13	16	27	30	35	37	38	44	103	142
y	4	7	8	8	10	11	16	26	21	9	31	30	75	90

a. Trace diagramas de caja para la carga másica y la eliminación de masa, y comente acerca de algunas características interesantes.

b. Construya un diagrama de dispersión de los datos y realice comentarios en relación con algunas características interesantes.

5. En el artículo "Objective Measurement of the Stretchability of Mozzarella Cheese" (*J. of Texture Studies*, 1992: 185-194) se describe un experimento para investigar cómo varía el comportamiento del queso mozzarella en función de la temperatura. Considere los siguientes datos de x = temperatura y y = alargamiento (%) a la ruptura del queso. (*Nota:* los investigadores eran italianos y usaron queso mozzarella *genuino*, no la mala imitación que se consigue en Estados Unidos.)

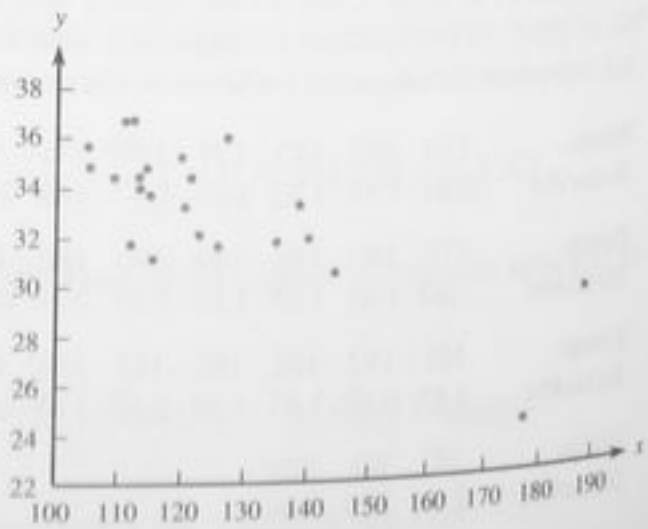
x	59	63	68	72	74	78	83
y	118	182	247	208	197	135	132

a. Elabore una gráfica de dispersión en la que los ejes se crucen en (0, 0). Marque 0, 20, 40, 60, 80 y 100 en el eje horizontal y 0, 50, 100, 150, 200 y 250 en el eje vertical.

b. Trace una gráfica de dispersión en la que los ejes se crucen en (55, 100), como se hizo en el artículo citado. ¿Esta gráfica parece preferible a la del inciso (a)? Explique por qué.

c. ¿Qué indican las gráficas de los incisos (a) y (b) acerca de la naturaleza de la relación entre las dos variables?

6. Un factor en el desarrollo del codo de tenista, un padecimiento que atemoriza a los jugadores de tenis, es la vibración inducida por impacto del sistema raqueta-brazo al contacto con la bola. Es bien conocido que la posibilidad de padecer la enfermedad depende de varias propiedades de la raqueta utilizada. Considere el diagrama de dispersión de x = frecuencia de resonancia de la raqueta (Hz) y y = suma de aceleración pico a pico (una característica de la vibración del brazo, en *m/s/s)* para $n = 23$ raquetas distintas ("Transfer of Tennis Racket Vibrations into the Human Forearm", *Medicine and Science in Sports and Exercise*, 1992: 1134-1140). Describa las características interesantes de los datos y la gráfica de dispersión.



7. En el artículo "Some Field Experience in the Use of an Accelerated Method in Estimating 28-Day Strength of Concrete" (*J. Amer. Concrete Institute*, 1969: 895) se consideró la regresión de la resistencia estándar de curado y = 28 días (lb/pulg^2) contra x = resistencia acelerada (lb/pulg^2). Suponga que la ecuación de la recta de regresión verdadera es $y = 1800 + 1.3x$.
   a. ¿Cuál es el valor esperado de la resistencia de 28 días cuando la resistencia acelerada es igual a 2500?
   b. ¿En cuánto se puede esperar que cambie la resistencia de 28 días esperada cuando la resistencia acelerada aumenta en 1 lb/pulg^2?
   c. Conteste el inciso (b) para un aumento de 100 lb/pulg^2.
   d. Conteste el inciso (b) para una disminución de 100 lb/pulg^2.

8. Con referencia al ejercicio 7, suponga que la desviación estándar de la desviación aleatoria ϵ es 350 lb/pulg^2.
   a. ¿Cuál es la probabilidad de que el valor observado de la resistencia de 28 días pase de 5000 lb/pulg^2 cuando el valor de la resistencia acelerada es 2000?
   b. Repita el inciso (a) con 2500 en lugar de 2000.
   c. Considere realizar dos observaciones independientes en la resistencia de 28 días, la primera para una resistencia acelerada de 2000 y la segunda para x = 2500. ¿Cuál es la probabilidad de que la segunda observación exceda a la primera en más de 1000 lb/pulg^2?
   d. Sean Y_1 y Y_2 las observaciones en la resistencia de 28 días cuando $x = x_1$ y $x = x_2$, respectivamente. ¿En cuánto tendría que exceder x_2 a x_1 a fin de que $P(Y_2 > Y_1) = 0.95$?

9. El flujo y (m^3/min) en un dispositivo empleado para medir la calidad del aire depende de la caída de presión x (pulg de agua) en el filtro del dispositivo. Suponga que para valores de x entre 5 y 20, las dos variables se relacionan de acuerdo con el modelo de regresión lineal simple con la recta de regresión verdadera $y = -0.12 + 0.095x$.
   a. ¿Cuál es el cambio esperado en el flujo relacionado con un aumento de una pulgada en la caída de presión? Explique.
   b. ¿Qué cambio se puede esperar en el flujo cuando la caída de presión disminuye en 5 pulgadas?
   c. ¿Cuál es el flujo esperado para una caída de presión de 10 pulg? ¿Para una caída de 15 pulgadas?
   d. Suponga que $\sigma = 0.025$ y considere una caída de presión de 10 pulg. ¿Cuál es la probabilidad de que el valor observado del flujo exceda 0.835? ¿De que el flujo observado pase de 0.840?
   e. ¿Cuál es la probabilidad de que una observación en el flujo, cuando la caída de presión sea de 10 pulg, exceda una observación en el flujo realizada cuando la caída de presión sea de 11 pulgadas?

10. Suponga que el costo esperado de un corrida de producción se relaciona con el tamaño de la corrida mediante la ecuación $y = 4000 + 10x$. Sea Y una observación en el costo de una corrida. Si las variables *tamaño* y *costo* están relacionadas según el modelo de regresión lineal simple, ¿podría suceder que $P(Y > 5500$ cuando $x = 100) = 0.05$ y $P(Y > 6500$ cuando $x = 200) = 0.10$? Explique.

11. Suponga que en un cierto proceso químico el tiempo de reacción y (h) se relaciona con la temperatura (°F) de la cámara donde tiene lugar la reacción según el modelo de regresión lineal simple con ecuación $y = 5.00 - 0.01x$ y $\sigma = 0.075$.
   a. ¿Cuál es el cambio esperado en el tiempo de reacción para un aumento de 1 °F en la temperatura? ¿Para un aumento de 10 °F?
   b. ¿Cuál es el tiempo de reacción esperado cuando la temperatura es de 200 °F. ¿Cuando sea 250 °F?
   c. Suponga que se hacen cinco observaciones independientemente del tiempo de reacción, cada una para una temperatura de 250 °F. ¿Cuál es la probabilidad de que los cinco tiempos estén entre 2.4 y 2.6 horas?
   d. ¿Cuál es la probabilidad de que dos tiempos de reacción observados de forma independiente para temperaturas separadas un grado sean tales que el tiempo a la mayor temperatura sea mayor que el tiempo a la temperatura menor?

12.2 Estimación de los parámetros del modelo

En esta sección y varias de las siguientes, se supondrá que las variables x y y se relacionan según un modelo de regresión lineal simple. El investigador casi nunca conocerá los valores de β_0, β_1 y σ^2. En cambio, estarán disponibles los datos muestrales que consisten en n pares observados $(x_1, y_1), \ldots, (x_n, y_n)$, a partir de los cuales se puede estimar los parámetros del modelo y la recta de regresión verdadera. Se considera que estas observaciones se obtuvieron de manera in-

dependiente entre sí. Es decir, y_i es el valor observado de una *va* Y_i, donde $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ y las n desviaciones $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ son *va* independientes. La independencia de $Y_1, Y_2, \ldots, Y_n$ se deduce de la independencia de las ϵ_i.

De acuerdo con el modelo, los puntos observados se distribuirán respecto a la recta de regresión verdadera de una manera aleatoria. En la figura 12.6 se muestra una gráfica característica de los pares observados junto con dos candidatos para la recta de regresión estimada, $y = a_0 + a_1 x$ y $y = b_0 + b_1 x$. De manera intuitiva, la recta $y = a_0 + a_1 x$ no es una estimación razonable de la recta verdadera $y = \beta_0 + \beta_1 x$ porque, si $y = a_0 + a_1 x$ fuera la recta verdadera, es casi seguro que los puntos observados habrían estado más cerca de esta recta. La recta $y = b_0 + b_1 x$ es una estimación más plausible porque los puntos observados se dispersan más cerca respecto a esta recta.

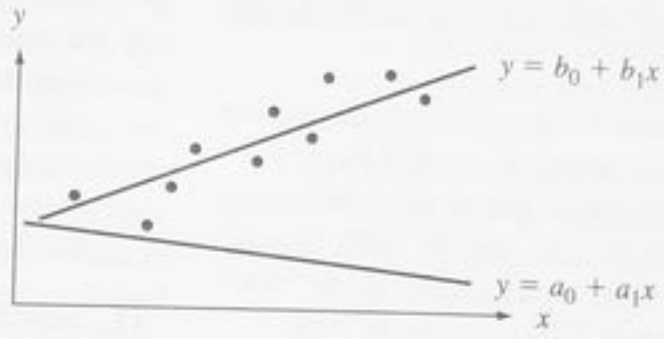


Figura 12.6 Dos estimaciones diferentes de la recta de regresión verdadera

Según la figura 12.6 y la explicación anterior, la estimación de $y = \beta_0 + \beta_1 x$ debe ser una recta que en cierto sentido proporciona un mejor ajuste a los puntos de datos observados. Lo anterior es lo que motiva el principio de los mínimos cuadrados, que se remonta al matemático alemán Gauss (1777-1855). De acuerdo con este principio, una recta proporciona un buen ajuste para los datos si las distancias verticales (desviaciones) de los puntos observados a la recta son pequeños (véase la figura 12.7). La medida de la bondad del ajuste es la suma de los cuadrados de estas desviaciones. Entonces, la recta del mejor ajuste es la que tiene la menor suma posible de desviaciones al cuadrado.

Principio de mínimos cuadrados

La desviación vertical del punto (x_i, y_i) de la recta $y = b_0 + b_1 x$ es

$$\text{altura del punto} - \text{altura de la recta} = y_i - (b_0 + b_1 x_i)$$

Entonces, la suma de desviaciones verticales al cuadrado de los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ a la recta es

$$f(b_0, b_1) = \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_i)]^2$$

Las estimaciones puntuales de β_0 y β_1, denotadas por $\hat{\beta}_0$ y $\hat{\beta}_1$ y conocidas como **estimaciones de mínimos cuadrados**, son los valores que minimizan $f(b_0, b_1)$. Es decir, $\hat{\beta}_0$ y $\hat{\beta}_1$ son tales que $f(\hat{\beta}_0, \hat{\beta}_1) \leq f(b_0, b_1)$ para cualquier b_0 y b_1. La **recta de regresión estimada** o **recta de mínimos cuadrados** es la recta cuya ecuación es $y = \hat{\beta}_0 + \hat{\beta}_1 x$.

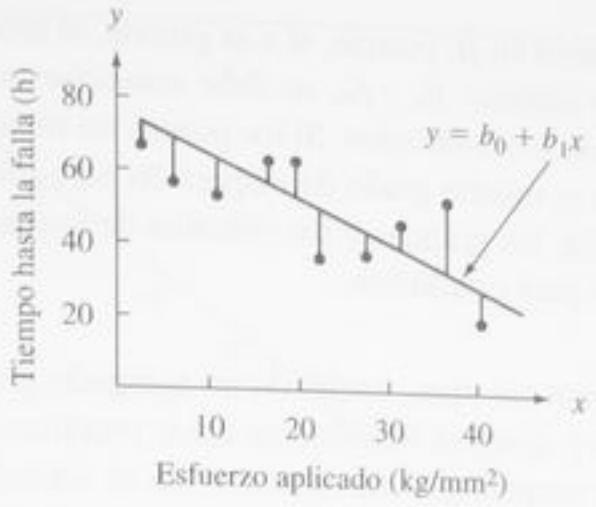


Figura 12.7 Desviaciones de los datos observados respecto a la recta $y = b_0 + b_1x$

Los valores óptimos de b_0 y b_1 se encuentran al tomar las derivadas parciales de $f(b_0, b_1)$ con respecto a b_0 y b_1 igualarlas a cero [de manera análoga a $f'(b) = 0$ en el cálculo de una variable], y resolver las ecuaciones

$$\frac{\partial f(b_0, b_1)}{\partial b_0} = \sum 2(y_i - b_0 - b_1x_i)(-1) = 0$$

$$\frac{\partial f(b_0, b_1)}{\partial b_1} = \sum 2(y_i - b_0 - b_1x_i)(-x_i) = 0$$

La cancelación del factor −2 y la nueva disposición produce el siguiente sistema de ecuaciones, conocidas como **ecuaciones normales:**

$$nb_0 + (\sum x_i)b_1 = \sum y_i$$
$$(\sum x_i)b_0 + (\sum x_i^2)b_1 = \sum x_iy_i$$

Las ecuaciones normales son lineales en las dos incógnitas b_0 y b_1. Siempre y cuando por lo menos dos de las x_i sean diferentes, las estimaciones de mínimos cuadrados son la única solución para este sistema.

La estimación de mínimos cuadrados del coeficiente β_1 (la pendiente) de la recta de regresión es

$$b_1 = \hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \qquad (12.2)$$

Las fórmulas de cálculo para el numerador y el denominador de $\hat{\beta}_1$ son

$$S_{xy} = \sum x_iy_i - (\sum x_i)(\sum y_i)/n \qquad S_{xx} = \sum x_i^2 - (\sum x_i)^2/n$$

La estimación de mínimos cuadrados de la ordenada al origen β_0 de la recta de regresión verdadera es

$$b_0 = \hat{\beta}_0 = \frac{\sum y_i - \hat{\beta}_1 \sum x_i}{n} = \bar{y} - \hat{\beta}_1\bar{x} \qquad (12.3)$$

Las fórmulas de cálculo para S_{xy} y S_{xx} requieren sólo los estadísticos de resumen $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_iy_i$ ($\sum y_i^2$ se necesitará en breve) y minimizar los efectos del redondeo. Al calcular $\hat{\beta}_0$, utilice

dígitos extra en $\hat{\beta}_1$ porque, si $\bar{x}$ es grande, el redondeo afecta la respuesta final. Se remarca que *antes de calcular $\hat{\beta}_1$ y $\hat{\beta}_0$, se debe examinar una gráfica de dispersión para ver si es posible un modelo probabilístico.* Si los puntos no tienden a agruparse alrededor de una recta con más o menos el mismo grado de dispersión para toda x, se deben investigar los otros modelos. En la práctica, las gráficas y los cálculos de regresión por lo común se hacen con un paquete de software para estadística.

Ejemplo 12.4 El concreto sin finos, hecho de un agregado grueso clasificado uniformemente y una pasta de cemento y agua, es benéfico en áreas proclives a la lluvia excesiva como resultado de sus excelentes propiedades de drenado. En el artículo "Pavement Thickness Design for No-Fines Concrete Parking Lots" (*J. of Transportation Engr.*, 1995: 476-484) se utilizó un análisis de mínimos cuadrados para estudiar cómo y = porosidad (%) se relaciona con x = peso unitario (lb/pie^3) en especímenes de concreto. Considere los siguientes datos representativos, mostrados en un formato tabular conveniente para calcular los valores de los estadísticos de resumen:

Obs	x	y	x^2	xy	y^2
1	99.0	28.8	9801.00	2851.20	829.44
2	101.1	27.9	10221.21	2820.69	778.41
3	102.7	27.0	10547.29	2772.90	729.00
4	103.0	25.2	10609.00	2595.60	635.04
5	105.4	22.8	11109.16	2403.12	519.84
6	107.0	21.5	11449.00	2300.50	462.25
7	108.7	20.9	11815.69	2271.83	436.81
8	110.8	19.6	12276.64	2171.68	384.16
9	112.1	17.1	12566.41	1916.91	292.41
10	112.4	18.9	12633.76	2124.36	357.21
11	113.6	16.0	12904.96	1817.60	256.00
12	113.8	16.7	12950.44	1900.46	278.89
13	115.1	13.0	13248.01	1496.30	169.00
14	115.4	13.6	13317.16	1569.44	184.96
15	120.0	10.8	14400.00	1296.00	116.64
Suma	1640.1	299.8	179 849.73	32 308.59	6430.06

Así, $\bar{x} = 109.34$, $\bar{y} = 19.986667$, y

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{32\,308.59 - (1640.1)(299.8)/15}{179\,849.73 - (1640.1)^2/15}$$

$$= \frac{-471.542}{521.196} = -0.90473066 \approx -0.905$$

$$\hat{\beta}_0 = 19.986667 - (-0.90473066)(109.34) = 118.909917 \approx 118.91$$

Se estima que el cambio de porosidad relacionado con un incremento de 1 lb/pie^3 en el peso unitario es −0.905% (una disminución de −0.905%). La ecuación de la recta de regresión estimada (recta de mínimos cuadrados) es $y = 118.91 - 0.905x$. La figura 12.8, generada mediante el paquete de software estadístico S-Plus, muestra que la recta de mínimos cuadrados proporciona un resumen excelente de la relación entre las dos variables.

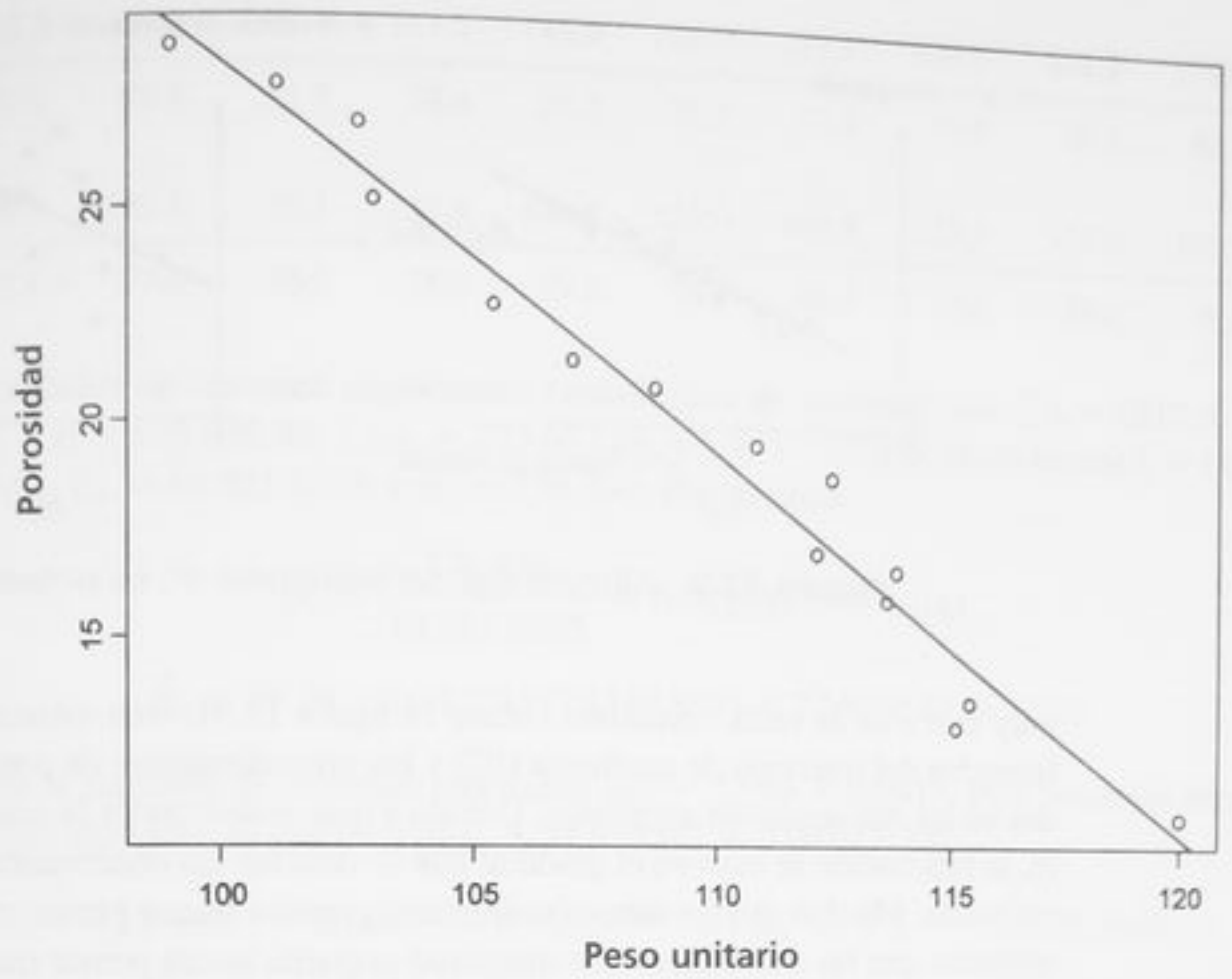


Figura 12.8 Gráfica de dispersión de los datos del ejemplo 12.4 con la recta de mínimos cuadrados superpuesta, de S-plus ■

La recta de regresión estimada se puede usar de inmediato para dos propósitos distintos. Para un valor de x fijo x^*, $\hat{\beta}_0 + \hat{\beta}_1 x^*$ (la altura de la recta arriba de x^*) proporciona ya sea (1) una estimación puntual del valor esperado de Y cuando $x = x^*$ o (2) una predicción puntual del valor de Y que da como resultado una nueva observación hecha en $x = x^*$.

Ejemplo 12.5 Refiérase a los datos de peso-porosidad del ejemplo previo. Una estimación puntual para la porosidad promedio verdadera de los especímenes cuyo peso unitario es 110 es

$$\hat{\mu}_{Y \cdot 110} = \hat{\beta}_0 + \hat{\beta}_1(110) = 118.91 - 0.905(110) = 19.4\%$$

Si se va a seleccionar un espécimen cuyo peso unitario es 110 lb/pie^3, 19.4% también es una predicción puntual para la porosidad de este espécimen. ■

La recta de mínimos cuadrados no se debe usar para hacer una predicción de un valor x mucho más allá del alcance de los datos, por ejemplo $x = 90$ o $x = 135$ en el ejemplo 12.4. El **riesgo de la extrapolación** es que la relación ajustada (en este caso una recta) podría no ser válida para tales valores. (En el ejemplo anterior, $x = 135$ da $\hat{y} = -3.3$, un valor de porosidad a todas luces ilógico, pero la extrapolación no siempre da como resultado tales inconsistencias.)

Estimación de σ^2 y σ

El parámetro σ^2 determina la cantidad de variabilidad inherente en el modelo de regresión. Un valor grande de σ^2 da lugar a que las (x_i, y_i) observadas que están bastante dispersas respecto a la recta de regresión, mientras que cuando σ^2 es pequeña, los puntos observados tenderán a caer

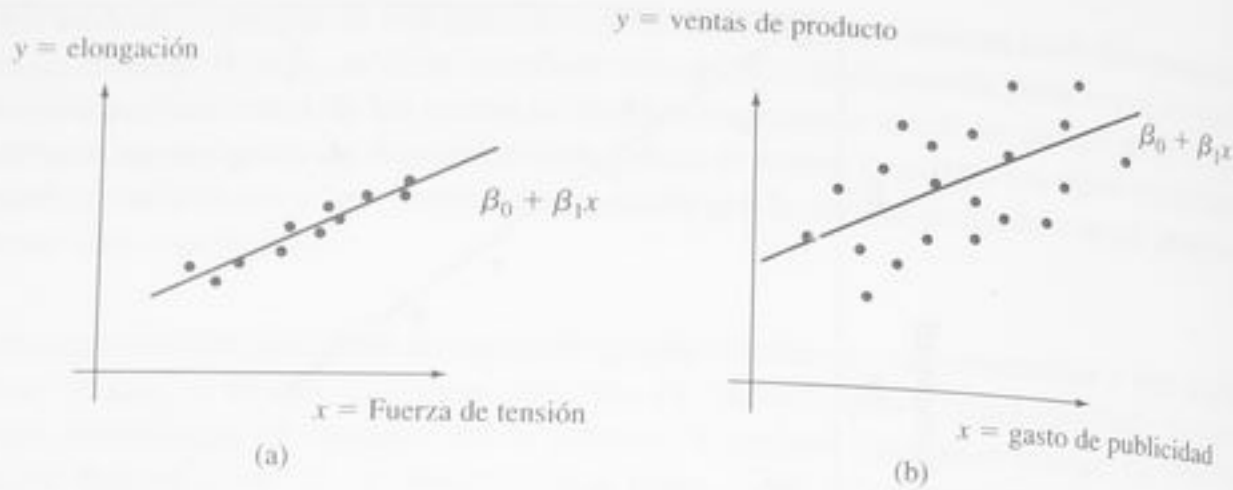


Figura 12.9 Muestra representativa para σ^2: (a) pequeña; (b) grande

muy cerca de la recta verdadera (véase la figura 12.9). Una estimación de σ^2 se usa en las fórmulas del intervalo de confianza (IC) y los procedimientos de prueba de hipótesis presentados en las dos secciones siguientes. Debido a que se desconoce la ecuación de la recta verdadera, la estimación se basa en el grado al que se desvían las observaciones muestrales de la recta estimada. Muchas desviaciones (residuales) grandes hacen pensar en un valor grande de σ^2, mientras que las desviaciones de magnitud pequeña hacen pensar que σ^2 es pequeña.

DEFINICIÓN

Los **valores ajustados** (o **predichos**) $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ se obtienen al sustituir en forma sucesiva $x_1, \ldots, x_n$ en la ecuación de la recta de regresión estimada: $\hat{y}_1 = \hat{\beta}_0 + \hat{\beta}_1 x_1$, $\hat{y}_2 = \hat{\beta}_0 + \hat{\beta}_1 x_2, \ldots, \hat{y}_n = \hat{\beta}_0 + \hat{\beta}_1 x_n$. Los **residuos** son las desviaciones verticales $y_1 - \hat{y}_1$, $y_2 - \hat{y}_2, \ldots, y_n - \hat{y}_n$ respecto a la recta estimada.

En palabras, el valor predicho $\hat{y}_i$ es el valor de y que se predeciría o esperaría al usar la recta de regresión estimada con $x = x_i$; $\hat{y}_i$ es la altura de la recta de regresión estimada arriba del valor x_i para el que se hizo la i-ésima observación. El residuo $y_i - \hat{y}_i$ es la diferencia entre la y_i observada y el valor predicho $\hat{y}_i$. Si los residuos son pequeños, entonces gran parte de la variabilidad en los valores de y observados al parecer se debe a la relación lineal entre x y y, mientras que muchos de los residuos grandes indican bastante variabilidad inherente en y relativa a la cantidad debida a la relación lineal. Suponiendo que la recta de la figura 12.7 es la recta de mínimos cuadrados, los residuos se identifican mediante segmentos de recta verticales de los puntos observados a la recta. Cuando la recta de regresión estimada se obtiene vía el principio de mínimos cuadrados, la suma de los residuos en teoría debe ser cero. En la práctica, la suma se desvía un poco de cero debido al redondeo.

Ejemplo 12.6 La alta densidad poblacional de Japón ha originado gran cantidad de problemas de uso de recursos. Una dificultad especialmente grave tiene que ver con la eliminación de desechos. En el artículo "Innovative Sludge Handling Through Pelletization Thickening" (*Water Research*, 1999: 3245-3252) se da a conocer el desarrollo de una nueva máquina de compresión para procesar aguas residuales. Una parte importante de la investigación requirió relacionar el control de humedad de gránulos comprimidos (y, en %) con la tasa de filtración de la máquina (x, en kg-DS/m/h). Los datos siguientes se leyeron de una gráfica del artículo:

x	125.3	98.2	201.4	147.3	145.9	124.7	112.2	120.2	161.2	178.9
y	77.9	76.8	81.5	79.8	78.2	78.3	77.5	77.0	80.1	80.2
x	159.5	145.8	75.1	151.4	144.2	125.0	198.8	132.5	159.6	110.7
y	79.9	79.0	76.7	78.2	79.5	78.1	81.5	77.0	79.0	78.6

Las cantidades de resumen importantes (*estadísticos de resumen*) son $\sum x_i = 2817.9$, $\sum y_i = 1574.8$, $\sum x_i^2 = 415\,949.85$, $\sum x_i y_i = 222\,657.88$, y $\sum y_i^2 = 124\,039.58$, de lo cual $\bar{x} = 140.895$, $\bar{y} = 78.74$, $S_{xx} = 18\,921.8295$ y $S_{xy} = 776.434$. Por lo tanto,

$$\hat{\beta}_1 = \frac{776.434}{18\,921.8295} = 0.04103377 \approx 0.041$$

$$\hat{\beta}_0 = 78.74 - (0.4103377)(140.895) = 72.958547 \approx 72.96$$

de donde la ecuación de mínimos cuadrados es $\hat{y} = 72.96 + 0.041x$. Para precisión numérica, los valores ajustados se calculan de $\hat{y}_i = 72.958547 + 0.04103377x_i$:

$$\hat{y}_1 = 72.958547 + 0.04103377(125.3) \approx 78.100, \quad y_1 - \hat{y}_1 \approx -0.200, \text{ etcétera.}$$

Un residuo positivo corresponde a un punto en la gráfica de dispersión que se ubica arriba de la recta de mínimos cuadrados, mientras que un residuo negativo resulta de un punto localizado abajo de la recta. Los valores predichos (ajustes) y los residuos aparecen en la tabla siguiente.

Obs	Filtrado	Cont. de humedad	Ajuste	Residuo
1	125.3	77.9	78.100	−0.200
2	98.2	76.8	76.988	−0.188
3	201.4	81.5	81.223	0.277
4	147.3	79.8	79.003	0.797
5	145.9	78.2	78.945	−0.745
6	124.7	78.3	78.075	0.225
7	112.2	77.5	77.563	−0.063
8	120.2	77.0	77.891	−0.891
9	161.2	80.1	79.573	0.527
10	178.9	80.2	80.299	−0.099
11	159.5	79.9	79.503	0.397
12	145.8	79.0	78.941	0.059
13	75.1	76.7	76.040	0.660
14	151.4	78.2	79.171	−0.971
15	144.2	79.5	78.876	0.624
16	125.0	78.1	78.088	0.012
17	198.8	81.5	81.116	0.384
18	132.5	77.0	78.396	−1.396
19	159.6	79.0	79.508	−0.508
20	110.7	78.6	77.501	1.099

■

Casi de la misma forma que las desviaciones de la media en el caso de una muestra se combinaron para obtener la estimación $s^2 = \sum(x_i - \bar{x})^2/(n-1)$, la estimación de σ^2 en el aná-

lisis de regresión se basa en elevar al cuadrado y sumar los residuos. Se continúa con el uso del símbolo s^2 para la varianza estimada, así que no lo confunda con la s^2 previa.

DEFINICIÓN

La **suma de cuadrados del error** (de manera equivalente, suma de cuadrados del residuo) denotada por SSE, es

$$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = \sum [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2$$

y la estimación de σ^2 es

$$\hat{\sigma}^2 = s^2 = \frac{\text{SSE}}{n-2} = \frac{\sum(y_i - \hat{y}_i)^2}{n-2}$$

El divisor $n - 2$ en s^2 es el número de grados de libertad (gl) relacionados con la estimación (o, de forma equivalente, con la suma de cuadrados del error). Esto se debe a que para obtener s^2, primero se deben estimar los dos parámetros β_0 y β_1, lo que da como resultado una pérdida de 2 gl (del mismo modo que μ se tuvo que estimar en problemas de una muestra, y originó una varianza estimada con base en $n - 1$ gl). Al reemplazar cada y_i en la fórmula para s^2 por la variable aleatoria Y_i se obtiene el estimador S^2. Se puede demostrar que S^2 es un estimador insesgado para σ^2 (aunque el estimador S no es insesgado para σ).

Ejemplo 12.7 (Continuación del ejemplo 12.6)

Los residuos para los datos de la tasa de filtración-contenido de humedad se calcularon antes. La suma de cuadrados del error correspondiente es

$$\text{SSE} = (-0.200)^2 + (-0.188)^2 + \cdots + (1.099)^2 = 7.968$$

Entonces la estimación de σ^2 es $\hat{\sigma}^2 = s^2 = 7.968/(20 - 2) = 0.4427$, y la desviación estándar estimada es $\hat{\sigma} = s = \sqrt{0.4427} = 0.665$. A grandes rasgos, 0.665 es la magnitud de una desviación representativa desde la recta de regresión estimada. ■

Los cálculos en SSE de la fórmula definida implica muchas operaciones aritméticas porque primero deben calcularse los valores pronosticados y los residuos. El uso de la siguiente fórmula de computadora no requiere de esas cantidades.

$$\text{SSE} = \sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i$$

Esta expresión resulta de sustituir $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ es $\sum(y_i - \hat{y}_i)^2$, elevar al cuadrado el sumando, aplicar la suma a los tres términos resultantes y simplificar. Esta fórmula de cálculo es particularmente sensible a los efectos de redondeo en $\hat{\beta}_0$ y $\hat{\beta}_1$, así que llevar tantos dígitos como sea posible en los cálculos intermedios protege contra el error por redondeo.

Ejemplo 12.8

En el artículo "Promising Quantitative Nondestructive Evaluation Techniques for Composite Materials" (*Materials Evaluation*, 1985: 561-565) se describe un estudio para investigar cómo la propagación de una onda de esfuerzo ultrasónica por una sustancia depende de las propie-

dades de la sustancia. Los datos siguientes de la resistencia a la fractura (x, como porcentaje de la resistencia máxima a la tensión) y la atenuación (y, en neper/cm, la disminución de la amplitud de la onda de esfuerzo) en compuestos de poliéster reforzado con fibra de vidrio se leyeron de una gráfica que apareció en el artículo. El patrón lineal sustancial de la gráfica de dispersión indica el modelo de regresión lineal simple.

x	12	30	36	40	45	57	62	67	71	78	93	94	100	105
y	3.3	3.2	3.4	3.0	2.8	2.9	2.7	2.6	2.5	2.6	2.2	2.0	2.3	2.1

Las cantidades de resumen necesarias son $n = 14$, $\sum x_i = 890$, $\sum x_i^2 = 67\,182$, $\sum y_i = 37.6$, $\sum y_i^2 = 103.54$ y $\sum x_i y_i = 2234.30$, de donde $S_{xx} = 10\,603.4285714$, $S_{xy} = -155.98571429$, $\hat{\beta}_1 = -0.0147109$ y $\hat{\beta}_0 = 3.6209072$. La fórmula de cálculo para SSE da

$$\begin{aligned} \text{SSE} &= 103.54 - (3.6209072)(37.6) - (-0.0147109)(2234.30) \\ &= 0.2624532 \end{aligned}$$

así que $s^2 = 0.2624532/12 = 0.0218711$ y $s = 0.1479$. Cuando $\hat{\beta}_0$ y $\hat{\beta}_1$ se redondean a tres decimales en la fórmula de cálculo para SSE, el resultado es

$$\text{SSE} = 103.54 - (3.621)(37.6) - (-0.015)(2234.30) = 0.905$$

que es más de tres veces el valor correcto. ■

Coeficiente de determinación

En la figura 12.10 se muestran tres diagramas de dispersión diferentes de datos bivariados. En las tres gráficas, las alturas de los distintos puntos varían en forma sustancial, lo cual indica que hay mucha variabilidad en los valores observados de y. Los puntos en la primera gráfica caen en una recta. En este caso, la variación muestral (100%) en y se puede atribuir al hecho de que x y y se relacionan linealmente en combinación con la variación en x. Los puntos de la figura 12.10(b) no caen en una recta, pero comparados con la variabilidad global de y, las desviaciones desde la recta de mínimos cuadrados son pequeñas. Es razonable concluir, en este caso, que gran parte de la variación observada de y se puede atribuir a la relación lineal aproximada entre las variables postuladas por el modelo de regresión lineal simple. Cuando la gráfica de dispersión se parece a la figura 12.10(c), hay una variación sustancial respecto a la recta de

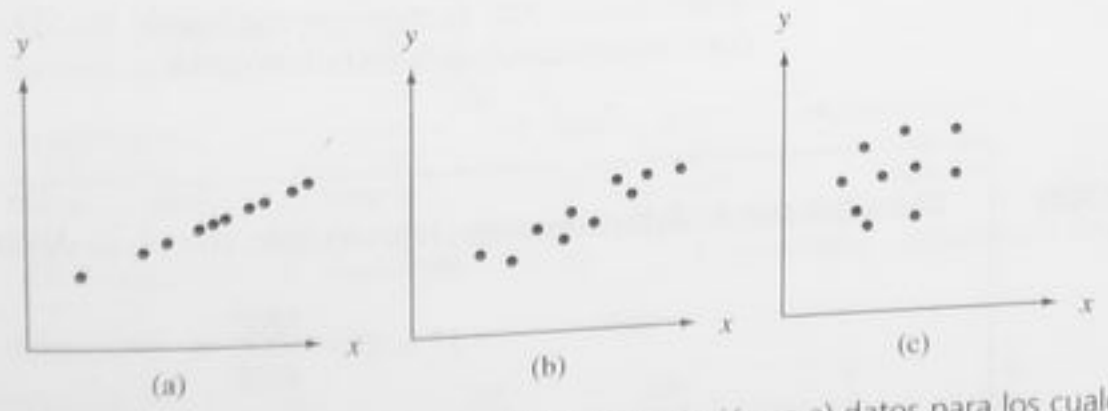


Figura 12.10 Uso del modelo para explicar la variación y: a) datos para los cuales se explica toda la variación; b) datos para los cuales se explica casi toda la variación; c) datos para los cuales se explica para variación

mínimos cuadrados en relación con la variación global de y, así que el modelo de regresión lineal simple no explica la variación en y al relacionar y con x.

La suma de cuadrados del error SSE se puede interpretar como una medida de la variación en y que el modelo deja sin explicar; es decir, cuánto no se puede atribuir a una relación lineal. En la figura 12.10(a), SSE = 0, y no hay variación inexplicada, mientras que la variación inexplicada es pequeña para los datos de la figura 12.10(b) y mucho mayor en la figura 12.10(c). Una medida cuantitativa de la cantidad total de variación en los valores observados de y la proporciona la **suma total de cuadrados.**

$$\text{SST} = S_{yy} = \sum(y_i - \bar{y})^2 = \sum y_i^2 - (\sum y_i)^2/n$$

La suma total de cuadrados es la suma de las desviaciones cuadradas respecto a la media muestral de los valores y observados. Así, el número $\bar{y}$ se resta de cada y_i en SST, mientras que SSE requiere restar cada valor esperado distinto $\hat{y}_i$ del valor observado y_i correspondiente. Al igual que SSE es la suma de las desviaciones cuadradas respecto a la recta de mínimos cuadrados $y = \hat{\beta}_0 + \hat{\beta}_1 x$, SST es la suma de desviaciones cuadradas respecto a la recta horizontal a la altura $\bar{y}$ (puesto que entonces las desviaciones verticales son $y_i - \bar{y}$), como se ilustra en la figura 12.11. Además, debido a que la suma de desviaciones cuadradas respecto a la recta de mínimos cuadrados es más pequeña que la suma de desviaciones cuadradas respecto a *cualquier* otra línea, SSE < SST a menos que una recta horizontal sea la recta de mínimos cuadrados. La relación SSE/SST es la proporción de la variación que el modelo de regresión lineal simple no aclara, y 1 − SSE/SST (un número entre 0 y 1) es la proporción de variación de y observada explicada con el modelo.

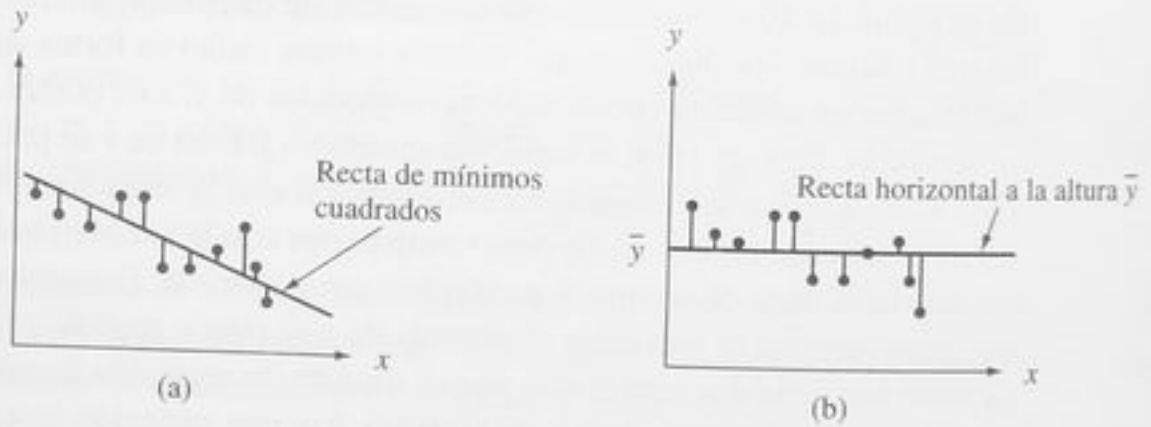


Figura 12.11 Sumas de cuadrados ilustradas: (a) SSE = suma de desviaciones cuadradas respecto a la recta de mínimos cuadrados: (b) SST = suma de desviaciones cuadradas respecto a la recta horizontal

DEFINICIÓN

El **coeficiente de determinación**, representado con r^2, se determina mediante

$$r^2 = 1 - \frac{\text{SSE}}{\text{SST}}$$

Se interpreta como la proporción de variación de y observada que se explica mediante el modelo de regresión lineal simple (atribuida a una relación lineal simple entre y y x).

Mientras mayor sea el valor de r^2, más exitoso es el modelo de regresión lineal simple en explicar la variación de y. Cuando el análisis de regresión se hace mediante software, ya sea r^2 o $100r^2$ (el porcentaje de variación que se explica mediante el modelo) es una parte sobresaliente del resultado. Si r^2 es pequeña, un analista por lo general desea buscar un modelo alternativo (ya sea un modelo no lineal o un modelo de regresión lineal simple que requiere más de una sola variable independiente) que explica con más eficacia la variación de y.

Ejemplo 12.9 (Continuación del ejemplo 12.4)

La gráfica de dispersión de los datos de concreto sin finos de la figura 12.8 sin duda augura un valor alto de r^2. Con

$$\hat{\beta}_0 = 118.909917 \qquad \hat{\beta}_1 = -0.90473066 \qquad \sum y_i = 299.8$$
$$\sum x_i y_i = 32\,308.59 \qquad \sum y_i^2 = 6430.06$$

se tiene

$$\text{SST} = 6430.06 - \frac{299.8^2}{15} = 438.057333 \approx 438.06$$

$$\text{SSE} = 6430.06 - (118.909917)(299.8) - (-0.90473066)(32\,308.59)$$
$$= 11.4388 \approx 11.44$$

Entonces el coeficiente de determinación es

$$r^2 = 1 - \frac{11.44}{438.06} = 1 - 0.026 = 0.974$$

Es decir, 97.4% de la variación observada en la porosidad es atribuible a (se explica mediante) la relación lineal aproximada entre la porosidad y el peso unitario, un resultado muy impresionante. (¡Muchos científicos sociales morirían por un valor de r^2 mucho mayor que 0.5!)

En la figura 12.12, se ilustra el resultado parcial de MINITAB para los datos de porosidad-peso unitario de los ejemplos 12.4 y 12.9; el software también proporciona los valores esperados y los residuos si se solicitan, así como otra información. Los formatos que utilizan otros paquetes de software difieren un poco del de MINITAB, pero el contenido de información es muy similar. Las cantidades como las desviaciones estándar, relaciones t y la tabla ANOVA se analizan en la sección 12.3.

```
La ecuación de regresión es
porosidad = 119 - 0.905 peso unitario
                          β̂₀   β̂₁
Predictor          Coef          Desv. est.     Relación t          p
Constante       118.910               4.499          26.43      0.000
Peso unitario -0.90473             0.04109         -22.02      0.000
                                    100r²
s = 0.9380      R cuadrada = 97.4%      R cuadrada (ajuste) = 97.2%

Análisis de varianza                SSE

                                                        F          p
Fuente          GL          SC          CM         484.84      0.000
Regresión        1      426.62      426.62
Error           13       11.44        0.88
                                        SST
Total           14      438.06
```

Figura 12.12 Resultado de MINITAB para la regresión de los ejemplos 12.4 y 12.9 ■

El coeficiente de determinación se puede escribir de una forma un poco distinta al introducir una tercera suma de cuadrados, **suma de cuadrados de regresión**, SSR, determinada mediante SSR = SST – SSE. La suma de cuadrados de la regresión se interpreta como la cantidad de variación total que *se* explica con el modelo. Entonces, se tiene

$$r^2 = 1 - \text{SSE/SST} = (\text{SST} - \text{SSE})/\text{SST} = \text{SSR/SST}$$

la relación entre la variación explicada y la variación total. La tabla ANOVA de la figura 12.12 muestra que SSR = 426.62, de donde $r^2 = 426.62/438.06 = 0.974$.

Terminología y ámbito del análisis de regresión

Francis Galton fue el primero que utilizó el término *análisis de regresión* en el siglo XIX por su trabajo sobre la relación entre la altura del padre x y la del hijo y. Después de reunir varios pares (x_i, y_i), Galton empleó el principio de los mínimos cuadrados para obtener la ecuación de la recta de regresión estimada con el objetivo de usarla para predecir la estatura del hijo a partir de la del padre. Al usar la recta obtenida, Galton encontró que si la estatura del padre estaba arriba del promedio, se esperaría que el hijo tuviera una estatura arriba del promedio, *aunque no tan alto como el padre*. De manera similar, el hijo de un padre más bajo que el promedio tendría que ser más bajo que el promedio, pero no tanto como el padre. Así que la estatura esperada de un hijo es "atraída" hacia la media; debido a que *regresión* significa venir o regresar, Galton adoptó la terminología *recta de regresión*. Este fenómeno de ser atraído hacia la media se observa en muchas situaciones (p. ej., los promedios de bateo de un año a otro en el béisbol) y se conoce como **efecto de regresión.**

En la descripción hecha hasta el momento, se supone que la variable independiente está bajo el control del investigador, así que sólo la variable dependiente Y es aleatoria. Sin embargo, éste no fue el caso con el experimento de Galton; las estaturas de los padres no se seleccionaron con anticipación, sino que X y Y fueron aleatorias. Los métodos y conclusiones del análisis de regresión se aplican cuando los valores de la variable independiente se fijan por anticipado y cuando son aleatorios, pero como las deducciones e interpretaciones son más directas en el primer caso, se sigue con valores fijos. Consulte más información en el excelente libro de John Neter *et al.*, listado en la bibliografía del capítulo.

Ejercicios | Sección 12.2 (12 a 29)

12. En el ejercicio 4 se muestran los datos de x = carga másica de DBO y y = eliminación másica de DBO. Los valores de cantidades de resumen relevantes son

$n = 14$ $\sum x_i = 517$
$\sum y_i = 346$ $\sum x_i^2 = 39\,095$
$\sum y_i^2 = 17\,454$ $\sum x_i y_i = 25\,825$

a. Obtenga la ecuación de la recta de mínimos cuadrados.

b. Prediga el valor de la eliminación másica de DBO para una sola observación hecha cuando la carga másica de DBO es 35 y calcule el valor del residuo correspondiente.

c. Calcule la SSE y a continuación un estimación puntual de σ.

d. ¿Qué proporción de la variación de eliminación observada se explica mediante la relación lineal aproximada entre las dos variables?

e. Los dos últimos valores de x, 103 y 142, son mucho mayores que los demás. ¿Cómo la eliminación de las dos observaciones correspondientes de la muestra afecta a la ecuación de la recta de mínimos cuadrados y al valor de r^2? Ajuste los valores de las

cantidades de resumen y utilice el hecho de que el nuevo valor de SSE es 311.79.

13. Los datos siguientes de x = densidad actual (mA/cm^2) y y = rapidez de deposición (μm/min) aparecieron en el artículo "Plating of 60/40 Tin/Lead Solder for Head Termination Metallurgy" (*Plating and Surface Finishing*, enero de 1997: 38-40). ¿Está de acuerdo con la afirmación del autor del artículo de que "se obtuvo una relación lineal para la rapidez de deposición de estaño-plomo en función de la densidad de corriente"? Explique su razonamiento.

x	20	40	60	80
y	0.24	1.20	1.71	2.22

14. Refiérase a los datos de la relación temperatura-eficiencia del tanque proporcionados en el ejercicio 1.
   a. Determine la ecuación de la recta de regresión estimada.
   b. Calcule una estimación puntual para la relación de eficiencia promedio real cuando la temperatura del tanque es 182.
   c. Calcule los valores de los residuos de la recta de mínimos cuadrados para las cuatro observaciones en las que la temperatura es 182. ¿Por qué no todos tienen el mismo signo?
   d. ¿Qué proporción de la variación observada de la relación en eficiencia se puede atribuir a la relación de regresión lineal simple entre las dos variables?

15. Se determinaron los valores del módulo de elasticidad (MOE, la relación del esfuerzo, es decir, fuerza por unidad de área, a la deformación, es decir, la deformación por unidad de longitud, en GPa) y la resistencia a la flexión (una medida de la capacidad para resistir la falla por flexión, en MPa) para una muestra de vigas de concreto de cierto tipo, y se obtuvieron los siguientes datos (leídos de una gráfica del artículo "Effects of Aggregates and Microfillers on the Flexural Properties of Concrete", *Magazine of Concrete Research*, 1997: 81-98).

MOE	29.8	33.2	33.7	35.3	35.5	36.1	36.2
Resistencia	5.9	7.2	7.3	6.3	8.1	6.8	7.0
MOE	36.3	37.5	37.7	38.7	38.8	39.6	41.0
Resistencia	7.6	6.8	6.5	7.0	6.3	7.9	9.0
MOE	42.8	42.8	43.5	45.6	46.0	46.9	48.0
Resistencia	8.2	8.7	7.8	9.7	7.4	7.7	9.7
MOE	49.3	51.7	62.6	69.8	79.5	80.0	
Resistencia	7.8	7.7	11.6	11.3	11.8	10.7	

   a. Elabore un diagrama de tallo y hojas de los valores del módulo de elasticidad y comente acerca de algunas características interesantes.
   b. ¿El valor de la resistencia se determina por completo y de manera única mediante el valor del módulo de elasticidad? Explique.
   c. Utilice el siguiente resultado de MINITAB a fin de obtener la ecuación de la recta de mínimos cuadrados para predecir la resistencia a partir del módulo de elasticidad y, luego, prediga la resistencia para una viga cuyo módulo de elasticidad es 40. ¿Confiaría en usar la recta de mínimos cuadrados para predecir la resistencia cuando el módulo de elasticidad es 100? Explique.

```
                                Desv.   Relación
Predictor          Coef         est.          t       p
Constante        3.2925       0.6008       5.48   0.000
mod. de
elast.          0.10748      0.01280       8.40   0.000

s = 0.8657      R cuadrada = 73.8%
R cuadrada      (ajustada) = 72.8%

Análisis de varianza

FUENTE        GL        SC        CM        F       p
Regresión      1    52.870    52.870    70.55   0.000
Error         25    18.736     0.749
Total         26    71.605
```

   d. ¿Cuáles son los valores de SSE, SST y el coeficiente de determinación? ¿Estos valores indican que el modelo de regresión lineal simple describe de manera eficaz la relación entre las dos variables? Explique.

16. En el artículo "Characterization of Highway Runoff in Austin, Texas, Area" (*J. of Envir. Engr.*, 1998: 131-137) se ilustra un diagrama de dispersión, junto con la recta de mínimos cuadrados, de x = volumen de lluvia (m^3) y y = volumen de escurrimiento (m^3) para determinado lugar. Los valores siguientes se tomaron de la gráfica.

x	5	12	14	17	23	30	40	47
y	4	10	13	15	15	25	27	46

x	55	67	72	81	96	112	127
y	38	46	53	70	82	99	100

   a. ¿El diagrama de dispersión de los datos respalda el uso del modelo de regresión lineal simple?
   b. Calcule las estimaciones puntuales de la pendiente y la ordenada al origen de la recta de regresión poblacional.
   c. Calcule una estimación puntual del volumen de escurrimiento promedio real cuando el volumen de lluvia es 50.
   d. Calcule una estimación puntual de la desviación estándar σ.
   e. ¿Qué proporción de la variación observada en el volumen de escurrimiento se puede atribuir a la rela-

ción de regresión lineal simple entre escurrimiento y lluvia?

17. En el artículo "Use of Fly Ash or Silica Fume to Increase the Resistance of Concrete to Feed Acids" (*Magazine of Concrete Research*, 1997· 337-344) se describe una regresión de y = contenido de calcio (g/L) en x = material disuelto (mg/cm^2). La ecuación de la recta de regresión estimada fue $y = 3.678 + 0.144x$, con $r^2 = 0.860$, basada en $n = 23$.

a. Interprete la pendiente estimada 0.144 y el coeficiente de determinación 0.860.

b. Calcule una estimación puntual del contenido promedio real de calcio cuando la cantidad de material disuelto es 50 mg/cm^2.

c. El valor de la suma total de cuadrados fue SST = 320.398. Calcule una estimación de la desviación estándar del error σ en el modelo de regresión lineal simple.

18. Las siguientes estadísticas de resumen se obtuvieron de un estudio en el que se utilizó el análisis de regresión para investigar la relación entre la flexión del pavimento y la temperatura superficial en varios lugares de una carretera estatal. Aquí x = temperatura (°F) y y = factor de ajuste de flexión ($y \geq 0$):

$n = 15 \quad \sum x_i = 1425 \quad \sum y_i = 10.68$

$\sum x_i^2 = 139\,037.25 \quad \sum x_i y_i = 987.645$

$\sum y_i^2 = 7.8518$

(Se hicieron más de 15 observaciones en el estudio; la referencia es "Flexible Pavement Evaluation and Rehabilitation", *Transportation Eng. J.*, 1977: 75-85.)

a. Calcule $\hat{\beta}_1$, $\hat{\beta}_0$ y la ecuación de la recta de regresión estimada. Dibuje la recta estimada.

b. ¿Cuál es la estimación del cambio esperado en el factor de ajuste de flexión cuando la temperatura aumenta 1 °F?

c. Suponga que la temperatura se midió en °C y no en °F. ¿Cuál sería la recta de regresión estimada? Conteste el inciso (b) para un aumento de 1 °C. (*Sugerencia*: °F = (9/5)°C + 32; ahora sustituya la "x anterior" en términos de la "x nueva".)

d. Si fuera posible una temperatura superficial de 200 °F, ¿utilizaría la recta estimada del inciso (a) para pronosticar el factor de flexión para esta temperatura? ¿Por qué sí o por qué no?

19. Los datos siguientes son representativos de lo que aparece en el artículo "An Experimental Correlation of Oxides of Nitrogen Emissions from Power Boilers Based on Field Data" (*J. Eng for Power*, julio de 1973: 165-170), con x = rapidez de liberación de calor del área del quemador (MBtu/h-pie^2) y y = proporción de emisiones de NO_x (ppm):

x	100	125	125	150	150	200	200
y	150	140	180	210	190	320	280

x	250	250	300	300	350	400	400
y	400	430	440	390	600	610	670

a. Suponga que es válido el modelo de regresión lineal simple y obtenga la estimación de mínimos cuadrados de la recta de regresión verdadera.

b. ¿Cuál es la estimación de la proporción de emisiones de NO_x cuando la cantidad de calor liberada del área del quemador es igual a 225?

c. Estime cuánto cambiaría la cantidad de emisiones de NO_x cuando la rapidez de liberación del área del quemador disminuye en 50.

d. ¿Utilizaría la recta de regresión estimada para pronosticar la proporción de emisiones para una rapidez de liberación de 500? ¿Por qué sí o por qué no?

20. En varios estudios se ha demostrado que los líquenes (ciertas plantas compuestas de un alga y un hongo) son excelentes indicadores biológicos de la contaminación del aire. En el artículo "The Epiphytic Lichen Hypogymnia Physodes as a Biomonitor of Atmospheric Nitrogen and Sulphur Deposition in Norway" (*Envir. Monitoring and Assessment*, 1993: 27-47) aparecen los siguientes datos (obtenidos de una gráfica) de x = depósito en húmedo de $x = NO_3^-$ en (g N/m^2) y y = N de liquen (% de peso en seco):

x	0.05	0.10	0.11	0.12	0.31	0.37	0.42
y	0.48	0.55	0.48	0.50	0.58	0.52	1.02

x	0.58	0.68	0.68	0.73	0.85	0.92
y	0.86	0.86	1.00	0.88	1.04	1.70

El autor utilizó la regresión lineal simple para analizar los datos. Utilice el resultado de MINITAB para contestar las siguientes preguntas:

a. ¿Cuáles son las estimaciones de mínimos cuadrados para β_0 y β_1?

b. Prediga el N de liquen para un valor de depósito de NO_3^- de 0.5.

c. ¿Cuál es la estimación de σ^2?

d. ¿Cuál es el valor de la variación total y cuánto de ésta se puede explicar mediante la relación de modelo?

```
La ecuación de regresión es N de liquen
= 0.365 + 0.967 depósito de NO;

                    Coef   Des Est   Relacion      P
Predictor
Constante        0.36510   0.09904       3.69  0.004
Depósito de NO;   0.9668    0.1829       5.29  0.000

s = 0.1932   R cuadrada = 71.7%   R cuadrada
(ajustada) = 69.2%

Análisis de varianza

FUENTE       GL       SC       CM       F       P
Regresión     1   1.0427   1.0427   27.94   0.000
Error        11   0.4106   0.0373
Total        12   1.4533
```

21. En el artículo "Effects of Bike Lanes on Driver and Bicyclist Behavior" (*ASCE Transportation Eng. J.*, 1977: 243-256) se describen los resultados de un análisis de regresión con x = espacio disponible de recorrido en pies (una medida conveniente del ancho de la calzada, definido como la distancia entre un ciclista y la línea del centro de la calzada) y la distancia de separación y entre una bicicleta y un automóvil que pase (determinada con una fotografía). A continuación se dan los datos para diez calles con carriles para bicicletas:

x	12.8	12.9	12.9	13.6	14.5
y	5.5	6.2	6.3	7.0	7.8

x	14.6	15.1	17.5	19.5	20.8
y	8.3	7.1	10.0	10.8	11.0

a. Compruebe que $\sum x_i = 154.20$, $\sum y_i = 80$, $\sum x_i^2 = 2452.18$, $\sum x_i y_i = 1282.74$ y $\sum y_i^2 = 675.16$.

b. Obtenga la ecuación de la recta de regresión estimada.

c. ¿Qué distancia de separación predeciría para otra calle cuyo valor de espacio de recorrido disponible es 15.0?

d. ¿Cuál sería la estimación de la distancia de separación esperada para todas las calles que tienen un valor de espacio de recorrido disponible de 15.0?

22. a. Utilice las cantidades de resumen del ejercicio 18 para estimar la desviación estándar de la desviación aleatoria ϵ en el modelo de regresión lineal simple.

b. Con base en las cantidades de resumen del ejercicio 18, ¿qué proporción de variación en el factor de ajuste de flexión se explica mediante la relación de regresión lineal simple entre el factor de ajuste y la temperatura?

23. a. Obtenga la SSE para los datos del ejercicio 19 a partir de la fórmula de definición [$SSE = \sum(y_i - \hat{y}_i)^2$] y compare con el valor calculado de la fórmula de cálculo.

b. Calcule el valor de la suma total de cuadrados. ¿El modelo de regresión lineal simple al parecer resulta eficaz para explicar la variación en la proporción de emisiones? Justifique su afirmación.

24. Los datos siguientes se tomaron de una gráfica del artículo "Reactions on Painted Steel Under the Influence of Sodium Chloride, and Combinations Thereof" (*Ind. Engr. Chem. Prod. Res. Dev.*, 1985: 375-378). La variable independiente es la rapidez de sedimentación de SO_2 (mg/m^2/día) y la variable dependiente es la pérdida de peso de acero (g/m^2).

x	14	18	40	43	45	112
y	280	350	470	500	560	1200

a. Construya una gráfica de dispersión. ¿El modelo de regresión lineal es razonable en esta situación?

b. Calcule la ecuación de la recta de regresión estimada.

c. ¿Qué porcentaje de variación observada en pérdida de peso de acero se puede atribuir a la relación de modelo en combinación con la variación en la rapidez de sedimentación?

d. Debido a que el valor máximo de x en la muestra excede por mucho a los otros, esta observación podría haber sido definitiva en la determinación de la ecuación de la recta estimada. Elimine esta observación y vuelva a calcular la ecuación. ¿La nueva ecuación difiere en forma considerable de la original (puede considerar valores esperados)?

25. Demuestre que b_1 y b_0 de las expresiones (12.2) y (12.3) satisfacen las ecuaciones normales.

26. Demuestre que el "punto de promedios" $(\bar{x}, \bar{y})$ está en la recta de regresión estimada.

27. Suponga que un investigador tiene datos de la cantidad de espacio de anaquel x destinado a exhibir un producto y el ingreso por ventas y de ese producto. Es posible que el investigador desee ajustar un modelo para el cual la recta de regresión verdadera pasa por (0, 0). El modelo adecuado es $Y = \beta_1 x + \epsilon$. Suponga que $(x_1, y_1), \ldots, (x_n, y_n)$ son pares observados generados a partir de este modelo y obtenga el estimador de mínimos cuadrados de β_1. (*Sugerencia*: escriba la suma de desviaciones cuadradas en función de b_1, un valor trivial, y por medio del cálculo determine el valor minimizante de b_1.)

28. a. Considere los datos del ejercicio 20. Suponga que en lugar de la recta de mínimos cuadrados que pasa por los puntos $(x_1, y_1), \ldots, (x_n, y_n)$, se desea la recta de mínimos cuadrados que pasa por $(x_1 - \bar{x}, y_1), \ldots, (x_n - \bar{x}, y_n)$. Construya una gráfica de dispersión de

los puntos (x_i, y_i) y luego la de los puntos $(x_i - \bar{x}, y_i)$. Por medio de las gráficas explica en forma intuitiva cómo se relacionan entre sí las dos rectas de mínimos cuadrados.

b. Suponga que en lugar del modelo $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ $(i = 1, \ldots, n)$, se desea ajustar un modelo de la forma $Y_i = \beta_0^* + \beta_1^*(x_i - \bar{x}) + \epsilon_i$ $(i = 1, \ldots, n)$. ¿Cuáles son los estimadores de mínimos cuadrados de β_0^* y β_1^* y cómo se relacionan con $\hat{\beta}_0$ y $\hat{\beta}_1$?

29. Considere los siguientes tres conjuntos de datos, en los que las variables de interés son x = distancia de ida y vuelta al trabajo y y = tiempo para ir y regresar del trabajo. Con base en una gráfica de dispersión y los valores de s y r^2, ¿en cuál situación la regresión lineal simple sería más eficaz, y en cuál menos efectiva y por qué?

Conjunto de datos	1		2		3	
	x	y	x	y	x	y
	15	42	5	16	5	8
	16	35	10	32	10	16
	17	45	15	44	15	22
	18	42	20	45	20	23
	19	49	25	63	25	31
	20	46	50	115	50	60
S_{xx}	17.50		1270.8333		1270.8333	
S_{xy}	29.50		2722.5		1431.6667	
$\hat{\beta}_1$	1.685714		2.142295		1.126557	
$\hat{\beta}_0$	13.666672		7.868852		3.196729	
SST	114.83		5897.5		1627.33	
SSE	65.10		65.10		14.48	

12.3 Inferencias acerca del parámetro de la pendiente β_1

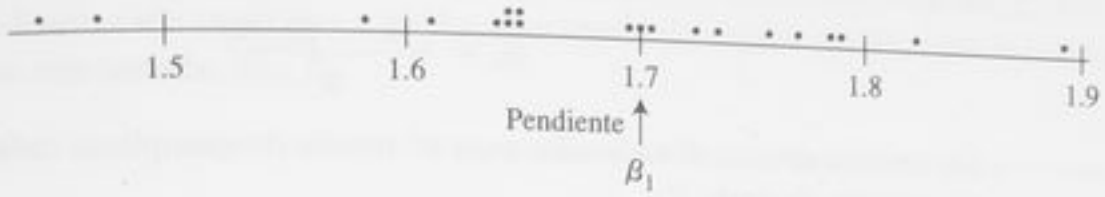
En casi todo el trabajo inferencial realizado hasta el momento, el concepto de variabilidad de muestreo ha sido dominante. En particular, las propiedades de las distribuciones de muestreo de varios estadísticos han sido la base para desarrollar fórmulas del intervalo de confianza y métodos de prueba de hipótesis. La idea clave es que el valor de casi cualquier cantidad calculada a partir de los datos muestrales, el valor de casi cualquier estadístico, va a variar de una muestra a otra.

Ejemplo 12.10 Reconsidere los datos de x = rapidez de liberación de calor del área del quemador y y = proporción de emisiones de NO_x del ejercicio 12.19 en la sección anterior. Hay 14 observaciones, hechas en los valores de x, 100, 125, 125, 150, 150, 200, 200, 250, 250, 300, 300, 350, 400 y 400, respectivamente. Suponga que la pendiente y la ordenada al origen de la recta de regresión verdadera son $\beta_1 = 1.70$ y $\beta_0 = -50$, con $\sigma = 35$ (consistente con los valores $\hat{\beta}_1 = 1.7114$, $\hat{\beta}_0 = -45.55$, $s = 36.75$). Se procede a generar una muestra de desviaciones aleatorias $\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_{14}$ de una distribución normal con media 0 y desviación estándar 35, y luego se añade $\tilde{\epsilon}_i$ a $\beta_0 + \beta_1 x_i$ para obtener los 14 valores correspondientes de y. Los cálculos de regresión se realizaron para obtener la pendiente estimada, la ordenada al origen y la desviación estándar. Este proceso se repitió un total de 20 veces y se obtuvieron los valores de la tabla 12.1.

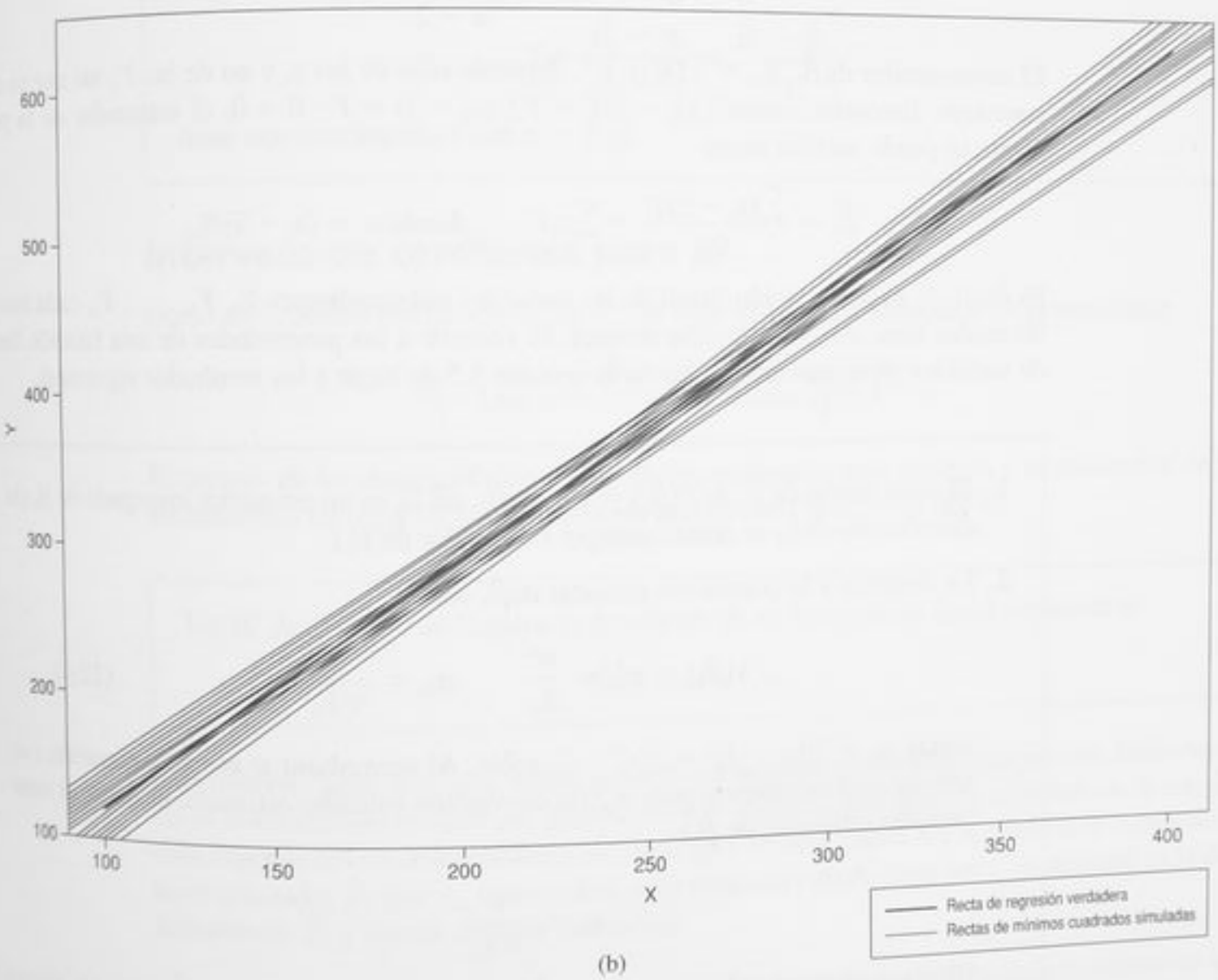
Tabla 12.1 Resultados de la simulación para el ejemplo 12.10

	$\hat{\beta}_1$	$\hat{\beta}_0$	s		$\hat{\beta}_1$	$\hat{\beta}_0$	s
1.	1.7559	−60.62	43.23	11.	1.7843	−67.36	41.80
2.	1.6400	−49.40	30.69	12.	1.5822	−28.64	32.46
3.	1.4699	−4.80	36.26	13.	1.8194	−83.99	40.80
4.	1.6944	−41.95	22.89	14.	1.6469	−32.03	28.11
5.	1.4497	5.80	36.84	15.	1.7712	−52.66	33.04
6.	1.7309	−70.01	39.56	16.	1.7004	−58.06	43.44
7.	1.8890	−95.01	42.37	17.	1.6103	−27.89	25.60
8.	1.6471	−40.30	43.71	18.	1.6396	−24.89	40.78
9.	1.7216	−42.68	23.68	19.	1.7857	−77.31	32.38
10.	1.7058	−63.31	31.58	20.	1.6342	−17.00	30.93

Hay variación clara de los valores estimados de la pendiente y la ordenada al origen, así como de la desviación estándar. Por lo tanto, la ecuación de mínimos cuadrados varía de una muestra a la siguiente. En la figura 12.13 se muestra un diagrama de puntos de las pendientes estimadas, así como las gráficas de la recta de regresión verdadera y las rectas de regresión de 20 muestras.



(a)



(b)

Figura 12.13 Resultados de simulación del ejemplo 12.10: (a) gráfica de puntos de las pendientes estimadas; (b) gráficas de la recta de regresión verdadera y 20 rectas de mínimos cuadrados (de S-Plus) ■

La pendiente β_1 de la recta de regresión poblacional es el cambio promedio real en la variable independiente y relacionada con un incremento unitario en la variable independiente x. La pendiente de la recta de mínimos cuadrados, $\hat{\beta}_1$, proporciona una estimación puntual de β_1. Del mismo modo que un intervalo de confianza para μ y los procedimientos de prueba respecto a μ se basaron en las propiedades de la distribución de muestreo de $\overline{X}$, más inferencias respecto a β_1 se basan en considerar a $\hat{\beta}_1$ como uno estadístico e investigar su distribución de muestreo.

Se supone que los valores de las x_i se eligen antes de realizar el experimento, de modo que sólo las Y_i son aleatorias. Los estimadores (estadísticos y, por consiguiente, variables aleatorias) para β_0 y β_1 se obtienen al reemplazar y_i por Y_i en (12.2) y (12.3):

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(Y_i - \bar{Y})}{\sum(x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \frac{\sum Y_i - \hat{\beta}_1 \sum x_i}{n}$$

De manera similar, el estimador para σ^2 resulta de reemplazar cada y_i en la fórmula para s^2 por la variable aleatoria Y_i:

$$\hat{\sigma}^2 = S^2 = \frac{\sum Y_i^2 - \hat{\beta}_0 \sum Y_i - \hat{\beta}_1 \sum x_i Y_i}{n - 2}$$

El denominador de $\hat{\beta}_1$, $S_{xx} = \sum(x_i - \bar{x})^2$, depende sólo de las x_i y no de las Y_i, así que es una constante. Entonces, como $\sum(x_i - \bar{x})\bar{Y} = \bar{Y}\sum(x_i - \bar{x}) = \bar{Y} \cdot 0 = 0$, el estimador de la pendiente se puede escribir como

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})Y_i}{S_{xx}} = \sum c_i Y_i \qquad \text{donde } c_i = (x_i - \bar{x})/S_{xx}$$

Es decir, $\hat{\beta}_1$ es una función lineal de las variables independientes $Y_1, Y_2, \ldots, Y_n$, cada una de las cuales tiene una distribución normal. El recurrir a las propiedades de una función lineal de variables aleatorias analizadas en la sección 5.5 da lugar a los resultados siguientes.

1. El valor medio de $\hat{\beta}_1$ es $E(\hat{\beta}_1) = \mu_{\hat{\beta}_1} = \beta_1$, así $\hat{\beta}_1$ es un estimador insesgado de β_1 (la distribución de $\hat{\beta}_1$ se centra siempre en el valor de β_1).
2. La varianza y la desviación estándar de $\hat{\beta}_1$ son

$$V(\hat{\beta}_1) = \sigma_{\hat{\beta}_1}^2 = \frac{\sigma^2}{S_{xx}} \qquad \sigma_{\hat{\beta}_1} = \frac{\sigma}{\sqrt{S_{xx}}} \tag{12.4}$$

   donde $S_{xx} = \sum(x_i - \bar{x})^2 = \sum x_i^2 - (\sum x_i)^2/n$. Al reemplazar σ por su estimación s se obtiene una estimación para $\sigma_{\hat{\beta}_1}$ (la desviación estándar estimada, es decir, el error estándar estimado, de $\hat{\beta}_1$):

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}}$$

   (Esta estimación también se puede denotar mediante $\hat{\sigma}_{\hat{\beta}_1}$.)
3. El estimador $\hat{\beta}_1$ tiene una distribución normal (debido a que es una función lineal de las variables aleatorias normales independientes).

De acuerdo con (12.4), la varianza de $\hat{\beta}_1$ es igual a la varianza de σ^2 del término de error aleatorio —o, en forma equivalente, de cualquier Y_i— dividida entre $\sum(x_i - \bar{x})^2$. Debido a que $\sum(x_i - \bar{x})^2$ es una medida de qué tan dispersas están las x_i respecto a $\bar{x}$, se concluye que hacer observaciones en los valores de x_i que están bastantes dispersos da como resultado un estimador más preciso del parámetro de la pendiente (menor varianza de $\hat{\beta}_1$), mientras que los valores

de x_i todos cercanos entre sí implican un estimador muy variable. Por supuesto, si las x_i están muy dispersas, podría ser inapropiado un modelo lineal en todo el intervalo de observación.

Muchos procedimientos inferenciales analizados antes se basaron en estandarizar un estimador al restar primero su valor medio y luego dividir entre su desviación estándar estimada. En particular, los procedimientos de prueba y un IC para la media μ de una población normal utilizaron el hecho de que la variable estandarizada $(\bar{X} - \mu)/(S/\sqrt{n})$ —es decir, $(\bar{X} - \mu)/S_{\bar{X}}$— tenía una distribución t con $n - 1$ gl. Aquí un resultado similar proporciona la clave para más inferencias relacionadas con β_1.

TEOREMA

Las suposiciones del modelo de regresión lineal simple implican que la variable estandarizada

$$T = \frac{\hat{\beta}_1 - \beta_1}{S/\sqrt{S_{xx}}} = \frac{\hat{\beta}_1 - \beta_1}{S_{\hat{\beta}_1}}$$

tiene una distribución t con $n - 2$ gl.

Intervalo de confianza para β_1

Como en la deducción previa de los IC, se empieza con un enunciado de probabilidad:

$$P\left(-t_{\alpha/2,n-2} < \frac{\hat{\beta}_1 - \beta_1}{S_{\hat{\beta}_1}} < t_{\alpha/2,n-2}\right) = 1 - \alpha$$

El manejo de las desigualdades dentro de los paréntesis para aislar β_1 y la sustitución de las estimaciones en lugar de los estimadores proporciona la fórmula del IC.

Un **IC** de $100(1 - \alpha)\%$ **para la pendiente** β_1 de la regresión lineal verdadera es

$$\hat{\beta}_1 \pm t_{\alpha/2,n-2} \cdot s_{\hat{\beta}_1}$$

Este intervalo tiene la misma forma general que muchos de los intervalos previos. Está centrado en la estimación puntual del parámetro, y la cantidad que abarca a cada lado de la estimación depende del nivel de confianza deseado (por el valor crítico t) y la cantidad de variabilidad en el estimador $\hat{\beta}_1$ (por $s_{\hat{\beta}_1}$ que tenderá a ser pequeña cuando haya poca variabilidad en la distribución de $\hat{\beta}_1$ y grande en caso contrario).

Ejemplo 12.11 Las variaciones en el peso de la mampostería con ladrillo de arcilla tienen implicaciones no sólo para el diseño estructural y acústico, sino también para el diseño de sistemas de calefacción, ventilación y de acondicionamiento de aire. En el artículo "Clay Brick Masonry Weight Variation" (*J. of Architectural Engr.*, 1996: 135-137) se ilustra una gráfica de dispersión de $y =$ densidad en seco del mortero (lb/pie^3) contra $x =$ contenido de aire del mortero (%) para una muestra de especímenes de mortero, de la cual se obtuvieron los siguientes datos representativos:

x	5.7	6.8	9.6	10.0	10.7	12.6	14.4	15.0	15.3
y	119.0	121.3	118.2	124.0	112.3	114.1	112.2	115.1	111.3

(continúa)

x	16.2	17.8	18.7	19.7	20.6	25.0
y	107.2	108.9	107.8	111.0	106.2	105.0

La gráfica de dispersión de estos datos en la figura 12.14 en realidad indica cuan apropiado es el modelo de regresión lineal simple; parece haber una relación lineal negativa sustancial entre el contenido de aire y la densidad, una en la que la densidad tiende a disminuir a medida que aumenta el contenido de aire.

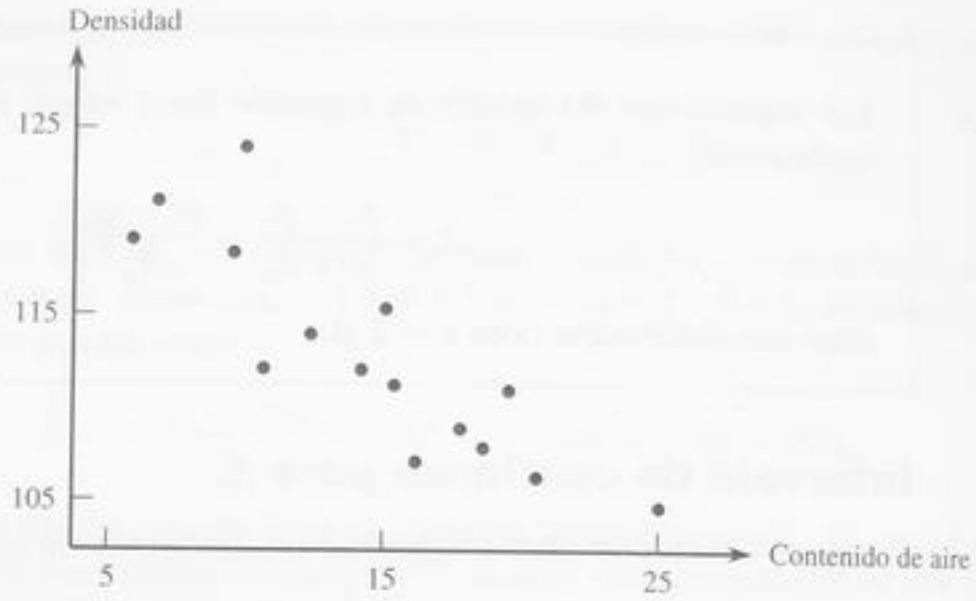


Figura 12.14 Gráfica de dispersión de los datos del ejemplo 12.11

Los valores de los estadísticos de resumen requeridos para el cálculo de las estimaciones por mínimos cuadrados son

$$\sum x_i = 218.1 \qquad \sum y_i = 1693.6 \qquad \sum x_i^2 = 3577.01$$
$$\sum x_i y_i = 24\,252.54 \qquad \sum y_i^2 = 191\,672.90$$

de lo cual $S_{xy} = -372.404$, $S_{xx} = 405.836$, $\hat{\beta}_1 = -0.917622$, $\hat{\beta}_0 = 126.248889$, SST = 454.1693, SSE = 112.4432 y $r^2 = 1 - 112.4432/454.1693 = 0.752$. Alrededor de 75% de la variación observada se puede atribuir a la relación del modelo de regresión lineal simple entre la densidad y el contenido de aire. Los gl del error son $15 - 2 = 13$, de lo cual se obtiene $s^2 = 112.4432/13 = 8.6495$ y $s = 2.941$.

La desviación estándar estimada de $\hat{\beta}_1$ es

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}} = \frac{2.941}{\sqrt{405.836}} = 0.1460$$

El valor crítico t para un nivel de confianza de 95% es $t_{0.025,13} = 2.160$. El intervalo de confianza es

$$-0.918 \pm (2.160)(0.1460) = -0.918 \pm 0.315 = (-1.233, -0.603)$$

Con un alto grado de confianza, se estima que una disminución promedio de densidad de entre 0.603 lb/pie^3 y 1.233 lb/pie^3 se relaciona con un incremento de 1% en el contenido de aire (por lo menos para los valores de contenido de aire entre 5 y 25%, que corresponde a valores de x en la muestra). El intervalo es razonablemente estrecho, lo cual indica que se estimó con precisión la pendiente de la recta poblacional. Observe que el intervalo incluye sólo valores ne-

gativos, así que es bastante seguro que la densidad tienda a disminuir a medida que aumenta el contenido de aire.

En referencia al resultado de SAS de la figura 12.15, se encuentra el valor de $s_{\hat{\beta}_1}$ bajo Estimaciones de los parámetros como el segundo número en la columna Error estándar. Todos los paquetes de estadística de uso común incluyen este error estándar estimado en el resultado. También hay un error estándar estimado para la estadística $\hat{\beta}_0$ de la cual se puede calcular un intervalo de confianza para la ordenada al origen β_0 de la recta de regresión poblacional.

Variable dependiente: DENSIDAD

Análisis de varianza

Fuente	GL	Suma de cuadrados	Cuadrado medio	Valor F	Prob > F
Modelo	1	341.72606	341.72606	39.508	0.0001
Error	13	112.44327	8.64948		
C Total	14	454.16933			

Raíz del ECM	2.94100	R cuadrada	0.7524
Media dep.	112.90667	R cuadrada ajustada	0.7334
V.C.	2.60481		

Estimaciones de los parámetros

Variable	GL	Estimación del parámetro	Error estándar	T para H0: parámetro = 0	Prob > |T|
ORD. AL ORIGEN	1	126.248889	2.25441683	56.001	0.0001
CONT. DE AIRE	1	-0.917622	0.14598888	-6.286	0.0001

Obs	Var. dep. DENSIDAD	Valor predicho	Residuo
1	119.0	121.0	-2.0184
2	121.3	120.0	1.2909
3	118.2	117.4	0.7603
4	124.0	117.1	6.9273
5	112.3	116.4	-4.1303
6	114.1	114.7	-0.5869
7	112.2	113.0	-0.8351
8	115.1	112.5	2.6154
9	111.3	112.2	-0.9093
10	107.2	111.4	-4.1834
11	108.9	109.9	-1.0152
12	107.8	109.1	-1.2894
13	111.0	108.2	2.8283
14	106.2	107.3	-1.1459
15	105.0	103.3	1.6917
Suma de residuos			0
Suma de residuos cuadrados			112.4433
SC de residuos predichos (sc pred.)			146.4144

Figura 12.15 Resultado de SAS para los datos del ejemplo 12.11 ■

Procedimientos de prueba de hipótesis

Como antes, la hipótesis nula en una prueba respecto a β_1 será un enunciado de igualdad. El valor nulo (valor de β_1 que en la hipótesis nula se afirma es verdadero) se denota con β_{10} (léase "beta uno cero", *no* "beta diez"). La estadística de prueba resulta de reemplazar β_1 en la variable T estandarizada por un valor nulo β_{10}; es decir, de estandarizar el estimador de β_1 bajo

el supuesto de que H_0 es verdadera. Así, la estadística de prueba tiene una distribución t con $n - 2$ gl cuando H_0 es verdadera, de modo que la probabilidad de error de tipo I se controla al nivel deseado α mediante un valor crítico apropiado t.

El par de hipótesis encontradas con más frecuencia respecto a β_1 es H_0: $\beta_1 = 0$ contra H_a: $\beta_1 \neq 0$. Cuando esta hipótesis nula es verdadera, $\mu_{Y \cdot x} = \beta_0$ independientemente de x, por lo que conocer x no proporciona información respecto al valor de la variable dependiente. Una prueba de estas dos hipótesis se llama *prueba de utilidad del modelo* en una regresión lineal simple. A menos que n sea bastante pequeña, H_0 se rechaza y la utilidad del modelo se confirma con precisión cuando r^2 es razonablemente grande. El modelo de regresión lineal simple no se debe usar para más inferencias (estimaciones del valor medio o predicciones de valores futuros), a menos que la prueba de utilidad del modelo dé como resultado el rechazo de H_0 para una α adecuadamente pequeña.

Hipótesis nula: H_0: $\beta_1 = \beta_{10}$

Valor del estadístico de prueba: $t = \dfrac{\hat{\beta}_1 - \beta_{10}}{s_{\hat{\beta}_1}}$

Hipótesis alternativa	Región de rechazo para la prueba de nivel α
H_a: $\beta_1 > \beta_{10}$	$t \geq t_{\alpha,n-2}$
H_a: $\beta_1 < \beta_{10}$	$t \leq -t_{\alpha,n-2}$
H_a: $\beta_1 \neq \beta_{10}$	ya sea $t \geq t_{\alpha/2,n-2}$ o $t \leq -t_{\alpha/2,n-2}$

Un valor P basado en $n - 2$ gl se calcula como se hizo antes para las pruebas t en los capítulos 8 y 9.

La **prueba de utilidad del modelo** es una prueba de H_0: $\beta_1 = 0$ contra H_a: $\beta_1 \neq 0$, en cuyo caso el valor del estadístico de prueba es la **relación** t $t = \hat{\beta}_1 / s_{\hat{\beta}_1}$.

Ejemplo 12.12 La limpieza del metal o aleación de aluminio fundido antes del colado se determina sobre todo mediante el contenido de hidrógeno e inclusión de la hornada. En el artículo "Effect of Melt Cleanliness on the Properties of an A1-10 Wt Pct Si-10 Vol Pct SiC(p) Composite" (*Metallurgical Trans.*, 1993: 1631-1645) se describe un estudio en el que varias propiedades de la tensión se relacionaron con x = fracción en volumen de óxidos/inclusiones (%). A continuación se presentan los datos (leídos de una gráfica) de y = elongación (%) de las barras de prueba. Los autores expresaron que la gráfica de dispersión muestra una relación lineal y proporciona la ecuación de la recta de mínimos cuadrados. Por medio del resultado de MINITAB de la figura 12.16, realice una prueba de utilidad del modelo al nivel de significación $\alpha = 0.01$.

x	0.10	0.16	0.31	0.37	0.37	0.46	0.50	0.50	0.60	0.70
y	0.96	1.10	0.80	0.84	0.77	0.87	0.60	0.87	0.60	0.61

x	0.75	0.80	0.90	1.00	1.07	1.08	1.11	1.30	1.37	1.54
y	0.70	0.41	0.40	0.41	0.45	0.59	0.25	0.25	0.08	0.10

```
La ecuación de regresión es
elon = 1.07 - 0.649 frac. vol.

Predictor          Coef  Desv. est.  Relación t      P      Valor P
Constante       1.06930     0.04966       21.53  0.000   para la prueba
Frac. en vol. -0.64884      0.05840      -11.11  0.000 ←de utilidad
                                                       del modelo

s = 0.1049      R cuadrada = 87.3%      R cuadrada (ajustada) = 86.6%

Análisis de varianza

FUENTE          GL          SC          CM          F          P
Regresión        1      1.3583      1.3583     123.42      0.000
Error           18      0.1981      0.0110
Total           19      1.5564
```

$s_{\hat{\beta}_1}$ apunta a Desv. est. Relación t; $t = \dfrac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$ apunta a Relación t

Figura 12.16 Resultado de MINITAB para el ejemplo 12.12

El parámetro de interés es β_1, el cambio esperado en el porcentaje de elongación relacionado con un incremento de 1% en la fracción en volumen de óxidos/inclusiones. H_0: $\beta_1 = 0$ se rechaza en favor de H_a: $\beta_1 \neq 0$ si la relación $t = \hat{\beta}_1/s_{\hat{\beta}_1}$ satisface $t \geq t_{\alpha/2,n-2} = t_{0.005,18} = 2.878$ o $t \leq -2.878$. De la figura 12.15, $\beta_1 = -0.64884$, $s_{\hat{\beta}_1} = 0.05840$, y

$$t = \frac{-0.64884}{0.05840} = -11.11 \qquad \text{(también en el resultado)}$$

Es claro que $-11.11 \leq -2.878$, por lo tanto se rechaza H_0. De modo alternativo, el valor de P es dos veces el área captada bajo la curva t de 18 gl a la izquierda de -11.11. MINITAB proporciona el valor $P = 0.000$, por lo tanto H_0 se debe rechazar con cualquier α razonable. Esta confirmación de la utilidad del modelo de regresión lineal simple proporciona licencia para calcular varias estimaciones y predicciones como se describe en la sección 12.4. ■

Regresión y ANOVA

La división de la suma total de cuadrados $\sum(y_i - \bar{y})^2$ en una parte SSE, que mide la variación inexplicada, y una parte SSR, que mide la variación explicada mediante la relación lineal, es una reminiscencia del ANOVA unidireccional. De hecho, la hipótesis nula H_0: $\beta_1 = 0$ se puede probar contra H_a: $\beta_1 \neq 0$ al construir una tabla de ANOVA (tabla 12.2) y rechazar H_0 si $f \geq F_{\alpha,1,n-2}$.

La prueba F proporciona el mismo resultado que la prueba t de utilidad del modelo $t^2 = f$ y $t^2_{\alpha/2,n-2} = F_{\alpha,1,n-2}$. Casi todos los paquetes de computadora que tienen opciones de regresión

Tabla 12.2 Tabla ANOVA para regresión lineal simple

Fuente de variación	gl	Suma de cuadrados	Cuadrado medio	f
Regresión	1	SSR	SSR	$\dfrac{\text{SSR}}{\text{SSE}/(n-2)}$
Error	$n-2$	SSE	$s^2 = \dfrac{\text{SSE}}{n-2}$	
Total	$n-1$	SST		

incluyen una tabla ANOVA en el resultado. Por ejemplo, en la figura 12.15 se ilustra el resultado de SAS para los datos de mortero del ejemplo 12.11. La tabla ANOVA en la parte superior del resultado tiene $f = 39.508$ con un valor P de 0.0001 para la prueba de utilidad del modelo. La tabla de estimaciones de parámetro da $t = -6.286$, de nuevo con $P = 0.0001$ y $(-6.286)^2 = 39.51$.

Ejercicios | Sección 12.3 (30 a 43)

30. Reconsidere la situación descrita en el ejercicio 7, donde x = resistencia acelerada del concreto y y = resistencia de curado de 28 días. Suponga que el modelo de regresión lineal simple es válido para x entre 1000 y 4000 y que $\beta_1 = 1.25$ y $\sigma = 350$. Considere un experimento en el que $n = 7$ y los valores x en los que se hacen las observaciones son $x_1 = 1000$, $x_2 = 1500$, $x_3 = 2000$, $x_4 = 2500$, $x_5 = 3000$, $x_6 = 3500$ y $x_7 = 4000$.

- **a.** Calcule $\sigma_{\hat{\beta}_1}$, la desviación estándar de $\hat{\beta}_1$.
- **b.** ¿Cuál es la probabilidad de que la pendiente estimada con base en estas observaciones esté entre 1.00 y 1.50?
- **c.** Suponga que también es posible realizar una sola observación en cada uno de los $n = 11$ valores $x_1 = 2000$, $x_2 = 2100, \ldots, x_{11} = 3000$. Si un objetivo importante es estimar β_1 con tanta precisión como sea posible, ¿sería preferible el experimento con $n = 11$ al que tiene $n = 7$?

31. Reconsidere las cantidades de resumen del ejercicio 18 para la regresión de y = factor de flexión en x = temperatura.

- **a.** Calcule la desviación estándar estimada $s_{\hat{\beta}_1}$.
- **b.** Calcule un intervalo de confianza de 95% para β_1, el cambio esperado en el factor de flexión correspondiente a un aumento de temperatura de 1 °F.

32. En el ejercicio 16 de la sección 12.2 se dan los datos de x = volumen de lluvia y y = volumen de escurrimiento (ambos en m^3). Utilice el siguiente resultado de MINITAB para decidir si hay una relación lineal útil entre lluvia y escurrimiento, y después calcule un intervalo de confianza para el cambio promedio real de volumen de escurrimiento relacionado con un aumento de 1 m^3 en el volumen de lluvia.

```
La ecuación de regresión es volumen de
escurrimiento = -1.13 + 0.827 volumen
de lluvia

                        Desv.   Relación
Predictor       Coef     est.          t       p
Constante     -1.128    2.368      -0.48   0.642
Lluvia      0.82697  0.03652      22.64   0.000
s = 5.240  R cuadrada = 97.5%  R cuadrada
(ajustada) = 97.3%
```

33. En el ejercicio 15 de la sección 12.2 se incluyó el resultado de MINITAB para una regresión de la resistencia a la flexión de vigas de concreto en el módulo de elasticidad.

- **a.** Utilice el resultado para calcular un intervalo de confianza a nivel de confianza de 95% para la pendiente β_1 de la recta de regresión poblacional, e interprete el intervalo resultante.
- **b.** Suponga que se tenía la creencia de que cuando el módulo de elasticidad se incrementaba en 1 GPa, el cambio promedio real de resistencia a la flexión sería a lo sumo 0.1 MPa. ¿Los datos muestrales contradicen esta creencia? Formule y pruebe las hipótesis pertinentes.

34. Refiérase al resultado de MINITAB del ejercicio 20, donde x = depósito en húmedo de NO_3^- y y = N de liquen (%).

- **a.** Realice la prueba de utilidad del modelo al nivel 0.01 por medio del método de la región de rechazo.
- **b.** Repita el inciso (a) utilizando el método de valor P.
- **c.** Suponga que antes se creía que cuando la deposición en húmedo de NO_3^- aumenta en 0.1 g de N/m^2, el cambio en N de liquen esperado es por lo menos 0.15%. Realice una prueba de hipótesis al nivel 0.01 para decidir si los datos contradicen esta creencia inicial.

35. ¿De qué manera la aceleración lateral, fuerzas laterales experimentadas en vueltas que en gran medida están bajo el control del conductor, afecta la náusea que perciben los pasajeros de un autobús? En el artículo "Motion Sickness in Public Road Transport: The Effect of Driver, Route, and Vehicle" (*Ergonomics*, 1999: 1646-1664) se muestran los datos de x = dosis para el mareo por movimiento (calculada de acuerdo con un estándar británico para evaluar el movimiento similar en el mar) y y = náusea reportada (%). Las cantidades de resumen importantes son

$n = 17$, $\sum x_i = 222.1$, $\sum y_i = 193$, $\sum x_i^2 = 3056.69$, $\sum x_i y_i = 2759.6$, $\sum y_i^2 = 2975$

Los valores de dosis en la muestra variaron de 6.0 a 17.6.

a. Suponiendo que el modelo de regresión lineal simple es válido para relacionar estas dos variables (los datos sin tratamiento respaldan la suposición), calcule e interprete una estimación del parámetro de la pendiente que transmita información acerca de la precisión y la confiabilidad de la estimación.
b. ¿Al parecer hay una relación lineal útil entre estas dos variables? Conteste la pregunta con el método del valor P.
c. ¿Sería conveniente usar el modelo de regresión lineal simple como base para predecir el % de náusea cuando la dosis es 5.0? Explique su razonamiento.
d. Cuando se utilizó MINITAB para ajustar el modelo de regresión lineal simple a los datos sin tratamiento, se identificó la observación (6.0, 2.50) como una que tiene un impacto importante en el ajuste. Elimine de la muestra esta observación y vuelva a calcular la estimación del inciso (a). Con base en esto, ¿la observación al parecer ejerce una influencia indebida?

36. La niebla (gotitas o aerosoles que transporta el aire) se genera cuando se emplean líquidos eliminadores de metal en operaciones de maquinado para enfriar y lubricar la herramienta y la pieza de trabajo. La generación de niebla es de interés para OSHA, que en fechas recientes disminuyó de forma importante la norma del lugar de trabajo. En el artículo "Variables Affecting Mist Generation from Metal Removal Fluids" (*Lubrication Engr.*, 2002: 10-17) se encuentran los datos siguientes de x = velocidad del flujo de líquido para un aceite soluble al 5% (cm/s) y y = la cantidad de gotitas de niebla que tienen diámetros menores de 10 μm (mg/m^3):

x	89	177	189	354	362	442	965
y	0.40	0.60	0.48	0.66	0.61	0.69	0.99

a. Los investigadores llevaron a cabo un análisis de regresión lineal simple para relacionar las dos variables. ¿Una gráfica de dispersión de los datos respalda esta estrategia?
b. ¿Qué proporción de la variación observada en la niebla se atribuye a la relación de regresión lineal simple entre la velocidad y la niebla?
c. En particular, el interés de los investigadores se centró en el impacto en la niebla de incrementar la velocidad de 100 a 1000 (un factor de 10 que corresponde a la diferencia entre los valores mínimo y máximo de la muestra). Cuando x se incrementa de esta manera, ¿hay evidencia importante de que el incremento promedio real en y sea menor que 0.6?
d. Estime el cambio promedio real en la niebla relacionado con un incremento de 1 cm/s en la velocidad, y procure que lleve información acerca de la precisión y la confiabilidad.

37. En el artículo "Hydrogen, Oxygen, and Nitrogen in Cobalt Metal" (*Metallurgia*, 1969: 121-127) se encuentra una gráfica de los siguientes pares de datos, donde x = presión de gas extraído (micrones) y y = tiempo de extracción (minutos):

x	40	130	155	160	260	275	325	370	420	480
y	2.5	3.0	3.1	3.3	3.7	4.1	4.3	4.8	5.0	5.4

a. Estime σ y la desviación estándar de $\hat{\beta}_1$.
b. Suponga que antes del experimento los investigadores creían que en promedio habría un incremento de 0.006 min en el tiempo de extracción relacionado con un incremento de 1 micrón en la presión. Use el método del valor P con un nivel de significación de 0.10 para decidir si los datos contradicen esta creencia inicial.

38. Refiérase a los datos de x = rapidez de liberación y y = proporción de emisiones de NO_x del ejercicio 19.
a. ¿El modelo de regresión lineal simple especifica una relación útil entre los dos índices? Utilice el procedimiento de prueba adecuado para obtener información sobre el valor P y después llegue a una conclusión al nivel de significación 0.01.
b. Calcule un intervalo de confianza de 95% para el cambio esperado en la proporción de emisiones que corresponde a un aumento de 10 MBtu/h-pie^2 en la rapidez de liberación.

39. Lleve a cabo la prueba de utilidad del modelo por medio del método ANOVA para los datos de densidad de rapidez de filtración, contenido de humedad del ejemplo 12.6. Compruebe que se obtiene un resultado equivalente al de la prueba t.

40. Utilice las reglas de valor esperado para demostrar que $\hat{\beta}_0$ es un estimador insesgado para β_0 (suponiendo que $\hat{\beta}_1$ es insesgado para β_1.

41. a. Verifique que $E(\hat{\beta}_1) = \beta_1$ mediante las reglas del valor esperado del capítulo 5.
b. Utilice las reglas de varianza del capítulo 5 a fin de verificar la expresión para $V(\hat{\beta}_1)$ presentada en esta sección.

42. Verifique que si cada x_i se multiplica por una constante positiva c y cada y_i se multiplica por otra constante positiva d, el estadístico t para probar H_0: $\beta_1 = 0$ contra H_a: $\beta_1 \neq 0$ no cambia de valor (cambia el valor de $\hat{\beta}_1$, lo cual demuestra que la magnitud de $\hat{\beta}_1$ por sí misma no es indicativa de la utilidad del modelo).

43. La probabilidad de un error tipo II para la prueba t de H_0: $\beta_1 = \beta_{10}$ se calcula de la misma manera como se hizo para las pruebas t del capítulo 8. Si el valor alternativo de β_1 se denota mediante β_1', se calcula primero el valor de

$$d = \frac{|\beta_{10} - \beta_1'|}{\sigma\sqrt{\dfrac{n-1}{\sum x_i^2 - (\sum x_i)^2/n}}}$$

después se introduce el conjunto apropiado de curvas de la tabla A.17 del apéndice en el eje horizontal al valor de d, y β se lee de la curva para $n - 2$ grados de libertad. En un artículo del *Journal of Public Health Engineering*, se encuentran los resultados de un análisis de regresión basado en $n = 15$ observaciones en las que x = temperatura de aplicación del filtro (°C) y y = % de eficiencia de eliminación de DBO. Las cantidades calculadas son $\sum x_i = 402$, $\sum x_i^2 = 11{,}098$, $s = 3.725$ y $\hat{\beta}_1 = 1.7035$. Considere probar al nivel 0.01 H_0: $\beta_1 = 1$, que expresa que el incremento esperado en el % de eliminación de DBO es 1 cuando la temperatura de aplicación del filtro se incrementa en 1°C, contra la alternativa H_a: $\beta_1 > 1$. Determine P(error de tipo II) cuando $\beta_1' = 2$, $\sigma = 4$.

12.4 Inferencias relacionadas con $\mu_{Y \cdot x^*}$ y predicción de valores de Y futuros

Sea x^* un valor especificado de la variable independiente x. Una vez calculadas las estimaciones de $\hat{\beta}_0$ y $\hat{\beta}_1$, $\hat{\beta}_0 + \hat{\beta}_1 x^*$ se considera como una estimación puntual de $\mu_{Y \cdot x^*}$ (el valor promedio esperado o verdadero de Y cuando $x = x^*$) o como una predicción del valor de Y que resulta de una sola observación hecha cuando $x = x^*$. La estimación puntual o predicción por sí misma no proporciona información de con cuánta precisión se estimó $\mu_{Y \cdot x^*}$ o se predijo Y. Esto se remedia al obtener un IC para $\mu_{Y \cdot x^*}$ y un intervalo de predicción (IP) para un solo valor de Y.

Antes de obtener los datos muestrales, tanto $\hat{\beta}_0$ como $\hat{\beta}_1$ se someten a variabilidad de muestreo, es decir, ambos son estadísticos cuyos valores variarán de una muestra a otra. Por ejemplo, suponga que $\beta_0 = 50$ y $\beta_1 = 2$. Entonces con una primera muestra de pares (x, y) se podría obtener $\hat{\beta}_0 = 52.35$, $\hat{\beta}_1 = 1.895$, una segunda muestra pudo dar lugar a $\hat{\beta}_0 = 46.52$, $\hat{\beta}_1 = 2.056$, etcétera. Se deduce que $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x^*$ cambia de valor de una muestra a otra, por lo tanto es un estadístico. Si la ordenada al origen y la pendiente de la recta poblacional son los valores antes mencionados 50 y 2, respectivamente, y $x^* = 10$, entonces este estadístico intenta estimar el valor $50 + 2(10) = 70$. La estimación de una primera muestra podría ser $52.35 + 1.895(10) = 71.30$, de una segunda muestra podría ser $46.52 + 2.056(10) = 67.08$, etcétera. De la misma forma que un intervalo de confianza para β_1 se basó en las propiedades de la distribución de muestreo de $\hat{\beta}_1$, un intervalo de confianza para un valor de y medio en la regresión se basa en las propiedades de la distribución de muestreo del estadístico $\hat{\beta}_0 + \hat{\beta}_1 x^*$.

La sustitución de las expresiones para $\hat{\beta}_0$ y $\hat{\beta}_1$ en $\hat{\beta}_0 + \hat{\beta}_1 x^*$ seguida de ciertas operaciones algebraicas conduce a la representación de $\hat{\beta}_0 + \hat{\beta}_1 x^*$ como una función lineal de las Y_i:

$$\hat{\beta}_0 + \hat{\beta}_1 x^* = \sum_{i=1}^{n}\left[\frac{1}{n} + \frac{(x^* - \bar{x})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}\right] Y_i = \sum_{i=1}^{n} d_i Y_i$$

Los coeficientes $d_1, d_2, \ldots, d_n$ en esta función lineal tienen que ver con las x_i y x^*, los cuales son fijos. La aplicación de las reglas de la sección 5.5 a esta función lineal da lugar a las propiedades siguientes.

Sea $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x^*$, donde x^* es algún valor fijo de x.

1. El valor medio de $\hat{Y}$ es

$$E(\hat{Y}) = E(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \mu_{\hat{\beta}_0 + \hat{\beta}_1 x^*} = \beta_0 + \beta_1 x^*$$

Por lo tanto $\hat{\beta}_0 + \hat{\beta}_1 x^*$ es un estimador insesgado para $\beta_0 + \beta_1 x^*$ (es decir, para $\mu_{Y\cdot x^*}$).

2. La varianza de $\hat{Y}$ es

$$V(\hat{Y}) = \sigma_{\hat{Y}}^2 = \sigma^2\left[\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum x_i^2 - (\sum x_i)^2/n}\right] = \sigma^2\left[\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}\right]$$

y la desviación estándar $\sigma_{\hat{Y}}$ es la raíz cuadrada de esta expresión. La desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_1 x^*$, denotada por $s_{\hat{Y}}$ o $s_{\hat{\beta}_0 + \hat{\beta}_1 x^*}$, resulta de reemplazar σ por su estimación s:

$$s_{\hat{Y}} = s_{\hat{\beta}_0 + \hat{\beta}_1 x^*} = s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

3. $\hat{Y}$ tiene una distribución normal.

La varianza de $\hat{\beta}_0 + \hat{\beta}_1 x^*$ es mínima cuando $x^* = \bar{x}$, y se incrementa cuando x^* se aleja de $\bar{x}$ en cualquier dirección. Por lo tanto, el estimador de $\mu_{Y\cdot x^*}$ es más preciso cuando x^* está cerca del centro de las x_i que cuando está lejos de los valores x en los que se hicieron las observaciones. Esto significa que tanto el IC como el IP son más estrechos para una x^* cercana a $\bar{x}$ que para una x^* alejada de $\bar{x}$. La mayor parte del software para estadística proporciona tanto $\hat{\beta}_0 + \hat{\beta}_1 x^*$ como $s_{\hat{\beta}_0 + \hat{\beta}_1 x^*}$ para alguna x^* siempre que se solicite.

Inferencias que tienen que ver con $\mu_{Y\cdot x^*}$

Al igual que los procedimientos inferenciales para β_1 se basaron en la variable t obtenida al estandarizar β_1, una variable t obtenida al estandarizar $\hat{\beta}_0 + \hat{\beta}_1 x^*$ origina un IC y procedimientos de prueba.

TEOREMA

La variable

$$T = \frac{\hat{\beta}_0 + \hat{\beta}_1 x^* - (\beta_0 + \beta_1 x^*)}{S_{\hat{\beta}_0 + \hat{\beta}_1 x^*}} = \frac{\hat{Y} - (\beta_0 + \beta_1 x^*)}{S_{\hat{Y}}} \qquad (12.5)$$

tiene una distribución t con $n - 2$ grados de libertad.

De la misma manera que para β_1 en la sección anterior, un enunciado de probabilidad en el que interviene esta variable estandarizada se opera de manera que produzca un intervalo de confianza para $\mu_{Y\cdot x^*}$.

Un **IC** de $100(1-\alpha)\%$ **para** $\mu_{Y\cdot x^*}$, el valor esperado de Y cuando $x = x^*$, es

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2,n-2} \cdot s_{\hat{\beta}_0+\hat{\beta}_1 x^*} = \hat{y} \pm t_{\alpha/2,n-2} \cdot s_{\hat{Y}} \qquad (12.6)$$

Este IC está centrado en la estimación puntual para $\mu_{Y\cdot x^*}$ y se extiende a cada lado por una cantidad que depende del nivel de confianza y el grado de variabilidad en el estimador en el que se basa la estimación puntual.

Ejemplo 12.13 La corrosión de las barras de acero de refuerzo es el problema de durabilidad más importante para estructuras de concreto reforzado. La carbonatación del concreto resulta de una reacción química que disminuye el valor de pH lo suficiente para iniciar la corrosión de las barras. A continuación se ilustran los datos representativos de x = profundidad de la carbonatación (mm) y y = resistencia (MPa) para una muestra de especímenes base tomados de un determinado edificio (leídos de una gráfica del artículo "The Carbonation of Concrete Structures in the Tropical Environment of Singapore", *Magazine of Concrete Res.*, 1996: 293-300).

x	8.0	15.0	16.5	20.0	20.0	27.5	30.0	30.0	35.0
y	22.8	27.2	23.7	17.1	21.5	18.6	16.1	23.4	13.4

x	38.0	40.0	45.0	50.0	50.0	55.0	55.0	59.0	65.0
y	19.5	12.4	13.2	11.4	10.3	14.1	9.7	12.0	6.8

Un diagrama de dispersión de los datos (véase la figura 12.17) respalda el uso del modelo de regresión lineal simple. Las cantidades importantes son las siguientes:

$$\sum x_i = 659.0 \quad \sum x_i^2 = 28\,967.50 \quad \bar{x} = 36.6111 \quad S_{xx} = 4840.7778$$
$$\sum y_i = 293.2 \quad \sum x_i y_i = 9293.95 \quad \sum y_i^2 = 5335.76$$
$$\hat{\beta}_1 = -0.297561 \quad \hat{\beta}_0 = 27.182936 \quad \text{SSE} = 131.2402$$
$$r^2 = 0.766 \quad s = 2.8640$$

Ahora se calcula un intervalo de confianza, con un nivel de confianza de 95%, para la resistencia media de los especímenes base que tienen una profundidad de carbonatación de 45 mm, es decir, un intervalo de confianza para $\beta_0 + \beta_1(45)$. El intervalo está centrado en

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(45) = 27.18 - 0.2976(45) = 13.79$$

La desviación estándar estimada del estadístico $\hat{Y}$ es

$$s_{\hat{Y}} = 2.8640\sqrt{\frac{1}{18} + \frac{(45 - 36.6111)^2}{4840.7778}} = 0.7582$$

El valor crítico t de 16 gl para un nivel de confianza de 95% es 2.120, a partir del cual se determina el intervalo deseado como

$$13.79 \pm (2.120)(0.7582) = 13.79 \pm 1.61 = (12.18, 15.40)$$

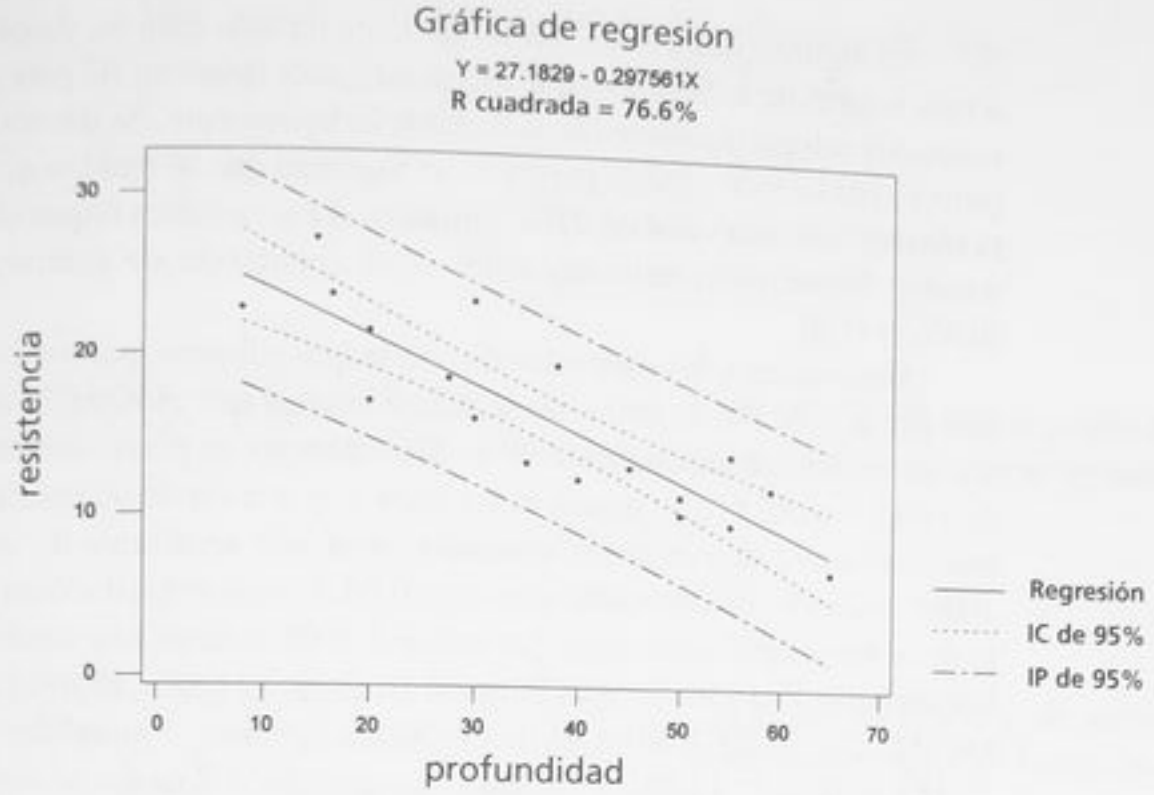


Figura 12.17 Diagrama de dispersión de MINITAB con intervalos de confianza e intervalos de predicción para los datos del ejemplo 12.13

La estrechez de este intervalo indica que se tiene información razonablemente precisa acerca del valor medio que está siendo estimado. Recuerde que si se calcula de nuevo este intervalo para una muestra tras otra, a la larga cerca de 95% de los intervalos calculados incluirían $\beta_0 + \beta_1(45)$. Sólo se puede esperar que este valor medio se ubique en el intervalo simple que se calculó.

En la figura 12.18 se muestra el resultado de MINITAB que se obtiene de solicitar ajustar el modelo de regresión lineal simple y calcular intervalos de confianza para el valor medio de la resistencia a profundidades de 45 y 35 mm. Los intervalos están en la parte inferior del resultado; observe que el segundo intervalo es más estrecho que el primero, porque 35 está mucho más cerca de $\bar{x}$ que 45. En la figura 12.17 se muestran (1) curvas que corresponden a los límites de confianza para cada valor x diferente y (2) límites de predicción, que se analizarán en breve. Observe cómo las curvas se apartan más y más a medida que x se aleja de $\bar{x}$.

```
La ecuación de regresión es resistencia = 27.2 - 0.298 profundidad

Predictor        Coef      Desv. est.  Relación t      P

Constante      27.183          1.651       16.46  0.000
  profundidad -0.29756        0.04116       -7.23  0.000
s = 2.864   R cuadrada = 76.6%   R cuadrada (ajustada) = 75.1%

Análisis de varianza

FUENTE          GL         SC          CM          F          P

Regresión        1     428.62      428.62      52.25      0.000
     Error      16     131.24        8.20
     Total      17     559.86

                                          95.0% I.C.          95.0% I.P.
   Ajuste  Ajuste de desv. est.
13.793                   0.758  (12.185, 15.401)  (7.510, 20.075)

                                          95.0% I.C.          95.0% I.P.
   Ajuste  Ajuste de desv. est.
16.768                   0.678  (15.330, 18.207)  (10.527, 23.009)
```

Figura 12.18 Resultado de regresión de MINITAB para los datos del ejemplo 12.13 ■

En algunas situaciones, un IC se desea no sólo para un simple valor de x, sino para dos o más valores de x. Suponga que un investigador desea un IC para $\mu_{Y\cdot v}$ y para $\mu_{Y\cdot w}$ donde v y w son dos valores diferentes de la variable independiente. Se intenta calcular el intervalo (12.6) primero para $x = v$ y luego para $x = w$. Suponga que se emplea $\alpha = 0.05$ en cada cálculo para obtener dos intervalos de 95%. Entonces si las variables requeridas para calcular los dos intervalos fueran independientes entre sí, el coeficiente de confianza conjunto sería $(0.95)\cdot(0.95) \approx 0.90$.

Sin embargo, los intervalos no son independientes, porque en cada uno se usan las mismas $\hat{\beta}_0$, $\hat{\beta}_1$ y S. Por lo tanto, no se puede afirmar que el nivel de confianza conjunto para los dos intervalos sea exactamente 90%. Sin embargo, se puede demostrar que si se calcula el IC de $100(1 - \alpha)\%$ (12.6) para $x = v$ y para $x = w$ a fin de obtener los IC para $\mu_{Y\cdot v}$ y $\mu_{Y\cdot w}$, entonces *el nivel de confianza conjunto en el par resultante de intervalos es por lo menos* $100(1 - 2\alpha)\%$. En particular, con $\alpha = 0.05$ da como resultado un nivel de confianza conjunto de *al menos* 90%, mientras que con $\alpha = 0.01$ se tiene una confianza de por lo menos 98%. Por ejemplo 12.13 un IC de 95% para $\mu_{Y\cdot 45}$ fue (12.185, 15.401) y un IC de 95% para $\mu_{Y\cdot 35}$ fue (15.330, 18.207). El nivel de confianza conjunto o simultáneo para las dos expresiones $12.185 < \mu_{Y\cdot 45} < 15.401$ y $15.330 < \mu_{Y\cdot 35} < 18.207$ es por lo menos 90%.

La validez de estos dos IC conjuntos o simultáneos se basa en un resultado de probabilidad conocido como **desigualdad de Bonferroni**, así que los intervalos de confianza conjuntos se denominan **intervalos de Bonferroni**. El método se generaliza sin dificultad para producir intervalos conjuntos para k con las $\mu_{Y\cdot x}$ distintas. *Usando por separado el intervalo* (12.6) *primero para* $x = x_1^*$, *luego para* $x = x_2^*, \ldots,$ *y, por último, para* $x = x_k^*$, *se obtiene un conjunto de k intervalos de confianza para los que se garantiza que el nivel de confianza conjunto o simultáneo es por lo menos* $100(1 - k\alpha)\%$.

Las pruebas de hipótesis respecto a $\beta_0 + \beta_1 x^*$ se basan en la estadística T de la prueba obtenida sustituyendo $\beta_0 + \beta_1 x^*$ en el numerador de (12.5) por el valor nulo de μ_0. Por ejemplo, H_0: $\beta_0 + \beta_1(45) = 15$ en el ejemplo 12.13 dice que cuando la profundidad de carbonatación es 45, la resistencia esperada (es decir, promedio real) es 15. Entonces, el valor del estadístico de prueba es $t = [\hat{\beta}_0 + \hat{\beta}_1(45) - 15]/s_{\hat{\beta}_0 + \hat{\beta}_1(45)}$, y la prueba es de extremo superior, inferior o de dos extremos según la desigualdad en H_a.

Intervalo de predicción para un valor futuro de *Y*

De modo análogo al IC (12.6) para $\mu_{Y\cdot x^*}$, con frecuencia se desea obtener un intervalo de valores posibles para el valor de Y relacionado con cierta observación futura cuando la variable independiente tiene valor x^*. Como ilustración, en el ejemplo en el que el tamaño del vocabulario y se relaciona con la edad de un niño, para $x = 6$ años (12.6) proporcionaría un IC para el tamaño de vocabulario promedio real de los niños de seis años de edad. De otra manera, se podría desear un intervalo de valores posibles para el tamaño de vocabulario de un determinado niño de seis años.

Un IC se refiere a un parámetro, o característica poblacional, cuyo valor es fijo pero desconocido. En contraste, un valor futuro de Y no es un parámetro, sino una variable aleatoria; por esta razón se hace referencia a un intervalo de valores posibles para una Y futura como un **intervalo de predicción,** y no como un intervalo de confianza. El error de estimación es $\beta_0 + \beta_1 x^* - (\hat{\beta}_0 + \hat{\beta}_1 x^*)$, una diferencia entre una cantidad fija (pero desconocida) y una variable aleatoria. El error de predicción es $Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)$, una diferencia entre dos variables aleatorias. Por consiguiente, hay más incertidumbre en la predicción que en la estimación, así que un IP será más amplio que un IC. Debido a que el valor futuro de Y es independiente de las Y_i observadas

$$\begin{aligned} V[Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)] &= \text{varianza del error de predicción.} \\ &= V(Y) + V(\hat{\beta}_0 + \hat{\beta}_1 x^*) \\ &= \sigma^2 + \sigma^2\left[\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}\right] \\ &= \sigma^2\left[1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}\right] \end{aligned}$$

Además, como $E(Y) = \beta_0 + \beta_1 x^*$ y $E(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \beta_0 + \beta_1 x^*$, el valor esperado del error de predicción es $E(Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)) = 0$. Se puede demostrar que la variable estandarizada

$$T = \frac{Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)}{S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}}$$

tiene una distribución t con $n - 2$ gl. Al sustituir esta T en la expresión de probabilidad $P(-t_{\alpha/2,n-2} < T < t_{\alpha/2,n-2}) = 1 - \alpha$ y realizar algunas operaciones para aislar Y entre las dos desigualdades se obtiene el siguiente intervalo.

Un **IP** $100(1 - \alpha)\%$ **para una observación Y futura que se realiza cuando** $x = x^*$ es

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2,n-2} \cdot s\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}} \qquad (12.7)$$

$$= \hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2,n-2} \cdot \sqrt{s^2 + s^2_{\hat{\beta}_0 + \hat{\beta}_1 x^*}}$$

$$= \hat{y} \pm t_{\alpha/2,n-2} \cdot \sqrt{s^2 + s^2_{\hat{Y}}}$$

La interpretación del nivel de predicción $100(1 - \alpha)\%$ es idéntica a la de los niveles de confianza previos, si se usa de forma repetida la ecuación (12.7), a la larga los intervalos resultantes en realidad contendrán los valores de y observados $100(1 - \alpha)\%$ del tiempo. Observe que el 1 debajo del símbolo de raíz cuadrada inicial hace al IP (12.7) más amplio que el IC (12.6), aunque ambos intervalos estén centrados en $\hat{\beta}_0 + \hat{\beta}_1 x^*$. También, cuando $n \to \infty$, la extensión del IC tiende a 0, pero no así la amplitud del IP (porque aún conociendo bien β_0 y β_1, habrá incertidumbre en la predicción).

Ejemplo 12.14 Se toman de nuevo los datos de la profundidad de carbonatación-resistencia del ejemplo 12.13 y se calcula un intervalo de predicción para un valor de resistencia que resultaría de seleccionar un solo espécimen base cuya profundidad de carbonatación es 45 mm. Las cantidades importantes de ese ejemplo son

$$\hat{y} = 13.79 \qquad s_{\hat{Y}} = 0.7582 \qquad s = 2.8640$$

Para un nivel de predicción de 95% basado en $n - 2 = 16$ gl, el valor crítico t es 2.120, justo lo que se usó antes para un nivel de confianza de 95%. Entonces, el intervalo de predicción es

$$\begin{aligned} 13.79 \pm (2.120)\sqrt{(2.8640)^2 + (0.7582)^2} &= 13.79 \pm (2.120)(2.963) \\ &= 13.79 \pm 6.28 = (7.51, 20.07) \end{aligned}$$

Los valores posibles para una sola observación de resistencia cuando la profundidad es de 45 mm, están (al nivel de predicción de 95%) entre 7.51 y 20.07 MPa. El intervalo de confianza de 95% para la resistencia media cuando la profundidad es 45 fue (12.18, 15.40). El intervalo de predicción es mucho más amplio que esto, debido al $(2.8640)^2$ extra debajo de la raíz cuadrada. En la figura 12.18, el resultado de MINITAB del ejemplo 12.13, se observa este intervalo, así como el intervalo de confianza. ■

La técnica de Bonferroni se emplea en el caso de intervalos de confianza. Si se calcula un IP de $100(1 - \alpha)\%$ para cada uno de k diferentes valores de x, el nivel de predicción simultáneo o conjunto para los k intervalos es por lo menos $100(1 - k\alpha)\%$.

Ejercicios | Sección 12.4 (44 a 56)

44. Al ajustar el modelo de regresión lineal simple a las $n = 27$ observaciones de x = módulo de elasticidad y y = resistencia a la flexión del ejercicio 15, sección 12.2, se obtuvo $\hat{y} = 7.592$, $s_{\hat{Y}} = 0.179$ cuando $x = 40$ y $\hat{y} = 9.741$, $s_{\hat{Y}} = 0.253$ para $x = 60$.

a. Explique por qué $s_{\hat{Y}}$ es mayor cuando $x = 60$ que cuando $x = 40$.

b. Calcule un intervalo de confianza con nivel de confianza de 95% para la resistencia promedio real de las vigas cuyo módulo de elasticidad es 40.

c. Calcule un intervalo de predicción al nivel 95% para la resistencia de una sola viga cuyo módulo de elasticidad es 40.

d. Si se calcula un intervalo de confianza de 95% para la resistencia promedio real cuando el módulo de elasticidad es 60, ¿cuál será el nivel de confianza simultáneo para este intervalo y el calculado en el inciso (b)?

45. Considere de nuevo los datos de rapidez de filtración-contenido de humedad del ejemplo 12.6 (véase también el ejemplo 12.7).

a. Calcule un IC de 90% para $\beta_0 + 125\beta_1$, el contenido de humedad promedio real cuando la rapidez de filtración es 125.

b. Prediga el valor del contenido de humedad para una sola corrida experimental en la que la tasa de filtración es 125, con un nivel de predicción de 90%. ¿Cómo se compara este intervalo con el del inciso (a)? ¿Por qué se tiene esta situación?

c. ¿Cómo se compararían los intervalos de los incisos (a) y (b) con un IC y un IP cuando la rapidez de filtración es 115? Conteste sin calcular en realidad estos nuevos intervalos.

d. Interprete las hipótesis H_0: $\beta_0 + 125\beta_1 = 80$ y H_0: $\beta_0 + 125\beta_1 < 80$, y luego realice una prueba de significación al nivel 0.01.

46. En el artículo "The Incorporation of Uranium and Silver by Hydrothermally Synthesized Galena" (*Econ. Geology*, 1964: 1003-1024) se describe la determinación del contenido de plata en cristales de galena en un sistema hidrotérmico cerrado sobre un intervalo de temperatura. Con x = temperatura de cristalización en °C y $y = Ag_2S$ en mol %, los datos son los siguientes:

x	398	292	352	575	568	450	550
	408	484	350	503	600	600	
y	0.15	0.05	0.23	0.43	0.23	0.40	0.44
	0.44	0.45	0.09	0.59	0.63	0.60	

de donde $\sum x_i = 6130$, $\sum x_i^2 = 3\,022\,050$, $\sum y_i = 4.73$, $\sum y_i^2 = 2.1785$, $\sum x_i y_i = 2418.74$, $\hat{\beta}_1 = 0.00143$, $\hat{\beta}_0 = -0.311$ y $s = 0.131$.

a. Estime el contenido promedio real de plata cuando la temperatura es 500° C con un intervalo de confianza de 95 por ciento.

b. ¿Cómo se compararía la extensión de un intervalo de confianza de 95% para el contenido promedio real de plata, cuando la temperatura es 400° C, con la extensión del inciso (a)? Responda sin calcular este nuevo intervalo.

c. Calcule un intervalo de confianza de 95% para el cambio promedio real de contenido de plata relacionado con un aumento de temperatura de 1° C.

d. Suponga que antes se creía que cuando la temperatura de cristalización era 400° C, el contenido promedio real de plata sería 0.25. Realice una prueba al nivel de significación 0.05 para decidir si los datos de la muestra contradicen la creencia anterior.

47. El modelo de regresión lineal simple proporciona un ajuste muy bueno para los datos de volumen de lluvia y escurrimiento del ejercicio 16, sección 12.2. La ecuación de la recta de mínimos cuadrados es $\hat{y} = -1.128 + 0.82697x$, $r^2 = 0.975$ y $s = 5.24$.

a. Utilice el hecho de que $s_{\hat{Y}} = 1.44$ cuando el volumen de lluvia es 40 m^3 para pronosticar el escurrimiento de una manera que transmita información acerca de la fiabilidad y la precisión. ¿El intervalo resultante in-

dica que se tiene información precisa acerca del valor de escurrimiento para esta observación futura? Explique su razonamiento.

b. Calcule un intervalo de predicción para el escurrimiento cuando la lluvia es 50 con el mismo intervalo de predicción que en el inciso (a). ¿Qué se puede decir acerca del nivel de predicción simultáneo para los dos intervalos calculados?

48. En un estudio descrito en el artículo "The Effects of Water Vapor Concentration on the Rate of Combustion of an Artificial Graphite in Humid Air Flow" (*Combustion and Flame*, 1983: 107-118) se ilustran los datos de x = temperatura de una mezcla de nitrógeno y oxígeno (en miles de °F) bajo condiciones especificadas, y y = difusividad del oxígeno. Las cantidades de resumen son $n = 9$, $\sum x_i = 12.6$, $\sum y_i = 27.68$, $\sum x_i^2 = 18.24$, $\sum x_i y_i = 40.968$ y $\sum y_i^2 = 93.3448$. Suponga que las dos variables se relacionan según un modelo de regresión lineal simple.

a. Estime la difusividad promedio real del oxígeno cuando la temperatura es 1500°F, de modo que transmita información acerca de la confiabilidad y la precisión.

b. Prediga la difusividad del oxígeno para una sola observación que se realiza cuando la temperatura es de 1500° F, de manera que lleve información acerca de la fiabilidad y la precisión. ¿Cómo se compara la predicción calculada aquí con la estimación calculada en el inciso (a)?

c. ¿Un intervalo de predicción para la difusividad cuando la temperatura es 1200°F, con el mismo nivel de predicción que en el inciso (b), sería más amplio o estrecho? Conteste sin calcular este segundo intervalo.

49. Se sabe que un intervalo de confianza de 95% para el contenido esperado de plomo cuando el flujo de tránsito es de 15, con base en una muestra de $n = 10$ observaciones, es (462.1, 597.7). Calcule un IC con nivel de confianza de 99% para el contenido esperado de plomo cuando el flujo de tránsito es 15.

50. En un experimento para medir el tiempo de relajación magnética macroscópica en cristales (μs) en función de la intensidad del campo magnético de polarización externo (KG) se obtuvieron los siguientes datos ("An Optical Faraday Rotation Technique for the Determination of Magnetic Relaxation Times", *IEEE Trans. Magnetics*, junio de 1968: 175-178. Los datos se leyeron de una gráfica que apareció en el artículo):

x	11.0	12.5	15.2	17.2	19.0	20.8
y	187	225	305	318	367	365
x	22.0	24.2	25.3	27.0	29.0	
y	400	435	450	506	558	

Las estadísticas de resumen son $\sum x_i = 223.2$, $\sum y_i = 4116$, $\sum x_i^2 = 4877.50$, $\sum x_i y_i = 90\,096.1$ y $\sum y_i^2 = 1\,666\,782$. Calcule lo siguiente:

a. Un intervalo de confianza de 95% para el tiempo de relajación esperado, cuando la intensidad del campo es igual a 18.

b. Un intervalo de pronóstico de 95% para tiempo de relajación futuro cuando la intensidad del campo es igual a 18.

c. Intervalos de confianza simultáneos para el tiempo de relajación esperado cuando la intensidad del campo es igual a 15, 18 y 20; el coeficiente de confianza conjunto debe ser por lo menos 97 por ciento.

51. Refiérase al ejemplo 12.12 donde x = fracción de volumen de óxidos/inclusiones y y = % de alargamiento.

a. Con MINITAB se obtuvo $s_{\hat{\beta}_0+\hat{\beta}_1(.40)} = 0.0311$ y $s_{\hat{\beta}_0+\hat{\beta}_1(1.20)} = 0.0352$. ¿Por qué la primera desviación estándar es menor que la segunda?

b. Utilice el resultado de MINITAB del ejemplo para calcular un intervalo de confianza de 95% para el porcentaje de alargamiento esperado cuando la fracción de volumen = 0.40.

c. Utilice el resultado de MINITAB para calcular un intervalo de pronóstico de 95% para un solo valor observado de % de alargamiento cuando la fracción de volumen es = 1.20.

52. El grabado con plasma es esencial para la transferencia de configuraciones de líneas finas en procesos actuales de semiconductores. En el artículo "Ion Beam-Assisted Etching of Aluminum with Chlorine" (*J. Electrochem. Soc.*, 1985: 2010-2012) se encuentran los datos siguientes (leídos de una gráfica) del flujo de cloro (x, en cm³ normales/min) por una boquilla utilizada en el mecanismo de grabado, y la rapidez de grabado (y, en 100A/min).

x	1.5	1.5	2.0	2.5	2.5	3.0	3.5	3.5	4.0
y	23.0	24.5	25.0	30.0	33.5	40.0	40.5	47.0	49.0

Las estadísticas de resumen son $\sum x_i = 24.0$, $\sum y_i = 312.5$, $\sum x_i^2 = 70.50$, $\sum x_i y_i = 902.25$, $\sum y_i^2 = 11\,626.75$, $\hat{\beta}_0 = 6.448718$ y $\beta_1 = 10.602564$.

a. ¿El modelo de regresión lineal simple especifica una relación útil entre el flujo de cloro y la rapidez de grabado?

b. Estime el cambio promedio real en la rapidez de grabado relacionado con un incremento de 1 cm³ normal/min en el flujo con un intervalo de confianza de 95%, e interprete el intervalo.

c. Calcule un IC de 95% para $\mu_{Y\cdot 3.0}$, la rapidez de grabado promedio real cuando el flujo es igual a 3.0. ¿Se estimó con precisión este promedio?

d. Calcule un IP de 95% para una sola observación futura en la rapidez de grabado que se realiza cuando

el flujo es igual a 3.0. ¿Es probable que la predicción sea precisa?

e. ¿El intervalo de confianza de 95% y el intervalo de predicción cuando el flujo es igual a 2.5 serían más amplios o estrechos que los intervalos correspondientes de los incisos (c) y (d)? Conteste sin calcular los intervalos?

f. ¿Recomendaría calcular un IP de 95% para un flujo de 6.0? Explique.

53. Considere los cuatro intervalos siguientes basados en los datos del ejemplo 12.4 (sección 12.2):

a. Un IC de 95% para la porosidad media cuando el peso unitario es 110.

b. Un IP de 95% para la porosidad cuando el peso unitario es 110.

c. Un IC de 95% para la porosidad media cuando el peso unitario es 115.

d. Un IP de 95% para la porosidad cuando el peso unitario es 115.

Sin calcular alguno de los intervalos, ¿qué se puede decir en relación con las extensiones de los intervalos entre sí?

54. La reducción del suministro de agua en ciertas regiones de Estados Unidos, ha creado la necesidad de comprender mejor las relaciones entre factores económicos, como el rendimiento de cosechas y factores hidrológicos y del suelo. En el artículo "Variability of Soil Water Properties and Crop Yield in a Sloped Watershed" (*Water Resources Bull.*, 1988: 281-288) se proporcionan datos del rendimiento del sorgo (y, en g/m-surco) y la distancia de pendiente (x, en m) en una cuenca inclinada. Las observaciones seleccionadas se ilustran en la siguiente tabla.

x	0	10	20	30	45	50	70
y	500	590	410	470	450	480	510

x	80	100	120	140	160	170	190
y	450	360	400	300	410	280	350

a. Construya una gráfica de dispersión. ¿El modelo de regresión lineal simple parece razonable?

b. Realice una prueba de la utilidad del modelo.

c. Estime el rendimiento promedio real cuando la distancia de pendiente es 75, mediante un intervalo de valores posibles.

55. Desde hace mucho tiempo, una de las preocupaciones de agricultores y científicos ha sido que los insectos infestan los cultivos. En el artículo "Cotton Square Damage by the Plant Bug, *Lygus hesperus*, and Abscission Rates" (*J. of Econ. Entom.*, 1988: 1328-1337) se dan datos de x = edad de una planta de algodón (días) y y = % de cuadros dañados. Considere las siguientes $n = 12$ observaciones (obtenidas de una gráfica de dispersión del artículo):

x	9	12	12	15	18	18
y	11	12	23	30	29	52

x	21	21	27	30	30	33
y	41	65	60	72	84	93

a. ¿Por qué no es determinística la relación entre x y y?

b. ¿La gráfica de dispersión indica que el modelo de regresión lineal simple describirá la relación entre las dos variables?

c. Las estadísticas de resumen son $\sum x_i = 246$, $\sum x_i^2 = 5742$, $\sum y_i = 572$, $\sum y_i^2 = 35\,634$ y $\sum x_i y_i = 14\,0222$. Determine la ecuación de la recta de mínimos cuadrados.

d. Prediga el porcentaje de cuadrados dañados cuando la edad es de 20 días mediante un intervalo de valores posibles.

56. Compruebe que $V(\hat{\beta}_0 + \hat{\beta}_1 x)$ se determina mediante la expresión del texto. [*Sugerencia*: $V(\sum d_i Y_i) = \sum d_i^2 \cdot V(Y_i)$.]

12.5 Correlación

Hay muchas situaciones en las que el objetivo de estudiar el comportamiento conjunto de dos variables es ver si están relacionadas, en vez de usar una para predecir el valor de la otra. En esta sección, primero se desarrolla el coeficiente de correlación muestral r como una medida de cuán estrecha es la relación entre las variables x y y en una muestra, y luego se relaciona r con el coeficiente de correlación x definido en el capítulo 5.

Coeficiente de correlación muestral *r*

Dados n pares de observaciones $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, es natural hablar de que x y y tienen una relación positiva si se forman pares de x y y grandes y x y y pequeñas. De manera similar, si las x grandes se relacionan con las y pequeñas y las x pequeñas con las y grandes, entonces se tiene una relación negativa entre las dos variables. Considere la cantidad

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\Big/ n$$

Entonces, si la relación es fuertemente positiva, una x_i arriba de la media $\bar{x}$ tenderá a relacionarse con una y_i arriba de la media $\bar{y}$, de modo que $(x_i - \bar{x})(y_i - \bar{y}) > 0$, y este producto también será positivo si x_i y y_i están abajo de sus respectivas medias. Por consiguiente, una relación positiva implica que S_{xy} será positiva. Un argumento análogo muestra que cuando la relación es negativa, S_{xy} será negativa, puesto que la mayor parte de los productos $(x_i - \bar{x})(y_i - \bar{y})$ serán negativos. Esto se ilustra en la figura 12.19.

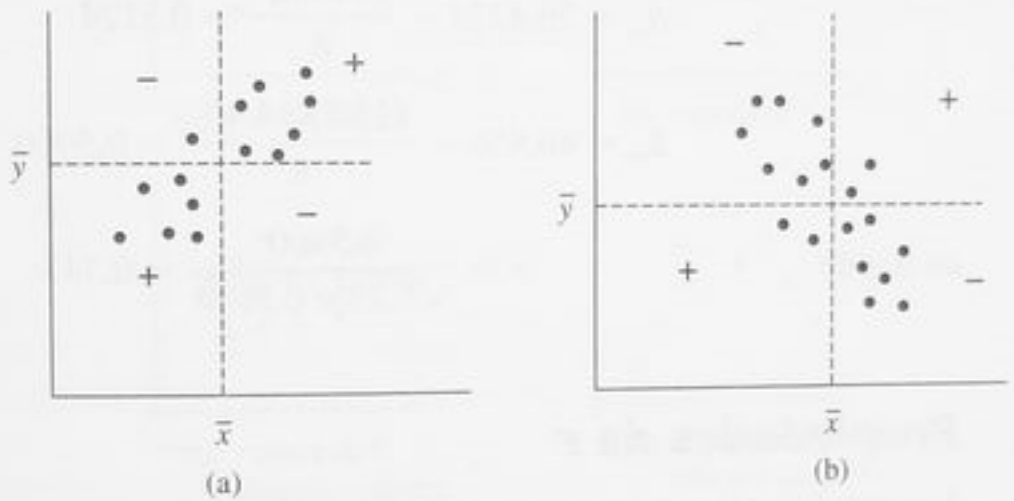


Figura 12.19 (a) Gráfica de dispersión con S_{xy} positiva; (b) gráfica de dispersión con S_{xy} negativa [+ significa $(x_i - \bar{x})(y_i - \bar{y}) > 0$ y − significa $(x_i - \bar{x})(y_i - \bar{y}) < 0$]

Aunque S_{xy} parece una medida plausible de la fuerza de una relación, aún no se tiene idea de cuán positiva o negativa puede ser. Infortunadamente, S_{xy} tiene un defecto grave: al cambiar las unidades de medición de x o y, se puede hacer que la magnitud de S_{xy} sea tan grande o cercana a cero como se quiera. Por ejemplo, si $S_{xy} = 25$ cuando x se mide en metros, entonces $S_{xy} = 25\,000$ cuando x se mide en milímetros y 0.025 cuando x se expresa en kilómetros. Una condición razonable a imponer en alguna medida de cuán fuerte es la relación entre x y y es que la medida calculada no debe depender de las medidas particulares utilizadas para medirla. Esta condición se logra al modificar S_{xy} para obtener el coeficiente de correlación muestral.

DEFINICIÓN

El **coeficiente de correlación muestral** para los n pares $(x_1, y_1), \ldots, (x_n, y_n)$ es

$$r = \frac{S_{xy}}{\sqrt{\sum (x_i - \bar{x})^2}\sqrt{\sum (y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}} \qquad (12.8)$$

Ejemplo 12.15 Una evaluación precisa de la productividad del suelo es crítica para la planificación racional del uso de la tierra. Infortunadamente, como argumenta el autor del artículo "Productivity Ratings Based on Soil Series" (*Prof. Geographer,* 1980: 158-163), no es tan fácil obtener un índice aceptable de productividad del suelo. Una dificultad es que la productividad se determina en parte por el cultivo que se plantó y es posible que la relación entre el rendimiento de dos cultivos plantados en el mismo suelo no sea muy fuerte. Como ilustración, el artículo presenta los datos siguientes del rendimiento x de maíz y el rendimiento y de maní (mT/Ha) para ochos tipos de suelo.

x	2.4	3.4	4.6	3.7	2.2	3.3	4.0	2.1
y	1.33	2.12	1.80	1.65	2.00	1.76	2.11	1.63

Con $\sum x_i = 25.7$, $\sum y_i = 14.40$, $\sum x_i^2 = 88.31$, $\sum x_i y_i = 46.856$ y $\sum y_i^2 = 26.4324$,

$$S_{xx} = 88.31 - \frac{(25.7)^2}{8} = 88.31 - 82.56 = 5.75$$

$$S_{yy} = 26.4324 - \frac{(14.40)^2}{8} = 0.5124$$

$$S_{xy} = 46.856 - \frac{(25.7)(14.40)}{8} = 0.5960$$

de donde

$$r = \frac{0.5960}{\sqrt{5.75}\sqrt{0.5124}} = 0.347$$

■

Propiedades de *r*

Las propiedades más importantes de r son las siguientes:

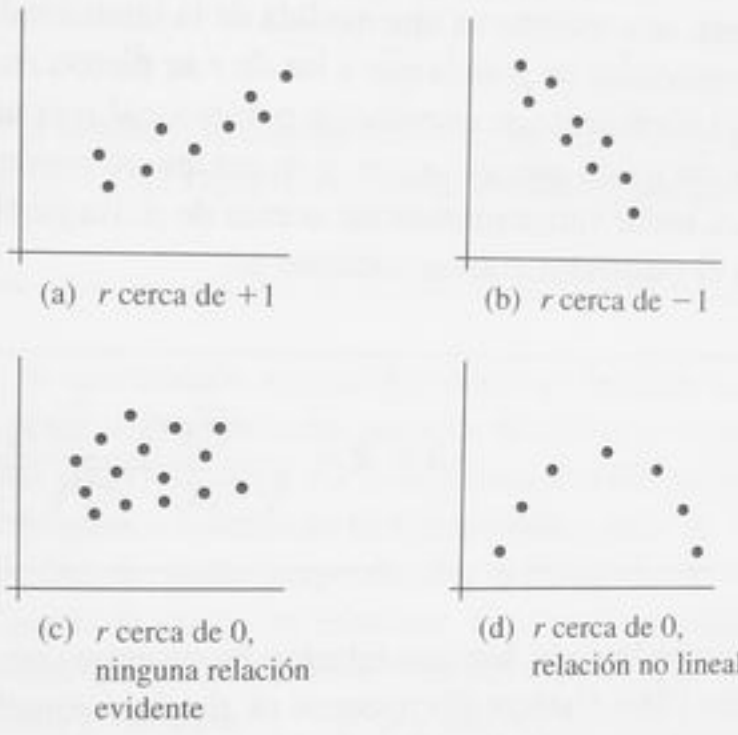
1. El valor de r no depende de cuál de las dos variables bajo estudio se marque con x y cuál con y.
2. El valor de r es independiente de las unidades en las que se midan x y y.
3. $-1 \le r \le 1$
4. $r = 1$ si y sólo si los pares (x_i, y_i) están en una recta con pendiente positiva, y $r = -1$ si y sólo si los pares (x_i, y_i) se localizan en una recta con pendiente negativa.
5. El cuadrado del coeficiente de correlación muestral da el valor del coeficiente de determinación que resultaría de ajustar el modelo de regresión lineal simple, en símbolos, $(r)^2 = r^2$.

La propiedad 1 contradice lo que sucede en el análisis de regresión, donde casi todas las cantidades de interés (la pendiente estimada, la ordenada al origen y estimada, s^2, etc.) dependen de cuál de las dos variables se trata como la variable dependiente. Sin embargo, la propiedad 5 muestra que la proporción de variación en la variable dependiente explicada al ajustar el modelo de regresión lineal simple no depende de cuál variable desempeñe este papel.

La propiedad 2 es equivalente a decir que r permanece sin cambio si cada x_i se reemplaza por cx_i, y si cada y_i se reemplaza por dy_i (un cambio en la escala de medición), así como si

cada x_i se reemplaza por $x_i - a$ y y_i por $y_i - b$ (lo cual cambia la ubicación del cero en el eje de medición). Esto implica, por ejemplo, que r es la misma si la temperatura se mide en °F o °C.

La propiedad 3 dice que el valor máximo de r, que corresponde al máximo grado posible de relación positiva, es $r = 1$, mientras que la relación más negativa se identifica con $r = -1$. De acuerdo con la propiedad 4, las correlaciones máximas positiva y negativa se logran sólo cuando los puntos están a lo largo de una recta. Cualquier otra configuración de puntos, incluso si la configuración indica una relación determinística entre variables, producirá un valor r menor que 1 en magnitud absoluta. Así que *r mide el grado de relación lineal* entre variables. Un valor de r cercano a 0 no es evidencia de la falta de una relación fuerte, sino sólo de la ausencia de una relación lineal, así que este valor de r se debe interpretar con precaución. En la figura 12.20 se ilustran varias configuraciones de puntos relacionados con valores de r distintos.



(a) r cerca de +1

(b) r cerca de −1

(c) r cerca de 0, ninguna relación evidente

(d) r cerca de 0, relación no lineal

Figura 12.20 Gráficas de datos para distintos valores de r

Una pregunta que se hace con frecuencia es, "¿cuándo se puede decir que hay una correlación fuerte entre las variables, y cuándo la correlación es débil?" Una regla empírica razonable es decir que la correlación es débil si $0 \leq |r| \leq 0.5$, fuerte si $0.8 \leq |r| \leq 1$ y moderada en caso contrario. Probablemente le sorprendería saber que $r = 0.5$ es considerada débil, pero $r^2 = 0.25$ significa que en una regresión de y sobre x, el modelo sólo explicaría 25% de la variación y observada. En el ejemplo 12.15, la correlación entre el rendimiento de maíz y el de maní se describe como débil.

Coeficiente de correlación poblacional ρ e inferencias acerca de la correlación

El coeficiente de correlación r es una medida de cuán fuerte es la relación entre x y y en la muestra observada. Se puede considerar que los pares (x_i, y_i) se extrajeron de una población de

pares bivariada, donde (X_i, Y_i) tienen una distribución de probabilidad conjunta $f(x, y)$. En el capítulo 5, se definió el coeficiente de correlación $\rho(X, Y)$ mediante la ecuación

$$\rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

donde

$$\text{Cov}(X, Y) = \begin{cases} \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y) & (X, Y) \text{ discreto} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y)\, dx\, dy & (X, Y) \text{ continuo} \end{cases}$$

Si se considera que $f(x, y)$ describen la distribución de pares de valores dentro de la población completa, se convierte en una medida de la intensidad de la relación de x y y en esa población. Las propiedades de ρ análogas a las de r se dieron en el capítulo 5.

El coeficiente de correlación poblacional ρ es un parámetro o característica poblacional, del mismo modo que μ_X, μ_Y, σ_X y σ_Y así que se puede usar el coeficiente de correlación muestral para hacer varias inferencias acerca de ρ. En particular, si r es una estimación puntual para ρ, y el estimador correspondiente es

$$\hat{\rho} = R = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum (X_i - \overline{X})^2} \sqrt{\sum (Y_i - \overline{Y})^2}}$$

Ejemplo 12.16 En algunos lugares, hay una relación fuerte entre concentraciones de dos contaminantes. En el artículo "The Carbon Component of the Los Angeles Aerosol: Source Apportionment and Contributions to the Visibility Budget" (*J. Air Pollution Control Fed.*, 1984: 643-650) se describen los datos acerca de la concentración de ozono x (ppm) y la concentración secundaria de carbono y (μg/m^3).

x	0.066	0.088	0.120	0.050	0.162	0.186	0.057	0.100
y	4.6	11.6	9.5	6.3	13.8	15.4	2.5	11.8
x	0.112	0.055	0.154	0.074	0.111	0.140	0.071	0.110
y	8.0	7.0	20.6	16.6	9.2	17.9	2.8	13.0

Las cantidades de resumen son $n = 16$, $\sum x_i = 1.656$, $\sum y_i = 170.6$, $\sum x_i^2 = 0.196912$, $\sum x_i y_i = 20.0397$ y $\sum y_i^2 = 2253.56$, de donde

$$r = \frac{20.0397 - (1.656)(170.6)/16}{\sqrt{0.196912 - (1.656)^2/16}\sqrt{2253.56 - (170.6)^2/16}}$$

$$= \frac{2.3826}{(0.1597)(20.8456)} = 0.716$$

La estimación puntual del coeficiente de correlación poblacional ρ entre la concentración de ozono y la concentración secundaria de carbono es $\hat{\rho} = r = 0.716$. ■

Los intervalos de muestra pequeña y los procedimientos de prueba presentados en los capítulos 7 a 9 se basaron en la suposición de normalidad poblacional. Para probar hipótesis acerca de ρ, se debe hacer una suposición análoga acerca de la distribución de pares de valores (x, y) en la población. Ahora se está suponiendo que *tanto* X como Y son aleatorios, mientras que mucho del trabajo de regresión se centró en un valor de x fijado por el experimentador.

SUPOSICIÓN

La distribución de probabilidad conjunta de (X, Y) se especifica mediante

$$f(x, y) = \frac{1}{2\pi \cdot \sigma_1 \sigma_2 \sqrt{1-\rho^2}} e^{-[((x-\mu_1)/\sigma_1)^2 - 2\rho(x-\mu_1)(y-\mu_2)/\sigma_1\sigma_2 + ((y-\mu_2)/\sigma_2)^2]/[2(1-\rho^2)]}$$

$$-\infty < x < \infty$$

$$-\infty < y < \infty \qquad (12.9)$$

donde μ_1 y σ_1 son la media y la desviación estándar de X y μ_2 y σ_2 son la media y la desviación estándar de Y; $f(x, y)$ se llama **distribución de probabilidad normal bivariada.**

Es evidente que la distribución normal bivariada es bastante complicada, pero para los fines que aquí se persiguen, sólo se necesita una leve noción con varias de sus propiedades. La superficie que determina $f(x, y)$ se ubica por completo arriba del plano x, y [$f(x, y) \geq 0$] y tiene una apariencia de campana o montículo tridimensional, como se ilustra en la figura 12.21. Si se corta la superficie con un plano perpendicular al plano x, y y se observa la forma de la curva dibujada en el "plano de corte", el resultado es una curva normal. Para ser precisos, si $X = x$, se puede demostrar que la distribución (condicional) de Y es normal con media $\mu_{Y \cdot x} = \mu_2 - \rho\mu_1\sigma_2/\sigma_1 + \rho\sigma_2 x/\sigma_1$ y varianza $(1 - \rho^2)\sigma_2^2$. Éste es el modelo empleado en la regresión lineal simple con $\beta_0 = \mu_2 - \rho\mu_1\sigma_2/\sigma_1$, $\beta_1 = \rho\sigma_2/\sigma_1$, y $\sigma^2 = (1 - \rho^2)\sigma_2^2$ independiente de x. La implicación es que *si los pares observados* (x_i, y_i) *se extraen en realidad de una distribución normal bivariada, entonces el modelo de regresión lineal simple es una forma apropiada de estudiar el comportamiento de Y para x fija.* Si $\rho = 0$, entonces $\mu_{Y \cdot x} = \mu_2$ independiente de x; de hecho, cuando $\rho = 0$ la función de densidad de probabilidad conjunta $f(x, y)$ de (12.9) se puede factorizar en una parte relacionada sólo con x y en una parte que tiene que ver sólo con y, lo que significa que X y Y son variables independientes.

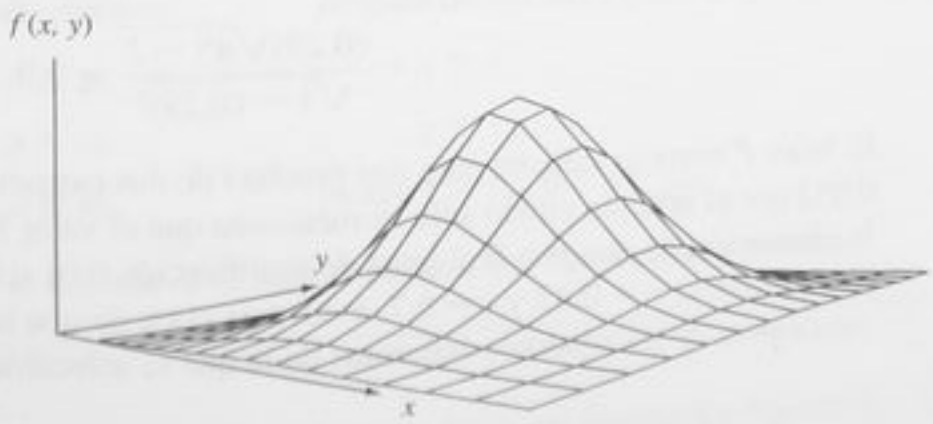


Figura 12.21 Gráfica de una fdp normal bivariada

Suponer que los pares se extraen de una distribución normal bivariada, permite probar hipótesis acerca de ρ y construir un IC. No existe una forma completamente satisfactoria de comprobar la posibilidad de la suposición de normalidad bivariada. Una comprobación parcial requiere construir dos gráficas de probabilidad normal separadas, una para las x_i y otra para las y_i puesto que la normalidad bivariada implica que las distribuciones marginales de X y Y son normales. Si alguna de las gráficas se desvía en forma importante del patrón de una recta, no se deben usar los siguientes patrones inferenciales cuando el tamaño de muestra n es pequeño.

Prueba para la ausencia de correlación

Cuando H_0: $\rho = 0$ es verdadera, el estadístico de prueba

$$T = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}}$$

tiene una distribución t con $n - 2$ gl.

Hipótesis alternativa	**Región de rechazo para la prueba de nivel α**
H_a: $\rho > 0$	$t \geq t_{\alpha,n-2}$
H_a: $\rho < 0$	$t \leq -t_{\alpha,n-2}$
H_a: $\rho \neq 0$	ya sea $t \geq t_{\alpha/2,n-2}$ o $t \leq -t_{\alpha/2,n-2}$

Un valor P basado en $n - 2$ gl se calcula como se describió antes,

Ejemplo 12.17 Los efectos neurotóxicos del manganeso son bien conocidos y, por lo general, son causados por la alta exposición ocupacional durante largos periodos. En las áreas de higiene ocupacional e higiene ambiental, no ha sido descrita antes la relación entre la peroxidación lipídica, a la que se debe el deterioro de los alimentos y el daño al tejido vivo, y la exposición ocupacional. En el artículo "Lipid Peroxidation in Workers Exposed to Manganese" (*Scand. J. Work and Environ. Health*, 1996: 381-386) se muestran datos de x = concentración de manganeso en la sangre (ppb) y y = concentración (μmol/L) de malondialdehído, que es un producto estable de la peroxidación lipídica, para una muestra de 22 trabajadores expuestos al manganeso y para una muestra de control de 45 individuos. El valor de r para la muestra de control fue 0.29, de donde

$$t = \frac{(0.29)\sqrt{45-2}}{\sqrt{1-(0.29)^2}} \approx 2.0$$

El valor P correspondiente para una prueba t de dos extremos basada en 43 gl es más o menos 0.052 (en el artículo citado sólo se menciona que el valor $P > 0.05$). No se desearía rechazar la afirmación de que $\rho = 0$ al nivel de significación 0.01 o 0.05. Para la muestra de trabajadores expuestos, $r = 0.83$ y $t \approx 6.7$, evidencia clara de que hay una relación lineal en la población entera de trabajadores expuestos de la que se seleccionó la muestra. ■

Debido a que ρ mide el grado al cual hay una relación lineal entre dos variables en la población, la hipótesis nula H_0: $\rho = 0$ expresa que no existe tal relación poblacional. En la sec-

ción 12.3, se usó la relación $t\ \hat{\beta}_1/s_{\hat{\beta}_1}$ para probar una relación lineal entre las dos variables en el contexto de análisis de regresión. Resulta que los dos procedimientos de prueba son equivalentes en su totalidad porque $r\sqrt{n-2}/\sqrt{1-r^2} = \hat{\beta}_1/s_{\hat{\beta}_1}$. Cuando sólo interesa evaluar la fuerza de alguna relación lineal en vez de ajustar un modelo y usarlo para estimar o predecir, la fórmula del estadístico de prueba recién presentada requiere menos cálculos que la relación t.

Otras inferencias relacionadas con ρ

El procedimiento para probar H_0: $\rho = \rho_0$ cuando $\rho_0 \neq 0$ no es equivalente a algún procedimiento de análisis de regresión. El estadístico de prueba se basa en una transformación de R llamada transformación de Fisher.

PROPOSICIÓN

Cuando $(X_1, Y_1), \ldots, (X_n, Y_n)$ es una muestra de una distribución normal bivariada, la *va*

$$V = \frac{1}{2}\ln\left(\frac{1+R}{1-R}\right) \tag{12.10}$$

tiene aproximadamente una distribución normal con media y varianza

$$\mu_V = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right) \qquad \sigma_V^2 = \frac{1}{n-3}$$

La razón fundamental para la transformación es obtener una función de R que tiene varianza independiente de ρ; éste no sería el caso con R. También, la transformación no se debe usar si n es bastante pequeña, puesto que la aproximación no será válida.

El estadístico de prueba para probar H_0: $\rho = \rho_0$ es

$$Z = \frac{V - \frac{1}{2}\ln[(1+\rho_0)/(1-\rho_0)]}{1/\sqrt{n-3}}$$

Hipótesis alternativa	**Región de rechazo para la prueba de nivel α**
H_a: $\rho > \rho_0$	$z \geq z_\alpha$
H_a: $\rho < \rho_0$	$z \leq -z_\alpha$
H_a: $\rho \neq \rho_0$	ya sea $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$

Un valor P se calcula de la misma manera que las pruebas z previas.

Ejemplo 12.18 En el artículo "A Study of a Partial Nutrient Removal System for Wastewater Treatment Plants" (*Water Research*, 1972: 1389-1397) se describe un método de eliminación de nitrógeno que requiere tratar las aguas madre de un digestor aeróbico. Tanto el nitrógeno total del afluente x

(mg/L) como el porcentaje y de nitrógeno eliminado se registraron durante 20 días, con estadísticos de resumen $\sum x_i = 285.90$, $\sum x_i^2 = 4409.55$, $\sum y_i = 690.30$, $\sum y_i^2 = 29\,040.29$ y $\sum x_i y_i = 10\,818.56$. ¿Los datos indican que el nitrógeno total del afluente y el porcentaje de nitrógeno eliminado tienen por lo menos una correlación positiva moderada?

La interpretación previa de la correlación positiva moderada fue $0.5 < \rho < 0.8$, por lo tanto se desea probar H_0: $\rho = 0.5$ contra H_a: $\rho > 0.5$. El valor calculado de r es 0.733, así que

$$\frac{1}{2}\ln\left(\frac{1+0.733}{1-0.733}\right) = 0.935 \qquad \frac{1}{2}\ln\left(\frac{1+0.5}{1-0.5}\right) = 0.549$$

Esto da $z = (0.935 - 0.549)\sqrt{17} = 1.59$. Puesto que $1.59 < 1.645$, al nivel 0.05 no se puede concluir que $\rho > 0.5$, por lo que no se ha demostrado que la relación sea incluso moderadamente fuerte (una conclusión un poco sorprendente puesto que $r = 0.73$, pero cuando n es pequeña podría resultar una r grande incluso cuando ρ es pequeña). ■

A fin de obtener un IC para ρ, primero se obtiene un intervalo para $\mu_V = \frac{1}{2}\ln[(1+\rho)/(1-\rho)]$. Al estandarizar V, escribir un enunciado de probabilidad y se realizan operaciones con las desigualdades resultantes, se obtiene

$$\left(v - \frac{z_{\alpha/2}}{\sqrt{n-3}},\ v + \frac{z_{\alpha/2}}{\sqrt{n-3}}\right) \tag{12.11}$$

como un intervalo de $100(1-\alpha)\%$ para μ_V, donde $v = \frac{1}{2}\ln[(1+r)/(1-r)]$. A partir de este intervalo se obtiene un IC para ρ.

> Un intervalo de confianza de $100(1-\alpha)\%$ para ρ es
>
> $$\left(\frac{e^{2c_1}-1}{e^{2c_1}+1},\ \frac{e^{2c_2}-1}{e^{2c_2}+1}\right)$$
>
> donde c_1 y c_2 son los puntos finales izquierdo y derecho, respectivamente, del intervalo (12.11).

Ejemplo 12.19 (Continuación del ejemplo 12.18) El coeficiente de correlación muestral entre el nitrógeno del afluente y el porcentaje de nitrógeno eliminado fue $r = 0.733$, con lo que se obtiene $v = 0.935$. Con $n = 20$, un intervalo de confianza de 95% para μ_V es $(0.935 - 1.96/\sqrt{17}, 0.935 + 1.96/\sqrt{17}) = (0.460, 1.410) = (c_1, c_2)$. El intervalo de 95% para ρ es

$$\left[\frac{e^{2(0.46)}-1}{e^{2(0.46)}+1},\ \frac{e^{2(1.41)}-1}{e^{2(1.41)}+1}\right] = (0.43, 0.89)$$ ■

En el capítulo 5, se advirtió que un valor grande del coeficiente de correlación (cerca de 1 o −1) sólo implica relación y no causa). Esto se aplica tanto a ρ como a r.

Ejercicios | Sección 12.5 (57 a 67)

57. En el artículo "Behavioral Effects of Mobile Telephone Use During Simulated Driving" (*Ergonomics*, 1995: 2536–2562) se describe que para una muestra de 20 sujetos experimentales, el coeficiente de correlación muestral para x = edad y y = tiempo desde que el individuo adquirió una licencia de manejo (años) fue 0.97. ¿Por qué considera que el valor de r es tan cercano a 1? (Los autores del artículo dieron una explicación.)

58. La prueba de oxidación de aceite de turbina (*Turbine Oil Oxidation Test*, TOST) y la prueba de oxidación de bomba rotatoria (*Rotating Bomb Oxidation Test*, RBOT) son dos procedimientos para evaluar la estabilidad de oxidación de los aceites para turbina de vapor. En el artículo "Dependence of Oxidation Stability of Steam Turbine Oil on Base Oil Composition" (*J. of the Society of Tribologists and Lubrication Engrs.*, octubre de 1997: 19-24) se muestran las siguientes observaciones de x = tiempo de TOST (h) y y = tiempo de RBOT (min) para 12 especímenes de aceite.

TOST	4200	3600	3750	3675	4050	2770
RBOT	370	340	375	310	350	200
TOST	4870	4500	3450	2700	3750	3300
RBOT	400	375	285	225	345	285

a. Calcule e interprete el valor del coeficiente de correlación muestral (como lo hicieron los autores del artículo).
b. ¿Cómo sería afectado el valor de r si x = tiempo de RBOT y y = tiempo de TOST?
c. ¿De qué manera se afectaría el valor de r si el tiempo de RBOT se expresara en horas?
d. Trace gráficas de probabilidad normal y comente acerca de características importantes.
e. Realice una prueba de hipótesis para decidir si el tiempo de RBOT y el tiempo de TOST están correlacionados linealmente.

59. La consistencia y lo fibroso del espárrago son de los principales determinantes de su calidad. En esto se centró un estudio descrito en "Post-Harvest Glyphosphate Application Reduces Toughening, Fiber Content, and Lignification of Stored Asparagus Spears" (*J. of the Amer. Soc. of Horticultural Science*, 1988: 569-572). En el artículo se encuentran los siguientes datos (leídos de una gráfica) de x = fuerza cortante (kg) y y = porcentaje de peso de fibra seca.

x	46	48	55	57	60	72	81	85	94
y	2.18	2.10	2.13	2.28	2.34	2.53	2.28	2.62	2.63

x	109	121	132	137	148	149	184	185	187
y	2.50	2.66	2.79	2.80	3.01	2.98	3.34	3.49	3.26

$n = 18 \quad \sum x_i = 1950 \quad \sum x_i^2 = 251\,970$
$\sum y_i = 47.92 \quad \sum y_i^2 = 130.6074 \quad \sum x_i y_i = 5530.92$

a. Calcule el valor del coeficiente de correlación muestral. Con base en este valor, ¿cómo describir a la naturaleza de la relación entre las dos variables?
b. Si un primer espécimen tiene mayor valor de fuerza cortante que un segundo espécimen, ¿qué tiende a cumplirse del porcentaje de peso de fibra seca para los dos especímenes?
c. Si la fuerza de corte se expresa en libras, ¿qué sucede con el valor de r? ¿Por qué?
d. Si el modelo de regresión lineal simple se ajustara a estos datos, ¿qué proporción de variación observada en el porcentaje de peso de fibra seca se podría explicar con la relación del modelo?
e. Realice una prueba al nivel de significación 0.01 para decidir si hay una relación lineal positiva entre las dos variables.

60. En el artículo "A Dual-Buffer Titration Method for Lime Requirement of Acid Mine-soils" (*J. of Environ. Qual.*, 1988: 452-456) se dan los resultados de un estudio en relación con la revegetación del suelo en sitios de recuperación de minas. Con x = aluminio extraíble con KCl y y = cantidad de cal necesaria para llevar el pH del suelo a 7.0, los datos del artículo produjeron las siguientes estadísticas de resumen: $n = 24$, $\sum x = 48.15$, $\sum x^2 = 155.4685$, $\sum y = 263.5$, $\sum y^2 = 3750.53$ y $\sum xy = 658.455$. Realice una prueba al nivel de significación 0.01 para ver si el coeficiente de correlación poblacional es diferente de 0.

61. Los autores del artículo "Objective Effects of a Six Months' Endurance and Strength Training Program in Outpatients with Congestive Hearth Failure" (*Medicine and Science in Sports and Exercise*, 1999: 1102-1107) presentaron un análisis de correlación para investigar la relación entre la concentración máxima de lactato x y la resistencia muscular y. Los datos siguientes se leyeron de una gráfica del artículo.

x	400	750	770	800	850	1025	1200
y	3.80	4.00	4.90	5.20	4.00	3.50	6.30
x	1250	1300	1400	1475	1480	1505	2200
y	6.88	7.55	4.95	7.80	4.45	6.60	8.90

$S_{xx} = 36.9839, \quad S_{yy} = 2\,628\,930.357, \quad S_{xy} = 7377.704.$

Un diagrama de dispersión muestra un patrón lineal.
a. Pruebe ver si hay una correlación positiva entre la concentración máxima de lactato y la resistencia muscular en la población de la que se seleccionaron estos datos.
b. Si se realizara un análisis de regresión para predecir la resistencia a partir de la concentración de lactato, ¿qué proporción de la variación observada en la resistencia se podría atribuir a la relación lineal aproximada? Conteste la pregunta análoga si se usa la regresión para predecir la concentración de lactato a

partir de la resistencia, y conteste ambas preguntas sin hacer ningún cálculo de regresión.

62. Se supone que el contenido de hidrógeno es un importante factor en la porosidad de coladas de aleaciones de aluminio. En el artículo "The Reduced Pressure Test as a Measuring Tool in the Evaluation of Porosity/Hidrogen Content in Al-7 Wt Pct Si-10 Vol Pct SiC(p) Metal Matrix Composite" (*Metallurgical Trans.*, 1993: 1857-1868), se presentan los siguientes datos sobre x = contenido y y = porosidad de gas para una técnica de medición en particular.

x	0.18	0.20	0.21	0.21	0.21	0.22	0.23
y	0.46	0.70	0.41	0.45	0.55	0.44	0.24
x	0.23	0.24	0.24	0.25	0.28	0.30	0.37
y	0.47	0.22	0.80	0.88	0.70	0.72	0.75

Con MINITAB se obtiene el siguiente resultado en respuesta a un comando CORRELATION:

```
Correlación de contenido de hidrógeno
y porosidad = 0.449
```

a. Pruebe al nivel 0.05 para ver si el coeficiente de correlación poblacional difiere de 0.

b. Si se hubiera realizado un análisis de regresión lineal simple, ¿qué porcentaje de la variación observada en la porosidad se podría atribuir a la relación del modelo?

63. Las propiedades físicas de seis muestras de tela retardante de flama se investigaron en el artículo "Sensory and Physical Properties of Inherently Flame Retardant Fabrics" (*Textile Research*, 1984: 61-68). Utilice los datos siguientes y un nivel de significación 0.05 para determinar si existe relación lineal entre rigidez x (mg-cm) y espesor y (mm). ¿Es sorprendente el resultado de la prueba en vista del valor de r?

x	7.98	24.52	12.47	6.92	24.11	35.71
y	0.28	0.65	0.32	0.27	0.81	0.57

64. En el artículo "Increases in Steroid Binding Globulins Induced by Tamoxifen in Patients with Carcinoma of the Breast" (*J. Endocrinology*, 1978: 219-226) se proporcionan datos de los efectos del fármaco Tamoxifén en el cambio en la concentración de globulina de enlace a Cortisol (*CBG*) de pacientes durante un tratamiento. Con edad = x y $\Delta CBG = y$, los valores de resumen son $n = 26$, $\sum x_i = 1613$, $\sum (x_i - \bar{x})^2 = 3756.96$, $\sum y_i = 281.9$, $\sum (y_i - \bar{y})^2 = 465.34$ y $\sum x_i y_i = 16{,}731$.

a. Calcule un intervalo de confianza de 90% para el coeficiente de correlación verdadero ρ.

b. Pruebe H_0: $\rho = -0.5$ contra H_a: $\rho < -0.5$ al nivel 0.05.

c. En un análisis de regresión de y en x, ¿qué proporción de la variación en el cambio de concentración de globulina de enlace a cortisol se podría explicar por la variación en la edad del paciente dentro de la muestra?

d. Si decide realizar un análisis de regresión con la edad como variable dependiente, ¿qué proporción de la variación en la edad se explica por la variación en ΔCBG?

65. En el artículo "Chronological Trend in Blood Lead Levels" (*N. Engl. J. Med.*, 1983: 1373-1377) se muestran los siguientes datos sobre y = concentración promedio de plomo en la sangre de niños de raza blanca de 6 meses a 5 años de edad, y x = cantidad de plomo utilizado en la producción de gasolina (en miles de toneladas) durante diez periodos de seis meses:

x	48	59	79	80	95
y	9.3	11.0	12.8	14.1	13.6
x	95	97	102	102	107
y	13.8	14.6	14.6	16.0	18.2

a. Trace las gráficas de probabilidad normal separadas para x y y. ¿Considera que es razonable suponer que los pares (x, y) son de una población normal bivariada?

b. ¿Los datos proporcionan evidencia suficiente para concluir que hay una relación lineal entre la concentración de plomo en la sangre y la cantidad de plomo que se emplea en la producción de gasolina? Utilice $\alpha = 0.01$.

66. Considere una serie de tiempo, es decir, una secuencia de observaciones $X_1, X_2, \ldots$ obtenida con el tiempo, con valores observados $x_1, x_2, \ldots, x_n$. Suponga que la serie no muestra tendencia ascendente o descendente con el tiempo. Un investigador desea saber cuán fuerte es la relación entre los valores de la serie separados por un número específico de unidades de tiempo. El *coeficiente de autocorrelación muestral r_1 atrasado una unidad es el valor del coeficiente de correlación muestral r para los pares* $(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)$, es decir, pares de valores separados por una unidad de tiempo. De manera similar, el *coeficiente de autocorrelación muestral r_2 atrasado dos unidades es r* para los $n - 2$ pares (x_1, x_3), $(x_2, x_4), \ldots, (x_{n-2}, x_n)$.

a. Calcule los valores de r_1, r_2 y r_3 para los datos de temperatura del ejercicio 82 del capítulo 1, y comente.

b. Análogo al coeficiente de correlación poblacional ρ, sean $\rho_1, \rho_2, \ldots$ indican los coeficientes de autoco-

rrelación teóricos o de largo plazo en los distintos retrasos. Si todos estos ρ son 0, no hay relación (lineal) en ningún retraso. En este caso, si n es grande, cada R_i tiene más o menos una distribución normal con media 0 y desviación estándar $1/\sqrt{n}$ y las distintas R_i son casi independientes. Por lo tanto, H_0: $\rho_i = 0$ se puede rechazar a un nivel de significación de aproximadamente 0.05 si $r_i \geq 2/\sqrt{n}$ o $r_i \leq -2/\sqrt{n}$. Si $n = 100$ y $r_1 = 0.16$, $r_2 = -0.09$ y $r_3 = -0.15$, ¿hay alguna evidencia de autocorrelación teórica en los tres primeros retrasos?

c. Si se prueba en forma simultánea la hipótesis nula del inciso (b) para más de un retraso, ¿por qué se desearía incrementar la constante de corte 2 en la región de rechazo?

67. Se recolectó una muestra de $n = 500$ pares (x, y) y se realizó una prueba de H_0: $\rho = 0$ contra H_a: $\rho \neq 0$. El valor P resultante se calculó como 0.00032.
   a. ¿Qué conclusión sería apropiada al nivel de significación 0.001?
   b. ¿Este pequeño valor P indica que hay una relación lineal muy fuerte entre x y y (un valor de ρ que difiere en forma considerable de 0)? Explique.
   c. Ahora componga una muestra de $n = 10\,000$ pares (x, y) que dan como resultado $r = 0.022$. Pruebe H_0: $\rho = 0$ contra H_a: $\rho \neq 0$ al nivel 0.05. ¿Es estadísticamente importante el resultado? Comente acerca de la importancia práctica de su análisis.

Ejercicios complementarios (68 a 85)

68. En el artículo "Refuse-Derived Fuel Evaluation in an Industrial Spreader-Stoker Boiler" (*J. Engr. For Gas Turbines and Power*, 1984: 782-788) aparecen los siguientes datos sobre x = % de aportación de calor de combustible obtenido de desechos (RDF) y y = % de eficiencia para cierta caldera.

x	37	30	48	29	27	16	0	20
y	78.0	77.2	74.4	77.7	76.9	79.0	82.1	76.5

   a. Obtenga la ecuación de la recta de regresión estimada.
   b. ¿El modelo de regresión lineal simple especifica una relación útil entre el porcentaje de calor aportado por el combustible obtenido de desechos y el porcentaje de eficiencia? Exprese y pruebe las hipótesis apropiadas.
   c. ¿Para obtener una estimación precisa de β_1, habría sido preferible hacer cuatro observaciones en $x = 0$ y cuatro en $x = 50$ (si se supone que el modelo es válido para x entre 0 y 50)? ¿Y si hubieran sido tres observaciones en $x = 0$ y tres en $x = 50$? Explique.
   d. Estime el porcentaje de eficiencia promedio real cuando el porcentaje de aportación de calor del combustible obtenido de desechos es 25 con un intervalo de confianza de 95%. ¿Al parecer se estimó con precisión el por ciento de eficiencia promedio real? Explique.

69. Los siguientes datos en relación con x = consumo de combustible diesel medido con el método de peso drenado y y = consumo medido con el método de rastreo de intervalos de confianza, ambos en g/h, se obtuvieron de una gráfica del artículo "A New Measurement Method of Diesel Engine Oil Consumption Rate" (*J. Society Auto Engr.*, 1985: 28-33):

x	4	5	8	11	12	16	17	20	22	28	30	31	39
y	5	7	10	10	14	15	13	25	20	24	31	28	39

   a. Si se supone que x y y se relacionan mediante el modelo de regresión lineal simple, pruebe H_0: $\beta_1 = 1$ contra H_a: $\beta_1 \neq 1$ con un nivel de significación de 0.05.
   b. Calcule el valor del coeficiente de correlación muestral para estos datos.

70. El siguiente resultado de SAS se basa en datos del artículo "Evidence for and the Rate of Denitrification in the Arabian Sea" (*Deep Sea Research*, 1978: 431-435). Las variables bajo estudio son x = nivel de salinidad (%) y y = concentración de nitrato (μM/L).
   a. ¿Cuál es el tamaño muestral n? [*Sugerencia*: determine los grados de libertad para SSE.]
   b. Calcule una estimación puntual de la concentración esperada de nitrato cuando el nivel de salinidad es 35.5.
   c. ¿Existe una relación lineal útil entre las dos variables?
   d. ¿Cuál es el valor del coeficiente de correlación muestral?
   e. ¿Utilizaría el modelo de regresión lineal simple para sacar conclusiones cuando el nivel de salinidad es 40?

Resultado de SAS para el ejercicio 70

Variable dependiente: CONC. DE NITRATO

Análisis de varianza

Fuente	GL	Suma de cuadrados	Cuadrado medio	Valor F	Prob > F
Modelo	1	64.49622	64.49622	63.309	0.0002
Error	6	6.11253	1.01875		
C Total	7	70.60875			

Raíz del ECM	1.00933	R cuadrada	0.9134
Media dep.	26.91250	R cuadrada ajustada	0.8990
V.C.	3.75043		

Estimaciones de parámetros

Variable	GL	Estimación del parámetro	Error estándar	T para HO: Parámetro = 0	Prob > |T|
ORDENADA AL ORIGEN	1	326.976038	37.71380243	8.670	0.0001
SALINIDAD	1	− 8.403964	1.05621381	− 7.957	0.0002

71. La presencia de carburos duros en aleaciones de hierro blanco con alto contenido de cromo, da como resultado una excelente resistencia a la abrasión, lo cual las hace adecuadas para el manejo de materiales en las industrias minera y de procesamiento de materiales. Los datos siguientes acerca de x = contenido de austenita retenida (%) y y = pérdida por desgaste abrasivo (mm^3), en pruebas de desgaste de pernos con granate como abrasivo, se obtuvieron de una gráfica del artículo "Microstructure-Property Relationships in High Chromium White Iron Alloys" (*Intl. Materials Reviews*, 1996: 59-82).

x	4.6	17.0	17.4	18.0	18.5	22.4	26.5	30.0	34.0
y	0.66	0.92	1.45	1.03	0.70	0.73	1.20	0.80	0.91

x	38.8	48.2	63.5	65.8	73.9	77.2	79.8	84.0
y	1.19	1.15	1.12	1.37	1.45	1.50	1.36	1.29

Utilice estos datos y el resultado de SAS para contestar las siguientes preguntas:

Resultado de SAS para el ejercicio 7.1

Variable dependiente: PÉRDIDA POR ABRASIÓN

Análisis de varianza

Fuente	GL	Suma de cuadrados	Cuadrado medio	Valor F	Prob > F
Modelo	1	0.63690	0.63690	15.444	0.0013
Error	15	0.61860	0.04124		
C Total	16	1.25551			

Raíz EMC	0.20308	R cuadrada	0.5073
Media dep.	1.10765	R cuadrada ajustada	0.4744
V.C.	18.33410		

Estimaciones de parámetros

Variable	GL	Estimación de parámetro	Error estándar	T para H0: Parámetro = 0	Prob > |T|
ORDENADA AL ORIGEN	1	0.787218	0.09525879	8.264	0.0001
CONT. DE AUSTENITA	1	0.007570	0.00192626	3.930	0.0013

a. ¿Qué proporción de la variación observada de pérdida por desgaste se puede atribuir a la relación del modelo de regresión lineal simple?
b. ¿Cuál es el valor del coeficiente de correlación muestral?
c. Pruebe la utilidad del modelo de regresión lineal simple con $\alpha = 0.01$.
d. Realice la estimación de la pérdida promedio real por desgaste cuando el contenido es 50% de modo que transmita información acerca de la confiabilidad y la precisión.
e. ¿Qué valor de pérdida por desgaste pronosticaría cuando el contenido es 30%, y cuál es el valor del residuo correspondiente?

72. Los datos siguientes se leyeron de un diagrama de dispersión en el artículo "Urban Emissions Measured with Aircraft" (*J. of the Air and Waste Mgmt. Assoc.*, 1998: 16-25). La variable de respuesta es ΔNO_y y la variable explicativa es ΔCO.

ΔCO	50	60	95	108	135
ΔNO_y	2.3	4.5	4.0	3.7	8.2
ΔCO	210	214	315	720	
ΔNO_y	5.4	7.2	13.8	32.1	

a. Ajuste un modelo adecuado a los datos y juzgue su utilidad.
b. Pronostique el valor de ΔNO_y que resultaría al realizar una observación más cuando ΔCO es 400, de manera que lleve información acerca de la precisión y la confiabilidad. ¿Se puede predecir con exactitud ΔNO_y? Explique.
c. El valor máximo de ΔCO es mucho mayor que los demás valores. ¿Al parecer esta observación tuvo un impacto importante en la ecuación ajustada?

73. Los datos siguientes son un subconjunto de los datos que aparecieron en el artículo "Radial Tension Strength of Pipe and Other Curved Flexural Members" (*J. Amer. Concrete Inst.*, 1980: 33-39). Las variables son edad de un espécimen de tubo (x, en días) y carga necesaria para obtener una primera grieta (y, en 1000 lb/pie).

x	20	20	20	25	25
y	11.45	10.42	11.14	10.84	11.17
x	25	31	31	31	
y	10.54	9.47	9.19	9.54	

a. Calcule la ecuación de la recta de regresión estimada.
b. Suponga que un modelo teórico indica que la disminución esperada en la carga relacionada con un aumento de un día en la edad es a lo sumo 0.10. ¿Los datos contradicen esta aseveración? Exprese y pruebe las hipótesis apropiadas al nivel de significación 0.05.
c. Con el fin de estimar la pendiente de la recta de regresión real de forma tan precisa como sea posible, ¿habría sido preferible realizar una sola observación en cada una de las edades 20, 21, 22, . . . , 30 y 31? Explique.
d. Estime la carga promedio real a la primera grieta cuando la edad es 28 días. Su estimación debe llevar información en relación con la precisión de la estimación.

74. Se llevó a cabo una investigación para estudiar la relación entre la velocidad (pie/s) y el ritmo de zancada (pasos/segundo) entre corredoras de maratón. Las cantidades de resumen resultantes son $n = 11$, $\sum$(velocidad) = 205.4, $\sum$(velocidad)^2 = 3880.08, $\sum$(ritmo) = 35.16, $\sum$(ritmo)^2 = 112.681 y $\sum$(velocidad)(ritmo) = 660.130.
a. Calcule la ecuación de la recta de mínimos cuadrados que utilizaría para pronosticar el ritmo de zancada a partir de la velocidad.
b. Calcule la ecuación de la recta de mínimos cuadrados que utilizaría para pronosticar la velocidad a partir del ritmo de zancada.
c. Calcule el coeficiente de determinación para la regresión de ritmo de zancada sobre la velocidad del inciso (a) y para la regresión de la velocidad sobre el ritmo de zancada del inciso (b). ¿Cómo se relacionan?

75. En el artículo "Photocharge Effects in Dye Sensitized Ag[Br,I] Emulsions at Millisecond Range Exposures" (*Photographic Sci. And Engr.*, 1981: 138-144) aparecen los siguientes datos relacionados con x = % de absorción de luz a 5 800 A y y = fotovoltaje máximo.

x	4.0	8.7	12.7	19.1	21.4
y	0.12	0.28	0.55	0.68	0.85
x	24.6	28.9	29.8	30.5	
y	1.02	1.15	1.34	1.29	

a. Construya un diagrama de dispersión de estos datos. ¿Qué indica?
b. Suponiendo que es apropiado el modelo de regresión lineal simple, obtenga la ecuación de la recta de regresión estimada.
c. ¿Qué proporción de la variación observada en el fotovoltaje máximo se explica mediante el modelo?

d. Prediga un fotovoltaje máximo cuando el porcentaje de absorción sea 19.1, y calcule el valor del residuo correspondiente.

e. Los autores del artículo afirman que hay una relación lineal útil entre el porcentaje de absorción y fotovoltaje máximo. ¿Está de acuerdo? Realice una prueba formal.

f. Dé una estimación del cambio en el fotovoltaje máximo esperado relacionado con 1% de aumento en la absorción de luz. Su evaluación debe llevar información acerca de la precisión de la estimación.

g. Repita el inciso (f) para el valor esperado de fotovoltaje máximo cuando el porcentaje de absorción de luz es 20.

76. En la sección 12.4 se presentó una fórmula para $V(\hat{\beta}_0 + \hat{\beta}_1 x^*)$ y un intervalo de confianza para $\beta_0 + \beta_1 x^*$. Con $x^* = 0$ se obtiene $\sigma^2_{\hat{\beta}_0}$ y un intervalo de confianza para β_0. Utilice los datos del ejemplo 12.11 para calcular la desviación estándar estimada de $\hat{\beta}_0$ y un intervalo de confianza de 95% para la ordenada al origen y de la recta de regresión verdadera.

77. Demuestre que $SSE = S_{yy} - \hat{\beta}_1 S_{xy}$, que proporciona una fórmula alternativa para calcular SSE.

78. Suponga que x y y son variables positivas y una muestra de n pares da como resultado $r \approx 1$. Si se calcula el coeficiente de correlación muestral para los pares (x, y^2), ¿el valor resultante también es más o menos 1? Explique.

79. Sean s_x y s_y las desviaciones estándar muestrales de las x y y observadas, respectivamente [de modo que $s_x^2 = \sum(x_i - \bar{x})^2/(n-1)$ y de manera similar para s_y^2].

a. Demuestre que otra expresión para la recta de regresión estimada $y = \hat{\beta}_0 + \hat{\beta}_1 x$ es

$$y = \bar{y} + r \cdot \frac{s_y}{s_x}(x - \bar{x})$$

b. Esta expresión para la recta de regresión se puede interpretar de la siguiente manera. Suponga que $r = 0.5$. ¿Cuál es entonces la y pronosticada para una x que está a 1 desviación estándar (unidades s_x) arriba de la media de las x_i? Si r fuera 1, el pronóstico sería que y está una desviación estándar arriba de su media $\bar{y}$, pero como $r = 0.5$, se pronostica una y que está a sólo 0.5 desviaciones estándar ($0.5s_y$ unidades) arriba de $\bar{y}$. Si se utilizan los datos del ejercicio 64 para un paciente cuya edad está una desviación estándar abajo del promedio de edad de la muestra, ¿a cuántas desviaciones estándar arriba o abajo del ΔCBG promedio muestral se ubica el ΔCBG pronosticado para el paciente?

80. Compruebe que el estadístico t para probar H_0: $\beta_1 = 0$ en la sección 12.3 es idéntico al estadístico t de la sección 12.5 para probar H_0: $\rho = 0$.

81. Utilice la fórmula para calcular SSE con el fin de comprobar que $r^2 = 1 - SSE/SST$.

82. En la biofiltración de aguas residuales, el aire descargado de una instalación de tratamiento pasa por una membrana porosa húmeda que hace que los contaminantes se disuelvan en agua y se transformen en productos inocuos. Los datos siguientes acerca de x = temperatura de entrada (°C) y y = eficiencia de eliminación (%) fueron la base para un diagrama de dispersión que apareció en el artículo "Treatment of Mixed Hydrogen Sulfide and Organic Vapors in a Rock Medium Biofilter" (*Water Environment Research*, 2001: 246-435).

Obs	Temp.	% de eliminación	Obs	Temp.	% de eliminación
1	7.68	98.09	17	8.55	98.27
2	6.51	98.25	18	7.57	98.00
3	6.43	97.82	19	6.94	98.09
4	5.48	97.82	20	8.32	98.25
5	6.57	97.82	21	10.50	98.41
6	10.22	97.93	22	16.02	98.51
7	15.69	98.38	23	17.83	98.71
8	16.77	98.89	24	17.03	98.79
9	17.13	98.96	25	16.18	98.87
10	17.63	98.90	26	16.26	98.76
11	16.72	98.68	27	14.44	98.58
12	15.45	98.69	28	12.78	98.73
13	12.06	98.51	29	12.25	98.45
14	11.44	98.09	30	11.69	98.37
15	10.17	98.25	31	11.34	98.36
16	9.64	98.36	32	10.97	98.45

Las cantidades de resumen calculadas son $\sum x_i = 384.26$, $\sum y_i = 3149.04$, $\sum x_i^2 = 5099.2412$, $\sum x_i y_i = 37\,850.7762$ y $\sum y_i^2 = 309\,892.6548$.

a. ¿El diagrama de dispersión de los datos indica idoneidad del modelo de regresión lineal simple?

b. Ajuste el modelo de regresión lineal simple, obtenga una predicción puntual de la eficiencia de elimi-

nación cuando la temperatura es igual a 10.50 y calcule el valor del residuo correspondiente.

c. ¿Más o menos cuál es el tamaño de una desviación característica de puntos en el diagrama de dispersión de la recta de mínimos cuadrados?

d. ¿Qué proporción de la variación observada en la eficiencia de eliminación se puede atribuir a la relación del modelo?

e. Estime la pendiente de manera que lleve información acerca de la confiabilidad y la precisión, e interprete su estimación.

f. La comunicación personal con los autores del artículo, reveló que hubo una observación más que no se incluyó en el diagrama de dispersión: (6.53, 96.55). ¿Qué impacto tiene esta otra observación en la recta de mínimos cuadrados y los valores de s y r^2?

83. Los procesos de incubación normal en acuacultura producen de forma inevitable estrés en los peces, lo cual podría producir un impacto negativo en el desarrollo, reproducción, calidad de la carne y susceptibilidad a las enfermedades. El estrés se pone de manifiesto en las concentraciones altas y sostenidas de corticoesteroides. En el artículo "Evaluation of Simple Instruments for the Measurement of Blood Glucose and Lactate, and Plasma Protein as Stress Indicators in Fish" (*J. of the World Aquaculture Society*, 1999: 276-284) se describe un experimento en el que se sometió a los peces a un protocolo de estrés, y luego se sacaron y se les realizaron pruebas varias veces después de haber aplicado el protocolo. Los datos siguientes de x = tiempo (min) y y = concentración de glucosa en la sangre (mmol/L) se leyeron de una gráfica.

x	2	2	5	7	12	13	17	18	23	24	26	28
y	4.0	3.6	3.7	4.0	3.8	4.0	5.1	3.9	4.4	4.3	4.3	4.4

x	29	30	34	36	40	41	44	56	56	57	60	60
y	5.8	4.3	5.5	5.6	5.1	5.7	6.1	5.1	5.9	6.8	4.9	5.7

Utilice los métodos de este capítulo para analizar los datos y escriba un breve informe que resuma sus conclusiones (suponga que los investigadores tienen un interés particular en la concentración de glucosa 30 minutos después del estrés).

84. En el artículo "Evaluating the BOD POD for Assessing Body Fat in Collegiate Fottball Players" (*Medicine and Science in Sports and Exercise*, 1999: 1350-1356) se describe un nuevo dispositivo de desplazamiento de aire para medir la grasa corporal. En el procedimiento ordinario se utiliza el dispositivo de pesaje hidrostático, que mide el porcentaje de grasa corporal por medio de desplazamiento de agua. A continuación se presentan los datos leídos de una gráfica del artículo.

BOD	2.5	4.0	4.1	6.2	7.1	7.0	8.3	9.2	9.3	12.0	12.2
HW	8.0	6.2	9.2	6.4	8.6	12.2	7.2	12.0	14.9	12.1	15.3

BOD	12.6	14.2	14.4	15.1	15.2	16.3	17.1	17.9	17.9
HW	14.8	14.3	16.3	17.9	19.5	17.5	14.3	18.3	16.2

a. Utilice varios métodos para decidir si es posible que las dos técnicas midan en promedio la misma cantidad de grasa.

b. Por medio de los datos desarrolle una forma de predecir una medición HW a partir de una medición BOD POD e investigue la eficacia de las predicciones.

85. Reconsidere la situación del ejercicio 68, donde x = % de aporte de calor de combustible obtenido de desechos y y = % de eficiencia para una caldera en particular, se relacionaron por medio del modelo de regresión lineal simple $Y = \beta_0 + \beta_1 x + \epsilon$. Suponga que para una segunda caldera, estas variables también se relacionan por medio del modelo de regresión lineal simple $Y = \gamma_0 + \gamma_1 x + \epsilon$ y que $V(\epsilon) = \sigma^2$ para ambas calderas. Si el conjunto de datos consiste en n_1 observaciones en la primera caldera y n_2 en la segunda y si SSE_1 y SSE_2 representan las dos sumas de cuadrados del error, entonces una estimación agrupada de σ^2 es $\hat{\sigma}^2 = (SSE_1 + SSE_2)/(n_1 + n_2 - 4)$. Sean SS_{x1} y SS_{x2} la $\sum(x_i - \bar{x})^2$ para los datos de las calderas primera y segunda, respectivamente. Una prueba de H_0: $\beta_1 - \gamma_1 = 0$ (pendientes iguales) está basada en el estadístico

$$T = \frac{\hat{\beta}_1 - \hat{\gamma}_1}{\hat{\sigma}\sqrt{\dfrac{1}{SS_{x1}} + \dfrac{1}{SS_{x2}}}}$$

Cuando H_0 es verdadera, T tiene una distribución t con $n_1 + n_2 - 4$ grados de libertad. Suponga que las seis observaciones relacionadas con la segunda caldera son (0, 81.3), (10, 78.4), (20, 78.2), (25, 79.1), (30, 77.6) y (40, 77.4). Con lo anterior y los datos del ejercicio 68 realice una prueba al nivel 0.05 para ver si el cambio esperado en el porcentaje de eficiencia relacionado con un aumento de 1% en el aporte de calor del combustible obtenido de desechos es idéntico para las dos calderas.

Bibliografía

Draper, Norman y Harry Smith, *Applied Regression Analysis* (3a. ed.), Wiley, Nueva York, 1999. El libro más completo y autorizado sobre el análisis de regresión actualmente en venta.

Neter, John, Michael Kutner, Christopher Nachtsheim y William Wasserman, *Applied Linear Statistical Models* (4a. ed.), Irwin, Homewood, IL, 1996. Los primeros 15 capítulos constituyen un estudio muy interesante e informativo del análisis de regresión.

13

Regresión no lineal y múltiple

Introducción

El modelo probabilístico estudiado en el capítulo 12, especificó que el valor observado de la variable dependiente Y se desvía de la función de regresión lineal $\mu_{Y \cdot x} = \beta_0 + \beta_1 x$ por una cantidad aleatoria. Aquí se consideran dos formas de generalizar el modelo de regresión lineal simple. La primera forma es reemplazar $\beta_0 + \beta_1 x$ por una función no lineal de x, y la segunda es usar una función de regresión relacionada con más de una sola variable independiente. Después de ajustar una función de regresión de la forma elegida para los datos proporcionados, es importante contar con métodos para hacer inferencias acerca de los parámetros del modelo elegido. Sin embargo, antes de usar estos métodos, el analista debe evaluar primero la validez del modelo elegido. En la sección 13.1, se analizan métodos que se basan sobre todo en un análisis gráfico de los residuos (y observadas menos pronosticadas) para comprobar la pertinencia de un modelo ajustado.

En la sección 13.2, se consideran funciones de regresión no lineales de una sola variable independiente x que son "intrínsecamente lineales". Esto quiere decir que es posible transformar una o ambas variables para que la relación entre las nuevas variables sea lineal. Otra clase de relaciones no lineales se obtiene al usar las funciones de regresión polinomiales de la forma $\mu_{Y \cdot x} = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k$; estos modelos polinomiales son el tema de la sección 13.3. El análisis de regresión múltiple requiere construir modelos para relacionar a y con dos o más variables

independientes. El tema central de la sección 13.4 es sobre la interpretación de varios modelos de regresión múltiple y sobre la comprensión y uso del resultado de la regresión de varios paquetes de software para estadística. En la última sección del capítulo se estudian algunas ampliaciones y desventajas del modelado de regresión múltiple.

13.1 Pertinencia y comprobación del modelo

Una gráfica de los pares observados (x_i, y_i) es un primer paso necesario para decidir la forma de una relación matemática entre x y y. Es posible ajustar a los datos muchas funciones distintas a la lineal ($y = b_0 + b_1x$), ya sea con el principio de mínimos cuadrados u otro método de ajuste. Una vez que ha sido ajustada una función de la forma elegida, es importante comprobar el ajuste del modelo para ver si en realidad es apropiado. Una forma de estudiar el ajuste es superponer una gráfica de la función del mejor ajuste sobre el diagrama de dispersión de los datos. Sin embargo, cualquier inclinación o curvatura de la función del mejor ajuste podría oscurecer algunos aspectos del ajuste que deben ser investigados. Además, la escala en el eje vertical podría dificultar la evaluación de la magnitud en que los valores observados se desvían de las funciones del mejor ajuste.

Residuos y residuos estandarizados

Un método más eficaz para evaluar la idoneidad del modelo, es calcular los valores $\hat{y}_i$ ajustados o pronosticados y los residuos $e_i = y_i - \hat{y}_i$ y después dibujar varias funciones de estas cantidades calculadas. Luego, se examinan las gráficas ya sea para confirmar la elección del modelo o por indicaciones de que el modelo no es apropiado. Suponga que el modelo de regresión lineal simple es correcto, y sea $y = \hat{\beta}_0 + \hat{\beta}_1x$ la ecuación de la recta de regresión estimada. Entonces el i-ésimo residuo es $e_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1x_i)$. Para obtener las propiedades de los residuos, sea $e_i = Y_i - \hat{Y}_i$ que representa el i-ésimo residuo como una variable aleatoria (*va*) (antes de que en realidad se hagan las observaciones). Entonces

$$E(Y_i - \hat{Y}_i) = E(Y_i) - E(\hat{\beta}_0 + \hat{\beta}_1x_i) = \beta_0 + \beta_1x_i - (\beta_0 + \beta_1x_i) = 0 \qquad (13.1)$$

Debido a que $\hat{Y}_i$ $(= \hat{\beta}_0 + \hat{\beta}_1x_i)$ es una función lineal de las Y_j, también lo es $Y_i - \hat{Y}_i$ (los coeficientes dependen de las x_j). Así que la normalidad de las Y_j implica que cada residuo esté distribuido normalmente. También se puede demostrar que

$$V(Y_i - \hat{Y}_i) = \sigma^2 \cdot \left[1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}}\right] \qquad (13.2)$$

Al sustituir σ^2 por s^2 y sacar la raíz cuadrada de la ecuación (13.2) se obtiene la desviación estándar estimada de un residuo.

Ahora se estandariza cada residuo restando el valor medio (cero) y luego se divide entre la desviación estándar estimada.

Los **residuos estandarizados** se determinan mediante

$$e_i^* = \frac{y_i - \hat{y}_i}{s\sqrt{1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}}}} \qquad i = 1, \ldots, n \tag{13.3}$$

Si, por ejemplo, un residuo estandarizado es 1.5, entonces el residuo es 1.5 desviaciones estándar (estimadas) más grande de lo que se esperaría de ajustar el modelo correcto. Observe que las varianzas de los residuos difieren entre sí. Sin embargo, si n es razonablemente grande, el término entre corchetes de la ecuación (13.2) será aproximadamente 1, por lo tanto algunas fuentes utilizan e_i/s como el residuo estandarizado. El cálculo de las e_i^* puede ser tedioso, pero los paquetes de software para estadística más comunes proporcionan de manera automática estos valores y (si se solicita) construyen varias gráficas en las que intervienen.

Ejemplo 13.1 En el ejercicio 19 del capítulo 12, se presentaron los datos de x = rapidez de liberación del área del quemador y y = emisiones de NO_x. Aquí se reproducen los datos y se proporcionan los valores ajustados, residuos y residuos estandarizados. La recta de regresión estimada es $y = -45.55 + 1.71x$ y $r^2 = 0.961$. Observe que los residuos estandarizados no son un múltiplo constante de los residuos (es decir, $e_i^* \neq e_i/s$).

x_i	y_i	$\hat{y}_i$	e_i	e_i^*
100	150	125.6	24.4	0.75
125	140	168.4	−28.4	−0.84
125	180	168.4	11.6	0.35
150	210	211.1	−1.1	−0.03
150	190	211.1	−21.1	−0.62
200	320	296.7	23.3	0.66
200	280	296.7	−16.7	−0.47
250	400	382.3	17.7	0.50
250	430	382.3	47.7	1.35
300	440	467.9	−27.9	−0.80
300	390	467.9	−77.9	−2.24
350	600	553.4	46.6	1.39
400	610	639.0	−29.0	−0.92
400	670	639.0	31.0	0.99

■

Gráficas de diagnóstico

Los diagramas básicos que recomiendan muchos estadísticos para una evaluación de la validez y utilidad del modelo son los siguientes:

1. e_i^* (o e_i) en el eje vertical contra x_i en el eje horizontal.
2. e_i^* (o e_i) en el eje vertical contra $\hat{y}_i$ en el eje horizontal.

3. $\hat{y}_i$ en el eje vertical contra y_i en el eje horizontal.
4. Una gráfica de probabilidad normal de los residuos estandarizados

Las gráficas 1 y 2 se llaman **diagramas residuales** (contra la variable independiente y los valores ajustados, respectivamente), mientras que la gráfica 3 se ajusta contra los valores observados.

Si la gráfica 3 produce puntos cercanos a la recta de 45° [pendiente +1 por (0, 0)], entonces la función de regresión estimada da predicciones precisas de los valores observados en realidad. Por consiguiente, la gráfica 3 proporciona una evaluación visual de la efectividad del modelo para hacer predicciones. Siempre que el modelo sea correcto, ninguna gráfica residual debe exhibir patrones distintos. Los residuos deben tener una distribución aleatoria respecto a 0, de acuerdo con una distribución normal, así que todos menos unos muy pocos residuos estandarizados deben quedar entre −2 y +2 (es decir, todos excepto algunos residuos dentro de dos desviaciones estándar de su valor esperado 0). La gráfica de los residuos estandarizados contra $\hat{y}$ es en realidad una combinación de las otras dos gráficas, lo que muestra de forma implícita cómo varían los residuos con x y cómo se comparan los valores ajustados con los valores observados. Esta última gráfica es la única que se recomienda con más frecuencia para el análisis de regresión múltiple. La gráfica 4 permite que el analista evalúe la posibilidad de la suposición de que ϵ tiene una distribución normal.

Ejemplo 13.2 (Continuación del ejemplo 13.1)

En la figura 13.1 se ilustra un diagrama de dispersión de los datos y las cuatro gráficas recomendadas antes. La gráfica de $\hat{y}$ contra y confirma la impresión determinada mediante r^2 de que x es efectiva para predecir y y también indica que no hay y observada para la cual el valor

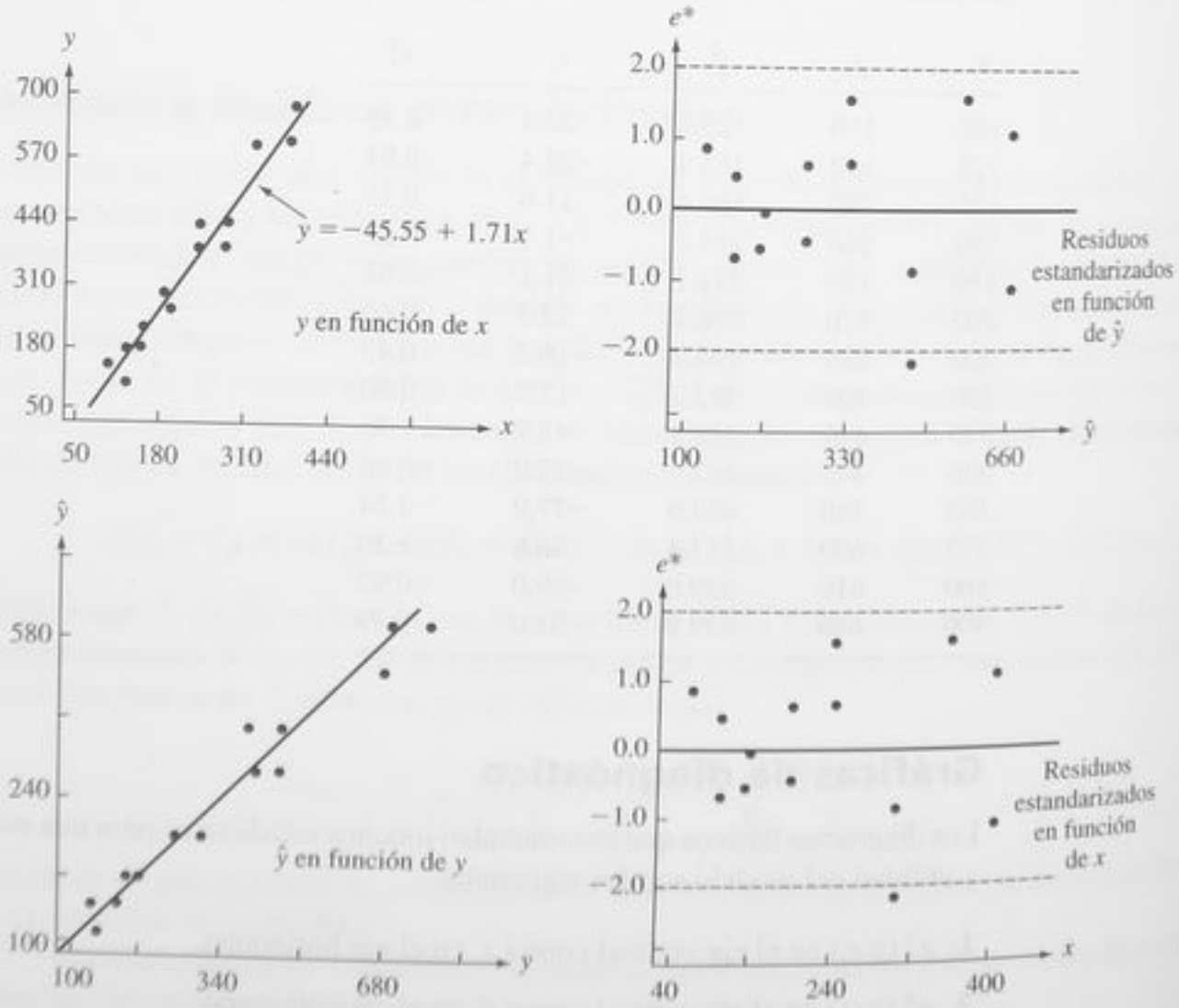


Figura 13.1 Gráficas para los datos del ejemplo 13.1

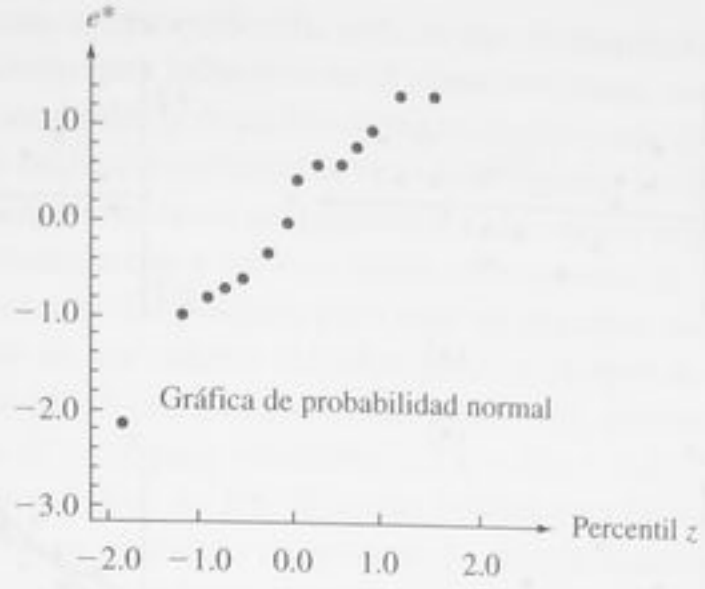


Figura 13.1 (continuación) Gráficas para los datos del ejemplo 13.1

pronosticado esté muy lejos de la marca. Ambas gráficas residuales no muestran un patrón inusual o valores discrepantes. Hay un residuo estandarizado un poco fuera del intervalo (−2, 2), pero esto no es sorprendente en una muestra de tamaño 14. La gráfica de probabilidad normal de los residuos estandarizados es razonablemente recta. En resumen, las gráficas no dejan duda de la idoneidad de una relación lineal simple o el ajuste a los datos proporcionados. ■

Dificultades y remedios

Aunque se espera que al realizar un análisis se obtengan gráficas como las de la figura 13.1, con bastante frecuencia las gráficas indican una o más de las siguientes dificultades:

1. Una relación probabilística no lineal entre x y y es apropiada.
2. La varianza de ϵ (y no de Y) no es una constante σ^2 sino que depende de x.
3. El modelo seleccionado ajusta bien los datos, excepto por muy pocos valores discrepantes o atípicos, lo cual podría influir bastante en la elección de la función del mejor ajuste.
4. El término de error ϵ no tiene una distribución normal.
5. Cuando el subíndice i indica el orden temporal de las observaciones, las ϵ_i exhiben dependencia con el tiempo.
6. Una o más variables independientes han sido omitidas del modelo.

En la figura 13.2 se ilustran las gráficas residuales que corresponden a los elementos 1 a 3, 5 y 6. En el capítulo 4, se analizan patrones de gráficas de probabilidad normal que arrojan dudas sobre la suposición de una distribución normal subyacente. Observe que los residuos de los datos de la figura 13.2(d) con el punto encerrado en un círculo incluido no necesariamente por sí mismos indicarían más análisis, aunque con ese punto borrado se ajuste una nueva recta, la nueva recta difiere en forma considerable de la original. Este tipo de comportamiento es más difícil de identificar en la regresión múltiple. Es más probable que surja cuando hay un solo dato (o muy pocos) con el valor de la variable independiente lejos del resto de los datos.

Ahora se indica de manera breve qué remedios están disponibles para los tipos de dificultades. Para un análisis más completo, se deben consultar una o más de las referencias del

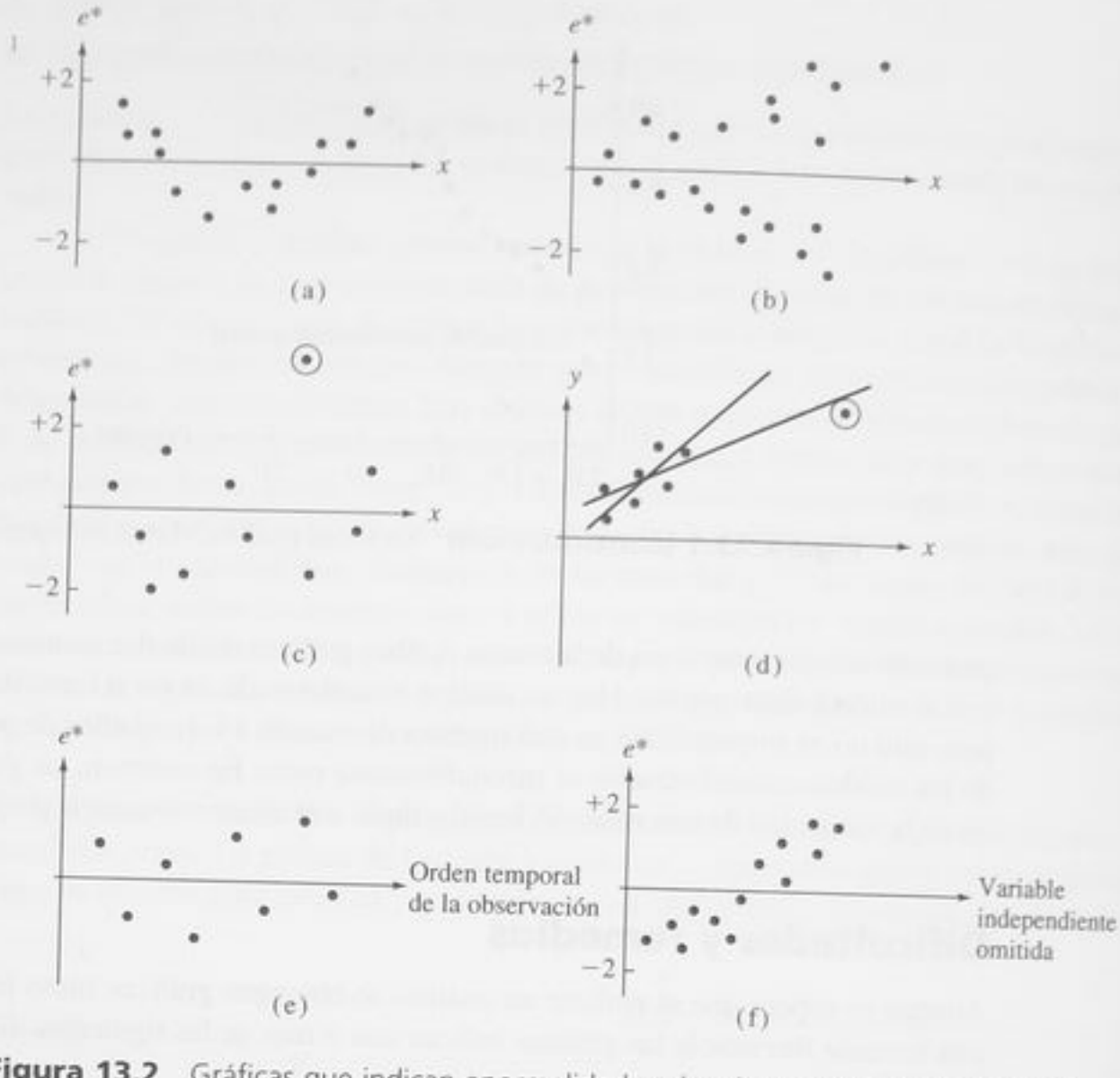


Figura 13.2 Gráficas que indican anormalidad en los datos: (a) relación no lineal; (b) varianza no constante; (c) observación discrepante; (d) observación con gran influencia; (e) dependencia de errores; (f) variable omitida.

análisis de regresión. Si la gráfica de residuos se parece un poco a la de la figura 13.2(a), que muestra un patrón curvo, entonces se podría ajustar una función lineal de x.

La gráfica residual de la figura 13.2(b) indica que, aunque parecería razonable una relación de línea recta, la suposición de que $V(Y_i) = \sigma^2$ para cada i es de validez incierta. Cuando las suposiciones del capítulo 12 son válidas, se puede demostrar que entre los estimadores insesgados de β_0 y β_1, los estimadores de mínimos cuadrados ordinarios tienen varianza mínima. Estos estimadores dan igual peso a cada (x_i, Y_i). Si la varianza de Y se incrementa con x, entonces las Y_i para x_i grande deben recibir menos peso que las que tienen x_i pequeña. Esto indica que β_0 y β_1 se deben estimar al minimizar

$$f_w(b_0, b_1) = \sum w_i[y_i - (b_0 + b_1 x_i)]^2 \tag{13.4}$$

donde las w_i son pesos que disminuyen con el incremento de x_i. La minimización de la expresión (13.4) produce estimaciones de **mínimos cuadrados ponderados**. Por ejemplo, si la desviación estándar de Y es proporcional a x (para $x > 0$), es decir, $V(Y) = kx^2$, entonces se puede demostrar que los pesos $w_i = 1/x_i^2$ producen mejores estimadores de β_0 y β_1. Los libros de John Neter *et al.* y S. Chatterjee y Bertram Price contienen más detalles (véase la bibliografía del capítulo). Los econometristas (economistas que emplean métodos estadísticos) utilizan con mucha frecuencia los mínimos cuadrados ponderados para estimar parámetros.

Cuando las gráficas, u otra evidencia, indican que el conjunto de datos contiene valores atípicos o puntos que tienen gran influencia en el ajuste resultante, un método posible es omitir estos puntos atípicos y recalcular la ecuación de regresión estimada. Esto sería correcto si se encontrara que los valores atípicos resultaron de errores al registrar los datos o errores experimentales. Si no se puede hallar una causa asignable para los valores atípicos, aún es conveniente informar la ecuación estimada con y sin los valores atípicos omitidos. Sin embargo, otro método es retener los valores atípicos posibles pero usar un principio de estimación que pone relativamente menos peso en los valores alejados que el principio de mínimos cuadrados. Un principio de este tipo es MAD (*minimize absolute deviations,* minimizar la desviaciones absolutas), que selecciona a $\hat{\beta}_0$ y $\hat{\beta}_1$ para minimizar $\sum | y_i - (b_0 + b_1 x_i) |$. A diferencia de las estimaciones de mínimos cuadrados, no hay fórmulas bonitas para las estimaciones de MAD; sus valores se deben determinar mediante un proceso de cálculo iterativo. Esta clase de procedimientos también se utilizan cuando se sospecha que las ϵ_i tienen una distribución que no es normal, sino que tiene "extremos densos" (lo cual hace mucho más probable que para la distribución normal que los valores discrepantes entren en la muestra); los procedimientos de regresión robustos son los que producen estimaciones confiables para una amplia variedad de distribuciones de error subyacentes. Los estimadores de mínimos cuadrados no son robustos de la misma forma que la media muestral $\bar{X}$ no es un estimador robusto para μ.

Cuando una gráfica indica dependencia con el tiempo en los términos de error, un análisis apropiado podría tener que ver con una transformación de las y o con un modelo que incluye en forma explícita una variable de tiempo. Por último, una gráfica como la de la figura 13.2(f), que muestra un patrón en los residuos cuando se grafican en función de una variable omitida, indica que se debe considerar un modelo de regresión múltiple que incluye la variable omitida antes.

Ejercicios | Sección 13.1 (1 a 14)

1. Suponga que las variables x = distancia de conmutación y y = tiempo de conmutación se relacionan de acuerdo con el modelo de regresión lineal simple con $\sigma = 10$.
   a. Si se hacen $n = 5$ observaciones en los valores de x, $x_1 = 5, x_2 = 10, x_3 = 15, x_4 = 20$ y $x_5 = 25$, calcule las desviaciones estándar de los cinco residuos correspondientes.
   b. Repita el inciso (a) para $x_1 = 5$, $x_2 = 10$, $x_3 = 15$, $x_4 = 20$ y $x_5 = 50$.
   c. ¿Qué implican los resultados de los incisos (a) y (b) acerca de la desviación de la recta estimada de la observación hecha en el valor muestreado máximo de x?

2. Los valores de x y los residuos estandarizados para los datos de flujo de cloro/rapidez de grabado del ejercicio 52 (sección 12.4) se muestran en la tabla siguiente. Construya una gráfica residual estandarizada y realice comentarios en relación con la apariencia de la gráfica.

x	1.50	1.50	2.00	2.50	2.50
e^*	0.31	1.02	−1.15	−1.23	0.23
x	3.00	3.50	3.50	4.00	
e^*	0.73	−1.36	1.53	0.07	

3. En el ejemplo 12.6 se presentaron los residuos de una regresión lineal simple del contenido de humedad y sobre la tasa de filtración x.
   a. Trace la gráfica de los residuos en función de x. ¿La gráfica resultante indica que una función de regresión de línea recta es una elección razonable del modelo? Explique su razonamiento.
   b. Con $s = 0.665$, calcule los valores de los residuos estandarizados. ¿Es $e_i^* = e_i/s$ para $i = 1, \ldots, n$ o no están cercanas las e_i^* para ser proporcionales a las e_i?
   c. Tace la gráfica de los residuos estandarizados contra x. ¿La gráfica difiere en forma significativa en cuanto a apariencia general de la gráfica del inciso (a)?

4. La resistencia al desgaste de ciertos componentes de reactor nuclear hechos de Zircaloy-2 se determina en parte mediante las propiedades de la capa de óxido. Los siguientes datos aparecen en un artículo que propuso un nuevo método de prueba no destructiva para monitorear el espesor de la capa ("Monitoring of Oxide

Layer Thickness on Zircaloy-2 by the Eddy Current Test Method", *J. of Testing and Eval.*, 1987: 333-336). Las variables son x = espesor de la capa de óxido (μm) y y = respuesta de la corriente de Foucault (unidades arbitrarias).

x	0	7	17	114	133
y	20.3	19.8	19.5	15.9	15.1
x	142	190	218	237	285
y	14.7	11.9	11.5	8.3	6.6

a. Los autores resumieron la relación mediante la ecuación de la recta de mínimos cuadrados $y = 20.6 - 0.047x$. Calcule y trace la gráfica de los residuos contra x y luego comente acerca de la idoneidad del modelo de regresión lineal simple.

b. Utilice $s = 0.7921$ para calcular los residuos estandarizados de una regresión lineal simple. Construya una gráfica residual estandarizada y comente acerca de características importantes. También construya una gráfica de probabilidad normal y haga comentarios al respecto.

5. Cuando disminuye la temperatura del aire, el agua de los ríos se vuelve superenfriada y se forman cristales de hielo. Este hielo afecta en forma importante la hidráulica de un río. En el artículo "Laboratory Study of Anchor Ice Growth" (*J. of Cold Regions Engr.*, 2001: 60-66) se describe un experimento en el que se estudió el espesor del hielo (mm) como función del tiempo transcurrido (h) en condiciones especificadas. Los datos siguientes se leyeron de una gráfica del artículo: $n = 33$; $x = 0.17, 0.33, 0.50, .67, \ldots, 5.50$; $y = 0.50, 1.25, 1.50, 2.75, 3.50, 4.75, 5.75, 5.60, 7.00, 8.00, 8.25, 9.50, 10.50, 11.00, 10.75, 12.50, 12.25, 13.25, 15.50, 15.00, 15.25, 16.25, 17.25, 18.00, 18.25, 18.15, 20.25, 19.50, 20.00, 20.50, 20.60, 20.50, 19.80$.

a. El valor de r^2 que resulta de un ajuste de mínimos cuadrados es 0.977. Interprete este valor y comente acerca de la idoneidad de suponer una relación lineal aproximada.

b. Los residuos, listados en el mismo orden que los valores de x, son

−1.03	−0.92	−1.35	−0.78	−0.68	−0.11	0.21
−0.59	0.13	0.45	0.06	0.62	0.94	0.80
−0.14	0.93	0.04	0.36	1.92	0.78	0.35
0.67	1.02	1.09	0.66	−0.09	1.33	−0.10
−0.24	−0.43	−1.01	−1.75	−3.14		

Trace la gráfica de los residuos en función del tiempo transcurrido. ¿Qué indica la gráfica?

6. Los siguientes datos de x = densidad verdadera (kg/mm^3) y y = contenido de humedad (% b. s.) se leyeron de una gráfica del artículo "Physical Properties of Cumin Seed" (*J. Agric. Engr. Res.*, 1996: 93-98).

x	7.0	9.3	13.2	16.3	19.1	22.0
y	1046	1065	1094	1117	1130	1135

La ecuación de la recta de mínimos cuadrados es $y = 1008.14 + 6.19268x$ (ésta difiere muy poco de la ecuación del artículo); $s = 7.265$ y $r^2 = 0.968$.

a. Realice una prueba de utilidad del modelo y comente.

b. Calcule los valores de los residuos y trace la gráfica en función de x. ¿La gráfica indica que es inapropiada una función de regresión lineal?

c. Calcule los valores de los residuos estandarizados y trace la gráfica en función de x. ¿Hay algunos residuos estandarizados inusualmente grandes (positivos o negativos)? ¿Esta gráfica transmite el mismo mensaje que la gráfica del inciso (b) en relación con la idoneidad de la función de regresión lineal?

7. En el artículo Effects of Gamma Radiation on Juvenile and Mature Cuttings of Quaking Aspen" (*Forest Science*, 1967: 240-245) aparecen los datos siguientes acerca del tiempo de exposición a la radiación (x, en kr/16 h) y el peso en seco de raíces (y, en mg $\times 10^{-1}$).

x	0	2	4	6	8
y	110	123	119	86	62

a. Construya una gráfica de dispersión. ¿La gráfica indica que es apropiada una relación probabilística lineal?

b. Una regresión lineal da como resultado la recta de mínimos cuadrados $y = 127 - 6.65x$, con $s = 16.94$. Calcule los residuos y los residuos estandarizados y luego trace las gráficas residuales. ¿Qué indican estas gráficas? ¿Qué tipo de función debe proporcionar un mejor ajuste para los datos que una recta?

8. El registro continuo de la frecuencia cardiaca (*heart rate*) se puede usar para obtener información acerca del nivel de intensidad de ejercicio o esfuerzo físico durante la participación en deportes, trabajo u otras actividades diarias. En el artículo "The Relationship between Heart Rate and Oxygen Uptake During Non-Steady State Exercise" (*Ergonomics*, 2000: 1578-1592) se describe un estudio para investigar el uso de la respuesta de frecuencia cardiaca (x, como porcentaje de la frecuencia máxima) para predecir la captación de oxígeno (y, como porcentaje de la captación máxima) durante el

ejercicio. Los datos siguientes se leyeron de una gráfica del artículo.

HR	43.5	44.0	44.0	44.5	44.0	45.0	48.0	49.0
VO_2	22.0	21.0	22.0	21.5	25.5	24.5	30.0	28.0
HR	49.5	51.0	54.5	57.5	57.7	61.0	63.0	72.0
VO_2	32.0	29.0	38.5	30.5	57.0	40.0	58.0	72.0

Utilice un paquete de software para estadística a fin de llevar a cabo un análisis de regresión lineal simple, con atención especial a la presencia de algunas observaciones inusuales o preponderantes.

9. Considere los siguientes cuatro conjuntos de datos (x, y); los primeros tres tienen los mismos valores de x, así que estos valores se listan sólo una vez Frank Anscombe, "Graphs in Statistical Analysis", *Amer. Statistician*, 1973: 17-21):

Conjunto de datos	1–3	1	2	3	4	4
Variable	x	y	y	y	x	y
	10.0	8.04	9.14	7.46	8.0	6.58
	8.0	6.95	8.14	6.77	8.0	5.76
	13.0	7.58	8.74	12.74	8.0	7.71
	9.0	8.81	8.77	7.11	8.0	8.84
	11.0	8.33	9.26	7.81	8.0	8.47
	14.0	9.96	8.10	8.84	8.0	7.04
	6.0	7.24	6.13	6.08	8.0	5.25
	4.0	4.26	3.10	5.39	19.0	12.50
	12.0	10.84	9.13	8.15	8.0	5.56
	7.0	4.82	7.26	6.42	8.0	7.91
	5.0	5.68	4.74	5.73	8.0	6.89

Para cada uno de estos cuatro conjuntos de datos, los valores de los estadísticos de resumen $\sum x_i$, $\sum x_i^2$, $\sum y_i$, $\sum y_i^2$ y $\sum x_i y_i$ son casi idénticos, así que las cantidades calculadas a partir de estos cinco serán casi idénticos para los cuatro conjuntos, la recta de mínimos cuadrados ($y = 3 + 0.5x$), SSE, s^2, r^2, intervalos t, estadísticos t, etcétera. Los estadísticos de resumen no proporcionan manera de distinguir entre los cuatro conjuntos de datos. Con base en un diagrama de dispersión y una gráfica residual para cada conjunto, comente acerca de la idoneidad o inadecuación de ajustar un modelo de línea recta; incluya en sus comentarios algunas sugerencias específicas de cómo se podría modificar o calificar un "análisis de línea recta".

10. a. Muestre que $\sum_{i=1}^{n} e_i = 0$ cuando las e_i son los residuos de una regresión lineal simple.

b. ¿Los residuos de una regresión lineal simple son independientes entre sí, están correlacionados de forma positiva o negativa? Explique.

c. Demuestre que $\sum_{i=1}^{n} x_i e_i = 0$ para los residuos de una regresión lineal simple. [Este resultado, junto con el inciso (a), muestra que hay dos restricciones lineales en las e_i, que dan como resultado una pérdida de dos grados de libertad cuando se emplean los residuos cuadrados para estimar σ^2.]

d. ¿Es cierto que $\sum_{i=1}^{n} e_i^* = 0$? Dé una demostración o un contraejemplo.

11. a. Exprese el i-ésimo residuo $Y_i - \hat{Y}_i$ (donde $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ en la forma $\sum c_j Y_j$ una función lineal de las Y_j. Luego, utilice las reglas de varianza para verificar que $V(Y_i - \hat{Y}_i)$ se determina mediante la expresión (13.2).

b. Se puede demostrar que Y_i y $Y_i - \hat{Y}_i$ (el valor i-ésimo pronosticado y el residuo) son independientes entre sí. Utilice este hecho, la relación $Y_i = \hat{Y}_i + (Y_i - \hat{Y}_i)$ y la expresión para $V(\hat{Y})$ de la sección 12.4 para comprobar de nuevo la expresión (13.2).

c. A medida que x_i se aleja de $\bar{x}$, ¿qué sucede con $V(\hat{Y}_i)$ y con $V(Y_i - \hat{Y}_i)$?

12. a. ¿Una regresión lineal podría dar como resultado los residuos 23, −27, 5, 17, −8, 9 y 15? ¿Por qué sí o por qué no?

b. ¿Una regresión lineal podría dar como resultado los residuos 23, −27, 5, 17, −8, −12 y 2 correspondientes a los valores de 3, −4, 8, 12, −14, −20 y 25? ¿Por qué si o por qué no? (*Sugerencia*: véase el ejercicio 10.)

13. Recuerde que $\hat{\beta}_0 + \hat{\beta}_1 x$ tiene una distribución normal con valor esperado $\beta_0 + \beta_1 x$ y varianza

$$\sigma^2\left\{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right\}$$

de modo que

$$Z = \frac{\hat{\beta}_0 + \hat{\beta}_1 x - (\beta_0 + \beta_1 x)}{\sigma\left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right)^{1/2}}$$

tiene una distribución normal estándar. Si $S = \sqrt{\text{SSE}/(n-2)}$ se sustituye por σ, la variable resultante tiene una distribución t con $n - 2$ gl. Por analogía, ¿cuál es la distribución de cualquier residuo estandarizado en particular? Si $n = 25$, ¿cuál es la probabilidad de que un determinado residuo estandarizado caiga fuera del intervalo $(-2.50, 2.50)$?

14. Si hay por lo menos un valor x en el que se ha hecho más de una observación, hay un procedimiento de prueba formal para probar

H_0: $\mu_{Y \cdot x} = \beta_0 + \beta_1 x$ para algunos valores β_0, β_1 (la función de regresión verdadera es lineal)

contra

H_a: H_0 no es verdadera (la función de regresión verdadera es no lineal)

Suponga que las observaciones se hacen en $x_1, x_2, \ldots, x_c$. Represente con $Y_{11}, Y_{12}, \ldots, Y_{1n_1}$ las n_1 observaciones cuando $x = x_1$; $\ldots$; $Y_{c1}, Y_{c2}, \ldots, Y_{cn_c}$ denotan las n_c observaciones cuando $x = x_c$. Con $n = \sum n_i$ (el número total de observaciones), SSE tiene $n - 2$ gl. Se divide SSE en dos partes, SSPE (error puro) y SSLF (falta de ajuste), como sigue:

$$\text{SSPE} = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i\cdot})^2$$
$$= \sum\sum Y_{ij}^2 - \sum n_i \bar{Y}_{i\cdot}^2$$
$$\text{SSLF} = \text{SSE} - \text{SSPE}$$

Las n_i observaciones en x_i contribuyen con $n_i - 1$ grados de libertad a SSPE, así que el número de grados de libertad para SSPE es $\sum(n_i - 1) = n - c$ y los grados de libertad para SSLF son $n - 2 - (n - c) = c - 2$. Sea MSPE = SSPE/$(n - c)$ y MSLF = SSLF/$(c - 2)$. Entonces se puede demostrar que mientras $E(\text{MSPE}) = \sigma^2$ ya sea que H_0 sea verdadera o no, $E(\text{MSLF}) = \sigma^2$ si H_0 es verdadera y $E(\text{MSLF}) > \sigma^2$ si H_0 es falsa.

Estadístico de prueba: $F = \dfrac{\text{MSLF}}{\text{MSPE}}$

Región de rechazo: $f \geq F_{\alpha, c-2, n-c}$

Los siguientes datos provienen del artículo "Changes in Growth Hormone Status Related to Body Weight of Growing Cattle" (*Growth*, 1977: 241-247), con x = peso corporal y y = rapidez de eliminación metabólica/peso corporal.

x	110	110	110	230	230	230	360
y	235	198	173	174	149	124	115
x	360	360	360	505	505	505	505
y	130	102	95	122	112	98	96

(Por lo tanto, $c = 4$, $n_1 = n_2 = 3$, $n_3 = n_4 = 4$.)

a. Pruebe H_0 contra H_a al nivel 0.05 con la prueba de falta de ajuste descrita antes.

b. ¿Una gráfica de dispersión de los datos indica que la relación entre x y y es lineal? ¿Cómo se compara esto con el resultado del inciso (a)? (En el artículo se utilizó una función de regresión no lineal.)

13.2 Regresión con variables transformadas

Un argumento teórico o el examen de las gráficas de diagnóstico de un análisis de regresión lineal podrían indican la necesidad de un modelo alternativo para el modelo probabilístico lineal $Y = \beta_0 + \beta_1 x + \epsilon$. En cualquier caso, es deseable decidirse por un modelo cuyos parámetros se pueden decidir con facilidad. Una clase importante de esta clase de modelos se especifica por medio de funciones que son "intrínsecamente lineales".

DEFINICIÓN

Una función que relaciona a y con x es **intrínsecamente lineal** si por medio de una transformación en x y/o y, o ambas, la función se puede expresar como $y' = \beta_0 + \beta_1 x'$, donde x' = la variable independiente transformada y y' = la variable dependiente transformada.

Cuatro de las funciones intrínsecamente lineales más útiles se dan en la tabla 13.1. En cada caso, la transformación apropiada es una transformación log, base 10 o logaritmo natural (base e),

Tabla 13.1 Funciones intrínsecamente lineales útiles*

Función	Transformación a linealizar	Forma lineal
a. Exponencial: $y = \alpha e^{\beta x}$	$y' = \ln(y)$	$y' = \ln(\alpha) + \beta x$
b. Potencia: $y = \alpha x^{\beta}$	$y' = \log(y), x' = \log(x)$	$y' = \log(\alpha) + \beta x'$
c. $y = \alpha + \beta \cdot \log(x)$	$x' = \log(x)$	$y = \alpha + \beta x'$
d. Recíproca: $y = \alpha + \beta \cdot \frac{1}{x}$	$x' = \frac{1}{x}$	$y = \alpha + \beta x'$

*Cuando aparece $\log(\cdot)$, se puede usar ya sea el logaritmo base 10 o el logaritmo base e.

o una transformación recíproca. Las gráficas representativas de las cuatro funciones se ilustran en la figura 13.3.

Para una relación de función exponencial, sólo y se transforma para lograr linealidad, mientras que para una relación de función de potencia, se transforman tanto x como y. Debido a que la variable x está en el exponente en una relación exponencial, y aumenta (si $\beta > 0$) o disminuye (si $\beta < 0$) mucho más rápido cuando x aumenta, que es el caso para la función de potencia, aunque en un intervalo corto de valores de x puede ser difícil diferenciar entre las dos funciones. Ejemplos de funciones que no son intrínsecamente lineales son $y = \alpha + \gamma e^{\beta x}$ y $y = \alpha + \gamma x^{\beta}$.

Las funciones intrínsecamente lineales conducen de manera directa a modelos probabilísticos que, aunque no son lineales en x como una función, tienen parámetros cuyos valores se estiman con facilidad por medio de mínimos cuadrados ordinarios.

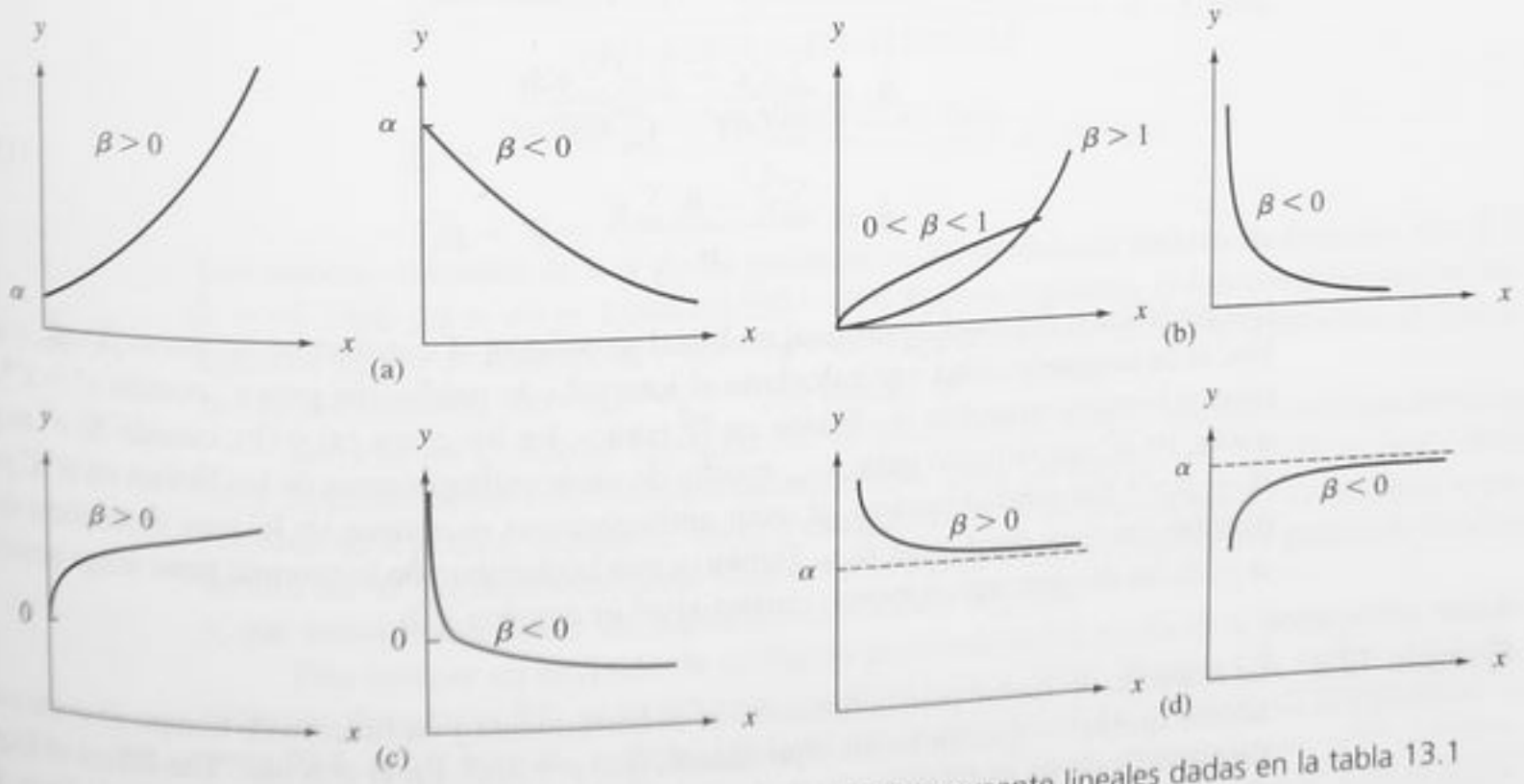


Figura 13.3 Gráficas de las funciones intrínsecamente lineales dadas en la tabla 13.1

DEFINICIÓN

> Un modelo probabilístico que relaciona a Y con x es **intrínsecamente lineal** si, por medio de una transformación en Y o x, o ambas, se puede reducir a un modelo probabilístico lineal $Y' = \beta_0 + \beta_1 x' + \epsilon'$.

Los modelos probabilísticos intrínsecamente lineales que corresponden a las cuatro funciones de la tabla 13.1 son como sigue:

a. $Y = \alpha e^{\beta x} \cdot \epsilon$, un modelo exponencial multiplicativo, así que $\ln(Y) = Y' = \beta_0 + \beta_1 x' + \epsilon'$ con $x' = x$, $\beta_0 = \ln(\alpha)$, $\beta_1 = \beta$ y $\epsilon' = \ln(\epsilon)$.

b. $Y = \alpha x^{\beta} \cdot \epsilon$, un modelo de potencia multiplicativo, de modo que $\log(Y) = Y' = \beta_0 + \beta_1 x' + \epsilon'$ con $x' = \log(x)$, $\beta_0 = \log(\alpha)$, $\beta_1 = \beta$ y $\epsilon' = \log(\epsilon)$.

c. $Y = \alpha + \beta \log(x) + \epsilon$, de modo que $x' = \log(x)$ linealiza de inmediato el modelo.

d. $Y = \alpha + \beta \cdot 1/x + \epsilon$, así que $x' = 1/x$ produce un modelo lineal.

Los modelos aditivos exponencial y de potencia, $Y = \alpha e^{\beta x} + \epsilon$ y $Y = \alpha x^{\beta} + \epsilon$, no son intrínsecamente lineales. Observe que tanto (a) como (b) requieren una transformación en y, y como resultado, una tranformacion en la variable de error ϵ. De hecho, si ϵ tiene una distribución lognormal (véase el capítulo 4) con $E(\epsilon) = e^{\sigma^2/2}$ y $V(\epsilon) = \tau^2$ independiente de x, entonces los modelos transformados para (a) y (b) satisfarán las suposiciones del capítulo 12 en relación con el modelo probabilístico lineal; esto a su vez implica que las inferencias para los parámetros del modelo transformado se basan en que serán válidas estas suposiciones. Si σ^2 es pequeña, $\mu_{Y \cdot x} \approx \alpha e^{\beta x}$ en (a) o αx^{β} en (b).

La ventaja principal de un modelo intrínsecamente lineal, es que los parámetros β_0 y β_1 del modelo transformado se pueden estimar de inmediato por medio del principio de mínimos cuadrados al sustituir x' y y' en las fórmulas de estimación

$$\hat{\beta}_1 = \frac{\sum x_i' y_i' - \sum x_i' \sum y_i'/n}{\sum (x_i')^2 - (\sum x_i')^2/n}$$

$$\hat{\beta}_0 = \frac{\sum y_i' - \hat{\beta}_1 \sum x_i'}{n} = \bar{y}' - \hat{\beta}_1 \bar{x}' \tag{13.5}$$

Los parámetros del modelo original no lineal se estiman al transformar de nuevo $\hat{\beta}_0$ o $\hat{\beta}_1$, o ambos, si es necesario. Una vez calculado el intervalo de predicción para y' cuando $x' = x'^*$, al invertir la transformación se obtiene un IP para y. En los casos (a) y (b), cuando Si σ^2 es pequeña, un IC aproximado para $\mu_{Y \cdot x^*}$ resulta de sacar antilogaritmos de los límites en el IC para $\beta_0 + \beta_1 x'^*$ (en sentido estricto, al sacar antilogaritmos se obtiene un IC para la *mediana* de la distribución Y, es decir, para $\tilde{\mu}_{Y \cdot x^*}$. Debido a que la distribución lognormal tiene sesgo positivo, $\mu > \tilde{\mu}$; las dos son más o menos iguales si σ^2 es cercana a 0).

Ejemplo 13.3 La ecuación de Taylor para la duración de la herramienta y en función de tiempo de corte x establece que $xy^c = k$ o, en forma equivalente, que $y = \alpha x^{\beta}$. En el artículo "The Effect of Experimental Error on the Determination of Optimum Metal Cutting Conditions" (*J. Engr. for Industry*, 1967: 315-322) se observa que la relación no es exacta (determinística) y que los parámetros α y β se deben estimar de los datos. Así, un modelo apropiado es el modelo de po-

tencia multiplicativo $Y = \alpha \cdot x^{\beta} \cdot \epsilon$, que el autor ajustó a los datos siguientes que consisten en 12 observaciones de la duración de herramientas de carburo (tabla 13.2). Además de los valores de x, y, x' y y' se proporcionan los valores transformados pronosticados ($\hat{y}'$) y los valores pronosticados de la escala original ($\hat{y}$, después de transformar de nuevo).

Tabla 13.2 Datos del ejemplo 13.3

	x	y	$x' = \ln(x)$	$y' = \ln(y)$	$\hat{y}'$	$\hat{y} = e^{\hat{y}'}$
1	600.	2.3500	6.39693	0.85442	1.12754	3.0881
2	600.	2.6500	6.39693	0.97456	1.12754	3.0881
3	600.	3.0000	6.39693	1.09861	1.12754	3.0881
4	600.	3.6000	6.39693	1.28093	1.12754	3.0881
5	500.	6.4000	6.21461	1.85630	2.11203	8.2650
6	500.	7.8000	6.21461	2.05412	2.11203	8.2650
7	500.	9.8000	6.21461	2.28238	2.11203	8.2650
8	500.	16.5000	6.21461	2.80336	2.11203	8.2650
9	400.	21.5000	5.99146	3.06805	3.31694	27.5760
10	400.	24.5000	5.99146	3.19867	3.31694	27.5760
11	400.	26.0000	5.99146	3.25810	3.31694	27.5760
12	400.	33.0000	5.99146	3.49651	3.31694	27.5760

Los estadísticos de resumen para ajustar una recta a los datos transformados son $\sum x_i' = 74.41200$, $\sum y_i' = 26.22601$, $\sum x_i'^2 = 461.75874$, $\sum y_i'^2 = 67.74609$ y $\sum x_i' y_i' = 160.84601$, por lo tanto

$$\hat{\beta}_1 = \frac{(160.84601) - (74.41200)(26.22601)/12}{461.75874 - (74.41200)^2/12} = -5.3996$$

$$\hat{\beta}_0 = \frac{26.22601 - (-5.3996)(74.41200)}{12} = 35.6684$$

Los valores estimados de α y β, los parámetros del modelo de función de potencia, son $\hat{\beta} = \hat{\beta}_1 = -5.3996$ y $\hat{\alpha} = e^{\hat{\beta}_0} = 3.094491530 \cdot 10^{15}$. Por consiguiente, la función de regresión estimada es $\mu_{Y \cdot x} \approx 3.094491530 \cdot 10^{15} \cdot x^{-5.3996}$. Para retomar la ecuación (estimada) de Taylor, se fija $y = 3.094491530 \cdot 10^{15} \cdot x^{-5.3996}$, de donde, $xy^{0.185} = 740$.

En la figura 13.4(a) de la página 568 se ilustra una gráfica de los residuos estandarizados de la regresión lineal con variables transformadas (para las cuales $r^2 = 0.922$); no hay patrón evidente en la gráfica, aunque un residuo estandarizado es un poco grande, y los residuos se parecen a los de una regresión lineal simple. En la figura 13.4(b) se ilustra una gráfica de $\hat{y}$ contra y, que indica predicciones satisfactorias en la escala original.

Para obtener un intervalo de confianza para la duración media de la herramienta cuando el tiempo de corte es 500, se transforma $x = 500$ a $x' = 6.21461$. Entonces $\hat{\beta}_0 + \hat{\beta}_1 x' = 2.1120$, y un 95% de IC para $\beta_0 + \beta_1(6.21461)$ es (de la sección 12.4) $2.1120 \pm (2.228)(0.0824) = (1.928, 2.296)$. El IC de 95% para $\tilde{\mu}_{Y \cdot 500}$ se obtiene al sacar antilogaritmos: $(e^{1.928}, e^{2.296}) = (6.876, 9.930)$. Se comprueba sin dificultad que para los datos transformados $s^2 = \hat{\sigma}^2 = 0.081$. Debido a que éste es bastante pequeño (6.876, 9.930) es un intervalo aproximado para $\mu_{Y \cdot 500}$.

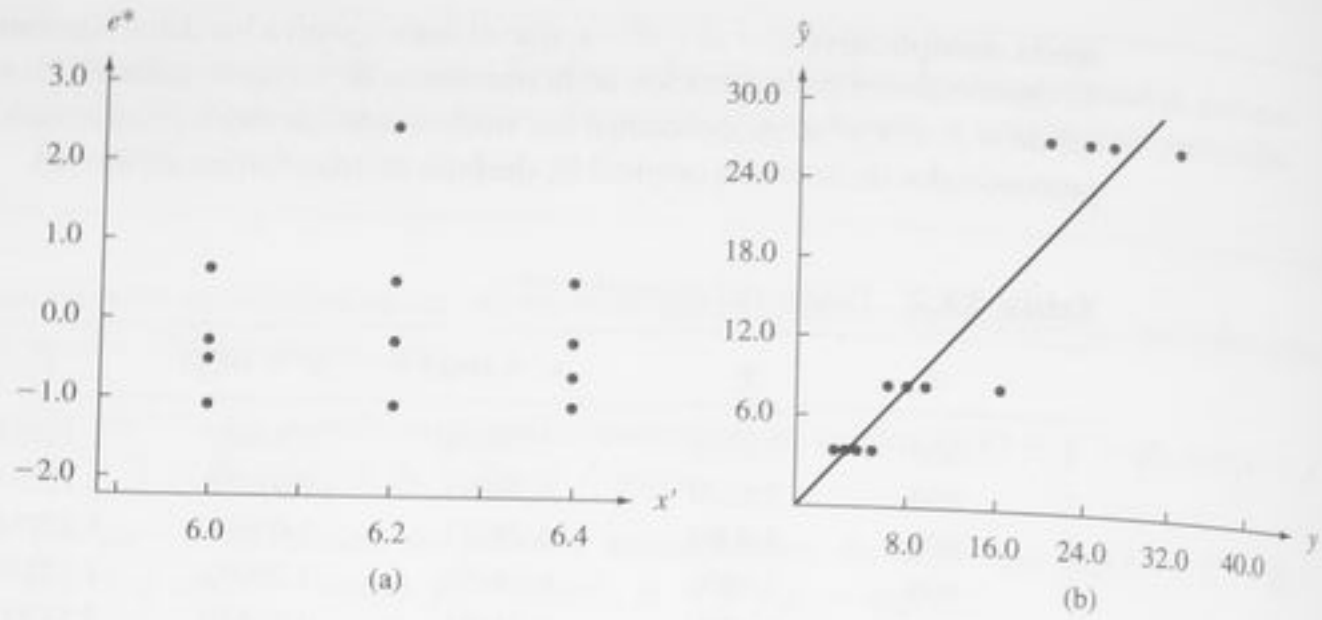


Figura 13.4 (a) Residuos estandarizados en función de x' del ejemplo 13.3; (b) $\hat{y}$ en función de y del ejemplo 13.3

■

Ejemplo 13.4 En el artículo "Ethylene Synthesis in Lettuce Seeds: Its Physiological Significance" (*Plant Physiology*, 1972: 719-722), se estudió el contenido de etileno de semillas de lechuga (y, en nL/g peso seco) en función del tiempo de exposición (x, en min) a un absorbente de etileno. En la figura 13.5 se ilustra un diagrama de dispersión de los datos y una gráfica de residuos generados de una regresión lineal de y sobre x. En ambas gráficas se observa un marcado patrón curvo, lo que hace pensar en que es apropiada una transformación para lograr linealidad. Además, una regresión lineal da predicciones negativas para $x = 90$ y $x = 100$.

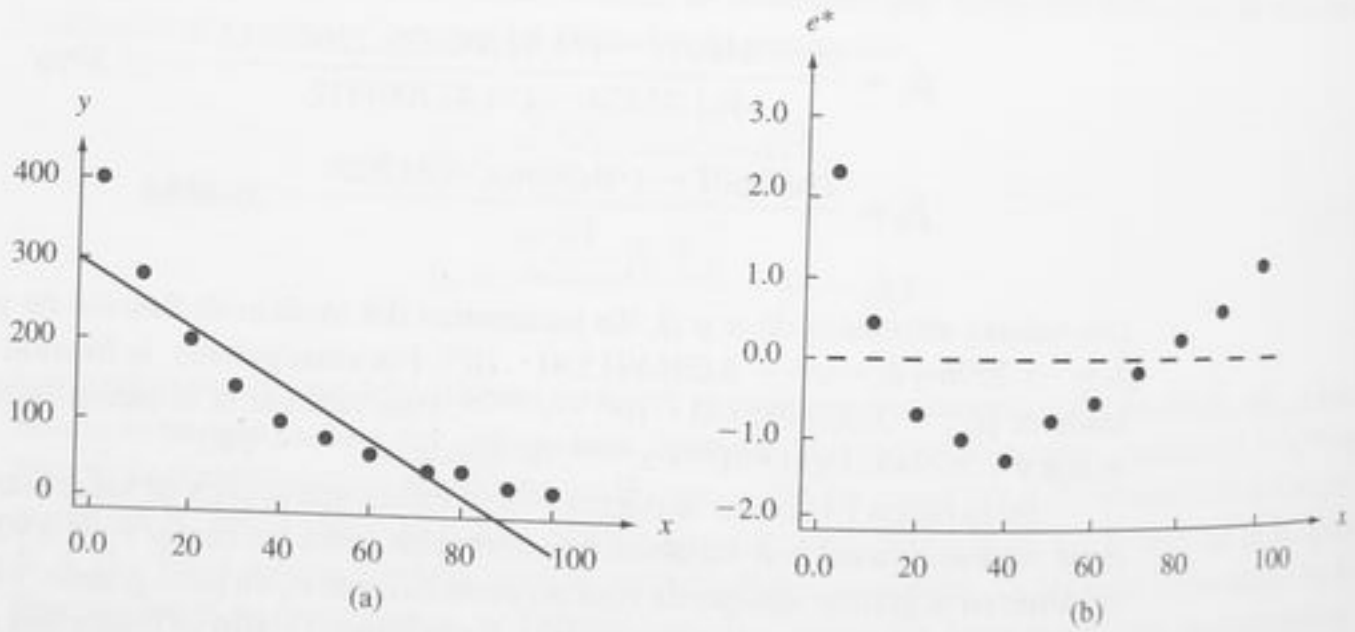


Figura 13.5 (a) Diagrama de dispersión; (b) gráfica residual de la regresión lineal para los datos del ejemplo 13.4

El autor no dio ningún argumento para el modelo teórico, pero esta gráfica de $y' = \ln(y)$ contra x muestra una fuerte relación lineal, de donde se deduce que una función exponencial proporciona un buen ajuste para los datos. En la tabla 13.3 se muestran los datos y otra información de una regresión lineal de y' sobre x. Las estimaciones de parámetros del modelo

Tabla 13.3 Datos del ejemplo 13.4

x	y	$y' = \ln(y)$	$\hat{y}'$	$\hat{y} = e^{\hat{y}'}$
2	408	6.01	5.876	353.32
10	274	5.61	5.617	275.12
20	196	5.28	5.294	199.12
30	137	4.92	4.971	144.18
40	90	4.50	4.647	104.31
50	78	4.36	4.324	75.50
60	51	3.93	4.001	54.64
70	40	3.69	3.677	39.55
80	30	3.40	3.354	28.62
90	22	3.09	3.031	20.72
100	15	2.71	2.708	15.00

lineal son $\hat{\beta}_1 = -0.0323$ y $\hat{\beta}_0 = 5.941$, con $r^2 = 0.995$. La función de regresión estimada para el modelo exponencial es $\hat{\mu}_{Y \cdot x} \approx e^{\hat{\beta}_0} \cdot e^{\hat{\beta}_1 x} = 380.32e^{-.0323x}$. Los valores pronosticados $\hat{y}_i$ se obtienen por sustitución de x_i $(i = 1, \ldots, n)$ en $\hat{\mu}_{Y \cdot x}$ o al calcular $y_i = e^{\hat{y}_i'}$, donde las $\hat{y}_i'$ son las predicciones del modelo de línea recta transformado. En la figura 13.6 se ilustra una gráfica de e'^* contra x (los residuos estandarizados de una regresión lineal) y una gráfica de $\hat{y}$ contra y. Estas gráficas respaldan la elección de un modelo exponencial.

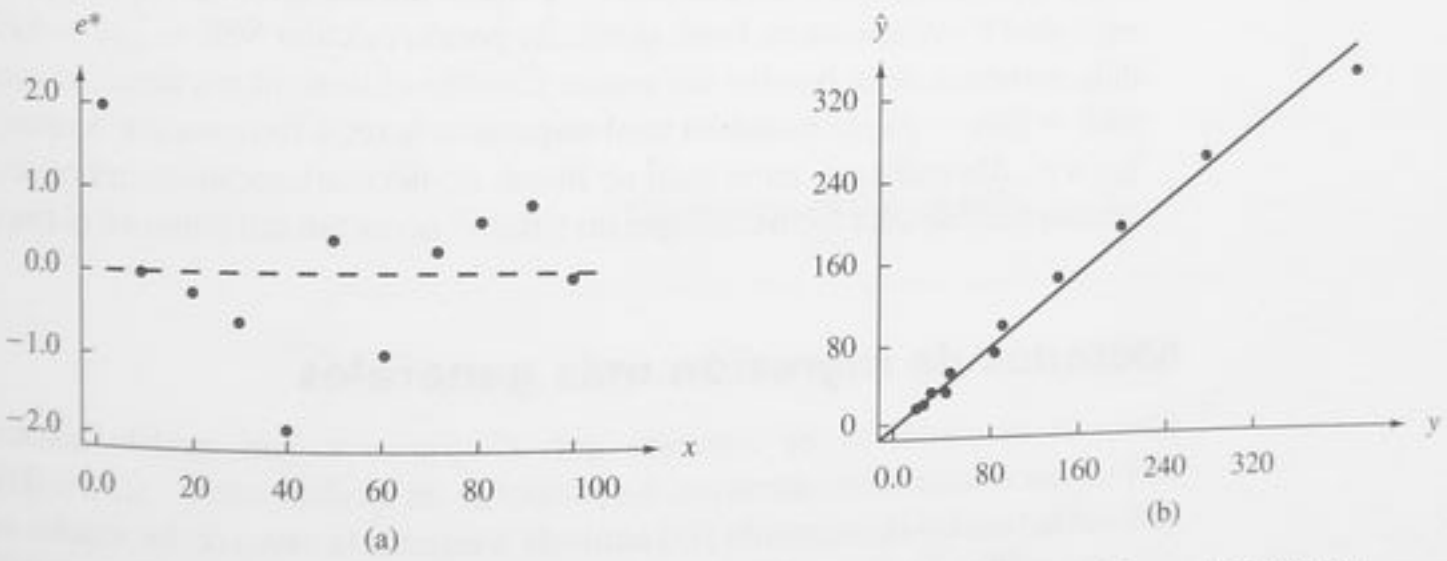


Figura 13.6 Gráfica de (a) residuos estandarizados (después de transformar) en función de x; (b) $\hat{y}$ en función de y para los datos del ejemplo 13.4 ■

Al analizar los datos transformados, se debe tener en mente los siguientes puntos:

1. Estimar β_1 y β_0 como en (13.5) y luego transformar de vuelta para obtener estimaciones de los parámetros originales no es equivalente a usar el principio de mínimos cuadrados directamente en el modelo original. Así, para el modelo experimental, se podría estimar α y β al minimizar $\sum(y_i - \alpha e^{\beta x_i})^2$. Sería necesario un cálculo iterativo. Las estimaciones resultantes no serían iguales: $\hat{\alpha} \neq e^{\hat{\beta}_0}$ y $\hat{\beta} \neq \hat{\beta}_1$.

2. Si el modelo elegido no es intrínsecamente lineal, no se puede usar el método resumido en (13.5). En cambio, se tendrían que aplicar mínimos cuadrados (o algún otro procedimiento de ajuste) al modelo no transformado. Así, para el modelo exponencial aditivo $Y = \alpha e^{\beta x} + \epsilon$, los mínimos cuadrados requerirían minimizar $\sum(y_i - \alpha e^{\beta x_i})^2$. Tomar derivadas parciales con respecto a α y β da como resultado dos ecuaciones normales no lineales en α y β; estas ecuaciones se deben resolver por medio de un procedimiento iterativo.

3. Cuando el modelo lineal transformado satisface las suposiciones listadas en el capítulo 12, el método de mínimos cuadrados produce mejores estimaciones de los parámetros transformados. Sin embargo, las estimaciones de los parámetros originales podrían no ser mejores en ningún sentido, aunque sean razonables. Por ejemplo, en el modelo exponencial, el estimador $\hat{\alpha} = e^{\hat{\beta}_0}$ no es insesgado, aunque será el estimador de máxima verosimilitud de α si la variable de error ϵ' tiene una distribución normal. Si se emplean de forma directa mínimos cuadrados (sin transformar) se obtendrían mejores estimaciones, aunque los cálculos serían bastante trabajosos.

4. Si se ha hecho una transformación en y y se desea usar fórmulas estándar para probar hipótesis o construir IC, ϵ' debe tener por lo menos una distribución normal aproximada. Para comprobar esto, se deben examinar los residuos de la regresión transformada.

5. Cuando se transforma y, la r^2 de la regresión resultante se refiere a la variación en las y'_i que se explica mediante el modelo de regresión transformado. Aunque aquí un valor alto de r^2 indica un buen ajuste de modelo no lineal original estimado para las y_i observadas, r^2 no se refiere a estas observaciones originales. Quizá la mejor forma de evaluar la calidad del ajuste es calcular los valores predichos $\hat{y}'_i$ con el modelo transformado, transformarlos de vuelta a la escala original y para obtener $\hat{y}_i$ y luego dibujar $\hat{y}$ contra y. Los puntos cercanos a la recta de 45° evidencian un buen ajuste. Se podría calcular $\text{SSE} = \sum(y_i - \hat{y}_i)^2$ como una medida numérica de la bondad del ajuste. Cuando el modelo era lineal, se comparó esto con $\text{SST} = \sum(y_i - \bar{y})^2$, la variación total respecto a la recta horizontal a la altura $\bar{y}$; esto dio lugar a r^2. Sin embargo, en el caso no lineal, no necesariamente es informativo medir la variación total de esta forma, así que un valor r^2 no es tan útil como en el caso lineal.

Métodos de regresión más generales

Hasta el momento se ha supuesto que $Y = f(x) + \epsilon$ (un modelo aditivo) o que $Y = f(x) \cdot \epsilon$ (un modelo multiplicativo). En el caso de un modelo aditivo, $\mu_{Y \cdot x} = f(x)$, de modo que estimar la función de regresión $f(x)$ equivale a estimar la curva de los valores y medios. A veces, una gráfica de dispersión de los datos indica que no hay expresión matemática simple para $f(x)$. Los estadísticos han elaborado en fechas recientes algunos métodos más flexibles que permiten modelar una amplia variedad de patrones con el mismo procedimiento de ajuste. Un método de este tipo es **LOWESS** (o LOESS), que es la abreviatura para *locally weighted scatter plot smoother*. Represente con (x^*, y^*) uno de los n pares (x, y) de la muestra. El valor $\hat{y}$ que corresponde a (x^*, y^*) se obtiene al ajustar una recta con sólo un porcentaje especificado de los datos (p. ej., 25%) cuyos valores x están cercanos a x^*. Además, en vez de usar los mínimos cuadrados "ordinarios", lo cual da igual peso a todos los puntos, los que tienen valores de x cercanos a x^* reciben más peso que aquéllos con valores de x apartados. La altura de la recta resultante arriba de x^* es el valor ajustado de $\hat{y}^*$. Este proceso se repite para cada uno de los n puntos, así que se ajustan n rectas diferentes (es seguro que el lector no querrá hacer esto a mano). Por último, se unen los puntos ajustados para obtener una curva LOWESS.

Ejemplo 13.5 Pesar grandes animales muertos encontrados en selvas por lo común no es factible, así que es adecuado tener un método para estimar el peso a partir de varias características de un animal que se pueden determinar con facilidad. MINITAB tiene almacenado un conjunto de datos que consiste en varias características para una muestra de $n = 143$ osos salvajes. En la figura 13.7(a) se muestra una gráfica de dispersión de y = peso en función de x = distancia alrededor del pecho (circunferencia alrededor del pecho). En primera instancia, parece como si una sola línea obtenida de mínimos cuadrados ordinarios resumiera de manera efectiva el patrón. En la figura 13.7(b) se muestra la curva LOWESS que produce MINITAB por medio de un intervalo de 50% [el ajuste en (x^*, y^*) se determina por el 50% más próximo de la muestra]. Al parecer, la curva consiste en dos segmentos de recta unidos arriba de aproximadamente $x = 38$. La recta más inclinada está a la derecha de 38, lo cual indica que el peso tiende a aumentar con más rapidez a medida que crece la circunferencia del pecho para perímetros que pasan de 38 pulgadas.

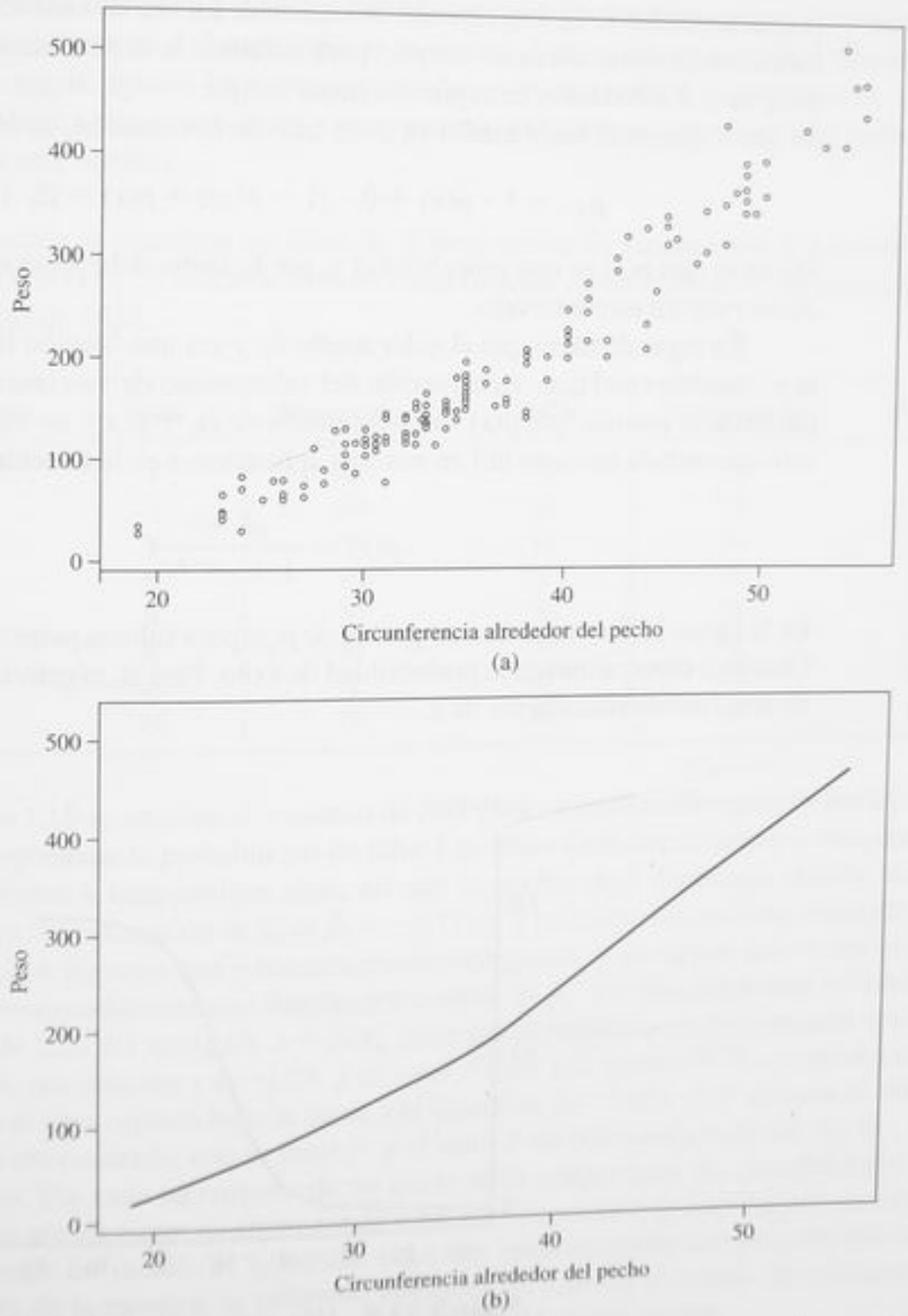


Figura 13.7 Gráfica de dispersión de MINITAB y curva LOWESS para los datos de peso de osos ■

Es complicado hacer inferencias (p. ej., obtener un IC para un valor y medio) con base en este tipo general de modelo de regresión. La técnica de inicialización mencionada antes se puede usar para este propósito.

Regresión logística

El modelo de regresión lineal simple es apropiado para relacionar una variable de respuesta cuantitativa y con un predictor cuantitativo x. Suponga que y es una variable dicótoma con valores posibles 1 y 0 que corresponden a éxito y fracaso. Sea $p = P(S) = P(y = 1)$. Con frecuencia, el valor de p dependerá del valor de cierta variable cuantitativa x. Por ejemplo, la probabilidad de que un automóvil requiera servicio de garantía de cierta clase podría depender del kilometraje del automóvil, o la probabilidad de evitar una infección de cierta clase podría depender de la dosis en una inoculación. En vez de usar sólo el símbolo p para la probabilidad de éxito, ahora se usa $p(x)$ para enfatizar la dependencia de esta probabilidad en el valor de x. La ecuación de regresión lineal simple $Y = \beta_0 + \beta_1 x + \epsilon$ ya no es apropiada, dado que al tomar el valor medio en cada lado de la ecuación, se obtiene

$$\mu_{Y \cdot x} = 1 \cdot p(x) + 0 \cdot (1 - p(x)) = p(x) = \beta_0 + \beta_1 x$$

Mientras que $p(x)$ es una probabilidad y, por lo tanto, debe estar entre 0 y 1, $\beta_0 + \beta_1 x$ no necesita estar en este intervalo.

En lugar de hacer que el valor medio de y sea una función lineal de x, ahora se considera un modelo en el que cierta función del valor medio de y es una función lineal de x. En otras palabras, se permite que $p(x)$ sea una función de $\beta_0 + \beta_1 x$ y no $\beta_0 + \beta_1 x$. Una función que ha sido encontrada bastante útil en muchas aplicaciones es la **función logit**

$$p(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

En la figura 13.8 se muestra una gráfica de $p(x)$ para valores particulares de β_0 y β_1 con $\beta_1 > 0$. Cuando x crece, aumenta la probabilidad de éxito. Para β_1 negativa, la probabilidad de éxito sería una función decreciente de x.

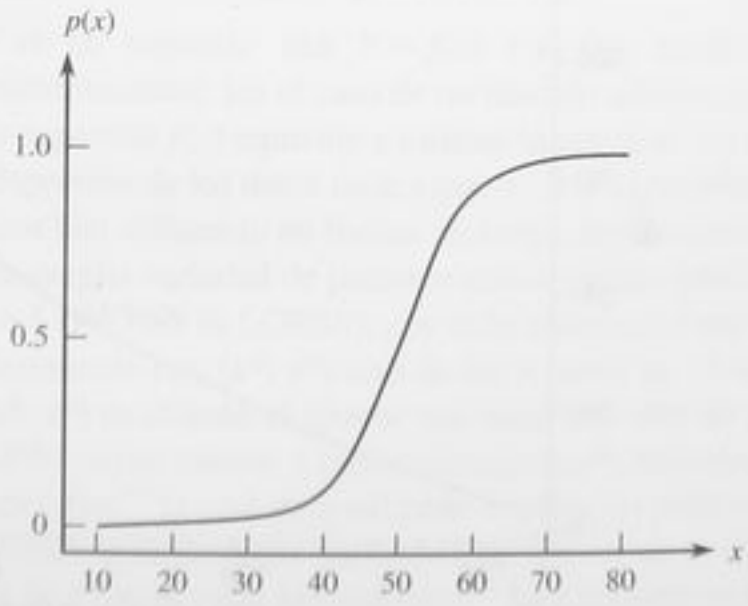


Figura 13.8 Una gráfica de una función logit

La *regresión logística* significa suponer que $p(x)$ se relaciona con x mediante la función logit. Con un poco de álgebra se muestra que

$$\frac{p(x)}{1 - p(x)} = e^{\beta_0 + \beta_1 x}$$

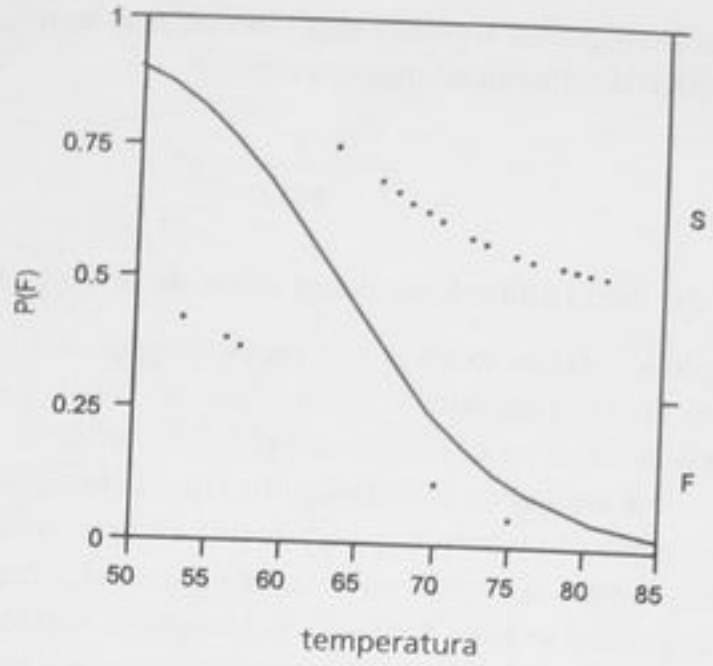
La expresión del lado izquierdo se llama *razón de ventaja*. Si, por ejemplo, $\frac{p(60)}{1 - p(60)} = 3$, entonces cuando $x = 60$ un éxito es tres veces tan probable como un fracaso. Ahora se ve que el logaritmo de la razón de ventaja es una función lineal del predictor. En particular, el parámetro de la pendiente β_1 es el cambio en el logaritmo de las probabilidades relacionadas con un incremento de 1 unidad en x. Esto significa que la relación de probabilidades cambia por el factor multiplicativo e^{β_1} cuando x se incrementa en una unidad.

Ajustar la regresión logística a los datos muestrales requiere estimar los parámetros β_0 y β_1. Esto por lo general se hace mediante la técnica de máxima verosimilitud descrita en el capítulo 6. Los detalles son bastante complicados, pero por fortuna los paquetes de software de estadística hacen esto si se solicita y proporcionan indicaciones cuantitativas y gráficas de cuan bien ajusta este modelo.

Ejemplo 13.6 A continuación se muestran los datos de la temperatura de lanzamiento y la incidencia de falla para anillos O en 24 lanzamientos del transbordador espacial antes del desastre del *Challenger* de enero de 1986.

Temperatura	Falla	Temperatura	Falla	Temperatura	Falla
53	Y	68	N	75	N
56	Y	69	N	75	Y
57	Y	70	N	76	N
63	N	70	Y	76	N
66	N	70	Y	78	N
67	N	70	Y	79	N
67	N	72	N	80	N
67	N	73	N	81	N

En la figura 13.9 se muestra el resultado de JMP para un análisis de regresión logístico. Se elige que p represente la probabilidad de falla. Las fallas tendieron a ocurrir a temperaturas bajas y los éxitos a temperaturas altas, así que la gráfica de $\hat{p}$ disminuye cuando aumenta la temperatura. La estimación de β_1 es $\hat{\beta}_1 = -0.1713$, y la desviación estándar estimada de β_1 es $s_{\hat{\beta}_1} = 0.08344$. Siempre que n sea suficientemente grande, y se supone que en este caso así es, $\hat{\beta}_1$ tiene aproximadamente una distribución normal. Si $\beta_1 = 0$ (la temperatura no afecta la probabilidad de falla del anillo O), $z = \hat{\beta}_1 / s_{\hat{\beta}_1}$ tiene una distribución normal estándar aproximada. El valor de esta relación z es -2.05, y el valor P para una prueba de dos extremos es 0.0404 (dos veces el área captada bajo la curva a la izquierda de -2.05). JMP informa el valor de un estadístico chi-cuadrada, que es justo z^2, y el valor P de chi-cuadrada difiere del de z sólo por el redondeo. Por cada incremento de un grado en la temperatura, las probabilidades de falla disminuyen por un factor de alrededor de 0.84. La temperatura de lanzamiento para la misión del *Challenger* fue de sólo 31°F. Debido a que este valor es mucho más pequeño que cualquier temperatura de la muestra, es peligroso extrapolar la relación estimada. No obstante, parece que para una temperatura tan baja, la falla del anillo O es casi segura.



Estimaciones de parámetro

Término	Estimación	Error est.	Chi cuadrada	Prob. > chi cuadrada
Ordenada al origen	10.8753321	5.7031291	3.64	0.0565
Temperatura	–0.1713202	0.0834419	4.22	0.0401

Figura 13.9 Resultado de JMP para la regresión logística

Ejercicios | Sección 13.2 (15 a 25)

15. A ningún aficionado a los nachos le gustan blandos, así que es importante hallar las características del proceso de producción con las que se obtiene una textura apetitosa. Los siguientes datos de x = tiempo de freído (s) y y = contenido de humedad (%) aparecieron en el artículo "Thermal and Physical Properties of Tortilla Chips as a Function of Frying Time" (*J. of Food Processing and Preservation*, 1995: 175-189).

x	5	10	15	20	25	30	45	60
y	16.3	9.7	8.1	4.2	3.4	2.9	1.9	1.3

a. Trace un diagrama de dispersión de y en función de x y comente acerca de características importantes.

b. Construya un diagrama de dispersión de los pares $(\ln(x), \ln(y))$ y comente.

c. ¿Qué relación probabilística entre x y y indica el patrón lineal en la gráfica del inciso (b)?

d. Prediga el valor del contenido de humedad cuando el tiempo de freído sea 20, de manera que lleve información acerca de la confiabilidad y la precisión.

e. Analice los residuos obtenidos de ajustar el modelo de regresión lineal simple a los datos transformados y comente.

16. En el artículo "The Luminosity-Spectral Index Relationship for Radio Galaxies" (*Nature*, 1972: 88-89) se sugiere que para las galaxias clase S, $\ln(L_{178})$ se relacionan linealmente con el índice espectral, donde L_{178} representa la luminosidad a 178 MHz. Los datos representativos se muestran a continuación:

Índice espectral	0.59	0.67	0.72	0.80	0.85	0.90
$\ln(L_{178})$	23.6	25.6	26.4	25.7	26.8	26.7
Índice espectral	0.94	0.66	1.00	0.86	1.03	0.70
$\ln(L_{178})$	27.0	24.9	27.1	27.2	26.9	25.2

a. Estime los parámetros del modelo exponencial que implica la relación lineal entre $\ln(L_{178})$ y el índice espectral.

b. ¿Qué valor de L_{178} pronosticaría para un índice espectral de 0.75?

c. Calcule un intervalo de pronóstico de 95% para la luminosidad cuando el índice espectral es 0.95.

17. Los siguientes datos sobre la rapidez de combustión x y la longitud de flama y son representativos de lo que apareció en el artículo "Some Burning Characteristics of Filter Paper" (*Combustion Science and Technology*, 1971: 103-120):

x	1.7	2.2	2.3	2.6	2.7	3.0	3.2
y	1.3	1.8	1.6	2.0	2.1	2.2	3.0

x	3.3	4.1	4.3	4.6	5.7	6.1
y	2.6	4.1	3.7	5.0	5.8	5.3

a. Estime los parámetros de un modelo de función de potencia.

b. Construya gráficas de diagnóstico para comprobar si una función de potencia es una selección apropiada de modelo.

c. Pruebe H_0: $\beta = \frac{4}{3}$ contra H_a: $\beta < \frac{4}{3}$, con una prueba de nivel 0.05.

d. Pruebe la hipótesis nula que expresa que la longitud de flama mediana cuando la rapidez de combustión es 5.0 es el doble de la longitud de flama mediana cuando la rapidez de combustión es 2.5 contra la alternativa de que éste no es el caso.

18. Las fallas en los motores de las turbinas de gas de aviones debidas a la fatiga de ciclo alto es un problema general. En el artículo "Effect of Crystal Orientation on Fatigue Failure of Single Crystal Nickel Base Turbine Blade Superalloys (*J. of Engineering for Gas Turbines and Power*, 2002: 161-176) aparecen los datos siguientes y el ajuste de un modelo de regresión no lineal, a fin de predecir la amplitud de esfuerzo de los ciclos hasta la falla. Ajuste un modelo apropiado, investigue la calidad del ajuste y prediga la amplitud cuando los ciclos hasta la falla son 5000.

Obs	Ciclos hasta la Falla	Amplitud del esfuerzo	Obs	Ciclos hasta la Falla	Amplitud del esfuerzo
1	1326	0.01495	11	7356	0.00576
2	1593	0.01470	12	7904	0.00580
3	4414	0.01100	13	79	0.01212
4	5673	0.01190	14	4175	0.00782
5	29 516	0.00873	15	34 676	0.00596
6	26	0.01819	16	114 789	0.00600
7	843	0.00810	17	2672	0.00880
8	1016	0.00801	18	7532	0.00883
9	3410	0.00600	19	30 220	0.00676
10	7101	0.00575			

19. Las pruebas de resistencia térmica se realizaron para estudiar la relación entre la temperatura y la duración de alambre recubierto con poliéster ("Thermal Endurance of Polyester Enameled Wires Using Twisted Wire Specimens", *IEEE Trans. Insulation*, 1965: 38-44), y se obtuvieron los siguientes datos.

Temperatura	200	200	200	200	200	200
Duración	5933	5404	4947	4963	3358	3878
Temperatura	220	220	220	220	220	220
Duración	1561	1494	747	768	609	777
Temperatura	240	240	240	240	240	240
Duración	258	299	209	144	180	184

a. ¿Una gráfica de dispersión de los datos indica una relación probabilística lineal entre la duración y la temperatura?

b. ¿Qué modelo implica una relación lineal entre ln(duración) y 1/temperatura esperados? ¿Una gráfica de dispersión de los datos transformados al parecer es consistente con esta relación?

c. Estime los parámetros del modelo sugerido en el inciso (b). ¿Qué duración predeciría para una temperatura de 220?

d. Debido a que hay múltiples observaciones en cada valor x, el método del ejercicio 14 se puede usar para probar la hipótesis nula que expresa que el modelo sugerido en el inciso (b) es correcto. Realice la prueba al nivel 0.01.

20. En el ejercicio 14 se presentan los datos del peso corporal x y la tasa de eliminación metabólica/peso corporal y. Considere las siguientes funciones intrínsecamente lineales para especificar la relación entre las dos variables: (a) $\ln(y)$ en función de x, (b) $\ln(y)$ en función de $\ln(x)$, (c) y en función de $\ln(x)$, (d) y en función de $1/x$ y (e) $\ln(y)$ en función de $1/x$. Utilice las gráficas de diagnóstico y análisis apropiados para decidir cuál de estas funciones elegiría para especificar un modelo probabilístico. Explique su razonamiento.

21. Una gráfica en el artículo "Thermal Conductivity of Polyethylene: The Effects of Crystal Size, Density, and Orientation on the Thermal Conductivity" (*Polymer Eng. and Science*, 1972: 204-208) indica que el valor esperado de la conductividad térmica y es una función lineal de $10^4 \cdot 1/x$, donde x es el espesor laminar.

x	240	410	460	490	520	590	745	8300
y	12.0	14.7	14.7	15.2	15.2	15.6	16.0	18.1

a. Estime los parámetros de la función de regresión y la función de regresión.
b. Prediga el valor de la conductividad térmica cuando el espesor laminar es 500Å.

22. En cada uno de los siguientes casos determine si la función es intrínsecamente lineal. En caso afirmativo, identifique x' y y' y después explique cómo se puede introducir el término de error aleatorio ϵ para obtener un modelo probabilístico intrínsecamente lineal.
a. $y = 1/(\alpha + \beta x)$
b. $y = 1/(1 + e^{\alpha+\beta x})$
c. $y = e^{e^{\alpha+\beta x}}$ (una curva de Gompertz)
d. $y = \alpha + \beta e^{\beta x}$

23. Suponga que x y y están relacionadas según un modelo exponencial probabilístico $Y = \alpha e^{\beta x} \cdot \epsilon$, con $V(\epsilon)$, una constante independiente de x (como fue el caso en el modelo lineal simple $Y = \beta_0 + \beta_1 x + \epsilon$). Es $V(Y)$ una constante independiente de x [como fue el caso para $Y = \beta_0 + \beta_1 x + \epsilon$, donde $V(Y) = \sigma^2$]? Explique su razonamiento. Dibuje una figura de la gráfica de dispersión prototipo a partir de este modelo. Conteste las mismas preguntas para el modelo de potencia $Y = \alpha x^{\beta} \cdot \epsilon$.

24. La cifosis se refiere a una flexión grave de la espina hacia delante después de una cirugía espinal correctiva. Las siguientes edades (en meses) para 40 personas al momento de la operación son de un estudio que se realizó a fin de determinar los factores de riesgo para la cifosis; los primeros 18 individuos tuvieron cifosis y los 22 restantes no.

Con cifosis	12	15	42	52	59	73
	82	91	96	105	114	120
	121	128	130	139	139	157
Sin cifosis	1	1	2	8	11	18
	22	31	37	61	72	81
	97	112	118	127	131	140
	151	159	177	206		

Utilice el siguiente resultado de regresión logística de MINITAB para decidir si al parecer la edad tiene un impacto importante en la presencia de cifosis.

25. Una gran empresa de asesoría administrativa obtuvo los siguientes datos en un estudio que realizó para investigar la relación entre la cantidad de experiencia en el puesto (meses) para un consultor principiante y la posibilidad de que el consultor pueda desempeñar cierta tarea compleja.

Éxito	8	13	14	18	20	21	21	22	25
	26	28	29	30	32				
Fracaso	4	5	6	6	7	9	10	11	11
	13	15	18	19	20	23	27		

Interprete el siguiente resultado de regresión logística de MINITAB y trace una gráfica de la probabilidad estimada de desempeño en la tarea en función de la experiencia.

Tabla de regresión logística para el ejercicio 24

Predictor	Coef	Desv. est.	Z	P	Razón de ventaja	95% Inferior	IC superior
Constante	−0.5727	0.6024	−0.95	0.342			
Edad	0.004296	0.005849	0.73	0.463	1.00	0.99	1.02

Tabla de regresión logística para el ejercicio 25

Predictor	Coef	Desv. est.	Z	P	Razón de ventaja	95% Inferior	IC superior
Constante	−3.211	1.235	−2.60	0.009			
Edad	0.17772	0.06573	2.70	0.007	1.19	1.05	1.36

13.3 Regresión polinomial

Los modelos no lineales, pero intrínsecamente lineales, de la sección 13.2, tienen que ver con funciones de la variable independiente x que fueron crecientes o decrecientes en sentido estricto. En muchas situaciones, el razonamiento teórico o una gráfica de dispersión de los datos indica que la función de regresión verdadera $\mu_{Y \cdot x}$ tiene uno o más picos o valles, es decir, por lo menos un mínimo o un máximo. En estos casos, una función polinomial $y = \beta_0 + \beta_1 x + \cdots + \beta_k x^k$ proporcionaría una aproximación satisfactoria para la función de regresión verdadera.

DEFINICIÓN:

La ecuación modelo de regresión polinomial de k-ésimo grado es

$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k + \epsilon \tag{13.6}$$

donde ϵ es una variable aleatoria distribuida normalmente con

$$\mu_\epsilon = 0 \qquad \sigma_\epsilon^2 = \sigma^2 \tag{13.7}$$

De (13.6) y (13.7), se deduce de inmediato que

$$\mu_{Y \cdot x} = \beta_0 + \beta_1 x + \cdots + \beta_k x^k \qquad \sigma_{Y \cdot x}^2 = \sigma^2 \tag{13.8}$$

En palabras, el valor esperado de Y es una función polinomial de x de k-ésimo grado, mientras que la varianza de Y, que controla la dispersión de valores observados respecto a la función de regresión, es la misma para cada valor de x. Se supone que los pares observados $(x_1, y_1), \ldots, (x_n, y_n)$ se generaron de forma independiente del modelo (13.6). En la figura 13.10 se ilustra el modelo cuadrático y el cúbico.

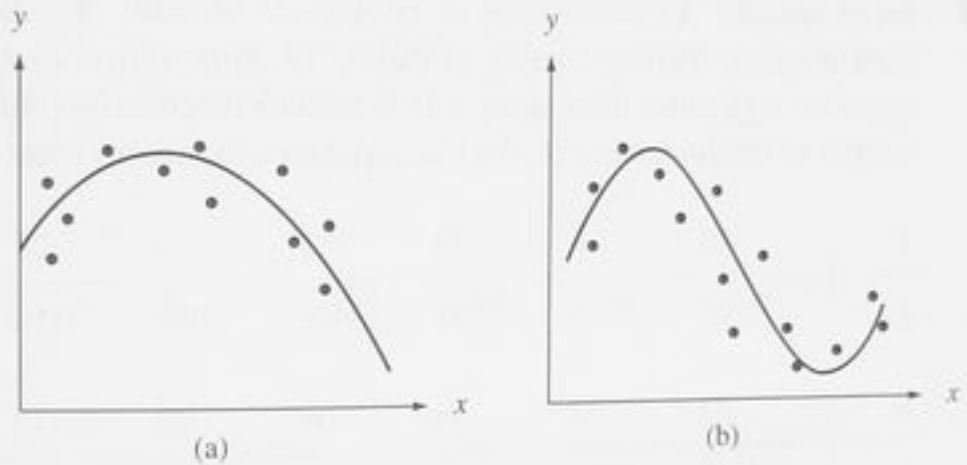


Figura 13.10 (a) Modelo de regresión cuadrático; (b) modelo de regresión cúbico.

Estimación de parámetros por medio de mínimos cuadrados

Para estimar $\beta_0, \beta_1, \ldots, \beta_k$, considere una función de regresión de ensayo $y = b_0 + b_1 x + \cdots + b_k x^k$. Entonces la bondad del ajuste de esta función para los datos observados se evalúa al calcular la suma de desviaciones cuadradas

$$f(b_0, b_1, \ldots, b_k) = \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_i + b_2 x_i^2 + \cdots + b_k x_i^k)]^2 \tag{13.9}$$

Según el principio de mínimos cuadrados, las estimaciones $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ son los valores de $b_0, b_1, \ldots, b_k$ que minimizan la expresión (13.9). Se debe notar que cuando $x_1, x_2, \ldots, x_n$ son diferentes, hay un polinomio de grado $n - 1$ que ajusta perfectamente los datos, así que el valor minimizante de (13.9) es 0 cuando $k = n - 1$. Sin embargo, en casi todas las aplicaciones, el modelo polinomial (13.6) con k grande es bastante irreal, y en la mayor parte de las aplicaciones $k = 2$ (cuadrático) o $k = 3$ (cúbico) es apropiada.

Para hallar los valores minimizantes de (13.9), se toman las $k + 1$ derivadas parciales $\partial f/\partial b_0, \partial f/\partial b_1, \ldots, \partial f/\partial b_k$ y se igualan a 0, lo que da como resultado el sistema de ecuaciones

normales para las estimaciones. Debido a que la función de prueba $b_0 + b_1x + \cdots + b_kx^k$ es lineal en $b_0, \ldots, b_k$ (aunque no en x), las $k + 1$ ecuaciones normales son lineales en estas incógnitas:

$$\begin{aligned} b_0 n + b_1 \sum x_i + b_2 \sum x_i^2 + \cdots + b_k \sum x_i^k &= \sum y_i \\ b_0 \sum x_i + b_1 \sum x_i^2 + b_2 \sum x_i^3 + \cdots + b_k \sum x_i^{k+1} &= \sum x_i y_i \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots & \\ b_0 \sum x_i^k + b_1 \sum x_i^{k+1} + \cdots + b_k \sum x_i^{2k} &= \sum x_i^k y_i \end{aligned} \tag{13.10}$$

Los paquetes de software para estadística normales resuelven de forma automática las ecuaciones de (13.10) y proporcionan las estimaciones así como mucha más información.*

Ejemplo 13.7 En el artículo "Determination of Biological Maturity and Effect of Harvesting and Drying Conditions on Milling Quality of Paddy" (*J. Agricultural Eng. Research*, 1975: 353-361) aparecen los siguientes datos acerca de la fecha x de cosecha (número de días después de la floración) y el rendimiento y (kg/ha) de arroz en cáscara, un grano cultivado en la India:

x	16	18	20	22	24	26	28	30
y	2508	2518	3304	3423	3057	3190	3500	3883

x	32	34	36	38	40	42	44	46
y	3823	3646	3708	3333	3517	3241	3103	2776

El diagrama de dispersión de la figura 13.11(a) indica que la elección de autor de un modelo cuadrático es razonable. En la figura 13.11(b) se ilustra el resultado de MINITAB de una regresión cuadrática, del cual

$$\hat{\beta}_0 = -1070.4 \qquad \hat{\beta}_1 = 293.48 \qquad \hat{\beta}_2 = -4.5358$$

y la función de regresión estimada es

$$y = -1070.4 + 293.48x - 4.5358x^2$$

Las funciones de regresión estimadas se pueden usar para calcular valores pronosticados (sustituya $x_1 = 16$, luego $x_2 = 18$, etc., para obtener $\hat{y}_1, \hat{y}_2, \ldots$) y luego los residuos $(y_i - \hat{y}_i)$. Una gráfica residual indica que una función de regresión cuadrática es apropiada, aunque hay un poco más de dispersión en los residuos para x pequeña que para x grande. Los mínimos cuadrados ponderados se usan para obtener un ajuste más satisfactorio, pero los autores no consideraron esto.

*En la sección 13.4 se verá que la regresión polinomial es un caso especial de la regresión múltiple, así que por lo común se emplea un comando apropiado para esta última tarea.

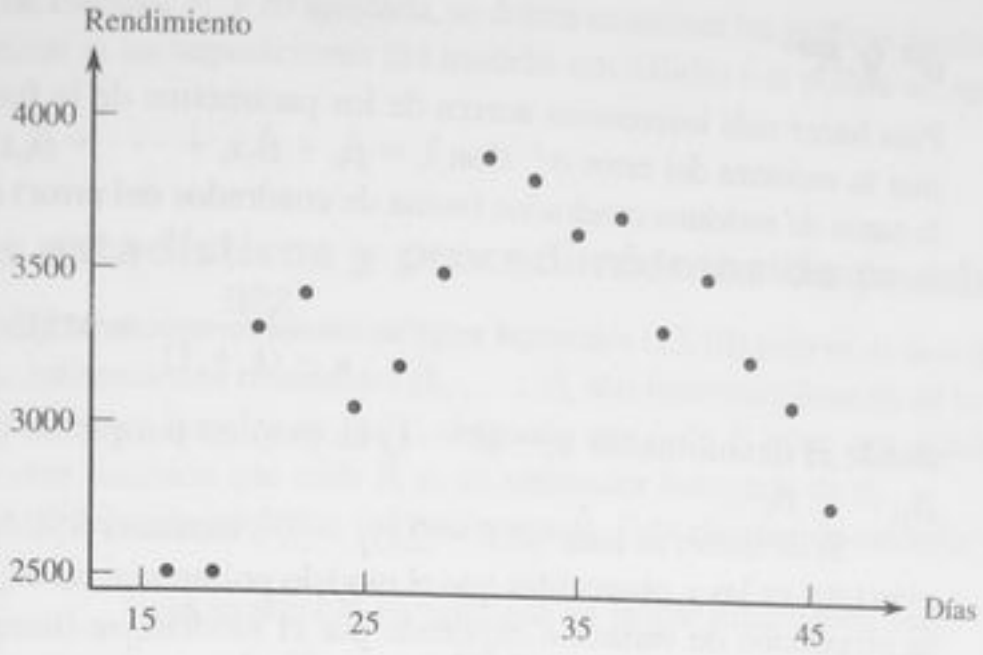


Figura 13.11(a) Diagrama de dispersión para el ejemplo 13.7

```
LA ECUACIÓN DE REGRESIÓN es RENDIMIENTO = -1070 + 293 DÍAS - 4.54
DÍAS AL CUADRADO

Predictor          Coef   Desv. est.   Relación t       p
Constante       -1070.4        617.3        -1.73   0.107
DÍAS             293.48        42.18         6.96   0.000
DÍAS AL
CUADRADO        -4.5358       0.6744        -6.73   0.000

s = 203.9  R cuadrada = 79.4%  R cuadrada (ajustada) = 76.2%

Análisis de varianza

FUENTE          GL          SC          CM         F       P
Regresión        2     2084779     1042389     25.08   0.000
Error           13      540388       41568
Total           15     2625167

                                      Ajuste                  Residuo
Obs.    DÍAS    Rendimiento  Ajuste  desv. est.  Residuo       est.
  1     16.0         2508.0  2464.2       135.6     43.8       0.29
  2     18.0         2518.0  2742.7       104.8   -224.7      -1.28
  3     20.0         3304.0  2984.9        83.0    319.1       1.71
  4     22.0         3423.0  3190.9        71.3    232.1       1.22
  5     24.0         3057.0  3360.6        68.4   -303.6      -1.58
  6     26.0         3190.0  3494.0        70.7   -304.0      -1.59
  7     28.0         3500.0  3591.1        74.2    -91.1      -0.48
  8     30.0         3883.0  3651.9        76.4    231.1       1.22
  9     32.0         3823.0  3676.4        76.4    146.6       0.78
 10     34.0         3646.0  3664.6        74.2    -18.6      -0.10
 11     36.0         3708.0  3616.6        70.7     91.4       0.48
 12     38.0         3333.0  3532.3        68.4   -199.3      -1.04
 13     40.0         3517.0  3411.6        71.3    105.4       0.55
 14     42.0         3241.0  3254.7        83.0    -13.7      -0.07
 15     44.0         3103.0  3061.5       104.8     41.5       0.24
 16     46.0         2776.0  2832.1       135.6    -56.1      -0.37

Ajuste      Ajuste de desv. est.      I.C. de 95%          I.P. de 95%
3431.8                  69.2     (3282.2, 3581.4)     (2966.5, 3897.1)
```

Figura 13.11(b) Resultado de MINITAB para el ejemplo 13.7 ■

$\hat{\sigma}^2$ y R^2

Para hacer más inferencias acerca de los parámetros de la función de regresión, se debe estimar la varianza del error σ^2. Con $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \cdots + \hat{\beta}_k x_i^k$ el i-ésimo residuo es $y_i - \hat{y}_i$, y la suma de residuos cuadrados (suma de cuadrados del error) es $\text{SSE} = \sum(y_i - \hat{y}_i)^2$. Entonces, la estimación de σ^2 es

$$\hat{\sigma}^2 = s^2 = \frac{\text{SSE}}{n - (k + 1)} = \text{MSE} \tag{13.11}$$

donde el denominador $n - (k + 1)$ es emplea porque se pierden $k + 1$ gl al estimar β_0, $\beta_1, \ldots, \beta_k$

Si de nuevo se hace $\text{SST} = \sum(y_i - \bar{y})^2$, entonces SSE/SST es la proporción de la variación total en las y_i observadas que el modelo polinomial no explica. La cantidad $1 - \text{SSE/SST}$, la proporción de variación explicada por el modelo, se llama **coeficiente de determinación múltiple** y se denota por R^2.

Suponga que se considera ajustar un modelo cúbico a los datos del ejemplo 13.7. Debido a que el modelo cúbico incluye al cuadrático como caso especial, el ajuste a un cúbico será por lo menos tan bueno como el ajuste a un cuadrático. En términos generales, con SSE_k = suma de cuadrados del error de un polinomio de k-ésimo grado, $\text{SSE}_{k'} \leq \text{SSE}_k$ y $R^2_{k'} \geq R^2_k$ siempre que $k' > k$. Debido a que el objetivo del análisis de regresión es encontrar un modelo que sea simple (relativamente pocos parámetros) y proporcione un buen ajuste a los datos, es posible que un polinomio de grado superior no especifique un mejor modelo que un modelo de grado de menor grado a pesar de su valor R^2 superior. Para equilibrar el costo de usar más parámetros contra la ganancia en R^2, muchos estadísticos emplean el **coeficiente ajustado de determinación múltiple**

$$R^2 \text{ ajustado} = 1 - \frac{n - 1}{n - (k + 1)} \cdot \frac{\text{SSE}}{\text{SST}} = \frac{(n - 1)R^2 - k}{n - 1 - k} \tag{13.12}$$

La R^2 ajustada ajusta la proporción de variación hacia arriba no explicada [puesto que la relación $(n - 1)/(n - k - 1)$ pasa de 1], lo cual da como resultado R^2 *ajustada* $< R^2$. Así, si $R_2^2 = 0.66$, $R_3^2 = 0.70$ y $n = 10$, entonces

$$R_2^2 \text{ ajustada} = \frac{9(0.66) - 2}{10 - 3} = 0.563 \qquad R_3^2 \text{ ajustada} = \frac{9(0.70) - 3}{10 - 4} = 0.550$$

de modo que la ganancia pequeña en R^2 al ir de un modelo cuadrático a un modelo cúbico es insuficiente para compensar el costo de agregar un parámetro extra al modelo.

Ejemplo 13.8 (Continuación del ejemplo 13.17)

SSE y SST por lo común se encuentran en un resultado de computadora en una tabla ANOVA. En la figura 13.11 se ilustra $\text{SSE} = 540\,388$ y $\text{SST} = 2\,625\,167$ para los datos de rendimiento del arroz con cáscara, de donde

$$R^2 = 1 - \frac{\text{SSE}}{\text{SST}} = 1 - \frac{540\,388}{2\,625\,167} = 1 - 0.206 = 0.794$$

Por lo tanto, 79.4% de la variación observada en el rendimiento se atribuye a la relación de modelo. La R^2 ajustada = 0.762, sólo un pequeño cambio descendente en R^2. Las estimaciones de σ^2 y σ son

$$s^2 = \frac{\text{SSE}}{n - (k + 1)} = \frac{540\,388}{16 - (2 + 1)} = 41\,568.3$$
$$s = \sqrt{s^2} = 203.9$$

■

Además de calcular R^2 y R^2 ajustada, se deben examinar las gráficas de diagnóstico usuales para determinar si las suposiciones del modelo son válidas o si podría ser apropiada la modificación.

Intervalos estadísticos y procedimientos de prueba

Debido a que las y_i aparecen en las ecuaciones normales (13.10) sólo en el lado derecho y en un modo lineal, las estimaciones resultantes $\hat{\beta}_0, \ldots, \hat{\beta}_k$ son funciones lineales de las y_i. Así, los estimadores son funciones lineales de las Y_i, de modo que cada $\hat{\beta}_i$ tiene una distribución normal. Se puede demostrar también que cada $\hat{\beta}_i$ es un estimador insesgado de β_i.

Sea $\sigma_{\hat{\beta}_i}$ la desviación estándar del estimador $\hat{\beta}_i$. Esta desviación estándar tiene la forma

$$\sigma_{\hat{\beta}_i} = \sigma \cdot \left\{ \begin{array}{c} \text{una expresión complicada en la que intervienen las} \\ x_j, x_j^2, \ldots, \text{y las } x_j^k \end{array} \right\}$$

Por fortuna, la expresión entre llaves ha sido programada en la mayor parte de los paquetes de software para estadística más comunes. La desviación estándar estimada de $\hat{\beta}_i$, $s_{\hat{\beta}_i}$ resulta de sustituir s en lugar de σ en la expresión para $\sigma_{\hat{\beta}_i}$. Estas desviaciones estándar estimadas $s_{\hat{\beta}_0}$, $s_{\hat{\beta}_1}, \ldots$ y $s_{\hat{\beta}_k}$ aparecen en el resultado de los paquetes de estadística antes mencionados. Sea $S_{\hat{\beta}_i}$ el estimador de $\sigma_{\hat{\beta}_i}$, es decir, la variable aleatoria cuyo valor observado es $s_{\hat{\beta}_i}$. Entonces, se puede demostrar que la variable estandarizada

$$T = \frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \tag{13.13}$$

tiene una distribución t basada en $n - (k + 1)$ gl. Esto da lugar a los siguientes procedimientos inferenciales.

Un IC de $100(1 - \alpha)\%$ para β_i el coeficiente de x^i en la función de regresión polinomial, es

$$\hat{\beta}_i \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{\beta}_i}$$

Una prueba de H_0: $\beta_i = \beta_{i0}$ se basa en el valor del estadístico t

$$t = \frac{\hat{\beta}_i - \beta_{i0}}{s_{\hat{\beta}_i}}$$

La prueba se basa en $n - (k + 1)$ gl y es de extremo superior, inferior o bilateral según si la desigualdad en H_a es $>$, $<$, o $\neq$.

Una estimación puntual de $\mu_{Y \cdot x}$, es decir, de $\beta_0 + \beta_1 x + \cdots + \beta_k x^k$, es $\hat{\mu}_{Y \cdot x} = \hat{\beta}_0 + \hat{\beta}_1 x + \cdots + \hat{\beta}_k x^k$. La desviación estándar estimada del estimador correspondiente es bastante complicada. Muchos paquetes de software producen la desviación estándar estimada para *cualquier* valor de x cuando el usuario lo solicita. Esto, junto con una variable t estandarizada, se puede usar para justificar los siguientes procedimientos.

Sea x^* un valor especificado de x. Un IC de $100(1 - \alpha)\%$ para $\mu_{Y \cdot x}$ es

$$\hat{\mu}_{Y \cdot x^*} \pm t_{\alpha/2,n-(k+1)} \cdot \left\{ \begin{matrix} \text{DE estimada de} \\ \hat{\mu}_{Y \cdot x^*} \end{matrix} \right\}$$

Con $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x^* + \cdots + \hat{\beta}_k (x^*)^k$, $\hat{y}$ que denota el valor calculado de $\hat{Y}$ para los datos proporcionados y $s_{\hat{Y}}$ que representa la desviación estándar estimada del estadístico $\hat{Y}$, la fórmula para el IC es mucho más parecida a la del caso de la regresión lineal simple:

$$\hat{y} \pm t_{\alpha/2,n-(k+1)} \cdot s_{\hat{Y}}$$

Un IP de $100(1 - \alpha)\%$ para un valor futuro y que se observa cuando $x = x^*$ es

$$\hat{\mu}_{Y \cdot x^*} \pm t_{\alpha/2,n-(k+1)} \cdot \left\{ s^2 + \left(\begin{matrix} \text{DE estimada} \\ \text{de } \hat{\mu}_{Y \cdot x^*} \end{matrix} \right)^2 \right\}^{1/2} = \hat{y} \pm t_{\alpha/2,n-(k+1)} \cdot \sqrt{s^2 + s_{\hat{Y}}^2}$$

Ejemplo 13.9 En la figura 13.11(b) se muestra que $\hat{\beta}_2 = -4.5358$ y $s_{\hat{\beta}_2} = 0.6744$ (de la columna Desv. est. en la parte superior del resultado). Para probar H_0: $\beta_2 = 0$ contra H_a: $\beta_2 \neq 0$ (aquí H_0 expresa que es innecesario el término cuadrático en el modelo), el estadístico de prueba es $T = \hat{\beta}_2 / S_{\hat{\beta}_2}$, con valor calculado $-4.5358/0.6744 = -6.73$. La prueba se basa en $n - (k + 1) = 16 - 3 = 13$ gl; con $t_{0.025,13} = 2.160$, H_0 se rechaza al nivel 0.05 si $t \geq 2.160$ o $t \leq -2.160$. Puesto que $-6.73 \leq -2.160$, H_0 se rechaza al nivel 0.05, lo que valida la inclusión del término cuadrático; el valor P correspondiente es 0.

La última línea del resultado en la figura 13.11(b) contiene información de la estimación y predicción para $x = 25$:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(25) + \hat{\beta}_2(25)^2 = \text{Ajuste} = 3431.8$$

$$\left(\begin{matrix} \text{DE estimada de} \\ \hat{\beta}_0 + \hat{\beta}_1(25) + \hat{\beta}_2(25)^2 \end{matrix} \right) = s_{\hat{Y}} = \text{ajuste de desv. est.} = 69.2$$

Así, un IC de 95% para el rendimiento medio cuando la fecha de cosecha = 25 es

$$3431.8 \pm (2.160)(69.2) = 3431.8 \pm 149.5 = (3282.3, 3581.3)$$

mientras que un IP de 95% para el rendimiento cuando $x = 25$ es

$$3431.8 \pm (2.160)[41\,568.3 + (69.2)^2]^{1/2} = 3431.8 \pm 465.1$$
$$= (2966.7, 3896.9)$$

Las ligeras discrepancias entre estos límites de intervalo y los que aparecen en el resultado se atribuyen al redondeo. Observe de nuevo que el IP es mucho más amplio que el IC correspondiente. Esto es porque s^2 es bastante grande en comparación con el ajuste de la desviación estándar. ■

Centrado de los valores de *x*

Para el modelo cuadrático con función de regresión $\mu_{Y \cdot x} = \beta_0 + \beta_1 x + \beta_2 x^2$, los parámetros β_0, β_1 y β_2 caracterizan el comportamiento de la función cerca de $x = 0$. Por ejemplo, β_0 es la altura en la que la función de regresión cruza el eje vertical $x = 0$, mientras que β_1 es la primera

derivada de la función en $x = 0$ (rapidez instantánea del cambio de $\mu_{Y \cdot x}$ en $x = 0$). Si las x_i están lejos de 0, no se puede tener información precisa acerca de los valores de estos parámetros. Sea $\bar{x}$ = promedio de las x_i para las que se hacen las observaciones, y considere el modelo

$$Y = \beta_0^* + \beta_1^*(x - \bar{x}) + \beta_2^*(x - \bar{x})^2 + \epsilon \tag{13.14}$$

En el modelo (13.14), $\mu_{Y \cdot x} = \beta_0^* + \beta_1^*(x - \bar{x}) + \beta_2^*(x - \bar{x})^2$, y los parámetros ahora describen el comportamiento de la función de regresión cerca de la $\bar{x}$ central de los datos.

Para estimar los parámetros de (13.14), simplemente se resta $\bar{x}$ de cada x_i a fin de obtener $x_i' = x_i - \bar{x}$, y luego se emplean las x_i' en lugar de las x_i. Un beneficio importante de esto es que los coeficientes de $b_0, \ldots, b_k$ en las ecuaciones normales (13.10) serán de magnitud mucho más pequeña de lo que sería si se emplearan las x_i originales. Cuando se utiliza una computadora para resolver el sistema, este centrado protege contra cualquier error de redondeo que pudiera resultar.

Ejemplo 13.10 En el artículo "A Method for Improving the Accuracy of Polynomial Regresion Analysis" (*J. Quality Technology*, 1971: 149-155) se muestran los siguientes datos acerca de x = temperatura de cura (°F) y y = resistencia de corte máxima de un compuesto de caucho (lb/pulg^2), con $\bar{x} = 297.13$:

x	280	284	292	295	298	305	308	315
x'	−17.13	−13.13	−5.13	−2.13	0.87	7.87	10.87	17.87
y	770	800	840	810	735	640	590	560

Un análisis de computadora produjo los resultados mostrados en la tabla 13.4.

Tabla 13.4 Coeficientes estimados y desviaciones estándar para el ejemplo 13.10

Parámetro	Estimación	D.E. estimada	Parámetro	Estimación	D.E. estimada
β_0	−26 219.64	11 912.78	β_0^*	759.36	23.20
β_1	189.21	80.25	β_1^*	−7.61	1.43
β_2	−0.3312	0.1350	β_2^*	−0.3312	0.1350

La función de regresión estimada usando el modelo original es $y = -26\,219.64 + 189.21x - 0.3312x^2$, mientras que para el modelo centrado la función es $y = 759.36 - 7.61(x - 297.13) - 0.3312(x - 297.13)^2$. Estas funciones estimadas son idénticas; la única diferencia es que se estimaron parámetros distintos para los dos modelos. Las desviaciones estándar estimadas indican con claridad que β_0^* y β_1^* se estimaron con más precisión que β_0 y β_1. Los parámetros cuadráticos son idénticos ($\beta_2 = \beta_2^*$), como se ve al comparar el término x^2 en (13.14) con el modelo original. Se remarca de nuevo que un beneficio importante de centrar es la ganancia en la precisión de cálculo, no sólo en los modelos cuadráticos, sino también en los de grado superior. ■

El libro de Neter *et al.*, listado en la bibliografía del capítulo, es una buena fuente para más información acerca de la regresión polinomial.

Ejercicios | Sección 13.3 (26 a 35)

26. Además de una regresión lineal de densidad verdadera en el contenido de humedad, en el artículo citado en el ejercicio 6 se consideró una regresión cuadrática de densidad en volumen en función del contenido de humedad. A continuación se presentan los datos de una gráfica del artículo, junto con un resultado de MINITAB del ajuste cuadrático.

```
La ecuación de regresión de densidad en
volumen = 403 + 16.2 cont. de humedad - 0.706 cont.
al cuadrado

Predictor              Coef  Desv. est.      T      P
Constante            403.24       36.45  11.06  0.002
Cont. de humedad     16.164       5.451   2.97  0.059
Cont. al cuadrado   -0.7063      0.1852  -3.81  0.032

S = 10.15   R cuadrada = 93.8% R cuadrada
(ajustada) = 89.6%

Análisis de varianza

Fuente           GL      SC      CM      F      P
Regresión         2  4637.7  2318.9  22.51  0.016
Error residual    3   309.1   103.0
Total             5  4946.8

   Cont. de   Densidad                Desv.              Residuo
Obs humedad de volumen  Ajuste  est.  Residuo    est.
  1     7.0    479.00  481.78  9.35    -2.78   -0.70
  2    10.3    503.00  494.79  5.78     8.21    0.98
  3    13.7    487.00  492.12  6.49    -5.12   -0.66
  4    16.6    470.00  476.93  6.10    -6.93   -0.85
  5    19.8    458.00  446.39  5.69    11.61    1.38
  6    22.0    412.00  416.99  8.75    -4.99   -0.97

        Ajuste de
Ajuste  desv.est.    IC de 95.0%        IP de 95.0%
491.10       6.52 (470.36, 511.83) (452.71, 529.48)
```

a. ¿Un diagrama de dispersión de los datos al parecer es congruente con el modelo de regresión cuadrático?

b. ¿Qué proporción de la variación observada en la densidad se puede atribuir a la relación de modelo?

c. ¿El modelo cuadrático al parecer es útil? Realice una prueba al nivel de significación 0.05.

d. El último renglón del resultado se obtiene al solicitar una estimación e información de pronóstico cuando el contenido de humedad es 14. Calcule un IP de 99% para la densidad cuando el contenido de humedad es 14.

e. ¿El predictor cuadrático al parecer proporciona información útil? Pruebe las hipótesis apropiadas al nivel de significación 0.05.

27. Los datos siguientes acerca de y = concentración de glucosa (g/L) y x = tiempo de fermentación (días) para una determinada mezcla de licor de malta se leyeron de una gráfica de dispersión del artículo "Improving Fermentation Productivity with Reverse Osmosis" *(Food Tech., 1984: 92-96)*:

x	1	2	3	4	5	6	7	8
y	74	54	52	51	52	53	58	71

a. Compruebe que un diagrama de dispersión es congruente con la elección de un modelo de regresión cuadrático.

b. La ecuación de regresión cuadrática estimada es $y = 84.482 - 15.875x + 1.7679x^2$. Prediga el valor de la concentración de glucosa para un tiempo de fermentación de 6 días y calcule el residuo correspondiente.

c. Con SSE = 61.77, ¿qué proporción de la variación observada se puede atribuir a la relación de regresión cuadrática?

d. Los $n = 8$ residuos estandarizados basados en el modelo cuadrático son 1.91, −1.95, −0.25, 0.58, 0.90, 0.04, −0.66 y 0.20. Construya una gráfica de los residuos estandarizados en función de x y una gráfica de probabilidad normal. ¿Las gráficas muestran algunas características problemáticas?

e. La desviación estándar estimada de $\hat{\mu}_{Y\cdot 6}$, es decir, $\hat{\beta}_0 + \hat{\beta}_1(6) + \hat{\beta}_2(36)$, es 1.69. Calcule un intervalo de confianza de 95% para $\mu_{Y\cdot 6}$.

f. Calcule un IP de 95% para una observación de concentración de glucosa realizada después de 6 días de tiempo de fermentación.

28. La viscosidad (y) de un aceite se midió con un viscosímetro de cono y placa a seis velocidades (x). Se supuso que era apropiado un modelo cuadrático de regresión y la función de regresión estimada resultante de las $n = 6$ observaciones fue

$$y = -113.0937 + 3.3684x - 0.01780x^2$$

a. Estime $\mu_{Y\cdot 75}$, la viscosidad esperada cuando la velocidad es 75 rpm.

b. ¿Qué viscosidad pronosticaría para una velocidad de cono de 60 rpm?

c. Si $\sum y_i^2 = 8386.43$, $\sum y_i = 210.70$, $\sum x_i y_i = 17\,002.00$ y $\sum x_i^2 y_i = 1\,419\,780$, calcule SSE $[= \sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i - \hat{\beta}_2 \sum x_i^2 y_i]$, s^2 y s.

d. Del inciso (c), SST = $8386.43 - (210.70)^2/6 = 987.35$. Con SSE calculada en el inciso (c), ¿cuál es el valor calculado de R^2?

e. Si la desviación estándar estimada de $\hat{\beta}_2$ es $s_{\hat{\beta}_2} = 0.00226$, pruebe H_0: $\beta_2 = 0$ contra H_a: $\beta_2 \neq 0$ al nivel 0.01 e interprete el resultado.

29. En el ejercicio 7 se muestran los siguientes datos sobre el tiempo de exposición a la radiación x y el peso en seco de raíces y:

x	0	2	4	6	8
y	110	123	119	86	62

La función de regresión cuadrática estimada es $y = 111.89 + 8.06x - 1.84x^2$.

a. Calcule los valores pronosticados y residuos. Después calcule SSE y s^2. [Véase el ejercicio 28(c).]
b. Calcule el coeficiente de determinación múltiple R^2.
c. La desviación estándar estimada de $\hat{\beta}_2$, el estimador del coeficiente cuadrático β_2, es $s_{\hat{\beta}_2} = 0.480$. ¿El término cuadrático pertenece al modelo? Exprese y pruebe las hipótesis apropiadas al nivel 0.05.
d. La desviación estándar estimada de $\hat{\beta}_1$ es $s_{\hat{\beta}_1} = 4.01$. Utilice esto y la información del inciso (c) para obtener intervalos de confianza conjuntos para β_1 y β_2 con nivel de confianza conjunto de (al menos) 95%.
e. La desviación estándar estimada de $\hat{\mu}_{Y\cdot 4}$ $(= \hat{\beta}_0 + 4\hat{\beta}_1 + 16\hat{\beta}_2)$ es 5.01. Calcule un intervalo de confianza de 90% para $\mu_{Y\cdot 4}$.
f. Estime el tiempo de exposición que maximice el peso seco esperado de las raíces.

30. En el artículo "A Simulation-Based Evaluation of Three Cropping Systems on Cracking-Clay Soils in a Summer Rainfall Environment" (*Agricultural Meteorology*, 1976: 211-229) propone un modelo cuadrático para la relación entre el índice de suministro de agua (x) y la producción de trigo (y). A continuación se muestran los datos representativos y el resultado de MINITAB:

x	1.2	1.3	1.5	1.8	2.1	2.3	2.5
y	790	950	740	1230	1000	1465	1370

x	2.9	3.1	3.2	3.3	3.9	4.0	4.3
y	1420	1625	1600	1720	1500	1550	1560

```
La ecuación de regresión es
producción = -252 + 1000 índice - 135
índice cuadrado

                          Desv.
Predictor          Coef    est.   Relación t      p
Constante        -251.6   285.1        -0.88  0.396
índice           1000.1   229.5         4.36  0.001
índice al
cuadrado        -135.44   41.97        -3.23  0.008

                R cuadrada          R cuadrada
s = 135.6         = 85.3%     (ajustada) = 82.6%

Análisis de varianza

FUENTE          GL       SC       CM       F      p
Regresión        2  1170208   585104   31.83  0.000
Error           11   202228    18384
Total           13  1372435
```

a. Interprete el valor del coeficiente de determinación múltiple.
b. Calcule un intervalo de confianza de 95% para el coeficiente del predictor cuadrático.
c. La desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$ cuando $x = 2.5$ es 53.5. Pruebe H_0: $\mu_{Y\cdot 2.5} = 1500$ contra H_a: $\mu_{Y\cdot 2.5} < 1500$ con $\alpha = 0.01$.
d. Obtenga un intervalo de predicción para la producción de trigo cuando el índice de suministro de agua es 2.5 mediante la información presentada en el inciso (c).

31. Los siguientes datos se obtuvieron de un estudio de cierto método para preparar alcohol puro de corrientes de refinerías ("Direct Hydration of Olefins", *Industrial and Eng. Chemistry*, 1961: 209-211). La variable independiente x es espacio velocidad por hora del volumen y la variable dependiente y es la cantidad de conversión de isobutileno.

x	1	1	2	4	4	4	6
y	23.0	24.5	28.0	30.9	32.0	33.6	20.0

a. Suponga que es apropiado un modelo probabilístico cuadrático y estime la función de regresión.
b. Determine los valores pronosticados y residuos, y trace una gráfica de residuos. ¿La gráfica se ve más o menos como se espera cuando el modelo cuadrático es correcto? ¿La gráfica indica que alguna observación tuvo gran influencia en el ajuste? ¿Una gráfica de dispersión identifica un punto que tiene gran influencia? Si es así, ¿de cuál punto se trata?
c. Obtenga s^2 y R^2. ¿El modelo cuadrático proporciona un buen ajuste para los datos?
d. En el ejercicio 11 se observó que el valor pronosticado $\hat{Y}_j$ y el residuo $Y_j - \hat{Y}_j$ son independientes entre sí, de modo que $\sigma^2 = V(Y_j) = V(\hat{Y}_j) + V(Y_j - \hat{Y}_j)$. En un resultado impreso de computadora se observa que las desviaciones estándar estimadas de los valores pronosticados son 0.955, 0.955, 0.712, 0.777, 0.777, 0.777 y 1.407. Utilice estos valores junto con s^2 para calcular la desviación estándar estimada de cada residuo. Después, calcule los residuos estandarizados y dibújelos en función de x. ¿Esta gráfica se asemeja a la del inciso (b)? Suponga que estandarizó los residuos utilizando sólo s en el denominador. ¿Los valores resultantes diferirían mucho de los valores correctos?
e. Con la información del inciso (d), calcule un IP de 90% para la conversión de isobutileno cuando el espacio velocidad por hora del volumen es 4.

32. Los siguientes datos son un subconjunto de los datos obtenidos en un experimento para estudiar la relación entre el pH x del suelo y y = A1. Concentración/conductividad eléctrica ("Root Responses of Three *Gramineae* Species to Soil Acidity in an Oxisol and an Ultisol", *Soil Science*, 1973: 295-302):

x	4.01	4.07	4.08	4.10	4.18
y	1.20	0.78	0.83	0.98	0.65
x	4.20	4.23	4.27	4.30	4.41
y	0.76	0.40	0.45	0.39	0.30
x	4.45	4.50	4.58	4.68	4.70	4.77
y	0.20	0.24	0.10	0.13	0.07	0.04

En el artículo se propuso un modelo cúbico, pero en la versión de MINITAB que utilizó el autor del presente texto no se pudo incluir el término x^3 en el modelo porque según el programa "x^3 tiene una correlación alta con otras variables predictoras". Para solucionar esto, $\bar{x} = 4.3456$ se restó de cada valor de x para obtener $x' = x - \bar{x}$. Entonces, se requirió una regresión cúbica para ajustar el modelo con función de regresión

$$y = \beta_0^* + \beta_1^* x' + \beta_2^* (x')^2 + \beta_3^* (x')^3$$

El resultado de computadora fue el siguiente:

Parámetro	Estimación	DE estimada
β_0^*	0.3463	0.0366
β_1^*	−1.2933	0.2535
β_2^*	2.3964	0.5699
β_3^*	−2.3968	2.4590

a. ¿Cuál es la función de regresión estimada para el modelo "centrado"?

b. ¿Cuál es el valor estimado del coeficiente β_3 en el modelo "no centrado", con función de regresión $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$? ¿Cuál es la estimación de β_2?

c. Con el modelo cúbico, ¿qué valor de y pronosticaría cuando el pH del suelo es 4.5?

d. Realice una prueba para determinar si el término cúbico se debe retener en el modelo.

33. En muchos problemas de regresión polinomial, en vez de ajustar una función de regresión "centrada" por medio de $x' = x - \bar{x}$, la precisión del cálculo se mejora si se utiliza una función de la variable independiente estandarizada $x' = (x - \bar{x})/s_x$, donde s_x es la desviación estándar de las x_i. Considere ajustar la función de regresión cúbica $y = \beta_0^* + \beta_1^* x' + \beta_2^* (x')^2 + \beta_3^* (x')^3$ a los siguientes datos que resultan de un estudio de la relación entre eficiencia de empuje y de cohetes propulsores supersónicos y el ángulo de semidivergencia x de la tobera del cohete ("More on Correlating Data", *CHEMTECH*, 1976: 266-270):

x	5	10	15	20	25	30	35
y	0.985	0.996	0.988	0.962	0.940	0.915	0.878

Parámetro	Estimación	DE estimada
β_0^*	0.9671	0.0026
β_1^*	−0.0502	0.0051
β_2^*	−0.0176	0.0023
β_3^*	0.0062	0.0031

a. ¿Cuál valor de y pronosticaría cuando el ángulo de semidivergencia es 20? ¿Cuándo $x = 25$?

b. ¿Cuál es la función de regresión estimada $\hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3$ para el modelo "no estandarizado"?

c. Utilice una prueba de nivel 0.05 para decidir si el término cúbico se debe eliminar del modelo.

d. ¿Qué se puede decir acerca de la relación entre las SSE y las R^2 para modelos estandarizados y no estandarizados? Explique.

e. La SSE para el modelo cúbico es 0.00006300, mientras que para un modelo cuadrático SSE es 0.00014367. Calcule R^2 para cada modelo. ¿La diferencia entre las dos indica que se puede eliminar el término cúbico?

34. Los siguientes datos se obtuvieron en un experimento para evaluar el potencial de desechos de mina de carbón sin quemar como medio para el crecimiento de plantas. Las variables son x = cationes ácidos extraíbles y y = acidez intercambiable/capacidad total de intercambio de cationes. ("Exchangeable Acidity in Unburnt Colliery Spoil", *Nature*, 1969: 161):

x	−23	−5	16	26	30	38	52
y	1.50	1.46	1.32	1.17	0.96	0.78	0.77
x	58	67	81	96	100	113	
y	0.91	0.78	0.69	0.52	0.48	0.55	

Al estandarizar la variable independiente x para obtener $x' = (x - \bar{x})/s_x$ y ajustar la función de regresión $y = \beta_0^* + \beta_1^* x' + \beta_2^* (x')^2$ se obtuvo el siguiente resultado de computadora:

Parámetro	Estimación	DE estimada
β_0^*	0.8733	0.0421
β_1^*	−0.3255	0.0316
β_2^*	0.0448	0.0319

a. Estime $\mu_{Y\cdot 50}$.
b. Calcule el valor del coeficiente de determinación múltiple. [Véase el ejercicio 28(c).]
c. ¿Cuál es la función de regresión estimada $\hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$ si se emplea la variable no estandarizada x?
d. ¿Cuál es la desviación estándar estimada de $\hat{\beta}_2$ calculada en el inciso (c)?
e. Realice una prueba con las estimaciones estandarizadas para determinar si el término cuadrático debe ser retenido en el modelo. Repita la prueba con las estimaciones no estandarizadas. ¿Difieren sus conclusiones?

35. En el artículo "The Respiration in Air and in Water of the Limpets *Patella caerulea* and *Patella lusitanica*" (*Comp. Biochemistry and Physiology*, 1975: 407-411) se propuso un modelo de potencia simple para la relación entre la frecuencia respiratoria y y la temperatura x para *P. caerulea* en el aire. Sin embargo, una gráfica de $\ln(y)$ en función de x muestra un patrón curvo. Ajuste el modelo de potencia cuadrático $Y = \alpha e^{\beta_0 + \beta_1 x + \beta_2 x^2} \cdot \epsilon$ a los siguientes datos:

x	10	15	20	25	30
y	37.1	70.1	109.7	177.2	222.6

13.4 Análisis de regresión múltiple

En la regresión múltiple, el objetivo es construir un modelo probabilístico que relaciona una variable dependiente y con más de una variable independiente o predictor. Sea k el número de variables predictoras ($k \geq 2$) y denote estos predictores con $x_1, x_2, \ldots, x_k$. Por ejemplo, en un intento por predecir el precio de venta de una casa, se podría tener $k = 3$ con x_1 = tamaño (pies^2), x_2 = edad (años) y x_3 = número de habitaciones.

DEFINICIÓN

La **ecuación del modelo de regresión múltiple aditiva general** es

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \tag{13.15}$$

donde $E(\epsilon) = 0$ y $V(\epsilon) = \sigma^2$. Además, para propósitos de probar las hipótesis y calcular IC o IP, se supone que ϵ tiene una distribución normal.

Sean $x_1^*, x_2^*, \ldots, x_k^*$ valores particulares de $x_1, \ldots, x_k$. Entonces (13.15) implica que

$$\mu_{Y\cdot x_1^*,\ldots,x_k^*} = \beta_0 + \beta_1 x_1^* + \cdots + \beta_k x_k^* \tag{13.16}$$

Así, del mismo modo que $\beta_0 + \beta_1 x$ describe el valor Y medio como función de x en la regresión lineal simple, la **función de regresión verdadera** (o **poblacional**) $\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$ da el valor esperado de Y como función de $x_1, \ldots, x_k$. Las β_i son los **coeficientes de regresión verdaderos** (o **poblacionales**). El coeficiente de regresión β_1 se interpreta como el cambio esperado en Y relacionado con un incremento de una unidad en x_1 *mientras que* $x_2, \ldots, x_k$ *se mantienen fijas*. Las interpretaciones análogas se cumplen para $\beta_2, \ldots, \beta_k$.

Modelos con interacción y predictores cuadráticos

Si un investigador ha obtenido observaciones en y, x_1 y x_2, un modelo posible es $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$. Sin embargo, otros modelos se construyen al formar predictores que son fun-

ciones matemáticas de x_1 y/o x_2, o ambas. Por ejemplo, con $x_3 = x_1^2$ y $x_4 = x_1x_2$, el modelo

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \epsilon$$

tiene la forma general de (13.15). En general, no sólo se permite que algunos predictores sean funciones matemáticas de otros, sino que también con frecuencia es muy deseable en el sentido de que el modelo resultante pudiera ser mucho más exitoso para explicar la variación en y que algún modelo sin esta clase de predictores. Esta explicación también muestra que la regresión polinomial es en realidad un caso especial de la regresión múltiple. Por ejemplo, el modelo cuadrático $Y = \beta_0 + \beta_1x + \beta_2x^2 + \epsilon$ tiene la forma de (13.15) con $k = 2$, $x_1 = x$ y $x_2 = x^2$.

Para el caso de dos variables independientes, x_1 y x_2, hay cuatro modelos de regresión múltiple útiles.

1. El modelo de primer orden:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$$

2. El modelo de segundo orden sin interacción:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2 + \epsilon$$

3. El modelo con predictores de primer orden e interacción:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \epsilon$$

4. El modelo de segundo orden completo o cuadrático total:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2 + \beta_5x_1x_2 + \epsilon$$

Entender las diferencias entre estos modelos es un primer paso importante en la construcción de modelos de regresión reales de las variables independientes bajo estudio.

El modelo de primer orden es la generalización más directa de la regresión lineal simple. Expresa que para un valor fijo de cualquier variable, el valor esperado de Y es una función lineal de la otra variable y que el cambio esperado en Y para un incremento unitario en x_1 (x_2) es β_1 (β_2) independiente del nivel de x_2 (x_1). Así, si se traza la gráfica de la función de regresión como una función de x_1 para varios valores de x_2, se obtiene como contorno de la función de regresión una colección de rectas paralelas, como se ilustra en la figura 13.12(a). La función $y = \beta_0 + \beta_1x_1 + \beta_2x_2$ especifica un plano en el espacio tridimensional; el modelo de primer orden establece que cada valor observado de la variable dependiente corresponde a un punto que se desvía verticalmente de este plano por una cantidad aleatoria ϵ.

Según el modelo de segundo orden sin interacción, si se fija x_2, el cambio esperado en Y para un incremento de una unidad en x_1 es

$$\beta_0 + \beta_1(x_1 + 1) + \beta_2x_2 + \beta_3(x_1 + 1)^2 + \beta_4x_2^2 - (\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2) = \beta_1 + \beta_3 + 2\beta_3x_1$$

Debido a que este cambio esperado no depende de x_2, los contornos de la función de regresión para distintos valores de x_2 aún son paralelos entre sí. Sin embargo, la dependencia del cambio esperado en el valor de x_1 significa que los contornos ahora son curvas en vez de rectas. Esto se ilustra en la figura 13.12(b). En este caso, la superficie de regresión ya no es un plano en el espacio tridimensional sino una superficie curvada.

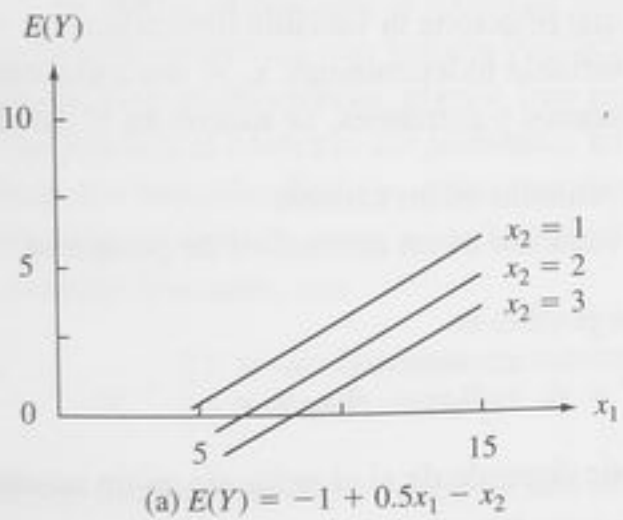
Los contornos de la función de regresión para el modelo de interacción de primer orden son rectas no paralelas. Esto es porque el cambio esperado en Y cuando x_1 se incrementa en uno es

$$\beta_0 + \beta_1(x_1 + 1) + \beta_2 x_2 + \beta_3(x_1 + 1)x_2 - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2) = \beta_1 + \beta_3 x_2$$

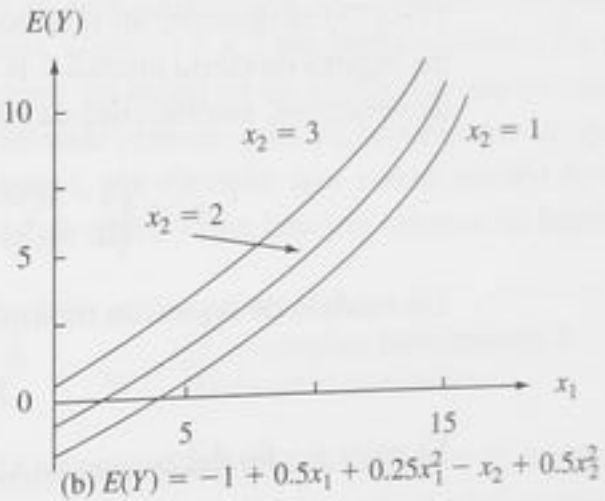
Este cambio esperado depende del valor de x_2, así que cada línea de contorno debe tener una pendiente distinta, como en la figura 13.12(c). La palabra *interacción* refleja el hecho de que un cambio esperado en Y cuando una variable aumenta de valor depende del valor de la otra variable.

Por último, para el modelo completo de segundo orden, el cambio esperado en Y cuando se mantiene fija x_2 mientras x_1 se incrementa en una unidad es $\beta_1 + \beta_3 + 2\beta_3 x_1 + \beta_5 x_2$, que es una función de x_1 y x_2. Esto significa que los contornos de la función de regresión son curvos y no paralelos entre sí, como se ilustra en la figura 13.12(d).

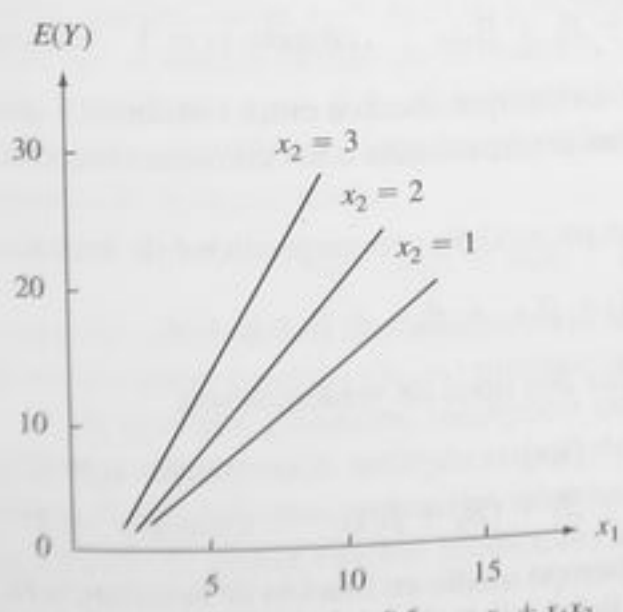
Consideraciones similares se aplican a modelos construidos de más de dos variables independientes. En general, la presencia de términos de interacción en el modelo significa que el cambio esperado en Y depende no sólo de la variable que aumenta o disminuye, sino también



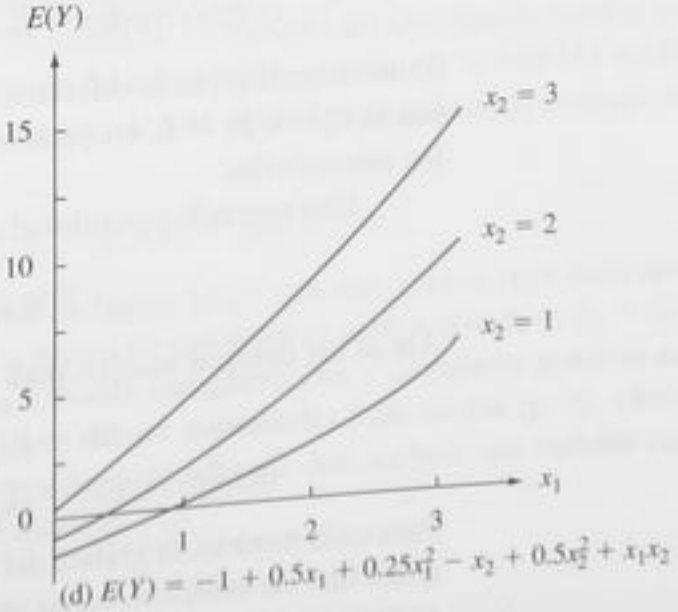
(a) $E(Y) = -1 + 0.5x_1 - x_2$



(b) $E(Y) = -1 + 0.5x_1 + 0.25x_1^2 - x_2 + 0.5x_2^2$



(c) $E(Y) = -1 + 0.5x_1 - x_2 + x_1x_2$



(d) $E(Y) = -1 + 0.5x_1 + 0.25x_1^2 - x_2 + 0.5x_2^2 + x_1x_2$

Figura 13.12 Contornos de cuatro funciones de regresión diferentes

de los valores de las variables fijas. Como en ANOVA, es posible tener términos de interacción multilaterales (p. ej., $x_1x_2x_3$), lo cual dificulta la interpretación del modelo.

Observe que si el modelo contiene interacción o cuadráticos predictores, no aplica la interpretación genérica de una β_i proporcionada antes. Esto se debe a que entonces no es posible incrementar x_i en una unidad y mantener fijos los valores de los otros predictores.

Modelos con predictores para variables categóricas

Hasta el momento se ha considerado en forma explícita la inclusión de sólo variables predictoras cuantitativas (numéricas) en un modelo de regresión múltiple. Si se emplea codificación numérica simple, las variables cualitativas (categóricas), como el material de cojinetes (aluminio o cobre/plomo) o tipo de madera (pino, encino o nogal), también se pueden incorporar en un modelo. Primero se ve el caso de una variable dicótoma, una con sólo dos categorías posibles, masculino o femenino, producto nacional o extranjero, etcétera. Con alguna variable de este tipo, se relaciona una **variable *x* ficticia** o **indicadora** cuyos valores posibles 0 y 1 indican cuál categoría es pertinente para una determinada observación.

Ejemplo 13.11 En el artículo "Estimating Urban Travel Times: A Comparative Study" (*Trans. Res.*, 1980: 173-175) se describe un estudio que relaciona la variable dependiente y = tiempo de viaje entre lugares de cierta ciudad y la variable independiente x_2 = distancia entre lugares. Dos tipos de vehículos, automóviles de pasajeros y camiones, se usaron en el estudio. Sea

$$x_1 = \begin{cases} 1 & \text{si el vehículo es un camión} \\ 0 & \text{si el vehículo es un automóvil de pasajeros} \end{cases}$$

Un modelo de regresión múltiple posible es

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

El valor medio del tiempo de viaje depende de si el vehículo es un automóvil o un camión:

$$\begin{aligned} &\text{tiempo medio} = \beta_0 + \beta_2 x_2 && \text{cuando } x_1 = 0 \quad \text{(automóviles)} \\ &\text{tiempo medio} = \beta_0 + \beta_1 + \beta_2 x_2 && \text{cuando } x_1 = 1 \quad \text{(camiones)} \end{aligned}$$

El coeficiente β_1 es la diferencia de tiempos medios entre camiones y automóviles con la distancia fija; si $\beta_1 > 0$, en promedio le tomará más a los camiones recorrer una distancia que a los automóviles.

Una segunda posibilidad es un modelo con un predictor de interacción:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon$$

Ahora los tiempos medios para los dos tipos de vehículos son

$$\begin{aligned} &\text{tiempo medio} = \beta_0 + \beta_2 x_2 && \text{cuando } x_1 = 0 \\ &\text{tiempo medio} = \beta_0 + \beta_1 + (\beta_2 + \beta_3) x_2 && \text{cuando } x_1 = 1 \end{aligned}$$

Para cada modelo, la gráfica del tiempo medio en función de la distancia es una recta para cualquier tipo de vehículo, como se ilustra en la figura 13.13. Las dos rectas son paralelas para el primer modelo (sin interacción), pero en general tendrán diferentes pendientes cuando el segundo modelo es correcto. Para este último modelo, el cambio en el tiempo de recorrido medio relacionado con un incremento de una milla depende de qué tipo de vehículo se trate, las dos

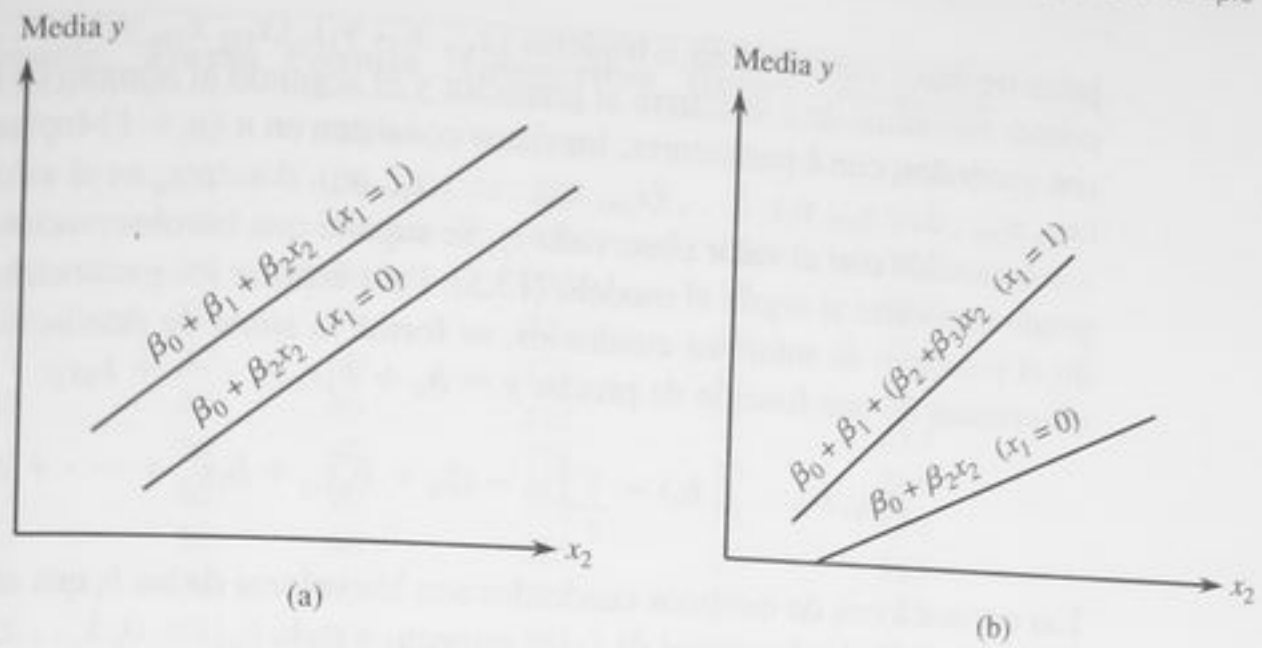


Figura 13.13 Funciones de regresión para modelos con una variable ficticia (x_1) y una variable cuantitativa x_2: (a) sin interacción (b) interacción

variables "tipo de vehículo" y "tiempo de viaje" interactúan. De hecho, los datos que reunieron los autores del artículo citado indican la presencia de interacción. ■

Se podría pensar que la forma de manejar una situación de tres categorías es definir una sola variable numérica con valores codificados, como 0, 1 y 2 que corresponden a las tres categorías. Esto es incorrecto, porque esto impone un orden en las categorías que no necesariamente implica el contexto del problema. El método correcto para incorporar tres categorías es definir *dos* variables ficticias diferentes. Suponga, por ejemplo, que y es el tiempo de vida de cierta herramienta de corte, x_1 es la velocidad de corte y que hay tres marcas de herramienta en estudio. Entonces, sea

$$x_2 = \begin{cases} 1 & \text{si usa herramienta marca A} \\ 0 & \text{en caso contrario} \end{cases} \qquad x_3 = \begin{cases} 1 & \text{si emplea herramienta B} \\ 0 & \text{en caso contrario} \end{cases}$$

Cuando se hace una observación en una herramienta marca A, $x_2 = 1$ y $x_3 = 0$, mientras que para una herramienta de la marca B, $x_2 = 0$ y $x_3 = 1$. Una observación hecha en una herramienta marca C tiene $x_2 = x_3 = 0$, y es imposible que $x_2 = x_3 = 1$ porque una herramienta no puede ser al mismo tiempo de la marca A y la marca B. El modelo sin interacción tendría sólo los predictores x_1, x_2 y x_3. El siguiente modelo de interacción permite que el cambio medio en el tiempo de vida relacionado con un incremento de una unidad en la velocidad dependa de la marca de la herramienta:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \epsilon$$

La construcción de una ilustración como la de la figura 13.13 con una gráfica para cada uno de los tres pares posibles (x_2, x_3) produce tres rectas no paralelas (a menos que $\beta_4 = \beta_5 = 0$).

En términos generales, incorporar una variable categórica con c categorías posibles en un modelo de regresión múltiple requiere el uso de $c - 1$ variables indicadoras (p. ej., cinco marcas de herramientas requerirían cuatro variables indicadoras). Así, incluso una variable categórica puede añadir muchos predictores al modelo.

Estimación de parámetros

Los datos de la regresión lineal simple consisten en n pares $(x_1, y_1), \ldots, (x_n, y_n)$. Suponga que el modelo de regresión múltiple contiene dos variables predictoras, x_1 y x_2. Entonces, el con-

junto de datos consistirá en n tripletas $(x_{11}, x_{21}, y_1), (x_{12}, x_{22}, y_2), \ldots, (x_{1n}, x_{2n}, y_n)$. Aquí el primer subíndice de x se refiere al predictor y el segundo al número de observación. En términos generales, con k predictores, los datos consisten en n $(k + 1)$-tuplas $(x_{11}, x_{21}, \ldots, x_{k1}, y_1)$, $(x_{12}, x_{22}, \ldots, x_{k2}, y_2), \ldots, (x_{1n}, x_{2n}, \ldots, x_{kn}, y_n)$, donde x_{ij} es el valor del i-ésimo predictor x_i relacionado con el valor observado y_j. Se supone que las observaciones para las y_j son independientes entre sí según el modelo (13.5). Para estimar los parámetros $\beta_0, \beta_1, \ldots, \beta_k$ usando el principio de mínimos cuadrados, se forma la suma de desviaciones cuadradas de las y_j observadas de una función de prueba $y = b_0 + b_1x_1 + \cdots + b_kx_k$:

$$f(b_0, b_1, \ldots, b_k) = \sum_j [y_j - (b_0 + b_1x_{1j} + b_2x_{2j} + \cdots + b_kx_{kj})]^2 \tag{13.17}$$

Las estimaciones de mínimos cuadrados son los valores de las b_i que minimizan $f(b_0, \ldots, b_k)$. Tomando la derivada parcial de f con respecto a cada b_i $(i = 0, 1, \ldots, k)$ e igualando las parciales a cero, se obtiene el siguientes sistema de **ecuaciones normales:**

$$\begin{aligned}
&b_0n + b_1\sum x_{1j} + b_2\sum x_{2j} + \cdots + b_k\sum x_{kj} = \sum y_j \\
&b_0\sum x_{1j} + b_1\sum x_{1j}^2 + b_2\sum x_{1j}x_{2j} + \cdots + b_k\sum x_{1j}x_{kj} = \sum x_{1j}y_j \\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&b_0\sum x_{kj} + b_1\sum x_{1j}x_{kj} + \cdots + b_{k-1}\sum x_{k-1,j}x_{kj} + b_k\sum x_{kj}^2 = \sum x_{kj}y_j
\end{aligned} \tag{13.18}$$

Estas ecuaciones son lineales en las incógnitas $b_0, b_1, \ldots, b_k$. Al resolver (13.18) se obtienen las estimaciones de mínimos cuadrados $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$. Lo mejor para realizar esto es usar un paquete de software para estadística.

Ejemplo 13.12 En el artículo "How to Optimize and Control the Wire Bonding Process: Part II" (*Solid State Technology*, enero de 1991: 67-72) se describe un experimento realizado para evaluar el impacto de las variables x_1 = fuerza (gm), x_2 = potencia (mW), x_3 = temperatura (°C) y x_4 = tiempo (ms) en y = resistencia cortante de la unión de bola (gm). Los siguientes datos* se generaron para ser congruentes con la información que se muestra en el artículo:

Observación	Fuerza	Potencia	Temperatura	Tiempo	Resistencia
1	30	60	175	15	26.2
2	40	60	175	15	26.3
3	30	90	175	15	39.8

(continúa)

*Del libro *Statistics Engineering Problem Solving* de Stephen Vardeman, una exposición excelente de los temas que se tratan en el presente texto, aunque a un nivel un poco más alto.

Observación	Fuerza	Potencia	Temperatura	Tiempo	Resistencia
4	40	90	175	15	39.7
5	30	60	225	15	38.6
6	40	60	225	15	35.5
7	30	90	225	15	48.8
8	40	90	225	15	37.8
9	30	60	175	25	26.6
10	40	60	175	25	23.4
11	30	90	175	25	38.6
12	40	90	175	25	52.1
13	30	60	225	25	39.5
14	40	60	225	25	32.3
15	30	90	225	25	43.0
16	40	90	225	25	56.0
17	25	75	200	20	35.2
18	45	75	200	20	46.9
19	35	45	200	20	22.7
20	35	105	200	20	58.7
21	35	75	150	20	34.5
22	35	75	250	20	44.0
23	35	75	200	10	35.7
24	35	75	200	30	41.8
25	35	75	200	20	36.5
26	35	75	200	20	37.6
27	35	75	200	20	40.3
28	35	75	200	20	46.0
29	35	75	200	20	27.8
30	35	75	200	20	40.3

Con un paquete de software para estadística se obtuvieron las siguientes estimaciones de mínimos cuadrados:

$$\hat{\beta}_0 = -37.48 \qquad \hat{\beta}_1 = 0.2117 \qquad \hat{\beta}_2 = 0.4983 \qquad \hat{\beta}_3 = 0.1297 \qquad \hat{\beta}_4 = 0.2583$$

Así, se estima que 0.1297 gm es el cambio promedio en la resistencia relacionada con un aumento de un grado en la temperatura cuando se mantienen fijos los otros tres predictores; los otros coeficientes estimados se interpretan de manera similar.

La ecuación de regresión estimada es

$$y = -37.48 + 0.2117x_1 + 0.4983x_2 + 0.1297x_3 + 0.2583x_4$$

Una predicción puntual de la resistencia resultante de una fuerza de 35 gm, potencia de 75 mW, temperatura de 200 grados y tiempo de 20 ms es

$$\hat{y} = -37.48 + (0.2117)(35) + (0.4983)(75) + (0.1297)(200) + (0.2583)(20)$$
$$= 38.41 \text{ gm}$$

Ésta es también una estimación puntual del valor medio de la resistencia para valores especificados de fuerza, potencia, temperatura y tiempo. ■

$\hat{\sigma}^2$ y R^2

Sustituyendo los valores de los predictores de las observaciones sucesivas en la ecuación para la regresión estimada se obtienen los **valores pronosticados** o **ajustados** $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$. Por ejemplo, puesto que los valores de los cuatro predictores para la última observación del ejemplo 13.12 son 35, 75, 200 y 20, respectivamente, el valor pronosticado correspondiente es $\hat{y}_{30} = 38.41$. Los **residuos** son las diferencias $y_1 - \hat{y}_1, \ldots, y_n - \hat{y}_n$. El último residuo del ejemplo 13.12 es $40.3 - 38.41 = 1.89$. Mientras más cerca de cero estén los residuos, mejor funciona la ecuación estimada para predecir los valores de y que corresponden a los valores de los predictores de la muestra.

Al igual que para la regresión lineal simple y la regresión polinomial, la estimación de σ^2 se basa en la suma de los residuos cuadrados:

$$\text{SSE} = \sum(y_j - \hat{y}_j)^2 = \sum[y_j - (\hat{\beta}_0 + \hat{\beta}_1 x_{1j} + \cdots + \hat{\beta}_k x_{kj})]^2$$

La mayor parte de los paquetes de software para estadística emplean una fórmula de cálculo eficaz para SSE. Debido a que han sido estimados $k + 1$ parámetros $(\beta_0, \beta_1, \ldots, \beta_k)$, se pierden $k + 1$ gl, así que $n - (k + 1)$ gl se relacionan con SSE, y

$$\hat{\sigma}^2 = s^2 = \frac{\text{SSE}}{n - (k + 1)} = \text{MSE}$$

Con $\text{SST} = \sum(y_i - \bar{y})^2$, la proporción de la variación total que se explica mediante el modelo de regresión es $R^2 = 1 - \text{SSE/SST}$, el **coeficiente de determinación múltiple.** Como en la regresión polinomial, R^2 suele ajustarse para el número de parámetros en el modelo con la fórmula

$$R_a^2 = [(n - 1)R^2 - k]/[n - (k + 1)]$$

La raíz cuadrada positiva del coeficiente de determinación múltiple se llama *coeficiente de correlación múltiple R*. Se puede demostrar que R es el coeficiente de correlación muestral r entre las y_j observadas y las $\hat{y}_j$ pronosticadas (es decir, con $x_j = \hat{y}_j$ en la fórmula para r se obtiene $r = R$).

Ejemplo 13.13 Unos investigadores realizaron un estudio para ver cómo varias características del concreto se ven afectadas por $x_1 = \%$ de polvo de piedra caliza y $x_2 =$ relación agua-cemento, y obtuvieron los datos siguientes ("Durability of Concrete with Addition of Limestone Powder", *Magazine of Concrete Research*, 1996: 131-137).

x_1	x_2	x_1x_2	Resistencia a la compresión de 28 días (MPa)	Adsorbabilidad (%)
21	0.65	13.65	33.55	8.42
21	0.55	11.55	47.55	6.26
7	0.65	4.55	35.00	6.74
7	0.55	3.85	35.90	6.59
28	0.60	16.80	40.90	7.28
0	0.60	0.00	39.10	6.90
14	0.70	9.80	31.55	10.80
14	0.50	7.00	48.00	5.63
14	0.60	8.40	42.30	7.43
			$\bar{y} = 39.317$, SST = 278.52	$\bar{y} = 7.339$, SST = 18.356

Considere a la resistencia compresiva como variable dependiente y. Al ajustar el modelo de primer orden, se obtiene

$$y = 84.82 + 0.1643x_1 - 79.67x_2 \quad \text{SSE} = 72.25 \text{ (gl} = 6) \quad R^2 = 0.741 \quad R_a^2 = 0.654$$

mientras que si se incluye un predictor de interacción el resultado es

$$y = 6.22 + 5.779x_1 + 51.33x_2 - 9.357x_1x_2$$
$$\text{SSE} = 29.35 \text{ (gl} = 5) \quad R^2 = 0.895 \quad R_a^2 = 0.831$$

Con base en este último ajuste, una predicción para la resistencia compresiva cuando el % de piedra caliza = 14 y la relación agua-cemento = 0.60 es

$$\hat{y} = 6.22 + 5.779(14) + 51.33(.60) - 9.357(8.4) = 39.32$$

El ajuste de la relación cuadrática total casi no produce ningún cambio en el valor de R^2. Sin embargo, cuando la variable dependiente es la adsorbabilidad, se obtienen los siguientes resultados: $R^2 = 0.747$ cuando se emplean sólo dos predictores, 0.802 cuando se añade el predictor de interacción y 0.889 cuando se emplean los cinco predictores para la relación cuadrática completa. ■

En general, $\hat{\beta}_i$ se interpreta como una estimación del cambio promedio en y relacionado con un incremento unitario en x_i mientras se mantienen fijos los valores de los otros predictores. Sin embargo, algunas veces es difícil o imposible incrementar el valor de un predictor mientras se mantiene fijos a los demás. En estos casos, hay otra interpretación de los coeficientes de regresión estimados. Para concretar, suponga que $k = 2$, y represente con $\hat{\beta}_1$ la estimación de β_1 en la regresión de y sobre los dos predictores x_1 y x_2. Entonces,

1. Realice la regresión de y en función de x_2 (como una regresión lineal simple) y represente con $g_1, g_2, \ldots, g_n$ los residuos resultantes. Estos residuos representan la variación en y después de eliminar o ajustar los efectos de x_2.
2. Efectúe la regresión de x_1 en función de x_2 (es decir, considera a x_1 como la variable dependiente y a x_2 como la variable independiente en esta regresión lineal simple), y denote los residuos mediante $f_1, \ldots, f_n$. Estos residuos representan la variación en x_1 después de eliminar o ajustar los efectos de x_2.

Ahora considere trazar lo gráfico de los residuos de la primera regresión contra los de la segunda; es decir, dibuje los pares $(f_1, g_1), \ldots, (f_n, g_n)$. El resultado se llama *gráfica residual parcial* o *gráfica residual ajustada*. Si se ajusta una recta de regresión a los puntos de esta gráfica, la pendiente resulta ser $\hat{\beta}_1$ (además, los residuos de esta recta son los residuos $e_1, \ldots, e_n$ de la regresión múltiple de y sobre x_1 y x_2). Así, $\hat{\beta}_1$ se interpreta como los cambios estimados en y relacionados con un incremento unitario en x_1 después de eliminar o ajustar los efectos de algunos otros predictores de modelo. La misma interpretación se cumple para otros coeficientes estimados sin importar el número de predictores en el modelo (no hay nada de especial acerca de $k = 2$; el argumento anterior aún es válido si se hace la regresión de y en función de los predictores distintos a x_1 en el paso 1 y se efectúa la regresión de x_1 en función de los otros $k - 1$ predictores en el paso 2).

Como ejemplo, suponga que y es el precio de venta de un edificio de departamentos, y que los predictores son número de departamentos, edad, tamaño del lote, número de espacios de estacionamiento y área de edificación total (pies²). Podría ser ilógico incrementar el núme-

ro de departamentos sin incrementar también el área total. Sin embargo, si $\hat{\beta}_5 = 16.00$, entonces se puede decir que un incremento de 16 dólares en el precio de venta se relaciona con cada pie cuadrado extra del área total después de ajustar los efectos de los otro cuatro predictores.

Prueba de utilidad del modelo

Con datos multivariados, no hay representación preliminar análoga a una gráfica de dispersión para indicar si se juzga útil un modelo de regresión múltiple particular. El valor de R^2 en realidad comunica un mensaje preliminar, pero este valor a veces es engañoso porque se puede incrementar en gran medida al usar un gran número de predictores (k grande) en relación con el tamaño muestral n (ésta es la razón fundamental de ajustar R^2).

La prueba de utilidad del modelo en la regresión lineal simple tuvo que ver con la hipótesis nula H_0: $\beta_1 = 0$, según lo cual no hay relación útil entre y y el simple predictor x. Aquí se considera la afirmación de que $\beta_1 = 0, \beta_2 = 0, \ldots, \beta_k = 0$, lo cual indica que no hay relación útil entre y y *alguno* de los k predictores. Si por lo menos una de estas β no es cero, resulta(n) útil(es) el(los) predictor(es) correspondiente(s). La prueba se basa en un estadístico que tiene una distribución F particular cuando H_0 es verdadera.

Hipótesis nula: H_0: $\beta_1 = \beta_2 = \cdots = \beta_k = 0$

Hipótesis alternativa: H_a: por lo menos una $\beta_i \neq 0$ $\quad (i = 1, \ldots, k)$

Valor del estadístico de prueba*:

$$f = \frac{R^2/k}{(1 - R^2)/[n - (k + 1)]} \tag{13.19}$$

$$= \frac{\text{SSR}/k}{\text{SSE}/[n - (k + 1)]} = \frac{\text{MSR}}{\text{MSE}}$$

Donde SSR = suma de cuadrados de la regresión = SST − SSE

La región de rechazo para una prueba de nivel α: $f \geq F_{\alpha,k,n-(k+1)}$

Excepto por un múltiplo constante, el estadístico de prueba aquí es $R^2/(1 - R^2)$, la relación de la variación explicada o inexplicada. Si la proporción de variación explicada es alta en relación con la inexplicada, se desearía rechazar H_0 y confirmar la utilidad del modelo. Sin embargo, si k es grande con respecto a n, el factor $[(n - (k + 1))/k]$ hace que f disminuya en forma considerable.

Ejemplo 13.14 Volviendo a los datos de la resistencia cortante de unión del ejemplo 13.12, se ajustó un modelo con $k = 4$ predictores, así que las hipótesis pertinentes son

H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_a: por lo menos una de estas cuatro β no es cero

En la figura 13.14 se muestra el resultado del paquete de estadística JMP. Los valores de s (error cuadrático medio), R^2 y R^2 ajustada en realidad indican un modelo útil. El valor de la relación F de utilidad del modelo es

$$f = \frac{R^2/k}{(1 - R^2)/[n - (k + 1)]} = \frac{0.713959/4}{0.286041/(30 - 5)} = 15.60$$

Este valor también aparece en la columna Relación F de la tabla ANOVA en la figura 13.14. El valor crítico F más grande para los 4 gl del numerador y 25 del denominador en la tabla A.9 del apéndice es 6.49, el cual capta un área de extremo superior de 0.001. Así que valor $P <$ 0.001. La tabla ANOVA en el resultado de JMP muestra que valor $P < 0.0001$. Éste es un resultado muy importante. La hipótesis nula se debe rechazar a algún nivel de significación razonable. Se concluye que hay una relación lineal útil entre y y *por lo menos uno* de los cuatro predictores en el modelo. Esto no significa que los cuatro predictores sean útiles; después se amplía la información al respecto.

Respuesta: resistencia

Resumen de ajuste

R cuadrada	0.713959
R cuadrada ajustada	0.668193
Error cuadrático medio	5.157979
Media de respuesta	38.40667
Observaciones (o pesos de suma)	30

Estimaciones de parámetros

Término	Estimación	Error est.	Relación t	Prob > |t|
Ordenada al origen	–37.47667	13.09964	–2.86	0.0084
Fuerza	0.2116667	0.210574	1.01	0.3244
Potencia	0.4983333	0.070191	7.10	<.0001
Temp.	0.1296667	0.042115	3.08	0.0050
Tiempo	0.2583333	0.210574	1.23	0.2313

Prueba de modelo total

Análisis de varianza

Fuente	GL	Suma de cuadrados	Cuadrado medio	Relación F
Modelo	4	1660.1400	415.035	15.6000
Error	25	665.1187	26.605	Prob > F
C Total	29	2325.2587		<0.0001

Figura 13.14 Resultado de la regresión múltiple del programa JMP para los datos del ejemplo 13.14 ■

Inferencias en la regresión múltiple

Antes de probar las hipótesis, construir los IC y hacer predicciones, se deben examinar primero las gráficas de diagnóstico para ver si el modelo requiere modificación, o si hay valores atípicos en los datos. Las gráficas recomendadas son residuos (estandarizados) en función de cada variable independiente, residuos en función de $\hat{y}$, $\hat{y}$ en función de y, y una gráfica de probabilidad normal de los residuos estandarizados. Los mismos patrones analizados en la sección 13.1 indican los posibles problemas. De particular importancia es la identificación de observaciones que tienen una gran influencia en el ajuste. En la siguiente sección, se describen varias herramientas de diagnóstico adecuadas para esta tarea.

Debido a que cada $\hat{\beta}_i$ es una función lineal de las y_i, la desviación estándar de cada $\hat{\beta}_i$ es el producto de σ y una función de las x_{ij}, así que se obtiene una estimación $s_{\hat{\beta}_i}$ al sustituir s por σ. La función de las x_{ij} es bastante complicada, pero los paquetes de computadora para regresión estándar calculan y muestran las $s_{\hat{\beta}_i}$. Las inferencias relacionadas con una sola β_i se basan en la variable estandarizada

$$T = \frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}}$$

que tiene una distribución t con $n - (k + 1)$ gl.

La estimación puntual de $\mu_{Y \cdot x_1^*, \ldots, x_k^*}$, el valor esperado de Y cuando $x_1 = x_1^*, \ldots, x_k = x_k^*$, es $\hat{\mu}_{Y \cdot x_1^*, \ldots, x_k^*} = \hat{\beta}_0 + \hat{\beta}_1 x_1^* + \cdots + \hat{\beta}_k x_k^*$. La desviación estándar estimada del estimador correspondiente es de nuevo una expresión complicada relacionada con las x_{ij} muestrales. Sin embargo, los mejores paquetes de software para estadística realizan el cálculo si se solicita. Las inferencias acerca de $\mu_{Y \cdot x_1^*, \ldots, x_k^*}$ se basan en estandarizar su estimador para obtener una variable t que tiene $n - (k + 1)$ gl.

1. Un IC de $100(1 - \alpha)\%$ para β_i, el coeficiente de x_i en la función de regresión, es

$$\hat{\beta}_i \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{\beta}_i}$$

2. Una prueba para H_0: $\beta_i = \beta_{i0}$ utiliza el valor del estadístico $t = (\hat{\beta}_i - \beta_{i0})/s_{\hat{\beta}_i}$ basado en $n - (k + 1)$ gl. La prueba es de extremo superior, inferior o bilateral según si H_a contiene la desigualdad $>$, $<$ o $\neq$.

3. Un IC de $100(1 - \alpha)\%$ para $\mu_{Y \cdot x_1^*, \ldots, x_k^*}$ es

$$\hat{\mu}_{Y \cdot x_1^*, \ldots, x_k^*} \pm t_{\alpha/2, n-(k+1)} \cdot \{\text{DE estimada de } \hat{\mu}_{Y \cdot x_1^*, \ldots, x_k^*}\} = \hat{y} \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{Y}}$$

donde $\hat{Y}$ es el estadístico $\hat{\beta}_0 + \hat{\beta}_1 x_1^* + \cdots + \hat{\beta}_k x_k^*$ y $\hat{y}$ es el valor calculado de $\hat{Y}$.

4. Un IP de $100(1 - \alpha)\%$ para un valor futuro de y es

$$\hat{\mu}_{Y \cdot x_1^*, \ldots, x_k^*} \pm t_{\alpha/2, n-(k+1)} \cdot \{s^2 + (\text{DE estimada de } \hat{\mu}_{Y \cdot x_1^*, \ldots, x_k^*})^2\}^{1/2} = \hat{y} \pm t_{\alpha/2, n-(k+1)} \cdot \sqrt{s^2 + s_{\hat{Y}}^2}$$

Los intervalos simultáneos para los cuales se controla el nivel simultáneo de confianza o predicción se obtiene al aplicar la técnica de Bonferroni.

Ejemplo 13.15 La adsorción de suelo y sedimento, el grado al cual los compuestos químicos se colectan en una forma condensada sobre la superficie, es una característica importante que afecta la efectividad de pesticidas y diversos compuestos utilizados en la agricultura. En el artículo "Adsorption of Phosphate, Arsenate, Methanearsonate, and Cacodylate by Lake and Stream Sediments: Comparisons with Soils" (*J. of Environ. Qual.*, 1984: 499-504) aparecen los siguientes datos (tabla 13.5) acerca de y = índice de adsorción de fosfato, x_1 = cantidad de hierro extraíble y x_2 = cantidad de aluminio extraíble.

Tabla 13.5 Datos para el ejemplo 13.15

Observación	x_1 = Hierro extraíble	x_2 = Alumnio extraíble	y = Índice de adsorción
1	61	13	4
2	175	21	18
3	111	24	14
4	124	23	18
5	130	64	26
6	173	38	26
7	169	33	21
8	169	61	30
9	160	39	28
10	244	71	36
11	257	112	65
12	333	88	62
13	199	54	40

En el artículo se propone el modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

Un análisis de computadora produjo la siguiente información:

Parámetro β_i	Estimación $\hat{\beta}_i$	DE estimada $s_{\hat{\beta}_i}$
β_0	−7.351	3.485
β_1	0.11273	0.02969
β_2	0.34900	0.07131

$R^2 = 0.948$ R^2 ajustada $= 0.938$ $s = 4.379$

$\hat{\mu}_{Y\cdot 160,39} = \hat{y} = -7.351 + (0.11273)(160) + (0.34900)(39) = 24.30$

DE estimada de $\hat{\mu}_{Y\cdot 160,39} = s_{\hat{Y}} = 1.30$

Un IC de 99% para β_1, el cambio en la adsorción esperada relacionado con un incremento unitario en el hierro extraíble mientras se mantiene fijo el aluminio extraíble, requiere $t_{0.005,13-(2+1)} = t_{0.005,10} = 3.169$. El IC es

$$0.11273 \pm (3.169)(0.02969) = 0.11273 \pm 0.09409 \approx (0.019, 0.207)$$

De manera similar, un intervalo de 99% para β_2 es

$$0.34900 \pm (3.169)(0.07131) = 0.34900 \pm 0.22598 \approx (0.123, 0.575)$$

La técnica de Bonferroni implica que el nivel de confianza simultáneo para ambos intervalos es por lo menos 98%.

Un IC de 95% para $\mu_{Y \cdot 160,39}$, la adsorción esperada cuando el hierro extraíble = 160 y el aluminio extraíble = 39, es

$$24.30 \pm (2.228)(1.30) = 24.30 \pm 2.90 = (21.40, 27.20)$$

Un IP de 95% para un valor futuro de adsorción que se observa cuando $x_1 = 160$ y $x_2 = 39$ es

$$24.30 \pm (2.228)\{(4.379)^2 + (1.30)^2\}^{1/2} = 24.30 \pm 10.18 = (14.12, 34.48)$$ ■

Con frecuencia la hipótesis de interés tiene la forma H_0: $\beta_i = 0$ para una i particular. Por ejemplo, después de ajustar el modelo de cuatro predictores del ejemplo 13.12, es posible que el investigador desee probar H_0: $\beta_4 = 0$. De acuerdo con H_0, siempre que los predictores x_1, x_2 y x_3 permanecen en el modelo, x_4 no contiene información útil acerca de y. El valor del estadístico de prueba es la **relación** $t\hat{\beta}_i/s_{\hat{\beta}_i}$. Muchos paquetes de software para estadística producen un informe de la relación t y el valor P correspondiente para cada predictor incluido en el modelo. Por ejemplo, en la figura 13.14 se muestra que siempre y cuando se retengan en el modelo potencia, temperatura y tiempo, se puede eliminar el predictor x_1 = fuerza.

Prueba *F* para grupo de predictores La prueba F de utilidad del modelo resultó apropiada para probar si hay información útil acerca de la variable dependiente en *cualquiera* de los k predictores (es decir, si $\beta_1 = \cdots = \beta_k = 0$). En muchas situaciones, primero se construye un modelo que contiene k predictores y luego se desea saber si alguno de los predictores en un subconjunto particular proporciona información útil acerca de Y. Por ejemplo, un modelo que se emplea para predecir las puntuaciones de exámenes de estudiantes podría incluir un grupo de variables de fondo como los niveles familiares de ingreso y educación y algunas variables escolares características como el tamaño del grupo y el gasto por alumno. Una hipótesis de interés es que se pueden eliminar del modelo los predictores escolares característicos.

Identifíquese a los predictores como $x_1, x_2, \ldots, x_l, x_{l+1}, \ldots, x_k$, de modo que sean los últimos $k - l$ los que se esté considerando eliminar del modelo. Entonces se desea probar

H_0: $\beta_{l+1} = \beta_{l+2} = \cdots = \beta_k = 0$
(por lo tanto el modelo "reducido" $Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_l x_l + \epsilon$ es correcto)

contra

H_a: por lo menos uno entre $\beta_{l+1}, \ldots, \beta_k$ no es 0
(así que en el modelo "completo" $Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$, al menos $k - l$ de los predictores proporciona información útil)

La prueba se realiza ajustando los modelos completo y reducido. Debido a que el modelo completo contiene no sólo los predictores del modelo reducido, sino también algunos predictores extra, éste debe ajustar los datos por lo menos tan bien como el modelo reducido. Es decir, si

se hace que SSE_k sea la suma de los residuos cuadrados para el modelo completo y SSE_l la suma correspondiente para el modelo reducido, entonces $SSE_k \leq SSE_l$*. De manera intuitiva, si SSE_k es mucho más pequeña que SSE_l el modelo completo proporciona un ajuste mucho mejor que el modelo reducido; entonces el estadístico de prueba apropiada debe depender de la reducción $SSE_l - SSE_k$ en la variación inexplicada. El procedimiento formal es

SSE_k = variación inexplicada para el modelo completo.

SSE_l = variación inexplicada para el modelo reducido.

Valor del estadístico de prueba: $$f = \frac{(SSE_l - SSE_k)/(k - l)}{SSE_k/[n - (k + 1)]} \qquad (13.20)$$

Región de rechazo: $f \geq F_{\alpha,k-l,n-(k+1)}$

Ejemplo 13.16 Los datos de la tabla 13.6 se tomaron del artículo "Applying Stepwise Multiple Regression Analysis to the Reaction of Formaldehyde with Cotton Cellulose" (*Textile Research J.*, 1984: 157-165). La variable dependiente y es el valor nominal de planchado permanente, una medida cuantitativa de la resistencia a las arrugas. Las cuatro variables independientes utilizadas en el proceso de construcción del modelo son x_1 = concentración de HCHO (formaldehído), x_2 = proporción de catalizador, x_3 = temperatura de curado y x_4 = tiempo de curado.

Tabla 13.6 Datos para el ejemplo 13.16

Observación	x_1	x_2	x_3	x_4	y	Observación	x_1	x_2	x_3	x_4	y
1	8	4	100	1	1.4	16	4	10	160	5	4.6
2	2	4	180	7	2.2	17	4	13	100	7	4.3
3	7	4	180	1	4.6	18	10	10	120	7	4.9
4	10	7	120	5	4.9	19	5	4	100	1	1.7
5	7	4	180	5	4.6	20	8	13	140	1	4.6
6	7	7	180	1	4.7	21	10	1	180	1	2.6
7	7	13	140	1	4.6	22	2	13	140	1	3.1
8	5	4	160	7	4.5	23	6	13	180	7	4.7
9	4	7	140	3	4.8	24	7	1	120	7	2.5
10	5	1	100	7	1.4	25	5	13	140	1	4.5
11	8	10	140	3	4.7	26	8	1	160	7	2.1
12	2	4	100	3	1.6	27	4	1	180	7	1.8
13	4	10	180	3	4.5	28	6	1	160	1	1.5
14	6	7	120	7	4.7	29	4	1	100	1	1.3
15	10	13	180	3	4.8	30	7	10	100	7	4.6

*Las estimaciones $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_l$ en general serán diferentes para los modelos completo y reducido, por lo que se deben ejecutar dos regresiones múltiples distintas para obtener SSE_l y SSE_k. Sin embargo, si las variables se listan en el orden sugerido, la mayor parte de los paquetes de computadora proporcionan una tabla ANOVA de "sumas de cuadrados secuenciales" para el modelo completo que se puede usar para evitar ajustar el modelo reducido.

Considere el modelo completo que consiste en $k = 14$ predictores: $x_1, x_2, x_3, x_4, x_5 = x_1^2, \ldots, x_8 = x_4^2, x_9 = x_1x_2, \ldots, x_{14} = x_3x_4$ (los predictores de primer y segundo orden). ¿Se justifica la inclusión de los predictores de segundo orden? Es decir, ¿se debe usar el modelo reducido que consiste sólo en los predictores x_1, x_2, x_3 y x_4 ($l = 4$)? El resultado de ajustar los dos modelos es el siguiente:

Parámetro	Estimación para el modelo reducido	Estimación para el modelo completo
β_0	−0.9122	−8.807
β_1	0.16073	0.1768
β_2	0.21978	0.7580
β_3	0.011226	0.10400
β_4	0.10197	0.5052
β_5	—	−0.04393
β_6	—	−0.035887
β_7	—	−0.00003271
β_8	—	−0.01646
β_9	—	0.00588
β_{10}	—	0.002702
β_{11}	—	0.01178
β_{12}	—	−0.0006547
β_{13}	—	0.00242
β_{14}	—	0.002526
R^2	0.692	0.921
SSE	17.4951	4.4782

Las hipótesis por probar son

$$H_0\colon\ \beta_5 = \beta_6 = \cdots = \beta_{14} = 0$$

contra

$$H_a\colon\ \text{al menos una entre las } \beta_5, \ldots, \beta_{14} \text{ no es cero}$$

Con $k = 14$ y $l = 4$, el valor crítico F para la prueba con $\alpha = 0.01$ is $F_{0.01,10,15} = 3.80$. El valor del estadístico de prueba es

$$f = \frac{(17.4951 - 4.4782)/10}{4.4782/15} = \frac{1.3017}{0.2985} = 4.36$$

Puesto que $4.36 \geq 3.80$, se rechaza H_0. Se concluye que el modelo apropiado debe incluir por lo menos uno de los predictores de segundo orden. ■

Evaluación de la adecuación del modelo

Los residuos estandarizados en la regresión múltiple resultan de dividir cada residuo entre su desviación estándar estimada; la fórmula para estas desviaciones estándar es sustancialmente más complicada que en el caso de la regresión lineal simple. Se recomienda una gráfica de probabilidad normal de los residuos estandarizados como base para validar la suposición de nor-

malidad. Las gráficas de los residuos estandarizados en función de cada predictor y en función de $\hat{y}$ no deben mostrar un patrón discernible. Las gráficas de residuos ajustados también pueden ser útiles en este cometido. El libro de Neter *et al.*, es una referencia muy útil.

Ejemplo 13.17 En la figura 13.15 se muestra una gráfica de probabilidad normal de los residuos estandarizados para los datos de adsorción y el modelo ajustado del ejemplo 13.15. La forma recta de la gráfica arroja pocas dudas acerca de la suposición de que la desviación aleatoria ϵ tiene una distribución normal.

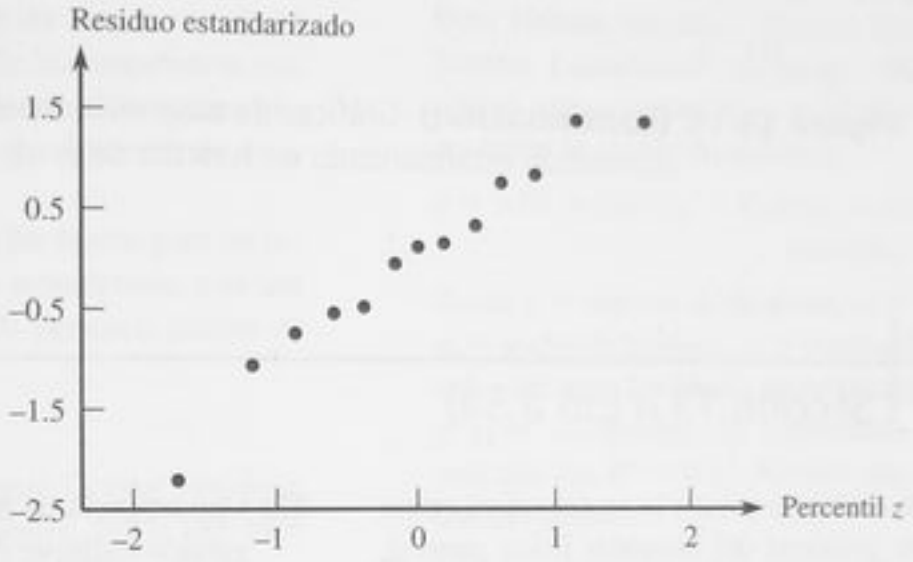


Figura 13.15 Gráfica de probabilidad normal de los residuos estandarizados para los datos y el modelo del ejemplo 13.15

En la figura 13.16 se muestran las otras gráficas sugeridas para los datos de adsorción. Dado que sólo hay 13 observaciones en el conjunto de datos, no hay mucha evidencia de un patrón en ninguna de las tres primeras gráficas que no sea aleatoriedad. El punto en la parte inferior de cada una de estas tres gráficas corresponde a la observación con el residuo grande. Más adelante se amplía la información acerca de estas observaciones. Por el momento, no hay razón convincente para la acción correctiva.

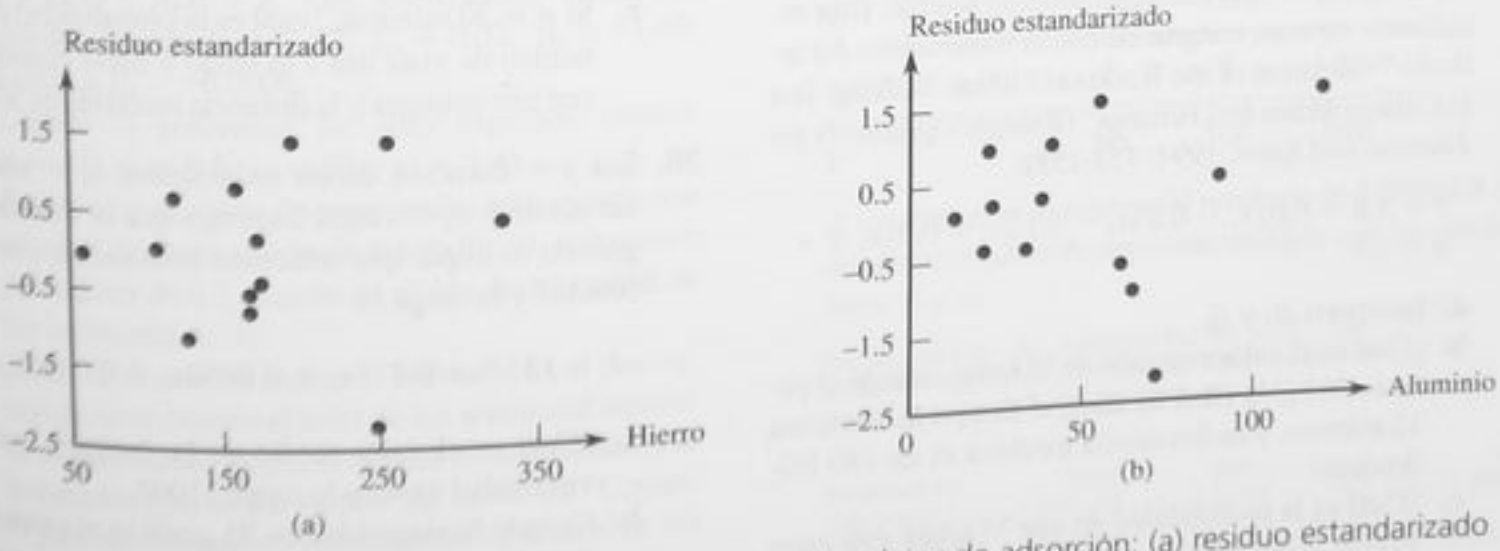
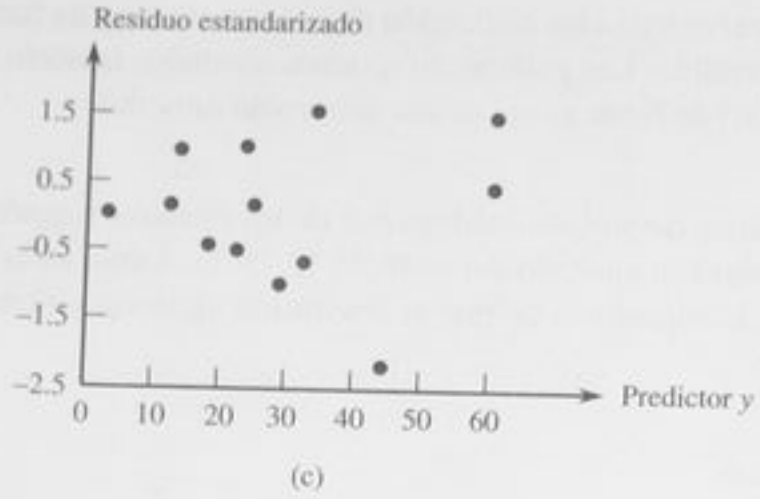
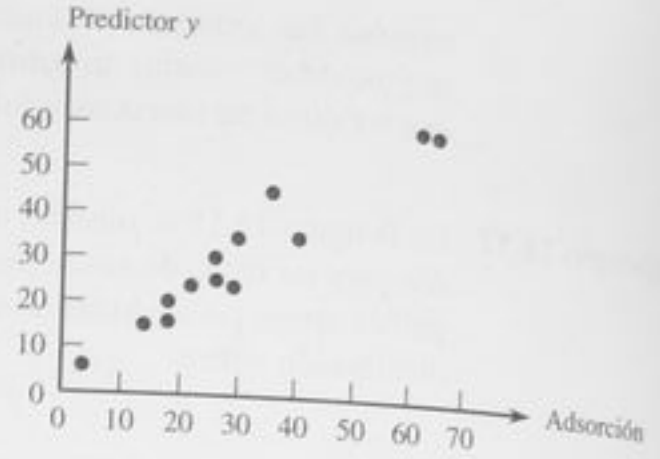


Figura 13.16 Gráfica de diagnóstico para los datos de adsorción: (a) residuo estandarizado en función de x_1; (b) Residuo estandarizado en función de x_2



(c)



(d)

Figura 13.16 (continuación) Gráficas de diagnóstico para los datos de adsorción: (c) residuo estandarizado en función de $\hat{y}$; (d) $\hat{y}$ en función de y ■

Ejercicios | Sección 13.4 (36 a 54)

36. La adecuación cardiorrespiratoria se reconoce como un componente principal del bienestar físico general. La medición directa de la captación máxima de oxígeno (VO_2máx) es la mejor medida de la adecuación, pero la medición directa es lenta y cara. En consecuencia, se desea tener una ecuación de predicción para VO_2máx en términos de cantidades obtenidas con facilidad. Considere las variables

$y = VO_2$máx (L/min) $\quad x_1 =$ peso (kg)
$x_2 =$ Edad (años)

$x_3 =$ tiempo necesario para caminar una milla (min).

$x_4 =$ frecuencia cardiaca al final de la caminata (latidos/min).

A continuación se presenta un modelo posible, para estudiantes varones, congruente con la información del artículo "Validation of the Rockport Fitness Walking Test in College Males and Females" (*Research Quarterly for Exercise and Sport*, 1994: 152-158):

$$Y = 5.0 + 0.01x_1 - 0.05x_2 - 0.13x_3 - 0.01x_4 + \epsilon$$
$$\sigma = 0.4$$

a. Interprete β_1 y β_3.

b. ¿Cuál es el valor esperado de VO_2máx cuando el peso es 76 kg, la edad 20 años, el tiempo de caminata 12 minutos y la frecuencia cardiaca es de 140 latidos/min?

c. ¿Cuál es la probabilidad de que VO_2máx esté entre 1.00 y 2.60 para una sola observación realizada cuando los valores de los predictores son como los mencionados en el inciso (b)?

37. Una empresa de transporte consideró un modelo de regresión múltiple para relacionar la variable dependiente $y =$ tiempo de recorrido diario total para uno de sus conductores (horas) con los predictores $x_1 =$ distancia recorrida (millas) y $x_2 =$ cantidad de entregas. Suponga que la ecuación del modelo es

$$Y = -0.800 + 0.060x_1 + 0.900x_2 + \epsilon$$

a. ¿Cuál es el valor medio del tiempo de viaje cuando la distancia recorrida es 50 millas y se hacen tres entregas?

b. ¿Cómo interpretaría $\beta_1 = 0.060$, el coeficiente del predictor x_1? ¿Cuál es la interpretación de $\beta_2 = 0.900$?

c. Si $\sigma = 30$ minutos, ¿cuál es la probabilidad de que el tiempo de viaje sea a lo sumo 6 horas cuando se hacen tres entregas y la distancia recorrida es 50 millas?

38. Sea $y =$ duración de un rodamiento, $x_1 =$ viscosidad del aceite y $x_2 =$ carga. Suponga que el modelo de regresión múltiple que relaciona la duración con la viscosidad y la carga es

$$Y = 125.0 + 0.775x_1 + 0.0950x_2 - 0.0090x_1x_2 + \epsilon$$

a. ¿Cuál es el valor medio de la duración cuando la viscosidad es 40 y la carga 1100?

b. Cuando la viscosidad es 30, ¿cuál es el cambio en la duración media relacionada con un incremento unitario en la carga? Cuando la viscosidad es 40, ¿cuál es el cambio en la vida media relacionada con un aumento unitario en la carga?

39. Sea y = ventas en un establecimiento de comida rápida (miles de dólares), x_1 = cantidad de establecimientos de la competencia dentro de un radio de una milla, x_2 = población dentro de un radio de una milla (miles de personas) y x_3 una variable indicadora que es igual a 1 si el establecimiento tiene ventanilla de autoservicio y 0 en caso contrario. Suponga que el modelo de regresión real es

$$Y = 10.00 - 1.2x_1 + 6.8x_2 + 15.3x_3 + \epsilon$$

   a. ¿Cuál es el valor promedio de las ventas cuando la cantidad de establecimientos de la competencia son 2, hay 8000 personas dentro de un radio de una milla y el establecimiento tiene una ventanilla de autoservicio?
   b. ¿Cuál es el valor promedio de las ventas para un establecimiento sin ventanilla de autoservicio, con tres negocios competidores y 5000 personas dentro de un radio de una milla?
   c. Interprete β_3.

40. En el artículo "Readability of Liquid Crystal Displays: A Response Surface" (*Human Factors*, 1983: 185-190) se utilizó un modelo de regresión múltiple con cuatro variables independientes para estudiar la precisión al leer pantallas de cristal líquido. Las variables fueron

y = porcentaje de error para individuos que leen en una pantalla de cristal líquido de cuatro dígitos.

x_1 = nivel de iluminación de fondo (0 a 122 cd/m^2).

x_2 = subtensión de carácter (0.025° a 1.34°).

x_3 = ángulo de visión (0° a 60°).

x_4 = nivel de luz ambiental (20 a 1500 lux).

El ajuste de modelo a los datos fue $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$. Los coeficientes estimados resultantes fueron $\hat{\beta}_0 = 1.52$, $\hat{\beta}_1 = 0.02$, $\hat{\beta}_2 = -1.40$, $\hat{\beta}_3 = 0.02$ y $\hat{\beta}_4 = -0.0006$.

   a. Calcule el porcentaje de error esperado cuando $x_1 = 10$, $x_2 = 0.5$, $x_3 = 50$ y $x_4 = 100$.
   b. Estime el porcentaje de error medio relacionado con un nivel de iluminación de fondo de 20, subtensión de carácter de 0.5, ángulo de visión de 10 y nivel de luz ambiental de 30.
   c. ¿Cuál es el cambio esperado estimado en el porcentaje de error cuando el nivel de luz ambiental aumenta en una unidad mientras que las demás variables se mantienen fijas en los valores del inciso (a)? Conteste para un aumento de 100 unidades en el nivel de luz ambiental.
   d. Explique por qué las respuestas del inciso (c) no dependen de valores fijos de x_1, x_2 y x_3. ¿Bajo cuáles condiciones habría tal dependencia?
   e. El modelo estimado se basó en $n = 30$ observaciones, con SST = 39.2 y SSE = 20.0. Calcule e interprete el coeficiente de determinación múltiple y después realice la prueba de utilidad del modelo mediante $\alpha = 0.05$.

41. La capacidad de los ecologistas para identificar regiones de máxima riqueza de especies podría tener un impacto en la conservación de la diversidad genética, uno de los objetivos principales de la Estrategia de Conservación Mundial. En el artículo "Prediction of Rarities from Habitat Variables: Coastal Plain Plants on Nova Scotian Lakeshores" (*Ecology*, 1992: 1852-1859) se utilizó una muestra de $n = 37$ lagos para obtener la ecuación de regresión estimada

$$y = 3.89 + 0.033x_1 + 0.024x_2 + 0.023x_3 - 0.0080x_4 - 0.13x_5 - 0.72x_6$$

donde y = riqueza de especies, x_1 = área de la cuenca, x_2 = ancho de la playa, x_3 = drenaje deficiente (%), x_4 = color del agua (unidades de color total), x_5 = arena (%) y x_6 = alcalinidad. El coeficiente de determinación múltiple fue $R^2 = 0.83$. Realice una prueba de la utilidad del modelo.

42. Una investigación sobre un proceso de colada por presión dio como resultado los siguientes datos de x_1 = temperatura del horno, x_2 = tiempo de cierre de la matriz y y = diferencia de temperatura en la superficie de la matriz ("A Multiple-Objective Decision-Making Approach for Assessing Simultaneous Improvement in Die Life and Casting Quality in a Die Casting Process", *Quality Engineering*, 1994: 371-383).

x_1	1250	1300	1350	1250	1300
x_2	6	7	6	7	6
y	80	95	101	85	92

x_1	1250	1300	1350	1350
x_2	8	8	7	8
y	87	96	106	108

A continuación se muestra el resultado de MINITAB de ajustar el modelo de regresión múltiple con los predictores x_1 y x_2.

```
La ecuación de regresión es dif. de
temp. = -200 + 0.210 temp. del horno
            + 3.00 tiempo de cierre

                                Desv.   Relación
                      Coef      est.          t      p
Predictor
Constante          -199.56     11.64     -17.14  0.000
temp. del
horno             0.210000  0.008642      24.30  0.000
tiempo de
cierre              3.0000    0.4321       6.94  0.000
```

(continúa)

```
s = 1.058  R cuadrada = 99.1%  R cuadrada (ajus-
tada) = 98.8%

Análisis de varianza

FUENTE        GL       SC        CM        F        P
Regresión      2   715.50   357.75   319.31   0.000
Error          6     6.72     1.12
Total          8   722.22
```

a. Realice una prueba de utilidad del modelo.

b. Calcule e interprete un intervalo de confianza de 95% para β_2, el coeficiente de regresión poblacional de x_2.

c. Cuando $x_1 = 1300$ y $x_2 = 7$, la desviación estándar estimada de $\hat{y}$ es $s_{\hat{y}} = 0.353$. Calcule un intervalo de confianza de 95% para la diferencia de temperatura promedio real cuando la temperatura del horno es 1300 y el tiempo de cierre de la matriz es 7.

d. Calcule un intervalo de predicción de 95% para la diferencia de temperatura que resulta de una sola corrida experimental con una temperatura de horno de 1300 y un tiempo de cierre de matriz de 7.

43. Un experimento realizado para estudiar el efecto de la concentración molar de cobalto (x_1) y la temperatura de calcinación (x_2) en el área superficial de un catalizador de hidróxido de hierro-cobalto (y) dio como resultado los siguientes datos ("Structural Changes and Surface Properties of $Co_xFe_{3-x}O_4$ Spinels", *J. of Chemical Tech. and Biotech.*, 1994: 161-170). Se solicitó al programa SAS para ajustar $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$, donde $x_3 = x_1 x_2$ (un predictor de interacción), y se obtuvo el siguiente resultado.

x_1	0.6	0.6	0.6	0.6	0.6	1.0	1.0
x_2	200	250	400	500	600	200	250
y	90.6	82.7	58.7	43.2	25.0	127.1	112.3

x_1	1.0	1.0	1.0	2.6	2.6	2.6	2.6
x_2	400	500	600	200	250	400	500
y	19.6	17.8	9.1	53.1	52.0	43.4	42.4

x_1	2.6	2.8	2.8	2.8	2.8	2.8	
x_2	600	200	250	400	500	600	
y	31.6	40.9	37.9	27.5	27.3	19.0	

a. Prediga el valor del área superficial cuando el contenido de cobalto es 2.6 y la temperatura es 250, y calcule el valor del residuo correspondiente.

b. Puesto que $\hat{\beta}_1 = -46.0$, es legítimo concluir que si el contenido de cobalto se incrementa en una unidad mientras que los valores de los otros predictores permanecen fijos, ¿se puede esperar que el área superficial disminuya en unas 46 unidades? Explique su razonamiento.

c. ¿Al parecer hay una relación útil entre y y los predictores?

d. Dado que la concentración molar y la temperatura de calcinación permanecen en el modelo, ¿el predictor de interacción x_3 proporciona información útil acerca de y? Exprese y pruebe las hipótesis apropiadas con un nivel de significación de 0.01.

e. La desviación estándar estimada de $\hat{Y}$ cuando la concentración molar es 2.0 y la temperatura de calcinación es 500 es $s_{\hat{y}} = 4.69$. Calcule un intervalo de confianza de 95% para el valor medio del área superficial en estas circunstancias.

Resultado de SAS para el ejercicio 43

```
Variable dependiente: ÁREA SUPERFICIAL

                    Análisis de varianza

  Fuente   GL    Suma de cuadrados   Cuadrado medio   Valor F   Prob>F
  Modelo    3          15223.52829       5074.50943    18.924   0.0001
  Error    16           4290.53971        268.15873
  C Total  19          19514.06800

ECM                16.37555      R cuadrada            0.7801
Media dep.         48.06000      R cuadrada ajustada   0.7389
V.C.               34.07314

                Estimaciones de parámetros

                        Estimación del          Error      T para H0:     Prob.
Variable          GL        parámetro       estándar   Parámetro =0     >|T|
ORDENADA AL ORIGEN  1       185.485740   21.19747682           8.750   0.0001
CONC. DE COBALTO    1       -45.969466   10.61201173          -4.332   0.0005
TEMP.               1        -0.301503    0.05074421          -5.942   0.0001
CONC. TEMP.         1         0.088801    0.02540388           3.496   0.0030
```

44. Los autores del artículo "An Ultracentrifuge Flour Absorption Method" (*Cereal Chemistry*, 1978: 96-101) estudiaron la relación entre la absorción de agua para la harina de trigo y varias características de la harina. En particular, los autores utilizaron un modelo de regresión lineal múltiple de primer orden para relacionar la absorción y (%) con la proteína de harina x_1 (%) y el daño al almidón x_2 (unidades Farrand). A continuación se muestran los datos y el resultado de SPSS.

x_1	x_2	y	x_1	x_2	y
8.5	2	30.9	12.9	24	47.0
8.9	3	32.7	12.0	25	46.8
10.6	3	36.7	12.9	28	45.9
10.2	20	41.9	13.1	28	48.8
9.8	22	40.9	11.4	32	46.2
10.8	20	42.9	13.2	28	47.8
11.6	31	46.3	11.6	35	49.2
12.0	32	47.6	12.1	34	48.3
12.5	31	47.2	11.3	35	48.6
10.9	28	44.0	11.1	40	50.2
12.2	36	47.7	11.5	45	49.6
11.9	28	43.9	11.6	50	53.2
11.3	30	46.8	11.7	55	54.3
13.0	27	46.2	11.7	57	55.8

a. Interprete $\hat{\beta}_1$ y $\hat{\beta}_2$.

b. ¿Qué proporción de la variación observada en la absorción se explica mediante la relación del modelo?

c. ¿El modelo elegido especifica una relación lineal útil entre la absorción y al menos uno de los dos predictores?

d. Si la proteína de harina permanece en el modelo, ¿se justifica la eliminación de daños al almidón como predictor?

e. Cuando $x_1 = 10$ y $x_2 = 25$, $\hat{y} = 42.253$ y $s_{\hat{y}} = 0.350$. Calcule e interprete un intervalo de confianza y un intervalo de predicción.

f. Al incluir un predictor de interacción x_3 se obtiene $\hat{\beta}_3 = -0.04304$ y $s_{\hat{\beta}_3} = 0.01773$. ¿Se debe retener este predictor al nivel de significación 0.01?

45. En el artículo "Analysis of the Modeling Methodologies for Predicting the Strength of Air-Jet Spun Yarns" (*Textile Res. J.*, 1997: 39-44) se describe un estudio realizado para relacionar la tenacidad de un hilo (y, en g/tex) con la cuenta de hilos (x_1 en tex), el porcentaje de poliéster (x_2), la presión en la primera boquilla (x_3, en kg/cm^2) y la presión en la segunda boquilla (x_4, en kg/cm^2). La estimación del término constante en la ecuación de regresión múltiple correspondiente fue 6.121. Los coeficientes estimados para los cuatro predictores fueron −0.082, 0.113, 0.256 y −0.219, respectivamente, y el coeficiente de determinación múltiple fue 0.946.

a. Suponiendo que el tamaño de la muestra fue $n = 25$, enuncie y pruebe las hipótesis adecuadas para decidir si el modelo ajustado especifica una relación lineal útil entre la variable dependiente y al menos uno de los cuatro predictores del modelo.

b. De nuevo con $n = 25$ calcule el valor de R^2 ajustada.

c. Calcule un intervalo de confianza de 99% para la tenacidad promedio real del hilo cuando la cuenta de hilos es 16.5, el hilo contiene 50% de poliéster, la presión de la primera boquilla es 3 y la de la segunda es 5 si la desviación estándar estimada de la tenacidad pronosticada en estas circunstancias es 0.350.

Resultado de SPSS para el ejercicio 44

```
R múltiple            0.98207
R cuadrada            0.96447
R cuadrada ajustada   0.96163
Error estándar        1.09412

Análisis de varianza
             GL Suma de cuadrados Cuadrado medio
Regresión     2         812.37959      406.18980
Residuo      25          29.92755        1.19710

F = 339.31092    F signif. = 0.0000

                    Variables en la ecuación

Variable            B      EE de B   Confianza de 95%   Intervalo de B
ALMIDÓN       0.33563      0.01814            0.29828          0.37298
HARINA        1.44228      0.20764            1.01465          1.86991
(Constante)  19.43976      2.18829           14.93290         23.94662
```

46. Al realizar un análisis de regresión para relacionar y = tiempo de reparación para un sistema de filtración de agua (horas) con x_1 = tiempo transcurrido desde el servicio anterior (meses) y x_2 = tipo de reparación (1 si es eléctrica y 0 si es mecánica), se obtuvo el siguiente modelo basado en $n = 12$ observaciones: $y = 0.950 + 0.400x_1 + 1.250x_2$. Además, SST = 12.72, SSE = 2.09 y $s_{\hat{\beta}_1} = 0.312$.

- **a.** ¿Al parecer hay una relación lineal útil entre el tiempo de reparación y los dos predictores de modelo? Realice una prueba de las hipótesis adecuadas con un nivel de significación de 0.05.
- **b.** Dado que el tiempo transcurrido desde el último servicio permanece en el modelo, ¿el tipo de reparación proporciona información útil acerca del tiempo de reparación? Enuncie y pruebe las hipótesis adecuadas con un nivel de significación de 0.01.
- **c.** Calcule e interprete un intervalo de confianza de 95% para β_2
- **d.** La desviación estándar estimada de una predicción para el tiempo de reparación cuando el tiempo transcurrido es 6 meses y la reparación eléctrica es 0.192. Prediga el tiempo de reparación en estas circunstancias calculando un intervalo de predicción de 99%. ¿El intervalo indica que el modelo estimado proporciona una predicción exacta? ¿Por qué sí o por qué no?

47. El diseño eficaz de ciertos tipos de incineradores para desperdicios municipales requiere contar con información acerca del contenido de energía de los desperdicios. Los autores del artículo "Modeling the Energy Content of Municipal Solid Waste Using Multiple Regression Analysis" (*J. of the Air and Waste Mgmt. Assoc.*, 1996: 650-656) proporcionaron amablemente los siguientes datos acerca de y = contenido de energía (kcal/kg), las tres variables físicas de composición x_1 = % de plásticos en peso, x_2 = % de papel en peso, x_3 = % de basura en peso y la variable de análisis aproximado x_4 = % de humedad en peso para especímenes de desperdicio obtenidos en cierta región.

Obs.	Plásticos	Papel	Basura	Agua	Contenido de energía
1	18.69	15.65	45.01	58.21	947
2	19.43	23.51	39.69	46.31	1407
3	19.24	24.23	43.16	46.63	1452
4	22.64	22.20	35.76	45.85	1553
5	16.54	23.56	41.20	55.14	989
6	21.44	23.65	35.56	54.24	1162
7	19.53	24.45	40.18	47.20	1466
8	23.97	19.39	44.11	43.82	1656
9	21.45	23.84	35.41	51.01	1254
10	20.34	26.50	34.21	49.06	1336
11	17.03	23.46	32.45	53.23	1097
12	21.03	26.99	38.19	51.78	1266
13	20.49	19.87	41.35	46.69	1401
14	20.45	23.03	43.59	53.57	1223
15	18.81	22.62	42.20	52.98	1216
16	18.28	21.87	41.50	47.44	1334
17	21.41	20.47	41.20	54.68	1155
18	25.11	22.59	37.02	48.74	1453
19	21.04	26.27	38.66	53.22	1278
20	17.99	28.22	44.18	53.37	1153
21	18.73	29.39	34.77	51.06	1225
22	18.49	26.58	37.55	50.66	1237
23	22.08	24.88	37.07	50.72	1327
24	14.28	26.27	35.80	48.24	1229
25	17.74	23.61	37.36	49.92	1205
26	20.54	26.58	35.40	53.58	1221
27	18.25	13.77	51.32	51.38	1138
28	19.09	25.62	39.54	50.13	1295
29	21.25	20.63	40.72	48.67	1391
30	21.62	22.71	36.22	48.19	1372

Al usar MINITAB para ajustar un modelo de regresión múltiple con las cuatro variables mencionadas como predictores del contenido de energía, se obtuvo el siguiente resultado:

```
La ecuación de regresión es cont. de
energía = 2245 + 28.9 plásticos
              + 7.64 papel + 4.30 basura
              -37.4 agua

Predictor       Coef    Desv. est.         T       p
Constante     2244.9         177.9     12.62   0.000
Plásticos     28.925         2.824     10.24   0.000
Papel          7.644         2.314      3.30   0.003
Basura         4.297         1.916      2.24   0.034
Agua         -37.354         1.834    -20.36   0.000

s = 31.48  R cuadrada = 96.4 % R cuadrada ajustada=
95.8%

Análisis de varianza

Fuente        GL        SC        CM         F       p
Regresión      4    664931    166233    167.71   0.000
Error         25     24779       991
Total         29    689710
```

- **a.** Interprete los valores de los coeficientes de regresión estimados $\hat{\beta}_1$ y $\hat{\beta}_4$.
- **b.** Enuncie y pruebe las hipótesis apropiadas para decidir si el ajuste del modelo a los datos especifica una relación lineal útil entre el consumo de energía y al menos uno de los cuatro predictores.
- **c.** Dado que % de plástico, % de papel y % de agua permanecen en el modelo, ¿el porcentaje de basura suministra información útil acerca del contenido de energía? Enuncie y pruebe las hipótesis adecuadas con un nivel de significación 0.05.

d. Utilice el hecho de que $s_{\hat{Y}} = 7.46$ cuando $x_1 = 20$, $x_2 = 25$, $x_3 = 40$ y $x_4 = 45$ para calcular un intervalo de confianza de 95% para el contenido de energía promedio real en estas circunstancias. ¿El intervalo obtenido indica que se ha estimado con precisión el contenido promedio de energía?

e. Utilice la información del inciso (d) para predecir el contenido de energía de una muestra de desperdicio que tiene las características especificadas de manera que lleve información acerca de la precisión y la confiabilidad.

48. Un experimento para investigar los efectos de una nueva técnica para desengomar hilo de seda se describió en el artículo "Some Studies in Degumming of Silk with Organic Acids" (*J. Society of Dyers and Colourists*, 1992: 79-86). Una variable de respuesta de interés fue y = pérdida de peso (%). Los experimentadores hicieron observaciones en relación con la pérdida de peso para varios valores de tres variables independientes: x_1 = temperatura (°C) = 90, 100, 110; x_2 = tiempo de tratamiento (min) = 30, 75, 120; x_3 = concentración de ácido tartárico (g/L) = 0, 8, 16. En los análisis de regresión, los tres valores de cada variable se codificaron como −1, 0 y 1, respectivamente, y se obtuvieron los datos siguientes (el valor y_8 del informe fue 19.3, pero al utilizar un valor $y_8 = 20.3$ se obtiene un resultado de regresión idéntico al que aparece en el artículo).

Obs.	1	2	3	4	5	6	7	8
x_1	−1	−1	1	1	−1	−1	1	1
x_2	−1	1	−1	1	0	0	0	0
x_3	0	0	0	0	−1	1	−1	1
y	18.3	22.2	23.0	23.0	3.3	19.3	19.3	20.3

Obs.	9	10	11	12	13	14	15
x_1	0	0	0	0	0	0	0
x_2	−1	−1	1	1	0	0	0
x_3	−1	1	−1	1	0	0	0
y	13.1	23.0	20.9	21.5	22.0	21.3	22.6

Un modelo de regresión múltiple con $k = 9$ predictores, $x_1, x_2, x_3, x_4 = x_1^2, x_5 = x_2^2, x_6 = x_3^2, x_7 = x_1x_2, x_8 = x_1x_3$ y $x_9 = x_2x_3$ se ajustó a los datos, y se obtuvo como resultado $\hat{\beta}_0 = 21.967$, $\hat{\beta}_1 = 2.8125$, $\hat{\beta}_2 = 1.2750$, $\hat{\beta}_3 = 3.4375$, $\hat{\beta}_4 = -2.208$, $\hat{\beta}_5 = 1.867$, $\beta_6 = -4.208$, $\hat{\beta}_7 = -0.975$, $\hat{\beta}_8 = -3.750$, $\hat{\beta}_9 = -2.325$, SSE = 23.379 y $R^2 = 0.938$.

a. ¿Este modelo especifica una relación útil? Exprese y pruebe las hipótesis pertinentes mediante un nivel de significación de 0.01.

b. La desviación estándar estimada de $\hat{\mu}_Y$ cuando $x_1 = \cdots = x_9 = 0$ (es decir, cuando la temperatura = 100, tiempo = 75 y concentración = 8), es 1.248. Calcule un intervalo de confianza de 95% para la pérdida de peso esperada cuando temperatura, tiempo y concentración tienen los valores especificados.

c. Calcule un intervalo de predicción de 95% para un solo valor de pérdida de peso que se observa cuando la temperatura, el tiempo y la concentración tienen valores de 100, 75 y 8, respectivamente.

d. Al ajustar el modelo con sólo x_1, x_2 y x_3 como predictores se obtiene $R^2 = 0.456$ y SSE = 203.82. ¿Hay una relación útil entre la pérdida de peso y por lo menos uno de los predictores de segundo orden $x_4, x_5, \ldots, x_9$? Exprese y pruebe las hipótesis pertinentes.

49. En el artículo "The Influence of Temperature and Sunshine on the Alpha-Acid Contents of Hops" (*Agricultural Meteorology*, 1974: 375-382) aparecen los siguientes datos sobre producción (y), temperatura media en el periodo entre la floración del lúpulo y la cosecha (x_1) y el porcentaje medio de luz solar durante el mismo periodo (x_2) para una variedad de lúpulo.

x_1	16.7	17.4	18.4	16.8	18.9	17.1
x_2	30	42	47	47	43	41
y	210	110	103	103	91	76
x_1	17.3	18.2	21.3	21.2	20.7	18.5
x_2	48	44	43	50	56	60
y	73	70	68	53	45	31

A continuación se muestra una parte del resultado de MINITAB del ajuste del modelo de primer orden $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$ que se empleó en el artículo:

```
Predictor      Coef.   Desv. est.   Relación t      p
Constante    415.11        82.52         5.03  0.000
Temp.         -6.593        4.859        -1.36  0.208
Luz solar     -4.504        1.071        -4.20  0.002

                                    R cuadrada
s = 24.45  R cuadrada = 76.8%  (ajustada) = 71.6%
```

a. ¿Cuál es $\hat{\mu}_{Y\cdot 18.9,43}$ y cuál es el residuo correspondiente?

b. Pruebe H_0: $\beta_1 = \beta_2 = 0$ contra H_a: ya sea β_1 o $\beta_2 \neq 0$ al nivel 0.05.

c. La desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2$ cuando $x_1 = 18.9$ y $x_2 = 43$ es 8.20. Utilice esto para obtener un intervalo de confianza de 95% para $\mu_{Y\cdot 18.9,43}$.

d. Utilice la información del inciso (c) para obtener un intervalo de predicción de 95% de la producción en un experimento futuro cuando $x_1 = 18.9$ y $x_2 = 43$.

e. El informe de MINITAB muestra que un intervalo de predicción de 95% para la producción cuando $x_1 = 18$ y $x_2 = 45$ es (35.94, 151.63). ¿Cuál es un IP de 90% en esta situación?

f. Dado que x_2 está en el modelo, ¿retendría x_1?

g. Cuando se ajusta el modelo $Y = \beta_0 + \beta_2 x_2 + \epsilon$, el valor resultante de R^2 es 0.721. Compruebe que el estadístico F para probar H_0: $Y = \beta_0 + \beta_2 x_2 + \epsilon$ contra H_a: $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ satisface $t^2 = f$, donde t es el valor del estadístico t del inciso (f).

50. a. Cuando el modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 + \epsilon$ se ajusta a los datos de lúpulo del ejercicio 49, la estimación de β_5 es $\hat{\beta}_5 = 0.557$ con desviación estándar estimada $s_{\hat{\beta}_5} = 0.94$. Pruebe H_0: $\beta_5 = 0$ contra H_a: $\beta_5 \neq 0$.

b. Cada relación t $\hat{\beta}_i / s_{\hat{\beta}_i}$ $(i = 1, 2, 3, 4, 5)$ para el modelo del inciso (a) es menor que 2 en valor absoluto, pero $R^2 = 0.861$ para este modelo. ¿Sería correcto cancelar cada término del modelo debido a su pequeña relación t? Explique.

c. Con $R^2 = 0.861$ para el modelo del inciso (a), pruebe H_0: $\beta_3 = \beta_4 = \beta_5 = 0$ (que dice que se pueden eliminar los términos de segundo orden).

51. En el artículo "The Undrained Strength of Some Thawed Permafrost Soils" (*Canadian Geotechnical J.*, 1979: 420-427) contiene los siguientes datos en relación con la resistencia al corte no drenada de suelos arenosos (y, en kPa), profundidad (x_1, en m) y contenido de agua (x_2 en %).

	y	x_1	x_2	$\hat{y}$	$y - \hat{y}$	e^*
1	14.7	8.9	31.5	23.35	−8.65	−1.50
2	48.0	36.6	27.0	46.38	1.62	0.54
3	25.6	36.8	25.9	27.13	−1.53	−0.53
4	10.0	6.1	39.1	10.99	−0.99	−0.17
5	16.0	6.9	39.2	14.10	1.90	0.33
6	16.8	6.9	38.3	16.54	0.26	0.04
7	20.7	7.3	33.9	23.34	−2.64	−0.42
8	38.8	8.4	33.8	25.43	13.37	2.17
9	16.9	6.5	27.9	15.63	1.27	0.23
10	27.0	8.0	33.1	24.29	2.71	0.44
11	16.0	4.5	26.3	15.36	0.64	0.20
12	24.9	9.9	37.8	29.61	−4.71	−0.91
13	7.3	2.9	34.6	15.38	−8.08	−1.53
14	12.8	2.0	36.4	7.96	4.84	1.02

Los valores pronosticados y residuos se calcularon al ajustar un modelo cuadrático completo, el cual dio como resultado la función de regresión estimada

$$y = -151.36 - 16.22x_1 + 13.48x_2 + 0.094x_1^2 - 0.253x_2^2 + 0.492x_1x_2$$

a. ¿Las gráficas de e^* en función de x_1, e^* en función de x_2 y e^* en función de $\hat{y}$ indican que debe ser modificado el modelo cuadrático completo? Explique su respuesta.

b. El valor de R^2 para el modelo cuadrático completo es 0.759. Pruebe al nivel 0.05 la hipótesis nula de que no hay relación lineal entre la variable dependiente y alguno de los cinco predictores.

c. Se puede demostrar que $V(Y) = \sigma^2 = V(\hat{Y}) + V(Y - \hat{Y})$. La estimación de σ es $\hat{\sigma} = s = 6.99$ (del modelo cuadrático completo). Primero obtenga la desviación estándar estimada de $Y - \hat{Y}$ y después estime la desviación estándar de $\hat{Y}$ (es decir, $\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_1^2 + \hat{\beta}_4 x_2^2 + \hat{\beta}_5 x_1 x_2$) cuando $x_1 = 8.0$ y $x_2 = 33.1$. Por último, calcule un intervalo de confianza de 95% para la resistencia promedio. [*Sugerencia*: ¿Cuál es $(y - \hat{y})/e^*$?]

d. Al ajustar el modelo de primer orden con función de regresión $\mu_{Y \cdot x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ se obtiene SSE = 894.95. Pruebe al nivel 0.05 la hipótesis nula que expresa que los términos cuadráticos se pueden eliminar del modelo.

52. La utilización de sucrosa como fuente de carbono para la producción de compuestos químicos resulta poco rentable. La melaza de remolacha es un sustituto asequible y de bajo precio. En el artículo "Optimization of the production of β-Carotene from Molasses by Blakeslea Trispora (*J. of Chemical Technology and Biotechnology*, 2002: 933-943) se realizó un análisis de regresión múltiple para relacionar la variable dependiente y = cantidad de β-caroteno (g/dm^3) con los tres predictores cantidad de ácido linoleico, cantidad de queroseno y cantidad de antioxidante (todas en g/dm^3).

Obs.	Linoleico	Queroseno	Antioxidante	Beta-caroteno
1	30.00	30.00	10.00	0.7000
2	30.00	30.00	10.00	0.6300
3	30.00	30.00	18.41	0.0130
4	40.00	40.00	5.00	0.0490
5	30.00	30.00	10.00	0.7000
6	13.18	30.00	10.00	0.1000
7	20.00	40.00	5.00	0.0400
8	20.00	40.00	15.00	0.0065

(continúa)

Obs.	Linoleico	Queroseno	Antioxidante	Beta-caroteno
9	40.00	20.00	5.00	0.2020
10	30.00	30.00	10.00	0.6300
11	30.00	30.00	1.59	0.0400
12	40.00	20.00	15.00	0.1320
13	40.00	40.00	15.00	0.1500
14	30.00	30.00	10.00	0.7000
15	30.00	46.82	10.00	0.3460
16	30.00	30.00	10.00	0.6300
17	30.00	13.18	10.00	0.3970
18	20.00	20.00	5.00	0.2690
19	20.00	20.00	15.00	0.0054
20	46.82	30.00	10.00	0.0640

a. Al ajustar el modelo completo de segundo orden en los tres predictores se obtuvo como resultado $R^2 = 0.978$ y R^2 ajustada $= 0.974$, mientras que al ajustar el modelo de primer orden se obtuvo $R^2 = 0.016$. ¿Qué concluiría acerca de los dos modelos?

b. Para $x_1 = x_2 = 30$, $x_3 = 10$, con un paquete de software para estadística se encontró que $\hat{y} = 0.66573$, $s_{\hat{Y}} = 0.01785$ con base en el modelo completo de segundo orden. Prediga la cantidad de β-caroteno que resultaría de una sola corrida experimental con valores designados de las variables independientes. Realice el cálculo de modo que lleve información acerca de la precisión y la confiabilidad.

53. Las zonas llenas de nieve contienen un amplio espectro de contaminantes que podrían representar riegos ambientales. En el artículo "Atmospheric PAH Deposition: Deposition Velocities and Washout Ratios" (*J. of Environmental Engineering*, 2002: 186-195) se trató el depósito de hidrocarburos poliaromáticos. Los autores propusieron un modelo de regresión múltiple para relacionar el depósito en un periodo especificado (y, en $\mu g/m^2$) para dos predictores bastante complicados x_1 ($\mu g\text{-}s/m^3$) y x_2 ($\mu g/m^2$) definidos en términos de concentraciones aéreas de PAH para varias especies, tiempo total y cantidad total de precipitación. A continuación se presentan los datos de la especie fluoranteno y el resultado de MINITAB correspondiente.

```
Obs.      x1          x2      Fluoranteno
  1    92017   0.0026900           278.78
  2    51830   0.0030000           124.53
  3    17236   0.0000196            22.65
  4    15776   0.0000360            28.68
  5    33462   0.0004960            32.66
  6   243500   0.0038900           604.70
  7    67793   0.0011200            27.69
  8    23471   0.0006400            14.18
  9    13948   0.0004850            20.64
 10     8824   0.0003660            20.60
 11     7699   0.0002290            16.61
 12    15791   0.0014100            15.08
 13    10239   0.0004100            18.05
 14    43835   0.0000960            99.71
 15    49793   0.0000896            58.97
 16    40656   0.0026000           172.58
 17    50774   0.0009530            44.25

La ecuación de regresión es

fluoranteno= -33.5 + 0.00205 x1 + 29836 x2

Predictor        Coef  Coef. de EE       T       P
Constante     -33.46        14.90   -2.25   0.041
x1         0.0020548    0.0002945    6.98   0.000
x2             29836        13654    2.19   0.046

S = 44.28  R cuadrada = 92.3%
R-cuadrada(ajustada) = 91.2%

Análisis de varianza

Fuente           GL      SC      CM      F      P
Regresión         2  330989  165495  84.39  0.000
Error residual   14   27454    1961
Total            16  358443
```

Formule preguntas y lleve a cabo el análisis apropiado para sacar conclusiones.

54. Los siguientes datos se obtuvieron de un estudio de la relación entre el brillo del papel acabado (y) y las variables % en peso de H_2O_2 (x_1), % en peso de NaOH (x_2), % en peso silicato (x_3) y temperatura del proceso (x_4) ("Advantages of CEHDP Bleaching for High Brightness Kraft Pulp Production", *TAPPI*, 1964: 170A-173A). Se permitió que cada variable independiente tomara cinco valores diferentes, y éstos se codificaron para el análisis de regresión como −2, −1, 0, 1 y 2.

Prueba No.	H_2O_2 (x_1)	Conc. de NaOH (x_2)	Conc. de silicato (x_3)	Temp. (x_4)	Brillo (y)
1	−1	−1	−1	−1	83.9
2	+1	−1	−1	−1	84.9
3	−1	+1	−1	−1	83.4
4	+1	+1	−1	−1	84.2
5	−1	−1	+1	−1	83.8
6	+1	−1	+1	−1	84.7
7	−1	+1	+1	−1	84.0
8	+1	+1	+1	−1	84.8
9	−1	−1	−1	+1	84.5
10	+1	−1	−1	+1	86.0
11	−1	+1	−1	+1	82.6
12	+1	+1	−1	+1	85.1
13	−1	−1	+1	+1	84.5
14	+1	−1	+1	+1	86.0

(continúa)

Prueba No.	H_2O_2 (x_1)	Conc. de NaOH (x_2)	Conc. de silicato (x_3)	Temp. (x_4)	Brillo (y)
15	−1	+1	+1	+1	84.0
16	+1	+1	+1	+1	85.4
17	−2	0	0	0	82.9
18	+2	0	0	0	85.5
19	0	−2	0	0	85.2
20	0	+2	0	0	84.5
21	0	0	−2	0	84.7
22	0	0	+2	0	85.0
23	0	0	0	−2	84.9
24	0	0	0	+2	84.0
25	0	0	0	0	84.5
26	0	0	0	0	84.7
27	0	0	0	0	84.6
28	0	0	0	0	84.9
29	0	0	0	0	84.9
30	0	0	0	0	84.5
31	0	0	0	0	84.6

Variables	−2	−1	0	+1	+2
x_1 Peróxido de hidrógeno (100%), % en peso	0.1	0.2	0.3	0.4	0.5
x_2 NaOH, % en peso	0.1	0.2	0.3	0.4	0.5
x_3 Silicato (41°Bé), % en peso	5	1.5	2.5	3.5	4.5
x_4 Temp. de proceso, °F	130	145	160	175	190

a. Cuando se ajustó un modelo (codificado) que tiene que ver con los términos lineales, los términos cuadráticos y los productos cruzados, la función de regresión estimada fue

$$\begin{aligned} y = {} & 84.67 + 0.650x_1 - 0.258x_2 + 0.133x_3 \\ & + 0.108x_4 - 0.135x_1^2 + 0.028x_2^2 + 0.028x_3^2 \\ & - 0.072x_4^2 + 0.038x_1x_2 - 0.075x_1x_3 \\ & + 0.213x_1x_4 + 0.200x_2x_3 - 0.188x_2x_4 \\ & + 0.050x_3x_4 \end{aligned}$$

Utilice este modelo estimado para pronosticar el brillo cuando se tiene 0.4% de H_2O_2, 0.4% de NaOH, 3.5% de silicato y la temperatura es 175. ¿Cuáles son los valores de los residuos para estos valores de las variables?

b. Exprese la función de región estimada en forma no codificada.

c. SST = 17.2567 y R^2 para el modelo del inciso (a) es 0.885. Cuando se ajusta un modelo que incluye sólo los cuatro términos lineales, el valor resultante de R^2 es 0.721. Exprese y pruebe al nivel 0.05 la hipótesis nula que especifica que los coeficientes de los términos cuadráticos y de producto cruzado de la función de regresión son cero.

d. La función de regresión (codificada) estimada, cuando sólo se incluyen términos lineales, es $\hat{\mu}_{Y\cdot x_1,x_2,x_3,x_4} = 85.5548 + 0.6500x_1 - 0.2583x_2 + 0.1333x_3 + 0.1083x_4$. Cuando $x_1 = x_2 = x_3 = x_4 = 0$, la desviación estándar estimada de $\hat{\mu}_{Y\cdot 0,0,0,0}$ es 0.0772. Suponga que se creía que el brillo esperado para estos valores de las x_i era al menos 85.0. ¿La información descrita contradice esta creencia? Exprese y pruebe las hipótesis pertinentes.

13.5 Otros temas de la regresión múltiple

En esta sección, se mencionan varios temas que podrían surgir cuando se lleva a cabo un análisis de regresión múltiple. Consulte las referencias del capítulo para un tratamiento más extenso de algún tema particular.

Transformaciones en la regresión múltiple

Con frecuencia, las consideraciones teóricas indican una relación no lineal, una variable dependiente y dos o más variables independientes, mientras que en otras ocasiones las gráficas de diagnóstico indican que se debe usar algún tipo de función no lineal. Es común que una transformación linealice el modelo.

Ejemplo 13.18 En un artículo de *Lubrication Eng.* ("Accelerated Testing of Solid Film Lubricants", 1972: 365-372) se describe una investigación de la durabilidad para lubricante de capa sólida. Se llevaron a cabo tres conjuntos de pruebas con cojinetes para muñón en una película tipo Mil-L-8937 en cada combinación de tres cargas (3000, 6000 y 10 000 lb/pulg^2) y tres velocidades (20, 60 y 100 rpm) y se registró la durabilidad (horas) para cada ejecución, como se ilustra en la tabla 13.7.

Tabla 13.7 Datos de la durabilidad para el ejemplo 13.18

s	l(1000's)	w	s	l(1000's)	w
20	3	300.2	60	6	65.9
20	3	310.8	60	10	10.7
20	3	333.0	60	10	34.1
20	6	99.6	60	10	39.1
20	6	136.2	100	3	26.5
20	6	142.4	100	3	22.3
20	10	20.2	100	3	34.8
20	10	28.2	100	6	32.8
20	10	102.7	100	6	25.6
60	3	67.3	100	6	32.7
60	3	77.9	100	10	2.3
60	3	93.9	100	10	4.4
60	6	43.0	100	10	5.8
60	6	44.5			

En el artículo se comenta que una distribución lognormal es apropiada para W, puesto que se sabe que $\ln(W)$ sigue una ley normal (recuerde del capítulo 4 que esto es lo que define a una distribución lognormal). El modelo que aparece es $W = (c/s^a l^b) \cdot \epsilon$, de donde $\ln(W) = \ln(c) - a\ln(s) - b\ln(l) + \ln(\epsilon)$; así que con $Y = \ln(W)$, $x_1 = \ln(s)$, $x_2 = \ln(l)$, $\beta_0 = \ln(c)$, $\beta_1 = -a$ y $\beta_2 = -b$, se tiene un modelo de regresión lineal múltiple. Después de calcular $\ln(w_i)$, $\ln(s_i)$ y $\ln(l_i)$ para los datos, con un modelo de primer orden en las variables transformadas se obtuvieron los resultados de la tabla 13.8.

Tabla 13.8 Coeficientes estimados y relaciones t para el ejemplo 13.18

Parámetro β_i	Estimación $\hat{\beta}_i$	DE estimada $s_{\hat{\beta}_i}$	$t = \hat{\beta}_i/s_{\hat{\beta}_i}$
β_0	10.8719	0.7871	13.81
β_1	−1.2054	0.1710	−7.05
β_2	−1.3979	0.2327	−6.01

El coeficiente de determinación múltiple (para el ajuste transformado) tiene un valor de $R^2 = 0.781$. La función de regresión estimada para las variables transformadas es

$$\ln(w) = 10.87 - 1.21\ln(s) - 1.40\ln(l)$$

de modo que la función de regresión original se estima como

$$w = e^{10.87} \cdot s^{-1.21} \cdot l^{-1.40}$$

El método de Bonferroni se puede usar para obtener IC simultáneos para β_1 y β_2 y, como $\beta_1 = -a$ y $\beta_2 = -b$, se tienen de inmediato intervalos para a y b. ■

En la sección 13.2, se presentó el modelo de regresión logística para relacionar una variable dicótoma y con un solo predictor. Este modelo se puede ampliar de una manera obvia para incorporar más de un predictor.

Estandarización de variables

En la sección 13.3, se consideró transformar x a $x' = x - \bar{x}$ antes de ajustar un polinomio. Para la regresión múltiple, en particular cuando los valores de las variables son de gran magnitud, es ventajoso realizar esta codificación un paso más allá. Sean $\bar{x}_i$ y s_i el promedio muestral y la desviación estándar muestral de las x_{ij} ($j = 1, \ldots, n$). Ahora se codifica cada variable x_i mediante $x_i' = (x_i - \bar{x}_i)/s_i$. La variable codificada x_i' simplemente vuelve a expresar cualquier valor x_i en unidades de desviación estándar arriba o abajo de la media. Así, si $\bar{x}_i = 100$ y $s_i = 20$, $x_i = 130$ se convierte en $x_i' = 1.5$ porque 130 está 1.5 desviaciones estándar arriba de la media de los valores de x_i. Por ejemplo, el modelo codificado de segundo orden completo con dos variables independientes tiene la función de regresión.

$$E(Y) = \beta_0 + \beta_1\left(\frac{x_1 - \bar{x}_1}{s_1}\right) + \beta_2\left(\frac{x_2 - \bar{x}_2}{s_2}\right) + \beta_3\left(\frac{x_1 - \bar{x}_1}{s_1}\right)^2$$
$$+ \beta_4\left(\frac{x_2 - \bar{x}_2}{s_2}\right)^2 + \beta_5\left(\frac{x_1 - \bar{x}_1}{s_1}\right)\left(\frac{x_2 - \bar{x}_2}{s_2}\right)$$
$$= \beta_0 + \beta_1 x_1' + \beta_2 x_2' + \beta_3 x_3' + \beta_4 x_4' + \beta_5 x_5'$$

Los beneficios de codificar son (1) mayor precisión numérica en los cálculos (menor error por redondeo en la computadora) y (2) estimación más precisa que para los parámetros del modelo no codificado, porque cada uno de los parámetros del modelo codificado caracterizan el comportamiento de la función de regresión cerca del centro de los datos y no cerca del origen.

Ejemplo 13.19 En el artículo "The Value and the Limitations of High-Speed Turbo-Exhausters for the Removal of Tar-Fog from Carburetted Water-Gas" (*Soc. Chemical Industriy J.*, 1946: 166-168) se presentan los datos (de la tabla 13.9) acerca de y = contenido de alquitrán (granos/100 pies^3) de una corriente de gas en función de x_1 = velocidad del rotor (rpm) y x_2 = temperatura de entrada del gas (°F). Los datos también se consideran en el artículo "Some aspects of Nonorthogonal Data Analysis" (*J. Quality Technology*, 1973: 67-79), que indica usar el modelo codificado descrito antes.

Tabla 13.9 Datos para el ejemplo 13.19

Corrida	y	x_1	x_2	x_1'	x_2'
1	60.0	2400	54.5	−1.52428	−0.57145
2	61.0	2450	56.0	−1.39535	−0.35543
3	65.0	2450	58.5	−1.39535	0.00461
4	30.5	2500	43.0	−1.26642	−2.22763
5	63.5	2500	58.0	−1.26642	−0.06740
6	65.0	2500	59.0	−1.26642	0.07662
7	44.0	2700	52.5	−0.75070	−0.85948

(continúa)

Corrida	y	x_1	x_2	x_1'	x_2'
8	52.0	2700	65.5	−0.75070	1.01272
9	54.5	2700	68.0	−0.75070	1.37276
10	30.0	2750	45.0	−0.62177	−1.93960
11	26.0	2775	45.5	−0.55731	−1.86759
12	23.0	2800	48.0	−0.49284	−1.50755
13	54.0	2800	63.0	−0.49284	0.65268
14	36.0	2900	58.5	−0.23499	0.00461
15	53.5	2900	64.5	−0.23499	0.86870
16	57.0	3000	66.0	0.02287	1.08472
17	33.5	3075	57.0	0.21627	−0.21141
18	34.0	3100	57.5	0.28073	−0.13941
19	44.0	3150	64.0	0.40966	0.79669
20	33.0	3200	57.0	0.53859	−0.21141
21	39.0	3200	64.0	0.53859	0.79669
22	53.0	3200	69.0	0.53859	1.51677
23	38.5	3225	68.0	0.60305	1.37276
24	39.5	3250	62.0	0.66752	0.50866
25	36.0	3250	64.5	0.66752	0.86870
26	8.5	3250	48.0	0.66752	−1.50755
27	30.0	3500	60.0	1.31216	0.22063
28	29.0	3500	59.0	1.31216	0.07662
29	26.5	3500	58.0	1.31216	−0.06740
30	24.5	3600	58.0	1.57002	−0.06740
31	26.5	3900	61.0	2.34360	0.36465

Las medias y las desviaciones estándar son $\bar{x}_1 = 2991.13$, $s_1 = 387.81$, $\bar{x}_2 = 58.468$ y $s_2 = 6.944$, por lo tanto $x_1' = (x_1 - 2991.13)/387.81$ y $x_2' = (x_2 - 58.468)/6.944$. Con $x_3' = (x_1')^2$, $x_4' = (x_2')^2$, $x_5' = x_1' \cdot x_2'$, ajustar el modelo de segundo orden completo requiere resolver el sistema de seis ecuaciones normales con seis incógnitas. Un análisis de computadora produjo$\hat{\beta}_0 = 40.2660$, $\hat{\beta}_1 = -13.4041$, $\hat{\beta}_2 = 10.2553$, $\hat{\beta}_3 = 2.3313$, $\hat{\beta}_4 = -2.3405$ y $\hat{\beta}_5 = 2.5978$. Entonces la ecuación de regresión estimada es

$$\hat{y} = 40.27 - 13.40x_1' + 10.26x_2' + 2.33x_3' - 2.34x_4' + 2.60x_5'$$

Por consiguiente, si $x_1 = 3200$ y $x_2 = 57.0$, $x_1' = 0.539$, $x_2' = -0.211$, $x_3' = (0.539)^2 = 0.2901$, $x_4' = (-0.211)^2 = 0.0447$ y $x_5' = (0.539)(-0.211) = -0.1139$, por lo tanto

$$\hat{y} = 40.27 - (13.40)(0.539) + (10.26)(-0.211) + (2.33)(0.2901) - (2.34)(0.0447) + (2.60)(-0.1139) = 31.16$$ ■

Selección de variable

Con frecuencia un experimentador tendrá datos de un gran número de predictores y deseará construir un modelo de regresión en el que interviene un subconjunto de predictores. El uso de un subconjunto hará que el modelo resultante sea menos difícil de manejar, en particular si después se reúnen más datos y también da lugar a un modelo que es más fácil de interpretar y entender que uno con muchos más predictores. Dos preguntas fundamentales en relación con la selección de variable son las siguientes:

1. Si se examinan regresiones en las que intervienen los subconjuntos posibles de los predictores para los que están disponibles los datos, ¿qué criterios se deben usar para seleccionar un modelo?
2. Si el número de predictores es demasiado grande para permitir que las regresiones sean examinadas, ¿hay una forma de examinar un número reducido de subconjuntos entre los que se encuentra un buen modelo (o modelos)?

Para tratar la primera pregunta, si el número de predictores es pequeño (p. ej., ≤ 5), entonces no sería demasiado tedioso examinar las regresiones posibles por medio de alguno de los paquetes de software para estadística de fácil acceso. Si están disponibles los datos de por lo menos seis predictores, las posibles regresiones tienen que ver con al menos 64 ($= 2^6$) modelos diferentes. Varios paquetes, para alguna m especificada entre 1 y 10, permiten obtener los m mejores modelos de un solo predictor, los m mejores modelos de dos predictores, etcétera (la palabra "mejor" en este caso significa la SSE más pequeña o, de manera equivalente, la R^2 máxima). Con MINITAB se realiza lo anterior para unos 20 predictores, mientras que BMDP maneja hasta 27. Las SSE correspondientes (o sus funciones) se comparan según alguno de los criterios descritos a continuación. La razón de especificar m mayor que uno es ver si los mejores modelos tienes valores similares de SSE o R^2.

Criterios para la selección de variables De nuevo se usa un subíndice k para denotar una cantidad (p. ej., SSE_k) calculada de un modelo con k predictores (y, por consiguiente, $k + 1$ β_i, porque siempre está incluida β_0). Para un valor fijo de k, es razonable identificar el mejor modelo como el que tiene un valor de SSE_k mínimo. El tema más difícil se relaciona con la comparación de las SSE_k para distintos valores de k. Tres criterios diferentes, cada uno una función simple de SSE_k, se emplean de manera extensa.

1. R_k^2, el coeficiente de determinación múltiple para un modelo de k predictores. Debido a que R_k^2 casi siempre se incrementa cuando k crece (y nunca disminuye), no se tiene interés en la k que maximiza a R_k^2. En vez de eso, se desea identificar una k pequeña para la cual R_k^2 es casi tan grande como R^2 para los predictores del modelo.
2. $MSE_k = SSE/(n - k - 1)$, el error cuadrado medio para un modelo de k predictores. Éste se utiliza por lo común en lugar de R_k^2, porque aunque R_k^2 nunca disminuye al aumentar k, una pequeña disminución en SSE_k obtenida con un predictor extra puede ser más que compensada por una disminución de uno en el denominador de MSE_k. Entonces el objetivo es encontrar un modelo con MSE_k mínimo. Puesto que la R_k^2 *ajustada* $= 1 - MSE_k/MST$, donde $MST = SST/(n - 1)$ es constante en k, el examen de R_k^2 *ajustada* es equivalente a la consideración de MSE_k.
3. El fundamento para este tercer criterio, C_k, es más difícil de entender, pero los analistas lo utilizan mucho. Suponga que m predictores especifican el modelo de regresión, es decir,

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \epsilon \qquad V(\epsilon) = \sigma^2$$

por lo tanto,

$$E(Y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m$$

Considere ajustar un modelo por medio de un subconjunto de k de estos m predictores; por simplicidad, suponga que se utiliza $x_1, x_2, \ldots, x_k$. Entonces, al resolver el sistema de ecuaciones normales, se obtienen las estimaciones $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ (pero, por supuesto, no las

estimaciones de alguna de las β que corresponden a los predictores ajenos al modelo ajustado). El valor esperado verdadero $E(Y)$ se estima mediante $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_k x_k$. Ahora considere el **error total esperado y normalizado de la estimación.**

$$\Gamma_k = \frac{E\left(\sum_{i=1}^{n} [\hat{Y}_i - E(Y_i)]^2\right)}{\sigma^2} = \frac{E(\text{SSE}_k)}{\sigma^2} + 2(k+1) - n \tag{13.21}$$

La segunda igualdad en (13.21) se debe aceptar sin discusión porque requiere un argumento artificioso de valor esperado. Entonces, un subconjunto particular es atractivo si su valor Γ_k es pequeño. Infortunadamente, se desconocen los valores de $E(\text{SSE}_k)$ y σ^2. Una forma de remediar esto es denotar con s^2 la estimación de σ^2 con base en el modelo que incluye los predictores para el cual están disponibles los datos y definir

$$C_k = \frac{\text{SSE}_k}{s^2} + 2(k+1) - n$$

Un modelo deseable se especifica mediante un subconjunto de predictores para el cual C_k es pequeña.

Ejemplo 13.20 En el artículo de revisión de Ron Hocking listado en la bibliografía del capítulo, se describe un análisis de datos tomados de los números de 1974 de la revista *Motor Tred*. La variable dependiente y fue el kilometraje, hubo $n = 32$ observaciones, y los predictores para los que se obtuvieron los datos fueron x_1 = forma de motor (1 = recto y 0 = V), x_2 = número de cilindros, x_3 = tipo de transmisión (1 = manual y 0 = automática), x_4 = número de velocidades de la transmisión, x_5 = tamaño del motor, x_6 = caballos de potencia, x_7 = número de gargantas del carburador, x_8 = relación de transmisión final, x_9 = peso y x_{10} = tiempo para recorrer un cuarto de milla. En la tabla 13.10, se presenta el resumen del análisis. En la tabla se describe para cada k el subconjunto que tiene el SSE_k mínimo; la lectura descendente de la columna de variables indica qué variable se añade al ir de k a $k + 1$ (al pasar de $k = 2$ a $k = 3$, se añaden tanto x_3 como x_{10} y se elimina x_2). La figura 13.17 contiene las gráficas de R_k^2, la R_k^2 ajustada y C_k en función de k; estas gráficas son una ayuda visual importante al seleccionar un subconjunto. La estimación de σ^2 es $s^2 = 6.24$, que es MSE_{10}. Un modelo simple altamente apreciado según los criterios es el que contiene los predictores x_3, x_9 y x_{10}.

Tabla 13.10 Mejores subconjuntos para los datos de kilometraje del ejemplo 13.20

k = número de predictores	Variables	SSE_k	R_k^2	R_k^2 ajustada	C_k
1	9	247.2	0.756	0.748	11.6
2	2	169.7	0.833	0.821	1.2
3	3, 10, −2	150.4	0.852	0.836	0.1
4	6	142.3	0.860	0.839	0.8
5	5	136.2	0.866	0.840	1.8
6	8	133.3	0.869	0.837	3.4
7	4	132.0	0.870	0.832	5.2
8	7	131.3	0.871	0.826	7.1
9	1	131.1	0.871	0.818	9.0
10	2	131.0	0.871	0.809	11.0

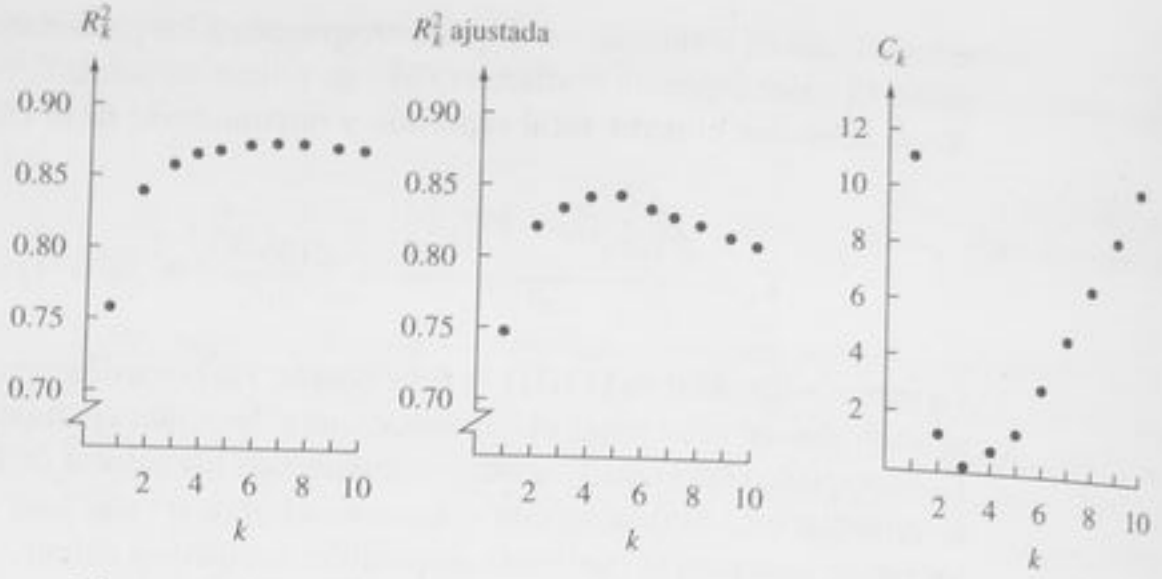


Figura 13.17 Gráfica de R_k^2 y C_k para los datos de kilometraje ■

En términos generales, cuando para ajustar un modelo se emplea un subconjunto de k predictores ($k < m$), los estimadores $\beta_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ serán sesgados para $\beta_0, \beta_1, \ldots, \beta_k$ y $\hat{Y}$ también será un estimador sesgado para la $E(Y)$ verdadera (todo esto porque se pierden $m - k$ predictores del modelo ajustado). Sin embargo, según el error esperado normalizado total Γ_p, las estimaciones basadas en un subconjunto proporcionan más precisión que la obtenida con todos los predictores posibles; en esencia, esta mayor precisión se logra al introducir un sesgo en los estimadores. Un valor de k para el cual $C_k \approx k + 1$ indica que sería pequeño el sesgo relacionado con este modelo de k predictores.

Ejemplo 13.21 En los datos de resistencia cortante de enlace presentados en el ejemplo 13.12 hay valores de cuatro variables independientes distintas x_1–x_4. Se encuentra que el modelo con sólo estas cuatro variables como predictores resultó útil, y no hay razón convincente para considerar la inclusión de los predictores de segundo orden. En la figura 13.18, se ilustra el resultado de MINITAB que se obtiene de solicitar la identificación de dos mejores modelos de cada tamaño.

El mejor modelo de dos predictores, potencia y temperatura, al parecer es una muy buena elección en los conteos: R^2 es significativamente mayor que para los modelos con menos predictores, aunque casi tan grande como para algunos modelos más grandes, la R^2 ajustada está casi a su máximo para estos datos y C_k es pequeña y cercana a $2 + 1 = 3$.

La respuesta es resistencia

Variables	R cuadrada	R cuadrada ajustada	C-p	s	fuerza	potencia	temperatura	tiempo
1	57.7	56.2	11.0	5.9289		X		
1	10.8	7.7	51.9	8.6045			X	
2	68.5	66.2	3.5	5.2070		X	X	
2	59.4	56.4	11.5	5.9136		X		X
3	70.2	66.8	4.0	5.1590		X	X	X
3	69.7	66.2	4.5	5.2078	X	X	X	
4	71.4	66.8	5.0	5.1580	X	X	X	X

Figura 13.18 Resultado de la opción de mejores subconjuntos de MINITAB ■

Regresión por pasos Cuando el número de predictores es demasiado grande para permitir el examen explícito o implícito de los subconjuntos posibles, los buenos modelos se identifican por lo común mediante varios procedimientos alternativos de selección. El procedimiento más simple es la **eliminación hacia atrás** (*backward elimination*, BE). Este método comienza con el modelo en el que se emplean los predictores en consideración. Sea $x_1, \ldots, x_m$ el conjunto de los predictores. Entonces se examina cada relación t $\hat{\beta}_i/s_{\hat{\beta}_i}$ ($i = 1, \ldots, m$) apropiada para probar H_0: $\beta_i = 0$ contra H_a: $\beta_i \neq 0$. Si la relación t con el valor absoluto más pequeño es menor que una constante especificada con anticipación t_{salida}, es decir, si

$$\min_{i=1,\ldots,m} \left| \frac{\hat{\beta}_i}{s_{\hat{\beta}_i}} \right| < t_{\text{salida}}$$

entonces el predictor correspondiente a la relación más pequeña se elimina del modelo. Ahora el modelo reducido está ajustado, se examinan de nuevo las $m - 1$ relaciones t y se elimina otro predictor si corresponde a la relación t absoluta más pequeña que t_{salida}. De este modo, el algoritmo continúa hasta que en algún paso, las relaciones t absolutas son al menos t_{salida}. El modelo utilizado es el que contiene los predictores que no se eliminaron. El valor $t_{\text{salida}} = 2$ se recomienda puesto que la mayor parte de valores $t_{0.05}$ son cercanos a 2. Algunos paquetes de software para computadora se centran en valores P y no en las relaciones t.

Ejemplo 13.22 (Continuación del ejemplo 13.19)

Para el modelo cuadrático modificado en el que y = contenido de alquitrán, los cinco predictores posibles son x_1', x_2', $x_3' = x_1'^2$, $x_4' = x_2'^2$ y $x_5' = x_1'x_2'$ (así que $m = 5$). Sin especificar t_{salida}, en cada etapa se elimina el predictor con la relación t absoluta más pequeña (marcada con el asterisco), y de esta manera se obtiene la secuencia de modelos mostrada en la tabla 13.11.

Tabla 13.11 Resultados de la eliminación hacia atrás para los datos del ejemplo 13.19

		| relación t |				
Etapa	**Predictores**	**1**	**2**	**3**	**4**	**5**
1	1, 2, 3, 4, 5	16.0	10.8	2.9	2.8	1.8*
2	1, 2, 3, 4	15.4	10.2	3.7	2.0*	—
3	1, 2, 3	14.5	12.2	4.3*	—	—
4	1, 2	10.9	9.1*	—	—	—
5	1	4.4*	—	—	—	—

Si se utiliza $t_{\text{salida}} = 2$, el modelo resultante se basaría en x_1', x_2' y x_3', puesto que en el paso 3 ningún predictor se podría eliminar. Se puede comprobar que cada subconjunto es en realidad el mejor subconjunto de su tamaño, aunque esto no siempre es así. ■

Una alternativa al procedimiento BE es la **selección hacia delante** (*forward selection*, FS). FS comienza sin ningún predictor en el modelo y a su vez considera ajustar el modelo con sólo x_1, solamente x_2, ... y, por último, sólo x_m. La variable que, cuando está ajustada produce la relación t absoluta máxima, entra al modelo siempre y cuando la relación exceda la constante especificada t_{entrada}. Suponga que x_1 entra al modelo. Entonces se consideran a su vez los modelos con (x_1, x_2), (x_1, x_3), ..., (x_1, x_m). Por lo tanto, el $|\hat{\beta}_j/s_{\hat{\beta}_j}|$ ($j = 2, \ldots, m$) especifica el

predictor de entrada siempre que su máximo exceda t_{entrada}. Esto continúa hasta que en algún paso ninguna relación absoluta exceda a t_{entrada}. Así, los predictores introducidos especifican el modelo. El valor $t_{\text{entrada}} = 2$ suele utilizarse por la misma razón que $t_{\text{salida}} = 2$ se emplea en BE. Para los datos del contenido de alquitrán, FS da como resultado la secuencia de modelos suministrada en los pasos 5, 4, . . . , 1 en la tabla 13.11 y, por consiguiente, concuerda con BE. Éste no siempre es el caso.

El procedimiento por pasos más utilizado es una combinación de FS y BE, que se representa con FB. Este procedimiento comienza como la selección hacia delante, añadiendo variables al modelo, pero después de cada adición examina las variables introducidas antes para ver si alguna es candidata a eliminación. Por ejemplo, si hay ocho predictores en consideración y el conjunto actual consiste en x_2, x_3, x_5, y x_6 con x_5 recién añadida, se examinan las relaciones t, $\hat{\beta}_2/s_{\hat{\beta}_2}$, $\hat{\beta}_3/s_{\hat{\beta}_3}$ y $\hat{\beta}_6/s_{\hat{\beta}_6}$. Si la relación absoluta más pequeña es menor que t_{salida}, entonces se elimina del modelo la variable correspondiente. La idea que respalda a FB es que con la selección hacia delante una sola variable podría tener una relación estrecha con y que cualquiera de dos o más variables por separado, pero después la combinación de estas variables podría hacer redundante a la variable única. Esto sucede en realidad con los datos de kilometraje analizados en el ejemplo 13.20, con x_2 que entra y después sale del modelo.

El procedimiento FB es parte de varios paquetes de software normales. El paquete BMDP especifica $t_{\text{entrada}} = 2$ y $t_{\text{salida}} = \sqrt{3.9}$ (la mayor parte de los paquetes en realidad utilizan $f = t^2$ en vez de t).

Aunque en muchas situaciones estos procedimientos de selección automática identifican un buen modelo, no hay garantía de que resulte el mejor modelo o uno que casi sea el mejor. Se deben examinar con detalle los conjuntos de datos en los que al parecer hay relaciones firmes entre algunos de los posibles predictores; en breve se amplía esta información.

Identificación de observaciones preponderantes

En una regresión lineal simple, es fácil localizar una observación cuyo valor de x es mucho mayor o mucho menor que otros valores de x en la muestra. Esa observación podría tener un gran impacto en la ecuación de regresión estimada (si depende en realidad de cuan consistente es el valor y correspondiente para el resto de los datos). En la regresión múltiple, también es conveniente saber si los valores de los predictores para una determinada observación son tales que tiene el potencial para ejercer gran influencia en la ecuación estimada. Un método para identificar las observaciones con posibilidades de ser preponderantes se basa en el hecho de que como cada $\hat{\beta}_i$ es una función lineal de $y_1, y_2, \ldots, y_n$, cada valor de y pronosticado de la forma $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_k x_k$ también es una función lineal de las y_j. En particular, los valores previstos que corresponden a observaciones muestrales, se escriben como sigue:

$$\begin{aligned}
\hat{y}_1 &= h_{11}y_1 + h_{12}y_2 + \cdots + h_{1n}y_n \\
\hat{y}_2 &= h_{21}y_1 + h_{22}y_2 + \cdots + h_{2n}y_n \\
&\vdots \quad\quad \vdots \quad\quad \vdots \quad\quad\quad \vdots \\
\hat{y}_n &= h_{n1}y_1 + h_{n2}y_2 + \cdots + h_{nn}y_n
\end{aligned}$$

Cada coeficiente h_{ij} es una función sólo de las x_{ij} en la muestra y no de las y_j. Se puede demostrar que $h_{ij} = h_{ji}$ y que $0 \le h_{jj} \le 1$.

Se centra la atención en los coeficientes de la diagonal $h_{11}, h_{22}, \ldots, h_{nn}$. El coeficiente h_{jj} es el peso que se da a y_j al calcular el valor previsto correspondiente $\hat{y}_j$. Esta cantidad también se puede expresar como una medida de la distancia entre el punto $(x_{1j}, \ldots, x_{kj})$ en el es-

pacio de k dimensiones y el centro de los datos $(\bar{x}_1, \ldots, \bar{x}_k)$. Por consiguiente es natural caracterizar una observación cuya h_{jj} es relativamente grande como una que posiblemente tiene gran influencia. A menos que haya una relación lineal perfecta entre los k predictores, $\sum_{j=1}^{n} h_{jj} = k + 1$, así que el promedio de las h_{jj} es $(k + 1)/n$. Algunos estadísticos sugieren que si $h_{jj} > 2(k + 1)/n$, la j-ésima observación sea citada como potencialmente importante; otros utilizan $3(k + 1)/n$ como línea divisoria.

Ejemplo 13.23 Los siguientes datos aparecieron en el artículo "Testing for the Inclusion of Variables in Linear Regression by a Randomization Technique" (*Technometrics*, 1966: 695-699) y se volvieron a analizar en Hoaglin y Welsch, "The Hat Matrix in Regression and ANOVA" (*Amer. Statistician*, 1978: 17-23). Después de los datos aparecen las h_{ij} (en donde por simetría se omiten los elementos debajo de la diagonal).

Número de viga	Densidad relativa (x_1)	Contenido de humedad (x_2)	Resistencia (y)
1	0.499	11.1	11.14
2	0.558	8.9	12.74
3	0.604	8.8	13.13
4	0.441	8.9	11.51
5	0.550	8.8	12.38
6	0.528	9.9	12.60
7	0.418	10.7	11.13
8	0.480	10.5	11.70
9	0.406	10.5	11.02
10	0.467	10.7	11.41

	1	2	3	4	5	6	7	8	9	10
1	0.418	−0.002	0.079	−0.274	−0.046	0.181	0.128	0.222	0.050	0.242
2		0.242	0.292	0.136	0.243	0.128	−0.041	0.033	−0.035	0.004
3			0.417	−0.019	0.273	0.187	−0.126	0.044	−0.153	0.004
4				0.604	0.197	−0.038	0.168	−0.022	0.275	−0.028
5					0.252	0.111	−0.030	0.019	−0.010	−0.010
6						0.148	0.042	0.117	0.012	0.111
7							0.262	0.145	0.277	0.174
8								0.154	0.120	0.168
9									0.315	0.148
10										0.187

Aquí $k = 2$, de modo que $(k + 1)/n = 3/10 = 0.3$; puesto que $h_{44} = 0.604 > 2(0.3)$, el cuarto dato se identifica como potencialmente influyente. ■

Otra técnica para evaluar la influencia de la j-ésima observación que toma en cuenta a y_j, así como a los predictores, requiere eliminar la j-ésima observación del conjunto de datos y llevar a cabo la regresión con base en las demás observaciones. Si los coeficientes estimados de la regresión con la "observación eliminada" difieren bastante de las estimaciones basadas en todos los datos, es evidente que la j-ésima observación tuvo un impacto sustancial en el ajus-

te. Una forma de juzgar si los coeficientes estimados cambian mucho es expresar cada cambio en relación con la desviación estándar estimada del coeficiente:

$$\frac{(\hat{\beta}_i \text{ antes de la eliminación}) - (\hat{\beta}_i \text{ despúes de la eliminación})}{s_{\hat{\beta}_i}} = \frac{\text{Cambio en } \hat{\beta}_i}{s_{\hat{\beta}_i}}$$

Hay fórmulas de cálculo eficaces que permiten obtener toda esta información de la regresión "sin eliminación", de manera que resultan innecesarias las otras n regresiones.

Ejemplo 13.24 (Continuación del ejemplo 13.23) Considere eliminar por separado las observaciones 1 a 6, cuyos residuos son los más grandes, y la cuarta observación, donde h_{jj} es grande. En la tabla 13.12 se suministra la información pertinente

Tabla 13.12 Cambios en los coeficientes estimados para el ejemplo 13.24

Parámetro	Estimaciones sin eliminaciones	DE estimada	Cambio cuando se elimina el punto j		
			$j = 1$	$j = 4$	$j = 6$
β_0	10.302	1.896	2.710	−2.109	−0.642
β_1	8.495	1.784	−1.772	1.695	0.748
β_2	0.2663	0.1273	−0.1932	0.1242	0.0329
		e_j:	−3.25	−0.96	2.20
		h_{jj}:	0.418	0.604	0.148

Para la eliminación del punto 1 y 4, el cambio en cada estimación está en el intervalo de 1 a 1.5 desviaciones estándar, lo cual es bastante considerable (esto no indica qué sucedería si al mismo tiempo se omitieran ambos puntos). Sin embargo, para el punto 6, el cambio es de alrededor de 0.25 desviaciones estándar. Por consiguiente, se podrían omitir los puntos 1 y 4, pero no el 6, al calcular una ecuación de regresión. ■

Multicolinealidad

En muchos conjuntos de datos de regresión múltiple, los predictores $x_1, x_2, \ldots, x_k$ son muy interdependientes. Suponga que se considera el modelo usual

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$$

con los datos $(x_{1j}, \ldots, x_{kj}, y_j)$ $(j = 1, \ldots, n)$ disponibles para el ajuste. Si se emplea el principio de mínimos cuadrados para llevar a cabo la regresión de x_i en los otros predictores $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$, y se obtiene

$$\hat{x}_i = a_0 + a_1 x_1 + \cdots + a_{i-1} x_{i-1} + a_{i+1} x_{i+1} + \cdots + a_k x_k$$

se puede demostrar que

$$V(\hat{\beta}_i) = \frac{\sigma^2}{\sum_{j=1}^{n} (x_{ij} - \hat{x}_{ij})^2} \tag{13.22}$$

Cuando se pueden prever muy bien los valores muestrales x_i a partir de los otros predictores, el denominador de (13.22) será pequeño, de modo que $V(\hat{\beta}_i)$ será bastante grande Si éste es el

caso para por lo menos un predictor, se dice que los datos muestran **multicolinealidad**. Por lo común, un resultado de regresión obtenido con computadora en el que R^2 es grande indica multicolinealidad, pero algunas de las relaciones $\hat{\beta}_i/s_{\hat{\beta}_i}$ son pequeñas para predictores que, con base en información previa y la intuición, parecen importantes. Otra pista de la presencia de multicolinealidad radica en un valor de $\hat{\beta}_i$ con el signo opuesto al que indicaría la intuición, señal de que otro predictor o colección de predictores sirve como "representante" de x_i.

Una estimación del grado de multicolinealidad se obtiene al realizar una regresión de cada predictor a su vez en los $k - 1$ predictores restantes. Represente con R_i^2 el valor de R^2 en la regresión con variable dependiente x_i y predictores $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$. Según si $R_i^2 > 0.9$ para cualquier i, la multicolinealidad es grave. El programa MINITAB no acepta un predictor en el modelo cuando su valor de R_i^2 es bastante cercano a la unidad.

Infortunadamente no hay consenso entre los estadísticos en cuanto a qué remedios son apropiados cuando hay multicolinealidad grave. Una posibilidad es continuar usando un modelo que incluye los predictores, pero la estimación de los parámetros se hace con otro método distinto al de mínimos cuadrados. Para más detalles consulte una referencia del capítulo.

Ejercicios | Sección 13.5 (55 a 64)

55. En el artículo "Bank Full Discharge of Rivers" (*Water Resources J.*, 1978: 1141-1154) aparecen los datos sobre la cantidad de descarga (q, en m^3/s), área de flujo (a en m^2) y pendiente de la superficie de agua (b, en m/m) obtenidos en varias estaciones de áreas de inundación. A continuación se ilustra un subconjunto de los datos. En el artículo se propone un modelo de potencia multiplicativo $Q = \alpha a^{\beta} b^{\gamma} \epsilon$.

q	17.6	23.8	5.7	3.0	7.5
a	8.4	31.6	5.7	1.0	3.3
b	0.0048	0.0073	0.0037	0.0412	0.0416

q	89.2	60.9	27.5	13.2	12.2
a	41.1	26.2	16.4	6.7	9.7
b	0.0063	0.0061	0.0036	0.0039	0.0025

a. Utilice una transformación adecuada para linealizar el modelo y después estimar los parámetros de regresión para el modelo transformado. Por último, estime α, β y γ (los parámetros del modelo original). ¿Qué predeciría para la cantidad de descarga cuando el área de flujo es 10 y la pendiente 0.01?

b. Sin hacer en realidad ningún análisis, ¿cómo ajustaría un modelo multiplicativo exponencial $Q = \alpha e^{\beta a} e^{\gamma b} \epsilon$?

c. Después de la transformación a linealidad en el inciso (a), un intervalo de confianza de 95% para el valor de la función de regresión transformada cuando $a = 3.3$ y $b = 0.0046$ se obtuvo de una salida impresa de computadora como (0.217, 1.755). Obtenga un intervalo de confianza de 95% para $\alpha a^{\beta} b^{\gamma}$ cuando $a = 3.3$ y $b = 0.0046$.

56. En un experimento para estudiar los factores que influyen en la densidad relativa de la madera ("Anatomical Factors Influencing Wood Specific Gravity of Slash Pilles and the Implications for the Development of a High-Quality Pulpwood", *TAPPI*, 1964: 401-404), se obtuvo una muestra de 20 especímenes de madera madura y se tomaron medidas del número de fibras/mm^2 en la madera de primavera (x_1), cantidad de fibras/mm^2 en madera de verano (x_2), % de madera de primavera (x_3), absorción de luz en madera de primavera (x_4) y absorción de luz en madera de verano (x_5).

a. Al ajustar la función de regresión $\mu_{Y \cdot x_1, x_2, x_3, x_4, x_5} = \beta_0 + \beta_1 x_1 + \cdots + \beta_5 x_5$ se obtuvo $R^2 = 0.769$. ¿Los datos indican que hay una relación lineal entre la densidad relativa y por lo menos uno de los predictores? Pruebe con $\alpha = 0.01$.

b. Cuando se elimina x_2 del modelo, el valor de R^2 aún es 0.769. Calcule la R^2 ajustada para el modelo completo y el modelo en el que se elimina x_2.

c. Cuando se eliminan x_1, x_2 y x_4 el valor resultante de R^2 es 0.654. La suma total de cuadrados es SST = 0.0196610. ¿Estos datos indican que x_1, x_2 y x_4 tienen coeficientes cero en el modelo de regresión real? Pruebe las hipótesis pertinentes al nivel 0.05.

d. La media y la desviación estándar de x_3 fueron 52.540 y 5.4447, respectivamente, mientras que las de x_5 fueron 89.195 y 3.6660, respectivamente. Una vez ajustado el modelo en el que intervienen estas dos variables estandarizadas, la ecuación de regresión estimada fue $y = 0.5255 - 0.0236x_3' + 0.0097x_5'$. ¿Qué valor de la densidad relativa se predeciría para una muestra de madera con % de madera de primavera = 50 y % de absorción de luz en la madera de verano = 90?

e. La desviación estándar estimada del coeficiente estimado $\hat{\beta}_3$ de x_3' (es decir, para $\hat{\beta}_3$ del modelo estandarizado) fue 0.0046. Obtenga un intervalo de confianza de 95% para β_3.

f. Con base en la información de los incisos (d) y (e), ¿cuál es el coeficiente estimado de x_3 del modelo no estandarizado (sólo con los predictores x_3 y x_5) y cuál es la desviación estándar estimada del estimador del coeficiente (es decir, $s_{\hat{\beta}_3}$ para $\hat{\beta}_3$ en el modelo no estandarizado)?

g. La estimación de σ para el modelo con dos predictores es $s = 0.02001$, mientras que la desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_3 x_3' + \hat{\beta}_5 x_5'$ cuando $x_3' = -0.3747$ y $x_5' = -0.2769$ (es decir, cuando $x_3 = 50.5$ y $x_5 = 88.9$) es 0.00482. Calcule un intervalo de predicción de 95% para la densidad relativa cuando el % de madera de primavera = 50.5 y el % de absorción de luz en la madera de verano = 88.9.

57. En la siguiente tabla se ilustra la SSE mínima de cada número de predictores k ($k = 1, 2, 3, 4$) para un problema de regresión donde y = calor acumulado de endurecimiento del cemento, x_1 = % de aluminato de tricalcio, x_2 = % de silicato de tricalcio, x_3 = % de ferrato de aluminio y x_4 = % de silicato de dicalcio.

Número de predictores k	Predictor(es)	SSE
1	x_4	880.85
2	x_1, x_2	58.01
3	x_1, x_2, x_3	49.20
4	x_1, x_2, x_3, x_4	47.86

Además, $n = 13$ y SST = 2715.76.

a. Utilice los criterios analizados en el texto para recomendar el uso de un modelo de regresión específico.

b. ¿La selección hacia delante daría como resultado el mejor modelo de dos predictores? Explique.

58. En el artículo "Creep and Fatigue Characteristics of Ferrocement Slabs" (*J. Ferrocement*, 1984: 309-322) aparecen los datos de y = esfuerzo a la tensión (MPa), x_1 = grosor de losa (cm), x_2 = carga (kg), x_3 = edad en carga (días) y x_4 = tiempo bajo prueba (días), que resultan de pruebas de esfuerzo de $n = 9$ losas de concreto reforzado. Los resultados al aplicar el método de eliminación hacia atrás de selección de variables se resumen en la tabla siguiente. Explique lo que ocurrió en cada paso del procedimiento.

Constante de etapa	1 8.496	2 12.670	3 12.989
x_1	−0.29	−0.42	−0.49
RELACIÓN T	−1.33	−2.89	−3.14
x_2	0.0104	0.0110	0.0116
RELACIÓN T	6.30	7.40	7.33
x_3	0.0059		
RELACIÓN T	0.83		
x_4	−0.023	−0.023	
RELACIÓN T	−1.48	−1.53	
S	0.533	0.516	0.570
R CUADRADA	95.81	95.10	92.82

59. La opción Best Regression de MINITAB se empleó en los datos de densidad relativa de la madera mencionados en el ejercicio 56, y se obtuvo el siguiente resultado. ¿Cuál o cuáles modelos recomienda investigar con mayor detalle?

La respuesta es densidad relativa

Variables	R cuadrada	R cuadrada (ajustada)	C-p	s	f b e n a	f i b e m	% m d e	A b s d l	A b s e l
1	56.4	53.9	10.6	0.021832			X		
1	10.6	5.7	38.5	0.031245					X
1	5.3	0.1	41.7	0.032155		X			
2	65.5	61.4	7.0	0.019975			X		X
2	62.1	57.6	9.1	0.020950		X	X		
2	60.3	55.6	10.2	0.021439	X		X		
3	72.3	67.1	4.9	0.018461	X		X		X
3	71.2	65.8	5.6	0.018807			X	X	X
3	71.1	65.7	5.6	0.018846		X	X		X
4	77.0	70.9	4.0	0.017353	X		X	X	X
4	74.8	68.1	5.4	0.018179		X	X	X	X
4	72.7	65.4	6.7	0.018919	X	X	X		X
5	77.0	68.9	6.0	0.017953	X	X	X	X	X

60. El siguiente resultado de MINITAB se obtuvo al aplicar el método de eliminación hacia atrás y el método de selección hacia adelante a los datos de densidad relativa mencionados en el ejercicio 56. Para cada método, explique lo que ocurrió en cada iteración del algoritmo.

```
La respuesta es dens. rel en 5 predictores,
con N = 20

Etapa              1         2         3         4
Constante     0.4421    0.4384    0.4381    0.5179

fib. de la
m. de p       0.00011   0.00011   0.00012
Valor T          1.17      1.95      1.98

fib. de la
m. de v.      0.00001
Valor T          0.12

% de m.
de p.         0.00531  -0.00526  -0.00498  -0.00438
Valor T         -5.70     -6.56     -5.96     -5.20

abs. de 1
de la m.
de p         -0.0018   -0.0019
Valor T        -1.63     -1.76

abs. de 1 de
la m. de v.   0.0044    0.0044    0.0031    0.0027
Valor T          3.01      3.31      2.63      2.12

R                0.0180    0.0174    0.0185    0.0200
R cuadrada      77.05     77.03     72.27     65.50
```

(*continúa en la parte superior de la siguiente columna*)

```
La respuesta es dens. rel en 5 predictores,
con N = 20

Etapa                  1         2
Constante         0.7585    0.5179

% de m. de p   -0.00444  -0.00438
Valor T           -4.82     -5.20

abs. de 1 de
la m. de v.                 0.0027
Valor T                       2.12

S                 0.0218    0.0200
R cuadrada         56.36     65.50
```

61. Reconsidere los datos de densidad relativa de la madera, mencionados en el ejercicio 56. Los siguientes valores de R^2 se obtuvieron al llevar a cabo la regresión de cada predictor en los otros cuatro predictores (en la primera regresión, la variable dependiente fue x_1 y los predictores fueron x_2 a x_5, etc.): 0.628, 0.711, 0.341, 0.403 y 0.403. ¿La multicolinealidad al parecer es un problema sustancial? Explique.

62. En un estudio que se realizó para investigar la relación entre una variable de respuesta, relativa a las caídas de presión en una columna de burbujeo de malla y placa, y los predictores x_1 = velocidad del líquido superficial, x_2 = viscosidad del líquido y x_3 = tamaño de malla, se obtuvieron los siguientes datos ("A Correlation of Two Phase Pressure Drops in Screen-Plate Bubble Column", *Canad. J. of Chem. Engr.*, 1993: 460-463). Los residuos estandarizados y los valores de h_{ii} se obtuvieron del modelo con sólo x_1, x_2 y x_3 como predictores. ¿Hay alguna observación extraordinaria?

Datos del ejercicio 62

Observación	Velocidad	Viscosidad	Tamaño de malla	Respuesta	Residuo estandarizado	h_{ii}
1	2.14	10.00	0.34	28.9	2.01721	0.202242
2	4.14	10.00	0.34	26.1	1.34706	0.066929
3	8.15	10.00	0.34	22.8	0.96537	0.274393
4	2.14	2.63	0.34	24.2	1.29177	0.224518
5	4.14	2.63	0.34	15.7	−0.68311	0.079651
6	8.15	2.63	0.34	18.3	0.23785	0.267959
7	5.60	1.25	0.34	18.1	0.06456	0.076001
8	4.30	2.63	0.34	19.1	0.13131	0.074927
9	4.30	2.63	0.34	15.4	−0.74091	0.074927
10	5.60	10.10	0.25	12.0	−1.38857	0.152317
11	5.60	10.10	0.34	19.8	−0.03585	0.068468
12	4.30	10.10	0.34	18.6	−0.40699	0.062849
13	2.40	10.10	0.34	13.2	−1.92274	0.175421
14	5.60	10.10	0.55	22.8	−1.07990	0.712933
15	2.14	112.00	0.34	41.8	−1.19311	0.516298
16	4.14	112.00	0.34	48.6	1.21302	0.513214
17	5.60	10.10	0.25	19.2	0.38451	0.152317
18	5.60	10.10	0.25	18.4	0.18750	0.152317
19	5.60	10.10	0.25	15.0	−0.64979	0.152317

63. El resultado de regresión múltiple de MINITAB para los datos de hidrocarburos poliaromáticos del ejercicio 53 en la sección anterior incluyó la siguiente información:

```
Observaciones inusuales

Obs.     x1  Fluor-                 Ajuste            Residuo
             ateno  Ajuste  del EE  Residuo    est.
  6 243500  604.7   582.9    40.7     21.8   1.25X
  7  67793   27.7   139.3    12.3   -111.6  -2.62R

R denota una observación con un gran residuo
  estandarizado.

X denota una observación cuyo valor de X le
  suministra gran influencia.
```

¿Qué indica lo anterior acerca de la conveniencia de usar la ecuación ajustada dada antes como base para las inferencias? Los investigadores eliminaron en realidad la observación #7 y volvieron a realizar la regresión. ¿Suena lógico lo anterior?

64. Refiérase a los datos de descarga de agua del ejercicio 55 y establezca $y = \ln(q)$, $x_1 = \ln(a)$ y $x_2 = \ln(b)$. Considere ajustar el modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$.

a. Las h_{ii} resultantes son 0.138, .302, 0.266, 0.604, 0.464, .360, 0.215, 0.153, 0.214 y 0.284. ¿Resulta preponderante alguna observación?

b. Los coeficientes estimados son $\hat{\beta}_0 = 1.5652$, $\hat{\beta}_1 = 0.9450$, $\hat{\beta}_2 = 0.1815$ y las desviaciones estándar estimadas correspondientes son $s_{\hat{\beta}_0} = 0.7328$, $s_{\hat{\beta}_1} = 0.1528$ y $s_{\hat{\beta}_2} = 0.1752$. El segundo residuo estandarizado es $e_2^* = 2.19$. Cuando la segunda observación se omite del conjunto de datos, los coeficientes estimados resultantes son $\hat{\beta}_0 = 1.8982$, $\hat{\beta}_1 = 1.025$ y $\hat{\beta}_2 = 0.3085$. ¿Alguno de estos cambios indica que la segunda observación es preponderante?

c. La eliminación de la cuarta observación produce $\hat{\beta}_0 = 1.4592$, $\hat{\beta}_1 = 0.9850$ y $\hat{\beta}_2 = 0.1515$ (¿por qué?). ¿Esta observación es preponderante?

Ejercicios complementarios (65 a 82)

65. Se sabe que el curado del concreto es vulnerable a las vibraciones de choque, que podrían causar agrietamiento o daño invisible al material. Como parte de un estudio de fenómenos de vibración, en el artículo "Shock Vibration Test of Concrete" (*ACI Materials J.*, 2002: 361-370) se describen los datos siguientes acerca de la velocidad de partícula máxima (mm/s) y la relación de la velocidad de impulso ultrasónico después del impacto con la anterior al impacto en prismas de concreto.

Obs.	vpm	Relación	Obs.	vpm	Relación
1	160	0.996	16	708	0.990
2	164	0.996	17	806	0.984
3	178	0.999	18	884	0.986
4	252	0.997	19	526	0.991
5	293	0.993	20	490	0.993
6	289	0.997	21	598	0.993
7	415	0.999	22	505	0.993
8	478	0.997	23	525	0.990
9	391	0.992	24	675	0.991
10	486	0.985	25	1211	0.981
11	604	0.995	26	1036	0.986
12	528	0.995	27	1000	0.984
13	749	0.994	28	1151	0.982
14	772	0.994	29	1144	0.962
15	532	0.987	30	1068	0.986

Las grietas transversales aparecieron en los 12 últimos prismas, en tanto que los primeros 18 permanecieron intactos.

a. Construya un diagrama de cuadro comparativo de la vpm para los prismas con fisura y sin fisura y comente al respecto. Luego, estime la diferencia entre la vpm promedio para prismas con fisura y sin fisura de modo que transmita información acerca de la precisión y la confiabilidad.

b. Los investigadores ajustaron el modelo de regresión lineal simple a todo el conjunto de datos que consiste en 30 observaciones, con vpm como la variable independiente y la relación como la variable dependiente. Utilice un paquete de software para estadística a fin de ajustar varios modelos de regresión e infiera lo apropiado.

66. Los autores del artículo "Long-Term Effects of Cathodic Protection on Prestressed Concrete Structures (*Corrosion*, 1997: 891-908) presentaron un diagrama de dispersión de y = flujo de permeación de régimen permanente ($\mu A/cm^2$) en función de x = espesor de hoja inverso (cm^{-1}); el patrón lineal se usó como base para llegar a una conclusión importante acerca del comportamiento del material. A continuación se muestra el resultado de MINITAB de ajustar el modelo de regresión lineal simple a los datos.

```
La ecuación de regresión es
flujo = -0.398 + 0.260 esp. inv.

Predictor      Coef.   Desv. est.   Relación t       p
Constante   -0.3982       0.5051        -0.79   0.460
esp. inv.   0.26042      0.01502        17.34   0.000

s = 0.4506 R cuadrada = 98.0% R cuadrada (ajustada) =
97.7%

Análisis de varianza

Fuente        GL        SC        CM         F       p
Regresión      1    61.050    61.050    300.64   0.000
Error          6     1.218     0.203
Total          7    62.269

                                Ajuste de
     esp.                        la desv.                Residuo
Obs. inv.  Flujo   Ajuste          est.  Residuo           est.
  1  19.8    4.3    4.758         0.242  -0.458          -1.20
  2  20.6    5.6    4.966         0.233   0.634           1.64
  3  23.5    6.1    5.722         0.203   0.378           0.94
  4  26.1    6.2    6.399         0.182  -0.199          -0.48
  5  30.3    6.9    7.493         0.161  -0.593          -1.41
  6  43.5   11.2   10.930         0.236   0.270           0.70
  7  45.0   11.3   11.321         0.253  -0.021          -0.06
  8  46.5   11.7   11.711         0.271  -0.011          -0.03
```

a. Interprete la pendiente estimada y el coeficiente de determinación.

b. Calcule una estimación puntual del flujo promedio real cuando el espesor de hoja inverso es 23.5.

c. ¿El modelo parece útil?

d. Prediga el flujo cuando el espesor inverso es 45 de forma que aporte información acerca de la precisión y la confiabilidad.

e. Investigue qué tan adecuado es el modelo.

67. En el artículo "Validation of the Rockport Fitness Walking Test in College Males and Females" (*Research Quarterly for Exercise and Sport*, 1994: 152-158) se recomienda la siguiente ecuación de regresión estimada para relacionar $y = VO_2$máx (L/min, una medida del buen funcionamiento cardiorrespiratorio) con los predictores x_1 = género (femenino = 0, masculino = 1), x_2 = peso (lb), x_3 = tiempo para caminar una milla (min) y x_4 = frecuencia cardiaca al final de la caminata (latidos/min):

$$y = 3.5959 + 0.6566x_1 + 0.0096x_2 - 0.0996x_3 - 0.0080x_4$$

a. ¿Cómo interpretaría el coeficiente estimado $\hat{\beta}_3 = -0.0996$?

b. ¿Cómo interpretaría el coeficiente estimado $\hat{\beta}_1 = 0.6566$?

c. Suponga una observación realizada en un varón cuyo peso fue de 170 libras, el tiempo de caminata fue 11 minutos y la frecuencia cardiaca fue de 140 pulsaciones/min dio como resultado VO_2máx = 3.15. ¿Qué habría pronosticado para VO_2máx en esta situación y cuál es el valor del residuo correspondiente?

d. Al utilizar SSE = 30.1033 y SST = 102.3922, ¿qué proporción de la variación observada de VO_2máx se puede atribuir a la relación de modelo?

e. Suponiendo un tamaño de muestra de $n = 20$, realice una prueba de hipótesis para decidir si el modelo elegido especifica una relación útil entre VO_2máx y al menos uno de los predictores.

68. El reconocimiento de características de modelos superficiales de partes complicadas es cada vez más importante en el desarrollo de sistemas eficaces de diseño asistido por computadora (*computer-aided design*, CAD). El artículo "A Computationally Efficient Approach to Feature Abstraction in Design-Manufacturing Integration" (*J. of Engr. for Industry*, 1995: 16-27) contenía una gráfica de $\log_{10}$(tiempo de reconocimiento total), con el tiempo en segundos, en función de $\log_{10}$(cantidad de aristas de una parte), de la cual se leyeron los siguientes valores representativos:

Log(aristas)	1.1	1.5	1.7	1.9	2.0	2.1
Log(tiempo)	0.30	0.50	0.55	0.52	0.85	0.98
Log(aristas)	2.2	2.3	2.7	2.8	3.0	3.3
Log(tiempo)	1.10	1.00	1.18	1.45	1.65	1.84
Log(aristas)	3.5	3.8	4.2	4.3		
Log(tiempo)	2.05	2.46	2.50	2.76		

a. ¿Un diagrama de dispersión de log(tiempo) en función de log(aristas) indica una relación lineal aproximada entre estas dos variables?

b. ¿Qué modelo probabilístico para relacionar y = tiempo de reconocimiento con x = cantidad de aristas, conlleva la relación de regresión lineal simple entre las variables transformadas?

c. Las cantidades de resumen calculadas a partir de los datos, son

$$n = 16 \quad \sum x_i' = 42.4 \quad \sum y_i' = 21.69$$
$$\sum (x_i')^2 = 126.34 \quad \sum (y_i')^2 = 38.5305$$
$$\sum x_i' y_i' = 68.640$$

Calcule estimaciones de los parámetros para el modelo del inciso (b) y después obtenga una predicción puntual del tiempo cuando el número de aristas es 300.

69. La presión del aire (lb/pulg^2) y la temperatura (°F) se midieron para un proceso de compresión, en cierto dispositivo de pistón y cilindro, y se obtuvieron los siguientes datos (de *Introduction to Engineering Experimentation*, Prentice-Hall, Inc., 1996, p. 153):

Presión	20.0	40.4	60.8	80.2	100.4
Temperatura	44.9	102.4	142.3	164.8	192.2

(continúa)

Presión	120.3	141.1	161.4	181.9	201.4
Temperatura	221.4	228.4	249.5	269.4	270.8
Presión	220.8	241.8	261.1	280.4	300.1
Temperatura	291.5	287.3	313.3	322.3	325.8
Presión	320.6	341.1	360.8		
Temperatura	337.0	332.6	342.9		

a. ¿Ajustaría el modelo de regresión lineal simple a los datos y lo usaría como base para predecir la temperatura a partir de la presión? ¿Por qué sí o por qué no?

b. Determine un modelo probabilístico adecuado y utilícelo como base para predecir el valor de la temperatura que se obtendría de una presión de 200 en la forma más informativa posible.

70. Un alumno de ingeniería aeronáutica realizó un experimento para estudiar cómo se relaciona el cociente y = sustentación/resistencia con las variables x_1 = posición de cierta superficie de sustentación hacia delante en relación con el ala principal, y x_2 = colocación de la cola respecto al ala principal. Los datos son los siguientes (*Statistics for Engineering Problem Solving*, p. 133):

x_1 (in.)	x_2 (in.)	y
−1.2	−1.2	0.858
−1.2	0	3.156
−1.2	1.2	3.644
0	−1.2	4.281
0	0	3.481
0	1.2	3.918
1.2	−1.2	4.136
1.2	0	3.364
1.2	1.2	4.018

$\bar{y} = 3.428$, SST = 8.55

a. Al ajustar el modelo de primer orden se obtiene SSE = 5.18, mientras que si se incluye $x_3 = x_1x_2$ como predictor, se obtiene SSE = 3.07. Calcule e interprete el coeficiente de determinación múltiple para cada modelo.

b. Aplique una prueba de la utilidad del modelo, mediante $\alpha = 0.05$ para cada uno de los modelos descritos en el inciso (a). ¿Le sorprende alguno de los resultados?

71. Un baño de amoniaco es lo más común para anodizar aleaciones de Pd y Ni. En el artículo "Modelling of Palladium and Nickel in an Ammonia Bath in a Rotary Device" (*Plating and Surface Finishing*, 1997: 102-104) se describe una investigación de cómo las características de composición del baño afectan las propiedades del recubrimiento. Considere los siguientes datos de x_1 = concentración de Pd (g/dm³), x_2 = concentración de Ni (g/dm³), x_3 = pH, x_4 = temperatura (°C), x_5 = densidad de corriente en el cátodo (A/dm²) y y = contenido de paladio (%) del recubrimiento.

Obs.	conc. de pd	conc. de ni	pH	temp	de. de corr.	cont. de pd
1	16	24	9.0	35	5	61.5
2	8	24	9.0	35	3	51.0
3	16	16	9.0	35	3	81.0
4	8	16	9.0	35	5	50.9
5	16	24	8.0	35	3	66.7
6	8	24	8.0	35	5	48.8
7	16	16	8.0	35	5	71.3
8	8	16	8.0	35	3	62.8
9	16	24	9.0	25	3	64.0
10	8	24	9.0	25	5	37.7
11	16	16	9.0	25	5	68.7
12	8	16	9.0	25	3	54.1
13	16	24	8.0	25	5	61.6
14	8	24	8.0	25	3	48.0
15	16	16	8.0	25	3	73.2
16	8	16	8.0	25	5	43.3
17	4	20	8.5	30	4	35.0
18	20	20	8.5	30	4	69.6
19	12	12	8.5	30	4	70.0
20	12	28	8.5	30	4	48.2
21	12	20	7.5	30	4	56.0
22	12	20	9.5	30	4	77.6
23	12	20	8.5	20	4	55.0
24	12	20	8.5	40	4	60.6
25	12	20	8.5	30	2	54.9
26	12	20	8.5	30	6	49.8
27	12	20	8.5	30	4	54.1
28	12	20	8.5	30	4	61.2
29	12	20	8.5	30	4	52.5
30	12	20	8.5	30	4	57.1
31	12	20	8.5	30	4	52.5
32	12	20	8.5	30	4	56.6

a. Ajuste el modelo de primer orden con 5 predictores y evalúe su utilidad. ¿Todos los predictores al parecer son importantes?

b. Ajuste el modelo completo de segundo orden y evalúe su utilidad.

c. ¿El grupo de predictores de segundo orden (con interacción y cuadráticos) al parecer aportan más información útil acerca de y que los predictores de primer orden? Realice una prueba de hipótesis adecuada.

d. Los autores del artículo citado recomendaron el uso de los cinco predictores de primer orden, más otro predictor $x_6 = (\text{pH})^2$. Ajuste este modelo. ¿Los seis predictores al parecer son importantes?

72. En el artículo "An Experimental Study of Resistance Spot Welding in 1 mm Thick Sheet of Low Carbon Steel" (*J. of Engr. Manufacture*, 1996: 341-348) se describe un análisis estadístico cuyo fin básico fue establecer una relación que podría explicar la variación en la resistencia de la soldadura (y) al relacionarla con las características de proceso corriente de soldadura (wc), tiempo de soldadura (wt) y fuerza del electrodo (ef).

a. SST = 16.18555, y al ajustar el modelo completo de segundo orden se obtuvo SSE = 0.80017. Calcule e interprete el coeficiente de determinación múltiple.
b. Si se supone que $n = 37$, realice una prueba de utilidad del modelo (en la tabla de análisis de varianza del artículo se afirma que $n - (k + 1) = 1$, pero otra información lo contradice y es congruente con el tamaño de muestra que se indica aquí).
c. La relación F para la interacción corriente-tiempo fue 2.32. Si se conservan en el modelo los demás predictores, ¿se puede eliminar este predictor de interacción? (*Sugerencia:* al igual que en la regresión lineal simple, una relación F para un coeficiente es el cuadrado de su relación t.)
d. Los autores propusieron eliminar dos predictores de interacción y uno cuadrático, además recomendaron la ecuación estimada $y = 3.352 + 0.098wc + 0.222wt + 0.297ef - 0.0102(wt)^2 - 0.037(ef)^2 + 0.0128(wc)(wt)$. Considere una corriente de soldadura de 10 kA, un tiempo de soldadura de 12 ciclos de ca y una fuerza de electrodo de 6 kN. Suponiendo que la desviación estándar estimada de la resistencia prevista en este caso es 0.0750, calcule un IP de 95% para la resistencia. ¿El intervalo indica que se puede anticipar con exactitud el valor de la resistencia?

73. Los datos siguientes sobre x = frecuencia (MHz) y y = potencia de salida (W) para cierta configuración láser se leyeron de una gráfica del artículo "Frequency Dependence in RF Discharge Excited Waveguide CO_2 Lasers" (*IEEE J. Quantum Electronics*, 1984: 509-514):

x	60	63	77	100	125	157	186	222
y	16	17	19	21	22	20	15	5

Mediante un análisis de computadora se obtuvo la siguiente información para un modelo de regresión cuadrático: $\hat{\beta}_0 = -1.5127$, $\hat{\beta}_1 = 0.391901$, $\hat{\beta}_2 = -0.00163141$, $s_{\hat{\beta}_2} = 0.00003391$, SSE = 0.29, SST = 202.88 y $s_{\hat{Y}} = 0.1141$ cuando $x = 100$.

a. ¿El modelo cuadrático parece ser apropiado para explicar la variación observada en la potencia de salida al relacionarla con la frecuencia?
b. ¿El modelo de regresión lineal simple sería casi tan satisfactorio como el modelo cuadrático?
c. ¿Considera que sería importante tomar en cuenta un modelo cúbico?
d. Calcule un IC de 95% para la salida de potencia esperada cuando la frecuencia es 100.
e. Utilice un IP de 95% para pronosticar la potencia de una sola corrida experimental cuando la frecuencia es 100.

74. La conductividad es una característica importante del vidrio. En el artículo "Structure and Properties of Rapidly Quenched Li_2O-Al_2O-Nb_2O_5 Glasses" (*J. Amer. Ceramic Soc.*, 1983: 890-892) aparecen los siguientes datos sobre x = contenido de Li_2O de cierto tipo de vidrio y y = conductividad a 500 K.

x	19	20	24	27	29	30
y	$10^{-8.0}$	$10^{-7.1}$	$10^{-7.2}$	$10^{-6.7}$	$10^{-6.2}$	$10^{-6.8}$
x	31	39	40	43	45	50
y	$10^{-5.8}$	$10^{-5.3}$	$10^{-6.0}$	$10^{-4.7}$	$10^{-5.4}$	$10^{-5.1}$

(Éste es un subconjunto de los datos mostrados en el artículo). Proponga un modelo apropiado para relacionar y con x, estime los parámetros del modelo y pronostique la conductividad cuando el contenido de Li_2O es 35.

75. El efecto del manganeso (Mn) sobre el crecimiento del trigo se examina en el artículo "Manganese Deficiency and Toxicity Effects on Growth, Development and Nutrient Composition in Wheat" (*Agronomy J.*, 1984: 213-217). Se utilizó un modelo de regresión cuadrática para relacionar y = altura de planta (cm) con $x = \log_{10}$ (Mn agregado), con μM como las unidades para el Mn agregado. Los siguientes datos se obtuvieron de un diagrama de dispersión del artículo.

x	−1.0	−0.4	0	0.2	1.0
y	32	37	44	45	46
x	2.0	2.8	3.2	3.4	4.0
y	42	42	40	37	30

Además, $\hat{\beta}_0 = 41.7422$, $\hat{\beta}_1 = 6.581$, $\hat{\beta}_2 = -2.3621$, $s_{\hat{\beta}_0} = 0.8522$, $s_{\hat{\beta}_1} = 1.002$, $s_{\hat{\beta}_2} = 0.3073$ y SSE = 26.98.

a. ¿El modelo cuadrático es útil para describir la relación entre x y y? [*Sugerencia:* La regresión cuadrática es un caso especial de regresión múltiple con $k = 2$, $x_1 = x$ y $x_2 = x^2$.] Aplique un procedimiento apropiado.
b. ¿Es necesario eliminar el predictor cuadrático?
c. Estime la altura esperada del trigo tratado con 10 μM de Mn mediante un intervalo de confianza de 90%. [*Sugerencia:* La desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2$ es 1.031.]

76. En el artículo "Chemithermomechanical Pulp from Mixed High Density Hardwoods" (*TAPPI*, julio de 1988: 145-146) se describe un estudio en el que se obtuvieron

los siguientes datos para relacionar y = área superficial específica (cm^2/g) con x_1 = % de NaOH utilizado como tratamiento químico previo y x_2 = tiempo de tratamiento (min) para un lote de pulpa.

x_1	x_2	y
3	30	5.95
3	60	5.60
3	90	5.44
9	30	6.22
9	60	5.85
9	90	5.61
15	30	8.36
15	60	7.30
15	90	6.43

El siguiente resultado de MINITAB se obtuvo de solicitar el ajuste del modelo $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$.

```
La ecuación de regresión
ÁREA = 6.05 + 0.142 NAOH - 0.0169 TIEMPO

Predictor        Coef  Desv. est.  Relación t      p
Constante      6.0483      0.5208       11.61  0.000
NAOH          0.14167     0.03301        4.29  0.005
TIEMPO     -0.016944    0.006601       -2.57  0.043

s = 0.4851    R cuadrada = 80.7%    R cuadrada
(ajustada)  74.2%

Análisis de varianza

FUENTE        GL        SC        CM        F        p
Regresión      2    5.8854    2.9427    12.51    0.007
Error          6    1.4118    0.2353
Total          8    7.2972
```

a. ¿Qué proporción de variación observada en el área superficial específica se explica mediante la relación del modelo?

b. ¿El modelo seleccionado al parecer especifica una relación útil entre la variable dependiente y los predictores?

c. Siempre que el % de NaOH permanezca en el modelo, ¿sugeriría que se elimine el predictor *tiempo de tratamiento*?

d. Calcule un IC de 95% para el cambio esperado en el área superficial específica relacionada con un aumento de 1% en NaOH cuando el tiempo de tratamiento se mantiene fijo.

e. En el informe de MINITAB se encuentra que la desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_1(9) + \hat{\beta}_2(60)$ es 0.162. Calcule un intervalo de predicción para el valor de área superficial específica que se observa cuando el % de NaOH = 9 y el tiempo de tratamiento = 60.

77. Se realizó un análisis de regresión múltiple para relacionar y = resistencia a la tensión de un espécimen de fibra sintética con las variables x_1 = porcentaje de algodón y x_1 = tiempo de secado. El conjunto de datos consistió en $n = 12$ observaciones.

a. Los coeficientes estimados fueron $\hat{\beta}_0 = 180.00$, $\hat{\beta}_1 = 1.000$ y $\hat{\beta}_2 = 10.500$. Calcule una estimación puntual de la resistencia a la tensión esperada cuando el porcentaje de algodón = 15 y el tiempo de secado = 3.5.

b. Las sumas de cuadrados fueron SST = 1210.30 y SSE = 117.40. ¿Qué proporción de la variación observada en la resistencia a la tensión se puede atribuir a la relación del modelo?

c. Utilice la información del inciso (b) para determinar si el modelo con variables x_1 y x_2 especifica una relación útil.

d. La desviación estándar estimada de $\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$ cuando $x_1 = 18$ y $x_2 = 3.0$ fue 1.20. Calcule un IP de 95% para la resistencia a la tensión de un espécimen de tejido para el cual $x_1 = 18$ y $x_2 = 3.0$.

78. Se realizó un estudio para relacionar tiempo de falla (y) de cierto componente de una máquina con las variables voltaje de operación (x_1), velocidad del motor (x_2) y temperatura de operación (x_3). El conjunto de datos resultante consistió en $n = 20$ observaciones. Cuando el modelo con las tres variables x_1, x_2 y x_3 se ajustó a los datos, el valor de la suma de cuadrados del error fue 8212.5. Al ajustar el modelo de interacción de segundo orden (con predictores cuadráticos y los productos de pares de variables) se obtuvo una suma de cuadrados del error de 5027.1. ¿Se debe retener en el modelo al menos uno de los predictores cuadráticos o de interacción? Exprese y pruebe las hipótesis pertinentes.

79. En el artículo "A Statistical Analysis of the Notch Toughness of 9% Nickel Steels Obtained for Production Heats" (*J. of Testing and Eval.*, 1987: 355-363) se ilustran los resultados de un análisis de regresión múltiple que relaciona la tenacidad de muesca en v de Charpy y (joules) con las siguientes variables: x_1 = espesor de placa (mm), x_2 = contenido de carbono (%), x_3 = contenido de manganeso (%), x_4 = contenido de fósforo (%), x_5 = contenido de azufre (%), x_6 = contenido de silicio (%), x_7 = contenido de níquel (%), x_8 = resistencia a la cedencia (Pa) y x_9 = resistencia a la tensión (Pa).

a. En los mejores subconjuntos posibles se requiere añadir las variables en el orden x_5, x_8, x_6, x_3, x_2, x_7, x_9, x_1 y x_4. Los valores de R_k^2, MSE_k y C_k son los siguientes:

No. de Predictores	1	2	3	4
R_i^2	0.354	0.453	0.511	0.550
MSE_i	2295	1948	1742	1607
C_i	314	173	89.6	35.7

No. de Predictores	5	6	7	8	9
R_i^2	0.562	0.570	0.572	0.575	0.575
MSE_i	1566	1541	1535	1530	1532
C_i	19.9	11.0	9.4	8.2	10.0

¿Cuál modelo recomendaría? Explique la razón fundamental de su selección.

b. Los autores también consideraron modelos de segundo orden donde intervienen los predictores x_i^2 y x_ix_j. La información respecto a los mejores modelos empezando con las variables x_2, x_3, x_5, x_6, x_7 y x_8 es la siguiente (al pasar del mejor modelo de cuatro predictores al mejor modelo de cinco predictores, se eliminó x_8 y se introdujeron x_2x_6 y x_7x_8 además x_8 se volvió a introducir en una etapa posterior):

No. de Predictores	1	2	3	4	5
R_i^2	0.415	0.541	0.600	0.629	0.650
MSE_i	2079	1636	1427	1324	1251
C_i	433	109	104	52.4	16.5

No. de Predictores	6	7	8	9	10
R_i^2	0.652	0.655	0.658	0.659	0.659
MSE_i	1246	1237	1229	1229	1230
C_i	14.9	11.2	8.5	9.2	11.0

¿Cuál de estos modelos recomendaría y por qué? (*Nota:* los modelos basados en ocho de las variables originales no produjeron marcada mejoría en los modelos considerados aquí.)

80. Se seleccionó una muestra de $n = 20$ compañías y se determinaron los valores de y = precio de las acciones y $k = 15$ variables predictoras (p. ej., dividendos trimestrales, ganancias del año anterior y relación de deuda). Cuando se ajustó a los datos el modelo de regresión múltiple mediante estos 15 predictores, se obtuvo $R^2 = 0.90$.

a. ¿El modelo al parecer especifica una relación útil entre y y las variables predictoras? Realice una prueba con un nivel de significación 0.05. [*Sugerencia:* El valor crítico F con 15 grados de libertad en el numerador y 4 en el denominador es 5.86.]

b. Con base en el resultado del inciso (a), ¿un valor de R^2 alto indica que un modelo es útil? ¿Bajo qué circunstancias se podría sospechar de un modelo con un valor de R^2 alto?

c. Según los valores de n y k, ¿qué tan grande debería ser R^2 para considerar útil al modelo al nivel de significación 0.05?

81. ¿La exposición a la contaminación atmosférica produce una menor esperanza de vida? Esta pregunta se examinó en el artículo "Does Air Pollution Shorten Lives?" (*Statistics and Public Policy*, Reading, MA, Addison Wesley, 1977). Se realizó un registro de datos acerca de

y = índice de mortalidad total (decesos por 10 000)

x_1 = lectura promedio de partículas suspendidas ($\mu g/m^3$)

x_2 = lectura mínima de sulfatos ([$\mu g/m^3$] × 10)

x_3 = densidad de población (habitantes/milla^2)

x_4 = (porcentaje de personas no blancas) × 10

x_5 = (porcentaje de personas con más de 65 años de edad) × 10

en 1960 para $n = 117$ áreas estadísticas metropolitanas estándar seleccionadas al azar. La ecuación de regresión estimada fue

$$y = 19.607 + 0.041x_1 + 0.071x_2 + 0.001x_3 + 0.041x_4 + 0.687x_5$$

a. Para este modelo $R^2 = 0.827$. Con un nivel de significación de 0.05, realice una prueba de utilidad de modelo.

b. La desviación estándar estimada de $\hat{\beta}_1$ fue 0.016. Calcule e interprete un IC de 90% para β_1.

c. Dado que la desviación estándar estimada de $\hat{\beta}_4$ es 0.007, determine si el porcentaje de personas no blancas es una variable importante en el modelo. Utilice un nivel de significación de 0.01.

d. En 1960 los valores de x_1, x_2, x_3, x_4 y x_5 para Pittsburgh fueron 166, 60, 788, 68 y 95, respectivamente. Utilice la ecuación de regresión anterior para pronosticar la tasa de mortalidad de Pittsburgh. ¿Cómo se compara esta predicción con el valor real de 1960 de 103 fallecimientos por 10 000?

82. Dado que $R^2 = 0.723$ para el modelo con los predictores x_1, x_4, x_5 y x_8 y $R^2 = 0.689$ para el modelo con los predictores x_1, x_3, x_5 y x_6, ¿qué se puede decir acerca de R^2 para el modelo con los predictores

a. x_1, x_3, x_4, x_5, x_6 y x_8? Explique su respuesta.

b. x_1 y x_4? Explique.

Bibliografía

Chatterjee, S. y Bertram Price, *Regression Analysis by Example* (2a. ed.), Wiley, Nueva York, 1990. Un análisis breve pero informativo de temas seleccionados, en particular la multicolinealidad y el uso de métodos de estimación sesgados.

Daniel, Cuthbert y Fred Wood, *Fitting Equations to Data* (2a. ed.), Wiley, Nueva York, 1980. Contiene muchas ideas y métodos que surgieron de la amplia experiencia de los autores como consultores.

Draper, Norman y Harry Smith, *Applied Regression Analysis* (3a. ed.), Wiley, Nueva York, 1999. Véase la bibliografía del capítulo 12.

Hoaglin, David y Roy Welsch. "The Hat Matrix in Regression and ANOVA", *American Statistician,* 1978: 17-23. Describe métodos para detectar observaciones preponderantes en un conjunto de datos de regresión.

Hocking, Ron, "The Analysis and Selection of Variables in Linear Regression", *Biometrics*, 1976: 1-49. Un estudio excelente de este tema.

Neter, John, Michael Kutner, Christopher Nachtsheim y William Wasserman, *Applied Linear Statistical Models* (4a. ed.), Irwin, Homewood, IL, 1996. Véase la bibliografía del capítulo 12.

14

Pruebas de bondad de ajuste y análisis de datos categóricos

Introducción

En el tipo de situación más simple considerada en este capítulo, cada observación de una muestra se clasifica como perteneciente a una de una cantidad finita de categorías (p. ej., el tipo de sangre podría ser una de las cuatro categorías O, A, B o AB). Con p_i que denota la probabilidad de que alguna observación pertenezca a la categoría i (o la proporción de la población que pertenece a la categoría i), se desea probar una hipótesis nula que especifica por completo los valores de las p_i (como p. ej., H_0: $p_1 = 0.45$, $p_2 = 0.35$, $p_3 = 0.15$, $p_4 = 0.05$, cuando hay cuatro categorías). El estadístico de prueba es una medida de la discrepancia entre los números observados en las categorías y los números esperados cuando H_0 es verdadera. Debido a que se llega a una decisión al comparar el valor calculado del estadístico de prueba con un valor crítico de la distribución chi-cuadrada, el procedimiento se llama prueba de bondad de ajuste de chi-cuadrada.

En ocasiones, la hipótesis nula especifica que las p_i dependen de algún número más pequeño de parámetros, sin especificar los valores de estos parámetros. Por ejemplo, con tres categorías la hipótesis nula podría expresar que $p_1 = \theta^2$, $p_2 = 2\theta(1 - \theta)^2$ y $p_3 = (1 - \theta)^2$. Para que se lleve a cabo una prueba chi-cuadrada, los valores de algunos parámetros no especificados se deben estimar a partir de los datos. Estos problemas se analizan en la sección 14.2. Después, los métodos se aplican para probar una hipótesis nula que establece que la muestra proviene de

una familia particular de distribuciones, como la de Poisson (con λ estimada de la muestra) o la familia normal (con μ y σ estimadas).

Las pruebas de chi-cuadrada para dos situaciones distintas se presentan en la sección 14.3. En la primera, la hipótesis nula expresa que las p_i son las mismas para varias poblaciones diferentes. El segundo tipo de situación requiere tomar una muestra de una sola población y clasificar cada individuo con respecto a dos factores categóricos diferentes (como p. ej., la preferencia religiosa y el registro de partido político). La hipótesis nula en esta situación es que los dos factores son independientes dentro de la población.

14.1 Pruebas de bondad de ajuste cuando se especifican por completo las probabilidades de categoría

Un experimento binomial consiste en una secuencia de ensayos independientes en los que cada ensayo produce uno de dos resultados posibles, S (para éxito) y F (para fracaso). Se supone que la probabilidad de éxito, representada por p, es constante de un ensayo a otro, y el número n de ensayos es fijo al inicio del experimento. En el capítulo 8, se presentó una prueba z con muestras grandes para probar H_0: $p = p_0$. Observe que esta hipótesis nula especifica tanto $P(S)$ como $P(F)$, puesto que si $P(S) = p_0$, entonces $P(F) = 1 - p_0$. Si $P(F)$ se denota con q y $1 - p_0$ con q_0, la hipótesis nula se puede escribir de otro modo como H_0: $p = p_0$, $q = q_0$. La prueba z es bilateral cuando la alternativa de interés es $p \neq p_0$.

Un **experimento multinomial** generaliza un experimento binomial permitiendo que cada ensayo dé lugar a uno de k resultados posibles, donde $k > 2$. Por ejemplo, suponga que un almacén acepta tres tipos de tarjeta de crédito. Un experimento multinomial resultaría de observar el tipo de tarjeta de crédito, tipo 1, tipo 2 o tipo 3, que utiliza cada uno de los n clientes siguientes que paga con tarjeta de crédito. En general, se hace referencia a los k resultados posibles en un determinado ensayo como categorías, y p_i representa la probabilidad de que un ensayo produzca la categoría i. Si el experimento consiste en seleccionar n individuos u objetos de una población y categorizar cada uno, entonces p_i es la probabilidad de la población que cae en la i-ésima categoría (esta clase de experimento es más o menos multinomial, siempre y cuando n sea mucho más pequeña que el tamaño de la población).

La hipótesis nula de interés especifica el valor de cada p_i. Por ejemplo, en el caso $k = 3$, se podría tener H_0: $p_1 = 0.5$, $p_2 = 0.3$, $p_3 = 0.2$. La hipótesis alternativa expresa que H_0 no es verdad, es decir que menos de los p_i tiene un valor diferente del afirmado en cuyo caso por lo menos dos deben ser diferentes, puesto que su suma es igual a 1). El símbolo p_{i0} representa el valor de p_i que se afirma en la hipótesis nula. En el ejemplo antes mencionado, $p_{10} = 0.5$, $p_{20} = 0.3$ y $p_{30} = 0.2$.

Antes de llevar a cabo el experimento multinomial, el número de ensayos que da lugar a la categoría i ($i = 1, 2, \ldots,$ o k) es una variable aleatoria, al igual que el número de éxitos

y el número de fracasos en un experimento binomial son variables aleatorias. Esta variable aleatoria se denota con N_i y su valor observado con n_i. Puesto que cada ensayo produce exactamente una de las k categorías, $\sum N_i = n$, y lo mismo resulta cierto para las n_i. Como ejemplo, un experimento con $n = 100$ y $k = 3$ podría producir $N_1 = 46$, $N_2 = 35$ y $N_3 = 19$.

El número esperado de éxitos y el número esperado de fracasos en un experimento binomial son np y nq, respectivamente. Cuando H_0: $p = p_0$, $q = q_0$ es verdad, los números previstos de aciertos y errores son np_0 y nq_0, respectivamente. De manera similar, en un experimento multinomial el número esperado de ensayos que dan lugar a la categoría i es $E(N_i) = np_i$ ($i = 1, \ldots, k$). Cuando H_0: $p_1 = p_{10}, \ldots, p_k = p_{k0}$ es verdadera, estos valores esperados se convierten en $E(N_1) = np_{10}$, $E(N_2) = np_{20}, \ldots, E(N_k) = np_{k0}$. Para el caso $k = 3$, H_0: $p_1 = 0.5$, $p_2 = 0.3$, $p_3 = 0.2$ y $n = 100$, $E(N_1) = 100(0.5) = 50$, $E(N_2) = 30$ y $E(N_3) = 20$ cuando H_0 es verdadera. Es común presentar las n_i en un formato tabular que consiste en un renglón de k celdas, una para cada categoría, como se ilustra en la tabla 14.1. Los valores esperados cuando H_0 es verdadera se muestran debajo de los valores observados. Las N_i y las n_i por lo común se denominan *cuentas de celda observadas* (o *frecuencias de celda observadas*), y $np_{10}, np_{20}, \ldots, np_{k0}$ son las *cuentas de celda esperadas* correspondientes bajo H_0.

Tabla 14.1 Cuentas de celda observadas y esperadas

Categoría	$i = 1$	$i = 2$	...	$i = k$	Total de renglón
Observada	n_1	n_2	...	n_k	n
Esperada	np_{10}	np_{20}	...	np_{k0}	n

Las n_i deben estar razonablemente cerca de las np_{i0} correspondientes cuando H_0 es verdadera. Por otro lado, varias de las cuentas observadas deben diferir sustancialmente de estas cuentas esperadas cuando los valores reales de las p_i difieren en forma notable de lo que se afirma en la hipótesis nula. El procedimiento de prueba requiere evaluar la discrepancia entre las n_i y las np_{i0}, con H_0 que se rechaza cuando la discrepancia es suficientemente grande. Es natural basar una medida de discrepancia en las desviaciones cuadradas $(n_1 - np_{10})^2$, $(n_2 - np_{20})^2, \ldots, (n_k - np_{k0})^2$. Una forma obvia de combinar éstas en una medida global es sumarlas para obtener $\sum(n_i - np_{i0})^2$. Sin embargo, suponga que $np_{10} = 100$ y $np_{20} = 10$. Entonces si $n_1 = 95$ y $n_2 = 5$, las dos categorías aportan las mismas desviaciones cuadradas a la medida propuesta. Sin embargo n_1 es sólo 5% menor de lo que se esperaría cuando H_0 es verdadera, mientras que n_2 es 50% menor. Para tomar en cuenta las magnitudes relativas de las desviaciones, se divide cada desviación cuadrada entre la cuenta esperada correspondiente y luego se combina.

Antes de dar una descripción más detallada, se debe analizar un tipo de distribución de probabilidad llamada *distribución chi-cuadrada*. Esta distribución se presentó en la sección 4.4 y se utilizó en el capítulo 7, a fin de obtener un intervalo de confianza para la varianza σ^2 de una población normal. La distribución chi-cuadrada tiene un solo parámetro ν, denominado número de grados de libertad (gl) de la distribución, con valores posibles 1, 2, 3, ... Análogo al valor crítico $t_{\alpha,\nu}$ para la distribución t, $\chi^2_{\alpha,\nu}$ es el valor tal que α del área bajo la curva χ^2

con ν gl se ubica a la derecha de $\chi^2_{\alpha,\nu}$ (véase la figura 14.1). Los valores seleccionados de $\chi^2_{\alpha,\nu}$ se dan en la tabla A.7 del apéndice.

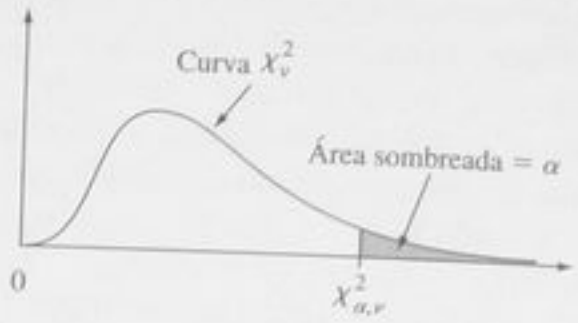


Figura 14.1 Valor crítico para una distribución chi-cuadrada

TEOREMA

Siempre y cuando $np_i \geq 5$ para cada i ($i = 1, 2, \ldots, k$), la variable

$$\chi^2 = \sum_{i=1}^{k} \frac{(N_i - np_i)^2}{np_i} = \sum_{\text{todas las celdas}} \frac{(\text{observada} - \text{esperada})^2}{\text{esperada}}$$

tiene una distribución chi-cuadrada aproximada con $k - 1$ gl.

El hecho de que gl $= k - 1$ es una consecuencia de la restricción $\sum N_i = n$. Aunque hay k cuentas de celda observadas, una vez que se conocen $k - 1$ el resto se determina unívocamente. Es decir, hay sólo $k - 1$ cuentas de celda "determinadas libremente" y, por lo tanto, $k - 1$ gl.

Si np_{i0} se sustituye por np_i en χ^2, el estadístico de prueba resultante tiene una distribución chi-cuadrada cuando H_0 es verdadera. El rechazo de H_0 es apropiado $\chi^2 \geq c$ (debido a que las discrepancias grandes entre las cuentas observadas y esperadas dan lugar a un valor grande de χ^2), y la elección de $c = \chi^2_{\alpha,k-1}$ produce una prueba con nivel de significación α.

Hipótesis nula: H_0: $p_1 = p_{10}, p_2 = p_{20}, \ldots, p_k = p_{k0}$

Hipótesis alternativa: H_a: al menos una p_i no es igual a p_{i0}

Valor del estadístico de prueba: $\chi^2 = \sum_{\text{todas las celdas}} \frac{(\text{observada} - \text{esperada})^2}{\text{esperada}} = \sum_{i=1}^{k} \frac{(n_i - np_{i0})^2}{np_{i0}}$

Región de rechazo: $\chi^2 \geq \chi^2_{\alpha,k-1}$

Ejemplo 14.1 Si se atiende a las diferentes características de un organismo, cada una controlada por un solo gen, y se cruza una cepa pura que tiene el genotipo AABB con una cepa pura que tiene el genotipo aabb (las letras mayúsculas denotan alelos dominantes y las minúsculas alelos recesivos), el genotipo resultante es AaBb. Si estos organismos de primera generación se cruzan después entre ellos mismos (una cruza dihíbrida), habrá cuatro genotipos dependiendo si está presente un alelo dominante de cualquier tipo. Las leyes de Mendel de la herencia implican que estos cuatro fenotipos deben tener probabilidades $\frac{9}{16}$, $\frac{3}{16}$, $\frac{3}{16}$ y $\frac{1}{16}$ de surgir en alguna cruza dihíbrida determinada.

En el artículo "Linkage Studies of the Tomato" (*Trans. Royal Canadian Institute,* 1931: 1-19) se muestran los siguientes datos de fenotipos de una cruza dihíbrida de tomates altos de hoja cortada con tomates enanos de hoja de papa. Hay $k = 4$ categorías que corresponden a los cuatro posibles fenotipos, con la hipótesis nula

$$H_0\colon\ p_1 = \frac{9}{16}, p_2 = \frac{3}{16}, p_3 = \frac{3}{16}, p_4 = \frac{1}{16}$$

Las cuentas de celda esperadas son $9n/16$, $3n/16$, $3n/16$ y $n/16$, y la prueba se basa en $k - 1 = 3$ gl. El tamaño de muestra total fue $n = 1611$. Las cuentas observada y esperada se dan en la tabla 14.2.

Tabla 14.2 Cuentas de celda observadas y esperadas para el ejemplo 14.1

	$i = 1$ **Alto, hoja cortada**	$i = 2$ **Alto, hoja de papa**	$i = 3$ **Enano, hoja cortada**	$i = 4$ **Enano, hoja de papa**
n_i	926	288	293	104
np_{i0}	906.2	302.1	302.1	100.7

La contribución a χ^2 de la primera celda es

$$\frac{(n_1 - np_{10})^2}{np_{10}} = \frac{(926 - 906.2)^2}{906.2} = 0.433$$

Las celdas 2, 3 y 4 contribuyen con 0.658, 0.274 y 0.108, respectivamente, de modo que $\chi^2 = 0.433 + 0.658 + 0.274 + 0.108 = 1.473$. Una prueba con nivel de significación 0.10 requiere $\chi^2_{0.10,3}$, el número en el renglón de 3 gl y columna 0.10 de la tabla A.7 del apéndice. Este valor crítico es 6.251. Puesto que 1.473 no es al menos 6.251, no se puede rechazar H_0 incluso a este nivel bastante grande de significación. Los datos son bastante consistentes con las leyes de Mendel. ■

Aunque la prueba chi-cuadrada se desarrolló para situaciones en las que $k > 2$, también se puede usar cuando $k = 2$. En este caso la hipótesis nula se prueba como $H_0\colon p_1 = p_{10}$, puesto que las relaciones $p_2 = 1 - p_1$ y $p_{20} = 1 - p_{10}$ hacen redundante la inclusión de $p_2 = p_{20}$ en H_0. La hipótesis alternativa es $H_a\colon p_1 \neq p_{10}$. Estas hipótesis también se pueden probar por medio de una prueba z de dos extremos con estadístico de prueba

$$Z = \frac{(N_1/n) - p_{10}}{\sqrt{\dfrac{p_{10}(1 - p_{10})}{n}}} = \frac{\hat{p}_1 - p_{10}}{\sqrt{\dfrac{p_{10}p_{20}}{n}}}$$

De modo sorprendente, los dos procedimientos de prueba son equivalentes por completo. Esto es porque se puede demostrar que $Z^2 = \chi^2$ y $(z_{\alpha/2})^2 = \chi^2_{1,\alpha}$, de modo que $\chi^2 \geq \chi^2_{1,\alpha}$ si y sólo si $|Z| \geq z_{\alpha/2}$.* Si la hipótesis alternativa es $H_a\colon p_1 > p_{10}$ o $H_a\colon p_1 < p_{10}$, no se puede usar la prueba chi-cuadrada. Entonces se debe volver a una prueba z de extremo superior o inferior.

*El hecho de que $(z_{\alpha/2})^2 = \chi^2_{1,\alpha}$ es una consecuencia de la relación entre la distribución normal estándar y la distribución chi-cuadrada con 1 gl: si $Z \sim N(0, 1)$, entonces Z^2 tiene una distribución chi-cuadrada con $\nu = 1$.

Como sucede con los procedimientos de prueba, se debe tener cuidado de no confundir la importancia estadística con la importancia práctica. Una χ^2 calculada que excede $\chi^2_{\alpha,k-1}$ podría ser un resultado de un tamaño de muestra muy grande y no de las diferencias prácticas entre las p_{i0} supuestas y las p_i verdaderas. Así, si $p_{10} = p_{20} = p_{30} = \frac{1}{3}$, pero los valores de las p_i verdaderas son 0.330, 0.340 y 0.330, un valor grande de χ^2 surge de seguro con una n suficientemente grande. Antes de rechazar H_0, se deben examinar las $\hat{p}_i$ para ver si indican un modelo diferente del de H_0 desde un punto de vista práctico.

Valores *P* para pruebas chi-cuadrada

Las pruebas chi-cuadrada de este capítulo son de extremo superior, así que se centra la atención en este caso. Del mismo modo que el valor P para una prueba t de extremo superior es el área bajo la curva t_ν a la derecha de la t calculada, el valor P de una prueba chi-cuadrada de extremo superior es el área bajo la curva χ^2_ν a la derecha de la χ^2 calculada. En la tabla A.7 del apéndice se proporciona información limitada del valor P, porque sólo se tabulan cinco valores críticos de extremo superior para cada ν distinta. Por lo tanto, se incluyó otra tabla en el apéndice, análoga a la tabla A.8, que facilita hacer declaraciones más precisas del valor P.

El hecho de que las curvas t se centraran en cero, permitía tabular las áreas de extremo de la curva t en una forma relativamente compacta, donde en el margen izquierdo aparecen los valores que van de 0.0 a 4.0 en la escala t horizontal y varias columnas muestran las áreas de extremo superior correspondientes para varios gl. El movimiento hacia la derecha de las curvas chi-cuadrada a medida que aumentan los grados de libertad, requiere un tipo de tabulación un poco diferente. En el margen izquierdo de la tabla A.11 del apéndice, se muestran varias áreas de extremo superior: 0.100, 0.095, 0.090, . . . , 0.005 y 0.001. Cada columna de la tabla es para un valor diferente de grados de libertad, y las entradas son valores del eje horizontal chi-cuadrada que capta estas áreas de extremo correspondientes. Por ejemplo, al ir en dirección vertical hasta el área de extremo 0.085 y luego en sentido horizontal hasta la columna de 2 gl, se ve que el área a la derecha de 4.93 bajo la curva chi-cuadrada de 2 gl es 0.085 (véase la figura 14.2).

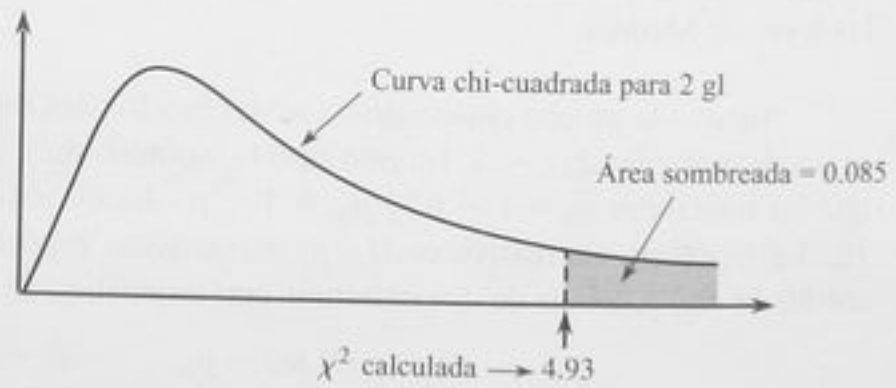


Figura 14.2 Valor P para una prueba chi-cuadrada de extremo superior

Para captar esta misma área de extremo superior bajo la curva de 10 gl, se debe ir hasta 16.54. En la columna de dos grados de libertad, en el renglón superior se observa que si el valor calculado de la variable chi-cuadrada es menor que 4.60, el área de extremo captada (valor P) pasa de 0.10. De manera similar, el renglón inferior de esta columna indica que si el valor calculado pasa de 13.81, el área de extremo es menor que 0.001 (valor $P < 0.001$).

χ^2 cuando las p_i son funciones de otros parámetros

Es común suponer que las p_i dependen de una cantidad menor de parámetros $\theta_1, \ldots, \theta_m$ ($m < k$). Entonces una hipótesis específica en relación con las θ_i produce p_{i0} específicas, las cuales se emplean en la prueba χ^2.

Ejemplo 14.2 En un famoso artículo de genética ("The Progeny in Generations F_{12} to F_{17} of a Cross Between a Yellow-Wrinkled and a Green-Round Seeded Pea", *J. Genetics*, 1923: 255-331), el primer estadístico G. U. Yule analizó los datos que resultan de cruzar chícharos. Los alelos dominantes del experimento fueron Y = color amarillo y R = forma redonda, que dan lugar al doble dominante YR. Yule examinó 269 vainas de cuatro semillas obtenidas de una cruza híbrida y contó el número de semillas YR en cada vaina. Si X representa la cantidad de YR en una vaina elegida al azar, los valores posibles de X son 0, 1, 2, 3, 4, que se identifican con las celdas 1, 2, 3, 4 y 5 de una tabla rectangular (así que, p. ej., una vaina con $X = 4$ produce una cuenta observada en la celda 5).

La hipótesis de que las leyes de Mendelian son funcionales y que los fenotipos de cada una de las semillas dentro de una vaina son independientes entre sí, significa que X tiene una distribución binomial con $n = 4$ y $\theta = \frac{9}{16}$. Por consiguiente, se desea probar H_0: $p_1 = p_{10}, \ldots, p_5 = p_{50}$, donde

$$p_{i0} = P(i - 1 \text{ YR entre 4 semillas cuando } H_0 \text{ es verdadera})$$
$$= \binom{4}{i-1} \theta^{i-1}(1-\theta)^{4-(i-1)} \qquad i = 1, 2, 3, 4, 5;\ \theta = \frac{9}{16}$$

Los datos de Yule y los cálculos se ilustran en la tabla 14.3 con cuentas de celda esperadas $np_{i0} = 269p_{i0}$.

Tabla 14.3 Cuentas de celda observadas y esperadas para el ejemplo 14.2

Celda *i*	**1**	**2**	**3**	**4**	**5**
Chícharos YR/vainas	**0**	**1**	**2**	**3**	**4**
Observada	16	45	100	82	26
Esperada	9.86	50.68	97.75	83.78	26.93
$\frac{(\text{observada} - \text{esperada})^2}{\text{esperada}}$	3.823	0.637	0.052	0.038	0.032

Así, $\chi^2 = 3.823 + \cdots + 0.032 = 4.582$. Puesto que $\chi^2_{0.01,k-1} = \chi^2_{0.01,4} = 13.277$, H_0 no se rechaza al nivel 0.01. En la tabla A.11 del apéndice se muestra que como $4.582 < 7.77$, el valor P para la prueba pasa de 0.10. No se debe rechazar H_0 a ningún nivel de significación razonable. ■

χ^2 cuando la distribución subyacente es continua

Se ha supuesto hasta el momento que las k categorías se definen de manera natural en el contexto del experimento en consideración. La prueba χ^2 se puede usar para probar si una mues-

tra proviene de una distribución continua subyacente específica. Sea X la variable que está siendo muestreada y suponga que la fdp de X supuesta es $f_0(x)$. Al igual que en la construcción de una distribución de frecuencia en el capítulo 1, subdivida la escala de medición de X en k intervalos $[a_0, a_1), [a_1, a_2), \ldots, [a_{k-1}, a_k)$, donde el intervalo $[a_{i-1}, a_i)$ incluye el valor a_{i-1} pero no al valor a_i. Las probabilidades de celda que especifica H_0 son

$$p_{i0} = P(a_{i-1} \le X < a_i) = \int_{a_{i-1}}^{a_i} f_0(x)\,dx$$

Las celdas se deben elegir de modo que $np_{i0} \ge 5$ para $i = 1, \ldots, k$. A menudo se seleccionan de modo que las np_{i0} sean iguales.

Ejemplo 14.3 Para ver si el tiempo de inicio de parto entre madres gestantes tiene una distribución uniforme las 24 horas de un día, se divide el día en k periodos, cada uno de duración $24/k$. La hipótesis nula establece que $f(x)$ es la fdp uniforme en el intervalo $[0, 24]$, de modo que $p_{i0} = 1/k$. En el artículo "The Hour of Birth" (*British J. Preventive and Social Medicine*, 1953: 43-59) se habla de 1186 tiempos de inicio, que se clasificaron en $k = 24$ intervalos de una hora comenzando a medianoche, con cuentas de celda de 52, 73, 89, 88, 68, 47, 58, 47, 48, 53, 47, 34, 21, 31, 40, 24, 37, 31, 47, 34, 36, 44, 78 y 59. Cada cuenta de celda esperada es $1186 \cdot \frac{1}{24} = 49.42$, y el valor de χ^2 es 162.77. Puesto que $\chi^2_{0.01,23} = 41.637$, el valor calculado es muy significativo, y la hipótesis nula se rechaza de forma rotunda. En términos generales, al parecer el alumbramiento tiene muchas más probabilidades de comenzar muy tarde en la noche que durante las horas de vigilia normales. ■

Para probar si una muestra proviene de una distribución normal específica, los parámetros fundamentales son $\theta_1 = \mu$ y $\theta_2 = \sigma$, y cada p_{i0} será una función de estos parámetros.

Ejemplo 14.4 En cierta universidad, se supone que los exámenes finales duran dos horas. El Departamento de Psicología elaboró un examen final departamental para un curso elemental que se cree satisface los siguientes criterios: 1) el tiempo real que toma completar el examen, tiene una distribución normal, 2) $\mu = 100$ min y 3) exactamente 90% de los estudiantes terminan dentro del periodo de dos horas. Para ver si en realidad esto es así, se elige al azar a 120 estudiantes y se registran sus tiempos para completar el examen. Se decidió que se deben usar $k = 8$ intervalos. Los criterios implican que el percentil 90 de la distribución del tiempo para completar el examen es $\mu + 1.28\sigma = 120$. Puesto que $\mu = 100$, esto significa que $\sigma = 15.63$.

Los ocho intervalos que dividen la escala normal estándar en ocho segmentos equiprobables son $[0, 0.32), [0.32, 0.675), [0.675, 1.15), [1.15, \infty)$, y sus cuatro contrapartes en el otro lado de 0. Para $\mu = 100$ y $\sigma = 15.63$, estos intervalos se convierten en $[100, 105), [105, 110.55), [110.55, 117.97)$ y $[117.97, \infty)$. Así $p_{i0} = \frac{1}{8} = 0.125$ $(i = 1, \ldots, 8)$, de modo que cada cuenta de celda esperada es $np_{i0} = 120(0.125) = 15$. Las cuentas de celda observadas fueron 21, 17, 12, 16, 10, 15, 19 y 10, de donde resulta que χ^2 es 7.73. Como $\chi^2_{0.10,7} = 12.017$ y 7.73 no es ≥ 12.017, no hay evidencia para concluir que no se cumplieron los criterios. ■

Ejercicios | Sección 14.1 (1 a 11)

1. ¿Qué conclusión sería apropiada para una prueba chi-cuadrada de extremo superior en cada una de las siguientes situaciones?
   a. $\alpha = 0.05$, gl $= 4$, $\chi^2 = 12.25$
   b. $\alpha = 0.01$, gl $= 3$, $\chi^2 = 8.54$
   c. $\alpha = 0.10$, gl $= 2$, $\chi^2 = 4.36$
   d. $\alpha = 0.01$, $k = 6$, $\chi^2 = 10.20$

2. Exprese cuanto pueda acerca del valor P para una prueba chi-cuadrada de extremo superior en cada una de las siguientes situaciones:
   a. $\chi^2 = 7.5$, gl $= 2$
   b. $\chi^2 = 13.0$, gl $= 6$
   c. $\chi^2 = 18.0$, gl $= 9$
   d. $\chi^2 = 21.3$, $k = 5$
   e. $\chi^2 = 5.0$, $k = 4$

3. Un departamento de estadística en una gran universidad mantiene un servicio de tutoría para estudiantes de sus cursos introductorios. Se dotó de personal al servicio con la esperanza de que 40% de sus clientes serían del curso de estadística empresarial, 30% de estadística en ingeniería, 20% del curso de estadística para estudiantes de ciencias sociales y el otro 10% del curso para estudiantes de agricultura. Una muestra aleatoria de $n = 120$ clientes reveló que hay 52, 38, 21 y 9 alumnos de los cuatro cursos, respectivamente. ¿Estos datos indican que los porcentajes en los que se basó la asignación de personal no son correctos? Formule y pruebe las hipótesis relevantes con $\alpha = 0.05$.

4. Se supone que cuando de alguna manera se desorientan las palomas, no muestran preferencia por ninguna dirección de vuelo después de alzar el vuelo (así que la dirección X debe tener una distribución uniforme en el intervalo de 0° a 360°). Para probar esto, se desorientó a 120 palomas, se soltaron y se registró la dirección de vuelo de cada una; a continuación se muestran los datos obtenidos. Utilice la prueba chi-cuadrada al nivel 0.10 para ver si los datos respaldan la hipótesis.

Dirección	0–<45°	45–<90°	90–<135°
Frecuencia	12	16	17
Dirección	135–<180°	180–<225°	225–<270°
Frecuencia	15	13	20
Dirección	270–<315°	315–<360°	
Frecuencia	17	10	

5. Un sistema de recuperación de información tiene 10 lugares de almacenamiento. La información se guarda con la esperanza de que la proporción de largo plazo de solicitudes para el lugar se determine mediante $p_i = (5.5 - \Omega i - 5.5\Omega)/30$. Una muestra de 200 solicitudes de recuperación aportó las siguientes frecuencias para los lugares 1 a 10, respectivamente: 4, 15, 23, 25, 38, 31, 32, 14, 10 y 8. Use una prueba chi-cuadrada al nivel de significación 0.10 para decidir si los datos son congruentes con las proporciones *a priori* (utilice el método del valor P).

6. El sorgo es un cereal importante cuya calidad y apariencia podría verse afectada por la presencia de pigmentos en el pericarpio (las paredes del ovario de la planta). En el artículo "A Genetic and Biochemical Study on Pericarp Pigments in a Cross Between Two Cultivars of Grain Sorghum, Sorghum Bicolor" (*Heredity*, 1976: 413-416) se describe un experimento en el que se hizo una cruza inicial entre sorgo CK60 (una variedad norteamericana con semillas blancas) y Abu Taima (una variedad etíope con semillas amarillas) para producir plantas con semillas rojas y después una autocruza de las plantas de semillas rojas. Según la teoría genética, esta cruza F_2 debe producir plantas con semillas rojas, amarillas o blancas en la proporción 9 : 3 : 4. A continuación se dan los datos del experimento. ¿Los datos confirman o contradicen la teoría genética? Pruebe al nivel 0.05 con el método del valor P.

Color de la semilla	Rojo	Amarilla	Blanca
Frecuencia observada	195	73	100

7. Durante mucho tiempo los criminalistas han debatido sobre si hay una relación entre las condiciones del clima y la incidencia de delitos violentos. El autor del artículo "Is There a Season for Homicide?" (*Criminology*, 1988: 287-296) clasificó 1361 homicidios según la estación del año, con los siguientes resultados. Pruebe la hipótesis nula de iguales proporciones con $\alpha = 0.01$ por medio de la tabla chi-cuadrada para decir lo más posible acerca del valor P.

Invierno	Primavera	Verano	Otoño
328	334	372	327

8. El artículo "Psychiatric and Alcoholic Admissions Do Not Occur Disproportionately Close to Patients' Birthdays" (*Psychological Reports*, 1992: 944-946) se centra en la existencia de alguna relación entre la fecha de admisión de un paciente para tratamiento de alcoholismo y su cumpleaños. Si se supone un año de 365 días (es decir, excluyendo al año bisiesto), en ausencia de alguna relación, la fecha de admisión de un paciente es equiprobable para alguno de los 365 días. Los investigadores

establecieron cuatro categorías de admisión: 1) dentro de los 7 días del cumpleaños, 2) entre 8 y 30 días, incluso, del cumpleaños, 3) entre 31 y 90 días, incluso, del cumpleaños y 4) más de 90 días del cumpleaños. Una muestra de 200 pacientes produjo frecuencias observadas de 11, 24, 69 y 96 para las categorías 1, 2, 3 y 4, respectivamente. Exprese y pruebe las hipótesis pertinentes con un nivel de significación de 0.01.

9. Se supone que el tiempo de respuesta de un sistema informático para solicitar cierto tipo de información tiene una distribución exponencial con parámetro $\lambda = 1$ s (de modo que si X = tiempo de respuesta, la fdp de X bajo H_0 es $f_0(x) = e^{-x}$ para $x \geq 0$).

a. Si se hubiera observado $X_1, X_2, \ldots, X_n$ y deseara usar la prueba chi-cuadrada con cinco intervalos de clase equiprobables bajo H_0, ¿cuáles serían los intervalos de clase resultantes?

b. Realice la prueba chi-cuadrada con los siguientes datos obtenidos de una muestra aleatoria de 40 tiempos de respuesta:

0.10	0.99	1.14	1.26	3.24	0.12	0.26	0.80
0.79	1.16	1.76	0.41	0.59	0.27	2.22	0.66
0.71	2.21	0.68	0.43	0.11	0.46	0.69	0.38
0.91	0.55	0.81	2.51	2.77	0.16	1.11	0.02
2.13	0.19	1.21	1.13	2.93	2.14	0.34	0.44

10. a. Demuestre que otra expresión para el estadístico chi-cuadrada es

$$\chi^2 = \sum_{i=1}^{k} \frac{N_i^2}{np_{i0}} - n$$

¿Por qué es más eficaz calcular χ^2 con esta fórmula?

b. Cuando la hipótesis nula es H_0: $p_1 = p_2 = \cdots = p_k = 1/k$ (es decir, $p_{i0} = 1/k$ para toda i), ¿cómo se simplifica la fórmula del inciso (a)? Utilice la expresión simplificada para calcular χ^2 para los datos de paloma/dirección del ejercicio 4.

11. a. Después de obtener una muestra aleatoria de una población, se desea utilizar una prueba chi-cuadrada para determinar si la distribución poblacional es normal estándar. Si la prueba se basa en seis intervalos de clase equiprobables bajo H_0, ¿cuáles deben ser los intervalos de clase?

b. Si se desea usar una prueba chi-cuadrada para probar H_0: la distribución poblacional es normal con $\mu = 0.5$, $\sigma = 0.002$ y la prueba se basará en seis intervalos de clase igualmente probables (bajo H_0) ¿cuáles deben ser estos intervalos?

c. Utilice la prueba chi-cuadrada con los intervalos del inciso (b) para determinar, con base en los siguientes 45 diámetros de tornillo, si el diámetro es una variable distribuida normalmente con $\mu = 0.5$ pulg, $\sigma = 0.002$ pulg.

0.4974	0.4976	0.4991	0.5014	0.5008	0.4993
0.4994	0.5010	0.4997	0.4993	0.5013	0.5000
0.5017	0.4984	0.4967	0.5028	0.4975	0.5013
0.4972	0.5047	0.5069	0.4977	0.4961	0.4987
0.4990	0.4974	0.5008	0.5000	0.4967	0.4977
0.4992	0.5007	0.4975	0.4998	0.5000	0.5008
0.5021	0.4959	0.5015	0.5012	0.5056	0.4991
0.5006	0.4987	0.4968			

14.2 Pruebas de bondad del ajuste para hipótesis compuestas

En la sección anterior se presentó una prueba de bondad del ajuste con base en el estadístico χ^2 para decidir entre H_0: $p_1 = p_{10}, \ldots, p_k = p_{k0}$ y la alternativa H_a que expresa que H_0 no es verdadera. La hipótesis nula fue una **hipótesis simple** en el sentido de que cada p_{i0} fue un número especificado, de modo que las cuentas de celda esperadas cuando H_0 fue verdadera resultaron ser números determinados de manera unívoca.

En muchas situaciones, hay k categorías que ocurren de manera natural, pero H_0 sólo establece que las p_i son funciones de otros parámetros $\theta_1, \ldots, \theta_m$ sin especificar los valores de estas θ. Por ejemplo, una población puede estar en equilibro con respecto a las proporciones de tres genotipos AA, Aa y aa. Si se utiliza p_1, p_2 y p_3 para denotar estas proporciones (probabilidades), se podría desear probar

$$H_0: p_1 = \theta^2, p_2 = 2\theta(1-\theta), p_3 = (1-\theta)^2 \qquad (14.1)$$

donde θ representa la proporción del gen A en la población. Esta hipótesis es **compuesta**, porque sabiendo que H_0 es verdadera, no determina de manera única las probabilidades de celda y las cuentas de celda esperadas, sino sólo su forma general. Para llevar a cabo una prueba χ^2 primero se deben estimar las θ_i desconocidas.

De manera similar, se podría estar interesado en probar si una muestra provino de una determinada familia de distribuciones sin especificar algún miembro en particular de la familia. Para usar la prueba χ^2 a fin de ver si la distribución es de Poisson, por ejemplo, se debe estimar el parámetro λ. Además, debido a que en realidad hay una cantidad infinita de valores posibles de una variable de Poisson, estos valores se deben agrupar de modo que haya un número finito de celdas. Si H_0 establece que la distribución subyacente es normal, la elección de las celdas y la estimación de μ y σ deben preceder al uso de la prueba χ^2.

χ^2 cuando se lleva a cabo la estimación de los parámetros

Como antes, k representa el número de categorías o celdas y p_i representa la probabilidad de que una observación caiga en la i-ésima celda. Ahora la hipótesis nula establece que cada p_i es una función de un pequeño número de parámetros $\theta_1, \ldots, \theta_m$ con las θ_i por lo demás no especificadas:

$$\begin{aligned} &H_0\text{: } p_1 = \pi_1(\boldsymbol{\theta}), \ldots, p_k = \pi_k(\boldsymbol{\theta}) \qquad \text{donde } \boldsymbol{\theta} = (\theta_1, \ldots, \theta_m) \\ &H_a\text{: la hipótesis nula } H_0 \text{ no es verdadera} \end{aligned} \tag{14.2}$$

Por ejemplo, para H_0 de (14.1), $m = 1$ (sólo hay una θ), $\pi_1(\theta) = \theta^2$, $\pi_2(\theta) = 2\theta(1 - \theta)$ y $\pi_3(\theta) = (1 - \theta)^2$.

En el caso $k = 2$, en realidad sólo hay una sola variable aleatoria, N_1 (puesto que $N_1 + N_2 = n$), que tiene una distribución binomial. Entonces, la probabilidad conjunta de que $N_1 = n_1$ y $N_2 = n_2$ es

$$P(N_1 = n_1, N_2 = n_2) = \binom{n}{n_1} p_1^{n_1} \cdot p_2^{n_2} \propto p_1^{n_1} \cdot p_2^{n_2}$$

donde $p_1 + p_2 = 1$ y $n_1 + n_2 = n$. Para k general, la distribución conjunta de $N_1, \ldots, N_k$ es la distribución multinomial (sección 5.1) con

$$P(N_1 = n_1, \ldots, N_k = n_k) \propto p_1^{n_1} \cdot p_2^{n_2} \cdot \cdots \cdot p_k^{n_k} \tag{14.3}$$

Cuando H_0 es verdadera, (14.3) se convierte en

$$P(N_1 = n_1, \ldots, N_k = n_k) \propto [\pi_1(\boldsymbol{\theta})]^{n_1} \cdot \cdots \cdot [\pi_k(\boldsymbol{\theta})]^{n_k} \tag{14.4}$$

Para aplicar la prueba chi-cuadrada, se debe estimar, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$.

MÉTODO DE ESTIMACIÓN

> Sean $n_1, n_2, \ldots, n_k$ los valores observados de $N_1, \ldots, N_k$. Entonces $\hat{\theta}_1, \ldots, \hat{\theta}_m$ son los valores de las θ_i que maximizan la ecuación (14.4).

Los estimadores resultantes $\hat{\theta}_1, \ldots, \hat{\theta}_m$ son los estimadores de máxima verosimilitud de $\theta_1, \ldots, \theta_m$; este principio de estimación se analizó en la sección 6.2.

Ejemplo 14.5 En los humanos hay un grupo sanguíneo, el grupo MN, que se compone de individuos que tienen uno de los tres tipos de sangre M, MN y N. El tipo se determina mediante dos alelos, y no

hay dominancia, así que los tres genotipos posibles dan lugar a tres genotipos. Una población que consiste en los individuos del grupo MN está en equilibrio si

$$P(\mathrm{M}) = p_1 = \theta^2$$
$$P(\mathrm{MN}) = p_2 = 2\theta(1 - \theta)$$
$$P(\mathrm{N}) = p_3 = (1 - \theta)^2$$

para alguna θ. Suponga que una muestra de una población de esta clase produjo los resultados mostrados en la tabla 14.4.

Tabla 14.4 Cuentas observadas para el ejemplo 14.5

Tipo	M	MN	N	
Observada	125	225	150	$n = 500$

Entonces

$$[\pi_1(\theta)]^{n_1}[\pi_2(\theta)]^{n_2}[\pi_3(\theta)]^{n_3} = [(\theta^2)]^{n_1}[2\theta(1 - \theta)]^{n_2}[(1 - \theta)^2]^{n_3}$$
$$= 2^{n_2} \cdot \theta^{2n_1 + n_2} \cdot (1 - \theta)^{n_2 + 2n_3}$$

Al maximizar esto con respecto a θ (o, de manera equivalente, maximizar el logaritmo natural de esta cantidad, que es más fácil derivar), se obtiene

$$\hat{\theta} = \frac{2n_1 + n_2}{[(2n_1 + n_2) + (n_2 + 2n_3)]} = \frac{2n_1 + n_2}{2n}$$

Con $n_1 = 125$ y $n_2 = 225$, $\hat{\theta} = 475/1000 = 0.475$. ■

Una vez que se estimó $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ mediante $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_m)$, las cuentas de celda esperadas estimadas son las $n\pi_i(\hat{\boldsymbol{\theta}})$. Éstas ahora se utilizan en lugar de las np_{i0} de la sección 14.1 para especificar un estadístico χ^2.

TEOREMA

En condiciones de "regularidad" generales en $\theta_1, \ldots, \theta_m$ y las $\pi_i(\boldsymbol{\theta})$, si $\theta_1, \ldots, \theta_m$ se estiman con el método de máxima verosimilitud, como se describió antes, y n es grande,

$$\chi^2 = \sum_{\text{todas las celdas}} \frac{(\text{observada} - \text{esperada estimada})^2}{\text{esperada estimada}} = \sum_{i=1}^{k} \frac{[N_i - n\pi_i(\hat{\boldsymbol{\theta}})]^2}{n\pi_i(\hat{\boldsymbol{\theta}})}$$

tiene una distribución chi-cuadrada aproximada con $k - 1 - m$ gl cuando H_0 de (14.2) es verdadera. Una prueba aproximadamente de nivel α de H_0 contra H_a es rechazar H_0 si $\chi^2 \geq \chi^2_{\alpha, k-1-m}$. En la práctica, la prueba se usa si $n\pi_i(\hat{\boldsymbol{\theta}}) \geq 5$ para toda i.

Observe que *el número de grados de libertad se reduce por la cantidad de θ_i estimadas.*

Ejemplo 14.6 (Continuación del ejemplo 14.5) Con $\hat{\theta} = 0.475$ y $n = 500$, las cuentas de celda esperadas estimadas son $n\pi_1(\hat{\theta}) = 500(\hat{\theta})^2 = 112.81$, $n\pi_2(\hat{\theta}) = (500)(2)(0.475)(1 - 0.475) = 249.38$ y $n\pi_3(\hat{\theta}) = 500 - 112.81 - 249.38 = 137.81$. Entonces

$$\chi^2 = \frac{(125 - 112.81)^2}{112.81} + \frac{(225 - 249.38)^2}{249.38} + \frac{(150 - 137.81)^2}{137.81} = 4.78$$

Puesto que $\chi^2_{.05,k-1-m} = \chi^2_{.05,3-1-1} = \chi^2_{.05,1} = 3.843$ y $4.78 \geq 3.843$, se rechaza H_0. En la tabla A.11 del apéndice se muestra que el valor $P \approx 0.029$. ■

Ejemplo 14.7 Considere una serie de juegos entre dos equipos, I y II, que termina tan pronto como un equipo haya ganado cuatro juegos (sin ninguna posibilidad de empate). Un modelo de probabilidad simple para esta serie supone que los resultados de juegos sucesivos son independientes, y que la probabilidad de que el equipo I gane un determinado juego es una constante θ. Se designa de manera arbitraria al equipo I como el mejor, de modo que $\theta \geq 0.5$. Entonces, cualquier serie puede terminar después de 4, 5, 6 o 7 juegos. Represente con $\pi_1(\theta)$, $\pi_2(\theta)$, $\pi_3(\theta)$, $\pi_4(\theta)$ la probabilidad de terminación en los juegos 4, 5, 6 y 7 respectivamente. Entonces

$$\begin{aligned}
\pi_1(\theta) &= P(\text{I gana en 4 juegos}) + P(\text{II gana en 4 juegos}) \\
&= \theta^4 + (1 - \theta)^4 \\
\pi_2(\theta) &= P(\text{I gana 3 de los primeros 4 y el quinto}) \\
&\quad + P(\text{I pierde 3 de los primeros 4 y el quinto}) \\
&= \binom{4}{3}\theta^3(1 - \theta) \cdot \theta + \binom{4}{1}\theta(1 - \theta)^3 \cdot (1 - \theta) \\
&= 4\theta(1 - \theta)[\theta^3 + (1 - \theta)^3] \\
\pi_3(\theta) &= 10\theta^2(1 - \theta)^2[\theta^2 + (1 - \theta)^2] \\
\pi_4(\theta) &= 20\theta^3(1 - \theta)^3
\end{aligned}$$

En el artículo "Seven-Game Series in Sports" by Groeneveld and Meeden (*Mathematics Magazine,* 1975: 187-192) se probó el ajuste de este modelo a los resultados de los juegos definitivos de la National Hockey League durante el periodo 1943-1967 (cuando se mantuvo estable la membresía de la liga). Los datos se muestran en la tabla 14.5.

Tabla 14.5 Cuentas observadas y esperadas para el modelo simple

Celda / **Número de juegos jugados**	**1** / **4**	**2** / **5**	**3** / **6**	**4** / **7**	
Frecuencia observada	15	26	24	18	$n = 83$
Frecuencia esperada estimada	16.351	24.153	23.240	19.256	

Las cuentas de celda esperadas estimadas son $83\pi_i(\hat{\theta})$, donde $\hat{\theta}$ es el valor de θ que maximiza a

$$\{\theta^4 + (1 - \theta)^4\}^{15} \cdot \{4\theta(1 - \theta)[\theta^3 + (1 - \theta)^3]\}^{26} \cdot \{10\theta^2(1 - \theta)^2[\theta^2 + (1 - \theta)^2]\}^{24} \cdot \{20\theta^3(1 - \theta)^3\}^{18} \qquad (14.5)$$

El método de cálculo estándar no produce una fórmula directa para maximizar el valor $\hat{\theta}$, de modo que se debe calcular por medio de métodos numéricos. El resultado es $\hat{\theta} = 0.654$, a partir del cual se calculan $\pi_i(\hat{\theta})$ y las cuentas de celda esperadas. El valor calculado de χ^2 es 0.360, y (puesto que $k - 1 - m = 4 - 1 - 1 = 2$) $\chi^2_{0.10,2} = 4.605$. Por lo tanto, no hay razón para rechazar el modelo simple cuando se aplica a la serie de juegos definitivos de la NHL.

En el artículo citado, también se consideraron los datos de la serie mundial para el periodo 1903-1973. Para este modelo simple, $\chi^2 = 5.97$, así que al parecer el modelo no es apropiado. La razón sugerida para esto es que para el modelo simple

$$P(\text{la serie dura seis juegos} \mid \text{la serie dura por lo menos seis juegos}) \geq 0.5 \qquad (14.6)$$

mientras que de las 38 series que en realidad duraron al menos seis juegos, sólo 13 duraron exactamente seis. Entonces se introduce el siguiente modelo alternativo:

$$\pi_1(\theta_1, \theta_2) = \theta_1^4 + (1 - \theta_1)^4$$
$$\pi_2(\theta_1, \theta_2) = 4\theta_1(1 - \theta_1)[\theta_1^3 + (1 - \theta_1)^3]$$
$$\pi_3(\theta_1, \theta_2) = 10\theta_1^2(1 - \theta_1)^2\theta_2$$
$$\pi_4(\theta_1, \theta_2) = 10\theta_1^2(1 - \theta_1)^2(1 - \theta_2)$$

Las dos primeras π_i son idénticas al modelo simple, mientras que θ_2 es la probabilidad condicional de (14.6) (que ahora puede ser cualquier número entre 0 y 1). Los valores de $\hat{\theta}_1$ y $\hat{\theta}_2$ que maximizan la expresión análoga a la expresión (14.5) se determinan numéricamente como $\hat{\theta}_1 = 0.614$, $\hat{\theta}_2 = 0.342$. En la tabla 14.6 se muestra un resumen, y $\chi^2 = 0.384$. Puesto que se estiman dos parámetros, gl $= k - 1 - m = 1$ con $\chi^2_{0.10,1} = 2.706$, lo que indica un buen ajuste de los datos a este nuevo modelo.

Ejemplo

Tabla 14.6 Cuentas observadas y esperadas para el modelo más complejo

Número de juegos jugados	4	5	6	7
Frecuencia observada	12	16	13	25
Frecuencia esperada estimada	10.85	18.08	12.68	24.39

■

Una de las condiciones de regularidad en las θ_i del teorema es que son funcionalmente independientes entre sí. Es decir, ninguna θ_i se puede determinar de los valores de las otras θ_i, así que m es el número de parámetros funcionalmente independientes estimados. Una regla empírica para los grados de libertad en una prueba chi-cuadrada es la siguiente:

$$\chi^2 \text{ gl} = \begin{pmatrix}\text{número de cuentas de celda}\\ \text{determinadas libremente}\end{pmatrix} - \begin{pmatrix}\text{número de parámetros}\\ \text{independientes estimados}\end{pmatrix}$$

Esta regla se usa junto con varias pruebas chi-cuadrada en la siguiente sección.

Bondad del ajuste para distribuciones discretas

Muchos experimentos requieren observar una muestra aleatoria $X_1, X_2, \ldots, X_n$ de alguna distribución discreta. Es posible que se desee investigar si la distribución subyacente es un miembro de una familia particular, como la de Poisson o la binomial negativa. En el caso de una distribución de Poisson y una binomial negativa, el conjunto de valores posibles es infinito, de modo que los valores se deben agrupar en k subconjuntos antes de poder usar una prueba chicuadrada. Los agrupamientos se deben hacer de modo que la frecuencia esperada en cada celda (grupo) sea al menos 5. Entonces la última celda corresponde a X valores de $c, c + 1, c + 2, \ldots$ para algún valor c.

Este agrupamiento puede complicar de forma considerable el cálculo de las $\hat{\theta}_i$ y las cuentas de celda esperadas estimadas. Esto se debe a que el teorema requiere que las $\hat{\theta}_i$ se obtengan de las cuentas de celda $N_1, \ldots, N_k$ y no de los valores muestrales $X_1, \ldots, X_n$.

Ejemplo 14.8 En la tabla 14.7 se muestran los datos de cuenta del número de plantas *Larrea divaricata* encontradas en cada uno de 48 bloques de muestreo, según el informe del artículo "Some Sampling Characteristics of Plants and Arthropods of the Arizona Desert" (*Ecology*, 1962: 567-571).

Tabla 14.7 Cuentas observadas para el ejemplo 14.8

Celda	**1**	**2**	**3**	**4**	**5**
Número de plantas	**0**	**1**	**2**	**3**	**≥4**
Frecuencia	9	9	10	14	6

El autor ajustó una distribución de Poisson a los datos. Denote con λ el parámetro de Poisson y suponga por el momento que las seis cuentas de la celda 5 fueron en realidad 4, 4, 5, 5, 6, 6. Entonces, si los valores muestrales se denotan por $x_1, \ldots, x_{48}$, nueve de las x_i fueron 0, nueve fueron 1, etcétera. La probabilidad de la muestra observada es

$$\frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \cdot \cdots \cdot \frac{e^{-\lambda}\lambda^{x_{48}}}{x_{48}!} = \frac{e^{-48\lambda}\lambda^{\sum x_i}}{x_1! \cdot \cdots \cdot} = \frac{e^{-48\lambda}\lambda^{101}}{x_1! \cdot \cdots \cdot}$$

El valor de λ para el cual esto se maximiza es $\hat{\lambda} = \sum x_i/n = 101/48 = 2.10$ (el valor reportado en el artículo).

Sin embargo, la $\hat{\lambda}$ requerida para χ^2 se obtiene al maximizar la expresión (14.4) en vez de la verosimilitud de la muestra completa. Las probabilidades de celda son

$$\pi_i(\lambda) = \frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!} \qquad i = 1, 2, 3, 4$$

$$\pi_5(\lambda) = 1 - \sum_{i=0}^{3} \frac{e^{-\lambda}\lambda^i}{i!}$$

así que el lado derecho de (14.4) se convierte en

$$\left[\frac{e^{-\lambda}\lambda^0}{0!}\right]^9 \left[\frac{e^{-\lambda}\lambda^1}{1!}\right]^9 \left[\frac{e^{-\lambda}\lambda^2}{2!}\right]^{10} \left[\frac{e^{-\lambda}\lambda^3}{3!}\right]^{14} \left[1 - \sum_{i=0}^{3} \frac{e^{-\lambda}\lambda^i}{i!}\right]^6$$

No hay fórmula simple para $\hat{\lambda}$, el valor de maximización de λ, en esta última expresión, por lo tanto se debe obtener de forma numérica. ■

Debido a que las estimaciones de parámetro por lo común son más difíciles de calcular a partir de los datos agrupados que dé la muestra completa, casi siempre se calculan con este último método. Cuando estos estimadores "completos" se utilizan en el estadístico chi-cuadrada, se modifica la distribución del estadístico y el valor crítico $\chi^2_{\alpha,k-1-m}$ ya no especifica una prueba de nivel α.

TEOREMA

Sean $\hat{\theta}_1, \ldots, \hat{\theta}_m$ los estimadores de máxima verosimilitud de $\theta_1, \ldots, \theta_m$ basados en la muestra completa $X_1, \ldots, X_n$, e indique con χ^2 el estadístico basado en estos estimadores. Entonces el valor crítico c_α que especifica una prueba de extremo superior al nivel α satisface

$$\chi^2_{\alpha,k-1-m} \leq c_\alpha \leq \chi^2_{\alpha,k-1} \tag{14.7}$$

Este teorema indica el siguiente procedimiento de prueba:

Si $\chi^2 \geq \chi^2_{\alpha,k-1}$, rechace H_0.
Si $\chi^2 \leq \chi^2_{\alpha,k-1-m}$, no rechace H_0.
Si $\chi^2_{\alpha,k-1-m} < \chi^2 < \chi^2_{\alpha,k-1}$, reserve su opinión. (14.8)

Ejemplo 14.9 (Continuación del ejemplo 14.8)

Con $\hat{\lambda} = 2.10$, las cuentas de celda esperadas estimadas se calculan de $n\pi_i(\hat{\lambda})$, donde $n = 48$. Por ejemplo,

$$n\pi_1(\hat{\lambda}) = 48 \cdot \frac{e^{-2.1}(2.1)^0}{0!} = (48)(e^{-2.1}) = 5.88$$

De manera similar, $n\pi_2(\hat{\lambda}) = 12.34$, $n\pi_3(\hat{\lambda}) = 12.96$, $n\pi_4(\hat{\lambda}) = 9.07$ y $n\pi_5(\hat{\lambda}) = 48 - 5.88 - \cdots - 9.07 = 7.75$. Entonces,

$$\chi^2 = \frac{(9 - 5.88)^2}{5.88} + \cdots + \frac{(6 - 7.75)^2}{7.75} = 6.31$$

Puesto que $m = 1$ y $k = 5$, al nivel 0.05 se requiere $\chi^2_{0.05,3} = 7.815$ y $\chi^2_{0.05,4} = 9.488$. Debido a que $6.31 \leq 7.815$, no se rechaza H_0; al nivel 5%, la distribución de Poisson proporciona un ajuste razonable de los datos. Observe que $\chi^2_{0.10,3} = 6.251$ y $\chi^2_{0.10,4} = 7.779$, por lo que al nivel 0.10 se tendría uno que abstener de emitir un juicio acerca de si fue apropiada la distribución de Poisson. ■

A veces incluso las estimaciones de máxima verosimilitud basadas en la muestra completa son bastante difíciles de calcular. Éste es el caso, por ejemplo, para la distribución binomial negativa (generalizada) de dos parámetros. En este tipo de situaciones, se usan por lo común las estimaciones del método de momentos y la chi-cuadrada resultante se compara con $\chi^2_{\alpha,k-1-m}$, aunque no se sabe hasta qué grado el uso de estimadores de momentos afecta el valor crítico verdadero.

Bondad del ajuste para distribuciones continuas

La prueba chi-cuadrada también se puede usar para probar si la muestra proviene de una familia especificada de distribuciones continuas, como la exponencial o la normal. La elección de celdas (intervalos de clase) es incluso más arbitraria en el caso continuo que en el caso discreto. Para asegurar que la prueba chi-cuadrada es válida, las celdas se deben elegir de manera independiente de las observaciones muestrales. Una vez elegidas las celdas, casi siempre es bastante difícil estimar parámetros no especificados (como μ y σ en el caso normal) a partir de cuentas de celda observadas, así que se calculan estimaciones de máxima verosimilitud con base en la muestra completa. El valor crítico c_a de nuevo satisface la ecuación (14.7), y el procedimiento de prueba se determina mediante (14.8).

Ejemplo 14.10 El Instituto de Nutrición de América Central y Panamá (INCAP) ha realizado estudios de gran alcance y proyectos de investigación acerca de la dieta en América Central. En un estudio publicado en el número de noviembre de 1964 del *American Journal of Clinical Nutrition* ("The Blood Viscosity of Various Socioeconomic Groups in Guatemala"), las medidas de colesterol total en suero para una muestra de 49 indígenas rurales de bajos ingresos fueron las siguientes (en mg/L):

204 108 140 152 158 129 175 146 157 174 192 194 144 152 135 223 145
231 115 131 129 142 114 173 226 155 166 220 180 172 143 148 171 143
124 158 144 108 189 136 136 197 131 95 139 181 165 142 162

¿Es posible que la concentración de colesterol sérico tenga una distribución normal para esta población? Suponga que antes del muestreo, se creía que los valores posibles para μ y σ eran 150 y 30, respectivamente. Los siete intervalos de clase equiprobables para la distribución normal estándar son $(-\infty, -1.07)$, $(-1.07, -0.57)$, $(-0.57, -0.18)$, $(-0.18, 0.18)$, $(0.18, 0.57)$, $(0.57, 1.07)$ y $(1.07, \infty)$, cada punto final también proporciona la distancia en las desviaciones estándar desde la media para cualquier otra distribución normal. Para $\mu = 150$ y $\sigma = 30$, estos intervalos se convierten en $(-\infty, 117.9)$, $(117.9, 132.9)$, $(132.9, 144.6)$, $(144.6, 155.4)$, $(155.4, 167.1)$, $(167.1, 182.1)$ y $(182.1, \infty)$.

Para obtener las probabilidades de celda estimadas $\pi_1(\hat{\mu}, \hat{\sigma}), \ldots, \pi_7(\hat{\mu}, \hat{\sigma})$, primero se estiman las estimaciones de máxima verosimilitud $\hat{\mu}$ y $\hat{\sigma}$. En el capítulo 6, se mostró que la estimación de máxima verosimilitud de σ es $[\sum(x_i - \bar{x})^2/n]^{1/2}$ (en vez de s), así que con $s = 31.75$,

$$\hat{\mu} = \bar{x} = 157.02 \qquad \hat{\sigma} = \left[\frac{\sum(x_i - \bar{x})^2}{n}\right]^{1/2} = \left[\frac{(n-1)s^2}{n}\right]^{1/2} = 31.42$$

Entonces, cada $\pi_i(\hat{\mu}, \hat{\sigma})$ es la probabilidad de que una variable aleatoria normal X con media 157.02 y desviación estándar 31.42 caiga en el i-ésimo intervalo de clase. Por ejemplo,

$$\pi_2(\hat{\mu}, \hat{\sigma}) = P(117.9 \le X \le 132.9) = P(-1.25 \le Z \le -0.77) = 0.1150$$

así que $n\pi_2(\hat{\mu}, \hat{\sigma}) = 49(0.1150) = 5.64$. Las cuentas de celda observadas y las esperadas estimadas se muestran en la tabla 14.8.

El valor χ^2 calculado es 4.60. Con $k = 7$ celdas y $m = 2$ parámetros estimados, $\chi^2_{0.05,k-1} = \chi^2_{0.05,6} = 12.592$ y $\chi^2_{0.05,k-1-m} = \chi^2_{0.05,4} = 9.488$. Puesto que $4.60 \le 9.488$, una distribución normal estándar proporciona un buen ajuste para los datos.

Tabla 14.8 Cuentas observadas y esperadas para el ejemplo 14.10

Celda	**$(-\infty, 117.9)$**	**(117.9, 132.9)**	**(132.9, 144.6)**	**(144.6, 155.4)**
Observada	5	5	11	6
Esperada estimada	5.17	5.64	6.08	6.64

Celda	**(155.4, 167.1)**	**(167.1, 182.1)**	**(182.1, ∞)**
Observada	6	7	9
Esperada estimada	7.12	7.97	10.38

■

Ejemplo 14.11 En el artículo "Some Studies on Tuft Weight Distribution in the Opening Room" (*Textile Research J.*, 1976: 567-573) aparecen los datos siguientes acerca de la distribución del peso de salida adicional X (mg) debido a la inserción de hilos en fibras de algodón para el peso de entrada $x_0 = 70$.

Intervalo	**0–8**	**8–16**	**16–24**	**24–32**	**32–40**	**40–48**	**48–56**	**56–64**	**64–70**
Frecuencia observada	20	8	7	1	2	1	0	1	0
Frecuencia esperada	18.0	9.9	5.5	3.0	1.8	0.9	0.5	0.3	0.1

Los autores postularon una distribución exponencial truncada:

$$H_0\colon f(x) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda x_0}} \qquad 0 \le x \le x_0$$

La media de esta distribución es

$$\mu = \int_0^{x_0} x f(x)\, dx = \frac{1}{\lambda} - \frac{x_0 e^{-\lambda x_0}}{1 - e^{-\lambda x_0}}$$

El parámetro estimado λ se estimó al reemplazar μ por $\bar{x} = 13.086$ y resolver la ecuación resultante para obtener $\hat{\lambda} = 0.0742$ (así que $\hat{\lambda}$ es una estimación de método de momentos y no una estimación de probabilidad máxima). Entonces, con $\hat{\lambda}$ en lugar de λ en $f(x)$, las frecuencias de celda esperadas estimadas, como se mostraron antes, se calculan como

$$40\pi_i(\hat{\lambda}) = 40P(a_{i-1} \le X < a_i) = 40\int_{a_{i-1}}^{a_i} f(x)\, dx = \frac{40(e^{-\hat{\lambda}a_{i-1}} - e^{-\hat{\lambda}a_i})}{1 - e^{-\hat{\lambda}x_0}}$$

donde $[a_{i-1}, a_i)$ es el i-ésimo intervalo de clase. Para obtener cuentas de celda esperadas de al menos 5, las últimas seis celdas se combinan para producir cuentas observadas de 20, 8, 7, 5 y cuentas esperadas de 18.0, 9.9, 5.5, 6.6. Entonces el valor calculado de chi-cuadrada es

$\chi^2 = 1.34$. Debido a que $\chi^2_{0.05,2} = 5.992$, no se rechaza H_0, de modo que el modelo exponencial truncado proporciona un buen ajuste. ■

Una prueba especial para la normalidad

Las gráficas de probabilidad se introdujeron en la sección 4.6 como un método informal para evaluar la verosimilitud de alguna distribución poblacional especificada como aquella de la que se eligió la muestra dada. Mientras más recta sea la gráfica de probabilidad, más plausible es la distribución en la que se basa la gráfica. Una gráfica de probabilidad normal se emplea para comprobar si es plausible *algún* miembro de la familia de distribuciones normales. Sea que $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ indican las x_i cuando se ordenan de mayor a menor. Entonces la gráfica sugerida para comprobar la normalidad fue una gráfica de los puntos $(x_{(i)}, y_i)$, donde $y_i = \Phi^{-1}((i - 0.5)/n)$.

Una medida cuantitativa del grado al que se agrupan los puntos respecto a la recta es el coeficiente de correlación de la muestra r presentado en el capítulo 12. Considere calcular r para los n pares $(x_{(1)}, y_1), \ldots, (x_{(n)}, y_n)$. Aquí las y_i no son valores observados en una muestra aleatoria de una población y, así que las propiedades de esta r son bastante diferentes de las descritas en la sección 12.5. Sin embargo, es cierto que mientras más se desvíe r de 1, menos se asemeja la gráfica de probabilidad a una recta (recuerde que una gráfica de probabilidad debe tener una pendiente ascendente). Esta idea se puede ampliar para producir un procedimiento de prueba formal: Rechace la hipótesis de normalidad poblacional si $r \leq c_\alpha$, donde c_α es un valor crítico elegido para producir el nivel de significación α deseado. Es decir, el valor crítico se elige a fin de que cuando la distribución poblacional sea en realidad formal, la probabilidad de obtener un valor r que sea a lo sumo c_α (y, por consiguiente, rechazar de manera incorrecta H_0) sea la α deseada. Los programadores del paquete de software para estadística MINITAB dieron valores críticos para $\alpha = 0.10$, 0.05 y 0.01 en combinación con diferentes tamaños de muestra. Estos valores críticos se basan en una definición un poco distinta de las y_i a la mencionada antes.

MINITAB también construye una gráfica de probabilidad normal con base en estas y_i. La gráfica será casi idéntica a la que se basa en las y_i previas. Cuando hay varias $x_{(i)}$ iguales, MINITAB calcula r con el promedio de las y_i correspondientes como el segundo número en cada par.

> Sea $y_i = \Phi^{-1}[(i - 0.375)/(n + 0.25)]$ y calcule el coeficiente de correlación muestral r para los n pares $(x_{(1)}, y_1), \ldots, (x_{(n)}, y_n)$. La prueba de Ryan-Joiner de
>
> H_0: la distribución poblacional es normal
>
> contra
>
> H_a: la distribución poblacional no es normal
>
> consiste en rechazar H_0 cuando $r \leq c_\alpha$. Los valores críticos c_α aparecen en la tabla A.12 del apéndice para varios niveles de significación α y tamaños de muestra n.

Ejemplo 14.12 La siguiente muestra de $n = 20$ observaciones en el voltaje disruptivo dieléctrico de una pieza de resina epóxica se presentaron por vez primera en el ejemplo 4.29.

y_i	−1.871	−1.404	−1.127	−0.917	−0.742	−0.587	−0.446	−0.313	−0.186	−0.062
$x_{(i)}$	24.46	25.61	26.25	26.42	26.66	27.15	27.31	27.54	27.74	27.94
y_i	0.062	0.186	0.313	0.446	0.587	0.742	0.917	1.127	1.404	1.871
$x_{(i)}$	27.98	28.04	28.28	28.49	28.50	28.87	29.11	29.13	29.50	30.88

Se pide a MINITAB que lleve a cabo la prueba de Ryan-Joiner, y el resultado aparece en la figura 14.3. El valor del estadístico de prueba es $r = 0.9881$, y en la tabla A.12 del apéndice se obtiene 0.9600 como el valor crítico que capta el área de extremo inferior 0.10 bajo la curva de distribución de muestreo r cuando $n = 20$ y la distribución subyacente es en realidad normal. Puesto que $0.9881 > 0.9600$, la hipótesis nula de normalidad no se puede rechazar incluso para un nivel de significación tan grande como 0.10.

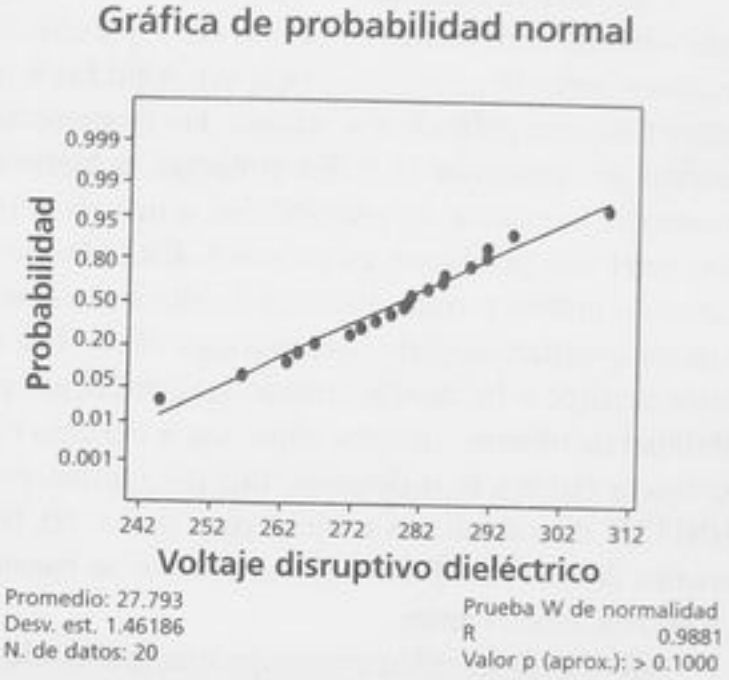


Figura 14.3 Resultado de MINITAB de la prueba de Ryan-Joiner para los datos del ejemplo 14.12 ■

Ejercicios | Sección 14.2 (12 a 23)

12. Considere una gran población de familias en las que cada una tiene exactamente tres hijos. Si el género de los tres hijos de alguna familia son independientes entre sí, el número de hijos varones en una familia seleccionada al azar tendrá una distribución binomial basada en tres ensayos.

a. Suponga que una muestra aleatoria de 160 familias produce los siguientes resultados. Pruebe las hipótesis pertinentes procediendo como en el ejemplo 14.5.

Número de hijos varones	0	1	2	3
Frecuencia	14	66	64	16

b. Suponga que una muestra aleatoria de familias de una población no humana dio como resultado frecuencias observadas de 15, 20, 12 y 3, respectivamente. ¿La prueba chi-cuadrada se basaría en el mismo número de grados de libertad que la prueba del inciso (a)? Explique.

13. En un estudio de esterilidad en la mosca de la fruta ("Hybrid Dysgenesis in *Drosophila melanogaster*: The Biology of Female and Male Sterility", *Genetics*, 1979: 161-174) se obtuvieron los siguientes datos sobre el número de ovarios que desarrolló cada mosca hembra en una muestra de tamaño 1388. Un modelo para la esterilidad unilateral establece que cada ovario se

desarrolla con cierta probabilidad p independientemente del otro ovario. Pruebe el ajuste de este modelo mediante χ^2.

x = Número de ovarios	0	1	2
Cuenta observada	1212	118	58

14. En el artículo "Feeding Ecology of the Red-Eyed Vireo and Associated Foliage-Gleaning Birds" (*Ecological Monographs*, 1971: 129-152) aparecen los siguientes datos sobre la variable X = número de saltos antes del primer vuelo y precedidos por un vuelo. Entonces, el autor propuso y ajustó una distribución de probabilidad geométrica [$p(x) = P(X = x) = p^{x-1} \cdot q$ para $x = 1, 2, \ldots$, donde $q = 1 - p$] para los datos. El tamaño muestral total fue $n = 130$.

x	1	2	3	4	5	6	7	8	9	10	11	12
Número de veces que se observó x	48	31	20	9	6	5	4	2	1	1	2	1

a. La probabilidad es $(p^{x_1-1} \cdot q) \cdot \cdots \cdot (p^{x_n-1} \cdot q) = p^{\sum x_i - n} \cdot q^n$. Demuestre que la estimación de máxima verosimilitud de p está dada por $\hat{p} = (\sum x_i - n)/\sum x_i$, y calcule $\hat{p}$ para los datos proporcionados.

b. Estime las cuentas de celda esperadas con $\hat{p}$ del inciso (a) [cuentas de celda esperadas $= n \cdot (\hat{p})^{x-1} \cdot \hat{q}$ para $x = 1, 2, \ldots$], y pruebe el ajuste del modelo mediante una prueba χ^2 combinando las cuentas para $x = 7, 8, \ldots$, y 12 en una celda ($x \geq 7$).

15. Cierto tipo de linterna de mano se vende con las cuatro pilas incluidas. Se obtiene una muestra aleatoria de 150 linternas y se determina el número de pilas defectuosas en cada una, con los siguientes resultados:

Número de pilas defectuosas	0	1	2	3	4
Frecuencia	26	51	47	16	10

Sea X el número de pilas defectuosas en una linterna elegida al azar. Pruebe la hipótesis nula de que la distribución de X es Bin(4, θ). Es decir, con $p_i = P(i$ pilas defectuosas), pruebe

$$H_0\colon\ p_i = \binom{4}{i}\theta^i(1-\theta)^{4-i} \qquad i = 0, 1, 2, 3, 4$$

[*Sugerencia*: para obtener la estimación de máxima verosimilitud de θ, escriba la probabilidad (función por maximizar) como $\theta^u(1-\theta)^v$, donde los exponentes u y v son funciones lineales de las cuentas de celda. Luego, tome el logaritmo natural, derive con respecto a θ, iguale el resultado a 0 y obtenga el valor de $\hat{\theta}$.]

16. En un experimento de genética, los investigadores consideraron 300 cromosomas de un tipo particular y contaron el número de intercambios de hermana cromátida en cada uno ("On the Nature of Sister-Chromatid Exchanges in 5-Bromodeoxyuridine-Substituted Chromosomes", *Genetics*, 1979: 1251-1264). Se estableció como hipótesis un modelo de Poisson para la distribución del número de intercambios. Pruebe el ajuste de una distribución de Poisson a los datos ajustando primero λ y combinando después las cuentas para $x = 8$ y $x = 9$ en una celda.

X = Número de intercambios	0	1	2	3	4	5	6	7	8	9
Cuentas observadas	6	24	42	59	62	44	41	14	6	2

17. En un artículo de *Annals of Mathematical Statistics* aparecen los siguientes datos del número de taladros en cada uno de 120 grupos de taladros. ¿La fmp de Poisson proporciona un modelo verosímil para la distribución del número de taladros en un grupo? (*Sugerencia*: agregue las frecuencias para 7, 8, . . . , 12 para establecer una sola categoría "≥ 7".)

Número de taladros	0	1	2	3	4	5	6	7	8	9	10	11	12
Frecuencia	24	16	16	18	15	9	6	5	3	4	3	0	1

18. En el artículo "A Probabilistic Analysis of Dissolved Oxygen–Biochemical Oxygen Demand Relationship in Streams" (*J. Water Resources Control Fed.*, 1969: 73-90) se muestran datos de la rapidez de oxigenación en corrientes a 20°C en cierta región. La media muestral y la desviación estándar se calcularon como $\bar{x} = 0.173$ y $s = 0.066$, respectivamente. Con base en la siguiente distribución de frecuencia, ¿se puede concluir que la rapidez de oxigenación es una variable distribuida normalmente? Utilice la prueba chi-cuadrada con $\alpha = 0.05$.

Rapidez (por día)	Frecuencia
Abajo de 0.100	12
0.100–abajo de 0.150	20
0.150–abajo de 0.200	23
0.200–abajo de 0.250	15
0.250 o más	13

19. Cada faro delantero de un automóvil al que se le practica una inspección anual, se puede enfocar demasiado alto (H), demasiado bajo (L), o de manera adecuada (N). Al probar los dos faros al mismo tiempo (y sin distinguir entre izquierdo y derecho) se obtienen los seis posibles resultados HH, LL, NN, HL, HN y LN. Si las probabilidades (proporciones poblacionales) para la

única dirección de enfoque de un faro son $P(H) = \theta_1$, $P(L) = \theta_2$ y $P(N) = 1 - \theta_1 - \theta_2$ y los dos faros se enfocan de forma independiente, las probabilidades de los seis resultados para un automóvil seleccionado al azar son las siguientes:

$$p_1 = \theta_1^2 \qquad p_2 = \theta_2^2 \qquad p_3 = (1 - \theta_1 - \theta_2)^2$$
$$p_4 = 2\theta_1\theta_2 \qquad p_5 = 2\theta_1(1 - \theta_1 - \theta_2)$$
$$p_6 = 2\theta_2(1 - \theta_1 - \theta_2)$$

Utilice los siguientes datos para probar la hipótesis nula

$$H_0: p_1 = \pi_1(\theta_1, \theta_2), \ldots, p_6 = \pi_6(\theta_1, \theta_2)$$

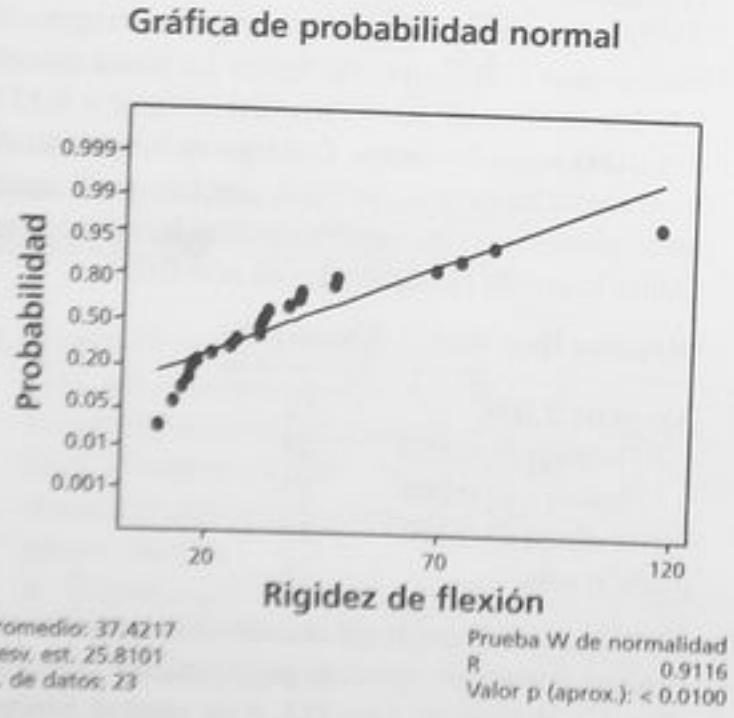
donde las $\pi_1(\theta_1, \theta_2)$ se proporcionaron antes.

Resultado	*HH*	*LL*	*NN*	*HL*	*HN*	*LN*
Frecuencia	49	26	14	20	53	38

(*Sugerencia*: escriba la probabilidad como una función de θ_1 y θ_2, tome el logaritmo natural, luego calcule $\partial/\partial\theta_1$ y $\partial/\partial\theta_2$, iguálelas a cero y determine $\hat{\theta}_1$, $\hat{\theta}_2$.)

20. En el artículo "Compatibility of Outer and Fusible Interlining Fabrics in Tailored Garments (*Textile Res. J.*, 1997: 137-142) aparecen las observaciones siguientes sobre la rigidez de flexión (μN · m) para especímenes de tela de mediana calidad, de las cuales se obtuvo el siguiente resultado de MINITAB:

24.6 12.7 14.4 30.6 16.1 9.5 31.5 17.2
46.9 68.3 30.8 116.7 39.5 73.8 80.6 20.3
25.8 30.9 39.2 36.8 46.6 15.6 32.3



¿Utilizaría un intervalo de confianza t con una muestra para estimar la rigidez de flexión promedio real? Explique su razonamiento.

21. En el artículo del que se obtuvieron los datos del ejercicio 20, también aparecen los siguientes datos sobre la relación masa compuesta/masa de la tela exterior para especímenes de tela de alta calidad:

1.15 1.40 1.34 1.29 1.36 1.26 1.22 1.40
1.29 1.41 1.32 1.34 1.26 1.36 1.36 1.30
1.28 1.45 1.29 1.28 1.38 1.55 1.46 1.32

Con MINITAB se obtuvo $r = 0.9852$ como el valor del estadístico de la prueba Ryan-Joiner y un valor $P > 0.10$. ¿Utilizaría la prueba t con una muestra para probar hipótesis acerca del valor de la relación promedio real? ¿Por qué sí o por qué no?

22. El siguiente conjunto de datos consiste en 25 observaciones sobre la resistencia a la fractura de la placa base de acero maraginizado, con 18% de níquel (de "Fracture Testing of Weldments", *ASTM Special Publ. No. 381*, 1965: 328-356). Utilice la prueba del coeficiente de correlación de la gráfica de probabilidad normal para decidir si una distribución normal proporciona un modelo verosímil para la resistencia a la fractura.

69.5 71.9 72.6 73.1 73.3 73.5 74.1 74.2 75.3
75.5 75.7 75.8 76.1 76.2 76.9 77.0 77.9 78.1
79.6 79.7 79.9 80.1 82.2 83.7 93.7

23. En el artículo "Nonbloated Burned Clay Aggregate Concrete" (*J. Materials*, 1972: 555-563) aparecen los siguientes datos sobre la resistencia a la flexión de 7 días de muestras de concreto sin dilatación complementadas con arcilla quemada (lb/pulg^2):

257 327 317 300 340 340 343 374 377 386
383 393 407 407 434 427 440 407 450 440
456 460 456 476 480 490 497 526 546 700

En el nivel de significación 0.10 determine si la resistencia a la flexión es una variable distribuida normalmente.

14.3 Tablas de contingencia bidireccionales

En las dos secciones previas se analizaron problemas inferenciales en los que los datos de cuenta se mostraron en una tabla rectangular de celdas. Cada tabla consistió en un renglón y un número especificado de columnas, donde las columnas correspondieron a las categorías en las que se dividió la población. Ahora se estudian problemas en los que los datos también consisten en cuentas o frecuencias, pero la tabla de datos ahora tiene I renglones ($I \geq 2$) y J columnas y, por lo tanto, IJ celdas. Hay dos situaciones comunes en las que surge este tipo de datos:

1. Hay I poblaciones de interés, cada una correspondiente a un renglón diferente de la tabla, y cada población se divide en las mismas J categorías. Se toma una muestra de la i-ésima población ($i = 1, \ldots, I$), y las cuentas se introducen en las celdas del i-ésimo renglón de la tabla. Por ejemplo, los clientes de cada uno de $I = 3$ cadenas de tiendas departamentales podrían tener disponibles las mismas $J = 5$ categorías de pago: efectivo, cheque, tarjeta de crédito del almacén, Visa y MasterCard.

2. Hay una sola población de interés, en la que cada individuo de la población se clasifica con respecto a dos factores distintos. Hay I categorías relacionadas con el primer factor y J categorías relacionadas con el segundo factor. Se toma una sola muestra, y el número de individuos que pertenecen a la categoría i del factor 1 y la categoría j del factor 2 se introducen en la celda del renglón i, columna j ($i = 1, \ldots, I$; $j = 1, \ldots, J$). Como ejemplo, los clientes que hacen una compra se podrían clasificar según el departamento en el que se hizo la compra, con $I = 6$ departamentos, y según el método de pago, con $J = 5$ como en (1).

Sea n_{ij} el número de individuos de la muestra que caen en la (i, j)-ésima celda (renglón i, columna j) de la tabla; es decir, la cuenta de la (i, j)-ésima celda. La tabla en la que aparecen las n_{ij} se llama **tabla de contingencia bidireccional**; un prototipo se muestra en la tabla 14.9.

Tabla 14.9 Tabla de contingencia bidireccional

	1	2	...	j	...	J
1	n_{11}	n_{12}	...	n_{1j}	...	n_{1J}
2	n_{21}					⋮
⋮	⋮					
i	n_{i1}	...		n_{ij}	...	
⋮	⋮					
I	n_{I1}	...				n_{IJ}

En situaciones del tipo 1, se desea investigar si las proporciones de las diferentes categorías son las mismas para todas las poblaciones. La hipótesis nula expresa que las poblaciones son *homogéneas* con respecto a estas categorías. En situaciones del tipo 2, se investiga si las categorías de los dos factores ocurren de manera independiente entre sí en la población.

Prueba de homogeneidad

Se supone que cada individuo en cada una de las I poblaciones pertenece exactamente a una de J categorías. Se toma una muestra de n_i individuos de la i-ésima población; sea $n = \sum n_i$ y

$$n_{ij} = \text{número de individuos en la } i\text{-ésima muestra que cae en la categoría } j$$

$$n_{\cdot j} = \sum_{i=1}^{I} n_{ij} = \begin{array}{l}\text{número total de individuos entre los } n \\ \text{muestreados que caen en la categoría } j\end{array}$$

Sea n_{ij} se registran en una tabla de contingencia bidireccional con I renglones y J columnas. La suma de las n_{ij} del i-ésimo renglón es n_i, mientras que la suma de los elementos de la j-ésima columna es $n_{\cdot j}$.

Sea

$$p_{ij} = \begin{array}{l}\text{proporción de individuos en la población } i \\ \text{que caen en la categoría } j\end{array}$$

Así, para la población 1, las J proporciones son $p_{11}, p_{12}, \ldots, p_{1J}$ (que suman 1) y de manera similar para las otras poblaciones. La **hipótesis nula de homogeneidad** expresa que la proporción de individuos en la categoría j es la misma para cada población, y que esto es cierto para cada categoría; es decir, para toda j, $p_{1j} = p_{2j} = \cdots = p_{Ij}$.

Cuando H_0 es verdadera, se puede usar $p_1, p_2, \ldots, p_J$ para denotar las proporciones poblacionales en las J diferentes categorías; estas proporciones son comunes a las I poblaciones. Entonces, el número esperado de individuos en la i-ésima muestra que caen en la j-ésima categoría cuando H_0 es verdadera es $E(N_{ij}) = n_i \cdot p_j$. Para estimar $E(N_{ij})$, se debe estimar primero p_j de la proporción en la categoría j. Entre la muestra total de n individuos, $N_{\cdot j}$ caen en la categoría j, así que se utiliza $\hat{p}_j = N_{\cdot j}/n$ como el estimador (se puede demostrar que éste es estimador de máxima verosimilitud de p_j). La sustitución de la estimación $\hat{p}_j$ por p_j en $n_i p_j$ produce una fórmula simple para las cuentas esperadas estimadas bajo H_0:

$$\begin{aligned}\hat{e}_{ij} &= \text{cuenta esperada estimada en la celda } (i, j) = n_i \cdot \frac{n_{\cdot j}}{n} \\ &= \frac{(\text{total del } i\text{-ésimo renglón})\ (\text{total de la } j\text{-ésima columna})}{n}\end{aligned} \qquad (14.9)$$

El estadístico de prueba también tiene la misma forma que en situaciones de problemas anteriores. El número de grados de libertad proviene de la regla empírica general. En cada renglón de la tabla 14.9 hay $J - 1$ cuentas de celda determinadas libremente (cada tamaño de muestra n_i es fijo), así que hay un total de $I(J - 1)$ celdas determinadas sin restricciones. Se estiman

los parámetros $p_1, \ldots, p_J$ pero como $\sum p_i = 1$, sólo $J - 1$ de éstos son independientes. Así, $\text{gl} = I(J - 1) - (J - 1) = (J - 1)(I - 1)$.

Hipótesis nula: H_0: $p_{1j} = p_{2j} = \cdots = p_{Ij} \qquad j = 1, 2, \ldots, J$

Hipótesis alternativa: H_a: H_0 no es verdadera

Valor del estadístico de prueba:

$$\chi^2 = \sum_{\text{todas las celdas}} \frac{(\text{observada} - \text{estimada esperada})^2}{\text{estimada esperada}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(n_{ij} - \hat{e}_{ij})^2}{\hat{e}_{ij}}$$

Región de rechazo: $\chi^2 \geq \chi^2_{\alpha,(I-1)(J-1)}$

La información del valor P se obtiene como se describió en la sección 14.1. La prueba se puede aplicar de forma segura siempre que $\hat{e}_{ij} \geq 5$ para todas las celdas.

Ejemplo 14.13 Una compañía envasa un determinado producto en latas de tres tamaños, cada una con una línea de producción distinta. La mayor parte de las latas cumplen con las especificaciones, pero un ingeniero de control de calidad identificó las siguientes incongruencias:

1. Mancha en la lata.
2. Fisura en la lata.
3. Ubicación inadecuada de la lengüeta para destapar la lata.
4. Falta de la lengüeta para destapar la lata.
5. Otra.

Se selecciona una muestra de unidades defectuosas de cada una de las tres líneas y cada unidad se clasifica según la razón por la que incumple las especificaciones, lo que da como resultado los siguientes datos de tabla de contingencia:

		Razón por la que se incumplen las especificaciones					Tamaño de la muestra
		Manchas	Fisura	Ubicación	Falta	Otra	
Línea de producción	1	34	65	17	21	13	150
	2	23	52	25	19	6	125
	3	32	28	16	14	10	100
	Total	89	145	58	54	29	375

¿Los datos indican que las proporciones que caen en las distintas categorías de incumplimiento de las especificaciones no son las mismas para las tres líneas? Los parámetros de interés son las distintas proporciones, y las hipótesis pertinentes son

H_0: las líneas de producción son homogéneas con respecto a las cinco categorías de incumplimiento de las especificaciones; es decir, $p_{1j} = p_{2j} = p_{3j}$ para $j = 1, \ldots, 5$

H_a: las líneas de producción no son homogéneas con respecto a las categorías

Ahora se deben calcular las frecuencias esperadas estimadas (suponiendo homogeneidad). Considere la primera categoría incumplimiento de las especificaciones para la primera línea de producción. Cuando las líneas son homogéneas,

$$\text{el número esperado estimado entre las 150 unidades seleccionadas que tienen mancha} = \frac{(\text{primer total de renglón})(\text{primer total de columna})}{\text{total de tamaños de muestra}} = \frac{(150)(89)}{375} = 35.60$$

Entonces la contribución de la celda de la esquina superior izquierda a χ^2 es

$$\frac{(\text{observada} - \text{esperada estimada})^2}{\text{esperada estimada}} = \frac{(34 - 35.60)^2}{35.60} = 0.072$$

Las otras contribuciones se calculan de manera similar. En la figura 14.4 se muestra el resultado de MINITAB para la prueba chi-cuadrada. La cuenta observada es el número de la parte superior de cada celda, y directamente abajo está la cuenta esperada estimada. La contribución de cada celda a χ^2 aparece abajo de las cuentas, y el valor de la estadística de prueba es $\chi^2 = 14.159$. Las cuentas esperadas estimadas son al menos 5, así que es innecesario combinar las categorías. La prueba se basa en $(3 - 1)(5 - 1) = 8$ gl. En la tabla A.11 del apéndice se muestra que los valores que captan las áreas de extremo superior de 0.08 y 0.075 bajo la curva de 8 gl son 14.06 y 14.26, respectivamente. Así que el valor P está entre 0.075 y 0.08; MINITAB da el valor = 0.079. La hipótesis nula de homogeneidad no se debe rechazar a los niveles de significación usuales de 0.05 o 0.01, pero se rechazaría para un valor de α mayor que 0.10.

```
Las cuentas esperadas se imprimen debajo de las cuentas observadas

             mancha     fisura   ubicación      falta      otra      Total

1                34         65          17         21         13        150
              35.60      58.00       23.20      21.60      11.60

2                23         52          25         19          6        125
              29.67      48.33       19.33      18.00       9.67

3                32         28          16         14         10        100
              23.73      38.67       15.47      14.40       7.73

Total            89        145          58         54         29        375

Chi-cuadrada = 0.072 + 0.845 + 1.657 + 0.017 + 0.169 + 1.498 + 0.278 +
               1.661 + 0.056 + 1.391 + 2.879 + 2.943 + 0.018 + 0.011 +
               0.664 = 14.159

gl = 8,  p = 0.079
```

Figura 14.4 Resultado de MINITAB para la prueba chi-cuadrada del ejemplo 14.13 ■

Prueba de independencia

Ahora se centra la atención en la relación entre dos factores distintos en una sola población. El número de categorías del primer factor se denota con I y el número de categorías del segundo factor con J. Se supone que cada individuo de la población pertenece exactamente a una de las I categorías relacionadas con el primer factor y exactamente a una de las J categorías relacionadas con el segundo factor. Por ejemplo, la población de interés podría consistir en los individuos que por lo común ven las noticias nacionales en la televisión, con la red preferida como el primer factor (ABC, CBS, NBC o PBS, así que $I = 4$) y la filosofía política como el segundo factor (liberal, moderado o conservador, de modo que $J = 3$).

Para una muestra de n individuos tomados de la población, sea n_{ij} el número entre los n que caen en la categoría i del primer factor y la categoría j del segundo factor. Las n_{ij} se pueden mostrar en una tabla de contingencia bidireccional con I renglones y J columnas. En el caso de homogeneidad para I poblaciones, los totales de renglón se fijaron por adelantado, y sólo los totales de columna J fueron aleatorios. Ahora sólo el tamaño de muestra total es fijo, y las $n_{i\cdot}$ y las $n_{\cdot j}$ son valores observados de variables aleatorias. Para expresar las hipótesis de interés, sea

p_{ij} = la proporción de individuos de la población que pertenecen a la categoría i del factor 1 y a la categoría j del factor 2

= P(un individuo elegido al azar cae en la categoría i del factor 1 y en la categoría j del factor 2)

Entonces

$$p_{i\cdot} = \sum_j p_{ij} = P(\text{un individuo elegido al azar cae en la categoría } i \text{ del factor 1})$$

$$p_{\cdot j} = \sum_i p_{ij} = P(\text{un individuo elegido al azar cae en la categoría } j \text{ del factor 2})$$

Recuerde que dos sucesos A y B son independientes si $P(A \cap B) = P(A) \cdot P(B)$. *La hipótesis nula en este caso dice que la categoría de un individuo con respecto al factor* 1 *es independiente de la categoría con respecto al factor* 2. En símbolos, esto se convierte en $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$ para cada par (i, j).

La cuenta esperada en la celda (i, j) es $n \cdot p_{ij}$, así que cuando la hipótesis nula es verdadera, $E(N_{ij}) = n \cdot p_{i\cdot} \cdot p_{\cdot j}$. Para obtener un estadístico chi-cuadrada, se deben estimar las $p_{i\cdot}$ $(i = 1, \ldots, I)$ y las $p_{\cdot j}$ $(j = 1, \ldots, J)$. Las estimaciones (de máxima verosimilitud) son

$$\hat{p}_{i\cdot} = \frac{n_{i\cdot}}{n} = \text{proporción muestral para la categoría } i \text{ del factor 1}$$

y

$$\hat{p}_{\cdot j} = \frac{n_{\cdot j}}{n} = \text{proporción muestral para la categoría } j \text{ del factor 2}$$

Esto proporciona cuentas de celda esperadas estimadas idénticas a las del caso de homogeneidad.

$$\hat{e}_{ij} = n \cdot \hat{p}_{i\cdot} \cdot \hat{p}_{\cdot j} = n \cdot \frac{n_{i\cdot}}{n} \cdot \frac{n_{\cdot j}}{n} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{n}$$

$$= \frac{(\text{total del } i\text{-ésimo renglón})(\text{total de la } j\text{-ésima columna})}{n}$$

El estadístico de prueba también es idéntico al que se emplea para la prueba de homogeneidad, al igual que el número de grados de libertad. Esto se debe a que el número de cuentas de celda determinadas sin restricciones es $IJ - 1$, ya que sólo se fija por adelantado el total n. Hay I $p_{i\cdot}$, estimadas, pero sólo $I - 1$ se estiman de forma independiente, puesto que $\sum p_{i\cdot} = 1$,

y de manera similar $J - 1$ $p_{\cdot j}$ se estiman de manera independiente, de modo que $I + J - 2$ parámetros se estiman del mismo modo. Ahora, la regla empírica produce gl $= IJ - 1 - (I + J - 2) = IJ - I - J + 1 = (I - 1) \cdot (J - 1)$.

Hipótesis nula: H_0: $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$ $\quad i = 1, \ldots, I; j = 1, \ldots, J$

Hipótesis alternativa: H_a: H_0 no es verdadera

Valor del estadístico de prueba:

$$\chi^2 = \sum_{\text{todas las celdas}} \frac{(\text{observada} - \text{esperada estimada})^2}{\text{esperada estimada}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(n_{ij} - \hat{e}_{ij})^2}{\hat{e}_{ij}}$$

Región de rechazo: $\chi^2 \geq \chi^2_{\alpha,(I-1)(J-1)}$

De nuevo, la información del valor P se obtiene como se describió en la sección 14.1. La prueba se puede aplicar con seguridad siempre y cuando $\hat{e}_{ij} \geq 5$ para todas las celdas.

Ejemplo 14.14 En un estudio de la relación entre las condiciones de las instalaciones en las gasolineras y la política de fijación de precios de la gasolina ("An Analysis of Price Aggressiveness in Gasoline Marketing", *J. Marketing Research*, 1970: 36-42) se obtuvieron los siguientes datos con base en una muestra de $n = 441$ gasolineras. Al nivel 0.01, ¿los datos indican que las condiciones de la instalación y la política de fijación de precios son independientes entre sí? Las cuentas observada y estimada se dan en la tabla 14.10.

Tabla 14.10 Cuentas observada y esperada estimada para el ejemplo 14.14

		Política de precios observada				Política de precios observada			
		Agresiva	Neutral	No agresiva	$n_{i\cdot}$				
Condición	Subestándar	24	15	17	56	17.02	22.10	16.89	56
	Estándar	52	73	80	205	62.29	80.88	61.83	205
	Moderna	58	86	36	180	54.69	71.02	54.29	180
	$n_{\cdot j}$	134	174	133	441	134	174	133	441

Así,

$$\chi^2 = \frac{(24 - 17.02)^2}{17.02} + \cdots + \frac{(36 - 54.29)^2}{54.29} = 22.47$$

y como $\chi^2_{0.01,4} = 13.277$, se rechaza la hipótesis de independencia.

Se concluye que conocer la política de precios de una gasolinera suministra información acerca de la condición de las instalaciones en la gasolinera. En particular, las gasolineras con una política agresiva de fijación de precios al parecer tienen más probabilidades de contar con instalaciones subestándar que las gasolineras con una política neutra o no agresiva de fijación de precios. ■

Los modelos y métodos para analizar los datos en los que se clasifica cada individuo con respecto a tres o más factores (tablas de contingencia multidimensionales) se analizan en varias de las referencias del capítulo.

Ejercicios | Sección 14.3 (24 a 36)

24. La siguiente tabla bidireccional se construyó con los datos del artículo "Television Viewing and Physical Fitness in Adults" (*Research Quarterly for Exercise and Sport*, 1990: 315-320). El autor deseaba determinar si el tiempo que se pasa viendo televisión se relaciona con el buen funcionamiento cardiovascular. Se preguntó a las personas acerca de sus hábitos de ver televisión y se clasificaron como con buena condición física si caían en la categoría de excelente o muy bueno en una prueba por etapas. Se incluye el resultado de MINITAB de un análisis chi-cuadrada. Los cuatro grupos de TV correspondieron a diferentes cantidades de tiempo por día que pasan viendo televisión (0, 1-2, 3-4 o 5 o más horas). Los 168 individuos representados en la primera columna, fueron considerados como con buena condición física. Las cuentas esperadas aparecen debajo de las cuentas observadas, y MINITAB muestra la contribución a χ^2 de cada celda. Exprese y pruebe las hipótesis pertinentes con $\alpha = 0.05$.

```
             1          2      Total
1           35        147        182
         25.48     156.52

2          101        629        730
        102.20     627.80

3           28        222        250
         35.00     215.00

4            4         34         38
          5.32      32.68

Total      168       1032       1200

Chi-cuadrada = 3.557 + 0.579 +
               0.014 + 0.002 +
               1.400 + 0.228 +
               0.328 + 0.053 = 6.161

gl = 3
```

25. Los siguientes datos se refieren a marcas de hojas encontradas en muestras de tréboles blancos seleccionados de áreas de pasto largo y áreas de pasto corto ("The Biology of the Leaf Mark Polymorphism in *Trifolium repens* L.", *Heredity*, 1976: 306-325). Utilice una prueba χ^2 para determinar si las proporciones verdaderas de marcas distintas son idénticas para los dos tipos de regiones.

	Tipo de marca					Tamaño de la muestra
	L	LL	Y + YL	O	Otras	
Áreas de pasto largo	409	11	22	7	277	726
Áreas de pasto corto	512	4	14	11	220	761

26. Los siguientes datos resultaron de un experimento para estudiar los efectos de la remoción de hojas sobre la capacidad de un fruto de cierto tipo para madurar ("Fruit Set, Herbivory, Fruit Reproduction, and the Fruiting Strategy of *Catalpa speciosa*", *Ecology*, 1980: 57-64):

Tratamiento	Número de frutos maduros	Número de frutos malogrados
Control	141	206
Dos hojas eliminadas	28	69
Cuatro hojas eliminadas	25	73
Seis hojas eliminadas	24	78
Ocho hojas eliminadas	20	82

¿Los datos indican que la probabilidad de maduración de una fruta resulta afectada por el número de hojas eliminadas? Exprese y pruebe las hipótesis apropiadas al nivel 0.01.

27. En el artículo "Human Lateralization from Head to Foot: Sex-Related Factors" (*Science*, 1978: 1291-1292) aparecen datos de una muestra de hombres diestros y de una muestra de mujeres diestras respecto al número de individuos cuyos pies fueron del mismo tamaño, tenían el pie izquierdo más grande que el derecho (una diferencia de medio número o más), o tenían el pie derecho más grande que el izquierdo.

	L > R	L = R	L < R	Tamaño de la muestra
Hombres	2	10	28	40
Mujeres	55	18	14	87

¿La información indica que el género tiene un efecto importante sobre el desarrollo de la asimetría en los pies? Exprese las hipótesis nula y alternativa apropiadas, calcule el valor de χ^2, y obtenga información acerca del valor P.

28. En el artículo "Susceptibility of Mice to Audiogenic Seizure Is Increased by Handling Their Dams During Gestation" (*Science*, 1976: 427-428) se habla de la investigación acerca del efecto de diferentes tratamientos con inyecciones sobre las frecuencias de convulsiones audiogénicas.

Tratamiento	Sin respuesta	Carrera desenfrenada	Convulsión clónica	Convulsión tónica
Tienilalanina	21	7	24	44
Solvente	15	14	20	54
Imitación	23	10	23	48
Sin tratamiento	47	13	28	32

¿Los datos indican que los porcentajes verdaderos en las diferentes categorías de respuesta dependen de la naturaleza del tratamiento con inyección? Exprese y pruebe las hipótesis pertinentes con $\alpha = 0.005$.

29. Los siguientes datos en relación con combinaciones de sexo de dos recombinantes que resultan de seis genotipos masculinos diferentes aparecen en el artículo "A New Method for Distinguishing Between Meiotic and Premeiotic Recombinational Events in *Drosophila melanogaster*" (*Genetics*, 1979: 543-554). ¿Los datos apoyan la hipótesis de que la distribución de frecuencia entre las tres combinaciones de sexo es homogénea con respecto a los diferentes genotipos? Defina los parámetros de interés, exprese las hipótesis H_0 y H_a, pertinentes y realice el análisis.

		Combinación de sexo		
		M/M	M/F	F/F
Genotipo masculino	1	35	80	39
	2	41	84	45
	3	33	87	31
	4	8	26	8
	5	5	11	6
	6	30	65	20

30. Para un componente particular, se consideran tres configuraciones de diseño distintas. Hay cuatro modos de falla posibles para el componente. Un ingeniero obtuvo los siguientes datos acerca de la cantidad de fallas en cada modo para cada una de las tres configuraciones. ¿La configuración parece tener un efecto en el tipo de falla?

		Modo de falla			
		1	2	3	4
Configuración	1	20	44	17	9
	2	4	17	7	12
	3	10	31	14	5

31. Se obtuvo una muestra aleatoria de individuos que conducen solos al trabajo en una gran zona metropolitana, y se clasificó a cada individuo de acuerdo con el tamaño de su automóvil y la distancia de ida y vuelta. ¿Los datos siguientes indican que la distancia de ida y vuelta y el tamaño del automóvil se relacionan en la población muestreada? Exprese las hipótesis pertinentes y utilice una prueba chi-cuadrada al nivel 0.05.

		Distancia recorrida		
		0–<10	10–<20	≥20
Tamaño del automóvil	Subcompacto	6	27	19
	Compacto	8	36	17
	Mediano	21	45	33
	Grande	14	18	6

32. Cada individuo de una muestra aleatoria de estudiantes de preparatoria y universidad, se clasificaron con respecto a su opinión política y el consumo de mariguana, y se obtuvieron los datos que se muestran en la siguiente tabla bidireccional ("Attitudes About Marijuana and Political Views", *Psychological Reports*, 1973: 1051-1054). ¿Los datos apoyan la hipótesis de que la opinión política y el nivel de consumo de mariguana son independientes dentro de la población? Pruebe las hipótesis pertinentes con el nivel de significación 0.01.

		Nivel de consumo		
		Nunca	Rara vez	Con frecuencia
Opinión política	Liberal	479	173	119
	Conservador	214	47	15
	Otro	172	45	85

33. Demuestre que el estadístico chi-cuadrada para la prueba de independencia se puede escribir en la forma

$$\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\frac{N_{ij}^2}{\hat{E}_{ij}}\right) - n$$

¿Por qué esta fórmula es más eficiente, desde el punto de vista del cálculo, que la fórmula de definición para χ^2?

34. Suponga que en el ejercicio 32 se clasificó a cada estudiante respecto a su punto de vista político, consumo de mariguana y preferencia religiosa, con las categorías de este último factor identificadas como protestante, católico y otro. Los datos se podrían mostrar en tres tablas bidireccionales diferentes correspondientes a cada categoría del tercer factor. Con $p_{ijk} = P$(categoría política i, categoría en relación con la mariguana j y categoría religiosa k), la hipótesis nula de independencia de los tres factores establece que $p_{ijk} = p_{i\cdot\cdot}\, p_{\cdot j\cdot} p_{\cdot\cdot k}$. Sea n_{ijk} la frecuencia observada en la celda (i, j, k). Muestre cómo estimar las cuentas de celda esperadas suponiendo que H_0 es verdadera ($\hat{e}_{ijk} = n\hat{p}_{ijk}$, de modo que se deben determinar las $\hat{p}_{ijk}$). Luego, utilice la regla empírica general para determinar el número de grados de libertad para el estadístico chi-cuadrada.

35. Suponga que en un estado en particular que consiste en cuatro regiones distintas, se obtiene una muestra aleatoria de n_k votantes de la k-ésima región para $k = 1, 2, 3, 4$. Cada votante se clasifica según el candidato de su preferencia (1, 2 o 3) y de acuerdo con su registro (1 = demócrata, 2 = republicano, 3 = independiente). Sea p_{ijk} la proporción de votantes de la región k que pertenecen a la categoría de candidato i y categoría de registro j. La hipótesis nula de regiones homogéneas es H_0: $p_{ij1} = p_{ij2} = p_{ij3} = p_{ij4}$ para toda i, j (es decir, la proporción donde cada combinación candidato/registro es la misma para las cuatro regiones). Si se supone que H_0 es verdadera, determine $\hat{p}_{ijk}$ y $\hat{e}_{ijk}$ como funciones de las n_{ijk} observadas y utilice la regla empírica general para obtener el número de grados de libertad para la prueba chi-cuadrada.

36. Considere la siguiente tabla de 2×3 en la que se ilustran las proporciones muestrales que caen en las diversas combinaciones de categorías (p. ej., 13% de los elementos de la muestra estuvieron en la primera categoría de ambos factores).

	1	2	3
1	0.13	0.19	0.28
2	0.07	0.11	0.22

a. Suponga que la muestra consistió en $n = 100$ personas. Utilice la prueba chi-cuadrada para determinar la independencia con nivel de significación 0.10.

b. Repita el inciso (a) suponiendo que el tamaño muestral fue $n = 1000$.

c. ¿Cuál es el mínimo tamaño muestral n para el cual estas proporciones observadas producirían el rechazo de la hipótesis de independencia?

Ejercicios complementarios (37 a 48)

37. De acuerdo con el artículo "Birth Order and Political Success" (*Psych. Reports*, 1971: 1239-1242) de entre 31 candidatos seleccionados al azar para cargos oficiales, provenientes de familias con cuatro hijos, 12 eran primogénitos, 11 intermedios y ocho nacieron al último. Utilice estos datos para probar la hipótesis nula de que un candidato político que pertenece a este tipo de familia, tiene las mismas probabilidades de ocupar alguna de las cuatro posiciones ordinales.

38. Los resultados de un experimento para evaluar el efecto de petróleo crudo en parásitos de peces, se describen en el artículo "Effects of Crude Oils on the Gastrointestinal Parasites of Two Species of Marine Fish" (*J. Wildlife Diseases*, 1983: 253-258). Se compararon tres tratamientos (correspondientes a poblaciones del procedimiento descrito): 1) sin contaminación, 2) contaminación por petróleo expuesto a la intemperie un año y 3) contaminación por petróleo nuevo. Para cada condición de tratamiento se tomó una muestra de peces y después se clasificó cada pez como con parásitos o sin parásitos. Se proporcionan datos compatibles con los del artículo. ¿Los datos indican que los tres tratamientos difieren con respecto a la proporción verdadera de peces con y sin parásitos? Pruebe con $\alpha = 0.01$.

Tratamiento	Infectado con parásitos	Sin parásitos
Control	30	3
Petróleo viejo	16	8
Petróleo nuevo	16	16

39. En el artículo "Sex Bias and the Validity of Believed Differences Between Male and Female Interscholastic Athletic Coaches" (*Research Quarterly for Exercise and Sport*, 1990: 259-267) se compararon las aptitudes de entrenadores universitarios, tanto jefes como ayudantes hombres y mujeres. Cada persona de muestras aleatorias de 2225 entrenadores y 1141 entrenadoras se clasificó según el número de años de experiencia a fin de obtener la siguiente tabla bidireccional. ¿Hay evidencia suficiente para concluir que las proporciones que caen en las categorías de experiencia son diferentes para hombres y mujeres? Use $\alpha = 0.01$.

	Años de experiencia				
Género	**1–3**	**4–6**	**7–9**	**10–12**	**13+**
Masculino	202	369	482	361	811
Femenino	230	251	238	164	258

40. Los autores del artículo "Predicting Professional Sports Game Outcomes from Intermediate Game Scores" (*Chance*, 1992: 18-22) utilizaron una prueba chi-cuadrada para determinar si tenía algún mérito la idea de que los juegos de baloncesto no se deciden sino hasta el último cuarto, mientras que los juegos de béisbol se dirimen en la séptima entrada. También se consideraron el fútbol y el hockey. Se recopilaron datos para 189 juegos de baloncesto, 92 de béisbol, 80 de hockey y 93 de fútbol. Se realizó un muestreo aletorio de los juegos de béisbol y fútbol realizados durante la temporada de 1990 y de los juegos de baloncesto y hockey de la temporada 1990-1991. Para cada juego se determinó el líder de la parte final del encuentro, y después se observó si éste en realidad terminaba ganando el juego. Los datos resultantes se resumen en la tabla siguiente.

Deporte	**Gana el líder del final del juego**	**Pierde el líder del final del juego**
Baloncesto	150	39
Béisbol	86	6
Hockey	65	15
Fútbol	72	21

Los autores expresan que "*líder de la parte final del juego* se define como el equipo que está adelante después de tres cuartos en baloncesto y fútbol, dos periodos en hockey y siete entradas en béisbol. El valor de chi-cuadrada con tres grados de libertad es 10.52 ($P < 0.015$)".

a. Exprese las hipótesis pertinentes y llegue a una conclusión con $\alpha = 0.05$.

b. ¿Considera que su conclusión extraordinaria en el inciso (a) se puede atribuir a un solo deporte?

41. La siguiente tabla de frecuencias bidireccional aparece en el artículo "Marijuana Use in College" (*Youth and Society*, 1979: 323-334). Se clasificó a cada uno de 445 estudiantes según la frecuencia de consumo de mariguana y el consumo de alcohol y de drogas psicotrópicas por parte de los padres. ¿Los datos indican que los hábitos de consumo de padres y estudiantes son independientes en la población de la que se tomó la muestra? Utilice el método del valor P para llegar a una conclusión.

		Nivel estándar del uso de drogas psicotrópicas		
		Nunca	**Ocasional**	**Regular**
Padres que consumen alcohol y drogas	**Ninguno**	141	54	40
	Uno	68	44	51
	Ambos	17	11	19

42. En un estudio de 2989 fallecimientos por cáncer, se registró el lugar de la muerte (casa, hospital de cuidados intensivos, instalación de atención crónica) y la edad al momento de morir, con los resultados mostrados en la tabla de frecuencias bidireccional ("Where Cancer Patients Die", *Public Health Reports*, 1983: 173). Con un nivel de significación 0.01, pruebe la hipótesis nula de que la edad y el lugar del fallecimiento son independientes.

		Localización		
		Casa	**Cuidados intensivos**	**Atención crónica**
Edad	**15–54**	94	418	23
	55–64	116	524	34
	65–74	156	581	109
	Más de 74	138	558	238

43. En un estudio para investigar hasta qué grado los individuos se percatan de los olores industriales en cierta región ("Annoyance and Health Reactions to Odor from Refineries and Other Industries in Carson, California", *Environmental Research*, 1978: 119-132), se obtuvo una muestra de individuos de cada una de tres áreas distintas cercanas a instalaciones industriales. A cada individuo se le preguntó si notaba los olores 1) todos los días, 2) al menos una vez a la semana, 3) por lo menos una vez al mes, 4) menos de una vez al mes o 5) nunca. A continuación se muestran los datos y el resultado de SPSS. Enuncie y pruebe las hipótesis adecuadas.

Comparación: ÁREA POR CATEGORÍA

Cuenta
Esp. Val.
CATEGORÍA → Renglón Porcentaje
ÁREA Col. Porcentaje

ÁREA	1.00	2.00	3.00	4.00	5.00	Total de renglón
1.00	20 12.7 20.6% 52.6%	28 24.7 28.9% 37.8%	23 18.0 23.7% 42.6%	14 16.0 14.4% 29.2%	12 25.7 12.4% 15.6%	97 33.3%
2.00	14 12.4 14.7% 36.8%	34 24.2 35.8% 45.9%	21 17.6 22.1% 38.9%	14 15.7 14.7% 29.2%	12 25.1 12.6% 15.6%	95 32.6%
3.00	4 12.9 4.0% 10.5%	12 25.2 12.1% 16.2%	10 18.4 10.1% 18.5%	20 16.3 20.2% 41.7%	53 26.2 53.5% 68.8%	99 34.0%
Total de columna	38 13.1%	74 25.4%	54 18.6%	48 16.5%	77 26.5%	291 100.0%

Chi-cuadrada	GL	Significación	F.E. mín	Celdas con F.E. < 5
70.64156	8	0.0000	12.405	Ninguna

44. Muchos compradores muestran descontento debido a que en las tiendas de comestibles ya no se colocan precios en cada uno de los productos. En el artículo "The Impact of Item Price Removal on Grocery Shopping Behavior" (*J. Marketing*, 1980: 73-93) se describe un estudio en el que cada comprador de una muestra se clasificó según su edad y si requería el precio del producto. Con base en los datos siguientes ¿la necesidad de ver el precio de un producto es independiente de la edad?

	Edad				
	<30	**30–39**	**40–49**	**50–59**	**≥60**
Número en la muestra	150	141	82	63	49
Número de personas que desean el precio de un producto	127	118	77	61	41

45. Sea p_1 la proporción de éxitos en una determinada población. El valor del estadístico de prueba en el capítulo 8 para probar H_0: $p_1 = p_{10}$ fue $z = (\hat{p}_1 - p_{10})/\sqrt{p_{10}p_{20}/n}$, donde $p_{20} = 1 - p_{10}$. Muestre que para el caso $k = 2$, el valor del estadístico de prueba chi-cuadrada de la sección 14.1 satisface $\chi^2 = z^2$. [*Sugerencia:* primero muestre que $(n_1 - np_{10})^2 = (n_2 - np_{20})^2$.]

46. El torneo de baloncesto NCAA comienza con 64 equipos que se reparten en cuatro torneos regionales, cada uno con 16 equipos. Luego, los 16 equipos de cada región se clasifican del 1 al 16. Durante el periodo de 12 años de 1991 a 2002, el equipo mejor clasificado ganó su torneo regional 22 veces, el equipo clasificado en segundo lugar ganó 10 veces, el clasificado en tercer lugar ganó cinco veces, y los 11 torneos regionales restantes los ganaron equipos clasificados abajo del tercer lugar. Sea P_{ij} la probabilidad de que el equipo clasificado en el lugar i en su región resulte victorioso en el juego en el que enfrenta al equipo clasificado en el lugar j. Una vez que se tienen las P_{ij} es posible calcular la probabilidad de que algún equipo clasificado en determinado lugar gane su torneo regional (un cálculo complicado porque el número de resultados en el espacio muestral es bastante grande). En el artículo "Probability Models for the NCAA Regional Basketball Tournaments" (*The American Statistician*, 1991: 35-38) se propusieron varios métodos para las P_{ij}.

a. Un modelo postulado $P_{ij} = 0.5 + \lambda(i - j)$ con $\lambda = 1/32$ (del cual $P_{16,1} = \lambda$, $P_{16,2} = 2\lambda$, etc.). Con base en esto, P(gana el clasificado #1) = 0.27477, P(gana el clasificado #2) = 0.20834 y P(gana el clasificado #3) = 0.15429. ¿Este modelo al parecer proporciona un buen ajuste para los datos?

b. Un modelo más complejo tiene $P_{ij} = 0.5 + 0.2813625(z_i - z_j)$, donde las z son medidas de resistencias relativas relacionadas con percentiles normales estándar (los percentiles para equipos sucesivos con clasificación alta están más cercanos que en el caso de equipos con menor clasificación, y

0.2813625 asegura que el intervalo de probabilidades es el mismo que para el modelo del inciso (a)). Las probabilidades resultantes de los clasificados 1, 2 o 3 que ganan sus torneos regionales son 0.45883, 0.18813 y 0.11032, respectivamente. Evalúe el ajuste de este modelo.

47. ¿Se ha preguntado alguna vez si los jugadores de fútbol sufren efectos adversos por golpear con la cabeza el balón? Los autores del artículo "No Evidence of Impaired Neurocognitive Performance in Collegiate Soccer Players" (*The Amer. J. of Sports Medicine*, 2002: 157-162) investigaron esta cuestión desde varias perspectivas.

a. En el artículo se afirmaba que 45 de los 91 jugadores de la muestra habían sufrido por lo menos una conmoción cerebral, 28 de 96 atletas de otros deportes habían sufrido al menos una conmoción cerebral, y sólo ocho de 53 del grupo de control habían sufrido por lo menos una conmoción. Analice estos datos y saque conclusiones apropiadas.

b. Para los jugadores de fútbol, el coeficiente de correlación muestral calculado de los valores de x = exposición al fútbol (número total de temporadas jugadas antes de participar en el estudio) y y = puntuación en un examen de memoria inmediata fue $r = -0.220$. Interprete este resultado.

c. A continuación se proporciona la información de resumen de las puntuaciones de un examen oral controlado de asociación de palabras para futbolistas y atletas de otros deportes:

$n_1 = 26,\ \bar{x}_1 = 37.50,\ s_1 = 9.13,\ n_2 = 56,$
$\bar{x}_2 = 39.63,\ s_2 = 10.19$

Analice estos datos y saque las conclusiones apropiadas.

d. Considerando el número de conmociones cerebrales de individuos que no juegan fútbol, los valores de la media $\pm$ d.e. para los tres grupos fueron 0.30 ± 0.67, 0.49 ± 0.87 y 0.19 ± 0.48. Analice estos datos y saque las conclusiones apropiadas.

48. ¿Los dígitos sucesivos del desarrollo decimal de π se comportan como si se seleccionaran de una tabla de números aleatorios (o provinieran del generador de números aleatorios de una computadora)?

a. Sea p_0 la proporción de largo plazo de los dígitos del desarrollo que son iguales a 0 y defina $p_1, \ldots, p_9$ de manera análoga. ¿Qué hipótesis acerca de estas proporciones se deben probar, y cuál es el número de gl para la prueba chi-cuadrada?

b. H_0 del inciso (a) no se rechazaría para la secuencia no aleatoria 012 . . . 901 . . . 901 . . . . Considere los grupos no superponibles de dos dígitos, y sea p_{ij} la proporción de largo plazo de grupos para los que el primer dígito es i y el segundo dígito es j. ¿Qué hipótesis acerca de estas proporciones se deben probar, y cuáles son los gl para la prueba chi-cuadrada?

c. Considere los grupos no superponibles de cinco dígitos. ¿Una prueba chi-cuadrada de las hipótesis apropiadas acerca de las p_{ijklm} se podrían basar en los primeros 100 000 dígitos? Explique.

d. En el artículo "Are the Digits of π an Independent and Identically Distributed Sequence?" (*The American Statistician*, 2000: 12-16) se consideraron los primeros 1 254 540 dígitos de π y se suministraron los siguientes valores de P para tamaños de grupo de $1, \ldots, 5$: 0.572, 0.078, 0.529, 0.691, 0.298. ¿Cuál sería su conclusión?

Bibliografía

Agresti, Alan, *An Introduction to Categorical Data Analysis*, Wiley, Nueva York, 1996. Un manejo excelente de varios aspectos del análisis de datos categóricos por uno de los investigadores más prominentes en esta área.

Everitt, B. S., *The Analysis of Contingency Tables* (2a. ed.), Halsted Press, Nueva York, 1992. Un estudio compacto pero ilustrativo de métodos para analizar datos categóricos, expuestos con un mínimo de matemáticas.

Mosteller, Frederick y Richard Rourke, *Sturdy Statistics*, Addison-Wesley, Reading, MA, 1973. Contiene varios capítulos muy amenos sobre los variados usos de chi-cuadrada.

15

Procedimientos libres de distribución

Introducción

Cuando una población o poblaciones subyacentes son no normales, las pruebas t y F y los intervalos de confianza de t de los capítulos 7 al 13 en general tendrán niveles reales de significación o niveles de confianza que difieren de los niveles nominales (los que prescribe el experimentador mediante la elección de, por ejemplo, $t_{0.025}$, $F_{0.01}$, etc.) α y $100(1 - \alpha)\%$, aunque es posible que la diferencia entre los niveles reales y nominales no sea grande cuando la desviación de la normalidad no es muy grave. Debido a que los procedimientos t y F requieren la suposición distribucional de normalidad, no son procedimientos "libres de distribución", otra posibilidad es que, como se basan en una familia paramétrica particular de distribuciones (normales), no son procedimientos "no paramétricos".

En este capítulo, se describen procedimientos que son válidos [nivel real α o nivel de confianza $100(1 - \alpha)\%$] de forma simultánea para muchos tipos de distribuciones subyacentes. Esta clase de procedimientos se llaman **libres de distribución** o **no paramétricos**. En la sección 15.1, se analiza un procedimiento de prueba para analizar una sola muestra de datos; en la sección 15.2 se presenta un procedimiento de prueba para uso en problemas de dos muestras. En la sección 15.3 se desarrollan intervalos de confianza libres de distribución para μ y $\mu_1 - \mu_2$. En la sección 15.4 se describen procedimientos ANOVA libres de distribución. Estos procedimientos son competidores de los procedimientos paramétricos (t y F) des-

> críticos en los capítulos anteriores, así que es importante comparar el desempeño de los dos tipos de procedimientos bajo modelos poblacionales normales y no normales. En términos generales, los procedimientos libres de distribución funcionan casi tan bien como sus contrapartes *t* y *F* en el "hábitat" de la distribución normal y con frecuencia producen una mejora considerable en condiciones no normales.

15.1 Prueba de Wilcoxon de los rangos con signo

Un químico investigador llevó a cabo un experimento particular un total de diez veces en condiciones idénticas, y obtuvo los siguientes valores ordenados de temperatura de reacción:

$$-0.57 \quad -0.19 \quad -0.05 \quad 0.76 \quad 1.30 \quad 2.02 \quad 2.17 \quad 2.46 \quad 2.68 \quad 3.02$$

La distribución de la temperatura de reacción es por supuesto continua. Suponga que el investigador está dispuesto a asumir que la distribución de la temperatura de reacción es simétrica; es decir, hay un punto de simetría tal, que la curva de densidad a la izquierda de ese punto es la imagen especular de la curva de densidad a su derecha. Este punto de simetría es la mediana de la distribución, como se ilustra en la figura 15.1 (y también es el valor medio μ siempre y cuando la media sea finita). La suposición de simetría en principio se podría considerar al parecer bastante obvia, pero recuerde que cualquier distribución normal es simétrica, de modo que la simetría es en realidad una suposición más débil que la normalidad.

Figura 15.1 Varias distribuciones simétricas

Ahora, considere probar la hipótesis nula de que la mediana de la distribución de la temperatura de reacción es cero; es decir, H_0: $\tilde{\mu} = 0$. Esto equivale a decir que una temperatura de determinada magnitud, por ejemplo, 1.50, no tiene mayores probabilidades de ser positiva (+1.50) que de ser negativa (−1.50). Un vistazo a los datos indica que esta hipótesis no es muy defendible; por ejemplo, la mediana muestral es 1.66, que es bastante más grande que la magnitud de alguna de las tres observaciones negativas.

En la figura 15.2 se muestran dos fdp simétricas distintas, una para la que H_0 es verdadera y otra para la cual H_a es verdadera. Cuando H_0 es verdadera, se espera que las magnitudes de las observaciones negativas en la muestra sean comparables con las magnitudes de las observaciones positivas. Sin embargo, si H_0 es "extremadamente" falsa como en la figura 15.2(b), entonces las observaciones de magnitud absoluta grande tienden a ser positivas en vez de negativas.

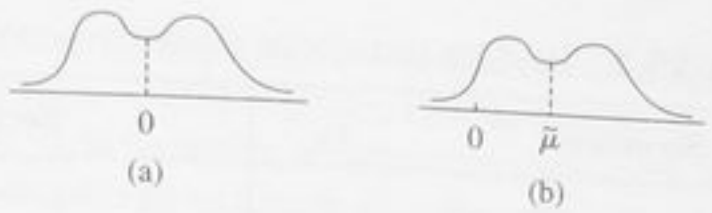


Figura 15.2 (a) $\tilde{\mu} = 0$; (b) $\tilde{\mu} > 0$

Para las diez observaciones de la muestra, suponga que se procede como sigue:

1. Sin considerar los signos de las observaciones, clasifique la colección de diez números en orden de magnitud creciente.
2. Utilice un estadístico de prueba S_+ = suma de los rangos relacionados con las observaciones positivas.

Magnitud absoluta	0.05	0.19	0.57	0.76	1.30	2.02	2.17	2.46	2.68	3.02
Rango	1	2	3	4	5	6	7	8	9	10
Rango con signo	−1	−2	−3	4	5	6	7	8	9	10

El valor observado de S_+ es $s_+ = 4 + 5 + 6 + 7 + 8 + 9 + 10 = 49$. La hipótesis nula H_0 ahora se rechaza en favor de H_a: $\tilde{\mu} > 0$ cuando s_+ es demasiado grande porque un valor grande de s_+ indica que la mayor parte de las observaciones con magnitud absoluta grande son positivas, lo que a su vez indica una mediana mayor que 0. Si por otro lado, las observaciones con signos negativos se combinan con observaciones positivas cuando se hace caso omiso de los signos, entonces la s_+ observada no será muy grande.

La distribución de S_+

Para decidir entre H_0 y H_a, se debe saber si 49 es un valor "sorprendentemente grande" de S_+ cuando la hipótesis nula es verdadera. Esto conlleva obtener la distribución de probabilidad de S_+ cuando H_0 es verdadera y seleccionar como la región de rechazo un conjunto de valores de s_+ grandes con probabilidad pequeña bajo H_0 (α pequeña). Si la s_+ observada está en la región de rechazo, entonces se puede rechazar H_0 al nivel α, deseado, de lo contrario no se rechaza H_0.

Considere el caso $n = 5$. La observación clave es que cuando H_0 es verdadera, *cualquier* colección de rangos con signos es *igual de probable* que cualquier otra colección. En otras palabras, la observación más pequeña en magnitud absoluta tiene las mismas probabilidades de ser positiva que negativa, la segunda observación más pequeña en magnitud absoluta tiene las mismas probabilidades de ser positiva que negativa, y de manera similar para las otras observaciones. Así, para $n = 5$, la secuencia −1, +2, +3, −4, +5 de rangos con signo tiene las mismas probabilidades de ser observada (cuando H_0 es verdadera) que la secuencia −1, −2, −3, −4, +5, o cualquier otra. En la tabla 15.1 se listan las posibles secuencias de rangos con signo para $n = 5$ y el valor s_+ relacionado con cada secuencia. Puesto que hay 32 secuencias de esta clase (dos signos posibles para la observación más pequeña, dos para la segunda más pequeña, etcétera, o $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 32$ secuencias), cada secuencia tiene probabilidad 1/32 de ser observada cuando H_0 es verdadera. A partir de esto, se puede obtener de inmediato la fdp de S_+ bajo H_0, como

Tabla 15.1 Posibles secuencias de rangos con signos para $n = 5$

Secuencia					s_+	Secuencia					s_+
−1	−2	−3	−4	−5	0	−1	−2	−3	+4	−5	4
+1	−2	−3	−4	−5	1	+1	−2	−3	+4	−5	5
−1	+2	−3	−4	−5	2	−1	+2	−3	+4	−5	6
−1	−2	+3	−4	−5	3	−1	−2	+3	+4	−5	7
+1	+2	−3	−4	−5	3	+1	+2	−3	+4	−5	7
+1	−2	+3	−4	−5	4	+1	−2	+3	+4	−5	8
−1	+2	+3	−4	−5	5	−1	+2	+3	+4	−5	9
+1	+2	+3	−4	−5	6	+1	+2	+3	+4	−5	10
−1	−2	−3	−4	+5	5	−1	−2	−3	+4	+5	9
+1	−2	−3	−4	+5	6	+1	−2	−3	+4	+5	10
−1	+2	−3	−4	+5	7	−1	+2	−3	+4	+5	11
−1	−2	+3	−4	+5	8	−1	−2	+3	+4	+5	12
+1	+2	−3	−4	+5	8	+1	+2	−3	+4	+5	12
+1	−2	+3	−4	+5	9	+1	−2	+3	+4	+5	13
−1	+2	+3	−4	+5	10	−1	+2	+3	+4	+5	14
+1	+2	+3	−4	+5	11	+1	+2	+3	+4	+5	15

en la tabla 15.2. Observe que para $n = 10$ observaciones hay $2^{10} = 1024$ secuencias posibles de rangos con signo, y sería muy tedioso enumerar estas secuencias. Sin embargo, cada secuencia recibiría una probabilidad de 1/1024 bajo H_0, de la cual se puede obtener con facilidad la distribución nula (distribución cuando H_0 es verdadera) de S_+.

Tabla 15.2 Distribución nula de S_+ cuando $n = 5$

s_+	0	1	2	3	4	5	6	7
$p(s_+)$	1/32	1/32	1/32	2/32	2/32	3/32	3/32	3/32
s_+	8	9	10	11	12	13	14	15
$p(s_+)$	3/32	3/32	3/32	2/32	2/32	1/32	1/32	1/32

Como un ejemplo de la forma en que se obtuvieron las probabilidades de la tabla 15.2, si fija la atención en el valor 7 de S_+ hay tres secuencias de rangos con signo para las cuales $S_+ = 7$, y cada una recibe probabilidad 1/32 bajo H_0, así que $P(S_+ = 7$ cuando H_0 es verdadera) = 3/32.

La tabla 15.2 ahora se puede usar para especificar una región de rechazo para H_0: $\tilde{\mu} = 0$ contra H_a: $\tilde{\mu} > 0$ para la cual se puede calcular α. Considere la región de rechazo $R = \{13, 14, 15\}$. Entonces

$$\begin{aligned}\alpha &= P(\text{se rechaza } H_0 \text{ cuando } H_0 \text{ es verdadera})\\ &= P(S_+ = 13, 14 \text{ o } 15 \text{ cuando } H_0 \text{ es verdadera})\\ &= 1/32 + 1/32 + 1/32 = 3/32\\ &= 0.094\end{aligned}$$

de manera que $R = \{13, 14, 15\}$ especifica una prueba con nivel aproximado 0.1. Para la región de rechazo $\{14, 15\}$, $\alpha = 2/32 = 0.063$. Para la muestra $x_1 = 0.58$, $x_2 = 2.50$, $x_3 = -0.21$, $x_4 = 1.23$, $x_5 = 0.97$, la secuencia de rangos con signo es $-1, +2, +3, +4, +5$, de modo que $s_+ = 14$ y al nivel 0.063 se rechazaría H_0.

Descripción general de la prueba de Wilcoxon de los rangos con signo

Debido a que se supone que la distribución subyacente es simétrica, $\mu = \tilde{\mu}$, así que se expresa la hipótesis de interés en términos de μ en vez de $\tilde{\mu}$.*

SUPOSICIÓN

$X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución de probabilidad continua y simétrica con media (y mediana) μ.

Cuando el valor supuesto de μ es μ_0, las diferencias absolutas $|x_1 - \mu_0|, \ldots, |x_n - \mu_0|$ se deben clasificar de menor a mayor.

Hipótesis nula: H_0: $\mu = \mu_0$

Valor del estadístico de prueba: s_+ = la suma de los rangos relacionados con $(x_i - \mu_0)$ positivas

Hipótesis alternativa	**Región de rechazo para la prueba de nivel α**
H_a: $\mu > \mu_0$	$s_+ \geq c_1$
H_a: $\mu < \mu_0$	$s_+ \leq c_2$ [donde $c_2 = n(n+1)/2 - c_1$]
H_a: $\mu \neq \mu_0$	ya sea $s_+ \geq c$ o $s_+ \leq n(n+1)/2 - c$

donde los valores críticos c_1 y c que se obtienen de la tabla A.13 del apéndice satisfacen $P(S_+ \geq c_1) \approx \alpha$ y $P(S_+ \geq c) \approx \alpha/2$ cuando H_0 es verdadera.

Ejemplo 15.1 Un fabricante de planchas eléctricas, que desea probar la precisión del termostato en la posición de 500°F, instruye a un ingeniero a probar las temperaturas reales en ese valor para 15 planchas que emplean un termostato. Las mediciones resultantes se muestran a continuación:

494.6 510.8 487.5 493.2 502.6 485.0 495.9 498.2
501.6 497.3 492.0 504.3 499.2 493.5 505.8

El ingeniero cree que es razonable suponer que una desviación de temperatura de 500° de alguna magnitud particular tiene las mismas probabilidades de ser positiva que negativa (la su-

*Si los extremos de la distribución son "demasiado densos", como en el caso de la distribución de Cauchy mencionada en el capítulo 6, entonces no existe μ. En esos casos, la prueba de Wilcoxon aún es válida para pruebas relacionadas con $\tilde{\mu}$.

posición de simetría), pero quiere proteger contra la posible no normalidad de la distribución de temperaturas reales, así que decide usar la prueba de Wilcoxon de los rangos con signo para ver si los datos indican una notable calibración incorrecta de la plancha.

Las hipótesis son H_0: $\mu = 500$ contra H_a: $\mu \neq 500$, donde μ = la temperatura real promedio verdadera en la posición de 500°F. Al restar 500 de cada x_i se obtiene

$$-5.6 \quad 10.8 \quad -12.5 \quad -6.8 \quad 2.6 \quad -15.0 \quad -4.1 \quad -1.8 \quad 1.6 \quad -2.7$$
$$-8.0 \quad 4.3 \quad -0.8 \quad -6.5 \quad 5.8$$

Los rangos se obtienen al ordenar estos valores de menor a mayor sin considerar el signo.

Magnitud absoluta	0.8	1.6	1.8	2.6	2.7	4.1	4.3	5.6	5.8	6.5	6.8	8.0	10.8	12.5	15.0
Rango	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Signo	−	+	−	+	−	−	+	−	+	−	−	−	+	−	−

Así, $s_+ = 2 + 4 + 7 + 9 + 13 = 35$. De la tabla A.13 del apéndice, $P(S_+ \geq 95) = P(S_+ \leq 25) = 0.024$ cuando H_0 es verdadera, de modo que la prueba de dos extremos con nivel aproximado 0.05 rechaza H_0 cuando $s_+ \geq 95$ o ≤ 25 [el valor exacto de α es $2(0.024) = 0.048$]. Puesto que $s_+ = 35$ está fuera de la región de rechazo, no se puede concluir al nivel 0.05 que μ es cualquier otro valor distinto de 500. Incluso al nivel 0.094 (casi 0.1), no se rechaza H_0, ya que $P(S_+ \leq 30) = 0.047$ indica que los valores de s_+ entre 30 y 90 son insignificantes a ese nivel. Por lo tanto, el valor P de los datos es mayor que 0.1. ■

Observaciones por pares

Cuando los datos consistieron en pares $(X_1, Y_1), \ldots, (X_n, Y_n)$ y la distribución de las diferencias $D_1 = X_1 - Y_1, \ldots, D_n = X_n - Y_n$ era normal, en el capítulo 9 se empleó una prueba t de pares para probar hipótesis acerca de la diferencia esperada μ_D. Si no se supone normalidad, las hipótesis acerca de μ_D se prueban por medio de la prueba de Wilcoxon de los rangos con signo en las D_i siempre que la distribución de las diferencias sea continua y simétrica. *Si X_i y Y_i tienen distribuciones continuas que difieren sólo con respecto a sus medias* (de modo que la distribución de Y es la distribución de X desplazada $\mu_1 - \mu_2 = \mu_D$), *entonces D_i tiene una distribución simétrica continua* (*no* es necesario que cada una de las distribuciones X y Y sean simétricas). La hipótesis nula es H_0: $\mu_D = \Delta_0$, y el estadístico de prueba S_+ es la suma de los rangos relacionados con las $(D_i - \Delta_0)$ positivas.

Ejemplo 15.2 Un experimento en el que se comparan las capacidades de dos disolventes para extraer creosol impregnado en troncos de prueba, requirió usar ocho troncos distintos. Después de dividir cada tronco en dos segmentos, se eligió al azar uno para la aplicación del primer disolvente, con el segundo disolvente utilizado en el otro segmento. Los resultados fueron los siguientes:

Tronco	1	2	3	4	5	6	7	8
Disolvente 1	3.92	3.79	3.70	4.08	3.87	3.95	3.55	3.76
Disolvente 2	4.25	4.20	4.41	3.89	4.39	3.75	4.20	3.90
Diferencia	−0.33	−0.41	−0.71	0.19	−0.52	0.20	−0.65	−0.14
Rangos con signo	−4	−5	−8	2	−6	3	−7	−1

En la actualidad se emplea el primer disolvente, y el segundo es una nueva formulación diseñada para dar como resultado una capacidad de extracción mejorada. ¿Estos datos indican que la cantidad promedio real extraída con el segundo disolvente es mayor que la del primero? Las hipótesis pertinentes son H_0: $\mu_D = 0$ contra H_a: $\mu_D < 0$. En la tabla A.13 del apéndice se observa que para una prueba con nivel de significación más o menos 0.05, la hipótesis nula se debe rechazar si $s_+ \leq (8)(9)/2 - 30 = 6$. El valor del estadístico de prueba es $2 + 3 = 5$, que cae en la región de rechazo. Por lo tanto, se rechaza H_0 al nivel de significación 0.05 en favor de la conclusión de que el nuevo disolvente no supera al que se usa actualmente. El siguiente resultado de MINITAB da el valor del estadístico de prueba y también el valor P correspondiente, que es $P(S_+ \leq 5$ cuando H_0 es verdadera).

```
Prueba de mediana = 0.000000 contra mediana < 0.000000

                N para      Estadístico                    Media
        N       prueba      de Wilcoxon          P       estimada
Dif.    8            8              5.0      0.040        -0.3025
```

■

Empates en la prueba de Wilcoxon de rangos con signo

Si bien una implicación teórica de la continuidad de la distribución subyacente es que no ocurrirán empates, en la práctica ocurren con frecuencia, debido a la naturaleza discreta de los instrumentos de medición. Si hay varios valores de datos con la misma magnitud absoluta, entonces se les asigna el promedio de los rangos que recibirían si difirieran muy poco entre sí. Como ilustración, si en el ejemplo 15.1 $x_8 = 498.2$ se cambia a 498.4, entonces dos valores de $(x_i - 500)$ diferentes tendrían magnitud absoluta 1.6. Los rangos por promediar serían 2 y 3, así que a cada uno se le asignaría el rango 2.5.

Aproximación de una gran muestra

En la tabla A.13 del apéndice, se encuentran valores críticos para las pruebas de nivel α sólo cuando $n \leq 20$. Para $n > 20$, se puede demostrar que S_+ tiene una distribución normal aproximada con

$$\mu_{S_+} = \frac{n(n+1)}{4} \qquad \sigma^2_{S_+} = \frac{n(n+1)(2n+1)}{24}$$

cuando H_0 es verdadera.

La media y la varianza resultan de observar que cuando H_0 es verdadera (la distribución simétrica se centra en μ_0), entonces el rango i tiene las mismas probabilidades de recibir un signo + que un signo −. Así

$$S_+ = W_1 + W_2 + W_3 + \cdots + W_n$$

donde

$$W_1 = \begin{cases} 1 & \text{con probabilidad } 0.5 \\ 0 & \text{con probabilidad } 0.5 \end{cases} \quad \cdots \quad W_n = \begin{cases} n & \text{con probabilidad } 0.5 \\ 0 & \text{con probabilidad } 0.5 \end{cases}$$

($W_i = 0$ es equivalente a clasificar i que se relaciona con un $-$, de modo que i no contribuye con S_+.)

Entonces S_+ es una suma de variables aleatorias, y cuando H_0 es verdadera, se puede demostrar que estas W_i son independientes. La aplicación de las reglas de valor esperado y varianza producen la media y la varianza de S_+. Debido a que las W_i no se distribuyen de manera idéntica es inaplicable el teorema del límite central, pero hay una versión más general del teorema que se emplea para justificar la conclusión de normalidad.

El estadístico de prueba con muestras grandes ahora se determina mediante

$$Z = \frac{S_+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \tag{15.1}$$

Para las tres alternativas estándar, los valores críticos para las pruebas de nivel α son los valores normales estándar comunes z_α, $-z_\alpha$ y $\pm z_{\alpha/2}$.

Ejemplo 15.3 Se diseñó un determinado tipo de viga de acero para tener una resistencia compresiva (lb/pulg^2) de por lo menos 50 000. Para cada viga en una muestra de 25, se determina la resistencia compresiva y se escribe en la tabla 15.3. Suponiendo que la resistencia compresiva real está distribuida de forma simétrica respecto al valor promedio real, utilice la prueba de Wilcoxon para decidir si la resistencia compresiva promedio es menor que el valor especificado. Es decir, pruebe H_0: $\mu = 50\,000$ contra H_a: $\mu < 50\,000$ (lo cual favorece la afirmación de que la resistencia compresiva promedio es por lo menos 50 000).

Tabla 15.3 Datos para el ejemplo 15.3

$x_i - 50\,000$	Rango con signo	$x_i - 50\,000$	Rango con signo	$x_i - 50\,000$	Rango con signo
−10	−1	−99	−10	165	+18
−27	−2	113	+11	−178	−19
36	+3	−127	−12	−183	−20
−55	−4	−129	−13	−192	−21
73	+5	136	+14	−199	−22
−77	−6	−150	−15	−212	−23
−81	−7	−155	−16	−217	−24
90	+8	−159	−17	−229	−25
−95	−9				

La suma de los rangos con signo positivo es $3 + 5 + 8 + 11 + 14 + 18 = 59$, $n(n+1)/4 = 162.5$ y $n(n+1)(2n+1)/24 = 1381.25$, por lo tanto

$$z = \frac{59 - 162.5}{\sqrt{1381.25}} = -2.78$$

La prueba de extremo inferior al nivel 0.01 rechaza H_0 si $z \leq -2.33$. Puesto que $-2.78 \leq -2.33$, H_0 se rechaza en favor de la conclusión de que la resistencia compresiva es menor que 50 000. ■

Cuando hay empates en las magnitudes absolutas, de modo que se deben usar rangos promedio, aún es correcto estandarizar S_+ al restar $n(n + 1)/4$, pero se debe usar la siguiente fórmula corregida para la varianza:

$$\sigma^2_{S_+} = \frac{1}{24}n(n + 1)(2n + 1) - \frac{1}{48}\sum(\tau_i - 1)(\tau_i)(\tau_i + 1) \tag{15.2}$$

donde τ_i es el número de empates en el i-ésimo conjunto de valores empatados. Si, por ejemplo, $n = 10$ y los rangos con signo son 1, 2, −4, −4, 4, 6, 7, 8.5, 8.5 y 10, entonces hay dos conjuntos empatados con $\tau_1 = 3$ y $\tau_2 = 2$, de manera que la suma es $(2)(3)(4) + (1)(2)(3) = 30$ y $\sigma^2_{S_+} = 96.25 - 30/48 = 95.62$. El denominador en (15.1) se debe reemplazar por la raíz cuadrada de (15.2), aunque como se muestra en este ejemplo, por lo común la corrección es insignificante.

Eficiencia de la prueba de Wilcoxon de los rangos con signo

Cuando la distribución subyacente muestreada es normal, la prueba t o la prueba de rangos con signo se emplea para probar una hipótesis respecto a μ. La prueba t es la mejor en esta situación porque entre las pruebas de nivel α es la que tiene un valor mínimo de β. Puesto que en general se acepta que hay muchas situaciones experimentales en las que de manera razonable se puede asumir normalidad, así como algunas en las que esto no debe ser, hay dos preguntas a las que hay que prestar atención en un intento por comparar las dos pruebas:

1. Cuando la distribución subyacente es normal (el "entorno" de la prueba t), ¿cuánto se pierde por usar la prueba de los rangos con signo?
2. Cuando la distribución subyacente no es normal, ¿se puede lograr una mejora importante al usar la prueba de rangos con signo?

Si la prueba de Wilcoxon no empeora mucho con respecto a la prueba t en el "entorno" de esta última, y funciona mucho mejor que la prueba t para un gran número de otras distribuciones, entonces se justifica el uso de la prueba de Wilcoxon.

Infortunadamente, no hay respuestas simples a las dos preguntas. Si se reflexiona al respecto, no es sorprendente que la prueba t funcione mal cuando la distribución subyacente tiene extremos densos (es decir, cuando los valores observados que están lejos de μ son relativamente más probables que cuando la distribución es normal). Esto se debe a que el comportamiento de la prueba t depende de la media muestral, que puede ser muy estable en presencia de extremos densos. La dificultad para dar respuestas a las dos preguntas es que β para la prueba de Wilcoxon es muy difícil de obtener y estudiar para *cualquier* distribución subyacente, y lo mismo se puede decir para la prueba t cuando la distribución no es normal. Aun cuando fuera muy fácil obtener β cualquier medida de eficiencia dependería de la distribución subyacente que se postulara. Los estadísticos han propuesto varias medidas de eficiencia; una que muchos estadísticos consideran creíble se llama **eficiencia relativa asintótica** (ERA). La ERA de una prueba con respecto a otra es en esencia la relación limitante de los tamaños de muestra necesarios para obtener probabilidades de error idénticas para las dos muestras. Así, la ERA de una prueba con respecto a una segunda es igual a 0.5, entonces cuando los tamaños de muestra son gran-

des, se requiere un tamaño de muestra del doble de la primera prueba para que funcione tan bien como la segunda. Si bien las ERA no caracterizan el desempeño de la prueba para tamaños de muestra pequeños, se puede demostrar que se cumplen los resultados siguientes:

1. Cuando la distribución subyacente es normal, la ERA de la prueba de Wilcoxon con respecto a la prueba t es de alrededor de 0.95.
2. Para cualquier distribución, la ERA será por lo menos 0.86 y para muchas distribuciones será mayor que 1.

Estos resultados se resumen diciendo que, en problemas con muestras grandes, la prueba de Wilcoxon nunca es mucho menos eficiente que la prueba t, y podría ser mucho más eficiente si la distribución subyacente dista de ser normal. Aunque el asunto está lejos de resolverse en el caso de tamaños de muestra obtenidos en la mayor parte de problemas prácticos, los estudios han mostrado que la prueba de Wilcoxon funciona de forma razonable y, por consiguiente, es una alternativa viable para la prueba t.

Ejercicios | Sección 15.1 (1 a 9)

1. Reconsidere la situación descrita en el ejercicio 32 de la sección 8.2 y utilice la prueba de Wilcoxon con $\alpha = 0.05$ para probar las hipótesis pertinentes.
2. Utilice la prueba de Wilcoxon para analizar los datos del ejemplo 8.9.
3. Los datos siguientes son un subconjunto de los datos que aparecen en el artículo "Synovial Fluid pH, Lactate, Oxygen and Carbon Dioxide Partial Pressure in Various Joint Diseases" (*Arthritis and Rheumatism,* 1971: 476-477). Las observaciones son valores de pH de líquido sinovial (el cual lubrica articulaciones y tendones) tomado de las rodillas de personas que padecen artritis. Si se supone que el pH promedio real para personas no artríticas es 7.39, pruebe al nivel 0.5 para ver si los datos indican una diferencia entre los valores de pH promedio para individuos con artritis y sin artritis.

   7.02 7.35 7.34 7.17 7.28 7.77 7.09

   7.22 7.45 6.95 7.40 7.10 7.32 7.14

4. Se seleccionó una muestra aleatoria de 15 mecánicos automotrices para trabajar en cierto tipo de automóvil, y se determinó el tiempo en minutos necesario para que cada uno diagnosticara un problema particular. Los datos que se obtuvieron son los siguientes:

   30.6 30.1 15.6 26.7 27.1 25.4 35.0 30.8

   31.9 53.2 12.5 23.2 8.8 24.9 30.2

   Utilice la prueba de Wilcoxon al nivel de significación 0.10 para decidir si los datos indican que el tiempo promedio real de diagnóstico es menor que 30 minutos.

5. Se someten a consideración tanto el método gravimétrico como el espectrofotométrico para determinar el contenido de fosfato de cierto material. Se obtienen 12 muestras del material, cada una se divide a la mitad y se hace una determinación en cada mitad mediante uno de los dos métodos. Los resultados se muestran a continuación:

Muestra	1	2	3	4
Gravimétrico	54.7	58.5	66.8	46.1
Espectrofotométrico	55.0	55.7	62.9	45.5

Muestra	5	6	7	8
Gravimétrico	52.3	74.3	92.5	40.2
Espectrofotométrico	51.1	75.4	89.6	38.4

Muestra	9	10	11	12
Gravimétrico	87.3	74.8	63.2	68.5
Espectrofotométrico	86.8	72.5	62.3	66.0

   Utilice la prueba de Wilcoxon para determinar si con una de las técnicas se obtiene en promedio un valor diferente que el obtenido con la otra técnica para este tipo de material.

6. Reconsidere la situación descrita en el ejercicio 41 de la sección 9.3, y utilice la prueba de Wilcoxon para probar las hipótesis adecuadas.
7. Utilice la versión de muestras grandes de la prueba de Wilcoxon al nivel de significación 0.05 en los datos del

ejercicio 37 de la sección 9.3 para decidir si la diferencia promedio real entre las concentraciones exterior e interior es mayor que 0.20.

8. Las siguientes 25 observaciones de resistencia a la fractura de placa base de acero maraginizado con 18% de níquel aparecen en el artículo "Fracture Testing of Weldments" (*ASTM Special Publ. No. 381*, 1965: 328-356). Suponga que una empresa conviene en comprar este acero para cierta aplicación sólo si se demuestra con claridad a partir de la evidencia experimental que la resistencia promedio real pasa de 75. Si se supone que la distribución de resistencia a la fractura es simétrica, exprese y pruebe las hipótesis pertinentes al nivel 0.05 y calcule el valor P.

69.5 71.9 72.6 73.1 73.3 73.5 74.1 74.2 75.3
75.5 75.7 75.8 76.1 76.2 76.2 76.9 77.0 77.9
78.1 79.6 79.7 80.1 82.2 83.7 93.7

9. Suponga que las observaciones $X_1, X_2, \ldots, X_n$ se realizan en un proceso en los tiempos $1, 2, \ldots, n$. Con base en estos datos se desea probar

H_0: las X_i constituyen una secuencia distribuida de manera idéntica e independiente

contra

H_a: X_{i+1} tiende a ser más grande que X_i para $i = 1, \ldots, n$ (una tendencia creciente)

Suponga que las X_i se clasifican de 1 a n. Entonces, cuando H_a es verdadera, los rangos más grandes tienden a presentarse al final de la secuencia, mientras que si H_0 es verdadera, los rangos grandes y pequeños tienden a mezclarse. Sea R_i el rango de X_i y considere el estadístico de prueba $D = \sum_{i=1}^{n} (R_i - i)^2$. Entonces los valores pequeños de D respaldan a H_a (p. ej., el valor mínimo es 0 para $R_1 = 1, R_2 = 2, \ldots, R_n = n$), así que se debe rechazar H_0 en favor de H_a si $d \leq c$. Cuando H_0 es verdadera, cualquier secuencia de rangos tiene probabilidad $1/n!$ Utilice esto a fin de determinar el valor de c para el cual la prueba tiene un nivel lo más cercano posible a 0.10 en el caso $n = 4$. [*Sugerencia:* enumere las 4! secuencias de rangos, calcule d para cada una y después obtenga la distribución nula de D. Para más información, consulte el libro de Lehmann, p. 290 (en la bibliografía de este capítulo).]

15.2 Prueba de Wilcoxon de la suma de rangos

Cuando en un problema con dos muestras por lo menos uno de los tamaños de muestra es pequeño, la prueba t requiere la suposición de normalidad (al menos en forma aproximada). Sin embargo, hay situaciones en las que un investigador desearía usar una prueba que es válida incluso si las distribuciones subyacentes son bastante normales. Ahora se describe una prueba de este tipo llamada **prueba de Wilcoxon de la suma de rangos.** Otro nombre para el procedimiento es prueba de Mann-Whitney, aunque el estadístico de prueba de Mann-Whitney a veces se expresa de una forma un poco distinta a la de la prueba de Wilcoxon. El procedimiento de la prueba de Wilcoxon está libre de distribución porque tendrá el nivel deseado de significación para una clase muy grande de distribuciones subyacentes.

SUPOSICIONES

$X_1, \ldots, X_m$ y $Y_1, \ldots, Y_n$ son dos muestras aleatorias independientes de distribuciones continuas con medias μ_1 y μ_2, respectivamente. Las distribuciones X y Y tienen la misma forma y dispersión, la única diferencia entre las dos radica en los valores de μ_1 y μ_2.

Cuando H_0: $\mu_1 - \mu_2 = \Delta_0$ es verdadera, la distribución X se desplaza por la cantidad Δ_0 a la derecha de la distribución Y, mientras que cuando H_0 es falsa, el desplazamiento es por una cantidad distinta de Δ_0.

Desarrollo de la prueba cuando $m = 3$, $n = 4$

Primero considere probar H_0: $\mu_1 - \mu_2 = 0$. Si μ_1 es en realidad mucho mayor que μ_2, entonces la mayor parte de las x observadas cae a la derecha de las y observadas. Sin embargo, si H_0 es verdadera, entonces deben combinarse los valores observados de las dos muestras. El estadístico de prueba proporciona una cuantificación de cuánta combinación hay en las dos muestras.

Considere el caso $m = 3$, $n = 4$. Entonces si las tres x observadas estuvieran a la derecha de las cuatro y observadas, esto proporcionaría evidencia convincente para rechazar H_0 en favor de H_a: $\mu_1 - \mu_2 \neq 0$, donde una conclusión similar es apropiada si las tres x caen abajo de las cuatro y. Suponga que se agrupan las X y Y en una muestra combinada de tamaño $m + n = 7$ y se clasifican estas observaciones de menor a mayor, donde la menor recibe la posición 1 y la mayor la posición 7. Si la mayor parte de las posiciones o rangos mayores o la mayor parte de las posiciones menores se relacionara con las X observaciones, se comenzaría a dudar de H_0. Esto indica que el estadístico de prueba

$$W = \text{suma de rangos en la muestra combinada relacionados con } X \text{ observaciones} \tag{15.3}$$

Para los valores de m y n en consideración, el menor valor posible de W es $w = 1 + 2 + 3 = 6$ (si las tres x son más pequeñas que las cuatro y), y el mayor valor posible de $w = 5 + 6 + 7 = 18$ (si las tres x son mayores que las cuatro y).

Como una ilustración, suponga que $x_1 = -3.10$, $x_2 = 1.67$, $x_3 = 2.01$, $y_1 = 5.27$, $y_2 = 1.89$, $y_3 = 3.86$ y $y_4 = 0.19$. Entonces la muestra ordenada agrupada es −3.10, 0.19, 1.67, 1.89, 2.01, 3.86 y 5.27. Los rangos de X para este ejemplo son 1 (para −3.10), 3 (para 1.67) y 5 (para 2.01), así que el valor calculado de W es $w = 1 + 3 + 5 = 9$.

El procedimiento de prueba basado en el estadístico (15.3) es rechazar H_0 si el valor calculado de w es "demasiado extremo", es decir, $\geq c$ para una prueba de extremo superior, $\leq c$ para una prueba de extremo inferior y $\geq c_1$ o $\leq c_2$ para una prueba de dos extremos. La(s) constante(s) crítica(s) c (c_1, c_2) se deben elegir de modo que la prueba tenga el nivel deseado de significación α. Para ver cómo se debe hacer esto, recuerde que cuando H_0 es verdadera, las siete observaciones provienen de la misma población. Esto significa que bajo H_0, cualquier terna posible de rangos relacionados con las x, como (1, 4, 5), (3, 5, 6) o (5, 6, 7) tiene la misma probabilidad que cualquier otra terna de rangos posible. Puesto que hay $\binom{7}{3} = 35$ ternas posibles de rangos, bajo H_0 cada terna de rangos tiene probabilidad $\frac{1}{35}$. De una lista de las 35 ternas de rangos y el valor de w relacionados con cada una, la distribución de probabilidad de W se determina de inmediato. Por ejemplo, hay cuatro rangos posibles que tienen valor 11 de w, (1, 3, 7), (1, 4, 6), (2, 3, 6) y (2, 4, 5), así que $P(W = 11) = \frac{4}{35}$. El resumen del listado y los cálculos se ilustran en la tabla 15.4.

Tabla 15.4 Distribución de probabilidad de W ($m = 3$, $n = 4$) cuando H_0 es verdadera

w	6	7	8	9	10	11	12	13	14	15	16	17	18
$P(W = w)$	$\frac{1}{35}$	$\frac{1}{35}$	$\frac{2}{35}$	$\frac{3}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{5}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{3}{35}$	$\frac{2}{35}$	$\frac{1}{35}$	$\frac{1}{35}$

La distribución de la tabla 15.4 es simétrica respecto al valor $w = (6 + 18)/2 = 12$, que es el valor medio de la lista ordenada de posibles valores W. Esto se debe a que las dos ternas de rangos (r, s, t) (con $r < s < t$) y $(8 - t, 8 - s, 8 - r)$ tienen valores de w simétricos res-

pecto a 12, de modo que para cada terna con valor w abajo de 12, hay una terna con valor w arriba de 12 por la misma cantidad.

Si la hipótesis alternativa es H_a: $\mu_1 - \mu_2 > 0$, entonces H_0 se debe rechazar en favor de H_a para valores grandes de W. Al elegir como la región de rechazo el conjunto de valores {17, 18}, $\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 \text{ cuando } H_0 \text{ es verdadera}) = P(W = 17 \text{ o } 18 \text{ cuando } H_0 \text{ es verdadera}) = \frac{1}{35} + \frac{1}{35} = \frac{2}{35} = 0.057$; así, la región {17, 18} especifica una prueba con nivel de significación de alrededor de 0.05. De forma similar, la región {6, 7}, que es apropiada para H_a: $\mu_1 - \mu_2 < 0$, tiene $\alpha = 0.057 \approx 0.05$. La región {6, 7, 17, 18}, que es apropiada para la segunda alternativa, tiene $\alpha = \frac{4}{35} = 0.114$. El valor de W para los datos mencionados varios párrafos antes fue $w = 9$, que es bastante cercano al valor medio 12, así que no se debe rechazar H_0 a ningún nivel razonable α para alguna de las tres H_a.

Descripción general de la prueba de Wilcoxon de la suma de rangos

La hipótesis nula H_0: $\mu_1 - \mu_2 = \Delta_0$ se maneja restando Δ_0 de cada X_i y usando las $(X_i - \Delta_0)$ como se usaron antes las X_i. Recordando que para cualquier entero positivo K, la suma de los primeros K enteros es $K(K + 1)/2$, el menor valor posible del estadístico W es $m(m + 1)/2$, que ocurre cuando las $(X_i - \Delta_0)$ se ubican a la izquierda de la muestra Y. El máximo valor posible de W ocurre cuando las $(X_i - \Delta_0)$ se ubican por completo a la derecha de las Y, en este caso, $W = (n + 1) + \cdots + (m + n) = (\text{suma de los primeros } m + n \text{ enteros}) - (\text{suma de los } n \text{ primeros enteros})$, con lo que se obtiene $m(m + 2n + 1)/2$. Al igual que con el caso $m = 3$, $n = 4$, la distribución de W es simétrica respecto al valor situado a la mitad entre los valores mínimo y máximo; este valor medio es $m(m + n + 1)/2$. Como resultado de esta simetría, las probabilidades relacionadas con valores críticos de extremo inferior se obtienen de los valores de extremo superior correspondientes.

Hipótesis nula: H_0: $\mu_1 - \mu_2 = \Delta_0$

Valor del estadístico de prueba: $w = \sum_{i=1}^{m} r_i$ donde r_i = rango de $(x_i - \Delta_0)$ en la muestra combinada de las $m + n$ $(x - \Delta_0)$ y las y

Hipótesis alternativa	Región de rechazo
H_a: $\mu_1 - \mu_2 > \Delta_0$	$w \geq c_1$
H_a: $\mu_1 - \mu_2 < \Delta_0$	$w \leq m(m + n + 1) - c_1$
H_a: $\mu_1 - \mu_2 \neq \Delta_0$	ya sea $w \geq c$ o bien, $w \leq m(m + n + 1) - c$

donde $P(W \geq c_1 \text{ cuando } H_0 \text{ es verdadera}) \approx \alpha$, $P(W \geq c \text{ cuando } H_0 \text{ es verdadera}) \approx \alpha/2$.

Debido a que W tiene una distribución de probabilidad discreta, no siempre existe un valor crítico que corresponde exactamente a uno de los niveles comunes de significación. En la tabla A.14 del apéndice aparecen los valores críticos para probabilidades cercanas a 0.05, 0.025, 0.01 y 0.005, a partir de los cuales se obtienen las pruebas de uno y dos extremos al nivel 0.05 o 0.01. En la tabla también se encuentra información sólo para $m = 3, 4, \ldots, 8$ y

$n = m, m + 1, \ldots, 8$ (es decir, $3 \leq m \leq n \leq 8$). Para valores de m y n que pasan de 8, se puede usar una aproximación normal. Sin embargo, para usar la tabla para m y n, pequeños, *las muestras X y Y se deben marcar de modo que* $m \leq n$.

Ejemplo 15.4 La concentración de fluoruro en la orina (partes por millón) se midió para una muestra de ganado que pace en una área expuesta antes a contaminación de fluoruro y para una muestra similar que pace en una región no contaminada:

Contaminada	21.3	18.7	23.0	17.1	16.8	20.9	19.7
No contaminada	14.2	18.3	17.2	18.4	20.0		

¿Los datos indican sin lugar a dudas que la concentración promedio real de fluoruro para ganado que pace en la región contaminada es mayor que para la región no contaminada? Utilice la prueba de Wilcoxon de la suma de rangos al nivel $\alpha = 0.01$.

Los tamaños de muestra en este caso son 7 y 5. Para obtener $m \leq n$, marque las observaciones de no contaminación como las x ($x_1 = 14.2, \ldots, x_5 = 20.0$) y las observaciones de contaminación como las y. Así, μ_1 es la concentración promedio real de fluoruro sin contaminación y μ_2 es la concentración promedio real con contaminación. La hipótesis alternativa es H_a: $\mu_1 - \mu_2 < 0$ (la contaminación causa un incremento en la concentración), por lo tanto es apropiada una prueba de extremo inferior. De la tabla A.14 del apéndice con $m = 5$ y $n = 7$, $P(W \geq 47$ cuando H_0 es verdadera) ≈ 0.01. Por consiguiente, el valor crítico para la prueba de extremo inferior es $m(m + n + 1) - 47 = 5(13) - 47 = 18$; ahora H_0 se rechaza si $w \leq 18$. Se deduce la muestra ordenada agrupada; la W calculada es $w = r_1 + r_2 + \cdots + r_5$ (donde r_i es el rango de x_i) $= 1 + 5 + 4 + 6 + 9 = 25$. Puesto que 25 no es ≤ 18, no se rechaza H_0 al nivel (aproximadamente) 0.01.

x	y	y	x	x	x	y	y	x	y	y	y
14.2	16.8	17.1	17.2	18.3	18.4	18.7	19.7	20.0	20.9	21.3	23.0
1	2	3	4	5	6	7	8	9	10	11	12

■

Una aproximación normal para *W*

Cuando m y n son mayores que 8, la distribución de W se aproxima mediante una curva normal apropiada, y esta aproximación se usa en lugar de la tabla A.14 del apéndice. Para obtener la aproximación, se requiere μ_W y σ_W^2 cuando H_0 es verdadera. En este caso, el rango R_i de $X_i - \Delta_0$ tiene las mismas probabilidades de ser alguno de los valores posibles $1, 2, 3, \ldots, m + n$ (R_i tiene una distribución uniforme discreta en los primeros $m + n$ enteros positivos), así que $\mu_{R_i} = (m + n + 1)/2$. Con esto se obtiene, puesto que $W = \sum R_i$,

$$\mu_W = \mu_{R_1} + \mu_{R_2} + \cdots + \mu_{R_m} = \frac{m(m + n + 1)}{2} \tag{15.4}$$

La varianza de R_i también se calcula fácilmente como $(m + n + 1)(m + n - 1)/12$. Sin embargo, como las R_i no son variables independientes, $V(W) \neq mV(R_i)$. Usando el hecho de que, para dos enteros distintos a y b entre 1 y $m + n$ inclusive, $P(R_i = a, R_j = b) = 1/[(m + n)(m + n - 1)]$ (se muestrean sin reemplazo dos enteros), $\text{Cov}(R_i, R_j) = -(m + n + 1)/12$, lo cual produce

$$\sigma_W^2 = \sum_{i=1}^{m} V(R_i) + \sum\sum_{i \neq j} \text{Cov}(R_i, R_j) = \frac{mn(m + n + 1)}{12} \tag{15.5}$$

También se puede usar un Teorema de límite central para concluir que cuando H_0 es verdadera, el *estadístico de prueba*

$$Z = \frac{W - m(m + n + 1)/2}{\sqrt{mn(m + n + 1)/12}}$$

tiene aproximadamente una distribución normal estándar. Este estadístico se usa junto con los valores críticos z_α, $-z_\alpha$ y $\pm z_{\alpha/2}$ para pruebas de extremo superior, inferior y dos extremos, respectivamente.

Ejemplo 15.5 En un artículo de *Journal of Applied Physiology* ("Histamine Content in Sputum from Allergic and Non-Allergic Individuals", 1969: 535-539) aparecen los siguientes datos acerca de la concentración de histamina en el esputo (μg/g peso seco de esputo) para una muestra de nueve individuos clasificados como alérgicos y otra muestra de 13 individuos clasificados como no alérgicos:

Alérgicos	67.6	39.6	1651.0	100.0	65.9	1112.0	31.0	102.4	64.7				
No alérgicos	34.3	27.3	35.4	48.1	5.2	29.1	4.7	41.7	48.0	6.6	18.9	32.4	45.5

¿Los datos indican que hay una diferencia en la concentración promedio real de histamina en el esputo entre alérgicos y no alérgicos?

Puesto que ambos tamaños de muestra pasan de ocho, se usa la aproximación normal. La hipótesis nula es H_0: $\mu_1 - \mu_2 = 0$, y los rangos observados de las x_i son $r_1 = 18$, $r_2 = 11$, $r_3 = 22$, $r_4 = 19$, $r_5 = 17$, $r_6 = 21$, $r_7 = 7$, $r_8 = 20$ y $r_9 = 16$, así que $w = \sum r_i = 151$. La media y la varianza de W se determinan mediante $\mu_W = 9(23)/2 = 103.5$ y $\sigma_W^2 = 9(13)(23)/12 = 224.25$. Por lo tanto,

$$z = \frac{151 - 103.5}{\sqrt{224.25}} = 3.17$$

La hipótesis alternativa es H_a: $\mu_1 - \mu_2 \neq 0$, así que al nivel 0.01 H_0 se rechaza si $z \geq 2.58$ o $z \leq -2.58$. Debido a que $3.17 \geq 2.58$, se rechaza H_0, y se concluye que hay una diferencia en las concentraciones promedio reales de histamina en el esputo (en el artículo también se utilizó la prueba de Wilcoxon). ■

Empates en la prueba de Wilcoxon de suma de rangos

En teoría, la suposición de continuidad de las dos distribuciones asegura que las $m + n$ x y y observadas tendrán valores distintos. No obstante, en la práctica con frecuencia habrá empates en los valores observados. Como con la prueba de Wilcoxon de los rangos con signo, la práctica

común al tratar con empates es asignar cada una de las observaciones empatadas de un conjunto particular de empates el promedio de rangos que recibirían si difirieran muy poco entre sí.

Si ambos tamaños de muestra pasan de 8, el numerador de Z aún es apropiado, pero el denominador se debe reemplazar por la raíz cuadrada de la varianza ajustada

$$\sigma_W^2 = \frac{mn(m+n+1)}{12} - \frac{mn}{12(m+n)(m+n-1)} \sum (\tau_i - 1)(\tau_i)(\tau_i + 1) \tag{15.6}$$

donde τ_i es el número de observaciones empatadas en el i-ésimo conjunto de empates y la suma es en todos lo conjuntos de empates. A menos que haya gran cantidad de empates, hay poca diferencia entre las ecuaciones (15.6) y (15.5).

Eficiencia de la prueba de Wilcoxon de la suma de rangos

Cuando ambas distribuciones en las que se realiza el muestreo son normales con $\sigma_1 = \sigma_2$, y, por lo tanto, tienen las mismas formas y dispersiones, se puede usar la prueba de t agrupada o la prueba de Wilcoxon (en la prueba t con dos muestras se supone normalidad pero no varianzas iguales, así que las suposiciones que sustentan su uso son más restrictivas en un sentido y menos en otro que las de la prueba de Wilcoxon). En esta situación, la prueba t agrupada es mejor entre las pruebas posibles en el sentido de minimizar β para alguna α fija. Sin embargo, uno nunca puede estar seguro en absoluto de que se satisfacen las suposiciones fundamentales. Por lo tanto, es pertinente preguntar 1) cuánto se pierde al usar la prueba de Wilcoxon en vez de la prueba t agrupada cuando las distribuciones son normales con varianzas iguales y 2) cómo se compara W con T en las situaciones no normales.

El concepto de eficiencia de prueba se analizó en la sección anterior junto con la prueba t de una muestra y la prueba de Wilcoxon de rangos con signo. Los resultados para las dos muestras son los mismos que para las pruebas con una muestra. Cuando se cumplen tanto la normalidad como varianzas iguales, la prueba de suma de rangos es aproximadamente 95% tan eficiente como la prueba t agrupada en muestras grandes. Es decir, la prueba t da las mismas probabilidades de error que la prueba de Wilcoxon con tamaños de muestra un poco más pequeños. Por otro lado, la prueba de Wilcoxon siempre será por lo menos 86% tan eficaz como la prueba t agrupada, y podría ser mucho más eficaz si las distribuciones subyacentes son muy poco normales, en particular con extremos densos. La comparación de la prueba de Wilcoxon con la prueba t bimuestral (no agrupada) es menos clara. Se desconoce que la prueba t sea la mejor en algún sentido, así que parece seguro concluir que siempre y cuando las distribuciones poblacionales tengan formas y dispersiones similares, el comportamiento de la prueba de Wilcoxon debe ser mucho mejor que el de la prueba t bimuestral.

Por último, se observa que los cálculos de β para la prueba de Wilcoxon son bastante difíciles. Esto es porque la distribución de W cuando H_0 es falsa, depende no sólo de $\mu_1 - \mu_2$ sino también de las formas de las dos distribuciones. Para la mayor parte de las distribuciones subyacentes, la distribución no nula de W es casi intratable. Por esta razón los estadísticos desarrollaron la eficiencia (relativa asintótica) con muestras grandes como medio para comparar pruebas. Con las capacidades del software actual, otro método de cálculo para β es realizar un experimento de simulación.

Ejercicios | Sección 15.2 (10 a 16)

10. En un experimento para comparar la fuerza de cohesión de dos adhesivos, cada uno se utilizó en cinco uniones de dos superficies y se determinó la fuerza necesaria para separar las superficies de cada unión. Para el adhesivo 1, los valores resultantes fueron 229, 286, 245, 299 y 250, mientras que las observaciones para el adhesivo 2 fueron 213, 179, 163, 247 y 225. Sea μ_i la fuerza de cohesión promedio del adhesivo tipo i. Utilice la prueba de Wilcoxon de suma de rangos al nivel 0.05 para probar H_0: $\mu_1 = \mu_2$ contra H_a: $\mu_1 > \mu_2$.

11. En el artículo "A Study of Wood Stove Particulate Emissions" (*J. Air Pollution Control Assn.*, 1979: 724-728) aparecen los siguientes datos sobre el tiempo de combustión (horas) para muestras de roble y pino. Pruebe al nivel 0.05 para ver si hay alguna diferencia en el tiempo de combustión promedio real de los dos tipos de madera.

Roble	1.72	0.67	1.55	1.56	1.42	1.23	1.77	0.48
Pino	0.98	1.40	1.33	1.52	0.73	1.20		

12. Se realizó una modificación al proceso para producir cierto tipo de película de "tiempo cero" (película que comienza el revelado tan pronto como se toma la fotografía). Debido a que la modificación implica un costo adicional, se incorpora sólo si los datos muestrales indican de manera clara que la modificación redujo el tiempo promedio real de revelado en más de un segundo. Si se supone que las distribuciones de tiempo de revelado difieren sólo con respecto a la localización si a caso, utilice la prueba de Wilcoxon de suma de rangos al nivel 0.05 en los datos siguientes para probar las hipótesis pertinentes.

Proceso original	8.6	5.1	4.5	5.4	6.3	6.6	5.7	8.5
Proceso modificado	5.5	4.0	3.8	6.0	5.8	4.9	7.0	5.7

13. Los datos siguientes resultaron de un experimento para comparar los efectos de la vitamina C del jugo de naranja y del ácido ascórbico sintético sobre la duración de odontoblastos en conejillos de indias durante un periodo de 6 semanas ("The Growth of the Odontoblasts of the Incisor Tooth as a Criterion of the Vitamin C Intake of the Guinea Pig", *J. Nutrition*, 1947: 491-504). Utilice la prueba de Wilcoxon de suma de rangos al nivel 0.01 para determinar si la duración promedio real difiere para los dos tipos de ingesta de vitamina C. Calcule también un valor P aproximado.

Jugo de naranja	8.2	9.4	9.6	9.7	10.0	14.5	15.2	16.1	17.6	21.5
Ácido ascórbico	4.2	5.2	5.8	6.4	7.0	7.3	10.1	11.2	11.3	11.5

14. Pruebe las hipótesis sugeridas en el ejercicio 13 por medio de los siguientes datos:

Jugo de naranja	8.2	9.5	9.5	9.7	10.0	14.5	15.2	16.1	17.6	21.5
Ácido ascórbico	4.2	5.2	5.8	6.4	7.0	7.3	9.5	10.0	11.5	11.5

15. En el artículo "Measuring the Exposure of Infants to Tobacco Smoke" (*N. Engl. J. Med.*, 1984: 1075-1078) se describe un estudio en el que se tomaron varias medidas de una muestra aleatoria de infantes expuestos a humo doméstico y de una muestra de niños no expuestos. Los siguientes datos consisten en observaciones de la concentración de cotinina en la orina, principal metabolito urinario de la nicotina (los valores son un subconjunto de los datos originales obtenidos de una gráfica del artículo). ¿Estos datos indican que la concentración promedio real de cotinina es mayor en niños expuestos que en no expuestos por más de 25? Realice una prueba al nivel de significación 0.05.

No expuestos	8	11	12	14	20	43	111	
Expuestos	35	56	83	92	128	150	176	208

16. Reconsidere la situación descrita en el ejercicio 79 del capítulo 9 y el siguiente resultado de MINITAB (la letra griega eta denota una mediana).

```
Intervalo de confianza y prueba de Mann-Whitney
Bueno              N = 8          Mediana = 0.540
Malo               N = 8          Mediana = 2.400
La estimación puntual para ETA1-ETA2 es   -1.155
El IC de 95.9 por ciento para ETA1-ETA2 es(-3.160,
-0.409)
W = 41.0
La prueba de ETA1 = ETA2 vs ETA1 < ETA2 es signi-
ficativa en 0.0027
```

a. Compruebe que el valor del estadístico de prueba de MINITAB es correcto.

b. Realice una prueba de hipótesis adecuada usando un nivel de significación de 0.01.

15.3 Intervalos de confianza sin distribución

El método utilizado hasta el momento para construir un intervalo de confianza (IC) se describe como sigue: Comience con una variable aleatoria (Z, T, χ^2, F, o algo parecido) que depende del parámetro de interés y una expresión de probabilidad en la que interviene la variable, realice operaciones con las desigualdades de la expresión para aislar el parámetro entre puntos finales aleatorios y, por último, sustituya los valores calculados para las variables aleatorias. Otro método general para obtener IC aprovecha una relación entre puntos finales aleatorios e intervalos de confianza. Un IC de $100(1 - \alpha)\%$ para un parámetro θ se obtiene a partir de una prueba de nivel α para H_0: $\theta = \theta_0$ contra H_a: $\theta \neq \theta_0$. Este método se utilizará para obtener intervalos de confianza relacionados con la prueba de Wilcoxon de los rangos con signo y la prueba de Wilcoxon de la suma de rangos.

Antes de usar el método para obtener intervalos nuevos, reconsidere la prueba t y el intervalo t. Suponga que una muestra aleatoria de $n = 25$ observaciones de una población normal produce los estadísticos de resumen $\bar{x} = 100$, $s = 20$. Entonces un IC de 90% para μ es

$$\left(\bar{x} - t_{0.05,24} \cdot \frac{s}{\sqrt{25}},\ \ \bar{x} + t_{0.05,24} \cdot \frac{s}{\sqrt{25}}\right) = (93.16,\ 106.84) \tag{15.7}$$

Suponga que en lugar de un IC, se hubiera deseado probar una hipótesis respecto a μ. Para H_0: $\mu = \mu_0$ contra H_a: $\mu \neq \mu_0$, la prueba t al nivel 0.10 especifica que se debe rechazar H_0 si t es ≥ 1.711 o ≤ -1.711, donde

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{25}} = \frac{100 - \mu_0}{20/\sqrt{25}} = \frac{100 - \mu_0}{4} \tag{15.8}$$

Ahora considere el valor nulo $\mu_0 = 95$. Entonces $t = 1.25$ de modo que no se rechaza H_0. De manera similar, si $\mu_0 = 104$, entonces $t = -1$ así que de nuevo no se rechaza H_0. Sin embargo, si $\mu_0 = 90$, entonces $t = 2.5$ y, por lo tanto, se rechaza H_0, y si $\mu_0 = 108$, entonces $t = -2$ de modo que de nuevo se rechaza H_0. Al considerar otros valores de μ_0 y la decisión resultante de cada uno, surge el siguiente hecho general: *todo número dentro del intervalo* (15.7) *especifica un valor de* μ_0 *para el cual t de* (15.8) *da lugar al no rechazo de* H_0, *mientras que todo número fuera del intervalo* (15.7) *corresponde a una t para la cual se rechaza* H_0. Es decir, para valores fijos de n $\bar{x}$ y s, el conjunto de los valores μ_0 para los que probar H_0: $\mu = \mu_0$ contra H_a: $\mu \neq \mu_0$ da como resultado que la aceptación de H_0 sea precisamente en el intervalo (15.7).

PROPOSICIÓN

Suponga que se tiene un procedimiento de prueba de nivel α para probar H_0: $\theta = \theta_0$ contra H_a: $\theta \neq \theta_0$. Para valores muestrales fijos, sea A el conjunto de los valores de θ_0 para los cuales no se rechaza H_0. Entonces A es un IC de $100(1 - \alpha)\%$ para θ.

Hay en realidad ejemplos patológicos en los que el conjunto A definido en la proposición no es un intervalo de valores θ sino el complemento de un intervalo o algo incluso más extraño. Para ser más precisos, se debe reemplazar el concepto de un IC por el de conjunto de confianza. En los casos de interés aquí, el conjunto A no resulta ser un intervalo.

Intervalo de rangos con signo de Wilcoxon

Para probar H_0: $\mu = \mu_0$ contra H_a: $\mu \neq \mu_0$ por medio de la prueba de Wilcoxon de los rangos con signo, donde μ es la media de una distribución simétrica continua, los valores absolutos $|x_1 - \mu_0|, \ldots, |x_n - \mu_0|$ se ordenan de menor a mayor, y se asigna el rango 1 a valor menor y el rango n al mayor. Luego, a cada rango se le asigna el signo de su $x_i - \mu_0$, correspondiente, y el estadístico de prueba es la suma de los rangos con signo positivo. La prueba de dos extremos rechaza H_0 si s_+ es $\geq c$ o bien, $\leq n(n+1)/2 - c$, donde c se obtiene de la tabla A.13 del apéndice una vez que se especifica el nivel de significación α deseado. Para $x_1, \ldots, x_n$, fijas, el intervalo de rangos con signo de $100(1-\alpha)\%$ consistirá en las μ_0 para las que H_0: $\mu = \mu_0$ no se rechaza al nivel α. Para identificar este intervalo, es conveniente expresar el estadístico de prueba S_+ de otra forma.

$$S_+ = \text{número de promedios por pares } (X_i + X_j)/2 \text{ con } i \leq j \text{ que son} \geq \mu_0 \tag{15.9}$$

Es decir, si se promedia cada x_j en la lista con cada x_i a su izquierda, incluso $(x_j + x_j)/2$ (que es sólo x_j), y se cuenta el número de estos promedios que son $\geq \mu_0$, se obtiene s_+. Al ir de izquierda a derecha en la lista de valores muestrales, sólo se está promediando cada par de observaciones de la muestra [de nuevo se incluye $(x_j + x_j)/2$] exactamente una vez, así que no importa el orden en que se listan las observaciones antes de promediar. Es fácil comprobar la equivalencia de los dos métodos para calcular s_+. El número de promedios por pares es $\binom{n}{2} + n$ (el primer término por promediar distintas observaciones y el segundo por promediar cada x_i con ella misma), que es igual a $n(n+1)/2$. Si demasiados promedios por pares o muy pocos de éstos son $\geq \mu_0$, se rechaza H_0.

Ejemplo 15.6 Las siguientes observaciones son valores de la tasa metabólica cerebral para monos rhesus: $x_1 = 4.51$, $x_2 = 4.59$, $x_3 = 4.90$, $x_4 = 4.93$, $x_5 = 6.80$, $x_6 = 5.08$, $x_7 = 5.67$. Los 28 promedios por pares son, en orden creciente,

4.51	4.55	4.59	4.705	4.72	4.745	4.76	4.795	4.835	4.90
4.915	4.93	4.99	5.005	5.08	5.09	5.13	5.285	5.30	5.375
5.655	5.67	5.695	5.85	5.865	5.94	6.235	6.80		

Algunos de los primeros y de los últimos de estos valores se ilustran en un eje de medición en la figura 15.3.

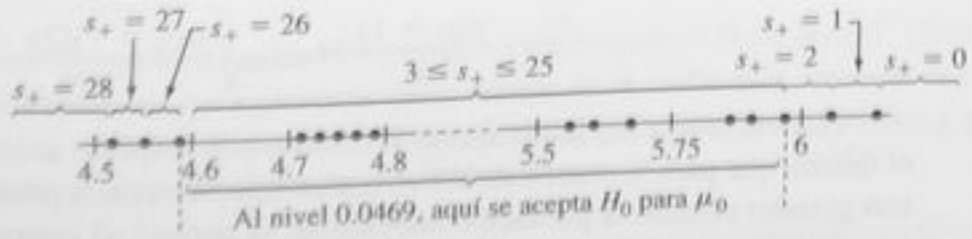


Figura 15.3 Gráfica de los datos para el ejemplo 15.6

Debido a la naturaleza discreta de la distribución de S_+, $\alpha = 0.05$ no se puede obtener de modo exacto. La región de rechazo $\{0, 1, 2, 26, 27, 28\}$ tiene $\alpha = 0.046$, que está lo más cerca posible a 0.05, así que el nivel es aproximadamente 0.05. Por tanto, si el número de promedios por pares $\geq \mu_0$ está entre 3 y 25, inclusive, no se rechaza H_0. De la figura 15.3 el IC de 95% (aproximado) para μ es (4.59, 5.94). ■

En general, una vez que los promedios por pares están ordenados de menor a mayor, los puntos finales del intervalo de Wilcoxon son dos de los promedios "extremos". Para expresar esto de manera precisa, represente con $\bar{x}_{(1)}$ el menor promedio por pares, el siguiente más pequeño con $\bar{x}_{(2)}$, . . . , y el mayor con $\bar{x}_{(n(n+1)/2)}$.

PROPOSICIÓN

Si la prueba de Wilcoxon de los rangos con signo el nivel α para H_0: $\mu = \mu_0$ contra H_a: $\mu \neq \mu_0$ es rechazar H_0 si $s_+ \geq c$ o $s_+ \leq n(n + 1)/2 - c$, entonces un IC de $100(1 - \alpha)\%$ para μ es

$$(\bar{x}_{(n(n+1)/2-c+1)}, \bar{x}_{(c)}) \tag{15.10}$$

En palabras, el intervalo abarca desde el d-ésimo promedio menor por pares hasta el d-ésimo promedio mayor, donde $d = n(n + 1)/2 - c + 1$. En la tabla A.15 del apéndice se encuentran los valores de c que corresponden a los niveles de confianza comunes para $n = 5, 6, \ldots, 25$.

Ejemplo 15.7 (Continuación del ejemplo 15.6)

Para $n = 7$, un intervalo de 89.1% (casi 90%) se obtiene al usar $c = 24$ (puesto que la región de rechazo $\{0, 1, 2, 3, 4, 24, 25, 26, 27, 28\}$ tiene $\alpha = 0.109$). El intervalo es $(\bar{x}_{(28-24+1)}, \bar{x}_{(24)}) = (\bar{x}_{(5)}, \bar{x}_{(24)}) = (4.72, 5.85)$, que se extiende desde el quinto menor promedio por pares hasta el quinto promedio mayor. ■

Otros usos y propiedades del intervalo de rangos con signo de Wilcoxon

La derivación del intervalo depende de tener una sola muestra de una distribución simétrica continua con media (mediana) μ. Cuando se forman pares de datos, el intervalo construido de las diferencias $d_1, d_2, \ldots, d_n$ es un IC para la diferencia media (mediana) μ_D. En este caso, no es necesario suponer la simetría de las distribuciones X y Y; siempre y cuando las distribuciones X y Y tengan la misma forma, la distribución $X - Y$ es simétrica, así que sólo se requiere continuidad.

Para $n > 20$, la aproximación de muestras grandes a la prueba de Wilcoxon basada en estandarizar S_+ proporciona una aproximación a c en (5.10). El resultado [para un intervalo de $100(1 - \alpha)\%$] es

$$c \approx \frac{n(n + 1)}{4} + z_{\alpha/2}\sqrt{\frac{n(n + 1)(2n + 1)}{24}}$$

La eficiencia del intervalo de Wilcoxon con respecto al intervalo t es aproximadamente el mismo que para la prueba de Wilcoxon en relación con la prueba t. En particular, para muestras grandes cuando la población subyacente es normal, el intervalo de Wilcoxon tenderá a ser un poco mayor que el intervalo t, pero si la población es bastante normal (simétrica pero con extremos densos), entonces el intervalo de Wilcoxon tenderá a ser mucho más corto que el intervalo t.

Intervalo de la suma de rangos de Wilcoxon

La prueba de la suma de rangos de Wilcoxon para probar H_0: $\mu_1 - \mu_2 = \Delta_0$ se realiza combinando primero las $(X_i - \Delta_0)$ y las Y_j en una muestra de tamaño $m + n$ y ordenándolas de menor (rango 1) a mayor (rango $m + n$). Entonces el estadístico de prueba es la suma de los rangos de las $(X_i - \Delta_0)$. Para la alternativa de dos extremos, se rechaza H_0 si w es demasiado pequeña o muy grande.

Para obtener el IC para las x_i y las y_j fijas, se debe determinar el conjunto de los valores de Δ_0 para los que no se rechaza H_0. Esto se realiza con mayor facilidad si primero se expresa el estadístico de prueba en una forma un poco distinta. El mínimo valor posible de W es $m(m + 1)/2$, que corresponde a toda $(X_i - \Delta_0)$ menor que toda Y_j, y hay mn diferencias de la forma $(X_i - \Delta_0) - Y_j$. Con unas cuantas operaciones se obtiene

$$\begin{aligned} W &= [\text{número de } (X_i - Y_j - \Delta_0)\text{'s} \geq 0] + \frac{m(m+1)}{2} \\ &= [\text{número de } (X_i - Y_j)\text{'s} \geq \Delta_0] + \frac{m(m+1)}{2} \end{aligned} \tag{15.11}$$

Así, rechazar H_0 si la cantidad de $(x_i - y_j)$'s $\geq \Delta_0$ es muy pequeña o demasiado grande es equivalente a rechazar H_0 para w pequeña o grande.

La expresión (15.11) indica que se calcula $x_i - y_j$ para cada i y j y se ordenan estas mn diferencias de menor a mayor. Entonces si el valor nulo Δ_0 no es menor que la mayor parte de las diferencias ni mayor, no se rechaza H_0: $\mu_1 - \mu_2 = \Delta_0$. Ahora, al variar Δ_0 se observa que un IC para $\mu_1 - \mu_2$ tendrá como punto final inferior una de las $(x_i - y_j)$ ordenadas y, de manera similar para el punto final superior.

PROPOSICIÓN

Sean $x_1, \ldots, x_m$ y $y_1, \ldots, y_n$ los valores observados en las dos muestras independientes de distribuciones continuas que difieren sólo en ubicación (y no en forma). Con $d_{ij} = x_i - y_j$ y las diferencias ordenadas denotadas por $d_{ij(1)}, d_{ij(2)}, \ldots, d_{ij(mn)}$, la forma general de un IC de $100(1 - \alpha)\%$ para $\mu_1 - \mu_2$ es

$$(d_{ij(mn-c+1)}, d_{ij(c)}) \tag{15.12}$$

donde c es la constante crítica para la prueba de la suma de rangos de Wilcoxon de dos extremos al nivel α.

Observe que la forma del intervalo de la suma de rangos de Wilcoxon (15.2) es muy similar a la del intervalo de rangos con signo de Wilcoxon (15.10); en el intervalo (15.10) se utilizan promedios por pares de una sola muestra, mientras que en (15.12) se emplean diferencias por pares de dos muestras. En la tabla A.16 del apéndice se encuentran los valores de c para valores seleccionados de m y n.

Ejemplo 15.8 En el artículo "Some Mechanical Properties of Impregnated Bark Board" (*Forest Products J.*, 1977: 31-38) aparecen los siguientes datos acerca de la resistencia máxima de compresión (lb/pulg^2) para una muestra de panel de corteza impregnado con resina epóxica y para una muestra de panel de corteza impregnado con otro polímero:

Resina epóxica (x)	10 860	11 120	11 340	12 130	14 380	13 070
Otro polímero (y)	4590	4850	6510	5640	6390	

Obtenga un IC de 95% para la diferencia promedio real en la resistencia de compresión entre el panel impregnado con resina epóxica y el otro tipo de panel.

De la tabla A.16 del apéndice, puesto que el tamaño de muestra más pequeño es 5 y el tamaño de muestra más grande es 6, $c = 26$ para un nivel de confianza de aproximadamente 95%. Las d_{ij} aparecen en la tabla 15.5. Las cinco d_{ij} más pequeñas $[d_{ij(1)}, \ldots, d_{ij(5)}]$ son 4350, 4470, 4610, 4730 y 4830; y las cinco d_{ij} más grandes son (en orden descendente) 9790, 9530, 8740, 8480 y 8220. Así, el IC es $(d_{ij(5)}, d_{ij(26)}) = (4830, 8220)$.

Tabla 15.5 Diferencias para el intervalo de la suma de rangos del ejemplo 15.8

	d_{ij}	y_j 4590	4850	5640	6390	6510
	10 860	6270	6010	5220	4470	4350
	11 120	6530	6270	5480	4730	4610
x_i	**11 340**	6750	6490	5700	4950	4830
	12 130	7540	7280	6490	5740	5620
	13 070	8480	8220	7430	6680	6560
	14 380	9790	9530	8740	7990	7870

■

Cuando m y n son grandes, el estadístico de la prueba de Wilcoxon tiene aproximadamente una distribución normal. Esto se puede usar para obtener una aproximación con muestras grandes para el valor c en el intervalo (15.12). El resultado es

$$c \approx \frac{mn}{2} + z_{\alpha/2}\sqrt{\frac{mn(m+n+1)}{12}} \tag{15.13}$$

Al igual que con el intervalo de los rangos con signo, el intervalo de suma de rangos (15.12) es bastante eficaz con respecto al intervalo t; en muestras grandes, (15.12) tiende a ser sólo un poco más extenso que el intervalo t cuando las poblaciones subyacentes son normales y podría ser considerablemente más corto que el intervalo t si las poblaciones subyacentes tienen extremos más densos que las poblaciones normales.

Ejercicios | Sección 15.3 (17 a 22)

17. En el artículo "The Lead Content and Acidity of Christchurch Precipitation" (*New Zealand J. Science*, 1980: 311-312) aparecen los siguientes datos sobre la concentración de plomo (μg/L) en muestras reunidas durante ocho lluvias de verano: 17.0, 21.4, 30.6, 5.0, 12.2, 11.8, 17.3 y 18.8. Si se supone que la distribución del contenido de plomo es simétrica, utilice el intervalo Wilcoxon de rangos con signo a fin de obtener un intervalo de confianza de 95% para μ.

18. Calcule el intervalo de rangos con signo de 99% para el pH promedio real μ (suponiendo simetría). Utilice los datos del ejercicio 3. [*Sugerencia:* pruebe calcular sólo los promedios por pares que tengan valores relativamente pequeños o grandes (en lugar de los 105 promedios).]

19. Calcule un intervalo de confianza para μ_D del ejemplo 15.2 con los datos suministrados en ese ejercicio; el nivel de confianza debe ser más o menos 95%.

20. Las siguientes observaciones son cantidades de emisiones de hidrocarburos que resultan del desgaste de neumáticos con cuerdas diagonales bajo una carga de 522 kg con presión de inflado de 228 kPa y velocidad de 64 km/h durante 6 horas ("Characterization of Tire Emissions Using an Indoor Test Facility", *Rubber Che-*

mistry and Technology, 1978: 7-25): 0.045, 0.117, 0.062 y 0.072. ¿Cuáles niveles de confianza son factibles para este tamaño de muestra utilizando el intervalo de rangos con signo? Seleccione un nivel de confianza apropiado y calcule el intervalo.

21. Calcule el intervalo de confianza de la suma de rangos de 90% para $\mu_1 - \mu_2$ mediante los datos del ejercicio 10.

22. Calcule un intervalo de confianza de 99% para $\mu_1 - \mu_2$ con los datos del ejercicio 11.

15.4 ANOVA sin distribución

En el modelo de ANOVA de un solo factor del capítulo 10 para comparar la población I o medias de tratamiento se supuso que para $i = 1, 2, \ldots, I$, una muestra aleatoria de tamaño J_i se extrajo de una población normal con media μ_i y varianza σ^2. Esto se puede escribir como

$$X_{ij} = \mu_i + \epsilon_{ij} \qquad j = 1, \ldots, J_i;\ i = 1, \ldots, I \tag{15.14}$$

donde las ϵ_{ij} son independientes y están distribuidas normalmente con media cero y varianza σ^2. Aunque la suposición de normalidad se requirió para la validez de la prueba F descrita en el capítulo 10, la validez de la prueba de Kruskal-Wallis para probar la igualdad de las μ_i depende sólo de las ϵ_{ij} que tienen la misma distribución de continuidad.

Prueba de Kruskal–Wallis

Sea $N = \sum J_i$, el número total de observaciones en el conjunto de datos, y suponga que las N observaciones se clasifican desde 1 (la menor de las X_{ij}) hasta N (la X_{ij} más grande). Cuando H_0: $\mu_1 = \mu_2 = \cdots = \mu_I$ es verdadera, las N observaciones provienen de la misma distribución, en cuyo caso las posibles asignaciones de los rangos $1, 2, \ldots, N$ a las I muestras son equiprobables y se espera que los rangos estén combinados en estas muestras. Sin embargo, si H_0 es falsa, entonces algunas muestras consistirán sobre todo en observaciones que tienen rangos pequeños en la muestra combinada, mientras que otras consistirán principalmente en observaciones que tienen rangos grandes. Para ser específicos, si R_{ij} denota el rango de X_{ij} entre las N observaciones, y $R_{i\cdot}$ y $\overline{R}_{i\cdot}$ denotan, respectivamente, el total y el promedio de los rangos en la i-ésima muestra, entonces cuando H_0 es verdadera

$$E(R_{ij}) = \frac{N+1}{2} \qquad E(\overline{R}_{i\cdot}) = \frac{1}{J_i}\sum_j E(R_{ij}) = \frac{N+1}{2}$$

El estadístico de prueba de Kruskal-Wallis es una medida del grado de desviación de las $\overline{R}_{i\cdot}$ respecto de su valor esperado común $(N+1)/2$, y H_0 se rechaza si el valor calculado del estadístico indica demasiada discrepancia entre los promedios de rangos observados y esperados.

ESTADÍSTICO DE PRUEBA

$$K = \frac{12}{N(N+1)}\sum_{i=1}^{I} J_i\left(\overline{R}_{i\cdot} - \frac{N+1}{2}\right)^2 = \frac{12}{N(N+1)}\sum_{i=1}^{I}\frac{R_{i\cdot}^2}{J_i} - 3(N+1) \tag{15.15}$$

La segunda expresión para K es la fórmula de cálculo; en ella intervienen los totales de rango ($R_{i\cdot}$) en lugar de los promedios y requiere sólo una resta.

Si se rechaza H_0 cuando $k \geq c$, entonces el valor de c se elige de tal modo que la prueba tenga nivel α. Es decir, c debe ser el valor crítico de extremo superior de la distribución de K cuando H_0 es verdadera. Bajo H_0, cada asignación posible de los rangos a las I muestras es equiprobable, de modo que en teoría todas las asignaciones de este tipo se pueden enumerar, el valor de K se determina para cada una y la distribución nula se obtiene al contar el número de veces que ocurre cada valor de K. Es evidente que este cálculo resulta tedioso, así que aunque haya tablas de la distribución nula exacta y valores críticos para valores pequeños de las J_i, se usará la siguiente aproximación para "muestras grandes".

PROPOSICIÓN

Cuando H_0 es verdadera y se cumple que

$$I = 3 \qquad J_i \geq 6 \qquad (i = 1, 2, 3)$$

o bien

$$I > 3 \qquad J_i \geq 5 \qquad (i = 1, \ldots, I)$$

entonces K tiene aproximadamente una distribución chi-cuadrada con $I - 1$ gl. Esto significa que una prueba con nivel de significación aproximado α rechaza H_0 si $k \geq \chi^2_{\alpha, I-1}$.

Ejemplo 15.9 Las siguientes observaciones (tabla 15.6) acerca del índice de rigidez axial resultaron de un estudio de armaduras unidas a placas metálicas en las que se usaron cinco longitudes de placa (pulgadas) distintas, 4, 6, 8, 10 y 12 ("Modeling Joints Made with Light-Gauge Metal Connector Plates", *Forest Products J.*, 1979: 39-44).

Tabla 15.6 Datos y rangos para el ejemplo 15.9

	$i = 1$ (4″):	309.2	309.7	311.0	316.8	326.5	349.8	409.5		
	$i = 2$ (6″):	331.0	347.2	348.9	361.0	381.7	402.1	404.5		
	$i = 3$ (8″):	351.0	357.1	366.2	367.3	382.0	392.4	409.9		
	$i = 4$ (10″):	346.7	362.6	384.2	410.6	433.1	452.9	461.4		
	$i = 5$ (12″):	407.4	410.7	419.9	441.2	441.8	465.8	473.4		
							$r_{i\cdot}$	$\bar{r}_{i\cdot}$		
	$i = 1$:	1	2	3	4	5	10	24	49	7.00
	$i = 2$:	6	8	9	13	17	21	22	96	13.71
Rangos	$i = 3$:	11	12	15	16	18	20	25	117	16.71
	$i = 4$:	7	14	19	26	29	32	33	160	22.86
	$i = 5$:	23	27	28	30	31	34	35	208	29.71

El valor calculado de K es

$$k = \frac{12}{35(36)}\left[\frac{(49)^2}{7} + \frac{(96)^2}{7} + \frac{(117)^2}{7} + \frac{(160)^2}{7} + \frac{(208)^2}{7}\right] - 3(36)$$
$$= 20.12$$

Al nivel 0.01, $\chi^2_{0.01,4} = 13.277$, y como $20.12 \geq 13.277$, se rechaza H_0 y se concluye que la rigidez axial esperada depende de la longitud de la placa. ■

Prueba de Friedman para un experimento de bloques aleatorizado

Suponga que $X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$, donde α_i es el i-ésimo efecto de tratamiento, β_j es el j-ésimo efecto de bloque y las ϵ_{ij} se extraen de forma independiente de la misma distribución continua (pero no necesariamente normal). Entonces para probar H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$, la hipótesis nula de ningún efecto de tratamiento, primero se clasifican por separado las observaciones de 1 hasta I dentro de cada bloque, y luego se calcula el promedio de rangos $\bar{r}_{i\cdot}$ para cada uno de los I tratamientos. Cuando H_0 es verdadera, las $\bar{r}_{i\cdot}$ deben estar cerca entre sí, puesto que dentro de cada bloque las $I!$ asignaciones de rangos a los tratamientos son equiprobables. El estadístico de prueba de Friedman mide la discrepancia entre el valor esperado $(I + 1)/2$ de cada promedio de rangos y las $\bar{r}_{i\cdot}$.

ESTADÍSTICO DE PRUEBA

$$F_r = \frac{12J}{I(I+1)} \sum_{i=1}^{I} \left(\bar{R}_{i\cdot} - \frac{I+1}{2} \right)^2 = \frac{12}{IJ(I+1)} \sum R_{i\cdot}^2 - 3J(I+1)$$

Como con la prueba de Kruskal-Wallis, la prueba de Friedman rechaza H_0 cuando el valor calculado del estadístico de prueba es demasiado grande. Para los casos $I = 3$, $J = 2, \ldots, 15$ e $I = 4$, $J = 2, \ldots, 8$, el libro de Lehmann (véase la bibliografía del capítulo) suministra los valores críticos de extremo superior para la prueba. De manera alternativa, incluso para valores moderados de J, el estadístico de prueba F_r tiene aproximadamente una distribución chi-cuadrada con $I - 1$ gl cuando H_0 es verdadera, así que se rechaza H_0 si $f_r \geq \chi^2_{\alpha,I-1}$.

Ejemplo 15.10 En el artículo "Physiological Effects During Hypnotically Requested Emotions" (*Psychosomatic Med.*, 1963: 334-343) aparecen los datos siguientes (tabla 15.7) acerca del potencial de la piel (mV) cuando se pidió de cada uno de ocho sujetos las emociones de temor, felicidad, depresión y tranquilidad.

Tabla 15.7 Datos de los rangos para el ejemplo 15.10

	Bloques (sujetos)									
x_{ij}	1	2	3	4	5	6	7	8		
Temor	23.1	57.6	10.5	23.6	11.9	54.6	21.0	20.3		
Felicidad	22.7	53.2	9.7	19.6	13.8	47.1	13.6	23.6		
Depresión	22.5	53.7	10.8	21.1	13.7	39.2	13.7	16.3		
Tranquilidad	22.6	53.1	8.3	21.6	13.3	37.0	14.8	14.8		
Rangos	1	2	3	4	5	6	7	8	$r_{i\cdot}$	$r_{i\cdot}^2$
Temor	4	4	3	4	1	4	4	3	27	729
Felicidad	3	2	2	1	4	3	1	4	20	400
Depresión	1	3	4	2	3	2	2	2	19	361
Tranquilidad	2	1	1	3	2	1	3	1	14	196
										1686

Así

$$f_r = \frac{12}{4(8)(5)}(1686) - 3(8)(5) = 6.45$$

Al nivel 0.05, $\chi^2_{0.05,3} = 7.815$, y como 6.45 no es ≥ 7.815, H_0 no se rechaza. No hay evidencia de que el potencial cutáneo promedio dependa de la emoción requerida. ■

En el libro de Myles Hollander y Douglas Wolfe (véase la bibliografía del capítulo) se analizan procedimientos de comparaciones múltiples relacionados con las pruebas de Kruskal-Wallis y Friedman, así como otros aspectos del ANOVA sin distribución.

Ejercicios | Sección 15.4 (23 a 27)

23. Los siguientes datos se refieren a la concentración del isótopo radiactivo estroncio 90 en muestras de leche obtenidas de cinco lecherías seleccionadas al azar, en cada una de cuatro regiones diferentes.

Región	**1**	6.4	5.8	6.5	7.7	6.1
	2	7.1	9.9	11.2	10.5	8.8
	3	5.7	5.9	8.2	6.6	5.1
	4	9.5	12.1	10.3	12.4	11.7

Pruebe al nivel 0.10 si el promedio verdadero de concentración de estroncio 90 difiere para por lo menos dos de las regiones.

24. En el artículo "Production of Gaseous Nitrogen in Human Steady-State Conditions" (*J. Applied Physiology*, 1972: 155-159) se encuentran las siguientes observaciones acerca de la cantidad de nitrógeno expirado (en litros) bajo cuatro regímenes dietéticos: 1) ayuno, 2) 23% de proteína, 3) 32% de proteína y 4) 67% de proteína. Utilice la prueba Kruskal-Wallis al nivel 0.05 para probar la igualdad de las μ_i correspondientes.

1.	4.079	4.859	3.540	5.047	3.298
2.	4.368	5.668	3.752	5.848	3.802
3.	4.169	5.709	4.416	5.666	4.123
4.	4.928	5.608	4.940	5.291	4.674
1.	4.679	2.870	4.648	3.847	
2.	4.844	3.578	5.393	4.374	
3.	5.059	4.403	4.496	4.688	
4.	5.038	4.905	5.208	4.806	

25. Los siguientes datos sobre la concentración de cortisol se encuentran en el artículo "Cortisol, Cortisone, and 11-Deoxycortisol Levels in Human Umbilical and Maternal Plasma in Relation to the Onset of Labor" (*J. Obstetric Gynaecology British Commonwealth*, 1974: 737-745). Los sujetos de experimento eran mujeres embarazadas que dieron a luz después de entre 38 y 42 semanas de gestación. Las mujeres del grupo 1 eligieron dar a luz por cesárea antes de comenzar el parto, el grupo 2 dio a luz por cesárea de emergencia durante el trabajo de parto inducido, y las del grupo 3 experimentaron trabajo de parto espontáneo. Utilice la prueba Kruskal-Wallis al nivel 0.05 para probar la igualdad de las tres medias poblacionales.

Grupo 1	262	307	211	323	454	339
	304	154	287	356		
Grupo 2	465	501	455	355	468	362
Grupo 3	343	772	207	1048	838	687

26. En una prueba para determinar si el suelo tratado previamente con pequeñas cantidades de Basic-H hace que éste sea más permeable al agua, se dividieron muestras de suelo en bloques, y a cada bloque se aplicó cada uno de los cuatro tratamientos en estudio. Los tratamientos fueron *(A)* agua con 0.001% de Basic-H inundado sobre el suelo de control, *(B)* agua sin Basic-H sobre suelo de control, (C) agua con Basic-H inundado sobre suelo tratado previamente con Basic-H y *(D)* agua sin Basic-H en suelo pretratado con Basic-H. Pruebe al nivel 0.01 si hay algunos efectos ocasionados por los diferentes tratamientos.

	Bloques				
	1	2	3	4	5
A	37.1	31.8	28.0	25.9	25.5
B	33.2	25.3	20.2	20.3	18.3
C	58.9	54.2	49.2	47.9	38.2
D	56.7	49.6	46.4	40.9	39.4
	6	**7**	**8**	**9**	**10**
A	25.3	23.7	24.4	21.7	26.2
B	19.3	17.3	17.0	16.7	18.3
C	48.8	47.8	40.2	44.0	46.4
D	37.1	37.5	39.6	35.1	36.5

27. En un experimento para estudiar la forma en que diferentes anestésicos afectan la concentración de epinefrina plasmática, se eligió a diez perros y se les midió la concentración mientras estaban bajo la influencia de los anestésicos isoflurano, halotano y ciclopropano ("Sympathoadrenal and Hemodynamic Effects of Isoflurane, Halothane, and Cyclopropane in Dogs", *Anesthesiology,* 1974: 465-470). Pruebe al nivel 0.05 si hay un efecto del anestésico en la concentración.

	Perro				
	1	2	3	4	5
Isoflurano	0.28	0.51	1.00	0.39	0.29
Halotano	0.30	0.39	0.63	0.38	0.21
Ciclopropano	1.07	1.35	0.69	0.28	1.24
	6	**7**	**8**	**9**	**10**
Isoflurano	0.36	0.32	0.69	0.17	0.33
Halotano	0.88	0.39	0.51	0.32	0.42
Ciclopropano	1.53	0.49	0.56	1.02	0.30

Ejercicios complementarios (28 a 36)

28. En el artículo "Effects of a Rice-Rich Versus Potato-Rich Diet on Glucose, Lipoprotein, and Cholesterol Metabolism in Noninsulin-Dependent Diabetics" (*Amer. J. Clinical Nutr.,* 1984: 598-606) aparecen los siguientes datos acerca de la tasa de síntesis de colesterol para ocho personas con diabetes sometidas a una dieta estandarizada con papas o arroz como fuente principal de carbohidratos. Los participantes recibieron ambas dietas durante periodos especificados, con la tasa de síntesis de colesterol (mmol/día) medida al término de cada periodo de dieta. En el análisis presentado en este artículo, se utilizó una prueba sin distribución. Utilice este tipo de prueba al nivel 0.05 para determinar si la tasa promedio real de síntesis de colesterol difiere de manera significativa de las dos fuentes de carbohidratos.

Tasa de síntesis de colesterol

Individuo	1	2	3	4	5	6	7	8
Papas	1.88	2.60	1.38	4.41	1.87	2.89	3.96	2.31
Arroz	1.70	3.84	1.13	4.97	0.86	1.93	3.36	2.15

29. Las tácticas de ventas en las que se presiona mucho al cliente o los vendedores de puerta en puerta, pueden ser bastante molestos. Muchas personas sucumben ante estas tácticas, firman un convenio de compra y después lamentan sus acciones. A mediados de la década de 1970, la Federal Trade Commission puso en práctica reglamentos que aclaran y amplían los derechos de los compradores para cancelar los convenios. Los siguientes datos son un subconjunto de los que se encuentran en el artículo "Evaluating the FTC Cooling-Off Rule" (*J. Consumer Affairs,* 1977: 101-106). Las observaciones individuales son tasas de cancelación para cada uno de nueve vendedores durante cada uno de cuatro años. Utilice una prueba adecuada al nivel 0.05 para ver si la tasa de cancelación promedio real depende del año.

	Vendedor								
	1	2	3	4	5	6	7	8	9
1973	2.8	5.9	3.3	4.4	1.7	3.8	6.6	3.1	0.0
1974	3.6	1.7	5.1	2.2	2.1	4.1	4.7	2.7	1.3
1975	1.4	0.9	1.1	3.2	0.8	1.5	2.8	1.4	0.5
1976	2.0	2.2	0.9	1.1	0.5	1.2	1.4	3.5	1.2

30. Los datos acerca de la concentración de fósforo en la superficie del suelo para cuatro tipos de tratamiento aparecen en el artículo "Fertilisers for Lotus and Clover Establishment on a Sequence of Acid Soils on the East Otago Uplands" (*N. Zeal. J. Exper. Ag.,* 1984: 119-129). Utilice un procedimiento sin distribución para probar la hipótesis nula de ninguna diferencia en concentración promedio real de fósforo (mg/g) para los cuatro tratamientos.

Tratamiento	**I**	8.1	5.9	7.0	8.0	9.0
	II	11.5	10.9	12.1	10.3	11.9
	III	15.3	17.4	16.4	15.8	16.0
	IV	23.0	33.0	28.4	24.6	27.7

31. Refiérase a los datos del ejercicio 30 y calcule un intervalo de confianza de 95% para la diferencia entre las concentraciones promedio reales para los tratamientos II y III.

32. El estudio descrito en "Gait Patterns During Free Choice Ladder Ascents" (*Human Movement Sci.*, 1983: 187-195) se inspiró en la publicidad relacionada con la mayor tasa de accidentes para personas que suben por escaleras. Individuos que ascendían por una escalera portátil recta utilizaron patrones de marcha de acuerdo con instrucciones especificadas. A continuación se muestran los tiempos de ascenso para siete individuos que utilizaron paso lateral y seis personas que usaron paso diagonal de cuatro tiempos.

Lateral	0.86	1.31	1.64	1.51	1.53	1.39	1.09
Diagonal	1.27	1.82	1.66	0.85	1.45	1.24	

a. Realice una prueba con $\alpha = 0.05$ para ver si los datos indican alguna diferencia en los tiempos de ascenso promedio verdaderos para los dos pasos.

b. Calcule un intervalo de confianza de 95% para la diferencia entre los tiempos de paso promedio reales.

33. La **prueba de los signos** es un procedimiento muy simple para probar hipótesis acerca de una mediana poblacional suponiendo sólo que la distribución subyacente es continua. Como ilustración, considere el siguiente ejemplo de 20 observaciones de duración de un componente (h):

1.7	3.3	5.1	6.9	12.6	14.4	16.4
24.6	26.0	26.5	32.1	37.4	40.1	40.5
41.5	72.4	80.1	86.4	87.5	100.2	

Se desea probar H_0: $\tilde{\mu} = 25.0$ contra H_a: $\tilde{\mu} > 25.0$. El estadístico de prueba es Y = número de observaciones mayores que 25.

a. Considere rechazar H_0 si $Y \geq 15$. ¿Cuál es el valor de α (la probabilidad de un error tipo I) para esta prueba? *Sugerencia:* Considere que una duración mayor que 25.0 es un "éxito". Entonces, Y es la cantidad de éxitos en la muestra. ¿Qué clase de distribución tiene Y cuando $\tilde{\mu} = 25.0$?

b. ¿Qué región de rechazo de la forma $Y \geq c$ especifica una prueba con nivel de significación tan cercano a 0.05 como sea posible? Use esta región a fin de efectuar la prueba para los datos.

Nota: El estadístico de prueba es el número de diferencias $X_i - 25.0$ que tienen signos positivos; de ahí el nombre de *prueba de los signos*.

34. Refiérase al ejercicio 33 y considere un intervalo de confianza relacionado con la prueba de los signos, el **intervalo de signo**. Ahora las hipótesis pertinentes son H_0: $\tilde{\mu} = \tilde{\mu}_0$ contra H_a: $\tilde{\mu} \neq \tilde{\mu}_0$. Utilícese la siguiente región de rechazo: ya sea $Y \geq 15$ o $Y \leq 5$.

a. ¿Cuál es el nivel de significación para esta prueba?

b. El intervalo de confianza consistirá en los valores $\tilde{\mu}_0$ para los que no se rechaza H_0. Determine el intervalo de confianza para los datos del ejercicio y exprese el nivel de confianza.

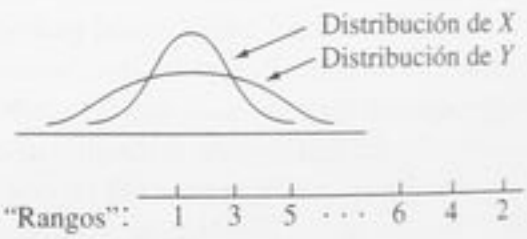
35. Suponga que se desea probar

H_0: las distribuciones de X y Y son idénticas

contra

H_a: la distribución de X está menos dispersa que la de Y

En la figura siguiente se ilustran las distribuciones de X y de Y para las cuales H_a es verdadera. La prueba de Wilcoxon de la suma de rangos no es apropiada en esta situación, porque cuando H_a es verdadera, como se ilustra, las Y tienden a quedar en los puntos extremos de la muestra combinada (lo cual da como resultado rangos de Y pequeños y grandes), por lo que la suma de rangos de X produce un valor W que no es grande ni pequeño.



Considere modificar el procedimiento para asignar rangos como sigue: Después de ordenar la muestra combinada de $m + n$ observaciones, la observación menor recibe el rango 1, la observación más grande el rango 2, la segunda más pequeña el rango 3, la segunda más grande el rango 4, y así sucesivamente. Entonces, si H_a es verdadera, como se ilustra, los valores X tienden a estar en la parte media de la muestra y, por lo tanto, reciben rangos grandes. Sea W' la suma de los rangos de X y consideremos rechazar H_0 en favor de H_a cuando $w' \geq c$. Cuando H_0 es verdadera, todo conjunto posible de rangos de X tiene la misma probabilidad, así que W' *tiene la misma distribución que* W *cuando* H_0 *es verdadera.* Por lo tanto, c se puede seleccionar de la tabla A.14 del apéndice a fin de obtener una prueba de nivel α. Los datos que se muestran a continuación se refieren

al grosor del músculo medial para arteriolas de los pulmones de niños que murieron por síndrome de muerte infantil repentina (*Sudden Infant Death Syndrome*, SIDS) (las x) y de un grupo de niños utilizados como control (las y). Realice la prueba de H_0 contra H_a al nivel 0.05.

SIDS	4.0	4.4	4.8	4.9	
Control	3.7	4.1	4.3	5.1	5.6

Para más información sobre esta prueba llamada *prueba de Siegel-Tukey*, consulte la obra de Lehmann (en la bibliografía del capítulo).

36. El procedimiento clasificación descrito en el ejercicio 35 es un poco asimétrico, porque la observación más pequeña recibe el rango 1, mientras que la más grande recibe el rango 2, y así sucesivamente. Suponga que las observaciones menor y mayor reciben el rango 1, la segunda más pequeña y la segunda más grande el rango 2, y así sucesivamente, y sea W'' la suma de los rangos de X. La distribución nula de W'' no es idéntica a la distribución nula de W, por lo que se necesitan tablas diferentes. Considere el caso $m = 3$, $n = 4$. Liste los 35 ordenamientos posibles de los tres valores de X entre las siete observaciones (p. ej., 1, 3, 7 o 4, 5, 6), asigne rangos en la forma descrita, calcule el valor de W'' para cada posibilidad y después tabule la distribución nula de W''. Para la prueba que rechaza si $w'' \geq c$, ¿cuál valor de c prescribe aproximadamente una prueba con nivel 0.10? Ésta es la *prueba de Ansari-Bradley*. Para más información, véase el libro de Hollander y Wolfe en la bibliografía del capítulo.

Bibliografía

Hollander, Myles y Douglas Wolfe, *Nonparametric Statistical Methods* (2a. ed.), Wiley, Nueva York, 1999. Una muy buena referencia sobre métodos sin distribución con una excelente colección de tablas.

Lehmann, Erich, *Nonparametrics: Statistical Methods Based on Ranks*, Holden-Day, San Francisco, 1975. Un excelente análisis de los métodos libres de distribución más importantes, presentados con gran cantidad de comentarios reveladores.

16

Métodos de control de calidad

Introducción

Las características de calidad de productos manufacturados han recibido mucha atención de los ingenieros de diseño y personal de producción, así como de los interesados en la administración financiera. Una cuestión de fe durante años, fue que los muy altos niveles de calidad y el bienestar económico eran objetivos incompatibles. Sin embargo, en fechas recientes, se ha vuelto cada vez más evidente que elevar los niveles de calidad puede dar lugar a costos reducidos, un mayor grado de satisfacción del cliente y, por lo tanto, mayor confiabilidad. Esto ha dado como resultado un énfasis renovado en las técnicas estadísticas para diseñar calidad en productos y para identificar problemas de calidad en varias etapas de la producción y distribución.

En la actualidad, las gráficas de control se utilizan de manera extensa en la industria como una técnica de diagnóstico para monitorear procesos de producción a fin de identificar inestabilidad y circunstancias inusuales. Después de una introducción a ideas básicas en la sección 16.1, se presentan varias gráficas de control en las siguientes cuatro secciones. La base para la mayor parte de éstas, se basa en el trabajo previo en relación con las distribuciones de probabilidad de varios estadísticos, como la media muestral $\overline{X}$ y la proporción muestral $\hat{p} = X/n$.

Otra situación encontrada con frecuencia en entornos industriales, tiene que ver con la decisión de un cliente en cuanto a si un lote de artículos que ofrece un proveedor es de calidad aceptable. En la última sección del capítulo, se estudian de

manera breve algunos métodos de muestreo de aceptación para decidir, con base en los datos de la muestra, sobre la disposición de un lote.

Además de las gráficas de control y los planes de muestreo de aceptación, que se desarrollaron por vez primera en las décadas de 1920 y 1930, los estadísticos e ingenieros han introducido en fechas recientes muchos métodos estadísticos nuevos para identificar tipos y niveles de entradas de producción que aseguran producto de alta calidad. Investigadores japoneses, y en particular el ingeniero y estadístico G. Taguchi y sus discípulos, han tenido mucho que ver en este respecto, y hoy en día está disponible una gran cantidad de material conocido como "Métodos de Taguchi". Las ideas de diseño experimental, y en particular experimentos factoriales fraccionarios, son elementos clave. Aún hay mucha controversia en la comunidad estadística en cuanto a qué diseños y métodos de análisis son los más adecuados para la tarea en cuestión. El artículo explicativo de George Box *et al.*, citado en la bibliografía del capítulo, contiene una crítica reciente; el libro de Thomas Ryan listado allí es una buena fuente de información.

16.1 Comentarios generales acerca de las gráficas de control

Un mensaje primordial en todo este libro, ha sido el dominio de la variación que ocurre de manera natural relacionada con alguna característica o atributo de individuos u objetos diferentes. En un contexto de manufactura, sin importar cuán cuidadosa sea la calibración de las máquinas, el control de los factores ambientales, el monitoreo de materiales y otros insumos y la capacitación de los trabajadores, el diámetro variará de un tornillo a otro, algunas láminas de plástico serán más resistentes que otras, algunos fusibles serán más eficaces mientras que otros no, etcétera. Se podría considerar a esta clase de variación aleatoria natural como un ruido de fondo incontrolable.

Sin embargo, hay otras fuentes de variación que podrían tener un impacto nocivo en la calidad de los artículos producidos en algunos procesos. La variación se podría atribuir a material contaminado, posición incorrecta de los controles de la máquina, uso inusual de las herramientas, y cosas por el estilo. Las fuentes de variación se denominan *causas asignables* en las publicaciones de control de calidad. Las **gráficas de control** proporcionan un mecanismo para reconocer situaciones donde las causas asignables podrían afectar de manera adversa la calidad del producto. Una vez que una gráfica indica una situación fuera de control, se podría lanzar una investigación para identificar las causas y emprender la acción correctiva.

Un elemento básico de las gráficas de control, es que las muestras han sido seleccionadas del proceso de interés en una secuencia de puntos del tiempo. Dependiendo del aspecto del proceso bajo investigación, se elige algun estadístico, como la media muestral o la proporción muestral de los artículos defectuosos. Entonces, el valor de este estadístico se calcula para cada muestra uno tras otro. Así, una gráfica de control común resulta de dibujar estos valores calculados en función del tiempo, como se ilustra en la figura 16.1.

Observe que además de los puntos trazados, la gráfica tiene una recta central y dos límites de control. La base para la elección de una recta central a veces es un valor objetivo o una

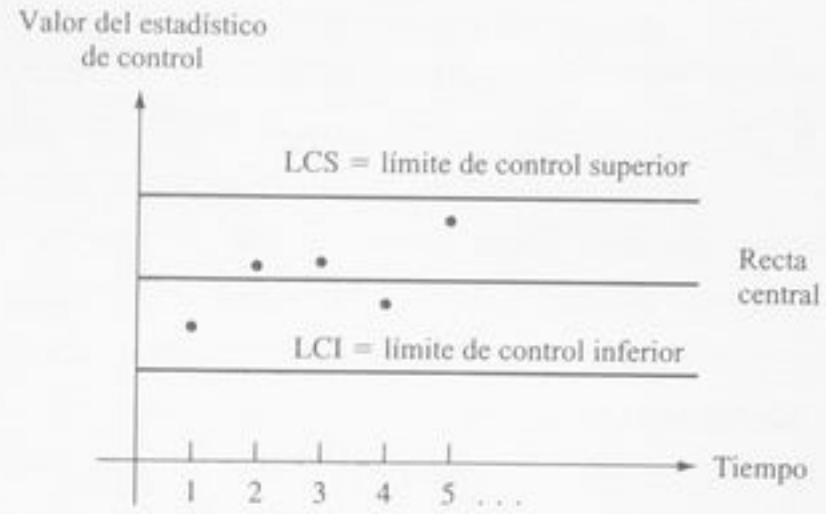


Figura 16.1 Prototipo de una gráfica de control

especificación de diseño, por ejemplo, un valor deseado del diámetro del cojinete. En otros casos, la altura de la recta central se estima a partir de los datos. Si los puntos de la gráfica yacen entre los dos límites de control, se considera que el proceso está bajo control. Es decir, se cree que el proceso está operando en un modo estable que refleja sólo variación aleatoria natural. Una "señal" fuera de control ocurre si un punto dibujado cae fuera de los límites. Se supone que esto es atribuible a alguna causa asignable, y comienza una búsqueda de tales causas. Los límites se diseñan de modo que un proceso bajo control genere muy pocas falsas alarmas, mientras que un proceso fuera de control rápidamente da lugar a un punto fuera de los límites.

Hay una fuerte analogía entre la lógica de las gráficas de control y el trabajo previo en la prueba de hipótesis. La hipótesis nula aquí es que el proceso está bajo control. Cuando un proceso bajo control produce un punto *fuera* de los límites de control (una señal fuera de control), ha ocurrido un error de tipo I. Por otro lado, un error de tipo II resulta cuando un proceso fuera de control produce un punto *dentro* de los límites de control. Una elección apropiada del tamaño de muestra y los límites de control (estos últimos correspondientes a especificar una región de rechazo en la prueba de hipótesis) hace que las probabilidades de error correspondientes sean adecuadamente pequeñas.

Se remarca que "bajo control" no es sinónimo de "satisface las especificaciones de diseño o tolerancia". El grado de variación natural podría ser tal que el porcentaje de elementos que incumplen la especificación puede ser mucho mayor que el que se puede tolerar. En estos casos, se requiere una mayor reestructuración del proceso, a fin de mejorar su capacidad. Un proceso bajo control es simplemente uno cuyo comportamiento con respecto a la variación es estable con el tiempo, sin indicaciones de causas extrañas inusuales.

Como comentario introductorio final, en la actualidad hay una amplia variedad de software para gráficas de control. La publicación *Quality Progress* contiene muchos anuncios de paquetes de software para control de calidad estadístico. Además, SAS y la más reciente versión de MINITAB, entre otros paquetes de aplicación general, tienen atractivas capacidades de control de calidad.

Ejercicios | Sección 16.1 (1 a 3)

1. Una gráfica de control para el espesor de láminas de acero rolado, se basa en un límite de control superior de 0.0520 pulgadas y un límite inferior de 0.0475 pulgadas. Los primeros diez valores del estadístico de calidad (en este caso $\bar{X}$, el espesor promedio muestral de $n = 5$ láminas de la muestra) son 0.0506, 0.0493, 0.0502, 0.0501, 0.0512, 0.0498, 0.0485, 0.0500, 0.0505 y 0.0483. Construya la parte inicial de la gráfica de control de calidad y comente acerca de la apariencia de la gráfica.

2. Refiérase al ejercicio 1 y suponga que los diez valores más recientes del estadístico de calidad son 0.0493, 0.0485, 0.0490, 0.0503, 0.0492, 0.0486, 0.0495, 0.0494, 0.0493 y 0.0488. Construya la parte importante de la gráfica de control correspondiente y comente acerca de características interesantes.

3. Suponga que se construye una gráfica de control de modo que la probabilidad de que un punto caiga fuera de los límites de control, cuando en realidad el proceso está bajo control, es 0.002. ¿Cuál es la probabilidad de que diez puntos sucesivos (basados en muestras elegidas de forma independiente) estén dentro de los límites de control? ¿Cuál es la probabilidad de que 25 puntos sucesivos queden dentro de los límites de control? ¿Cuál es el número más pequeño de puntos sucesivos trazados para los cuales la probabilidad de observar por lo menos uno fuera de los límites de control es mayor que 0.10?

16.2 Gráficas de control para localización de proceso

Suponga que la característica de calidad de interés se relaciona con una variable cuyos valores observados resultan de hacer mediciones. Por ejemplo, la característica podría ser resistencia de cable eléctrico (ohms), diámetro interno de juntas de expansión de caucho moldeado (cm) o la dureza de cierta aleación (unidades Brinell). Un uso importante de las gráficas de control es ver si alguna medida de ubicación de la distribución de la variable permanece estable con el tiempo. El diagrama más popular para este propósito es la gráfica de $\bar{X}$.

Gráfica de $\bar{X}$ basada en valores de parámetros conocidos

Debido a que hay incertidumbre acerca del valor de la variable para algún elemento particular o espécimen, esta variable *aleatoria* se representa mediante X. Suponga que para un proceso bajo control, X tiene una distribución normal con valor medio μ y desviación estándar σ. Entonces, si $\bar{X}$ denota la media muestral para una muestra aleatoria de tamaño n seleccionada en un determinado punto del tiempo, se sabe que

1. $E(\bar{X}) = \mu$
2. $\sigma_{\bar{X}} = \sigma/\sqrt{n}$
3. $\bar{X}$ tiene una distribución normal.

Se deduce que

$$P(\mu - 3\sigma_{\bar{X}} \le \bar{X} \le \mu + 3\sigma_{\bar{X}}) = P(-3.00 \le Z \le 3.00) = 0.9974$$

donde Z es una va normal estándar.* Por lo tanto, es muy probable que para un proceso bajo control, la media muestral caiga dentro de tres desviaciones estándar ($3\sigma_{\bar{X}}$) de la media del proceso μ.

Considere primero el caso en el que se conocen los valores de μ y σ. Suponga que en cada uno de los puntos de tiempo 1, 2, 3, . . . , se tiene una muestra aleatoria de tamaño n. Sean $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots$ los valores calculados de las medias muestrales correspondientes. Una gráfica de $\bar{X}$ resulta de graficar estas $\bar{x}_i$ en función del tiempo; es decir, graficar los puntos $(1, \bar{x}_1)$, $(2, \bar{x}_2)$, $(3, \bar{x}_3)$, etcétera, y luego trazar rectas horizontales de un lado a otro de la gráfica en

$$\text{LCI} = \text{límite de control inferior} = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}}$$

$$\text{LCS} = \text{límite de control superior} = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}}$$

*El uso de gráficas basadas en límites de tres desviaciones estándar es tradicional, pero de hecho la tradición no es inviolable.

A este diagrama se le denomina gráfica 3-sigma. Cualquier punto fuera de los límites de control indica que el proceso podría haber estado fuera de control en ese momento, así que se inicia una búsqueda de causas asignables.

Ejemplo 16.1 Una vez al día, se escogen al azar tres especímenes de aceite para motor del proceso de producción, y se analiza cada uno para determinar la viscosidad. Los datos siguientes (tabla 16.1) son para un periodo de 25 días. La amplia experiencia con este proceso indica que cuando el proceso está bajo control, la viscosidad de un espécimen tiene una distribución normal con media 10.5 y desviación estándar 0.18. Así, $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.18/\sqrt{3} = 0.104$, de modo que los límites de control de tres desviaciones estándar son

$$\text{LCI} = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} = 10.5 - 3(0.104) = 10.188$$

$$\text{LCS} = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} = 10.5 + 3(0.104) = 10.812$$

Los puntos en la gráfica de control mostrados en la figura 16.2 están entre los límites de control, lo que indica comportamiento estable de la media del proceso durante este tiempo (la desviación estándar y el intervalo para cada muestra se emplean en la siguiente subsección).

Tabla 16.1 Datos de viscosidad del ejemplo 16.1

Día	Observaciones de viscosidad			$\bar{x}$	s	Intervalo
1	10.37	10.19	10.36	10.307	0.101	0.18
2	10.48	10.24	10.58	10.433	0.175	0.34
3	10.77	10.22	10.54	10.510	0.276	0.55
4	10.47	10.26	10.31	10.347	0.110	0.21
5	10.84	10.75	10.53	10.707	0.159	0.31
6	10.48	10.53	10.50	10.503	0.025	0.05
7	10.41	10.52	10.46	10.463	0.055	0.11
8	10.40	10.38	10.69	10.490	0.173	0.31
9	10.33	10.35	10.49	10.390	0.087	0.16
10	10.73	10.45	10.30	10.493	0.218	0.43
11	10.41	10.68	10.25	10.447	0.217	0.43
12	10.00	10.60	10.71	10.437	0.382	0.71
13	10.37	10.50	10.34	10.403	0.085	0.16
14	10.47	10.60	10.75	10.607	0.140	0.28
15	10.46	10.46	10.56	10.493	0.058	0.10
16	10.44	10.68	10.32	10.480	0.183	0.36
17	10.65	10.42	10.26	10.443	0.196	0.39
18	10.73	10.72	10.83	10.760	0.061	0.11
19	10.39	10.75	10.27	10.470	0.250	0.48
20	10.59	10.23	10.35	10.390	0.183	0.36
21	10.47	10.67	10.64	10.593	0.108	0.20
22	10.40	10.55	10.38	10.443	0.093	0.17
23	10.24	10.71	10.27	10.407	0.263	0.47
24	10.37	10.69	10.40	10.487	0.177	0.32
25	10.46	10.35	10.37	10.393	0.059	0.11

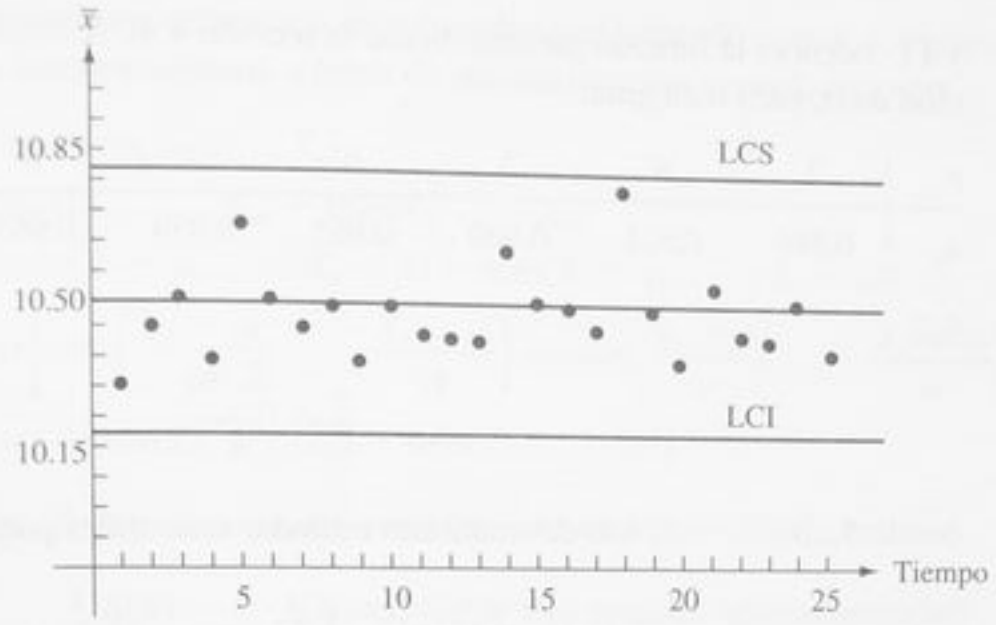


Figura 16.2 Gráfica de $\bar{X}$ para los datos de viscosidad del ejemplo 16.1 ■

Gráficas de $\bar{X}$ basadas en los parámetros estimados

En la práctica sucede con frecuencia que se ignoran los valores de μ y σ, así que se deben estimar a partir de los datos muestrales antes de determinar los límites de control. Esto es cierto en particular cuando un proceso se somete primero a un análisis de control de calidad. De nuevo represente con n el número de observaciones de cada muestra y con k el número disponible de muestras. Los valores representativos de n son 3, 4, 5 o 6, y se recomienda que k sea por lo menos 20. Se supone que las k muestras se reunieron durante un periodo cuando se consideró que el proceso estuvo bajo control. En breve se amplía la información acerca de esta suposición.

Con $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ que representa las k medias muestrales calculadas, la estimación usual de μ es simplemente el promedio de estas medias:

$$\hat{\mu} = \bar{\bar{x}} = \frac{\sum_{i=1}^{k} \bar{x}_i}{k}$$

Se presentan dos métodos distintos de uso común para estimar σ, uno basado en las k desviaciones estándar muestrales y el otro en los k intervalos de la muestra (recuerde que el intervalo de la muestra es la diferencia entre las observaciones máxima y mínima). Antes de la extensa disponibilidad de buenas calculadoras y software para estadística, una consideración primordial era la facilidad de cálculo a mano, así que predominaba el método del intervalo. Sin embargo, en el caso de una distribución poblacional normal, se sabe que el estimador insesgado de s basado en S tiene variación menor que el que se basa en el intervalo de la muestra. Los estadísticos opinan que el primer estimador es más *eficaz* que el último. La pérdida de eficiencia para el estimador es ligera cuando n es muy pequeña, pero se vuelve importante cuando $n > 4$.

Recuerde que la desviación estándar muestral no es un estimador insesgado para σ. Cuando $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución normal, se puede demostrar (refiérase al ejercicio 6.37) que

$$E(S) = a_n \cdot \sigma$$

donde

$$a_n = \frac{\sqrt{2}\Gamma(n/2)}{\sqrt{n-1}\Gamma[(n-1)/2]}$$

y $\Gamma(\cdot)$ denota la función gamma (véase la sección 4.4). A continuación se muestra una tabulación de a_n para n elegida:

n	3	4	5	6	7	8
a_n	0.886	0.921	0.940	0.952	0.959	0.965

Sea

$$\bar{S} = \frac{\sum_{i=1}^{k} S_i}{k}$$

donde $S_1, S_2, \ldots, S_k$ son desviaciones estándar muestrales para las k muestras. Entonces

$$E(\bar{S}) = \frac{1}{k} E\left(\sum_{i=1}^{k} S_i\right) = \frac{1}{k} \sum_{i=1}^{k} E(S_i) = \frac{1}{k} \sum_{i=1}^{k} a_n \cdot \sigma = a_n \cdot \sigma$$

Así

$$E\left(\frac{\bar{S}}{a_n}\right) = \frac{1}{a_n} E(\bar{S}) = \frac{1}{a_n} \cdot a_n \cdot \sigma = \sigma$$

por lo tanto $\hat{\sigma} = \bar{S}/a_n$ es un estimador insesgado de σ.

Límites de control basados en las desviaciones estándar muestrales

$$\text{LCI} = \bar{\bar{x}} - 3 \cdot \frac{\bar{s}}{a_n \sqrt{n}}$$

$$\text{LCS} = \bar{\bar{x}} + 3 \cdot \frac{\bar{s}}{a_n \sqrt{n}}$$

donde

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{k} \bar{x}_i}{k} \qquad \bar{s} = \frac{\sum_{i=1}^{k} s_i}{k}$$

Ejemplo 16.2 Con referencia a los datos de viscosidad del ejemplo 16.1, se tenía $n = 3$ y $k = 25$. Los valores de $\bar{x}_i$ y s_i $(i = 1, \ldots, 25)$ aparecen en la tabla 16.1, de los cuales se deduce que $\bar{\bar{x}} = 261.896/25 = 10.476$ y $\bar{s} = 3.834/25 = 0.153$. Con $a_3 = 0.886$, se tiene

$$\text{LCI} = 10.476 - 3 \cdot \frac{0.153}{0.886\sqrt{3}} = 10.476 - 0.299 = 10.177$$

$$\text{LCS} = 10.476 + 3 \cdot \frac{0.153}{0.886\sqrt{3}} = 10.476 + 0.299 = 10.775$$

Estos límites difieren un poco de los límites previos basados en $\mu = 10.5$ y $\sigma = 0.18$ porque ahora $\hat{\mu} = 10.476$ y $\hat{\sigma} = \bar{s}/a_3 = 0.173$. La inspección de la tabla 16.1 muestra que toda $\bar{x}_i$ está entre estos nuevos límites, así que de nuevo ninguna situación fuera de control es evidente. ■

Para obtener una estimación de σ basada en el intervalo muestral, observe que si $X_1, \ldots, X_n$ forma una muestra aleatoria a partir de una distribución normal, entonces

$$\begin{aligned} R &= \text{intervalo}(X_1, \ldots, X_n) \\ &= \text{máx}(X_1, \ldots, X_n) - \text{mín}(X_1, \ldots, X_n) \\ &= \text{máx}(X_1 - \mu, \ldots, X_n - \mu) - \text{mín}(X_1 - \mu, \ldots, X_n - \mu) \\ &= \sigma\left\{\text{máx}\left(\frac{X_1 - \mu}{\sigma}, \ldots, \frac{X_n - \mu}{\sigma}\right) - \text{mín}\left(\frac{X_1 - \mu}{\sigma}, \ldots, \frac{X_n - \mu}{\sigma}\right)\right\} \\ &= \sigma \cdot \{\text{máx}(Z_1, \ldots, Z_n) - \text{mín}(Z_1, \ldots, Z_n)\} \end{aligned}$$

donde $Z_1, \ldots, Z_n$ son va normales estándar independientes. Por lo tanto

$$\begin{aligned} E(R) &= \sigma \cdot E(\text{intervalo de una muestra normal estándar}) \\ &= \sigma \cdot b_n \end{aligned}$$

de modo que R/b_n es un estimador insesgado de σ.

Ahora se denotan los intervalos de las k muestras del conjunto de datos de control de calidad mediante $r_1, r_2, \ldots, r_k$. El argumento antes mencionado indica que la estimación

$$\hat{\sigma} = \frac{\frac{1}{k}\sum_{i=1}^{k} r_i}{b_n} = \frac{\bar{r}}{b_n}$$

proviene de un estimador insesgado para σ. Los valores seleccionados de b_n aparecen en la siguiente tabla [su cálculo se basa en usar la teoría estadística y la integración numérica para determinar $E(\text{mín}(Z_1, \ldots, Z_n))$ y $E(\text{máx}(Z_1, \ldots, Z_n))$].

n	3	4	5	6	7	8
b_n	1.693	2.058	2.325	2.536	2.706	2.844

Límites de control basados en intervalos muestrales

$$\text{LCI} = \bar{\bar{x}} - 3 \cdot \frac{\bar{r}}{b_n\sqrt{n}}$$

$$\text{LCS} = \bar{\bar{x}} + 3 \cdot \frac{\bar{r}}{b_n\sqrt{n}}$$

donde $\bar{r} = \sum_{i=1}^{k} r_i/k$ y $r_1, \ldots, r_k$ son los k intervalos muestrales individuales.

Ejemplo 16.3 (Continuación del ejemplo 16.2) Es fácil comprobar de la tabla 16.1 que $\bar{r} = 0.292$, de modo que $\hat{\sigma} = 0.292/b_3 = 0.292/1.693 = 0.172$ y

$$\text{LCI} = 10.476 - 3 \cdot \frac{0.292}{1.693\sqrt{3}} = 10.476 - 0.299 = 10.177$$

$$\text{LCS} = 10.476 + 3 \cdot \frac{0.292}{1.693\sqrt{3}} = 10.476 + 0.299 = 10.775$$

Estos límites son idénticos a los que se basan en $\bar{s}$, y de nuevo toda $\bar{x}_i$ se ubica entre los límites. ■

Recálculo de los límites de control

Se ha supuesto que los datos muestrales empleados para estimar μ y σ se obtuvieron de un proceso bajo control. Sin embargo, suponga que uno de los puntos de la gráfica de control resultante queda fuera de los límites de control. Entonces, si es posible encontrar y comprobar una causa asignable para esta situación fuera de control, se recomienda que se calculen nuevos límites después de eliminar la muestra correspondiente del conjunto de datos. De manera similar, si más de un punto cae fuera de los límites originales, se deben determinar nuevos límites después de eliminar algún punto para el cual se puede identificar y tratar una causa asignable. Podría suceder que uno o más puntos quedaran fuera de los nuevos límites, en cuyo caso se debe repetir el proceso de eliminación y recálculo.

Características de desempeño de gráficas de control

A grandes rasgos, una gráfica de control es eficaz si suministra muy pocas señales fuera de control cuando el proceso está controlado, pero muestra un punto fuera de los límites tan pronto como el proceso sale de control. Una estimación de la efectividad de una gráfica se basa en el concepto de "probabilidades de error". Suponga que la variable de interés tiene una distribución normal con σ conocida (el mismo valor para un proceso bajo control o fuera de control). Además, considere una gráfica 3-sigma basada en el valor objetivo μ_0, con $\mu = \mu_0$ cuando el proceso está bajo control. Una probabilidad de error es

$$\begin{aligned}\alpha &= P(\text{una sola muestra proporciona un punto fuera de los límites de control cuando } \mu = \mu_0)\\ &= P(\bar{X} > \mu_0 + 3\sigma/\sqrt{n} \text{ o } \bar{X} < \mu_0 - 3\sigma/\sqrt{n} \text{ cuando } \mu = \mu_0)\\ &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > 3 \text{ o } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -3 \text{ cuando } \mu = \mu_0\right)\end{aligned}$$

La variable estandarizada $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ tiene una distribución normal estándar cuando $\mu = \mu_0$, por lo tanto

$$\alpha = P(Z > 3 \text{ o } Z < -3) = \Phi(-3.00) + 1 - \Phi(3.00) = 0.0026$$

Si se utiliza 3.09 en vez de 3 para determinar los límites de control (esto se acostumbra en Gran Bretaña), entonces

$$\alpha = P(Z > 3.09 \text{ o } Z < -3.09) = 0.0020$$

El uso de límites 3-sigma hace que sea muy improbable que una señal fuera de control resulte de un proceso bajo control.

Ahora suponga que un proceso se sale de control porque μ ha cambiado a $\mu + \Delta\sigma$ (Δ podría ser positiva o negativa); Δ es el número de desviaciones estándar por las que ha cambiado μ. Una segunda probabilidad de error es

$$\begin{aligned}\beta &= P\left(\begin{array}{l}\text{una sola muestra proporciona un punto dentro}\\ \text{de los límites de control cuando } \mu = \mu_0 + \Delta\sigma\end{array}\right)\\ &= P(\mu_0 - 3\sigma/\sqrt{n} < \bar{X} < \mu_0 + 3\sigma/\sqrt{n} \text{ cuando } \mu = \mu_0 + \Delta\sigma)\end{aligned}$$

Ahora se estandariza restando primero $\mu_0 + \Delta\sigma$ de cada término dentro del paréntesis y dividiendo después entre $\sigma/\sqrt{n}$. Esto da

$$\begin{aligned}\beta &= P(-3 - \sqrt{n}\Delta < \text{va normal estándar} < 3 - \sqrt{n}\Delta) \\ &= \Phi(3 - \sqrt{n}\Delta) - \Phi(-3 - \sqrt{n}\Delta)\end{aligned}$$

Esta probabilidad de error depende de Δ, que determina el tamaño del cambio, y del tamaño de la muestra n. En particular, para Δ fija, β disminuye cuando n aumenta (mientras más grande sea el tamaño de la muestra, hay más probabilidades de que resulte una señal fuera de control), y para n, β disminuye cuando $|\Delta|$ aumenta (mientras más grande sea la magnitud del cambio, más probable es que resulte una señal fuera de control). En la tabla siguiente se proporciona β para valores seleccionados de Δ cuando $n = 4$.

Δ	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
β *cuando* $n = 4$	0.9936	0.9772	0.9332	0.8413	0.5000	0.1587	0.0668	0.0013

Es claro que un cambio pequeño tiene bastantes probabilidades de pasar inadvertido en una sola muestra.

Si el 3 se sustituye por 3.09 en los límites de control, entonces α disminuye de 0.0026 a 0.002, pero para cualquier n y σ fijas β disminuye. Ésta es sólo una manifestación de la relación inversa entre los dos tipos de probabilidades de error en la prueba de hipótesis. Por ejemplo, cambiar 3 a 2.5 aumenta α y disminuye β.

Las probabilidades de error analizadas hasta el momento se calculan bajo la suposición de que la variable de interés tiene una distribución normal. Si la distribución es sólo un poco no normal, el efecto del teorema del límite central indica que $\overline{X}$ tendrá aproximadamente una distribución normal incluso cuando n es pequeña, en cuyo caso las probabilidades de error expresadas serán más o menos correctas. Por supuesto, éste ya no es el caso cuando la distribución de la variable se desvía en forma considerable de la normalidad.

Una segunda evaluación de desempeño tiene que ver con la duración esperada o promedio de la corrida necesaria para observar una señal fuera de control. Cuando el proceso está bajo control, se debe esperar observar muchas muestras antes de ver una cuya $\bar{x}$ quede fuera de los límites de control. Por otro lado, si el proceso sale de control, el número esperado de muestras necesario para detectar esto debe ser pequeño.

Sea p la probabilidad de que una sola muestra produzca un valor $\bar{x}$ fuera de los límites de control, es decir,

$$p = P(\overline{X} < \mu_0 - 3\sigma/\sqrt{n} \text{ o bien, } \overline{X} > \mu_0 + 3\sigma/\sqrt{n})$$

Considere primero un proceso bajo control, de modo que $\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots$ están distribuidas normalmente con un valor medio μ_0 y desviación estándar $\sigma/\sqrt{n}$. Defina una va Y mediante

$$Y = \text{la primera } i \text{ para la cual } \overline{X}_i \text{ cae fuera de los límites de control}$$

Si se considera que cada número de muestra es un ensayo y una muestra fuera de control es un éxito, entonces Y es el número de ensayos (independientes) necesarios para observar un éxito. Esta Y tiene una distribución geométrica, y en el ejemplo 3.18 se mostró que $E(Y) = 1/p$. Las siglas ARL (para *average run lenght*, duración promedio de la corrida) se utilizan con frecuencia en lugar de $E(Y)$. Debido a que $p = \alpha$ para un proceso bajo control, se tiene

$$\text{ARL} = E(Y) = \frac{1}{p} = \frac{1}{\alpha} = \frac{1}{0.0026} = 384.62$$

Al sustituir el 3 en los límites de control por 3.09, se obtiene ARL = 1/0.002 = 500.

Ahora se supone que, en un determinado punto del tiempo, la media del proceso cambia a $\mu = \mu_0 + \Delta\sigma$. Si se define a Y como la primer i posterior al cambio para la cual una muestra genera una señal fuera de control, de nuevo se cumple que ARL = $E(Y) = 1/p$, pero ahora $p = 1 - \beta$. En la tabla siguiente se muestran las ARL seleccionadas para una gráfica 3-sigma cuando $n = 4$. Estos resultados muestran de nuevo la efectividad de la gráfica para detectar cambios grandes, pero también su capacidad para identificar con rapidez cambios pequeños. Cuando el muestreo se hace en forma bastante esporádica, es probable que se produzca gran cantidad de elementos antes que se detecte un cambio pequeño en μ. Los procedimientos CUSUM analizados en la sección 16.5 se elaboraron para atender esta deficiencia.

Δ	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
ARL cuando $n = 4$	156.25	43.86	14.97	6.30	2.00	1.19	1.07	1.0013

Reglas complementarias para gráficas de $\overline{X}$

La incapacidad de las gráficas de $\overline{X}$ con límites 3-sigma para detectar cambios pequeños en la media del proceso ha motivado a los investigadores a elaborar procedimientos que proporcionan un comportamiento mejorado en este respecto. Un método requiere introducir condiciones adicionales que causan que se genere una señal fuera de control. Western Electric (entonces una subsidiaria de AT&T) recomendó las siguientes condiciones. Una intervención para emprender una acción correctiva es apropiada si se satisface una de estas condiciones:

1. Dos de tres puntos sucesivos caen fuera de los límites 2-sigma en el mismo lado de la recta central.
2. Cuatro de cinco puntos sucesivos caen fuera de los límites 1-sigma en el mismo lado de la recta central.
3. Tres puntos sucesivos caen en el mismo lado de la recta central.

Se debe consultar un texto de control de calidad para tener una descripción de éstas y otras reglas complementarias. En la sección 16.5, se presenta un tipo de procedimiento distinto para superar la deficiencia de la gráfica de $\overline{X}$.

Gráficas de control robustas

La presencia de valores atípicos en la muestra tiende a reducir la sensibilidad de procedimientos de gráficas de control cuando se deben estimar los parámetros. Esto es porque los límites de control se mueven fuera de la recta central, lo cual hace más difícil la identificación de puntos inusuales. *No* se quieren los estadísticos cuyos valores al dibujarse sean resistentes a los valores atípicos, porque eso ocultaría cualquier señal fuera de control. Por ejemplo, trazar las medianas muestrales sería menos eficaz que trazar $\overline{x}_1, \overline{x}_2, \ldots$ como se hizo en una gráfica de $\overline{X}$.

En el artículo "Robust Control Charts" de David M. Rocke (*Technometrics*, 1989: 173 a 184) se presenta un estudio de procedimientos para los que los límites de control se basan en estadísticos resistentes a los efectos de los valores atípicos. Rocke recomienda límites de control calculados a partir del *intervalo intercuartil* (IQR), que es muy similar a la cuarta dispersión presentada en el capítulo 1. En particular,

$$\text{IQR} = \begin{cases} (\text{2a. } x_i \text{ más grande}) - (\text{2a. } x_i \text{ más pequeña}) & n = 4, 5, 6, 7 \\ (\text{3a. } x_i \text{ más grande}) - (\text{3a. } x_i \text{ más pequeña}) & n = 8, 9, 10, 11 \end{cases}$$

Para una muestra aleatoria de una distribución normal, $E(\text{IQR}) = k_n\sigma$; los valores de k_n se muestran en la siguiente tabla.

n	4	5	6	7	8
k_n	0.596	0.990	1.282	1.512	0.942

Los límites de control sugeridos son

$$\text{LCI} = \tilde{\bar{x}} - 3 \cdot \frac{\text{IQR}}{k_n\sqrt{n}} \qquad \text{LCS} = \tilde{\bar{x}} + 3 \cdot \frac{\text{IQR}}{k_n\sqrt{n}}$$

Se dibujan los valores de $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots$ Las simulaciones descritas en el artículo indican que el desempeño de la gráfica con estos límites es superior al de la gráfica de $\bar{X}$ comun.

Ejercicios | Sección 16.2 (4 a 13)

4. En el caso de μ y σ, conocidas, ¿cuáles límites de control son necesarios para que la probabilidad de un solo punto, que está fuera de los límites en un proceso bajo control, sea 0.005?

5. Considere una gráfica de control de 3-sigma con línea central en μ_0 y basada en $n = 5$. Suponiendo normalidad, calcule la probabilidad de que un solo punto caiga fuera de los límites de control cuando la media real del proceso es

   a. $\mu_0 + 0.5\sigma$
   b. $\mu_0 - \sigma$
   c. $\mu_0 + 2\sigma$

6. En la siguiente tabla se muestran datos acerca del contenido de humedad para especímenes de cierto tipo de tela. Determine los límites de control para una gráfica con línea central en la altura 13.00 con base en $\sigma = 0.600$, construya la gráfica de control y comente sobre sus características.

Datos para el ejercicio 6

Núm. de muestra	Observaciones de contenido de humedad					$\bar{x}$	s	Intervalo
1	12.2	12.1	13.3	13.0	13.0	12.72	0.536	1.2
2	12.4	13.3	12.8	12.6	12.9	12.80	0.339	0.9
3	12.9	12.7	14.2	12.5	12.9	13.04	0.669	1.7
4	13.2	13.0	13.0	12.6	13.9	13.14	0.477	1.3
5	12.8	12.3	12.2	13.3	12.0	12.52	0.526	1.3
6	13.9	13.4	13.1	12.4	13.2	13.20	0.543	1.5
7	12.2	14.4	12.4	12.4	12.5	12.78	0.912	2.2
8	12.6	12.8	13.5	13.9	13.1	13.18	0.526	1.3
9	14.6	13.4	12.2	13.7	12.5	13.28	0.963	2.4
10	12.8	12.3	12.6	13.2	12.8	12.74	0.329	0.9
11	12.6	13.1	12.7	13.2	12.3	12.78	0.370	0.9
12	13.5	12.3	12.8	13.1	12.9	12.92	0.438	1.2
13	13.4	13.3	12.0	12.9	13.1	12.94	0.559	1.4
14	13.5	12.4	13.0	13.6	13.4	13.18	0.492	1.2
15	12.3	12.8	13.0	12.8	13.5	12.88	0.432	1.2
16	12.6	13.4	12.1	13.2	13.3	12.92	0.554	1.3
17	12.1	12.7	13.4	13.0	13.9	13.02	0.683	1.8
18	13.0	12.8	13.0	13.3	13.1	13.04	0.182	0.5
19	12.4	13.2	13.0	14.0	13.1	13.14	0.573	1.6
20	12.7	12.4	12.4	13.9	12.8	12.84	0.619	1.5
21	12.6	12.8	12.7	13.4	13.0	12.90	0.316	0.8
22	12.7	13.4	12.1	13.2	13.3	12.94	0.541	1.3

7. Refiérase a los datos del ejercicio 6 y construya una gráfica de control con una recta central estimada y límites basados en el uso de desviaciones estándar muestrales para estimar σ. ¿Hay alguna evidencia de que el proceso está fuera de control?

8. Consulte los ejercicios 6 y 7, y ahora utilice límites de control, basados en el uso de intervalos muestrales para estimar σ. ¿El proceso parece estar bajo control?

9. En la tabla siguiente aparecen las medias muestrales y desviaciones estándar, cada una basada en $n = 6$ observaciones de índice de refracción de cable de fibra óptica. Construya una gráfica de control y comente sobre su aspecto. [*Sugerencia:* $\sum \bar{x}_i = 2317.07$ y $\sum s_i = 30.34$.]

Día	$\bar{x}$	s	Día	$\bar{x}$	s
1	95.47	1.30	9	96.63	1.48
2	97.38	0.88	10	96.50	0.80
3	96.85	1.43	11	97.22	1.42
4	96.64	1.59	12	96.55	1.65
5	96.87	1.52	13	97.02	1.28
6	96.52	1.27	14	95.55	1.14
7	96.08	1.16	15	96.29	1.37
8	96.48	0.79	16	96.80	1.40
17	96.01	1.58	21	97.06	1.34
18	95.39	0.98	22	98.34	1.60
19	96.58	1.21	23	96.42	1.22
20	96.43	0.75	24	95.99	1.18

10. Refiérase al ejercicio 9. Se encontró una causa asignable del índice de refracción promedio muestral inusualmente alto del día 22. Vuelva a calcular los límites de control, después de eliminar los datos de este día. ¿Cuál es su conclusión?

11. Considere la gráfica de control basada en límites de control $\mu_0 \pm 2.81\ \sigma/\sqrt{n}$.
    a. ¿Cuál es la ARL cuando el proceso está bajo control?
    b. ¿Cuál es la ARL cuando $n = 4$ y la media del proceso se ha desplazado $\mu = \mu_0 + \sigma$?
    c. ¿Cómo se comparan los valores de los incisos (a) y (b) con los valores correspondientes para una gráfica de 3-sigma?

12. Aplique las reglas complementarias sugeridas en el texto a los datos del ejercicio 6. ¿Hay algunas señales fuera de control?

13. Calcule límites de control para los datos del ejercicio 6, utilizando el procedimiento robusto presentado en esta sección.

16.3 Gráficas de control para variación de proceso

Las gráficas de control analizadas en la sección anterior se diseñaron para controlar la ubicación (en otras palabras, la tendencia central) de un proceso, con especial atención en la media como una medida de ubicación. Es igualmente importante asegurar que un proceso está bajo control con respecto a la variación. De hecho, la mayor parte de los profesionales recomendaron que se estableciera el control en la variación *antes de* construir una gráfica de $\bar{X}$ o cualquier otra gráfica para controlar la ubicación. En esta sección, se consideran gráficas para variación basadas en la desviación estándar muestral S y también gráficas basadas en el rango muestral R. Las primeras tienen mayor preferencia porque la desviación estándar proporciona una evaluación más eficaz de la variación que el intervalo, pero las gráficas de R se utilizaron primero y es difícil acabar con la tradición.

La gráfica de *S*

De nuevo suponga que están disponibles k muestras seleccionadas de manera independiente, cada una constituida por n observaciones de una variable con distribución normal. Denote las variaciones estándar muestrales mediante $s_1, s_2, \ldots, s_k$, con $\bar{s} = \sum s_i/k$. Los valores $s_1, s_2, s_3, \ldots$ se trazan en secuencia en una diagrama de S. La recta central de la gráfica estará a la altura $\bar{s}$, y los límites 3-sigma necesitan determinar $3\sigma_S$ (del mismo modo que los límites 3-sigma de una gráfica de $\bar{X}$ requirieron $3\sigma_{\bar{X}} = 3\sigma/\sqrt{n}$, con sigma estimada de los datos).

Recuerde que para cualquier va Y, $V(Y) = E(Y^2) - [E(Y)]^2$, y que una varianza muestral S^2 es un estimador insesgado de σ^2, es decir, $E(S^2) = \sigma^2$. Por lo tanto,

$$V(S) = E(S^2) - [E(S)]^2 = \sigma^2 - (a_n\sigma)^2 = \sigma^2(1 - a_n^2)$$

donde los valores de a_n para $n = 3, \ldots, 8$ se tabularon en la sección anterior. Entonces la desviación estándar de S es

$$\sigma_S = \sqrt{V(S)} = \sigma\sqrt{1 - a_n^2}$$

Es natural estimar σ por medio de $s_1, \ldots, s_k$ como se hizo en la sección anterior, a saber, $\hat{\sigma} = \bar{s}/a_n$. Al sustituir $\hat{\sigma}$ por σ en la expresión para σ_S se obtiene la cantidad usada para calcular los límites 3-sigma.

Los límites de control 3-sigma para una gráfica de control S son

$$\text{LCI} = \bar{s} - 3\bar{s}\sqrt{1 - a_n^2}/a_n$$
$$\text{LCS} = \bar{s} + 3\bar{s}\sqrt{1 - a_n^2}/a_n$$

La expresión para LCI será negativa si $n \leq 5$, en cuyo caso se acostumbra usar LCI = 0.

Ejemplo 16.4 En la tabla 16.2 se muestran observaciones sobre la resistencia a la deformación de láminas de plástico (la fuerza, en lb/pulg^2, necesaria para romper una lámina). Hay $k = 22$ muestras, obtenidas en puntos del tiempo igualmente espaciados, y $n = 4$ observaciones en cada muestra. Es sencillo comprobar que $\sum s_i = 51.10$ y $\bar{s} = 2.32$, así que el centro de la gráfica de S estará en 2.32 (aunque como $n = 4$, LCI = 0 y la recta central no estará a la mitad entre los límites de control). De la sección anterior, $a_4 = 0.921$, de donce el LCS es

$$\begin{aligned}\text{LCS} &= 2.32 + 3(2.32)(\sqrt{1 - (0.921)^2})/0.921 \\ &= 2.32 + 2.94 \\ &= 5.26\end{aligned}$$

Tabla 16.2 Datos de resistencia a la deformación para el ejemplo 16.4

Núm. de muestra	Observaciones				DE	Rango
1	29.7	29.0	28.8	30.2	0.64	1.4
2	32.2	29.3	32.2	32.9	1.60	3.6
3	35.9	29.1	32.1	31.3	2.83	6.8
4	28.8	27.2	28.5	35.7	3.83	8.5
5	30.9	32.6	28.3	28.3	2.11	4.3
6	30.6	34.3	34.8	26.3	3.94	8.5
7	32.3	27.7	30.9	27.8	2.30	4.6
8	32.0	27.9	31.0	30.8	1.76	4.1
9	24.2	27.5	28.5	31.1	2.85	6.9
10	33.7	24.4	34.3	31.0	4.53	9.9
11	35.3	33.2	31.4	28.0	3.09	7.3

(continúa)

Núm. de muestra	Observaciones				DE	Rango
12	28.1	34.0	31.0	30.8	2.41	5.9
13	28.7	28.9	25.8	29.7	1.71	3.9
14	29.0	33.0	30.2	30.1	1.71	4.0
15	33.5	32.6	33.6	29.2	2.07	4.4
16	26.9	27.3	32.1	28.5	2.37	5.2
17	30.4	29.6	31.0	33.8	1.83	4.2
18	29.0	28.9	31.8	26.7	2.09	5.1
19	33.8	30.9	31.7	28.2	2.32	5.6
20	29.7	27.9	29.1	30.1	0.96	2.2
21	27.9	27.7	30.2	32.9	2.43	5.2
22	30.0	31.4	27.7	28.1	1.72	3.7

La gráfica de control resultante se muestra en la figura 16.3. Los puntos trazados están dentro de los límites de control, lo cual indica un comportamiento estable con respecto a la variación.

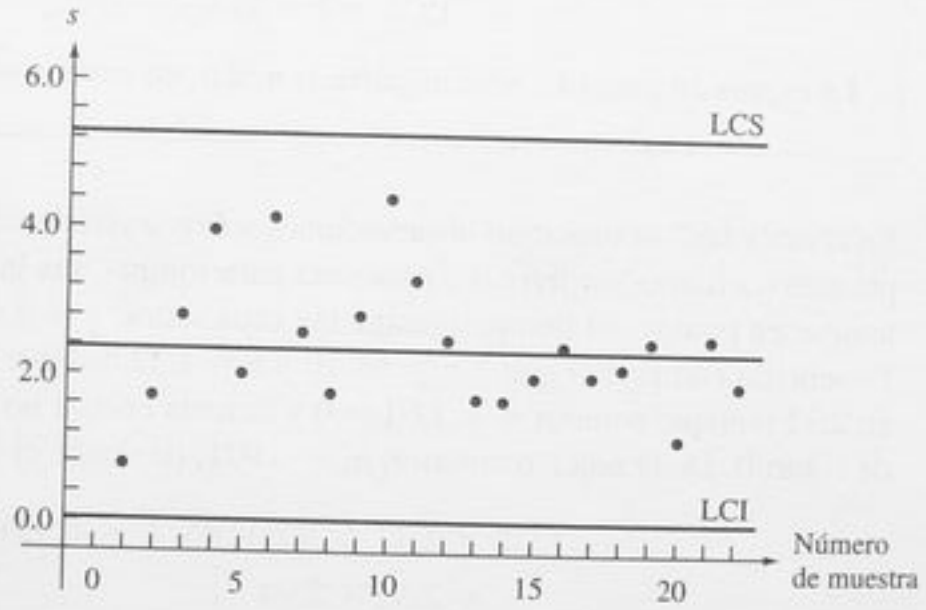


Figura 16.3 Gráfica de S para los datos de resistencia del ejemplo 16.4 ■

Gráfica de *R*

Sean $r_1, r_2, \ldots, r_k$ los k rangos muestrales y $\bar{r} = \sum r_i/k$. La recta central de una gráfica de R se encuentra a la derecha de $\bar{r}$. La determinación de los límites de control requiere σ_R, donde R denota el rango (antes de hacer las observaciones, como una variable aleatoria) de una muestra aleatoria de tamaño n de una distribución normal con valor medio μ y desviación estándar σ. Debido a que,

$$\begin{aligned} R &= \text{máx}(X_1, \ldots, X_n) - \text{mín}(X_1, \ldots, X_n) \\ &= \sigma\{\text{máx}(Z_1, \ldots, Z_n) - \text{mín}(Z_1, \ldots, Z_n)\} \end{aligned}$$

donde $Z_i = (X_i - \mu)/\sigma$, y las Z_i son variables aleatorias normales estándar, se deduce que

$$\begin{aligned} \sigma_R &= \sigma \cdot \begin{pmatrix} \text{desviación estándar del rango de una muestra aleatoria} \\ \text{de tamaño } n \text{ de una distribución normal estándar} \end{pmatrix} \\ &= \sigma \cdot c_n \end{aligned}$$

Los valores de c_n para $n = 3, \ldots, 8$ aparecen en la tabla siguiente.

n	3	4	5	6	7	8
c_n	0.888	0.880	0.864	0.848	0.833	0.820

Se acostumbra estimar σ mediante $\hat{\sigma} = \bar{r}/b_n$ como se describió en la sección anterior. Esto da $\hat{\sigma}_R = c_n\bar{r}/b_n$ como la desviación estándar estimada de R.

Los límites 3-sigma para una gráfica de R son

$$\text{LCI} = \bar{r} - 3c_n\bar{r}/b_n$$

$$\text{LCS} = \bar{r} + 3c_n\bar{r}/b_n$$

La expresión para LCI es negativa si $n \leq 6$, en cuyo caso se debe usar LCI = 0.

Ejemplo 16.5 (Continuación del ejemplo 16.4)

En la tabla 16.2 se encuentra que $\sum r_i = 115.3$ y $\bar{r} = 5.24$. Puesto que $n = 4$, LCI = 0. Con $b_4 = 2.058$ y $c_4 = 0.880$,

$$\text{LCS} = 5.24 + 3(0.880)(5.24)/2.058 = 11.96$$

La gráfica de R se muestra en la figura 16.4. Al igual que con la gráfica de S, los puntos están entre los límites, lo cual indica que un proceso está bajo control en lo que respecta a la variación.

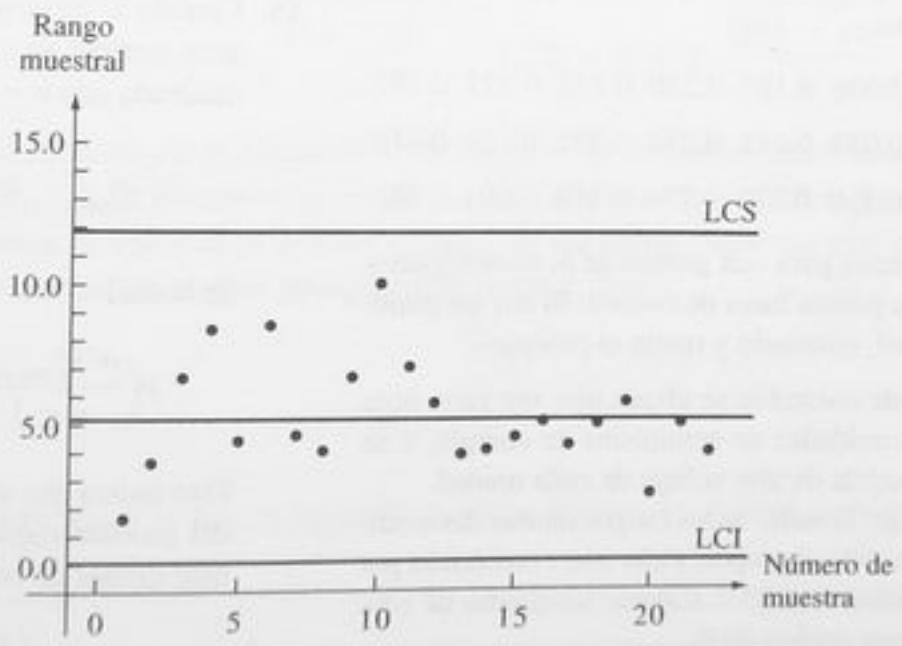


Figura 16.4 Gráfica de R para los datos de resistencia a la deformación del ejemplo 16.5 ■

Gráficas basadas en límites de probabilidad

Considere una gráfica de $\bar{X}$ basada en el valor (objetivo) bajo control μ_0 y σ conocida. Cuando la variable de interés tiene una distribución normal y el proceso está bajo control,

$$P(\bar{X}_i > \mu_0 + 3\sigma/\sqrt{n}) = 0.0013 = P(\bar{X}_i < \mu_0 - 3\sigma/\sqrt{n})$$

Es decir, la probabilidad de que un punto sobre la gráfica caiga arriba del LCS es 0.0013, al igual que la probabilidad de que caiga debajo del LCI (con 3.09 en lugar de 3, se obtiene 0.001 para cada probabilidad). Cuando los límites de control se basan en estimaciones de μ y σ, estas probabilidades serán más o menos correctas siempre que n no sea demasiado pequeña y k sea por lo menos 20.

Por el contrario, *no* es el caso para una gráfica de S 3-sigma que $P(S_i > \text{LCS}) = P(S_i < \text{LCI}) = 0.0013$, y tampoco se cumple para una gráfica de R 3-sigma que $P(R_i > \text{LCS}) = P(R_i < \text{LCI}) = 0.0013$. Esto se debe a que ni la desviación estándar muestral S ni el rango muestral R tienen una distribución normal estándar cuando la distribución poblacional es normal. En cambio, S y R tienen distribuciones sesgadas. Lo mejor que se puede decir para gráficas 3-sigma S y R es que un proceso bajo control tiene muy pocas probabilidades de producir un punto en algún tiempo particular que esté fuera de los límites de control. Algunos autores se han avocado al uso de límites de control para los que la "probabilidad de extralimitación" de cada límite es más o menos 0.001. El libro *Statistical Methods for Quality Improvement* (véase la bibliografía del capítulo) contiene más información acerca de este tema.

Ejercicios | Sección 16.3 (14 a 18)

14. Un fabricante de tizas que no desprenden polvo inició un programa de control de calidad para supervisar la densidad de la tiza. A continuación se muestran las desviaciones estándar muestrales de densidades para 24 subgrupos diferentes, cada uno formado por $n = 8$ especímenes de tiza:

0.204 0.315 0.096 0.184 0.230 0.212 0.322 0.287
0.145 0.211 0.053 0.145 0.272 0.351 0.159 0.214
0.388 0.187 0.150 0.229 0.276 0.118 0.091 0.056

Calcule los límites para una gráfica de S, constrúyala y compruebe los puntos fuera de control. Si hay un punto fuera de control, elimínelo y repita el proceso.

15. De una línea de ensamble se eligen una vez cada hora subgrupos de unidades de suministro de energía, y se determina la salida de alto voltaje de cada unidad.

a. Suponga que la suma de los rangos muestrales resultantes para 30 subgrupos, cada uno constituido por cuatro unidades, es 85.2. Calcule los límites de control para una gráfica de R.

b. Repita el inciso (a) si cada subgrupo consta de ocho unidades y la suma es de 106.2.

16. Calcule límites de control para una gráfica de S y una de R por medio de los datos de contenido de humedad del ejercicio 6. Después, compruebe la presencia de alguna señal fuera de control.

17. Calcule los límites de control para una gráfica de S de los datos de índice de refracción del ejercicio 9. ¿Al parecer el proceso está bajo control con respecto a la variabilidad? ¿Por qué sí o por qué no?

18. Cuando S^2 es la varianza muestral de una muestra aleatoria normal $(n-1)S^2/\sigma^2$ tiene una distribución chi-cuadrada con $n-1$ grados de libertad, por lo tanto

$$P\left(\chi^2_{0.999,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{0.001,n-1}\right) = 0.998$$

de la cual

$$P\left(\frac{\sigma^2\chi^2_{0.999,n-1}}{n-1} < S^2 < \frac{\sigma^2\chi^2_{0.001,n-1}}{n-1}\right) = 0.998$$

Esto indica que otra gráfica para controlar la variación del proceso requiere trazar las varianzas muestrales y usar límites de control

$$\text{LCI} = \bar{s}^2\chi^2_{0.999,n-1}/(n-1)$$
$$\text{LCS} = \bar{s}^2\chi^2_{0.001,n-1}/(n-1)$$

Construya la gráfica correspondiente para los datos del ejercicio 9. [*Sugerencia:* los valores críticos chi-cuadrada de extremo inferior y superior para 5 grados de libertad son 0.210 y 20.515, respectivamente.]

16.4 Gráficas de control para atributos

El término *datos de atributo* se emplea en las publicaciones de control de calidad para describir dos situaciones:

1. Cada elemento producido es defectuoso o no defectuoso (cumple con las especificaciones o no).
2. Un solo elemento podría tener uno o más defectos, y se determina el número de defectos.

En el primer caso, una gráfica control se basa en la distribución binomial; en el último caso, la distribución de Poisson es la base para una gráfica.

Gráfica *p* para la fracción de artículos defectuosos

Suponga que cuando un proceso está bajo control, la probabilidad de que algún elemento particular esté defectuoso es p (de forma equivalente, p es la proporción a largo plazo de artículos defectuosos para un proceso bajo control) y que diferentes artículos son independientes entre sí con respecto a sus condiciones. Considere una muestra de n artículos obtenidos en un determinado momento, y sea X el número de artículos defectuosos y $\hat{p} = X/n$. Debido a que X tiene una distribución binomial, $E(X) = np$ y $V(X) = np(1 - p)$, por lo tanto

$$E(\hat{p}) = p \qquad V(\hat{p}) = \frac{p(1-p)}{n}$$

También, si $np \geq 10$ y $n(1 - p) \geq 10$, $\hat{p}$ tiene aproximadamente una distribución normal.

En el caso de p desconocida (o una gráfica basada en un valor objetivo), los límites de control son

$$\text{LCI} = p - 3\sqrt{\frac{p(1-p)}{n}} \qquad \text{LCS} = p + 3\sqrt{\frac{p(1-p)}{n}}$$

Si cada muestra consiste en n elementos, el número de artículos defectuosos en la i-ésima muestra es x_i, y $\hat{p}_i = x_i/n$, entonces $\hat{p}_1, \hat{p}_2, \hat{p}_3, \ldots$ se dibujan en la gráfica de control.

Por lo común, el valor de p se debe estimar de los datos. Suponga que están disponibles k muestras de lo que se cree es un proceso bajo control, y sea

$$\bar{p} = \frac{\sum_{i=1}^{k} \hat{p}_i}{k}$$

Entonces, la estimación $\bar{p}$ se usa en lugar de p en los límites de control antes mencionados.

La grafica p para la fracción de artículos defectuosos tiene su línea central a la altura $\bar{p}$ y límites de control

$$\text{LCI} = \bar{p} - 3\sqrt{\frac{p(1-\bar{p})}{n}}$$

$$\text{LCS} = \bar{p} + 3\sqrt{\frac{p(1-\bar{p})}{n}}$$

Si el LCI es negativo, se sustituye por 0.

Ejemplo 16.6 Se selecciona una muestra de 100 tazas de un determinado modelo de vajilla en cada uno de 25 días sucesivos, y cada una se examina en busca de defectos. Los números resultantes de tazas inaceptables y las proporciones muestrales correspondientes son como se muestra a continuación:

Día (i)	1	2	3	4	5	6	7	8	9	10	11	12	13
x_i	7	4	3	6	4	9	6	7	5	3	7	8	4
$\hat{p}_i$	0.07	0.04	0.03	0.06	0.04	0.09	0.06	0.07	0.05	0.03	0.07	0.08	0.04

Día (i)	14	15	16	17	18	19	20	21	22	23	24	25
x_i	6	2	9	7	6	7	11	6	7	4	8	6
$\hat{p}_i$	0.06	0.02	0.09	0.07	0.06	0.07	0.11	0.06	0.07	0.04	0.08	0.06

Suponiendo que el proceso estuvo bajo control durante este periodo, se establecen límites de control y se construye una grafica p. Se tiene que $\sum \hat{p}_i = 1.52$, así que $\bar{p} = 1.52/25 = 0.0608$ y

$$\text{LCI} = 0.0608 - 3\sqrt{(0.0608)(0.9392)/100} = 0.0608 - 0.0717 = -0.0109$$
$$\text{LCS} = 0.0608 + 3\sqrt{(0.0608)(0.9392)/100} = 0.0608 + 0.0717 = 0.1325$$

Por lo tanto, el LCI se fija en 0. La gráfica de control ilustrada en la figura 16.5 muestra que los puntos están dentro de los límites de control. Esto es coherente con la suposición de un proceso bajo control.

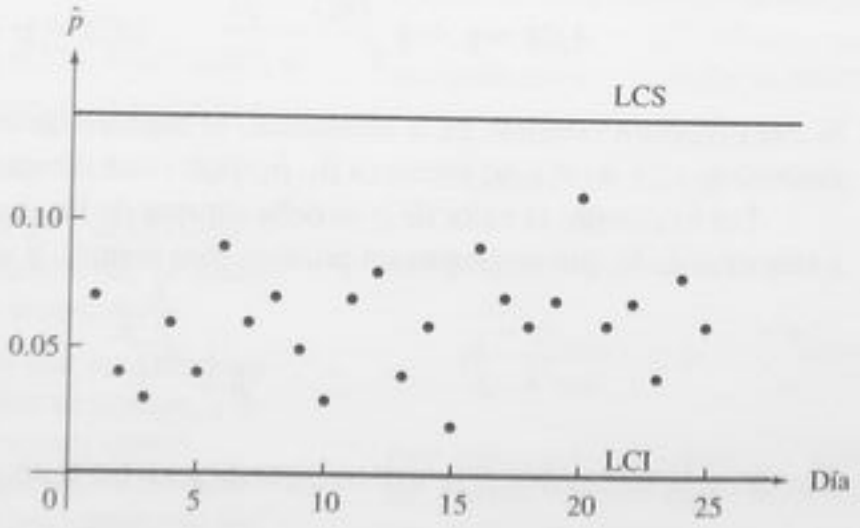


Figura 16.5 Gráfica de control para los datos de la fracción de artículos defectuosos del ejemplo 16.6 ■

Gráfica *c* para el número de defectos

Ahora se consideran situaciones en las que la observación en cada punto del tiempo es el número de defectos en una unidad de algún tipo. La unidad podría consistir en un solo artículo (p. ej., un automóvil) o un grupo de artículos (p. ej., imperfecciones en un grupo de cuatro neumáticos). En el segundo caso, se supone que el tamaño de grupo es el mismo en cada punto.

La gráfica de control para el número de defectos se basa en la distribución de probabilidad de Poisson. Recuerde que si Y es una variable aleatoria de Poisson con parámetro θ, entonces

$$E(Y) = \theta \qquad V(Y) = \theta \qquad \sigma_Y = \sqrt{\theta}$$

Asimismo, Y tiene aproximadamente una distribución normal cuando θ es grande ($\theta \geq 10$ será suficiente para la mayoría de los propósitos). Además, si $Y_1, Y_2, \ldots, Y_n$ son variables de Poisson independientes con parámetros $\theta_1, \theta_2, \ldots, \theta_n$, se puede demostrar que $Y_1 + \cdots + Y_n$ tiene una distribución de Poisson con parámetro $\theta_1 + \cdots + \theta_n$. En particular, si $\theta_1 = \cdots = \theta_n = \theta$ (la distribución del número de defectos por artículo es el mismo para cada artículo), entonces el parámetro de Poisson es $\lambda = n\theta$.

Sea que λ denota el parámetro de Poisson para el número de defectos en una unidad (es el número esperado de defectos por unidad). En el caso de λ conocida (o una gráfica basada en un valor objetivo),

$$\text{LCI} = \lambda - 3\sqrt{\lambda} \qquad \text{LCS} = \lambda + 3\sqrt{\lambda}$$

Con x_i que denota el número total de defectos en la i-ésima unidad ($i = 1, 2, 3, \ldots$), entonces los puntos en las alturas $x_1, x_2, x_3, \ldots$ se bosquejan en la gráfica. Por lo común, el valor de λ se debe estimar de los datos. Puesto que $E(X_i) = \lambda$, es natural usar la estimación $\hat{\lambda} = \bar{x}$ (basada en $x_1, x_2, \ldots, x_k$).

La gráfica de c para el número de defectos en una unidad tiene una línea central a la altura $\bar{x}$ y

$$\text{LCI} = \bar{x} - 3\sqrt{\bar{x}}$$
$$\text{LCS} = \bar{x} + 3\sqrt{\bar{x}}$$

Si LCI es negativo, se reemplaza por 0.

Ejemplo 16.7 Una compañía fabrica paneles de metal, los cuales son horneados después de ser recubiertos primero con una lechada de cerámica en polvo. A veces aparecen defectos en el terminado de estos paneles y la compañía desea establecer una gráfica de control para el número de defectos. A continuación se muestra la cantidad de defectos en cada uno de los 24 paneles muestreados en intervalos de tiempo regulares:

7 10 9 12 13 6 13 7 5 11 8 10
13 9 21 10 6 8 3 12 7 11 14 10

con $\sum x_i = 235$ y $\hat{\lambda} = \bar{x} = 235/24 = 9.79$. Los límites de control son

$$\text{LCI} = 9.79 - 3\sqrt{9.79} = 0.40 \qquad \text{LCS} = 9.79 + 3\sqrt{9.79} = 19.18$$

La gráfica de control se ilustra en la figura 16.6. El punto que corresponde al panel decimoquinto se ubica arriba del LCS. Al realizar una investigación se encontró que la lechada utilizada en ese panel tenía una viscosidad inusualmente baja (una causa asignable). Al eliminar esa observación del conjunto de datos, se obtiene $\bar{x} = 214/23 = 9.30$ y los nuevos límites de control

$$\text{LCI} = 9.30 - 3\sqrt{9.30} = 0.15 \qquad \text{LCS} = 9.30 + 3\sqrt{9.30} = 18.45$$

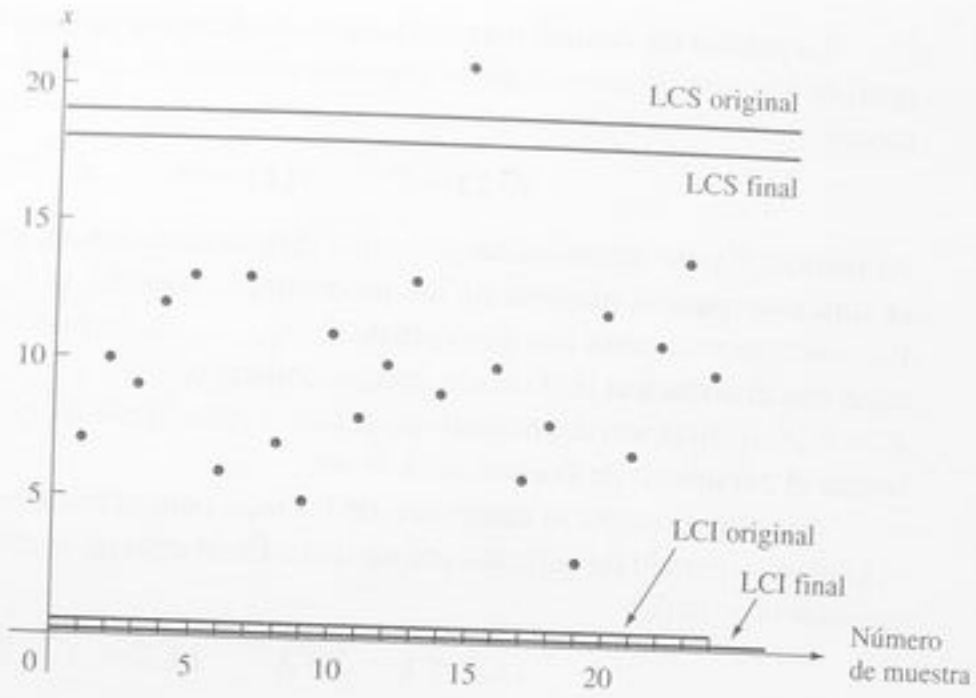


Figura 16.6 Gráfica de control para los datos de número de defectos en un proceso bajo control 16.7

Las 23 observaciones restantes están entre estos límites, lo cual indica que un proceso está bajo control. ■

Gráficas de control basadas en datos transformados

Se presume que el uso de límites de control 3-sigma da como resultado $P(\text{estadístico} < \text{LCI}) \approx P(\text{estadístico} > \text{LCS}) \approx 0.0013$ cuando el proceso está bajo control. Sin embargo, cuando p es pequeña, la aproximación normal a la distribución $\hat{p} = X/n$ suele no ser muy precisa en la parte final de los extremos. En la tabla 16.3 se muestra evidencia de este comportamiento para valores seleccionados de p y n (el valor de p se emplea para calcular los límites de control). En muchos casos, la probabilidad de que un solo punto caiga fuera de los límites de control es muy diferente de la probabilidad nominal de 0.0026.

Tabla 16.3 Probabilidades bajo control para una grafica p

p	n	$P(\hat{p} < \text{LCI})$	$P(\hat{p} > \text{LCS})$	P(punto fuera de control)
0.10	100	0.00003	0.00198	0.00201
0.10	200	0.00048	0.00299	0.00347
0.10	400	0.00044	0.00171	0.00215
0.05	200	0.00004	0.00266	0.00270
0.05	400	0.00020	0.00207	0.00227
0.05	600	0.00031	0.00189	0.00220
0.02	600	0.00007	0.00275	0.00282
0.02	800	0.00036	0.00374	0.00410
0.02	1000	0.00023	0.00243	0.00266

Este problema se remedia si se aplica una transformación a los datos. Sea $h(X)$ una función aplicada para transformar la variable binomial X. Entonces se debe elegir $h(\cdot)$ de modo que $h(X)$ tenga aproximadamente una distribución normal y esta aproximación sea precisa en los extremos. Una transformación recomendada se basa en la función arco seno (es decir, sen^{-1}):

$$Y = h(X) = \text{sen}^{-1}(\sqrt{X/n})$$

Entonces Y es aproximadamente normal con valor medio $\text{sen}^{-1}(\sqrt{p})$ y varianza $1/(4n)$; observe que la varianza es independiente de p. Sea $y_i = \text{sen}^{-1}(\sqrt{x_i/n})$. Luego, los puntos en la gráfica de control se trazan en las alturas $y_1, y_2, \ldots$. Para n conocida, los límites de control son

$$\text{LCI} = \text{sen}^{-1}(\sqrt{p}) - 3\sqrt{1/(4n)} \qquad \text{LCS} = \text{sen}^{-1}(\sqrt{p}) + 3\sqrt{1/(4n)}$$

Cuando se desconoce el valor de p, $\text{sen}^{-1}(\sqrt{p})$ se sustituye por $\bar{y}$.

Comentarios similares se aplican a la distribución de Poisson cuando λ es pequeña. La transformación sugerida es $Y = h(X) = 2\sqrt{X}$, que tiene valor medio $2\sqrt{\lambda}$ y varianza 1. Los límites de control resultantes son $2\sqrt{\lambda} \pm 3$ cuando se conoce el valor de λ y $\bar{y} \pm 3$ en caso contrario. En el libro *Statistical Methods for Quality Improvement* listado en la bibliografía del capítulo se analizan estos temas con mayor detalle.

Ejercicios | Sección 16.4 (19 a 26)

19. En cada uno de los 25 días anteriores se seleccionaron al azar 100 dispositivos electrónicos de cierto tipo y se sometieron a pruebas rigurosas de resistencia al calor. El número total de piezas que no pasaron la prueba fue 578.
    a. Determine los límites de control para una grafica p 3-sigma.
    b. El número máximo de piezas que no pasaron la prueba en un determinado día fue 39, y el mínimo 13. ¿Alguno de éstos corresponde a un punto fuera de control? Explique.

20. Se seleccionó una muestra de 200 chips de computadora ROM en cada uno de 30 días consecutivos, y el número de estas piezas que incumplió con las especificaciones en cada día fue como sigue: 10, 18, 24, 17, 37, 19, 7, 25, 11, 24, 29, 15, 16, 21, 18, 17, 15, 22, 12, 20, 17, 18, 12, 24, 30, 16, 11, 20, 14, 28. Construya una gráfica p y examínela en busca de puntos fuera de control.

21. Cuando $n = 150$, ¿cuál es el valor mínimo de $\bar{p}$ para el cual el LCI de una gráfica p es positivo?

22. Refiérase a los datos del ejercicio 20 y construya una gráfica de control por medio de la transformación sen^{-1} como se sugiere en el texto.

23. Las siguientes observaciones son la cantidad de defectos en 25 especímenes de una yarda cuadrada de tejido de cierto tipo: 3, 7, 5, 3, 4, 2, 8, 4, 3, 3, 6, 7, 2, 3, 2, 4, 7, 3, 2, 4, 4, 1, 5, 4, 6. Construya una gráfica c para el número de defectos.

24. ¿Para qué valores de $\bar{x}$ será negativo el LCI en una gráfica c?

25. En algunas situaciones, los tamaños de especímenes muestreados varían, y se espera que los especímenes más grandes tengan más defectos que los más pequeños. Por ejemplo, los tamaños de muestras de tejidos inspeccionadas en busca de fallas, pueden variar con el tiempo. Por otro lado, el número de piezas inspeccionadas podría cambiar con el tiempo. Sea

$$u_i = \frac{\text{número de defectos observados en el tiempo } i}{\text{tamaño de entidad inspeccionada en el tiempo}} = \frac{x_i}{g_i}$$

donde "tamaño" podría referirse al área, longitud, volumen o, simplemente, al número de piezas inspeccionadas. Entonces, en una **gráfica u** se trazan los valores $u_1, u_2, \ldots$, tiene línea central $\bar{u}$, y los límites de control para las i-ésimas observaciones son $\bar{u} \pm 3\sqrt{\bar{u}/g_i}$.

Se examinaron paneles pintados en tiempos sucesivos, y para cada uno se determinó el número de defectos en una región de muestreo especificada. El área superficial (pies^2) de la región examinada varió de un panel a otro. Construya una gráfica u.

Panel	Área examinada	No. de defectos
1	0.8	3
2	0.6	2
3	0.8	3
4	0.8	2
5	1.0	5
6	1.0	5
7	0.8	10
8	1.0	12
9	0.6	4
10	0.6	2
11	0.6	1
12	0.8	3
13	0.8	5
14	1.0	4
15	1.0	6
16	1.0	12
17	0.8	3
18	0.6	3
19	0.6	5
20	0.6	1

26. Construya una gráfica de control para los datos del ejercicio 23 mediante la transformación sugerida en el texto.

16.5 Procedimientos CUSUM

Un defecto de la gráfica X usual es su incapacidad para detectar un cambio relativamente pequeño en una media de proceso. Esto en gran medida es una consecuencia del hecho de que si un proceso se juzga fuera de control en determinado momento, depende sólo de la muestra en ese instante y no de lo sucedido antes con el proceso. Para remediar este defecto, se han diseñado gráficas de control y procedimientos de **suma acumulada (CUSUM)**.

Hay dos versiones equivalentes de un procedimiento CUSUM para una media de proceso, uno gráfico y el otro computacional. La versión computacional se utiliza casi exclusivamente en la práctica, pero la base lógica del procedimiento se comprende con más facilidad si primero se considera la forma gráfica.

Máscara V

Denote con μ_0 un valor objetivo para la media del proceso, y defina *sumas acumuladas* mediante

$$S_1 = \bar{x}_1 - \mu_0$$
$$S_2 = (\bar{x}_1 - \mu_0) + (\bar{x}_2 - \mu_0) = \sum_{i=1}^{2} (\bar{x}_i - \mu_0)$$
$$\vdots$$
$$S_l = (\bar{x}_1 - \mu_0) + \cdots + (\bar{x}_l - \mu_0) = \sum_{i=1}^{l} (\bar{x}_i - \mu_0)$$

(en ausencia de un valor objetivo, $\bar{\bar{x}}$ se usa en lugar de μ_0). Estas sumas acumuladas se trazan en función del tiempo. Es decir, en el instante l, se traza un punto a la altura S_l. En el punto de tiempo actual r, los puntos trazados son $(1, S_1), (2, S_2), (3, S_3), \ldots, (r, S_r)$.

Ahora se superpone una "máscara" en forma de V sobre la gráfica, como se muestra en la figura 16.7. El punto 0, que está a una distancia d detrás del punto en el que se intersecan los dos brazos de la máscara, se ubica en el punto CUSUM actual (r, S_r). En el tiempo r, se juzga que el proceso está fuera de control si alguno de los puntos trazados queda fuera de la máscara V, ya sea arriba del brazo superior o abajo del brazo inferior. Cuando el proceso está bajo control, las $\bar{x}_i$ varían en torno al valor objetivo μ_0, así que las S_i sucesivas deben variar alrededor de

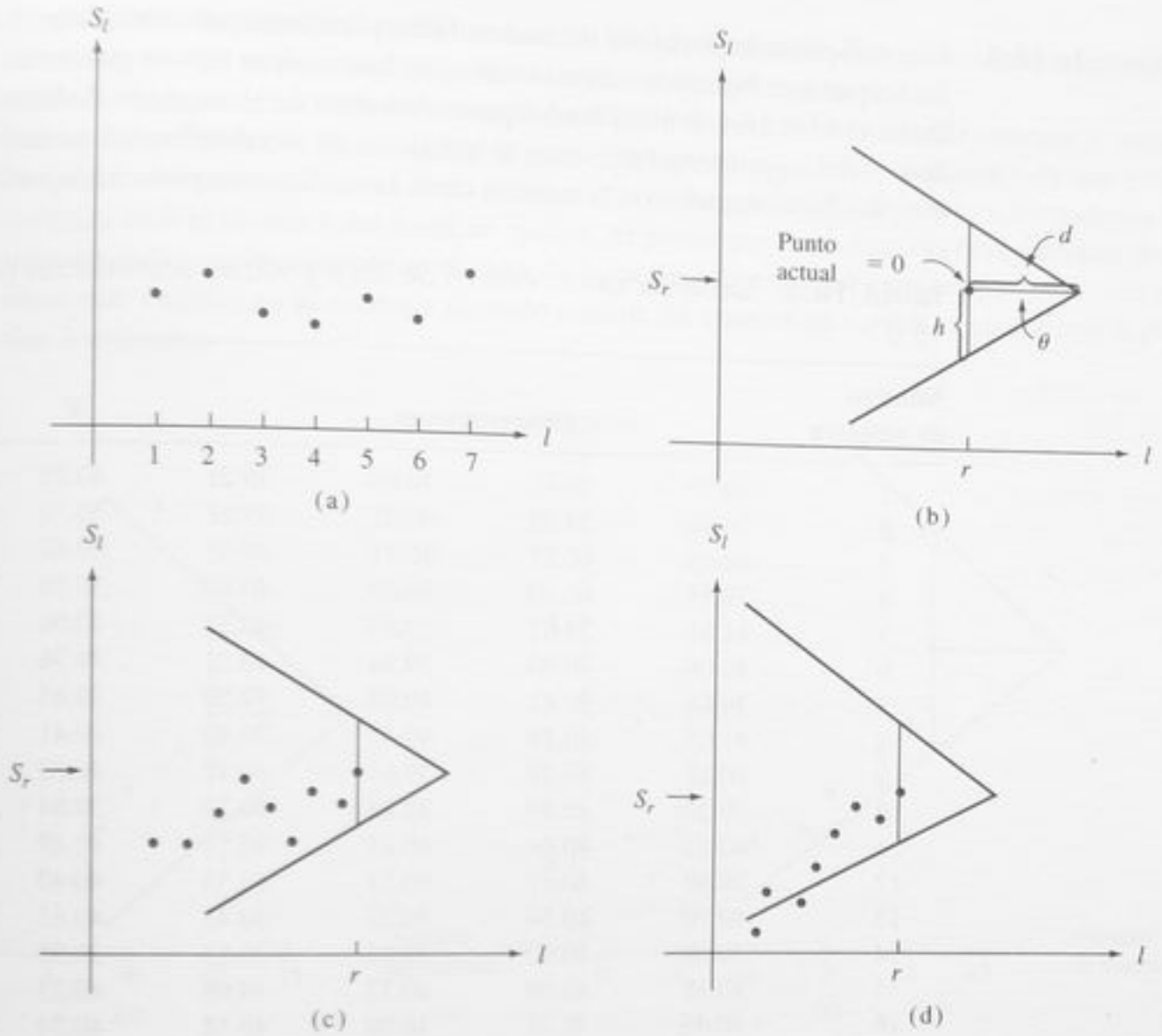


Figura 16.7 Gráficas CUSUM: (a) puntos sucesivos (l, S_l) en una gráfica CUSUM; (b) Una máscara en V con $0 = (r, S_r)$; (c) un proceso bajo control; (d) un proceso fuera de control

0. Sin embargo, suponga que en determinado momento, la media del proceso cambia a un valor más grande que el objetivo. De ese punto en adelante, las diferencias $\bar{x}_i - \mu_0$ tienden a ser positivas, de modo que las S_l sucesivas aumentan y los puntos trazados se desplazan hacia arriba. Si ha ocurrido un cambio antes del punto de tiempo actual r, hay una buena posibilidad de que (r, S_r) sea sustancialmente mayor que algunos otros puntos en la gráfica, en cuyo caso estos otros puntos estarán abajo del brazo inferior de la máscara. De manera similar, un cambio a un valor más pequeño que el objetivo después da como resultado puntos arriba del brazo superior de la máscara.

Cualquier máscara en forma de V se determina al especificar la "distancia directora" d y "medio ángulo" θ, o equivalente especificando d y la longitud h del segmento de recta vertical de 0 al brazo inferior (o al superior) de la máscara. Un método para decidir cuál máscara usar, requiere especificar el tamaño de un cambio en la media del proceso que es de particular interés para el investigador. Entonces se eligen los parámetros de la máscara para obtener los valores deseados de α y β, la probabilidad de falsa alarma y la probabilidad de no detectar el cambio especificado, respectivamente. En otro método, es necesario seleccionar la máscara que produce valores especificados de la ARL (duración de corrida promedio) tanto para un proceso bajo control como para uno en el que la media ha cambiado por una cantidad designada. Después de desarrollar la forma computacional del procedimiento CUSUM, se ilustra el segundo método de construcción.

Ejemplo 16.8 Una compañía de productos de madera fabrica briquetas para parrilladas. La compañía empaca las briquetas en bolsas de varios tamaños, de las cuales se supone que la más grande contiene 40 lb. En la tabla 16.4 se muestran los pesos de bolsas de 16 muestras distintas, cada una de tamaño $n = 4$. Las primeras 10 de éstas se extrajeron de una distribución normal con $\mu = \mu_0 = 40$ y $\sigma = 0.5$. Comenzando con la muestra once, la media se desplaza hacia arriba hasta $\mu = 40.3$.

Tabla 16.4 Observaciones, valores de las $\bar{x}$ y sumas acumuladas para el ejemplo 16.8

Número de muestra	Observaciones				$\bar{x}$	$\sum(\bar{x}_i - 40)$
1	40.77	39.95	40.86	39.21	40.20	0.20
2	38.94	39.70	40.37	39.88	39.72	−0.08
3	40.43	40.27	40.91	40.05	40.42	0.34
4	39.55	40.10	39.39	40.89	39.98	0.32
5	41.01	39.07	39.85	40.32	40.06	0.38
6	39.06	39.90	39.84	40.22	39.76	0.14
7	39.63	39.42	40.04	39.50	39.65	−0.21
8	41.05	40.74	40.43	39.40	40.41	0.20
9	40.28	40.89	39.61	40.48	40.32	0.52
10	39.28	40.49	38.88	40.72	39.84	0.36
11	40.57	40.04	40.85	40.51	40.49	0.85
12	39.90	40.67	40.51	40.53	40.40	1.25
13	40.70	40.54	40.73	40.45	40.61	1.86
14	39.58	40.90	39.62	39.83	39.98	1.84
15	40.16	40.69	40.37	39.69	40.23	2.07
16	40.46	40.21	40.09	40.58	40.34	2.41

En la figura 16.8 se muestra una gráfica de $\bar{X}$ con línea central a la altura 40 y límites de control en

$$\mu_0 \pm 3\sigma_{\bar{X}} = 40 \pm 3 \cdot (0.5/\sqrt{4}) = 40 \pm 0.75$$

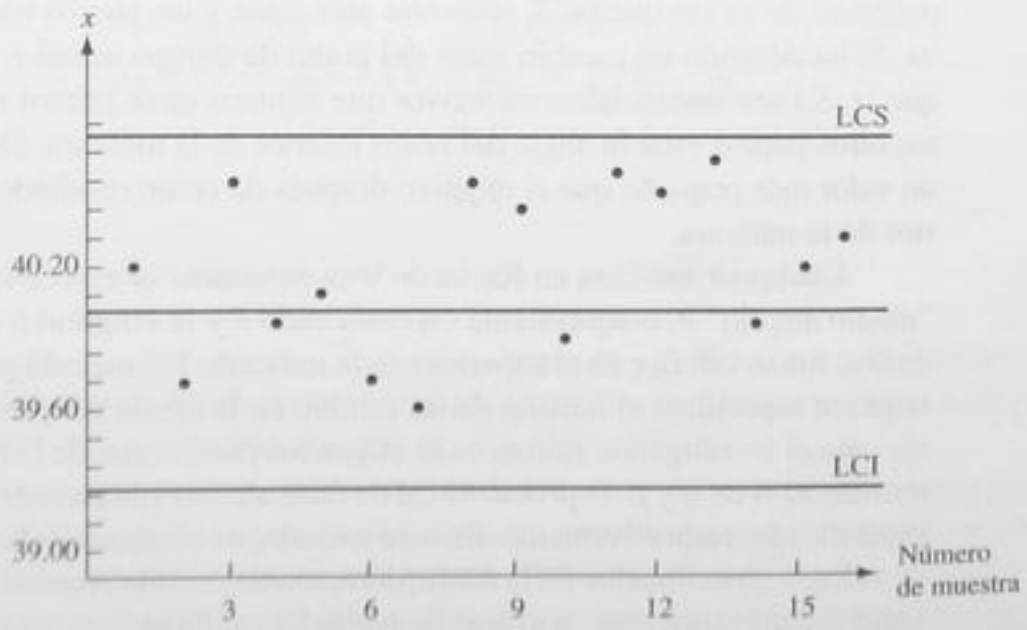


Figura 16.8 Gráfica de control de $\bar{X}$ para los datos del ejemplo 16.8

Ningún punto de la gráfica queda fuera de los límites de control. Esta gráfica indica un proceso estable para el cual la media ha permanecido en el objetivo.

En la figura 16.9 se muestran gráficas CUSUM con una determinada máscara V superpuesta. La gráfica de la figura 16.9(a) es para el tiempo actual $r = 12$. Los puntos de esta gráfica quedan dentro de los brazos de la máscara. Sin embargo, la gráfica para $r = 13$ mostrada en la figura 16.9(b) da una señal fuera de control. El punto que cae debajo del brazo inferior de la máscara indica un incremento en el valor de la media del proceso. La máscara en $r = 16$ es incluso más enfática en su mensaje fuera de control. Se observa un notable contraste con la gráfica $\bar{X}$ ordinaria.

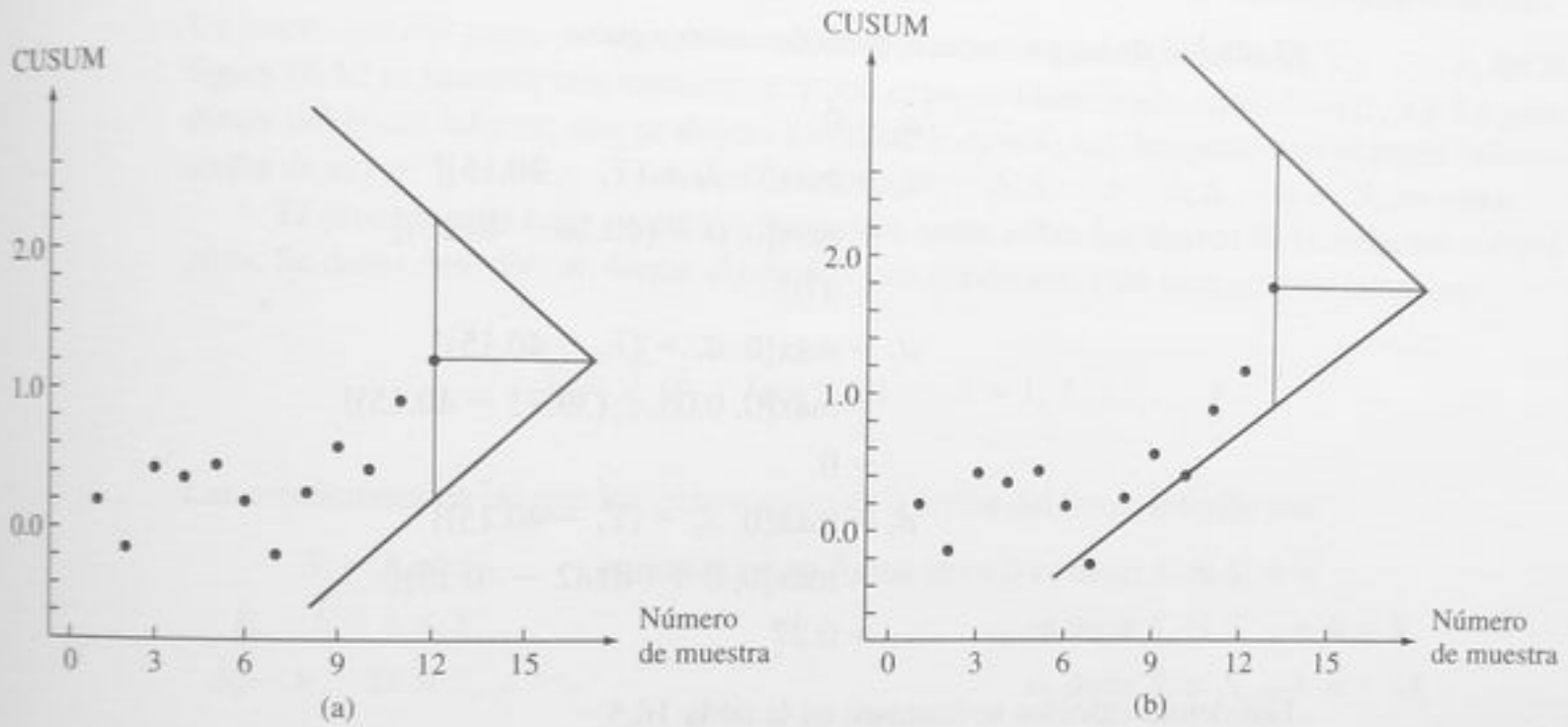


Figura 16.9 Gráficas CUSUM y máscaras en V para los datos del ejemplo 16.8: (a) Máscara en V en el tiempo $r = 12$, proceso bajo control; (b) máscara en V en el tiempo $r = 13$, señal fuera de control ■

Forma computacional para el procedimiento CUSUM

Primero se describe la versión computacional del procedimiento CUSUM, luego se muestra su equivalencia con la forma gráfica y, por último, se analiza diseñar un procedimiento para satisfacer los criterios especificados.

Sea $d_0 = e_0 = 0$, y calcule $d_1, d_2, d_3, \ldots$ y $e_1, e_2, e_3, \ldots$ de forma repetitiva por medio de las relaciones

$$d_i = \text{máx}[0, d_{i-1} + (\bar{x}_i - (\mu_0 + k))] \qquad (i = 1, 2, 3, \ldots)$$
$$e_i = \text{máx}[0, e_{i-1} - (\bar{x}_i - (\mu_0 - k))]$$

Aquí el símbolo k denota la pendiente del brazo inferior de la máscara en V y, por lo común, su valor se toma como $\Delta/2$ (donde Δ es el tamaño de un cambio en μ en el que se centra la atención).

Si en un tiempo actual r, ya sea que $d_r > h$ o $e_r > h$, se considera que el proceso está fuera de control. La primera desigualdad indica que la media del proceso se ha desplazado a un valor mayor que el objetivo, mientras que $e_r > h$ indica un cambio a un valor más pequeño.

Ejemplo 16.9 Reconsidere los datos de las briquetas mostrados en la tabla 16.4 del ejemplo 16.8. El valor objetivo es $\mu_0 = 40$, y el tamaño de un cambio que se detecta con rapidez es $\Delta = 0.3$. Así,

$$k = \frac{\Delta}{2} = 0.15 \qquad \mu_0 + k = 40.15 \qquad \mu_0 - k = 39.85$$

por lo tanto

$$d_i = \text{máx}[0, d_{i-1} + (\bar{x}_i - 40.15)]$$
$$e_i = \text{máx}[0, e_{i-1} - (\bar{x}_i - 39.85)]$$

El cálculo de las primeras d_i procede como sigue:

$$\begin{aligned} d_0 &= 0 \\ d_1 &= \text{máx}[0, d_0 + (\bar{x}_1 - 40.15)] \\ &= \text{máx}[0, 0 + (40.20 - 40.15)] \\ &= 0.05 \\ d_2 &= \text{máx}[0, d_1 + (\bar{x}_2 - 40.15)] \\ &= \text{máx}[0, 0.05 + (39.72 - 40.15)] \\ &= 0 \\ d_3 &= \text{máx}[0, d_2 + (\bar{x}_3 - 40.15)] \\ &= \text{máx}[0, 0 + (40.42 - 40.15)] \\ &= 0.27 \end{aligned}$$

Los demás cálculos se resumen en la tabla 16.5.

Tabla 16.5 Cálculos CUSUM para el ejemplo 16.9

Número de muestra	$\bar{x}_i$	$\bar{x}_i - 40.15$	d_i	$\bar{x}_i - 39.85$	e_i
1	40.20	0.05	0.05	0.35	0
2	39.72	−0.43	0	−0.13	0.13
3	40.42	0.27	0.27	0.57	0
4	39.98	−0.17	0.10	0.13	0
5	40.06	−0.09	0.01	0.21	0
6	39.76	−0.39	0	−0.09	0.09
7	39.65	−0.50	0	−0.20	0.29
8	40.41	0.26	0.26	0.56	0
9	40.32	0.17	0.43	0.47	0
10	39.84	−0.31	0.12	−0.01	0.01
11	40.49	0.34	0.46	0.64	0
12	40.40	0.25	0.71	0.55	0
13	40.61	0.46	1.17	0.76	0
14	39.98	−0.17	1.00	0.13	0
15	40.23	0.08	1.08	0.38	0
16	40.34	0.19	1.27	0.49	0

El valor $h = 0.95$ da un procedimiento CUSUM con propiedades deseables, rara vez ocurren falsas alarmas (señales incorrectas de fuera de control), sin embargo un cambio de $\Delta = 0.3$ se detecta con rapidez. Con este valor de h, la primera señal de fuera de control resulta después que está disponible la muestra 13. Puesto que $d_{13} = 1.17 > 0.95$, al parecer la media se desplazó a un valor más grande que el objetivo. Éste es el mismo mensaje que se obtiene con la máscara en V de la figura 16.9(b). ■

Equivalencia de la máscara en V y la forma computacional

De nuevo sea r el punto de tiempo actual, de modo que estén disponibles $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_r$ En la figura 16.10 se muestra una máscara en V con el punto identificado como 0 en (r, S_r). La pendiente del brazo inferior, que se denota mediante k, es h/d. Así, los puntos en el brazo inferior arriba de r, $r-1$, $r-2$, . . . están en las alturas $S_r - h$, $S_r - h - k$, $S_r - h - 2k$, etcétera.

El proceso está bajo control si los puntos están sobre los brazos de la máscara o entre ellos. Se desea describir en forma algebraica esta condición. Para realizar esta tarea, sea

$$T_l = \sum_{i=1}^{l} [\bar{x}_i - (\mu_0 + k)] \qquad l = 1, 2, 3, \ldots, r$$

Las condiciones en las que los puntos están en o arriba del brazo inferior son

$S_r - h \leq S_r$	(se satisface en forma trivial)	es decir $S_r \leq S_r + h$
$S_r - h - k \leq S_{r-1}$		es decir $S_r \leq S_{r-1} + h + k$
$S_r - h - 2k \leq S_{r-2}$		es decir $S_r \leq S_{r-2} + h + 2k$
.		.
.		.
.		.

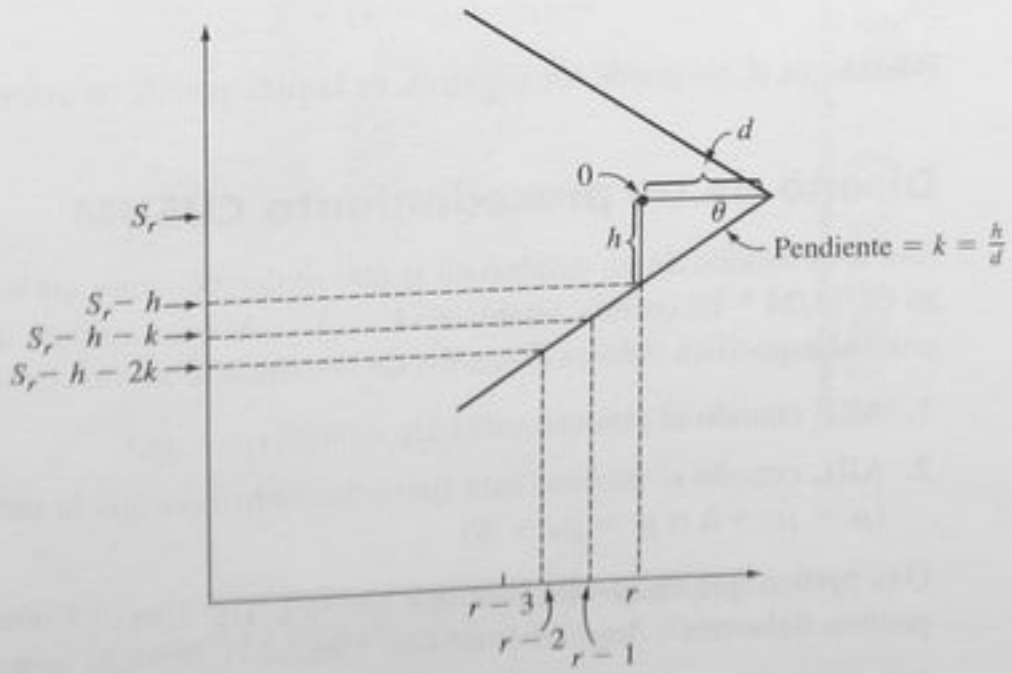


Figura 16.10 Máscara en V con pendiente del brazo inferior = k

Ahora se resta rk de ambos lados de cada desigualdad para obtener

$$\begin{aligned} S_r - rk &\le S_r - rk + h && \text{es decir } T_r \le T_r + h \\ S_r - rk &\le S_{r-1} - (r-1)k + h && \text{es decir } T_r \le T_{r-1} + h \\ S_r - rk &\le S_{r-2} - (r-2)k + h && \text{es decir } T_r \le T_{r-2} + h \\ &\vdots && \vdots \end{aligned}$$

Así, los puntos trazados quedan arriba o abajo del brazo inferior si y sólo si $T_r - T_r \le h$, $T_r - T_{r-1} \le h$, $T_r - T_{r-2} \le h$, etcétera. Esto es equivalente a

$$T_r - \text{mín}(T_1, T_2, \ldots, T_r) \le h$$

De manera similar, si se permite que

$$V_r = \sum_{i=1}^{r} [\bar{x}_i - (\mu_0 - k)] = S_r + rk$$

se puede demostrar que los puntos se ubican en o abajo del brazo superior si y sólo si

$$\text{máx}(V_1, \ldots, V_r) - V_r \le h$$

Si ahora se permite que

$$\begin{aligned} d_r &= T_r - \text{mín}(T_1, \ldots, T_r) \\ e_r &= \text{máx}(V_1, \ldots, V_r) - V_r \end{aligned}$$

se observa que $d_1, d_2, \ldots$ y $e_1, e_2, \ldots$ se calculan en forma recursiva como se ilustró antes. Por ejemplo, la expresión para d_r se deduce de la consideración de dos casos:

1. $\text{mín}(T_1, \ldots, T_r) = T_r$, de donde $d_r = 0$
2. $\text{mín}(T_1, \ldots, T_r) = \text{mín}(T_1, \ldots, T_{r-1})$, de modo que

$$\begin{aligned} d_r &= T_r - \text{mín}(T_1, \ldots, T_{r-1}) \\ &= \bar{x}_r - (\mu_0 + k) + T_{r-1} - \text{mín}(T_1, \ldots, T_{r-1}) \\ &= \bar{x}_r - (\mu_0 + k) + d_{r-1} \end{aligned}$$

Puesto que d_r no puede ser negativa, es la más grande de estas dos cantidades.

Diseño de un procedimiento CUSUM

Sea Δ el tamaño de un cambio en μ que se detecta con rapidez por medio de un procedimiento CUSUM.* Es común establecer $k = \Delta/2$. Ahora suponga que un profesional de control de calidad especifica valores deseados de dos duraciones de corrida promedio:

1. ARL cuando el proceso está bajo control ($\mu = \mu_0$)
2. ARL cuando el proceso está fuera de control porque la media se desplazó una cantidad Δ ($\mu = \mu_0 + \Delta$ o $\mu = \mu_0 - \Delta$)

Una gráfica que desarrolló Kenneth Kemp ("The Use of Cumulative Sums for Sampling Inspection Schemes", *Applied Statistics*, 1962: 23), llamada *nomograma*, se usa para determinar

*Esto contrasta con la notación previa, donde Δ representa el número de desviaciones estándar por las que cambia μ.

los valores de h y n que logran las duraciones de corrida promedio.* Esta gráfica se muestra como la figura 16.11. El método para usar la gráfica se describe en el cuadro siguiente. Se debe conocer el valor de σ, o bien, se usa una estimación.

Cómo utilizar el nomograma de Kemp

1. Localice las ARL deseadas en las escalas bajo control y fuera de control. Una estos dos puntos con una recta.
2. Observe dónde la recta cruza la escala k' y determine el valor de n a partir de la ecuación

$$k' = \frac{\Delta/2}{\sigma/\sqrt{n}}$$

   Luego, redondee n hasta el entero más próximo.
3. Una el punto de la escala k' con el punto en la escala ARL bajo control por medio de una segunda línea, y note donde ésta cruza la escala h'. Entonces $h = (\sigma/\sqrt{n}) \cdot h'$.

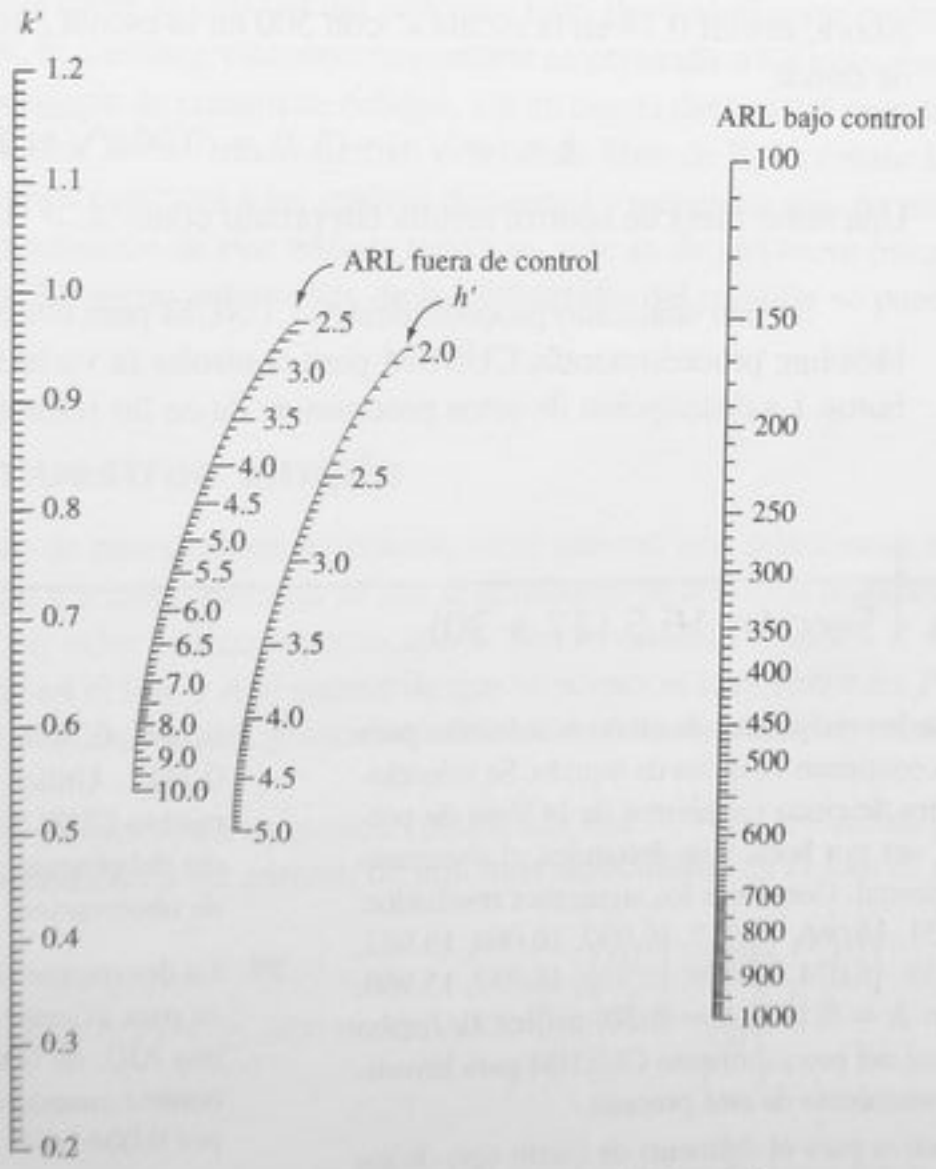


Figura 16.11 Nomograma de Kemp†

*La palabra *nomograma* no es específica para esta gráfica; los nomogramas se utilizan para muchos otros propósitos.

†FUENTE: Kemp, Kenneth W., "The Use of Cumulative Sums for Sampling Inspection Schemes", *Applied Statistics*, Vol. XI, 1962: 23. Con autorización de Blackwell Publishing.

El valor de $h = 0.95$ se usó en el ejemplo 16.9. En esa situación, se deduce que la ARL bajo control es 500 y la ARL fuera de control (para $\Delta = 0.3$) es 7.

Ejemplo 16.10 El valor objetivo para el diámetro del núcleo interior de una bomba hidráulica es 2.250 pulg. Si la desviación estándar del diámetro del núcleo es $\sigma = 0.004$, ¿qué procedimiento CUSUM produce una ARL bajo control de 500 y una ARL de 5 cuando el diámetro medio del núcleo cambia por la cantidad 0.003 pulg.?

Al unir el punto 500 en la escala ARL bajo control con el punto 5 de la escala ARL fuera de control y extender la línea hasta la escala k' en el lado izquierdo de la figura 16.11, se obtiene $k' = 0.74$. Por lo tanto,

$$k' = 0.74 = \frac{\Delta/2}{\sigma/\sqrt{n}} = \frac{0.0015}{0.004/\sqrt{n}} = 0.375\sqrt{n}$$

entonces,

$$\sqrt{n} = \frac{0.74}{0.375} = 1.973 \qquad n = (1.973)^2 = 3.894$$

Por consiguiente el procedimiento CUSUM se debe basar en el tamaño de muestra $n = 4$. Ahora, al unir 0.74 en la escala k' con 500 en la escala ARL bajo control se obtiene $h' = 3.2$, de donde

$$h = (\sigma/\sqrt{n}) \cdot (3.2) = (0.004/\sqrt{4})(3.2) = 0.0064$$

Una señal fuera de control resulta tan pronto como $d_t > 0.0064$ o $e_t > 0.0064$. ■

Se han analizado procedimientos CUSUM para controlar la ubicación del proceso. También hay procedimientos CUSUM para controlar la variación de proceso y para datos de atributos. La descripción de estos procesos se da en las referencias del capítulo.

Ejercicios | Sección 16.5 (27 a 30)

27. Se supone que los recipientes de cierto tratamiento para fosas sépticas contienen 16 onzas de líquido. Se selecciona una muestra de cinco recipientes de la línea de producción, una vez por hora, y se determina el contenido promedio muestral. Considere los siguientes resultados: 15.992, 16.051, 16.066, 15.912, 16.030, 16.060, 15.982, 15.899, 16.038, 16.074, 16.029, 15.935, 16.032, 15.960, 16.055. Con $\Delta = 0.10$ y $h = 0.20$, utilice la forma computacional del procedimiento CUSUM para investigar el comportamiento de este proceso.

28. El valor objetivo para el diámetro de cierto tipo de eje impulsor es 0.75 pulg. El tamaño del cambio en el diámetro promedio que se considera importante detectar es 0.002 pulg. Los diámetros promedio muestrales para grupos sucesivos de $n = 4$ ejes son los siguientes: 0.7507, 0.7504, 0.7492, 0.7501, 0.7503, 0.7510, 0.7490, 0.7497, 0.7488, 0.7504, 0.7516, 0.7472, 0.7489, 0.7483, 0.7471, 0.7498, 0.7460, 0.7482, 0.7470, 0.7493, 0.7462, 0.7481. Utilice la forma computacional del procedimiento CUSUM con $h = 0.003$ para discernir si la media del proceso permaneció en el objetivo todo el tiempo de observación.

29. La desviación estándar de cierta dimensión en una pieza para aviones es 0.005 cm. ¿Cuál procedimiento dará una ARL de 600 bajo control y una ARL de 4 fuera de control cuando el valor medio de la dimensión cambia por 0.004 cm?

30. Cuando la ARL fuera de control corresponde a un cambio de una desviación estándar en la media del proceso, ¿cuáles son las características del procedimiento CUSUM con duraciones de corrida promedio de 250 y 4.8 para las condiciones bajo control y fuera de control, respectivamente?

16.6 Muestreo de aceptación

Los artículos que provienen de un proceso de producción suelen ser enviados en grupos a otra compañía o establecimiento comercial. Un grupo podría consistir en las unidades de determinada corrida de producción o turno, en un recipiente de embarque de algún tipo, enviado en respuesta a una orden particular, etcétera. El grupo de artículos por lo común se llama *lote*, a la persona que envía se le llama *productor* y a quien recibe el lote se le conoce como *consumidor*. Se centra la atención en situaciones en las que cada artículo es defectuoso o no defectuoso, con p que denota la proporción de unidades defectuosas en el lote. Por supuesto que el consumidor acepta el lote si el valor de p es adecuadamente pequeño. El muestreo de aceptación es la parte de la estadística aplicada que trata con métodos para decidir si el consumidor acepta o rechaza un lote.

Hasta hace poco, los profesionales consideraban los procedimientos de gráficas de control y las técnicas de muestreo de aceptación de partes igualmente importantes de la metodología de control de calidad. Esto ya no es así. La razón es que el uso de las gráficas de control y otras estrategias recién desarrolladas ofrecen la oportunidad de diseñar calidad en un producto, mientras que el muestreo de aceptación trata con lo que ya se produjo y, por lo tanto, no proporciona ningún control directo sobre la calidad del proceso. Esto llevó al experto en control de calidad, recién fallecido, W. E. Deming, elemento importante en persuadir a los japoneses a hacer uso sustancial de la metodología de control de calidad, a ir en contra del uso del muestreo de aceptación en muchas situaciones. De un modo similar, el reciente libro de Ryan (véase la bibliografía del capítulo) dedica varios capítulos a las gráficas de control y menciona sólo de paso el muestreo de aceptación. Como reflexión de este énfasis reducido, sólo se da una breve introducción de estos conceptos básicos. En varias referencias de la bibliografía del capítulo se puede encontrar más información.

Planes de muestreo simple

El tipo más directo de muestreo de aceptación tiene que ver con seleccionar una sola muestra aleatoria de tamaño n y luego rechazar el lote si el número de artículos defectuosos en la muestra es mayor que un valor crítico especificado c. Sea la variable aleatoria X el número de artículos defectuosos en el lote y A el suceso de que se acepta el lote. Entonces $P(A) = P(X \leq c)$ es una función de p; mientras más grande sea el valor de p, menor será la probabilidad de aceptar el lote.

Si el tamaño de muestra n es grande comparado con N, $P(A)$ se calcula por medio de la distribución hipergeométrica (el número de artículos defectuosos en el lote es Np):

$$P(X \leq c) = \sum_{x=0}^{c} h(x; n, Np, N) = \sum_{x=0}^{c} \frac{\binom{Np}{x} \cdot \binom{N(1-p)}{n-x}}{\binom{N}{n}}$$

Cuando n es pequeña con respecto a N (la regla empírica sugerida antes fue $n \leq 0.05N$, aunque algunos autores emplean la regla menos conservadora $n \leq 0.10N$), se puede usar la distribución binomial:

$$P(X \leq c) = \sum_{x=0}^{c} b(x; n, p) = \sum_{x=0}^{c} \binom{n}{x} p^x (1-p)^{n-x}$$

Por último, si $P(A)$ es grande sólo cuando p es pequeña (esto depende del valor de c), se justifica la aproximación de Poisson a la distribución binomial:

$$P(X \le c) \approx \sum_{x=0}^{c} p(x; np) = \sum_{x=0}^{c} \frac{e^{-np}(np)^x}{x!}$$

El comportamiento de un plan de muestreo se puede resumir muy bien al trazar $P(A)$ como una función de p. A esta clase de gráfica se le conoce como **curva característica de operación (CO)** para el plan.

Ejemplo 16.11 Considere el plan de muestreo con valor crítico $c = 2$ y tamaño de muestra $n = 50$, y suponga que el tamaño del lote es mayor que 1000, de manera que se puede usar la distribución binomial. Esto da

$$P(A) = P(X \le 2) = (1 - p)^{50} + 50p(1 - p)^{49} + 1255p^2(1 - p)^{48}$$

En la tabla siguiente se muestra $P(A)$ para valores seleccionados de p, y la curva característica de operación (CO) correspondiente se muestra en la figura 16.12.

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.12	0.15
$P(A)$	0.986	0.922	0.811	0.677	0.541	0.416	0.311	0.226	0.161	0.112	0.051	0.014

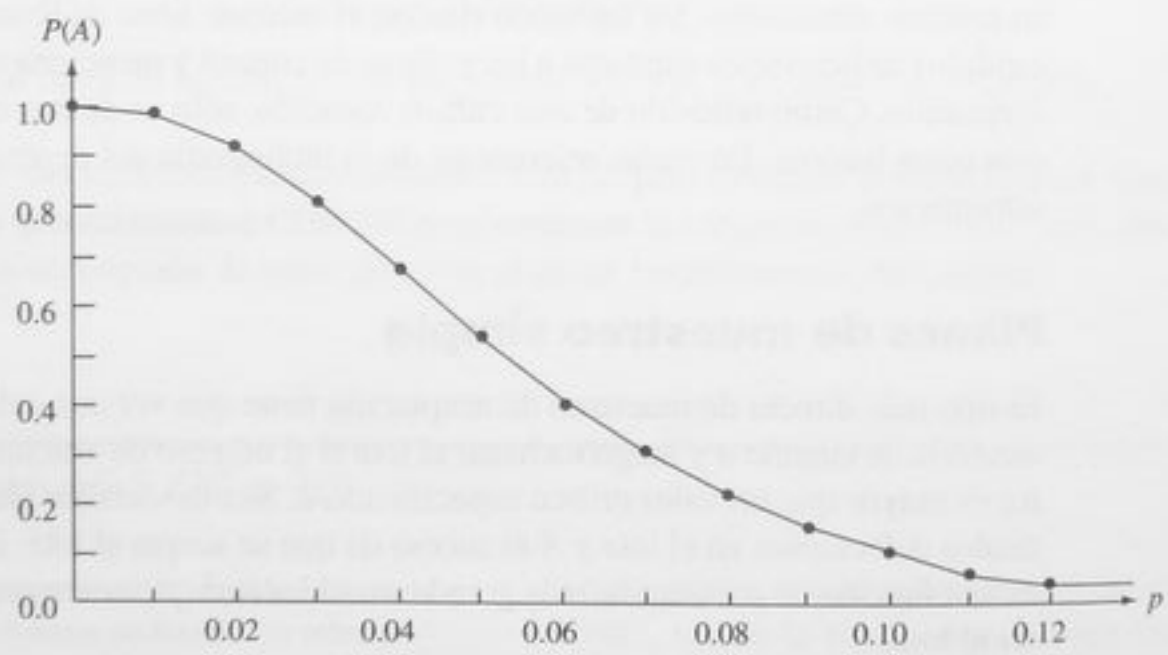


Figura 16.12 Curva CO para el plan de muestreo con $c = 2$, $n = 50$ ■

La curva CO para el plan del ejemplo 16.11 tiene $P(A)$ cercana a 1 para p muy próxima a 0. Sin embargo, en muchas aplicaciones una proporción de artículos defectuosos de 8% [para la cual $P(A) = 0.226$] o incluso sólo 5% [$P(A) = 0.541$] sería considerada excesiva, en cuyo caso las probabilidades de aceptación son demasiado altas. Incrementando el valor crítico c mientras se mantiene fija n, se obtiene un plan para el cual $P(A)$ se incrementa en cada p (excepto 0 y 1), así que la nueva curva CO queda arriba de la anterior. Esto es deseable para p cerca de 0, pero no para valores de p más grandes. Mantener c constante mientras se incrementa n produce una curva CO menor, que está bien para p más grande, pero no para p cercana a 0. Se quiere una curva CO que sea mayor para p pequeña, y menor para p más grande. Esto se logra al incrementar n y ajustar c.

Diseño de un plan de una sola muestra

Un plan de muestreo efectivo es el que tiene las características siguientes:

1. Tiene una alta probabilidad especificada de aceptar lotes que el productor considera de buena calidad.
2. Tiene una baja probabilidad especificada de aceptar lotes que el productor considera de mala calidad.

Un plan de este tipo se desarrolla si se procede como sigue: Desígnense dos valores diferentes de p, uno para el que $P(A)$ es un valor especificado cercano a 1 y el otro para el cual $P(A)$ es un valor especificado cercano a 0. Estos dos valores de p, por ejemplo p_1 y p_2 por lo común reciben el nombre de **nivel de calidad aceptable (NCA)** y el **porcentaje de tolerancia de artículos defectuosos en el lote (PTDL)**. Es decir, se requiere un plan para el cual

1. $P(A) = 1 - \alpha$ cuando $p = p_1 = \text{NCA}$ (α pequeña)
2. $P(A) = \beta$ cuando $p = p_2 = \text{PTDL}$ (β pequeña)

Esto es análogo a buscar un procedimiento de prueba con probabilidad α especificada de error de tipo I y probabilidad β especificada de error de tipo II al probar las hipótesis. Por ejemplo, se podría tener

$$\text{NCA} = 0.01 \qquad \alpha = 0.05 \ (P(A) = 0.95)$$
$$\text{PTDL} = 0.045 \qquad \beta = 0.10 \ (P(A) = 0.10)$$

Debido a que X es discreta, por lo común se debe estar satisfecho con los valores de n y c que satisfacen más o menos estas condiciones.

En la tabla 16.6 se muestra información a partir de la cual se determinan n y c en el caso $\alpha = 0.05$, $\beta = 0.10$.

Tabla 16.6 Factores para determinar n y c para un plan de una sola muestra con $\alpha = 0.05$, $\beta = 0.10$

c	np_1	np_2	p_2/p_1	c	np_1	np_2	p_2/p_1
0	0.051	2.30	45.10	8	4.695	12.99	2.77
1	0.355	3.89	10.96	9	5.425	14.21	2.62
2	0.818	5.32	6.50	10	6.169	15.41	2.50
3	1.366	6.68	4.89	11	6.924	16.60	2.40
4	1.970	7.99	4.06	12	7.690	17.78	2.31
5	2.613	9.28	3.55	13	8.464	18.86	2.24
6	3.285	10.53	3.21	14	9.246	20.13	2.18
7	3.981	11.77	2.96	15	10.040	21.29	2.12

Ejemplo 16.12 Determínese un plan para el cual $\text{NCA} = p_1 = 0.01$ y $\text{PTDL} = p_2 = 0.045$. La relación de p_2 a p_1 es

$$\frac{\text{PTDL}}{\text{NCA}} = \frac{p_2}{p_1} = \frac{0.045}{0.01} = 4.50$$

Este valor se ubica entre la relación 4.89 dada en la tabla 16.6, para la cual $c = 3$, y 4.06, para la cual $c = 4$. Una vez elegido uno de estos valores de c, n se determina ya sea al dividir el valor np_1 de la tabla 16.6, entre p_1 o vía np_2/p_2. Así, cuatro planes diferentes (dos valores de c, y para cada uno dos valores de n) dan aproximadamente el valor especificado de α y β. Considere, por ejemplo, usar $c = 3$ y

$$n = \frac{np_1}{p_1} = \frac{1.366}{0.01} = 136.6 \approx 137$$

Entonces,

$$\alpha = 1 - P(X \leq 3 \text{ cuando } p = p_1)$$

$$= 1 - \sum_{x=0}^{3} \binom{137}{x} (0.01)^x (0.99)^{137-x} = 0.050$$

(la aproximación de Poisson con $\lambda = 1.37$ también da 0.050) y

$$\beta = P(X \leq 3 \text{ cuando } p = p_2) = 0.131$$

El plan con $c = 4$ y n determinada de $np_2 = 7.99$ tiene $n = 178$, $\alpha = 0.034$ y $\beta = 0.094$. El tamaño de muestra más grande da como resultado un plan con α y β más pequeñas que los valores especificados correspondientes. ■

El libro de Douglas Montgomery citado en la bibliografía del capítulo, contiene una gráfica de la cual se determinan c y n para valores específicos de α y β.

Podría suceder que el número de artículos defectuosos en una muestra llegue a $c + 1$ antes de examinarlos. Por ejemplo, en el caso $c = 3$ y $n = 137$, podría ser que el artículo 125 examinado sea el cuarto artículo defectuoso, así que es innecesario examinar los 12 artículos restantes. Sin embargo, se recomienda examinar los artículos, incluso si esto ocurre a fin de proveer una historia de calidad lote por lote y estimaciones de p con el tiempo.

Planes de muestreo doble

En un plan de muestreo doble, se determina el número de artículos defectuosos x_1 en una muestra inicial de tamaño n_1. Entonces hay tres posibles maneras de actuar: aceptar el lote de inmediato, rechazarlo de inmediato o tomar una segunda muestra de n_2 artículos y rechazar o aceptar el lote dependiendo del número total de artículos defectuosos $x_1 + x_2$ en las dos muestras. Además de los dos tamaños de muestra, un plan específico se caracteriza por otros tres números, c_1, r_1 y c_2, como sigue:

1. Rechazar el lote si $x_1 \geq r_1$.
2. Aceptar el lote si $x_1 \leq c_1$.
3. Si $c_1 < x_1 < r_1$, tomar una segunda muestra; luego, aceptar el lote si $x_1 + x_2 \leq c_2$ y rechazarlo en caso contrario.

Ejemplo 16.13 Considere el plan de muestreo doble con $n_1 = 80$, $n_2 = 80$, $c_1 = 2$, $r_1 = 5$ y $c_2 = 6$. Así, el lote será aceptado si (1) $x_1 = 0$, 1, o 2; (2) $x_1 = 3$ y $x_2 = 0$, 1, 2 o 3; o bien (3) $x_1 = 4$ y $x_2 = 0$, 1 o 2.

Suponiendo que el tamaño del lote es suficientemente grande para aplicar la aproximación binomial, la probabilidad $P(A)$ de aceptar el lote es

$$\begin{aligned} P(A) &= P(X_1 = 0, 1 \text{ o } 2) + P(X_1 = 3, X_2 = 0, 1, 2 \text{ o } 3) \\ &\quad + P(X_1 = 4, X_2 = 0, 1 \text{ o } 2) \\ &= \sum_{x_1=0}^{2} b(x_1; 80, p) + b(3; 80, p) \sum_{x_2=0}^{3} b(x_2; 80, p) \\ &\quad + b(4; 80, p) \sum_{x_2=0}^{2} b(x_2; 80, p) \end{aligned}$$

De nuevo la gráfica de $P(A)$ en función de p es la curva CO del plan. La curva CO para plan se ilustra en la figura 16.13.

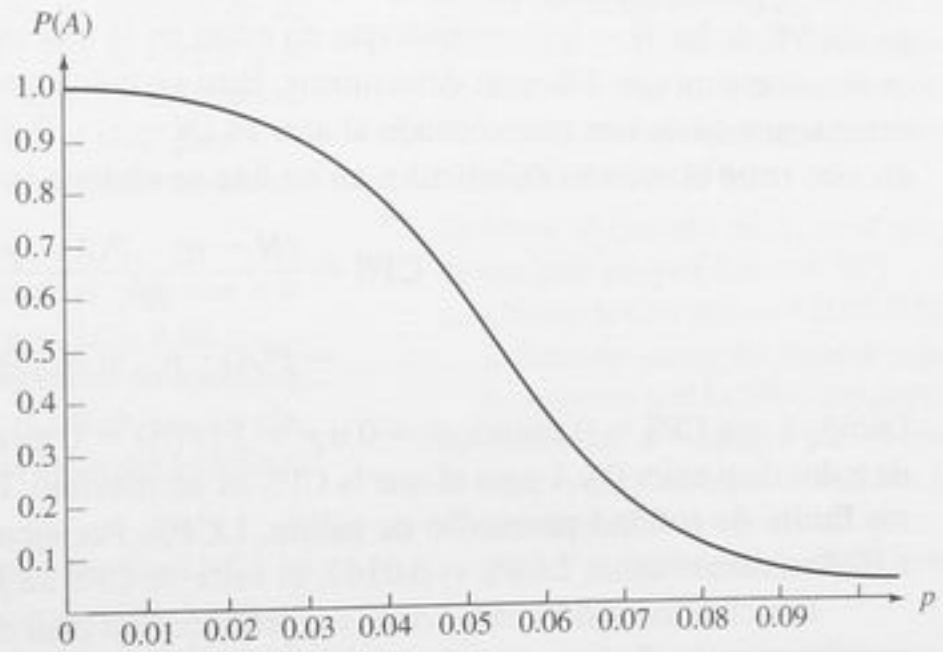


Figura 16.13 Curva CO para el plan de muestreo doble del ejemplo 16.13 ■

Un método para diseñar un plan de muestreo doble, requiere proceder como se sugirió antes para planes de una sola muestra. Especifique los valores p_1 y p_2 junto con las probabilidades de aceptación correspondientes $1 - \alpha$ y β. Luego encuentre un plan que satisface estas condiciones. El libro de Montgomery contiene tablas similares a la tabla 16.6 para este propósito en los casos $n_2 = n_1$ y $n_2 = 2n_1$ con $1 - \alpha = 0.95$, $\beta = 0.10$. En otras fuentes se encuentran tablas más completas de planes.

Análoga a la práctica estándar con los planes de una sola muestra, se recomienda examinar los artículos de la primera muestra, aun cuando se descubra el artículo defectuoso $(r_1 + 1)$ antes de la inspección del artículo n_1. Sin embargo, se acostumbra terminar la inspección de la segunda muestra si el número de artículos defectuosos es suficiente para justificar el rechazo antes de haber examinado todos los artículos. Esto se conoce como *restricción* en la segunda muestra. Bajo la restricción, se puede demostrar que el número esperado de artículos inspeccionados en un plan de muestreo doble sea menor que el número de artículos examinados en un plan de muestreo simple cuando las curvas CO de los dos planes estén próximas a ser idénticas. Ésta es la virtud principal de los planes de muestreo doble. Es necesario consultar un libro de control de calidad a fin de profundizar en estas cuestiones, así como en el análisis de

planes de muestreo múltiples y secuenciales (en los que se requiere seleccionar artículos para inspección uno por uno en vez de en grupos).

Inspección de rectificación y otros criterios de diseño

En algunas situaciones, la inspección de muestreo se realiza por medio de *rectificación*. Para planes de una sola muestra, esto significa que cada artículo defectuoso de la muestra se reemplaza por uno satisfactorio, y si el número de artículos defectuosos en la muestra es mayor que el límite de aceptación c, se examinan *todos* los artículos del lote y cada artículo defectuoso se sustituye por uno bueno. Sea N el tamaño del lote. Una característica importante de un plan de muestreo con inspección de rectificación es la **calidad promedio de salida**, denotada por **CPS**. Ésta es la proporción a largo plazo de artículos defectuosos entre los que son enviados o después que se empleó el plan de muestreo. Ahora los artículos defectuosos aparecerán entre los $N - n$ artículos no inspeccionados en un lote que se consideró aceptable con base en una muestra. Por ejemplo, suponga que $P(A) = P(X \leq c) = 0.985$ cuando $p = 0.01$. Entonces, a la larga, 98.5% de los $N - n$ elementos que no están en la muestra quedarán sin inspección, de los cuales se espera que 1% sean defectuosos. Esto significa que el número esperado de artículos defectuosos en un lote seleccionado al azar es $(N - n) \cdot P(A) \cdot p = 0.00985(N - n)$. Al dividir esto entre el número de artículos en un lote se obtiene la calidad promedio de salida:

$$\begin{aligned} \text{CPS} &= \frac{(N-n)}{N} \cdot P(A) \cdot p \\ &\approx P(A) \cdot p \quad \text{si } N \gg n \end{aligned}$$

Debido a que CPS $= 0$ cuando $p = 0$ o $p = 1$ [$P(A) = 0$ en el último caso], se deduce que hay un valor de p entre 0 y 1 para el que la CPS es un máximo. El valor máximo de la CPS se llama **límite de calidad promedio de salida, LCPS.** Por ejemplo, para el plan con $n = 137$ y $c = 3$ analizado antes, LCPS $= 0.0142$, el valor de CPS en $p \approx 0.02$.

Las elecciones adecuadas de n y c producen un plan de muestreo para el cual LCPS es un número pequeño especificado. Sin embargo, tal plan no es único, así que se puede imponer otra condición. A menudo esta segunda condición requiere el **número total inspeccionado promedio** (es decir, esperado), representado por **NTIP**. El número total de artículos inspeccionados es un lote elegido al azar es una variable aleatoria que toma el valor n con probabilidad $P(A)$ y N con probabilidad $1 - P(A)$. Así que el número esperado de artículos inspeccionados en un lote seleccionado al azar es

$$\text{NTIP} = n \cdot P(A) + N \cdot (1 - P(A))$$

Es práctica común seleccionar un plan de muestreo que tiene un LCPS y, además, un NTIP mínimo a un determinado nivel de calidad p.

Planes de muestreo estándar

Parecería como si la determinación de un plan de muestreo que satisface simultáneamente varios criterios sería bastante difícil. Por fortuna, otros han sentado las bases de tales planes en la forma de tabulaciones extensas. Después de la Segunda Guerra Mundial, el ejército desarrolló MIL STD 105D, que es el conjunto de planes más utilizado. Una versión civil, ANSI/ASQC Z1.4, es bastante similar a la versión militar. Dos expertos en estadística aplicada llamados Dodge y Romig, elaboraron antes de la Segunda Guerra Mundial en los laboratorios Bell un tercer

conjunto de planes que es bastante popular. El libro de Montgomery (véase la bibliografía del capítulo) contiene una introducción al uso de estos planes que vale la pena leer.

Ejercicios | Sección 16.6 (31 a 38)

31. Considere el plan de una sola muestra con $c = 2$ y $n = 50$, como se analizó en el ejemplo 16.11, pero ahora suponga que el tamaño del lote es $N = 500$. Calcule $P(A)$, la probabilidad de aceptar el lote, para $p = 0.01, 0.02, \ldots, 0.10$ por medio de la distribución hipergeométrica. ¿En este caso la aproximación binomial produce resultados satisfactorios?

32. Se va a seleccionar una muestra de 50 artículos de un lote que consta de 5000 piezas. El lote se acepta si la muestra contiene a lo sumo un artículo defectuoso. Calcule la probabilidad de aceptar el lote para $p = 0.01, 0.02, \ldots, 0.10$, trace la curva característica de operación.

33. Refiérase al ejercicio 32 y considere el plan con $n = 100$ y $c = 2$. Calcule $P(A)$ para $p = 0.01, 0.02, \ldots, 0.05$, y trace las dos curvas características de operación sobre el mismo conjunto de ejes. ¿Cuál de los dos planes es preferible (dejando aparte el costo del muestreo) y por qué?

34. Elabore un plan de una sola muestra para el que NCA = 0.02 y PTDL = 0.07 en el caso $\alpha = 0.05$, $\beta = 0.10$. Una vez determinados los valores de n y c calcule los valores de α y β alcanzados para el plan.

35. Considere el plan de muestreo doble para el cual ambos tamaños de muestra son 50; el lote se acepta después de la primera muestra si el número de piezas defectuosas es a lo sumo uno, se rechaza si el número de artículos defectuosos es por lo menos cuatro, y se rechaza después de la segunda muestra si el número total de artículos defectuosos es seis o más. Calcule la probabilidad de aceptar el lote cuando $p = 0.01$, 0.05 y 0.10.

36. Algunas fuentes apoyan un tipo de plan de muestreo doble un poco más restrictivo en el cual $r_1 = c_2 + 1$; es decir, el lote es rechazado si en alguna etapa el número (total) de piezas defectuosas es por lo menos r_1 (véase el libro de Montgomery). Considere este tipo de plan de muestreo con $n_1 = 50$, $n_2 = 100$, $c_1 = 1$ y $r_1 = 4$. Calcule la probabilidad de aceptación del lote cuando $p = 0.02$, 0.05 y 0.10.

37. Refiérase al ejemplo 16.11, en el que se utilizó un plan de una sola muestra con $n = 50$ y $c = 2$.
    a. Calcule la CPS para $p = 0.01, 0.02, \ldots, 0.10$. ¿Qué indica esto acerca del valor de p para el cual CPS es un máximo y el LCPS correspondiente?
    b. Determine el valor de p para el cual CPS sea un máximo y el LCPS correspondiente. [*Sugerencia:* utilice cálculo].
    c. Para $N = 2000$, calcule el ATI para los valores p del inciso (a).

38. Considere el plan de una sola muestra que utiliza $n = 50$ y $c = 1$ cuando $N = 2000$. Determine los valores de CPS y NTIP para los valores seleccionados de p, y dibuje cada uno de éstos en función de p. También determine el valor de LCPS.

Ejercicios complementarios (39 a 44)

39. Las observaciones sobre la resistencia al corte para 26 subgrupos de puntos de soldadura de prueba, cada uno formado por seis soldaduras, producen $\sum \bar{x}_i = 10\,980$, $\sum s_i = 402$ y $\sum r_i = 1074$. Calcule límites de control para alguna gráfica de control pertinente.

40. Se determina el número de rayas en la superficie de cada una de 24 placas metálicas rectangulares, y se obtienen los siguientes datos: 8, 1, 7, 5, 2, 0, 2, 3, 4, 3, 1, 2, 5, 7, 3, 4, 6, 5, 2, 4, 0, 10, 2, 6. Construya una gráfica de control apropiada y comente acerca de las características importantes.

41. Los siguientes números son observaciones acerca de la resistencia a la tensión de especímenes de tela sintética seleccionados de un proceso de producción a intervalos de tiempo igualmente separados. Construya gráficas de control apropiadas y coméntelas (suponga que una causa asignada es identificable para alguna observación fuera de control).

1.	51.3	51.7	49.5
2.	51.0	50.0	49.3
3.	50.8	51.1	49.0
4.	50.6	51.1	49.0
5.	49.6	50.5	50.9
6.	51.3	52.0	50.3
7.	49.7	50.5	50.3
8.	51.8	50.3	50.0
9.	48.6	50.5	50.7
10.	49.6	49.8	50.5
11.	49.9	50.7	49.8
12.	49.6	48.4	50.0
13.	49.8	51.2	49.7
14.	50.4	49.9	50.7
15.	49.4	49.5	49.0
16.	50.7	49.0	50.0
17.	50.8	49.5	50.9
18.	48.5	50.3	49.3
19.	49.6	50.6	49.4
20.	50.9	49.4	49.7
21.	54.1	49.8	48.5
22.	50.2	49.6	51.5

42. Una alternativa para la gráfica de p de la fracción de artículos defectuosos, es la *gráfica np para el número de artículos defectuosos*. Esta gráfica tiene LCS $= n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$, LCI $= n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})}$, y el *número* de artículos defectuosos de cada muestra se bosqueja en una gráfica. Construya este tipo de gráfica para los datos del ejemplo 16.6. ¿El uso de una gráfica *np* siempre transmite el mismo mensaje que el uso de una gráfica p (es decir, son equivalentes las dos gráficas)?

43. Las observaciones de resistencia (en ohms) para subgrupos de cierto tipo de registro, dieron las siguientes cantidades de resumen:

i	n_i	$\bar{x}_i$	s_i	i	n_i	$\bar{x}_i$	s_i
1	4	430.0	22.5	7	4	420.8	25.4
2	4	418.2	20.6	8	4	431.4	24.0
3	3	435.5	25.1	9	4	428.7	21.2
4	4	427.6	22.3	10	4	440.1	25.8
5	4	444.0	21.5	11	4	445.2	27.3
6	3	431.4	28.9	12	4	430.1	22.2
13	4	427.2	24.0	17	3	447.0	19.8
14	4	439.6	23.3	18	4	434.4	23.7
15	3	415.9	31.2	19	4	422.2	25.1
16	4	419.8	27.5	20	4	425.7	24.4

(*continúa*)

Construya los límites de control apropiados. [*Sugerencia:* $\bar{\bar{x}} = \sum n_i\bar{x}_i/\sum n_i$ y $s^2 = \sum(n_i - 1)s_i^2/\sum(n_i - 1)$.]

44. Sea α un número entre 0 y 1 y defina una secuencia W_1, W_2, W_3, . . . mediante $W_0 = \mu$ y $W_t = \alpha\bar{X}_t + (1-\alpha)W_{t-1}$ para $t = 1, 2, \ldots$. Sustituya W_{t-1} por su representación en términos de $\bar{X}_{t-1}$ y W_{t-2}, luego sustituya W_{t-2}, etc. El resultado es

$$W_t = \alpha\bar{X}_t + \alpha(1-\alpha)\bar{X}_{t-1} + \cdots + \alpha(1-\alpha)^{t-1}\bar{X}_1 + (1-\alpha)^t\mu$$

El hecho de que W_t dependa no sólo de $\bar{X}_t$ sino también de promedios para puntos de tiempo anteriores, aunque con ponderaciones decrecientes (exponencialmente), indica que los cambios en la media del proceso se reflejan con más rapidez en las W_t que en cada una de las $\bar{X}_t$.

a. Demuestre que $E(W_t) = \mu$.

b. Sea $\sigma_t^2 = V(W_t)$, y demuestre que

$$\sigma_t^2 = \frac{\alpha[1-(1-\alpha)^{2t}]}{2-\alpha} \cdot \frac{\sigma^2}{n}$$

c. En una *gráfica de control de promedios móviles ponderados exponencialmente*, se trazan las W_t y se utilizan límites de control $\mu_0 \pm 3\sigma_t$ (en lugar de $\bar{\bar{x}}$ μ_0). Construya esta gráfica para los datos del ejemplo 16.9 con $\mu_0 = 40$.

Bibliografía

Box, George, Soren Bisgaard y Conrad Fung, "An Explanation and Critique of Taguchi's Contributions to Quality Engineering", *Quality and Reliability Engineering International*, 1988: 123-131.

Montgomery, Douglas C., *Introduction to Statistical Quality Control* (3a. ed.), Wiley, Nueva York, 1996. Una introducción completa a muchos aspectos del control de calidad casi al mismo nivel de este libro.

Ryan, Thomas P., *Statistical Methods for Quality Improvement*, Wiley, Nueva York, 1989. Capta muy bien el sentido moderno de control de calidad con requerimientos mínimos de fundamentos de los lectores.

Vardeman, Stephen B. y J. Marcus Jobe, *Statistical Quality Assurance Methods for Engineers*, Wiley, Nueva York, 1999. Incluye temas tradicionales de control de calidad, y también material de diseño experimental relacionado con cuestiones de calidad; informal y fiable.

Apéndice de tablas

Tabla A.1 Probabilidades binomiales acumuladas

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

a. $n = 5$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	0.951	0.774	0.590	0.328	0.237	0.168	0.078	0.031	0.010	0.002	0.001	0.000	0.000	0.000	0.000
	1	0.999	0.977	0.919	0.737	0.633	0.528	0.337	0.188	0.087	0.031	0.016	0.007	0.000	0.000	0.000
x	2	1.000	0.999	0.991	0.942	0.896	0.837	0.683	0.500	0.317	0.163	0.104	0.058	0.009	0.001	0.000
	3	1.000	1.000	1.000	0.993	0.984	0.969	0.913	0.812	0.663	0.472	0.367	0.263	0.081	0.023	0.001
	4	1.000	1.000	1.000	1.000	0.999	0.998	0.990	0.969	0.922	0.832	0.763	0.672	0.410	0.226	0.049

b. $n = 10$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	0.904	0.599	0.349	0.107	0.056	0.028	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.996	0.914	0.736	0.376	0.244	0.149	0.046	0.011	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.988	0.930	0.678	0.526	0.383	0.167	0.055	0.012	0.002	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.999	0.987	0.879	0.776	0.650	0.382	0.172	0.055	0.011	0.004	0.001	0.000	0.000	0.000
	4	1.000	1.000	0.998	0.967	0.922	0.850	0.633	0.377	0.166	0.047	0.020	0.006	0.000	0.000	0.000
x	5	1.000	1.000	1.000	0.994	0.980	0.953	0.834	0.623	0.367	0.150	0.078	0.033	0.002	0.000	0.000
	6	1.000	1.000	1.000	0.999	0.996	0.989	0.945	0.828	0.618	0.350	0.224	0.121	0.013	0.001	0.000
	7	1.000	1.000	1.000	1.000	1.000	0.998	0.988	0.945	0.833	0.617	0.474	0.322	0.070	0.012	0.000
	8	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.989	0.954	0.851	0.756	0.624	0.264	0.086	0.004
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.972	0.944	0.893	0.651	0.401	0.096

c. $n = 15$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	0.860	0.463	0.206	0.035	0.013	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.990	0.829	0.549	0.167	0.080	0.035	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.964	0.816	0.398	0.236	0.127	0.027	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.995	0.944	0.648	0.461	0.297	0.091	0.018	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.999	0.987	0.836	0.686	0.515	0.217	0.059	0.009	0.001	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.998	0.939	0.852	0.722	0.402	0.151	0.034	0.004	0.001	0.000	0.000	0.000	0.000
	6	1.000	1.000	1.000	0.982	0.943	0.869	0.610	0.304	0.095	0.015	0.004	0.001	0.000	0.000	0.000
x	7	1.000	1.000	1.000	0.996	0.983	0.950	0.787	0.500	0.213	0.050	0.017	0.004	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.999	0.996	0.985	0.905	0.696	0.390	0.131	0.057	0.018	0.000	0.000	0.000
	9	1.000	1.000	1.000	1.000	0.999	0.996	0.966	0.849	0.597	0.278	0.148	0.061	0.002	0.000	0.000
	10	1.000	1.000	1.000	1.000	1.000	0.999	0.991	0.941	0.783	0.485	0.314	0.164	0.013	0.001	0.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.982	0.909	0.703	0.539	0.352	0.056	0.005	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.973	0.873	0.764	0.602	0.184	0.036	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.965	0.920	0.833	0.451	0.171	0.010
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.987	0.965	0.794	0.537	0.140

(continúa)

Tabla A.1 Probabilidades binomiales acumuladas (*cont.*)

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

d. $n = 20$

	p														
x	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	0.818	0.358	0.122	0.012	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.983	0.736	0.392	0.069	0.024	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.999	0.925	0.677	0.206	0.091	0.035	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	1.000	0.984	0.867	0.411	0.225	0.107	0.016	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	1.000	0.997	0.957	0.630	0.415	0.238	0.051	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	1.000	1.000	0.989	0.804	0.617	0.416	0.126	0.021	0.002	0.000	0.000	0.000	0.000	0.000	0.000
6	1.000	1.000	0.998	0.913	0.786	0.608	0.250	0.058	0.006	0.000	0.000	0.000	0.000	0.000	0.000
7	1.000	1.000	1.000	0.968	0.898	0.772	0.416	0.132	0.021	0.001	0.000	0.000	0.000	0.000	0.000
8	1.000	1.000	1.000	0.990	0.959	0.887	0.596	0.252	0.057	0.005	0.001	0.000	0.000	0.000	0.000
9	1.000	1.000	1.000	0.997	0.986	0.952	0.755	0.412	0.128	0.017	0.004	0.001	0.000	0.000	0.000
10	1.000	1.000	1.000	0.999	0.996	0.983	0.872	0.588	0.245	0.048	0.014	0.003	0.000	0.000	0.000
11	1.000	1.000	1.000	1.000	0.999	0.995	0.943	0.748	0.404	0.113	0.041	0.010	0.000	0.000	0.000
12	1.000	1.000	1.000	1.000	1.000	0.999	0.979	0.868	0.584	0.228	0.102	0.032	0.000	0.000	0.000
13	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.942	0.750	0.392	0.214	0.087	0.002	0.000	0.000
14	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.979	0.874	0.584	0.383	0.196	0.011	0.000	0.000
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.949	0.762	0.585	0.370	0.043	0.003	0.000
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.984	0.893	0.775	0.589	0.133	0.016	0.000
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.965	0.909	0.794	0.323	0.075	0.001
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.976	0.931	0.608	0.264	0.017
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997	0.988	0.878	0.642	0.182

(*continúa*)

Tabla A.1 Probabilidades binomiales acumuladas (*cont.*)

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

e. $n = 25$

	p														
	0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
0	0.778	0.277	0.072	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.974	0.642	0.271	0.027	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.998	0.873	0.537	0.098	0.032	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	1.000	0.966	0.764	0.234	0.096	0.033	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	1.000	0.993	0.902	0.421	0.214	0.090	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	1.000	0.999	0.967	0.617	0.378	0.193	0.029	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	1.000	1.000	0.991	0.780	0.561	0.341	0.074	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	1.000	1.000	0.998	0.891	0.727	0.512	0.154	0.022	0.001	0.000	0.000	0.000	0.000	0.000	0.000
8	1.000	1.000	1.000	0.953	0.851	0.677	0.274	0.054	0.004	0.000	0.000	0.000	0.000	0.000	0.000
9	1.000	1.000	1.000	0.983	0.929	0.811	0.425	0.115	0.013	0.000	0.000	0.000	0.000	0.000	0.000
10	1.000	1.000	1.000	0.994	0.970	0.902	0.586	0.212	0.034	0.002	0.000	0.000	0.000	0.000	0.000
11	1.000	1.000	1.000	0.998	0.980	0.956	0.732	0.345	0.078	0.006	0.001	0.000	0.000	0.000	0.000
12	1.000	1.000	1.000	1.000	0.997	0.983	0.846	0.500	0.154	0.017	0.003	0.000	0.000	0.000	0.000
13	1.000	1.000	1.000	1.000	0.999	0.994	0.922	0.655	0.268	0.044	0.020	0.002	0.000	0.000	0.000
14	1.000	1.000	1.000	1.000	1.000	0.998	0.966	0.788	0.414	0.098	0.030	0.006	0.000	0.000	0.000
15	1.000	1.000	1.000	1.000	1.000	1.000	0.987	0.885	0.575	0.189	0.071	0.017	0.000	0.000	0.000
16	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.946	0.726	0.323	0.149	0.047	0.000	0.000	0.000
17	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.978	0.846	0.488	0.273	0.109	0.002	0.000	0.000
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	0.926	0.659	0.439	0.220	0.009	0.000	0.000
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.971	0.807	0.622	0.383	0.033	0.001	0.000
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991	0.910	0.786	0.579	0.098	0.007	0.000
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.967	0.904	0.766	0.236	0.034	0.000
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991	0.968	0.902	0.463	0.127	0.002
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.993	0.973	0.729	0.358	0.026
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.928	0.723	0.222

Tabla A.2 Probabilidades de Poisson acumuladas

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^{y}}{y!}$$

	λ									
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3		1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4				1.000	1.000	1.000	0.999	0.999	0.998	0.996
5							1.000	1.000	1.000	0.999
6										1.000

(continúa)

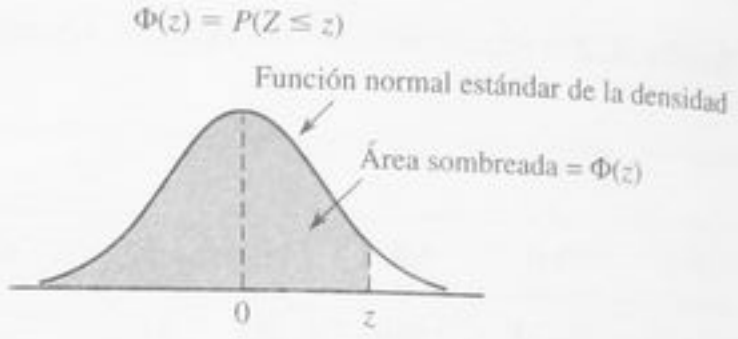
x

Tabla A.2 Probabilidades de Poisson acumuladas (*cont.*)

$$F(x;\lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^{y}}{y!}$$

	λ										
	2.0	**3.0**	**4.0**	**5.0**	**6.0**	**7.0**	**8.0**	**9.0**	**10.0**	**15.0**	**20.0**
0	0.135	0.050	0.018	0.007	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1	0.406	0.199	0.092	0.040	0.017	0.007	0.003	0.001	0.000	0.000	0.000
2	0.677	0.423	0.238	0.125	0.062	0.030	0.014	0.006	0.003	0.000	0.000
3	0.857	0.647	0.433	0.265	0.151	0.082	0.042	0.021	0.010	0.000	0.000
4	0.947	0.815	0.629	0.440	0.285	0.173	0.100	0.055	0.029	0.001	0.000
5	0.983	0.916	0.785	0.616	0.446	0.301	0.191	0.116	0.067	0.003	0.000
6	0.995	0.966	0.889	0.762	0.606	0.450	0.313	0.207	0.130	0.008	0.000
7	0.999	0.988	0.949	0.867	0.744	0.599	0.453	0.324	0.220	0.018	0.001
8	1.000	0.996	0.979	0.932	0.847	0.729	0.593	0.456	0.333	0.037	0.002
9		0.999	0.992	0.968	0.916	0.830	0.717	0.587	0.458	0.070	0.005
10		1.000	0.997	0.986	0.957	0.901	0.816	0.706	0.583	0.118	0.011
11			0.999	0.995	0.980	0.947	0.888	0.803	0.697	0.185	0.021
12			1.000	0.998	0.991	0.973	0.936	0.876	0.792	0.268	0.039
13				0.999	0.996	0.987	0.966	0.926	0.864	0.363	0.066
14				1.000	0.999	0.994	0.983	0.959	0.917	0.466	0.105
15					0.999	0.998	0.992	0.978	0.951	0.568	0.157
16					1.000	0.999	0.996	0.989	0.973	0.664	0.221
17						1.000	0.998	0.995	0.986	0.749	0.297
18							0.999	0.998	0.993	0.819	0.381
19							1.000	0.999	0.997	0.875	0.470
20								1.000	0.998	0.917	0.559
21									0.999	0.947	0.644
22									1.000	0.967	0.721
23										0.981	0.787
24										0.989	0.843
25										0.994	0.888
26										0.997	0.922
27										0.998	0.948
28										0.999	0.966
29										1.000	0.978
30											0.987
31											0.992
32											0.995
33											0.997
34											0.999
35											0.999
36											1.000

Tabla A.3 Áreas de la curva normal estándar



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0038
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3482
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

(continúa)

Tabla A.3 Áreas de la curva normal estándar (*cont.*)

$\Phi(z) = P(Z \leq z)$

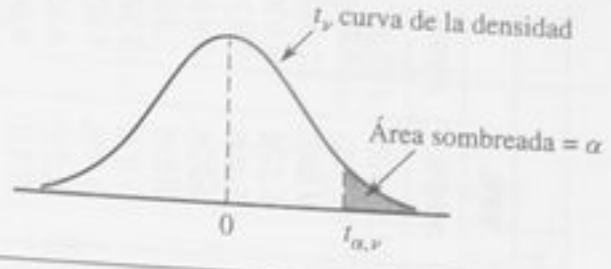
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Tabla A.4 Función gamma incompleta

$$F(x;\alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y}\, dy$$

x \ α	1	2	3	4	5	6	7	8	9	10
1	0.632	0.264	0.080	0.019	0.004	0.001	0.000	0.000	0.000	0.000
2	0.865	0.594	0.323	0.143	0.053	0.017	0.005	0.001	0.000	0.000
3	0.950	0.801	0.577	0.353	0.185	0.084	0.034	0.012	0.004	0.001
4	0.982	0.908	0.762	0.567	0.371	0.215	0.111	0.051	0.021	0.008
5	0.993	0.960	0.875	0.735	0.560	0.384	0.238	0.133	0.068	0.032
6	0.998	0.983	0.938	0.849	0.715	0.554	0.394	0.256	0.153	0.084
7	0.999	0.993	0.970	0.918	0.827	0.699	0.550	0.401	0.271	0.170
8	1.000	0.997	0.986	0.958	0.900	0.809	0.687	0.547	0.407	0.283
9		0.999	0.994	0.979	0.945	0.884	0.793	0.676	0.544	0.413
10		1.000	0.997	0.990	0.971	0.933	0.870	0.780	0.667	0.542
11			0.999	0.995	0.985	0.962	0.921	0.857	0.768	0.659
12			1.000	0.998	0.992	0.980	0.954	0.911	0.845	0.758
13				0.999	0.996	0.989	0.974	0.946	0.900	0.834
14				1.000	0.998	0.994	0.986	0.968	0.938	0.891
15					0.999	0.997	0.992	0.982	0.963	0.930

Tabla A.5 Valores críticos para la distribución



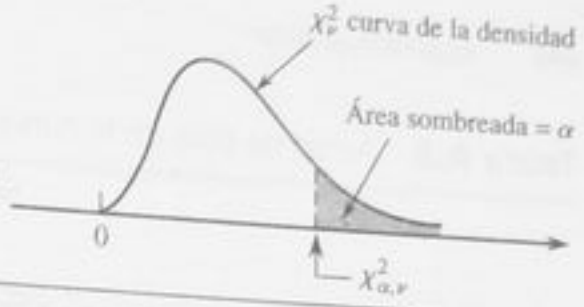
	α						
ν	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Tabla A.6 Valores críticos de la tolerancia para las distribuciones normales de la población

	Intervalos bilaterales						Intervalos unilaterales					
Nivel de confianza	95%			99%			95%			99%		
% de la población capturada	≥90%	≥95%	≥99%	≥90%	≥95%	≥99%	≥90%	≥95%	≥99%	≥90%	≥95%	≥99%
2	32.019	37.674	48.430	160.193	188.491	242.300	20.581	26.260	37.094	103.029	131.426	185.617
3	8.380	9.916	12.861	18.930	22.401	29.055	6.156	7.656	10.553	13.995	17.370	23.896
4	5.369	6.370	8.299	9.398	11.150	14.527	4.162	5.144	7.042	7.380	9.083	12.387
5	4.275	5.079	6.634	6.612	7.855	10.260	3.407	4.203	5.741	5.362	6.578	8.939
6	3.712	4.414	5.775	5.337	6.345	8.301	3.006	3.708	5.062	4.411	5.406	7.335
7	3.369	4.007	5.248	4.613	5.488	7.187	2.756	3.400	4.642	3.859	4.728	6.412
8	3.136	3.732	4.891	4.147	4.936	6.468	2.582	3.187	4.354	3.497	4.285	5.812
9	2.967	3.532	4.631	3.822	4.550	5.966	2.454	3.031	4.143	3.241	3.972	5.389
10	2.839	3.379	4.433	3.582	4.265	5.594	2.355	2.911	3.981	3.048	3.738	5.074
11	2.737	3.259	4.277	3.397	4.045	5.308	2.275	2.815	3.852	2.898	3.556	4.829
12	2.655	3.162	4.150	3.250	3.870	5.079	2.210	2.736	3.747	2.777	3.410	4.633
13	2.587	3.081	4.044	3.130	3.727	4.893	2.155	2.671	3.659	2.677	3.290	4.472
14	2.529	3.012	3.955	3.029	3.608	4.737	2.109	2.615	3.585	2.593	3.189	4.337
15	2.480	2.954	3.878	2.945	3.507	4.605	2.068	2.566	3.520	2.522	3.102	4.222
16	2.437	2.903	3.812	2.872	3.421	4.492	2.033	2.524	3.464	2.460	3.028	4.123
17	2.400	2.858	3.754	2.808	3.345	4.393	2.002	2.486	3.414	2.405	2.963	4.037
18	2.366	2.819	3.702	2.753	3.279	4.307	1.974	2.453	3.370	2.357	2.905	3.960
19	2.337	2.784	3.656	2.703	3.221	4.230	1.949	2.423	3.331	2.314	2.854	3.892
20	2.310	2.752	3.615	2.659	3.168	4.161	1.926	2.396	3.295	2.276	2.808	3.832
25	2.208	2.631	3.457	2.494	2.972	3.904	1.838	2.292	3.158	2.129	2.633	3.601
30	2.140	2.549	3.350	2.385	2.841	3.733	1.777	2.220	3.064	2.030	2.516	3.447
35	2.090	2.490	3.272	2.306	2.748	3.611	1.732	2.167	2.995	1.957	2.430	3.334
40	2.052	2.445	3.213	2.247	2.677	3.518	1.697	2.126	2.941	1.902	2.364	3.249
45	2.021	2.408	3.165	2.200	2.621	3.444	1.669	2.092	2.898	1.857	2.312	3.180
50	1.996	2.379	3.126	2.162	2.576	3.385	1.646	2.065	2.863	1.821	2.269	3.125
60	1.958	2.333	3.066	2.103	2.506	3.293	1.609	2.022	2.807	1.764	2.202	3.038
70	1.929	2.299	3.021	2.060	2.454	3.225	1.581	1.990	2.765	1.722	2.153	2.974
80	1.907	2.272	2.986	2.026	2.414	3.173	1.559	1.965	2.733	1.688	2.114	2.924
90	1.889	2.251	2.958	1.999	2.382	3.130	1.542	1.944	2.706	1.661	2.082	2.883
100	1.874	2.233	2.934	1.977	2.355	3.096	1.527	1.927	2.684	1.639	2.056	2.850
150	1.825	2.175	2.859	1.905	2.270	2.983	1.478	1.870	2.611	1.566	1.971	2.741
200	1.798	2.143	2.816	1.865	2.222	2.921	1.450	1.837	2.570	1.524	1.923	2.679
250	1.780	2.121	2.788	1.839	2.191	2.880	1.431	1.815	2.542	1.496	1.891	2.638
300	1.767	2.106	2.767	1.820	2.169	2.850	1.417	1.800	2.522	1.476	1.868	2.608
∞	1.645	1.960	2.576	1.645	1.960	2.576	1.282	1.645	2.326	1.282	1.645	2.326

Tamaño *n* de la muestra

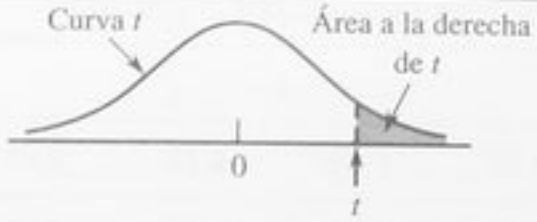
Tabla A.7 Valores críticos para la distribución de chi cuadrada



	α									
ν	**0.995**	**0.99**	**0.975**	**0.95**	**0.90**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

Para $\nu > 40$, $\chi^2_{\alpha,\nu} \approx \nu\left(1 - \frac{2}{9\nu} + z_\alpha\sqrt{\frac{2}{9\nu}}\right)^3$

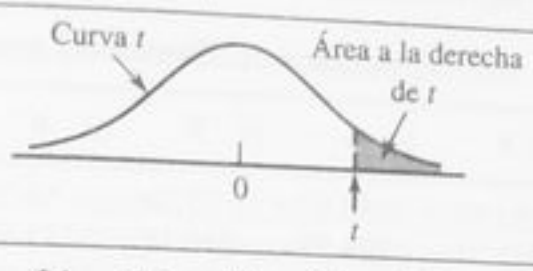
Tabla A.8 Áreas de cola de la curva t



t \ ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0.0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
0.1	0.468	0.465	0.463	0.463	0.462.	0.462	0.462	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461
0.2	0.437	0.430	0.427	0.426	0.425	0.424	0.424	0.423	0.423	0.423	0.423	0.422	0.422	0.422	0.422	0.422	0.422	0.422
0.3	0.407	0.396	0.392	0.390	0.388	0.387	0.386	0.386	0.386	0.385	0.385	0.385	0.384	0.384	0.384	0.384	0.384	0.384
0.4	0.379	0.364	0.358	0.355	0.353	0.352	0.351	0.350	0.349	0.349	0.348	0.348	0.348	0.347	0.347	0.347	0.347	0.347
0.5	0.352	0.333	0.326	0.322	0.319	0.317	0.316	0.315	0.315	0.314	0.313	0.313	0.313	0.312	0.312	0.312	0.312	0.312
0.6	0.328	0.305	0.295	0.290	0.287	0.285	0.284	0.283	0.282	0.281	0.280	0.280	0.279	0.279	0.279	0.278	0.278	0.278
0.7	0.306	0.278	0.267	0.261	0.258	0.255	0.253	0.252	0.251	0.250	0.249	0.249	0.248	0.247	0.247	0.247	0.247	0.246
0.8	0.285	0.254	0.241	0.234	0.230	0.227	0.225	0.223	0.222	0.221	0.220	0.220	0.219	0.218	0.218	0.218	0.217	0.217
0.9	0.267	0.232	0.217	0.210	0.205	0.201	0.199	0.197	0.196	0.195	0.194	0.193	0.192	0.191	0.191	0.191	0.190	0.190
1.0	0.250	0.211	0.196	0.187	0.182	0.178	0.175	0.173	0.172	0.170	0.169	0.169	0.168	0.167	0.167	0.166	0.166	0.165
1.1	0.235	0.193	0.176	0.167	0.162	0.157	0.154	0.152	0.150	0.149	0.147	0.146	0.146	0.144	0.144	0.144	0.143	0.143
1.2	0.221	0.177	0.158	0.148	0.142	0.138	0.135	0.132	0.130	0.129	0.128	0.127	0.126	0.124	0.124	0.124	0.123	0.123
1.3	0.209	0.162	0.142	0.132	0.125	0.121	0.117	0.115	0.113	0.111	0.110	0.109	0.108	0.107	0.107	0.106	0.105	0.105
1.4	0.197	0.148	0.128	0.117	0.110	0.106	0.102	0.100	0.098	0.096	0.095	0.093	0.092	0.091	0.091	0.090	0.090	0.089
1.5	0.187	0.136	0.115	0.104	0.097	0.092	0.089	0.086	0.084	0.082	0.081	0.080	0.079	0.077	0.077	0.077	0.076	0.075
1.6	0.178	0.125	0.104	0.092	0.085	0.080	0.077	0.074	0.072	0.070	0.069	0.068	0.067	0.065	0.065	0.065	0.064	0.064
1.7	0.169	0.116	0.094	0.082	0.075	0.070	0.065	0.064	0.062	0.060	0.059	0.057	0.056	0.055	0.055	0.054	0.054	0.053
1.8	0.161	0.107	0.085	0.073	0.066	0.061	0.057	0.055	0.053	0.051	0.050	0.049	0.048	0.046	0.046	0.045	0.045	0.044
1.9	0.154	0.099	0.077	0.065	0.058	0.053	0.050	0.047	0.045	0.043	0.042	0.041	0.040	0.038	0.038	0.038	0.037	0.037
2.0	0.148	0.092	0.070	0.058	0.051	0.046	0.043	0.040	0.038	0.037	0.035	0.034	0.033	0.032	0.032	0.031	0.031	0.030
2.1	0.141	0.085	0.063	0.052	0.045	0.040	0.037	0.034	0.033	0.031	0.030	0.029	0.028	0.027	0.027	0.026	0.025	0.025
2.2	0.136	0.079	0.058	0.046	0.040	0.035	0.032	0.029	0.028	0.026	0.025	0.024	0.023	0.022	0.022	0.021	0.021	0.021
2.3	0.131	0.074	0.052	0.041	0.035	0.031	0.027	0.025	0.023	0.022	0.021	0.020	0.019	0.018	0.018	0.018	0.017	0.017
2.4	0.126	0.069	0.048	0.037	0.031	0.027	0.024	0.022	0.020	0.019	0.018	0.017	0.016	0.015	0.015	0.014	0.014	0.014
2.5	0.121	0.065	0.044	0.033	0.027	0.023	0.020	0.018	0.017	0.016	0.015	0.014	0.013	0.012	0.012	0.012	0.011	0.011
2.6	0.117	0.061	0.040	0.030	0.024	0.020	0.018	0.016	0.014	0.013	0.012	0.012	0.011	0.010	0.010	0.010	0.009	0.009
2.7	0.113	0.057	0.037	0.027	0.021	0.018	0.015	0.014	0.012	0.011	0.010	0.010	0.009	0.008	0.008	0.008	0.008	0.007
2.8	0.109	0.054	0.034	0.024	0.019	0.016	0.013	0.012	0.010	0.009	0.009	0.008	0.008	0.007	0.007	0.006	0.006	0.006
2.9	0.106	0.051	0.031	0.022	0.017	0.014	0.011	0.010	0.009	0.008	0.007	0.007	0.006	0.005	0.005	0.005	0.005	0.005
3.0	0.102	0.048	0.029	0.020	0.015	0.012	0.010	0.009	0.007	0.007	0.006	0.006	0.005	0.004	0.004	0.004	0.004	0.004
3.1	0.099	0.045	0.027	0.018	0.013	0.011	0.009	0.007	0.006	0.006	0.005	0.005	0.004	0.004	0.004	0.003	0.003	0.003
3.2	0.096	0.043	0.025	0.016	0.012	0.009	0.008	0.006	0.005	0.005	0.004	0.004	0.003	0.003	0.003	0.003	0.003	0.002
3.3	0.094	0.040	0.023	0.015	0.011	0.008	0.007	0.005	0.005	0.004	0.004	0.003	0.003	0.002	0.002	0.002	0.002	0.002
3.4	0.091	0.038	0.021	0.014	0.010	0.007	0.006	0.005	0.004	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002
3.5	0.089	0.036	0.020	0.012	0.009	0.006	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
3.6	0.086	0.035	0.018	0.011	0.008	0.006	0.004	0.004	0.003	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001
3.7	0.084	0.033	0.017	0.010	0.007	0.005	0.004	0.003	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001
3.8	0.082	0.031	0.016	0.010	0.006	0.004	0.003	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.9	0.080	0.030	0.015	0.009	0.006	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
4.0	0.078	0.029	0.014	0.008	0.005	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000

(continúa)

Tabla A.8 Áreas de cola de la curva t (*cont.*)



t \ ν	19	20	21	22	23	24	25	26	27	28	29	30	35	40	60	120	$\infty(=z)$
0.0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
0.1	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.461	0.460	0.460	0.460	0.460	0.460
0.2	0.422	0.422	0.422	0.422	0.422	0.422	0.422	0.422	0.421	0.421	0.421	0.421	0.421	0.421	0.421	0.421	0.421
0.3	0.384	0.384	0.384	0.383	0.383	0.383	0.383	0.383	0.383	0.383	0.383	0.383	0.383	0.383	0.383	0.382	0.382
0.4	0.347	0.347	0.347	0.347	0.346	0.346	0.346	0.346	0.346	0.346	0.346	0.346	0.346	0.346	0.345	0.345	0.345
0.5	0.311	0.311	0.311	0.311	0.311	0.311	0.311	0.311	0.311	0.310	0.310	0.310	0.310	0.310	0.309	0.309	0.309
0.6	0.278	0.278	0.278	0.277	0.277	0.277	0.277	0.277	0.277	0.277	0.277	0.277	0.276	0.276	0.275	0.275	0.274
0.7	0.246	0.246	0.246	0.246	0.245	0.245	0.245	0.245	0.245	0.245	0.245	0.245	0.244	0.244	0.243	0.243	0.242
0.8	0.217	0.217	0.216	0.216	0.216	0.216	0.216	0.215	0.215	0.215	0.215	0.215	0.215	0.214	0.213	0.213	0.212
0.9	0.190	0.189	0.189	0.189	0.189	0.189	0.188	0.188	0.188	0.188	0.188	0.188	0.187	0.187	0.186	0.185	0.184
1.0	0.165	0.165	0.164	0.164	0.164	0.164	0.163	0.163	0.163	0.163	0.163	0.163	0.162	0.162	0.161	0.160	0.159
1.1	0.143	0.142	0.142	0.142	0.141	0.141	0.141	0.141	0.141	0.140	0.140	0.140	0.139	0.139	0.138	0.137	0.136
1.2	0.122	0.122	0.122	0.121	0.121	0.121	0.121	0.120	0.120	0.120	0.120	0.120	0.119	0.119	0.117	0.116	0.115
1.3	0.105	0.104	0.104	0.104	0.103	0.103	0.103	0.103	0.102	0.102	0.102	0.102	0.101	0.101	0.099	0.098	0.097
1.4	0.089	0.089	0.088	0.088	0.087	0.087	0.087	0.087	0.086	0.086	0.086	0.086	0.085	0.085	0.083	0.082	0.081
1.5	0.075	0.075	0.074	0.074	0.074	0.073	0.073	0.073	.073	.072	0.072	0.072	0.071	0.071	0.069	0.068	0.067
1.6	0.063	0.063	0.062	0.062	0.062	0.061	0.061	0.061	0.061	0.060	0.060	0.060	0.059	0.059	0.057	0.056	0.055
1.7	0.053	0.052	0.052	0.052	0.051	0.051	0.051	0.051	0.050	0.050	0.050	0.050	0.049	0.048	0.047	0.046	0.045
1.8	0.044	0.043	0.043	0.043	0.042	0.042	0.042	0.042	0.042	0.041	0.041	0.041	0.040	0.040	0.038	0.037	0.036
1.9	0.036	0.036	0.036	0.035	0.035	0.035	0.035	0.034	0.034	0.034	0.034	0.034	0.033	0.032	0.031	0.030	0.029
2.0	0.030	0.030	0.029	0.029	0.029	0.028	0.028	0.028	0.028	0.028	0.027	0.027	0.027	0.026	0.025	0.024	0.023
2.1	0.025	0.024	0.024	0.024	0.023	0.023	0.023	0.023	0.023	0.022	0.022	0.022	0.022	0.021	0.020	0.019	0.018
2.2	0.020	0.020	0.020	0.019	0.019	0.019	0.019	0.018	0.018	0.018	0.018	0.018	0.017	0.017	0.016	0.015	0.014
2.3	0.016	0.016	0.016	0.016	0.015	0.015	0.015	0.015	0.015	0.015	0.014	0.014	0.014	0.013	0.012	0.012	0.011
2.4	0.013	0.013	0.013	0.013	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.011	0.011	0.011	0.010	0.009	0.008
2.5	0.011	0.011	0.010	0.010	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.009	0.008	0.008	0.007	0.006
2.6	0.009	0.009	0.008	0.008	0.008	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.007	0.006	0.005	0.005
2.7	0.007	0.007	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.004	0.004	0.003
2.8	0.006	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.004	0.004	0.004	0.003	0.003	0.003
2.9	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.003	0.003	0.003	0.003	0.002	0.002
3.0	0.004	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.001
3.1	0.003	0.003	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
3.2	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001
3.3	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000
3.4	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000
3.5	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000
3.6	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000
3.7	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.8	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Tabla A.9 Valores críticos para la distribución de F

ν_2 =grados de libertad del denominador	α	ν_1 =grados de libertad del numerador								
		1	2	3	4	5	6	7	8	9
1	0.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	0.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
	0.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
	0.001	405284	500000	540379	562500	576405	585937	592873	598144	602284
2	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	0.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
3	0.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	0.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	0.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
4	0.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	0.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	0.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	0.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
5	0.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	0.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	0.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	0.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
6	0.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	0.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	0.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	0.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
7	0.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33
8	0.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.001	25.41	18.49	15.83	14.39	13.48	12.86	12.40	12.05	11.77
9	0.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	0.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	0.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	0.001	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11
10	0.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	0.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	0.001	21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96
11	0.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	0.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	0.001	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12
12	0.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	0.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	0.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	0.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48

(continúa)

Tabla A.9 Valores críticos para la distribución de F (*cont.*)

ν_1 =grados de libertad del numerador										
10	**12**	**15**	**20**	**25**	**30**	**40**	**50**	**60**	**120**	**1000**
60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.69	62.79	63.06	63.30
241.88	243.91	245.95	248.01	249.26	250.10	251.14	251.77	252.20	253.25	254.19
6055.8	6106.3	6157.3	6208.7	6239.8	6260.6	6286.8	6302.5	6313.0	6339.4	6362.7
605621	610668	615764	620908	624017	626099	628712	630285	631337	633972	636301
9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.47	9.48	9.49
19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48	19.48	19.49	19.49
99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.48	99.49	99.50
999.40	999.42	999.43	999.45	999.46	999.47	999.47	999.48	999.48	999.49	999.50
5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15	5.15	5.14	5.13
8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.58	8.57	8.55	8.53
27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.35	26.32	26.22	26.14
129.25	128.32	127.37	126.42	125.84	125.45	124.96	124.66	124.47	123.97	123.53
3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.80	3.79	3.78	3.76
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.70	5.69	5.66	5.63
14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.69	13.65	13.56	13.47
48.05	47.41	46.76	46.10	45.70	45.43	45.09	44.88	44.75	44.40	44.09
3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.15	3.14	3.12	3.11
4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.44	4.43	4.40	4.37
10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.24	9.20	9.11	9.03
26.92	26.42	25.91	25.39	25.08	24.87	24.60	24.44	24.33	24.06	23.82
2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.77	2.76	2.74	2.72
4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.75	3.74	3.70	3.67
7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.09	7.06	6.97	6.89
18.41	17.99	17.56	17.12	16.85	16.67	16.44	16.31	16.21	15.98	15.77
2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.52	2.51	2.49	2.47
3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.32	3.30	3.27	3.23
6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.86	5.82	5.74	5.66
14.08	13.71	13.32	12.93	12.69	12.53	12.33	12.20	12.12	11.91	11.72
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.35	2.34	2.32	2.30
3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.02	3.01	2.97	2.93
5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.07	5.03	4.95	4.87
11.54	11.19	10.84	10.48	10.26	10.11	9.92	9.80	9.73	9.53	9.36
2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.22	2.21	2.18	2.16
3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.80	2.79	2.75	2.71
5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.52	4.48	4.40	4.32
9.89	9.57	9.24	8.90	8.69	8.55	8.37	8.26	8.19	8.00	7.84
2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.12	2.11	2.08	2.06
2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.64	2.62	2.58	2.54
4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.12	4.08	4.00	3.92
8.75	8.45	8.13	7.80	7.60	7.47	7.30	7.19	7.12	6.94	6.78
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.04	2.03	2.00	1.98
2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.51	2.49	2.45	2.41
4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.81	3.78	3.69	3.61
7.92	7.63	7.32	7.01	6.81	6.68	6.52	6.42	6.35	6.18	6.02
2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.97	1.96	1.93	1.91
2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.40	2.38	2.34	2.30
4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.57	3.54	3.45	3.37
7.29	7.00	6.71	6.40	6.22	6.09	5.93	5.83	5.76	5.59	5.44

(*continúa*)

Tabla A.9 Valores críticos para la distribución de *F* (*cont.*)

ν_2 =grados de libertad del denominador	α	ν_1 =grados de libertad del numerador								
		1	2	3	4	5	6	7	8	9
13	0.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	0.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	0.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	0.001	17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
14	0.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	0.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	0.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	0.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58
15	0.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	0.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	0.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	0.001	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	0.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	0.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	0.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	0.001	16.12	10.97	9.01	7.94	7.27	6.80	6.46	6.19	5.98
17	0.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	0.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	0.010	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68
	0.001	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75
18	0.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	0.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	0.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	0.001	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	0.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	0.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	0.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	0.001	15.08	10.16	8.28	7.27	6.62	6.18	5.85	5.59	5.39
20	0.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	0.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	0.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	0.001	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24
21	0.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	0.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	0.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	0.001	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	0.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	0.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	0.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	0.001	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	0.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	0.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	0.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	0.001	14.20	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.89
24	0.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	0.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	0.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	0.001	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80

(continúa)

Tabla A.9 Valores críticos para la distribución de *F* (*cont.*)

ν_1 = grados de libertad del numerador										
10	12	15	20	25	30	40	50	60	120	1000
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.92	1.90	1.88	1.85
2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.31	2.30	2.25	2.21
4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.38	3.34	3.25	3.18
6.80	6.52	6.23	5.93	5.75	5.63	5.47	5.37	5.30	5.14	4.99
2.10	2.05	2.01	1.96	1.93	1.91	1.89	1.87	1.86	1.83	1.80
2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.24	2.22	2.18	2.14
3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.22	3.18	3.09	3.02
6.40	6.13	5.85	5.56	5.38	5.25	5.10	5.00	4.94	4.77	4.62
2.06	2.02	1.97	1.92	1.89	1.87	1.85	1.83	1.82	1.79	1.76
2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.18	2.16	2.11	2.07
3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.08	3.05	2.96	2.88
6.08	5.81	5.54	5.25	5.07	4.95	4.80	4.70	4.64	4.47	4.33
2.03	1.99	1.94	1.89	1.86	1.84	1.81	1.79	1.78	1.75	1.72
2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.12	2.11	2.06	2.02
3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.97	2.93	2.84	2.76
5.81	5.55	5.27	4.99	4.82	4.70	4.54	4.45	4.39	4.23	4.08
2.00	1.96	1.91	1.86	1.83	1.81	1.78	1.76	1.75	1.72	1.69
2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.08	2.06	2.01	1.97
3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.87	2.83	2.75	2.66
5.58	5.32	5.05	4.78	4.60	4.48	4.33	4.24	4.18	4.02	3.87
1.98	1.93	1.89	1.84	1.80	1.78	1.75	1.74	1.72	1.69	1.66
2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.04	2.02	1.97	1.92
3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.78	2.75	2.66	2.58
5.39	5.13	4.87	4.59	4.42	4.30	4.15	4.06	4.00	3.84	3.69
1.96	1.91	1.86	1.81	1.78	1.76	1.73	1.71	1.70	1.67	1.64
2.38	2.31	2.23	2.16	2.11	2.07	2.03	2.00	1.98	1.93	1.88
3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.71	2.67	2.58	2.50
5.22	4.97	4.70	4.43	4.26	4.14	3.99	3.90	3.84	3.68	3.53
1.94	1.89	1.84	1.79	1.76	1.74	1.71	1.69	1.68	1.64	1.61
2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.97	1.95	1.90	1.85
3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.64	2.61	2.52	2.43
5.08	4.82	4.56	4.29	4.12	4.00	3.86	3.77	3.70	3.54	3.40
1.92	1.87	1.83	1.78	1.74	1.72	1.69	1.67	1.66	1.62	1.59
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.94	1.92	1.87	1.82
3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.58	2.55	2.46	2.37
4.95	4.70	4.44	4.17	4.00	3.88	3.74	3.64	3.58	3.42	3.28
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.60	1.57
2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.91	1.89	1.84	1.79
3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.53	2.50	2.40	2.32
4.83	4.58	4.33	4.06	3.89	3.78	3.63	3.54	3.48	3.32	3.17
1.89	1.84	1.80	1.74	1.71	1.69	1.66	1.64	1.62	1.59	1.55
2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.88	1.86	1.81	1.76
3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.48	2.45	2.35	2.27
4.73	4.48	4.23	3.96	3.79	3.68	3.53	3.44	3.38	3.22	3.08
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.57	1.54
2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.86	1.84	1.79	1.74
3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.44	2.40	2.31	2.22
4.64	4.39	4.14	3.87	3.71	3.59	3.45	3.36	3.29	3.14	2.99

(continúa)

Tabla A.9 Valores críticos para la distribución de *F* (*cont.*)

ν_2 =grados de libertad del denominador	α	ν_1 =grados de libertad del numerador								
		1	2	3	4	5	6	7	8	9
25	0.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	0.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	0.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	0.001	13.88	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.71
26	0.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	0.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	0.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	0.001	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	0.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	0.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	0.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	0.001	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57
28	0.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	0.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	0.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	0.001	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	0.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	0.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	0.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	0.001	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	0.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	0.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	0.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	0.001	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
40	0.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	0.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	0.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	0.001	12.61	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02
50	0.100	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76
	0.050	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	0.010	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78
	0.001	12.22	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82
60	0.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	0.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	0.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	0.001	11.97	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
100	0.100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69
	0.050	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97
	0.010	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
	0.001	11.50	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.44
200	0.100	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66
	0.050	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93
	0.010	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50
	0.001	11.15	7.15	5.63	4.81	4.29	3.92	3.65	3.43	3.26
1000	0.100	2.71	2.31	2.09	1.95	1.85	1.78	1.72	1.68	1.64
	0.050	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89
	0.010	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43
	0.001	10.89	6.96	5.46	4.65	4.14	3.78	3.51	3.30	3.13

(continúa)

Tabla A.9 Valores críticos para la distribución de *F* (*cont.*)

ν_1 =grados de libertad del numerador									
10	12	15	20	25	30	40	50	60	120	1000
1.87	1.82	1.77	1.72	1.68	1.66	1.63	1.61	1.59	1.56	1.52
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.84	1.82	1.77	1.72
3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.40	2.36	2.27	2.18
4.56	4.31	4.06	3.79	3.63	3.52	3.37	3.28	3.22	3.06	2.91
1.86	1.81	1.76	1.71	1.67	1.65	1.61	1.59	1.58	1.54	1.51
2.22	2.15	2.07	1.99	1.94	1.90	1.85	1.82	1.80	1.75	1.70
3.09	2.96	2.81	2.66	2.57	2.50	2.42	2.36	2.33	2.23	2.14
4.48	4.24	3.99	3.72	3.56	3.44	3.30	3.21	3.15	2.99	2.84
1.85	1.80	1.75	1.70	1.66	1.64	1.60	1.58	1.57	1.53	1.50
2.20	2.13	2.06	1.97	1.92	1.88	1.84	1.81	1.79	1.73	1.68
3.06	2.93	2.78	2.63	2.54	2.47	2.38	2.33	2.29	2.20	2.11
4.41	4.17	3.92	3.66	3.49	3.38	3.23	3.14	3.08	2.92	2.78
1.84	1.79	1.74	1.69	1.65	1.63	1.59	1.57	1.56	1.52	1.48
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.79	1.77	1.71	1.66
3.03	2.90	2.75	2.60	2.51	2.44	2.35	2.30	2.26	2.17	2.08
4.35	4.11	3.86	3.60	3.43	3.32	3.18	3.09	3.02	2.86	2.72
1.83	1.78	1.73	1.68	1.64	1.62	1.58	1.56	1.55	1.51	1.47
2.18	2.10	2.03	1.94	1.89	1.85	1.81	1.77	1.75	1.70	1.65
3.00	2.87	2.73	2.57	2.48	2.41	2.33	2.27	2.23	2.14	2.05
4.29	4.05	3.80	3.54	3.38	3.27	3.12	3.03	2.97	2.81	2.66
1.82	1.77	1.72	1.67	1.63	1.61	1.57	1.55	1.54	1.50	1.46
2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.76	1.74	1.68	1.63
2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.25	2.21	2.11	2.02
4.24	4.00	3.75	3.49	3.33	3.22	3.07	2.98	2.92	2.76	2.61
1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.42	1.38
2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.66	1.64	1.58	1.52
2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.06	2.02	1.92	1.82
3.87	3.64	3.40	3.14	2.98	2.87	2.73	2.64	2.57	2.41	2.25
1.73	1.68	1.63	1.57	1.53	1.50	1.46	1.44	1.42	1.38	1.33
2.03	1.95	1.87	1.78	1.73	1.69	1.63	1.60	1.58	1.51	1.45
2.70	2.56	2.42	2.27	2.17	2.10	2.01	1.95	1.91	1.80	1.70
3.67	3.44	3.20	2.95	2.79	2.68	2.53	2.44	2.38	2.21	2.05
1.71	1.66	1.60	1.54	1.50	1.48	1.44	1.41	1.40	1.35	1.30
1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.56	1.53	1.47	1.40
2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.88	1.84	1.73	1.62
3.54	3.32	3.08	2.83	2.67	2.55	2.41	2.32	2.25	2.08	1.92
1.66	1.61	1.56	1.49	1.45	1.42	1.38	1.35	1.34	1.28	1.22
1.93	1.85	1.77	1.68	1.62	1.57	1.52	1.48	1.45	1.38	1.30
2.50	2.37	2.22	2.07	1.97	1.89	1.80	1.74	1.69	1.57	1.45
3.30	3.07	2.84	2.59	2.43	2.32	2.17	2.08	2.01	1.83	1.64
1.63	1.58	1.52	1.46	1.41	1.38	1.34	1.31	1.29	1.23	1.16
1.88	1.80	1.72	1.62	1.56	1.52	1.46	1.41	1.39	1.30	1.21
2.41	2.27	2.13	1.97	1.87	1.79	1.69	1.63	1.58	1.45	1.30
3.12	2.90	2.67	2.42	2.26	2.15	2.00	1.90	1.83	1.64	1.43
1.61	1.55	1.49	1.43	1.38	1.35	1.30	1.27	1.25	1.18	1.08
1.84	1.76	1.68	1.58	1.52	1.47	1.41	1.36	1.33	1.24	1.11
2.34	2.20	2.06	1.90	1.79	1.72	1.61	1.54	1.50	1.35	1.16
2.99	2.77	2.54	2.30	2.14	2.02	1.87	1.77	1.69	1.49	1.22

Tabla A.10 Valores críticos para la distribución del rango estudentizado

		m										
ν	α	2	3	4	5	6	7	8	9	10	11	12
5	0.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32
	0.01	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70
6	0.05	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79
	0.01	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.48
7	0.05	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43
	0.01	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71
8	0.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18
	0.01	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03	8.18
9	0.05	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98
	0.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65	7.78
10	0.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83
	0.01	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.49
11	0.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71
	0.01	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25
12	0.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61
	0.01	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06
13	0.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53
	0.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90
14	0.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46
	0.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77
15	0.05	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40
	0.01	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66
16	0.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35
	0.01	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56
17	0.05	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21	5.31
	0.01	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48
18	0.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27
	0.01	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41
19	0.05	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23
	0.01	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34
20	0.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20
	0.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.28
24	0.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10
	0.01	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11
30	0.05	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92	5.00
	0.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93
40	0.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82	4.90
	0.01	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69	5.76
60	0.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81
	0.01	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60
120	0.05	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71
	0.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37	5.44
∞	0.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62
	0.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29

Tabla A.11 Áreas de colas de la curva chi cuadrada

Área de cola superior	$\nu = 1$	$\nu = 2$	$\nu = 3$	$\nu = 4$	$\nu = 5$
> 0.100	< 2.70	< 4.60	< 6.25	< 7.77	< 9.23
0.100	2.70	4.60	6.25	7.77	9.23
0.095	2.78	4.70	6.36	7.90	9.37
0.090	2.87	4.81	6.49	8.04	9.52
0.085	2.96	4.93	6.62	8.18	9.67
0.080	3.06	5.05	6.75	8.33	9.83
0.075	3.17	5.18	6.90	8.49	10.00
0.070	3.28	5.31	7.06	8.66	10.19
0.065	3.40	5.46	7.22	8.84	10.38
0.060	3.53	5.62	7.40	9.04	10.59
0.055	3.68	5.80	7.60	9.25	10.82
0.050	3.84	5.99	7.81	9.48	11.07
0.045	4.01	6.20	8.04	9.74	11.34
0.040	4.21	6.43	8.31	10.02	11.64
0.035	4.44	6.70	8.60	10.34	11.98
0.030	4.70	7.01	8.94	10.71	12.37
0.025	5.02	7.37	9.34	11.14	12.83
0.020	5.41	7.82	9.83	11.66	13.38
0.015	5.91	8.39	10.46	12.33	14.09
0.010	6.63	9.21	11.34	13.27	15.08
0.005	7.87	10.59	12.83	14.86	16.74
0.001	10.82	13.81	16.26	18.46	20.51
< 0.001	> 10.82	> 13.81	> 16.26	> 18.46	> 20.51

Área de cola superior	$\nu = 6$	$\nu = 7$	$\nu = 8$	$\nu = 9$	$\nu = 10$
> 0.100	< 10.64	< 12.01	< 13.36	< 14.68	< 15.98
0.100	10.64	12.01	13.36	14.68	15.98
0.095	10.79	12.17	13.52	14.85	16.16
0.090	10.94	12.33	13.69	15.03	16.35
0.085	11.11	12.50	13.87	15.22	16.54
0.080	11.28	12.69	14.06	15.42	16.75
0.075	11.46	12.88	14.26	15.63	16.97
0.070	11.65	13.08	14.48	15.85	17.20
0.065	11.86	13.30	14.71	16.09	17.44
0.060	12.08	13.53	14.95	16.34	17.71
0.055	12.33	13.79	15.22	16.62	17.99
0.050	12.59	14.06	15.50	16.91	18.30
0.045	12.87	14.36	15.82	17.24	18.64
0.040	13.19	14.70	16.17	17.60	19.02
0.035	13.55	15.07	16.56	18.01	19.44
0.030	13.96	15.50	17.01	18.47	19.92
0.025	14.44	16.01	17.53	19.02	20.48
0.020	15.03	16.62	18.16	19.67	21.16
0.015	15.77	17.39	18.97	20.51	22.02
0.010	16.81	18.47	20.09	21.66	23.20
0.005	18.54	20.27	21.95	23.58	25.18
0.001	22.45	24.32	26.12	27.87	29.58
< 0.001	> 22.45	> 24.32	> 26.12	> 27.87	> 29.58

(continúa)

Tabla A.11 Áreas de colas de la curva chi cuadrada (*cont.*)

Área de cola superior	$\nu = 11$	$\nu = 12$	$\nu = 13$	$\nu = 14$	$\nu = 15$
> 0.100	< 17.27	< 18.54	< 19.81	< 21.06	< 22.30
0.100	17.27	18.54	19.81	21.06	22.30
0.095	17.45	18.74	20.00	21.26	22.51
0.090	17.65	18.93	20.21	21.47	22.73
0.085	17.85	19.14	20.42	21.69	22.95
0.080	18.06	19.36	20.65	21.93	23.19
0.075	18.29	19.60	20.89	22.17	23.45
0.070	18.53	19.84	21.15	22.44	23.72
0.065	18.78	20.11	21.42	22.71	24.00
0.060	19.06	20.39	21.71	23.01	24.31
0.055	19.35	20.69	22.02	23.33	24.63
0.050	19.67	21.02	22.36	23.68	24.99
0.045	20.02	21.38	22.73	24.06	25.38
0.040	20.41	21.78	23.14	24.48	25.81
0.035	20.84	22.23	23.60	24.95	26.29
0.030	21.34	22.74	24.12	25.49	26.84
0.025	21.92	23.33	24.73	26.11	27.48
0.020	22.61	24.05	25.47	26.87	28.25
0.015	23.50	24.96	26.40	27.82	29.23
0.010	24.72	26.21	27.68	29.14	30.57
0.005	26.75	28.29	29.81	31.31	32.80
0.001	31.26	32.90	34.52	36.12	37.69
< 0.001	> 31.26	> 32.90	> 34.52	> 36.12	> 37.69

Área de cola superior	$\nu = 16$	$\nu = 17$	$\nu = 18$	$\nu = 19$	$\nu = 20$
> 0.100	< 23.54	< 24.77	< 25.98	< 27.20	< 28.41
0.100	23.54	24.76	25.98	27.20	28.41
0.095	23.75	24.98	26.21	27.43	28.64
0.090	23.97	25.21	26.44	27.66	28.88
0.085	24.21	25.45	26.68	27.91	29.14
0.080	24.45	25.70	26.94	28.18	29.40
0.075	24.71	25.97	27.21	28.45	29.69
0.070	24.99	26.25	27.50	28.75	29.99
0.065	25.28	26.55	27.81	29.06	30.30
0.060	25.59	26.87	28.13	29.39	30.64
0.055	25.93	27.21	28.48	29.75	31.01
0.050	26.29	27.58	28.86	30.14	31.41
0.045	26.69	27.99	29.28	30.56	31.84
0.040	27.13	28.44	29.74	31.03	32.32
0.035	27.62	28.94	30.25	31.56	32.85
0.030	28.19	29.52	30.84	32.15	33.46
0.025	28.84	30.19	31.52	32.85	34.16
0.020	29.63	30.99	32.34	33.68	35.01
0.015	30.62	32.01	33.38	34.74	36.09
0.010	32.00	33.40	34.80	36.19	37.56
0.005	34.26	35.71	37.15	38.58	39.99
0.001	39.25	40.78	42.31	43.81	45.31
< 0.001	> 39.25	> 40.78	> 42.31	> 43.81	> 45.31

Tabla A.12 Valores críticos para la prueba de normalidad de Ryan-Joiner

		α		
		0.10	**0.05**	**0.01**
n	**5**	0.9033	0.8804	0.8320
	10	0.9347	0.9180	0.8804
	15	0.9506	0.9383	0.9110
	20	0.9600	0.9503	0.9290
	25	0.9662	0.9582	0.9408
	30	0.9707	0.9639	0.9490
	40	0.9767	0.9715	0.9597
	50	0.9807	0.9764	0.9664
	60	0.9835	0.9799	0.9710
	75	0.9865	0.9835	0.9757

Tabla A.13 Valores críticos para la prueba de rangos con signo de Wilcoxon

$P_0(S_+ \geq c_1) = P(S_+ \geq c_1$ cuando H_0 es verdadera)

n	c_1	$P_0(S_+ \geq c_1)$
3	6	0.125
4	9	0.125
	10	0.062
5	13	0.094
	14	0.062
	15	0.031
6	17	0.109
	19	0.047
	20	0.031
	21	0.016
7	22	0.109
	24	0.055
	26	0.023
	28	0.008
8	28	0.098
	30	0.055
	32	0.027
	34	0.012
	35	0.008
	36	0.004
9	34	0.102
	37	0.049
	39	0.027
	42	0.010
	44	0.004
10	41	0.097
	44	0.053
	47	0.024
	50	0.010
	52	0.005
11	48	0.103
	52	0.051
	55	0.027
	59	0.009
	61	0.005
12	56	0.102
	60	0.055
	61	0.046
	64	0.026
	68	0.010
	71	0.005
13	64	0.108
	65	0.095
	69	0.055
	70	0.047
	74	0.024
14	78	0.011
	79	0.009
	81	0.005
	73	0.108
	74	0.097
	79	0.052
	84	0.025
	89	0.010
	92	0.005
15	83	0.104
	84	0.094
	89	0.053
	90	0.047
	95	0.024
	100	0.011
	101	0.009
	104	0.005
16	93	0.106
	94	0.096
	100	0.052
	106	0.025
	112	0.011
	113	0.009
	116	0.005
17	104	0.103
	105	0.095
	112	0.049
	118	0.025
	125	0.010
	129	0.005
18	116	0.098
	124	0.049
	131	0.024
	138	0.010
	143	0.005
19	128	0.098
	136	0.052
	137	0.048
	144	0.025
	152	0.010
	157	0.005
20	140	0.101
	150	0.049
	158	0.024
	167	0.010
	172	0.005

Tabla A.14 Valores críticos para la prueba de suma de rangos de Wilcoxon

$P_0(W \geq c) = P(W \geq c$ cuando H_0 es verdadera)

m	n	c	$P_0(W \geq c)$
3	3	15	0.05
	4	17	0.057
		18	0.029
	5	20	0.036
		21	0.018
	6	22	0.048
		23	0.024
		24	0.012
	7	24	0.058
		26	0.017
		27	0.008
	8	27	0.042
		28	0.024
		29	0.012
		30	0.006
4	4	24	0.057
		25	0.029
		26	0.014
	5	27	0.056
		28	0.032
		29	0.016
		30	0.008
	6	30	0.057
		32	0.019
		33	0.010
		34	0.005
	7	33	0.055
		35	0.021
		36	0.012
		37	0.006
	8	36	0.055
		38	0.024
		40	0.008
		41	0.004
5	5	36	0.048
		37	0.028
		39	0.008
		40	0.004
	6	40	0.041
		41	0.026
		43	0.009
		44	0.004
	7	43	0.053
		45	0.024
		47	0.009
		48	0.005
	8	47	0.047
		49	0.023
		51	0.009
		52	0.005
6	6	50	0.047
		52	0.021
		54	0.008
		55	0.004
	7	54	0.051
		56	0.026
		58	0.011
		60	0.004
	8	58	0.054
		61	0.021
		63	0.01
		65	0.004
7	7	66	0.049
		68	0.027
		71	0.009
		72	0.006
	8	71	0.047
		73	0.027
		76	0.01
		78	0.005
8	8	84	0.052
		87	0.025
		90	0.01
		92	0.005

Tabla A.15 Valores críticos para el intervalo de rangos con signo de Wilcoxon $(\bar{x}_{(n(n+1)/2-c+1)}, \bar{x}_{(c)})$

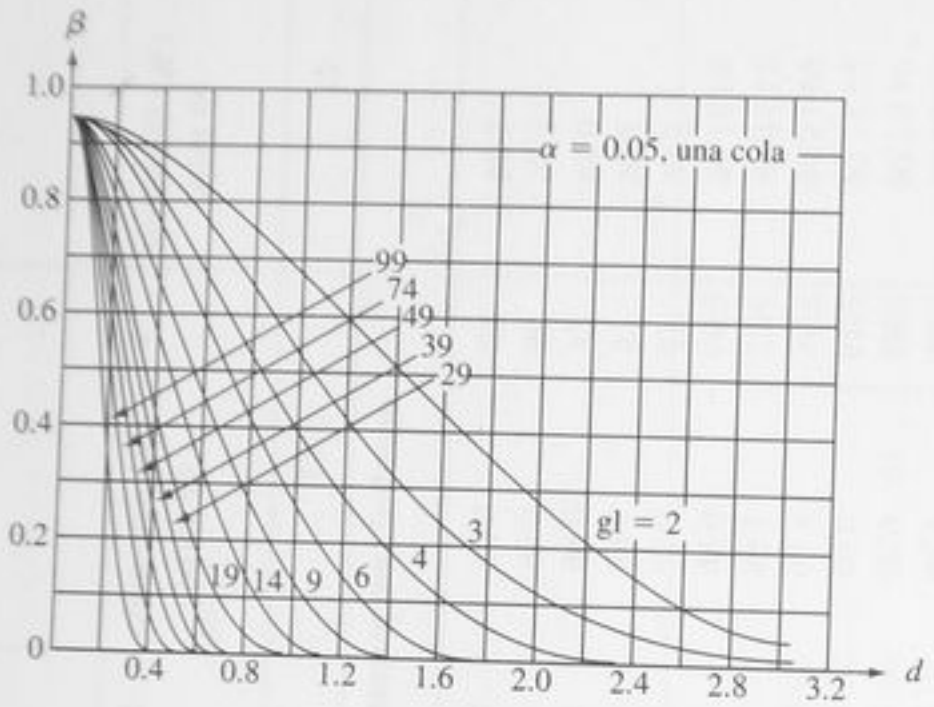
n	Nivel de confianza (%)	c	n	Nivel de confianza (%)	c	n	Nivel de confianza (%)	c
5	93.8	15	13	99.0	81	20	99.1	173
	87.5	14		95.2	74		95.2	158
6	96.9	21		90.6	70		90.3	150
	93.7	20	14	99.1	93	21	99.0	188
	90.6	19		95.1	84		95.0	172
7	98.4	28		89.6	79		89.7	163
	95.3	26	15	99.0	104	22	99.0	204
	89.1	24		95.2	95		95.0	187
8	99.2	36		90.5	90		90.2	178
	94.5	32	16	99.1	117	23	99.0	221
	89.1	30		94.9	106		95.2	203
9	99.2	44		89.5	100		90.2	193
	94.5	39	17	99.1	130	24	99.0	239
	90.2	37		94.9	118		95.1	219
10	99.0	52		90.2	112		89.9	208
	95.1	47	18	99.0	143	25	99.0	257
	89.5	44		95.2	131		95.2	236
11	99.0	61		90.1	124		89.9	224
	94.6	55	19	99.1	158			
	89.8	52		95.1	144			
12	99.1	71		90.4	137			
	94.8	64						
	90.8	61						

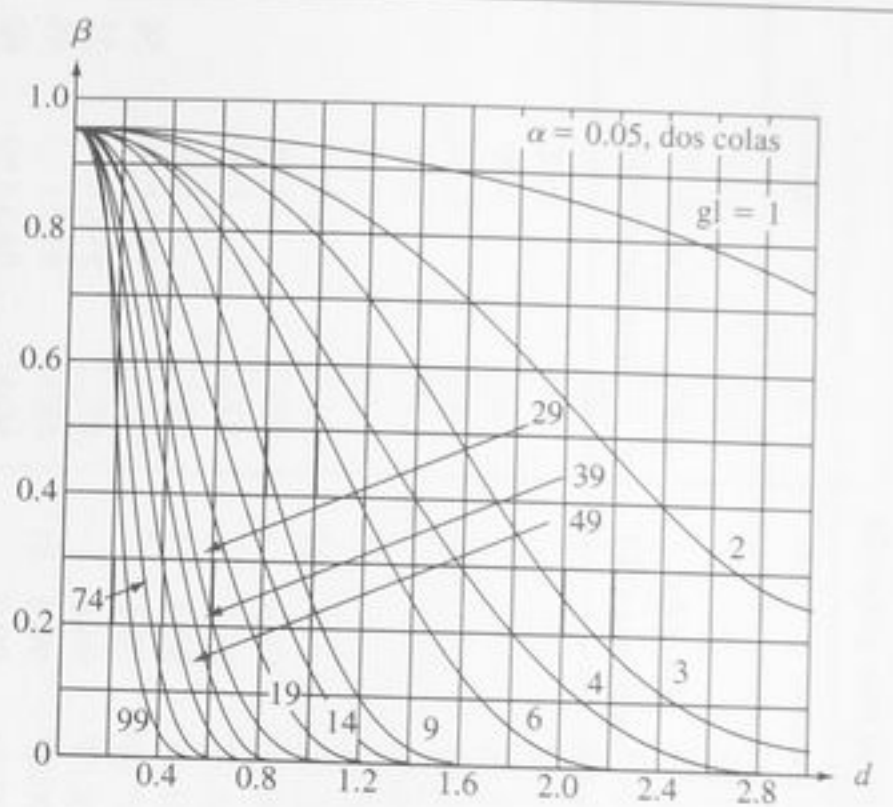
Tabla A.16 Valores críticos para el intervalo de suma de rangos de Wilcoxon $(d_{(g(mn-c+1))}, d_{(g(c))})$

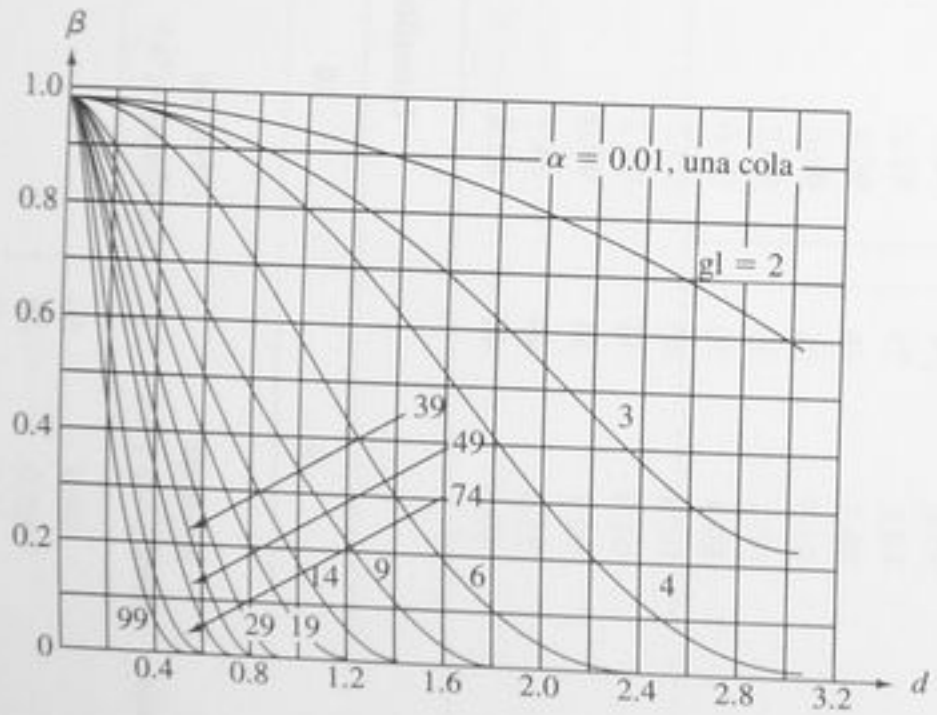
	Tamaño muestral más pequeño							
	5		6		7		8	
Tamaño muestral más grande	Nivel de confianza (%)	c	Nivel de confianza (%)	c	Nivel de confianza (%)	c	Nivel de confianza (%)	c
5	99.2	25						
	94.4	22						
	90.5	21						
6	99.1	29	99.1	34				
	94.8	26	95.9	31				
	91.8	25	90.7	29				
7	99.0	33	99.2	39	98.9	44		
	95.2	30	94.9	35	94.7	40		
	89.4	28	89.9	33	90.3	38		
8	98.9	37	99.2	44	99.1	50	99.0	56
	95.5	34	95.7	40	94.6	45	95.0	51
	90.7	32	89.2	37	90.6	43	89.5	48
9	98.8	41	99.2	49	99.2	56	98.9	62
	95.8	38	95.0	44	94.5	50	95.4	57
	88.8	35	91.2	42	90.9	48	90.7	54
10	99.2	46	98.9	53	99.0	61	99.1	69
	94.5	41	94.4	48	94.5	55	94.5	62
	90.1	39	90.7	46	89.1	52	89.9	59
11	99.1	50	99.0	58	98.9	66	99.1	75
	94.8	45	95.2	53	95.6	61	94.9	68
	91.0	43	90.2	50	89.6	57	90.9	65
12	99.1	54	99.0	63	99.0	72	99.0	81
	95.2	49	94.7	57	95.5	66	95.3	74
	89.6	46	89.8	54	90.0	62	90.2	70

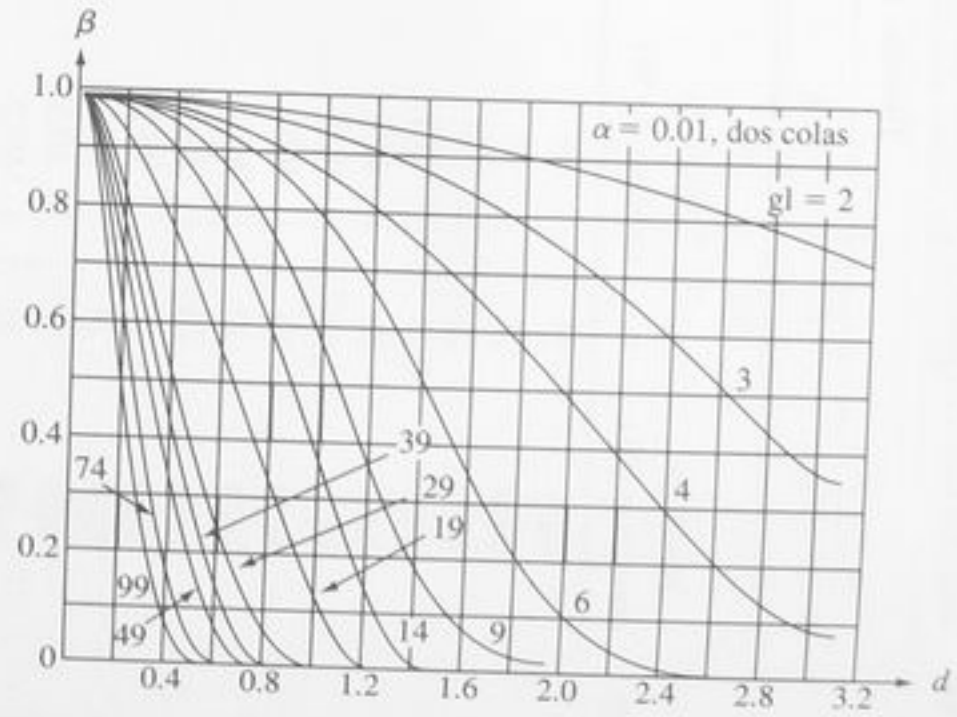
	Tamaño muestral más pequeño							
	9		10		11		12	
Tamaño muestral más grande	Nivel de confianza (%)	c	Nivel de confianza (%)	c	Nivel de confianza (%)	c	Nivel de confianza (%)	c
9	98.9	69						
	95.0	63						
	90.6	60						
10	99.0	76	99.1	84				
	94.7	69	94.8	76				
	90.5	66	89.5	72				
11	99.0	83	99.0	91	98.9	99		
	95.4	76	94.9	83	95.3	91		
	90.5	72	90.1	79	89.9	86		
12	99.1	90	99.1	99	99.1	108	99.0	116
	95.1	82	95.0	90	94.9	98	94.8	106
	90.5	78	90.7	86	89.6	93	89.9	101

Tabla A.17 Curvas β para pruebas t









gl = grados de libertad

Respuestas a ejercicios con número impar

Capítulo 1

1. **a.** *Los Angeles Times, Oberlin Tribune, Gainesville Sun, Washington Post*
   b. Duke Energy, Clorox, Seagate, Neiman Marcus
   c. Vince Correa, Catherine Miller, Michael Cutler, Ken Lee
   d. 2.97, 3.56, 2.20, 2.97

3. **a.** ¿Qué tan probable es que más de la mitad de las videograbadoras muestreadas requieran o hayan necesitado servicio de garantía? ¿Cuál es el número esperado entre las 100 que necesitan servicio de garantía? ¿Cuán probable es que el número que requiere servicio de garantía sea mayor que el número esperado en más de 10?
   b. Suponga que 15 de las 100 muestreadas necesitan servicio de garantía. ¿Qué tan confiado se puede estar de que la proporción de *todas* las videograbadoras que necesitan servicio esté entre 0.08 y 0.22? ¿La muestra proporciona evidencia convincente para concluir que más de 10% de las videograbadoras necesitan servicio de garantía?

5. **a.** No. Los alumnos del grupo grande de estadística que participan en un programa de instrucción complementaria de esta clase.
   b. La aleatorización protege contra diversos sesgos y ayuda a asegurar que quienes están en el grupo de instrucción complementaria sean tan parecidos como sea posible a los alumnos del grupo de control.
   c. No habría una base firme para evaluar la eficacia de la instrucción complementaria (nada con qué comparar en forma razonable las calificaciones de la instrucción complementaria).

7. Se podría generar una muestra aleatoria simple de los hogares con una sola familia en la ciudad, o una muestra aleatoria estratificada al tomar una muestra aleatoria simple de cada uno de los 10 distritos. De cada uno de los hogares seleccionados se determinarán los valores de todas las variables deseadas. Esto sería un estudio enumerativo porque existe una población finita e identificable de objetos en los que se realizaría el muestreo.

9. **a.** Es posible que haya error de medición, error de registro, diferencias en las condiciones ambientales al momento de la medición, etcétera.
   b. No. No hay marco de muestreo.

11.

6L	430
6H	769689
7L	42014202
7H	
8L	011211410342
8H	9595578
9L	30
9H	58

El espacio en los datos; ninguna calificación entre 70 y 80.

13. **a.**
```
12 2                    hoja: dígito de unidades
12 445
12 6667777
12 889999
13 00011111111
13 22222222233333333333333
13 44444444444444444455555555555555555555
13 6666666666667777777777
13 888888888888999999
14 0000001111
14 2333333
14 444
14 77
simetría
```
   b. Se asemeja a la forma de campana, centro □ 135, ninguna dispersión insignificante, ningún espacio o valor atípico.

15.

Crujiente		Cremosa
	2	2
644	3	0069
77220	4	00145
6320	5	003666
222	6	258
55	7	
0	8	

tallo: decenas
hoja: unidades

Ambos conjuntos de puntuaciones están bastante dispersos. Parece que no hay valores atípicos. La distribución de las puntuaciones de la mantequilla crujiente al parecer se desplaza a la derecha (hacia valores más grandes) de las puntuaciones de la mantequilla cremosa en más o menos 10.

17. a. # de transductores defectuosos

	Frecuencia	Frec. rel.
0	7	0.117
1	12	0.200
2	13	0.217
3	14	0.233
4	6	0.100
5	3	0.050
6	3	0.050
7	1	0.017
8	1	0.017
	60	1.001

b. 0.917, 0.867, $1 - 0.867 = 0.133$

c. El histograma tiene un sesgo positivo sustancial. Está centrado en algún lugar entre 2 y 3, y se dispersa mucho respecto de su centro.

19. a. 0.99 (99%), 0.71 (71%) **b.** 0.64 (64%), 0.44 (44%)

c. En sentido estricto, el histograma no es unimodal, pero está muy cerca de serlo con un sesgo positivo moderado. Probablemente con una muestra mucho más grande se obtendría una ilustración más uniforme.

21. a.

y	Frec.	Frec. rel.
0	17	0.362
1	22	0.468
2	6	0.128
3	1	0.021
4	0	0.000
5	1	0.021
	47	1.000

0.362, 0.638

b.

z	Frec.	Frec. rel.
0	13	0.277
1	11	0.234
2	3	0.064
3	7	0.149
4	5	0.106
5	3	0.064
6	3	0.064
7	0	0.000
8	2	0.043
	47	1.001

0.894, 0.830

23. a.

Clase	Frec.	Frec. rel.
0–<100	21	0.21
100–<200	32	0.32
200–<300	26	0.26
300–<400	12	0.12
400–<500	4	0.04
500–<600	3	0.03
600–<700	1	0.01
700–<800	0	0.00
800–<900	1	0.01
	100	1.00

b.

Clase	Frec.	Frec. rel.	Densidad
0–<50	8	0.08	0.0016
50–<100	13	0.13	0.0026
100–<150	11	0.11	0.0022
150–<200	21	0.21	0.0042
200–<300	26	0.26	0.0026
300–<400	12	0.12	0.0012
400–<500	4	0.04	0.0004
500–<600	3	0.03	0.0003
600–<900	2	0.02	0.00007
	100	1.00	

c. 0.79

25.

Clase	Frec.	Clase	Frec.
10–<20	8	1.1–<1.2	2
20–<30	14	1.2–<1.3	6
30–<40	8	1.3–<1.4	7
40–<50	4	1.4–<1.5	9
50–<60	3	1.5–<1.6	6
60–<70	2	1.6–<1.7	4
70–<80	1	1.7–<1.8	5
	40	1.8–<1.9	1
			40

Original: con sesgo positivo;
Transformado: mucho más simétrico, no lejos de la forma de campana.

27. a. La observación 50 cae en el límite de clase.

b.

Clase	Frec.	Frec. rel.
0–<50	9	0.18
50–<100	19	0.38
100–<150	11	0.22
150–<200	4	0.08
200–<300	4	0.08
300–<400	2	0.04
400–<500	0	0.00
500–<600	1	0.02
	50	1.00

Un valor representativo (central) está un poco abajo o ligeramente arriba de 100, dependiendo de cómo se centre una de las medidas. Hay mucha variabilidad en los tiempos de vida, en particular en los valores del extremo derecho de los datos. Hay varios candidatos para los valores atípicos.

c.

Clase	Frec.	Frec. rel.
2.25–<2.75	2	0.04
2.75–<3.25	2	0.04
3.25–<3.75	3	0.06
3.75–<4.25	8	0.16
4.25–<4.75	18	0.36
4.75–<5.25	10	0.20
5.25–<5.75	4	0.08
5.75–<6.25	3	0.06
	50	1.00

Hay mucha más simetría en la distribución de los valores $\ln(x)$ que en los valores de x, y menos variabilidad. Ya no hay espacios o valores atípicos evidentes.

d. 0.38, 0.14

29.

Dolencia	Frec.	Frec. rel.
J	10	0.1667
F	9	0.1500
B	7	0.1167
M	4	0.0667
C	3	0.0500
N	6	0.1000
O	21	0.3500
	60	1.0001

31.

Clase	Frec.	Frec. acumulada	Frec. rel. acumulada
0–<4	2	2	0.050
4–<8	14	16	0.400
8–<12	11	27	0.675
12–<16	8	35	0.875
16–<20	4	39	0.975
20–<24	0	39	0.975
24–<28	1	40	1.000

33. a. $\bar{x} = 192.57$, $\tilde{x} = 189.0$

b. Nueva $\bar{x} = 189.71$; $\tilde{x}$ sin cambio

c. 191.0, 7.14% **d.** 122.6

35. a. $\bar{x} = 12.55$, $\tilde{x} = 12.50$, $\bar{x}_{rec(12.5)} = 12.40$. La eliminación de la observación más grande (18.0) hace que $\bar{x}$ y $\tilde{x}_{rec}$ sean un poco más pequeñas que $\bar{x}$.

b. En a lo sumo 4.0 **c.** No; se multiplican los valores de $\bar{x}$ y $\tilde{x}$ por el factor de conversión 1/2.2.

37. $\bar{x}_{rec(10)} = 11.46$

39. a. $\bar{x} = 1.0297$, $\tilde{x} = 1.009$ **b.** 0.383

41. a. 0.7 **b.** También 0.7 **c.** 13

43. $\tilde{x} = 68.0$, $\bar{x}_{rec(20)} = 66.2$, $\bar{x}_{rec(30)} = 67.5$

45. a. $\bar{x} = 115.58$; las desviaciones son 0.82, 0.32, −0.98, −0.38, 0.22

b. 0.482, 0.694 **c.** 0.482 **d.** 0.482

47. $\bar{x} = 116.2$, $s = 25.75$. La magnitud de s indica que hay una cantidad considerable de variación respecto al centro (una desviación "representativa" aproximadamente igual a 25).

49. a. 0.56.80, 197.8040 **b.** 0.5016, 0.708

51. a. 1264.766, 35.564 **b.** 0.351, 0.593

53. a. 2.74, 3.88 **b.** 1.14 **c.** Sin cambio

d. A lo sumo 0.40 **e.** 1.19

55. a. 33 **b.** No

c. Ligero sesgo positivo en la mitad intermedia, pero en general bastante simétrico. Al parecer es considerable el grado de variabilidad.

d. A lo sumo 32

57. a. Sí. 125.8 es un valor atípico extremo y 250.2 es un valor atípico moderado.

b. Además de la presencia de valores atípicos, hay sesgo positivo tanto en el 50% intermedio de los datos como en general, a excepción de los valores atípicos. Excepto por dos valores atípicos, parece haber una cantidad relativamente pequeña de variabilidad en los datos.

59. a. DE: 0.4, 0.10, 2.75, 2.65;
Sin DE: 1.60, 0.30, 7.90, 7.60

b. DE: 8.9 y 9.2 son valores atípicos moderados, y 11.7 y 21.0 son valores atípicos extremos.
No hay valores atípicos en la muestra de sin DE.

c. Cuatro valores atípicos para DE, ninguno para sin DE. Sesgo positivo considerable en ambas muestras; menos variabilidad en DE (menor f_s), y las observaciones sin DE tienden a ser un poco más grandes que las observaciones DE.

61. Los valores atípicos, tanto leves como extremos, sólo a las 6 A.M. Las distribuciones en otros tiempos son bastante simétricas. La variabilidad aumenta un poco hasta las 2 P.M. y después disminuye ligeramente, y lo mismo se cumple para valores del coeficiente de vapor de la gasolina "típica".

63. Con base en un diagrama de caja comparativo, al parecer las tres muestras difieren de manera notable respecto al centro. Hay un poco menos de variabilidad para el flujo 160 que para los otros dos. No hay valores atípicos. Hay evidencia de sesgo positivo leve en el 50% intermedio de cada muestra, pero no para la de flujo 160 ni para toda la muestra de flujo 200.

65. **a.** 9.59, 59.41
b. $CV = 0.396$ para HC y $= 0.323$ para CO, por lo que hay más variabilidad relativa en los datos de HC aun cuando su desviación estándar es mucho menor.

67. 10.65

69. **a.** $\bar{y} = a\bar{x} + b$, $s_y^2 = a^2 s_x^2$ **b.** 189.14, 1.87

71. **a.** La media, la mediana y la media recortada son casi idénticas, lo que indica una cantidad sustancial de simetría en los datos; el hecho de que los cuartiles estén más o menos a la misma distancia de la media y que las observaciones menor y mayor estén aproximadamente a la misma distancia del centro proporciona un respaldo adicional para la simetría. La desviación estándar es bastante pequeña en relación con la media y la mediana.
b. Véanse los comentarios del inciso (a). Además, si se usa $1.5(Q3 - Q1)$ como criterio, las dos observaciones más grandes y las tres más pequeñas son valores atípicos moderados.

73. $\bar{x} = 0.9255$, $s = 0.0809$, $\tilde{x} = 0.93$, variabilidad pequeña y un poco de sesgo.

75. **a.** Los "resúmenes de 5 números" ($\tilde{x}$, los dos cuartiles y las observaciones menor y mayor) son idénticos y no hay valores atípicos, por lo que los tres diagramas de caja son idénticos.
b. Diferencias de variabilidad, naturaleza de los espacios y existencia de grupos para las tres muestras.
c. No. Se pierde el detalle.

77. **a.**

0	2355566777888	
1	0000135555	
2	00257	
3	0033	
4	0057	tallo: unidades
5	044	hoja: decenas
6		
7	05	
8	8	
9	0	
10	3	
HI	22.0, 24.5	

b.

Clase	Frec.	Frec. rel.	Densidad
0–<2	23	0.500	0.250
2–<4	9	0.196	0.098
4–<6	7	0.152	0.076
6–<10	4	0.087	0.022
10–<20	1	0.022	0.002
20–<30	2	0.043	0.004

79. **a.** $\bar{x}_{n+1} = (n\bar{x}_n + x_{n+1})/(n + 1)$
c. 12.53, 0.532

81. Sesgo positivo considerable (suponiendo unimodalidad).

83. **a.** Todos los puntos están en una recta de 45°. Los puntos caen abajo de una recta de 45°. **b.** Los puntos están muy abajo de una recta de 45°, lo cual indica sesgo positivo considerable.

Capítulo 2

1. **a.** $\mathcal{S} = \{1324, 3124, 1342, 3142, 1423, 1432, 4123, 4132, 2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$
b. $A = \{1324, 1342, 1423, 1432\}$
c. $B = \{2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$
d. $A \cup B = \{1324, 1342, 1423, 1432, 2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$,
$A \cap B$ no contienen resultados (A y B son disjuntos),
$A' = \{3124, 3142, 4123, 4132, 2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$

3. **a.** $A = \{SSF, SFS, FSS\}$
b. $B = \{SSF, SFS, FSS, SSS\}$
c. $C = \{SFS, SSF, SSS\}$
d. $C' = \{FFF, FSF, FFS, FSS, SFF\}$,
$A \cup C = \{SSF, SFS, FSS, SSS\}$,
$A \cap C = \{SSF, SFS\}$,
$B \cup C = \{SSF, SFS, FSS, SSS\} = B$,
$B \cap C = \{SSF, SFS, SSS\} = C$

5. **a.** $\mathcal{S} = \{(1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 2, 1), (1, 2, 2), (1, 2, 3), (1, 3, 1), (1, 3, 2), (1, 3, 3), (2, 1, 1), (2, 1, 2), (2, 1, 3), (2, 2, 1), (2, 2, 2), (2, 2, 3), (2, 3, 1), (2, 3, 2), (2, 3, 3), (3, 1, 1), (3, 1, 2), (3, 1, 3), (3, 2, 1), (3, 2, 2), (3, 2, 3), (3, 3, 1), (3, 3, 2), (3, 3, 3)\}$ **b.** $\{(1, 1, 1), (2, 2, 2), (3, 3, 3)\}$ **c.** $\{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$ **d.** $\{(1, 1, 1), (1, 1, 3), (1, 3, 1), (1, 3, 3), (3, 1, 1), (3, 1, 3), (3, 3, 1), (3, 3, 3)\}$

7. **a.** Hay 35 resultados en $\mathcal{S}$. **b.** $\{AABABAB, AABAABB, AAABBAB, AAABABB, AAAABBB\}$

11. **a.** 0.07 **b.** 0.30 **c.** 0.57

13. **a.** 0.36 **b.** 0.64 **c.** 0.53
d. 0.47 **e.** 0.17 **f.** 0.75

15. **a.** 0.572 **b.** 0.879

17. **a.** Hay paquetes se software para estadística distintos a SPSS y SAS.
b. 0.70 **c.** 0.80 **d.** 0.20

19. **a.** 0.8841 **b.** 0.0435

21. **a.** 0.10 **b.** 0.18, 0.19 **c.** 0.41 **d.** 0.59 **e.** 0.31 **f.** 0.69

23. **a.** 0.067 **b.** 0.400 **c.** 0.933 **d.** 0.533

25. **a.** 0.98 **b.** 0.02 **c.** 0.03 **d.** 0.24

27. **a.** 0.1 **b.** 0.7 **c.** 0.6

29. **a.** 20 **b.** 60 **c.** 10

31. **a.** 243 **b.** 3645 días (casi 10 años)

33. **a.** 53,130 **b.** 1190 **c.** 0.022 **d.** 0.023

35. 0.20

37. 0.0456

39. **a.** 0.0839 **b.** 0.24975

41. **a.** 0.929 **b.** 0.0714 **c.** 0.99997520

43. 0.000394, 0.00394, 0.00001539

45. **a.** 0.447, 0.500, 0.200 **b.** 0.400, 0.447 **c.** 0.211

47. **a.** 0.50 **b.** 0.50 **c.** 0.625 **d.** 0.375 **e.** 0.769

49. 0.217, 0.178

51. 0.436, 0.581

53. 0.083

55. **a.** 0.0111 **b.** 0.333 **c.** 0.2

59. **a.** 0.21 **b.** 0.455 **c.** 0.264, 0.274

61. **a.** 0.578, 0.278, 0.144 **b.** 0, 0.457, 0.543

63. **b.** 0.54 **c.** 0.68 **d.** 0.74 **e.** 0.7941

65. $P(\text{Mean} \mid S) = 0.3922$, $P(\text{Median} \mid S) = 0.2941$, así que Mean y Median son el más y el menos probable, respectivamente.

67. **a.** 0.126 **b.** 0.05 **c.** 0.1125 **d.** 0.2725 **e.** 0.5325 **f.** 0.2113

69. **a.** 0.300 **b.** 0.820 **c.** 0.146

73. 0.401, 0.722

75. **a.** 0.06235 **b.** 0.00421

77. 0.0059

79. **a.** 0.95

81. **a.** 0.10, 0.20 **b.** 0

83. **a.** $p(2-p)$ **b.** $1-(1-p)^n$ **c.** $(1-p)^3$ **d.** $0.9 + (1-p)^3(0.1)$ **e.** $0.1(1-p)^3/[0.9 + 0.1(1-p)^3] = 0.0137$ para $p = 0.5$

85. 0.8588, 0.9897

87. $[2\pi(1-\pi)]/(1-\pi^2)$

89. **a.** 0.333, 0.444 **b.** 0.150 **c.** 0.291

91. 0.45, 0.32

93. **a.** 0.0083 **b.** 0.2 **c.** 0.2

95. 0.905

97. **a.** 0.956 **b.** 0.994

99. 0.926

101. **a.** 0.018 **b.** 0.601

103. **a.** 0.883, 0.117 **b.** 23 **c.** 0.156

105. $1-(1-p_1)(1-p_2)\cdot\cdots\cdot(1-p_n)$

107. **a.** 0.0417 **b.** 0.375

109. $P(\text{contratar \#1}) = 6/24$ para $s = 0$, $= 11/24$ para $s = 1$, $= 10/24$ para $s = 2$ e $= 6/24$ para $s = 3$, por lo que $s = 1$ es mejor.

111. $1/4 = P(A_1 \cap A_2 \cap A_3) \neq P(A_1)\cdot P(A_2)\cdot P(A_3) = 1/8$

Capítulo 3

1. $x = 0$ para *FFF*; $x = 1$ para *SFF*, *FSF*, y *FFS*; $x = 2$ para *SSF*, *SFS* y *FSS*; y $x = 3$ para *SSS*

3. Z = promedio de los dos números, con valores posibles 2/2, 3/2, . . . , 12/2; W = valor absoluto de la diferencia, con valores posibles 0, 1, 2, 3, 4 y 5

5. No. En el ejemplo 3.4 sea $Y = 1$ si se examinan a lo sumo 3 baterías, y sea $Y = 0$ en caso contrario. Entonces Y sólo tiene dos valores.

7. **a.** {0, 1, . . . , 12}; discreta **c.** {1, 2, 3, . . . }; discreta **e.** {0, c, $2c$, . . . , 10 000c}, donde c son las regalías por libro; discreta **g.** $\{x: m \le x \le M\}$ donde m (M) es la tensión mínima (o máxima) posible; continua

9. **a.** {2, 4, 6, 8, . . . }, es decir, {2(1), 2(2), 2(3), 2(4), . . . }, una sucesión infinita; discreta **b.** {2, 3, 4, 5, 6, . . . }, es decir, {1 + 1, 1 + 2, 1 + 3, 1 + 4, . . . }, una sucesión infinita; discreta

11. **a.** $p(4) = 0.45$, $p(6) = 0.40$, $p(8) = 0.15$, $p(x) = 0$ para $x \neq 4, 6$ u 8 **c.** 0.55, 0.15

13. **a.** 0.70 **b.** 0.45 **c.** 0.55 **d.** 0.71 **e.** 0.65 **f.** 0.45

15. **a.** (1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5) **b.** $p(0) = 0.3$, $p(1) = 0.6$, $p(2) = 0.1$ **c.** $F(x) = 0$ para $x < 0$, $= 0.3$ para $0 \le x < 1$, $= 0.9$ para $1 \le x < 2$ e $= 1$ para $2 \le x$

17. **a.** 0.81 **b.** 0.162 **c.** Es A: *AUUUA, UAUUA, UUAUA, UUUAA;* 0.00405

19. $p(1) = p(2) = p(3) = p(4) = 0.25$

21. $F(x) = 0$ para $x < 0$, $= 0.10$ para $0 \le x < 1$, $= 0.25$ para $1 \le x < 2$, $= 0.45$ para $2 \le x < 3$, $= 0.70$ para $3 \le x < 4$, $= 0.90$ para $4 \le x < 5$, $= 0.96$ para $5 \le x < 6$ e $= 1$ para $6 \le x$

23. **a.** $p(1) = 0.30$, $p(3) = 0.10$, $p(4) = 0.05$, $p(6) = 0.15$, $p(12) = 0.40$ **b.** 0.30, 0.60

25. **a.** $p(x) = (\frac{2}{3})^{x-1} \cdot (\frac{1}{3})$ para $x = 1, 2, 3, \ldots$

b. $p(y) = (\frac{2}{3})^{y-2} \cdot (\frac{1}{3})$ para $y = 2, 3, 4, 5, \ldots$

c. $p(z) = (\frac{25}{54}) \cdot (\frac{2}{3})^{2z-2}$ para $z = 1, 2, 3, \ldots$

29. **a.** 0.60 **b.** \$110

31. **a.** 16.38, 272.298, 3.9936 **b.** 401 **c.** 2496 **d.** 13.66

33. Sí, porque $\sum_{x=1}^{\infty} 1/x^2 < \infty$.

35. \$700

37. $E(h(X)) = 0.408$, así que hay que jugar.

39. $V(-X) = V(X)$

41. **a.** 32.5 **b.** 7.5

c. $V(X) = E[X(X-1)] + E(X) - [E(X)]^2$

43. **a.** 0.25, 0.11, 0.06, 0.04, 0.01 **b.** 2.64, 1.54, 0.04, conservador **c.** Son iguales. **d.** $p(-1) = 0.02 = p(1)$, $p(0) = 0.96$

45. **a.** 0.850 **b.** 0.200 **c.** 0.200 **d.** 0.701 **e.** 0.851 **f.** 0.000 **g.** 0.570

47. **a.** 0.354 **b.** 0.115 **c.** 0.918

49. **a.** 0.403 **b.** 0.787 **c.** 0.774

51. 0.1478

53. 0.407, independencia

55. **a.** 0.017 **b.** 0.811, 0.425 **c.** 0.006, 0.902, 0.586

57. Cuando $p = 0.9$, la probabilidad es 0.99 para A y 0.9963 para B. Si $p = 0.5$, estas probabilidades son 0.75 y 0.6875, respectivamente.

59. La tabulación para $p > 0.5$ es innecesaria.

61. **a.** 20, 16 **b.** 70, 21

63. $P(|X - \mu| \ge 2\sigma) = 0.042$ cuando $p = 0.5$ y $= 0.065$ cuando $p = 0.75$, en comparación con el límite superior de 0.25. Con $k = 3$ en lugar de $k = 2$, estas probabilidades son 0.002 y 0.004, respectivamente, mientras que el límite superior es 0.11.

65. **a.** 0.114 **b.** 0.879 **c.** 0.121 **d.** Utilice la distribución binomial con $n = 15$, $p = 0.10$

67. **a.** $h(x; 15, 10, 20)$ para $x = 5, \ldots, 10$ **b.** 0.0325 **c.** 0.697

69. **a.** $h(x; 10, 10, 20)$ **b.** 0.033

71. **a.** $nb(x; 2, 0.5)$ **b.** 0.188 **c.** 0.688 **d.** 2, 4

73. $nb(x; 6, 0.5)$, 6

75. **a.** 0.932 **b.** 0.065 **c.** 0.068 **d.** 0.492 **e.** 0.251

77. **a.** 0.011 **b.** 0.441 **c.** 0.554, 0.459 **d.** 0.945

79. Poisson(5) **a.** 0.492 **b.** 0.133

81. **a.** 0.313, 0.809, 0.283 **b.** 12, 3.464 **c.** 0.530, 0.011

83. **a.** 0.099 **b.** 0.135 **c.** 2

85. **a.** 4 **b.** 0.215 **c.** Por lo menos $-\ln(0.1)/2 \approx 1.1513$ años

87. **a.** 0.221 **b.** 6 800 000 **c.** $p(x; 20.106)$

91. **b.** 3.114, 0.405, 0.636

93. **a.** $b(x; 15, 0.75)$ **b.** 0.686 **c.** 0.313 **d.** 11.25, 2.81 **e.** 0.310

95. 0.991

97. **a.** $p(x; 2.5)$ **b.** 0.067 **c.** 0.109

99. 1.813, 3.05

101. $p(2) = p^2$, $p(3) = (1-p)p^2$, $p(4) = (1-p)p^2$, $p(x) = [1 - p(2) - \cdots - p(x-3)](1-p)p^2$ para $x = 5, 6, 7, \ldots$; 0.99950841

103. **a.** 0.0029 **b.** 0.0767, 0.9702

105. **a.** 0.135 **b.** 0.00144 **c.** $\sum_{x=0}^{\infty}[p(x; 2)]^5$

107. 3.590

109. **a.** No **b.** 0.0273

111. **b.** $0.6p(x; \lambda) + 0.4p(x; \mu)$ **c.** $(\lambda + \mu)/2$ **d.** $(\lambda - \mu)^2/4 + (\lambda + \mu)/2$

113. $\sum_{i=1}^{10}(p_{i+j+1} + p_{i-j-1})p_i$, donde $p_k = 0$ si $k < 0$ o bien $k > 10$.

Capítulo 4

1. a. 0.25 b. 0.50 c. 0.4375

3. b. 0.5 c. 0.6875 d. 0.6328

5. a. 0.375 b. 0.125 c. 0.297 d. 0.578

7. a. $f(x) = 0.1$ para $25 \le x \le 35$ y 0 en caso contrario
b. 0.20 c. 0.40 d. 0.20

9. a. 0.562 b. 0.438, 0.438 c. 0.071

11. a. 0.25 b. 0.1875 c. 0.9375 d. 1.4142
e. $f(x) = x/2$ para $0 < x < 2$
f. 1.33 g. 0.222, 0.471 h. 2

13. a. 3 b. 0 para $x \le 1$, $1 - x^{-3}$ para $x > 1$
c. 0.125, 0.088 d. 1.5, 0.866 e. 0.924

15. a. $F(x) = 0$ para $x \le 0$, $= 90[\frac{x^5}{9} - \frac{x^{10}}{10}]$ para $0 < x < 1$, $=$ 1 para $x \ge 1$ b. 0.0107 c. 0.0107, 0.0107
d. 0.9036 e. 0.818, 0.111 f. 0.3137

17. a. Para $2 \le x \le 4$, $F(x) = 0.25[3x - 7 - (x - 3)^3]$
b. 3 c. 3, 0.2

19. a. 0.597 b. 0.369
c. $f(x) = 0.3466 - 0.25 \ln(x)$ para $0 < x < 4$

21. 314.79

23. 248, 3.60

25. b. 1.8(percentil 90 para X) + 32
c. $a(X$ percentil$) + b$

27. a. 2.14 b. 0.81 c. 1.17
d. 0.97 e. 2.41

29. a. 2.54 b. 1.34 c. −0.42

31. a. 0.9918 b. 0.0082 c. 0.9544

33. a. 0.3336 b. Aproximadamente 0
c. 0.5795 d. 6.524 e. 0.8028

35. a. 36.7 b. 22.225 c. 3.179

37. 0.002

39. 10, 0.2

41. 7.3%

43. 21.155

45. a. 0.1190, 0.6969 b. 0.0021 c. 0.7054
d. >5020 o <1844 (con $z_{0.0005} = 3.295$)
e. Normal, $\mu = 7.576$, $\sigma = 1.064$, 0.7054

47. 0.3174 para $k = 1$, 0.0456 para $k = 2$, 0.0026 para $k = 3$, en comparación con los límites de 1, 0.25 y 0.111, respectivamente.

49. a. Exacto: 0.212, 0.577, 0.573; Aproximado: 0.211, 0.567, 0.596 b. Exacto: 0.885, 0.575, 0.017; Aproximado: 0.885, 0.579, 0.012 c. Exacto: 0.002, 0.029, 0.617; Aproximado: 0.003, 0.033, 0.599

51. a. 0.9666 b. 0.0099, 0

53. b. Normal, $\mu = 239$, $\sigma^2 = 12.96$

55. a. 120 b. 1.329 c. 0.371
d. 0.735 e. 0

57. a. 5, 4 b. 0.715 c. 0.411

59. a. 1 b. 1 c. 0.982 d. 0.129

61. a. 0.449, 0.699, 0.148 b. 0.05, 0.018

63. a. $\cap A_i$ b. Exponencial con $\lambda = 0.05$
c. Exponencial con parámetro $n\lambda$

67. a. 0.826, 0.826, 0.0636 b. 0.664 c. 172.727

71. a. 0.930 b. 0.298 c. 98.18

73. a. 68.0, 122.1 b. 0.3204
c. 0.7257, sesgo

75. a. 149.157, 223.595 b. 0.9573 c. 0.0414
d. 148.41 e. 9.57 f. 125.90

77. $\alpha = \beta$

79. b. $[\Gamma(\alpha + \beta) \cdot \Gamma(m + \beta)]/[\Gamma(\alpha + \beta + m) \cdot \Gamma(\beta)]$, $\beta/(\alpha + \beta)$

81. Sí, puesto que el patrón de la gráfica es bastante lineal.

83. Sí

85. Sí

87. Grafique $\ln(x)$ en función del percentil z. El patrón es recto, así que es factible una distribución poblacional lognormal.

89. El patrón es bastante lineal; es muy posible que la resistencia esté distribuida normalmente.

91. Hay curvatura considerable en la gráfica. λ es un parámetro de escala (al igual que σ para la familia normal).

93. a. $F(y) = \frac{1}{48}(y^2 - y^3/18)$ para $0 \le y \le 12$
b. 0.259, 0.5, 0.241 c. 6, 43.2, 7.2
d. 0.518 e. 3.75

95. a. $f(x) = x^2$ para $0 \le x < 1$ y $= \frac{7}{4} - \frac{3}{4}x$ para $1 \le x \le \frac{7}{3}$
b. 0.917 c. 1.213

97. a. 0.9162 b. 0.9549 c. 1.3374

99. a. 0.3859 b. 0.0663 c. (72.97, 119.03)

101. b. $F(x) = 0$ para $x < -1$, $= (4x - x^3/3)/9 + \frac{11}{27}$ para $-1 \le x \le 2$ y $= 1$ para $x > 2$

c. No. $F(0) < 0.5 \Rightarrow \tilde{\mu} > 0$
d. $Y \sim \text{Bin}(10, \frac{3}{11})$

103. a. 0.368, 0.828, 0.460 b. 352.53
c. $1/\beta \cdot \exp[-\exp(-(x-\alpha)/\beta)] \cdot \exp(-(x-\alpha)/\beta)$
d. α e. $\mu = 201.95$, moda = 150, $\tilde{\mu} = 182.99$

105. a. μ b. No c. 0
d. $(\alpha - 1)\beta$ e. $\nu - 2$

107. b. $p(1 - \exp(-\lambda_1 x)) + (1-p)(1 - \exp(-\lambda_2 x))$ para $x \geq 0$ c. $p/\lambda_1 + (1-p)/\lambda_2$
d. $V(X) = 2p/\lambda_1^2 + 2(1-p)/\lambda_2^2 - \mu^2$
e. 1, $CV > 1$ f. $CV < 1$

109. a. Lognormal b. 1 c. 2.72, 0.0185

113. a. Exponencial con $\lambda = 1$
c. Gamma con parámetros α y $c\beta$

115. a. $(1/365)^3$ b. $(1/365)^2$ c. 0.000002145

117. b. Sea $u_1, u_2, u_3, \ldots$ una secuencia de observaciones de una distribución Unif[0, 1] una sucesión de números aleatorios). Entonces, con $x_i = (-0.1)\ln(1 - u_i)$, las x_i son observaciones de una distribución exponencial con $\lambda = 10$.

119. $g(E(X)) \leq E(g(X))$

Capítulo 5

1. a. 0.20 b. 0.42 c. Al menos una manguera está en uso en cada bomba; 0.70. d. $p_X(x) = 0.16, 0.34, 0.50$ para $x = 0, 1, 2$, respectivamente; $p_Y(y) = 0.24, 0.38, 0.38$ para $y = 0, 1, 2$, respectivamente; 0.50
e. No; $p(0, 0) \neq p_X(0) \cdot p_Y(0)$

3. a. 0.15 b. 0.40 c. 0.22 d. 0.17, 0.46

5. a. 0.054 b. 0.00018

7. a. 0.030 b. 0.120 c. 0.300
d. 0.380 e. Sí

9. a. 3/380 000 b. 0.3024 c. 0.3593
d. $10Kx^2 + 0.05$ para $20 \leq x \leq 30$ e. No

11. a. $e^{-\lambda-\mu} \cdot \lambda^x \cdot \mu^y/x!y!$ b. $e^{-\lambda-\mu} \cdot [1 + \lambda + \mu]$
c. $e^{-(\lambda+\mu)} \cdot (\lambda + \mu)^m/m!$; Poisson $(\lambda + \mu)$

13. a. e^{-x-y} para $x \geq 0, y \geq 0$ b. 0.400 c. 0.594
d. 0.330

15. a. $F(y) = 1 - e^{-\lambda y} + (1 - e^{-\lambda y})^2 - (1 - e^{-\lambda y})^3$ para $y \geq 0$ b. $2/3\lambda$

17. a. 0.25 b. 0.318 c. 0.637
d. $f_X(x) = 2\sqrt{R^2 - x^2}/\pi R^2$ para $-R \leq x \leq R$; no

19. a. $K(x^2 + y^2)/(10Kx^2 + 0.05)$; $K(x^2 + y^2)/(10Ky^2 + 0.05)$
b. 0.556, 0.549 c. 25.37, 2.87

21. $= f_Y(y)$ para toda x

23. 0.15

25. L^2

27. $\frac{1}{3}$ hr

29. $-\frac{1}{3}$

31. a. −0.1082 b. −0.0131

37. a.

$\bar{x}$	25	32.5	40	45	52.5	65
$p(\bar{x})$	0.04	0.20	0.25	0.12	0.30	0.09

$E(\bar{X}) = \mu = 44.5$

b.

s^2	0	112.5	312.5	800
$p(s^2)$	0.38	0.20	0.30	0.12

$E(S^2) = 212.25 = \sigma^2$

39.

Proporción	0	0.1	0.2	0.3	0.4	0.5
Probabilidad	0.000	0.000	0.000	0.001	0.005	0.027
Proporción	0.6	0.7	0.8	0.9	1.0	
Probabilidad	0.088	0.201	0.302	0.269	0.107	

41. a.

$\bar{x}$	1	1.5	2	2.5	3	3.5	4
$p(\bar{x})$	0.16	0.24	0.25	0.20	0.10	0.04	0.01

b. 0.85 c.

r	0	1	2	3
$p(r)$	0.30	0.40	0.22	0.08

47. a. 0.6826 b. 0.1056

49. a. 0.6026 b. 0.2981

51. 0.7720

53. a. 0.0062 b. 0

55. a. 0.0968 b. 0.8882

57. 0.9616

59. a. 0.9986, 0.9986 b. 0.9015, 0.3970
c. 0.8357 d. 0.9525, 0.0003

61. a. 3.5, 2.27, 1.51 b. 15.4, 75.94, 8.71

63. a. 0.695 b. 4.0675 > 2.6775

65. a. 0.9232 b. 0.9660

67. 0.1588

69. a. 2400 b. 1205; independencia
c. 2400, 41.77

71. a. 158, 430.25 b. 0.9788

73. a. Aproximadamente normal con media = 105, DE = 1.2649; aproximadamente normal con media = 100, DE = 1.0142
b. Aproximadamente normal con media = 5, DE = 1.6213
c. 0.0068 d. 0.0010, sí

75. a. 0.2, 0.5, 0.3 para $x = 12, 15, 20$; 0.10, 0.35, 0.55 para $y = 12, 15, 20$ b. 0.25 c. No d. 33.35
e. 3.85

77. a. 3/81 250 b. $f_X(x) = k(250x - 10x^2)$ para $0 \leq x \leq 20$ y $= k(450x - 30x^2 + \frac{1}{2}x^3)$ para $20 < x \leq 30$; $f_Y(y)$ resulta de sustituir y por x en $f_X(x)$. No son independientes.
c. 0.355 d. 25.969
e. 204.6154, −0.894 f. 7.66

79. ≈ 1

81. a. 400 mín b. 70

83. 97

85. 0.9973

89. *b, c.* Chi-cuadrado con $v = n$.

91. a. $\sigma_W^2/(\sigma_W^2 + \sigma_E^2)$ b. 0.9999

93. 26, 1.64

Capítulo 6

1. a. 8.14, $\overline{X}$ b. 0.77, $\tilde{X}$ c. 1.66, S
d. 0.148 e. 0.204, $S/\overline{X}$

3. a. 1.348, $\overline{X}$ b. 1.348, $\overline{X}$
c. 1.781, $\overline{X} + 1.28S$
d. 0.6736 e. 0.0905

5. 1 703 000; 1 591 300; 1 601 438.281

7. a. 120.6 b. 1 206 000 c. 0.80 d. 120.0

9. a. 2.11 b. 0.119

11. b. $\left[\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}\right]^{1/2}$ c. Utilice $\hat{p}_i = x/n_i$ y $\hat{q}_i = 1 - \hat{p}_i$ en lugar de p_i y q_i en el inciso (b) para $i = 1, 2$.
d. −0.245 e. 0.041

15. a. $\hat{\theta} = \sum X_i^2/2n$ b. 74.505

17. b. 0.444

19. a. $\hat{p} = 2\hat{\lambda} - 0.30 = 0.20$ b. $\hat{p} = (100\hat{\lambda} - 9)/70$

21. b. $\hat{\alpha} = 5$, $\hat{\beta} = 28.0/\Gamma(1.2)$

23. $\hat{\lambda}_1 = \bar{x}$, $\hat{\lambda}_2 = \bar{y}$, la estimación de $(\lambda_1 - \lambda_2)$ es $\bar{x} - \bar{y}$.

25. a. 384.4, 18.86 b. 415.42

29. a. $\hat{\theta} = \text{mín}(X_i)$, $\hat{\lambda} = n/\sum[X_i - \text{mín}(X_i)]$
b. 0.64, 0.202

33. Con x_i = tiempo entre el nacimiento $i - 1$ y el nacimiento i, $\hat{\lambda} = 6/\sum_{i=1}^{6} ix_i = 0.0436$.

35. 29.5

37. 1.0132

Capítulo 7

1. a. 99.5% b. 85% c. 2.96 d. 1.15

3. a. Más reducido b. No c. No d. No

5. a. (4.52, 5.18) b. (4.12, 5.00)
c. 0.55 d. 94

7. Por un factor de 4; la amplitud se reduce en un factor de 5.

9. a. $(\bar{x} - 1.645\sigma/\sqrt{n}, \infty)$; (4.57, ∞)
b. $(\bar{x} - z_\alpha \cdot \sigma/\sqrt{n}, \infty)$ c. $(-\infty, \bar{x} + z_\alpha \cdot \sigma/\sqrt{n})$; $(-\infty, 59.7)$

11. 950, 0.8714

13. a. (0.990, 1.066) b. 158

15. a. 80% b. 98% c. 75%

17. 134.53

19. (0.513, 0.615)

21. 0.218

23. a. (0.438, 0.814) b. 659

25. a. 381 b. 339

29. a. 2.228 b. 2.086 c. 2.845 d. 2.680
e. 2.485 f. 2.571

31. **a.** 1.812 **b.** 1.753 **c.** 2.602 **d.** 3.747
e. 2.1716 (de MINITAB) **f.** Alrededor de 2.43

33. **a.** Grado razonable de simetría; sin valores atípicos
b. Sí (se basa en una gráfica de probabilidad normal)
c. (430.5, 446.1), sí, no

35. **a.** 95% CI: (2783.3, 2991.9)
b. 95% PI: (2632.1, 3143.1), más o menos 2.5 veces el doble de extensión

37. **a.** (0.888, 0.964) **b.** (0.752, 1.100)
c. (0.634, 1.218)

39. **a.** Sí **b.** (6.45, 98.01)
c. (18.63, 85.83)

41. Todos 70%; (c), porque es el más corto

43. **a.** 18.307 **b.** 3.940 **c.** 0.95 **d.** 0.10

45. (3.6, 8.1); no

47. **a.** 95% CI: (6.702, 9.456) **b.** (0.166, 0.410)

49. **a.** Al parecer hay un poco de sesgo positivo en la mitad intermedia de la muestra, pero el extremo inferior es mucho más largo que el extremo superior. El grado de variabilidad es bastante considerable, aunque no hay valores atípicos.
b. Sí. El patrón de puntos en una gráfica de probabilidad normal es razonablemente lineal.
c. (33.53, 43.79)

51. **a.** (0.624, 0.732) **b.** 1080 **c.** No

53. $(-0.84, -0.16)$

55. 246

57. $(2t_r/\chi^2_{1-\alpha/2,2r}, 2t_r/\chi^2_{\alpha/2,2r}) = (65.3, 232.5)$

59. **a.** $(\text{máx}(x_i)/(1-\alpha/2)^{1/n}, \text{máx}(x_i)/(\alpha/2)^{1/n})$
b. $(\text{máx}(x_i), \text{máx}(x_i)/\alpha^{1/n})$ **c.** (b); (4.2, 7.65)

61. (73.6, 78.8) contra (75.1, 79.6)

Capítulo 8

1. **a.** Sí **b.** No **c.** No
d. Sí **e.** No **f.** Sí

5. H_0: $\sigma = 0.05$ contra H_a: $\sigma < 0.05$. I: se concluye que la variabilidad en el espesor es satisfactoria cuando no lo es. II: se concluye que la variabilidad en espesor no es satisfactoria cuando sí lo es.

7. I: concluir que la planta no cumple con las normas cuando en realidad sí lo hace. II: concluir que la planta cumple con las normas cuando en realidad no es así.

9. **a.** R_1 **b.** I: juzgar que se favorece a una de las dos empresas cuando no es el caso. II: considerar que no se favorece a ninguna empresa cuando en realidad se prefiere a una de las dos. **c.** 0.044
d. $\beta(0.3) = \beta(0.7) = 0.488$, $\beta(0.4) = \beta(0.6) = 0.845$
e. Rechazar H_0 en favor de H_a.

11. **a.** H_0: $\mu = 10$ contra H_a: $\mu \neq 10$ **b.** 0.01
c. 0.5319, 0.0078 **d.** 2.58
e. Se sustituye 10.1032 por 10.124 y 9.8968 por 9.876.
f. $\bar{x} = 10.020$, por lo que no se debe rechazar H_0.
g. $z \geq 2.58$ o ≤ -2.58

13. **b.** 0.0004, 0, menor que 0.01

15. **a.** 0.0301 **b.** 0.003 **c.** 0.004

17. **a.** $z = 2.56 \geq 2.33$, así que se rechaza H_0. **b.** 0.8413
c. 143 **d.** 0.0052

19. **a.** $z = -2.27$, así que no rechace H_0. **b.** 0.2266
c. 22

21. **a.** $t_{0.005,12} = 2.179 > 1.6$, por lo tanto no rechace H_0: $\mu = 0.5$.
b. $-1.6 > -2.179$, así que no se rechaza H_0.
c. No se rechaza H_0.
d. Se rechaza H_0 en favor de H_a: $\mu \neq 0.5$.

23. $t = 2.24 \geq 1.708$, así que se debe rechazar H_0. Los datos indican una contradicción de la hipótesis inicial.

25. **a.** $z = -3.33 \leq -2.58$, así que se rechaza H_0.
b. 0.1056 **c.** 217

27. $-2.09 > -2.33$, por lo que no se rechaza H_0.

29. **a.** $0.498 < 1.895$, por lo tanto no se rechaza H_0.
b. 0.72

31. $-1.24 > -1.397$, así que al parecer se contradice la opinión previa.

35. Sí, porque $-2.47 \leq -1.96$.

37. $z = 3.67 \geq 2.58$, por lo tanto se rechaza H_0: $p = 0.40$. No.

39. -6.1 es mucho menor que -2.33, de modo que los datos respaldan con firmeza esa conclusión.

41. **a.** $\{15, \ldots, 20\}$ **b.** 0.021, sí **c.** 0.874, 0.196
d. Con $R = \{14, \ldots, 20\}$, no.

43. $z = -0.61 > -2.33$; no

45. **a.** No rechace H_0. **b.** No rechace H_0.
c. No rechace H_0. **d.** Rechace H_0.
e. No rechace H_0. **f.** No rechace H_0.

47. **a.** 0.0358 **b.** 0.0802 **c.** 0.5824
d. 0.1586 **e.** 0

49. **a.** Valor $P = 0.003 < 0.05$, así que se rechaza H_0.
b. Valor $P = 0.055 > 0.01$, por lo tanto no se rechaza H_0.
c. Valor $P > 0.5$, así que no se puede rechazar H_0 a algún nivel α razonable.

51. Valor $P < 0.0004 < 0.01$, así que se debe rechazar H_0: $\mu = 5$ en favor de H_a: $\mu \neq 5$.

53. No; Valor $P \approx 0.2$

55. $t \approx 1.9$, así que valor $P \approx 0.041$. Como el valor $P \leq \alpha$, se debe rechazar H_0: $\mu = 25$ en favor de H_a: $\mu > 25$.

57. $t \approx 1.9$, por lo tanto el valor $P \approx 0.116$. En consecuencia, no se debe rechazar H_0.

59. **a.** 0.8980, 0.1049, 0.0014 **b.** Valor $P \approx 0$. Sí. **c.** No

61. $z = -3.12 \leq -1.96$, de manera que se debe rechazar H_0.

63. **a.** H_0: $\mu = 0.85$ contra H_a: $\mu \neq 0.85$
b. No se puede rechazar H_0 por ninguna α.

65. **a.** Sí, porque $t = 12.9 \geq 2.228$.
b. Distribución normal de la población

67. **a.** No; no
b. No, porque $z = 0.44$ y el valor $P = 0.33 > 0.10$.

69. **a.** Aproximadamente 0.6; más o menos 0.2 (de la tabla A. 17 del apéndice) **b.** $n = 28$

71. **a.** $z = 1.64 < 1.96$, por lo tanto no se puede rechazar H_0; tipo II **b.** 0.10. Sí.

73. Sí. $z = -3.32 \leq -3.08$, así que se debe rechazar H_0.

75. No, puesto que $z = 1.33 < 2.05$.

77. No, porque $12.11 < 21.665$.

79. **a.** $\overline{X} + 2.33S$ **b.** $3.7145\sigma^2/n$; $1.927s/\sqrt{n}$
c. $z = -1.22$, por consiguiente no se puede rechazar H_0: $\mu + 2.33\sigma = 6.75$ en favor de H_a: $\mu + 2.33\sigma < 6.75$.

81. **b.** $\Phi(z_\gamma + (\mu_0 - \mu)/(\sigma/\sqrt{n})) - \Phi(-z_{\alpha-\gamma} + (\mu_0 - \mu)/(\sigma/\sqrt{n}))$ **c.** $\gamma > \alpha/2$

Capítulo 9

1. **a.** −0.4 h; no depende **b.** 0.0724, 0.2691 **c.** No

3. $z = 1.76 < 2.33$, así que no se rechaza H_0.

5. **a.** $z = -2.90$, por lo tanto se rechaza H_0. **b.** 0.0019
c. 0.8212 **d.** 66

7. Sí, puesto que $z = 1.83 \geq 1.645$.

9. **a.** 6.2; sí **b.** $z = 1.14$, valor $P \approx 0.25$, no
c. No **d.** A 95% IC es (10.0, 21.8).

11. Un IC de 95% es (0.99, 2.41).

13. 50

15. b. Aumenta.

17. **a.** 17 **b.** 21 **c.** 18 **d.** 26

19. $t = -1.20 > -t_{0.01,9} = -2.821$, así que no se rechaza H_0.

21. Sí, $-2.64 \leq -2.602$, de modo que se rechaza H_0.

23. b. No **c.** $t = -0.38 > -t_{\alpha/2,10}$ para algún valor razonable de α, así que no se rechaza H_0 (valor $P \approx 0.7$).

25. (0.3, 6.1), sí, sí

27. (6.5, 31.3) basado en 9 gl; sí, sí

29. $t = -2.10$, gl = 25, valor $P = 0.023$. Al nivel de significación de 0.05 se concluiría que la bebida de cola tiene mayor fuerza promedio, pero no al nivel de significación 0.01.

31. **a.** Centros casi idénticos, bastante más variabilidad en observaciones intermedias que en observaciones mayores.
b. (−7.9, 9.6), basado en 23 gl; no

33. $t = -2.2$, valor $P = 0.022$, no se rechaza H_0.

35. $t = -2.2$, gl = 16, valor $P = 0.021 > 0.01 = \alpha$, así que no se rechaza H_0.

37. **a.** (−0.561, −0.287) **b.** Entre −1.224 y 0.376

39. **a.** Sí
b. $t = 2.7$, valor $P = 0.018 < 0.05 = \alpha$, así que se debe rechazar H_0.

41. $t = 1.9$, valor $P = 0.047$. No se puede rechazar H_0 al nivel de significación 0.01, pero apenas se rechaza con $\alpha = 0.05$.

43. **a.** No **b.** −49.1 **c.** 49.1

45. (−10.499, −4.589)

47. Se rechaza H_0 porque $-4.18 \leq -2.33$

49. Valor $P = 0.4247$, así que no se puede rechazar H_0.

51. **a.** $z = 1.48 < 1.645$, de modo que no debe rechazarse H_0. **b.** $n = 6582$

53. **a.** El intervalo de confianza para $\ln(\theta)$ es $\ln(\hat{\theta}) \pm z_{\alpha/2}[(m-x)/(mx) + (n-y)/(ny)]^{1/2}$. Al calcular los antilogaritmos de los límites inferior y superior se obtiene un intervalo de confianza para θ.
b. (1.43, 2.31); la aspirina parece ser benéfica.

55. $(-0.35, 0.07)$

57. **a.** 3.69 **b.** 4.82 **c.** 0.207 **d.** 0.271
e. 4.30 **f.** 0.212 **g.** 0.95 **h.** 0.94

59. $f = 0.384$; como $0.167 < 0.384 < 3.63$, no se rechaza H_0.

61. $f = 2.85 \geq 2.08$, así que se rechaza H_0; parece que hay más variabilidad en la ganancia de peso con dosis baja.

63. $(s_2^2 F_{1-\alpha/2}/s_1^2,\ s_2^2 F_{\alpha/2}/s_1^2)$; (0.023, 1.99)

65. No. $t = 3.2$, gl = 15, valor $P = 0.006$, por lo tanto se rechaza H_0: $\mu_1 - \mu_2 = 0$ ya sea con $\alpha = 0.05$ o 0.01.

67. $z > 0 \Rightarrow$ valor $P > 0.5$, así que no se puede rechazar H_0: $p_1 - p_2 = 0$.

69. $(-299.3, 1517.9)$

71. (1024.0, 1336.0), sí

73. Sí. $t = -2.25$, gl = 57, valor $P \approx 0.028$

75. **a.** No. $t = -2.84$, gl = 18, valor $P \approx 0.012$
b. No. $t = -0.56$, valor $P \approx 0.29$

77. No al nivel de significación 0.05. $t = -1.76 > -t_{0.05,4} = -2.015$

79. No, ni se debe usar la prueba t de dos muestras, porque una gráfica de probabilidad normal indica que no es normal la distribución de buena visibilidad.

81. No agrupada: gl = 15, $t = -1.8$, valor $P \approx 0.092$
Agrupada: gl = 24, $t = -1.9$, valor $P \approx 0.070$

83. **a.** $m = 141$, $n = 47$ **b.** $m = 240$, $n = 160$

85. **a.** No; $-t_{0.0005,11} = -4.437 < 3.3 < 4.437$, así que no se rechaza H_0. **b.** Valor $P = 0.008$, por lo tanto se rechazaría H_0 si $\alpha = 0.05$ o 0.01.

87. 0.9015, 0.8264, 0.0294, 0.0000; coeficientes de inteligencia promedio reales; no

89. Sí; $z = 4.2$, valor $P \approx 0$

91. **a.** Sí. $t = -6.4$, gl = 57, y valor $P \approx 0$
b. $t = 1.1$, valor $P = 0.14$, así que no se rechaza H_0.

93. $(-1.29, -0.59)$

Capítulo 10

1. **a.** $f = 1.85 < 3.06 = F_{0.05,4,15}$, así que no se rechaza H_0.
b. Valor $P > 0.10$

3. $f = 1.30 < 2.57 = F_{0.10,2,21}$, por lo que el valor $P > 0.10$. No se puede rechazar H_0 en ningún nivel de significación razonable.

5. $f = 1.73 < 5.49 = F_{0.01,2,27}$, así que al parecer no difieren los tres grados.

7. $f = 1.70 < 2.46 = F_{0.10,3,16}$, así que el valor $P > 0.10$. No se puede rechazar H_0 en ningún nivel de significación razonable.

9. $f = 3.96$ y $F_{0.05,3,20} = 3.10 < 3.96 < 4.94 = F_{0.01,3,20}$, de manera que $0.01 <$ valor $P < 0.05$. Por consiguiente, se puede rechazar H_0 al nivel de significación 0.05; al parecer hay diferencias entre los granos.

11. $w = 36.09$

3	1	4	2	5
437.5	462.0	469.3	512.8	532.1

Las marcas 2 y 5 al parecer no difieren, y tampoco parece haber diferencia alguna entre las marcas 1, 3 y 4, pero cada marca del primer grupo al parecer difiere bastante de las marcas del segundo grupo.

13.

3	1	4	2	5
427.5	462.0	469.3	502.8	532.1

15. $w = 5.94$

2	1	3	4
24.69	26.08	29.95	33.84

Las únicas diferencias significativas son entre 4 y 1 y entre 4 y 2.

17. $(-0.029, 0.379)$

19. Funciona cualquier valor de SSE entre 422.16 y 431.88.

21. **a.** $f = 22.6$ y $F_{0.01,5,78} \approx 3.3$, así que se rechaza H_0.
b. $(-99.16, -35.64)$, $(29.34, 94.16)$

23.

	1	2	3	4
1	–	2.88 ± 5.81	7.43 ± 5.81	12.78 ± 5.48
2	–	–	4.55 ± 6.13	9.90 ± 5.81
3	–	–	–	5.35 ± 5.81
4	–	–	–	–

4	3	2	1

25. **a.** Normal, varianzas iguales
b. SSTr = 8.33, SSE = 77.79, $f = 1.7$, no se debe rechazar H_0 (valor $P > 0.10$)

27. **a.** $f = 3.75 \geq 3.10 = F_{0.05,3,20}$, así que al parecer difieren las marcas. **b.** Es bastante factible la normalidad (una gráfica de probabilidad normal de los residuos $x_{ij} - \bar{x}_{i\cdot}$ muestra un patrón lineal).
c.

4	3	2	1

Sólo parecen diferir en forma significativa las marcas 1 y 4.

31. Aproximadamente 0.62

33. $\operatorname{arcsen}(\sqrt{x/n})$

35. **a.** $3.68 < 4.94$, así que no se rechaza H_0.
b. $0.029 > 0.01$, así que tampoco se rechaza H_0.

37. $f = 8.44 > 6.49 = F_{0.001}$, de modo que el valor $P < 0.001$ y se debe rechazar H_0.
5 3 1 4 2 Este patrón de subrayado es un poco difícil de interpretar.

39. El intervalo de confianza es (−0.144, 0.474), el cual no incluye al 0.

41. $f = 3.96 < 4.07$, así que no se debe rechazar H_0: $\sigma_A^2 = 0$.

43. (−3.70, 1.04), (−4.83, −.33), (−3.77, 1.27), (−3.99, 0.15). Sólo $\mu_1 - \mu_3$ parece diferir significativamente de cero entre estos cuatro contrastes.

45. Son idénticos.

Capítulo 11

1. **a.** $f_A = 1.55$, así que no se rechaza H_{0A}.
b. $f_B = 2.98$, así que no se rechaza H_{0B}.

3. **a.** $f_A = 12.987 \geq F_{0.01,3,9}$, así que se concluye que hay un efecto del flujo de gas; $f_B = 105.31$, entonces hay un efecto del flujo de líquido. **b.** $w = 95.44$; 231.75 325.25 441.0 613.25, por lo que sólo los dos flujos menores no difieren entre sí en forma significativa.
c. 336.75 382.25 419.25 473 por lo que sólo los flujos mínimo y máximo parecen diferir entre sí en forma significativa.

5. $f_A = 2.56$, $F_{0.01,3,12} = 5.95$, así que al parecer no hay efecto debido al ángulo de tiro.

7. **a.**

Fuente	gl	SC	CM	f
Tratamientos	2	28.78	14.39	1.04
Bloques	17	2977.67	175.16	12.68
Error	34	469.55	13.81	
Total	53	3476.00		

La calificación promedio real de adaptación al parecer no depende del tratamiento aplicado. **b.** Sí; f_B es bastante grande, lo que indica gran variabilidad entre sujetos.

9.

Fuente	gl	SC	CM	f	$F_{0.05}$
Tratamientos	3	81.19	27.06	22.4	3.01
Bloques	8	66.50	8.31		
Error	24	29.06	1.21		
Total	35	176.75			

1	4	3	2
8.56	9.22	10.78	12.44

11. Los residuos son 0.0350, 0.0117, −0.0750, 0.0283, 0.0875, −0.0758, −0.0825, 0.0708, −0.1225, 0.0642, 0.1575 y −0.0992. El patrón en la gráfica de probabilidad normal es bastante lineal.

13. **b.** Cada suma de cuadrados se multiplica por c^2, pero f_A y f_B no cambian.

15. **a.** Aproximadamente 0.20, 0.43 **b.** Más o menos 0.30

17. **a.** $f_A = 3.76$, $f_B = 6.82$, $f_{AB} = 0.74$ y $F_{0.05,2,9} = 4.26$, de manera que parece importante la cantidad de fibra de carbono agregada. **b.** $f_A = 6.54$, $f_B = 5.33$, $f_{AB} = 0.27$

19. **a.**

Fuente	gl	SC	CM	f
Carbón	2	1.00241	0.50121	29.49
NaOH	2	0.12431	0.06216	3.66
Interacción	4	0.01456	0.00364	0.21
Error	9	0.15295	0.01699	
Total	17	1.29423		

El tipo de carbón al parecer afecta la acidez total.
b. Los carbones 1 y 3 no parecen diferir de forma importante entre sí, pero ambos difieren bastante del carbón 2.

21. **a, b.**

Fuente	gl	SC	CM	f
A	2	22941.80	11470.90	22.98
B	4	22765.53	5691.38	5.60
AB	8	3993.87	499.23	0.49
Error	15	15253.50	1016.90	
Total	29	64954.70		

Se rechaza tanto H_{0A} como H_{0B}.

23.

Fuente	gl	SC	CM	f
A	2	11 573.38	5786.69	$\frac{\text{MSA}}{\text{MSAB}} = 26.70$
B	4	17 930.09	4482.52	$\frac{\text{MSB}}{\text{MSE}} = 28.51$
AB	8	1734.17	216.77	$\frac{\text{MSAB}}{\text{MSE}} = 1.38$
Error	30	4716.67	157.22	
Total	44	35 954.31		

Puesto que $F_{0.01,8,30} = 3.17$, $F_{0.01,2,8} = 8.65$, y $F_{0.01,4,30} = 4.02$, no se rechaza H_{0AB} pero se rechazan H_{0A} y H_{0B}.

25. (−1.39, −1.05)

27. a.

Fuente	gl	SC	CM	f	$F_{0.05}$
A	2	14 144.44	7 072.22	61.06	3.35
B	2	5 511.27	2 755.64	23.79	3.35
C	2	244 696.39	122 348.20	1 056.27	3.35
AB	4	1 069.62	267.41	2.31	2.73
AC	4	62.67	15.67	0.14	2.73
BC	4	331.67	82.92	0.72	2.73
ABC	8	1 080.77	135.10	1.17	2.31
Error	27	3 127.50	115.83		
Total	53	270 024.33			

d. $Q_{0.05,3,27} = 3.51$, $w = 8.90$, y los tres niveles difieren significativamente entre sí.

29.

Fuente	gl	SC	CM	f
A	2	12.896	6.448	1.04
B	1	100.041	100.041	16.10
C	3	393.416	131.139	21.10
AB	2	1.646	0.823	<1
AC	6	71.021	11.837	1.905
BC	3	1.542	0.514	<1
ABC	6	9.771	1.629	<1
Error	72	447.500	6.215	
Total	95	1037.833		

a. No hay efectos de interacción significativos.

b. Los principales efectos de factor B y factor C son significativos.

c. $w = 1.89$; sólo las máquinas 2 y 4 no difieren significativamente entre sí.

31. La columna de valor P muestra que varios efectos de interacción son significativos al nivel 0.01.

33.

Fuente	gl	SC	CM	f
A	6	67.32	11.02	
B	6	51.06	8.51	
C	6	5.43	0.91	0.61
Error	30	44.26	1.48	
Total	48	168.07		

$F_{0.05,6,30} = 2.42$, $f_C = 0.61$, así que no se rechaza H_{0C}.

35.

Fuente	gl	SC	CM	f
A	4	28.88	7.22	10.7
B	4	23.70	5.93	8.79
C	4	0.62	0.155	<1
Error	12	8.10	0.675	
Total	24	61.30		

Puesto que $F_{0.05,4,12} = 3.26$, tanto A y B son importantes.

37.

Fuente	gl	CM	f
A	2	2207.329	2259*
B	1	47.255	48.4*
C	2	491.783	503*
D	1	0.044	<1
AB	2	15.303	15.7*
AC	4	275.446	282*
AD	2	0.470	<1
BC	2	2.141	2.19
BD	1	0.273	<1
CD	2	0.247	<1
ABC	4	3.714	3.80
ABD	2	4.072	4.17*
ACD	4	0.767	<1
BCD	2	0.280	<1
$ABCD$	4	0.347	<1
Error	36	0.977	
Total	71	93.621	

*Indica una relación F significativa.

39. a. $\hat{\beta}_1 = 54.38$, $\hat{\gamma}_{11}^{AC} = -2.21$, $\hat{\gamma}_{21}^{AC} = 2.21$.

b.

Fuente	Contraste de efecto	CM	f
A	1307	71 177.04	436.7
B	1305	70 959.34	435.4
C	529	11 660.04	71.54
AB	199	1 650.04	10.12
AC	−53	117.04	<1
BC	57	135.38	<1
ABC	27	30.38	<1
Error		162.98	

41.

Fuente	SC	f
A	136 640.02	1 007.6
B	139 644.19	1 029.8
C	24 616.02	181.5
D	20 377.52	150.3
AB	2 173.52	16.0
AC	2.52	<1
AD	58.52	<1
BC	165.02	1.2
BD	9.19	<1
CD	17.52	<1
ABC	42.19	<1
ABD	117.19	<1
ACD	188.02	1.4
BCD	13.02	<1
$ABCD$	204.19	1.5
Error	4339.33	
Total	328 607.98	

$F_{0.05,1,32} = 4.15$, de modo que sólo los cuatro efectos principales y la interacción AB parecen importantes.

43.

Fuente	gl	SC	f
A	1	0.436	<1
B	1	0.099	<1
C	1	0.109	<1
D	1	414.12	851
AB	1	0.003	<1
AC	1	0.078	<1
AD	1	0.017	<1
BC	1	1.404	3.62
BD	1	0.456	<1
CD	1	2.190	4.50
Error	5	2.434	

$F_{0.05,1,5} = 6.61$, de modo que se considera importante el efecto principal del factor D.

45. **a.** 1: (1), *ab, cd, abcd;* 2: *a, b, acd, bcd;* 3: *c, d, abc, abd;* 4: *ac, bc, ad, bd.*

b.

Fuente	gl	SC	f
A	1	14 028.13	53.89
B	1	92 235.13	345.33
C	1	3.13	<1
D	1	18.00	<1
AC	1	105.13	<1
AD	1	200.00	<1
BC	1	91.13	<1
BD	1	420.50	1.62
ABC	1	276.13	1.06
ABD	1	2.00	<1
ACD	1	450.00	1.73
BCD	1	2.00	<1
Bloques	7	898.88	<1
Error	12	3 123.72	
Total	31	111 853.88	

$F_{0.01,1,12} = 9.33$, entonces sólo son importantes los efectos principales de A y B.

47. **a.** *ABFG:* (1), *ab, cd, ce, de, fg, acf, adf, adg, aef, acg, aeg, bcg, bcf, bdf, bdg, bef, beg, abcd, abce, abde, abfg, cdfg, cefg, defg, acdef, acdeg, bcdef, bcdeg, abcdfg, abcefg, abdefg.* {*A, BCDE, ACDEFG, BFG*}, {*B, ACDE, BCDEFG, AFG*}, {*C, ABDE, DEFG, ABCFG*}, {*D, ABCE, CEFG, ABDFG*}, {*E, ABCD, CDFG, ABEFG*}, {*F, ABCDEF, CDEG, ABG*}, {*G, ABCDEG, CDEF, ABF*}. **b.** 1: (1), *aef, beg, abcd, abfg, cdfg, acdeg, bcdef;* 2: *ab, cd, fg, aeg, bef, acdef, bcdeg, abcdfg;* 3: *de, acg, adf, bcf, bdg, abce, cefg, abdefg;* 4: *ce, acf, adg, bcg, bdf, abde, defg, abcefg.*

49. $SSA = 2.250$, $SSB = 7.840$, $SSC = 0.360$, $SSD = 52.563$, $SSE = 10.240$, $SSAB = 1.563$, $SSAC = 7.563$, $SSAD = 0.090$, $SSAE = 4.203$, $SSBC = 2.103$, $SSBD = 0.010$, $SSBE = 0.123$, $SSCD = 0.010$, $SSCE = 0.063$, $SSDE = 4.840$. SC del error = suma de la SC de dos factores = 20.568, CM del error = 2.057, $F_{0.01,1,10} = 10.04$, por lo que sólo el efecto principal D es significativo.

51.

Fuente	gl	SC	CM	f
Efectos principales de A	1	322.667	322.667	980.38
Efectos principales de B	3	35.623	11.874	36.08
Interacción	3	8.557	2.852	8.67
Error	16	5.266	0.329	
Total	23	372.113		

$F_{0.05,3,16} = 3.24$, así que al parecer están presentes las interacciones.

53.

Fuente	gl	SC	CM	f
A	1	30.25	30.25	6.72
B	1	144.00	144.00	32.00
C	1	12.25	12.25	2.72
AB	1	1 122.25	1 122.25	249.39
AC	1	1.00	1.00	0.22
BC	1	12.25	12.25	2.72
ABC	1	16.00	16.00	3.56
Error	4	36.00	4.50	
Total	7			

Sólo el efecto principal para B y el efecto de interacción AB son significativos al nivel $\alpha = 0.01$.

55. **a.** $\hat{\alpha}_1 = 9.00$, $\hat{\beta}_1 = 2.25$, $\hat{\delta}_1 = 17.00$, $\hat{\gamma}_1 = 21.00$, $(\hat{\alpha\beta})_{11} = 0$, $(\hat{\alpha\delta})_{11} = 2.00$, $(\hat{\alpha\gamma})_{11} = 2.75$, $(\hat{\beta\delta})_{11} = 0.75$, $(\hat{\beta\gamma})_{11} = 0.50$, $(\hat{\delta\gamma})_{11} = 4.50$

b. Una gráfica de probabilidad normal indica que los efectos principales A, C y D son bastante importantes, y quizá la interacción CD. De hecho, al agrupar las cuatro SC de interacción de tres factores y la SC de interacción de cuatro factores para obtener un error de la suma de cuadrados con base en 5 gl, y luego construir una tabla ANOVA se encuentra que éstos son los efectos más importantes.

57.

Fuente	gl	SC	CM	f	P
A	2	34 436	17 218	436.92	0.000
B	2	105 793	52 897	1342.30	0.000
C	2	516 398	258 199	6552.04	0.000
AB	4	6868	1717	43.57	0.000
AC	4	10 922	2731	69.29	0.000
BC	4	10 178	2545	64.57	0.000
ABC	8	6713	839	21.30	0.000
Error	27	1064	39		
Total	53	692 372			

Todos los efectos son importantes.

59. Con base en los valores P en la tabla de análisis de varianza, los factores estadísticamente significativos al nivel $\alpha = 0.01$ son el tipo de adhesivo y el tiempo de curado. El material conductor no tiene un efecto estadísticamente significativo sobre la resistencia de la unión. No hay interacciones importantes.

61.

Fuente	gl	SC	CM	f
A	4	285.76	71.44	0.594
B	4	227.76	56.94	0.473
C	4	2867.76	716.94	5.958
D	4	5536.56	1384.14	11.502
Error	8	962.72	120.34	$F_{0.05,4,8} = 3.84$
Total	24			

No se pueden rechazar H_{0A} ni H_{0B} mientras que sí se rechazan H_{0C} y H_{0D}.

Capítulo 12

1. a. Los diagramas siguientes se basan en repetir cinco veces cada valor de tallo (una vez para las hojas 0 y 1, una segunda vez para las hojas 2 y 3, etcétera).

```
17 | 0
17 | 2 3
17 | 4 4 5
17 | 6 7
17 |                          tallo: centenas y decenas
18 | 0 0 0 0 0 1 1            hoja: unidades
18 | 2 2 2 2
18 | 4 4 5
18 | 6
18 | 8
```

No hay valores atípicos ni espacios significativos, y la distribución es casi en forma de campana con una concentración razonablemente alta respecto a su centro, más o menos en 180.

```
0 | 8 8 9
1 | 0 0 0 0
1 | 3
1 | 4 4 4 4
1 | 6 6
1 | 8 8 8 9                  tallo: unidades
2 | 1 1                      hoja: decenas
2 |
2 | 5
2 | 6
2 |
3 | 0 0
```

Un valor característico aproximado es 1.6, y hay una dispersión razonable en torno a este valor. La distribución está un poco sesgada hacia valores grandes, los dos valores más grandes que podrían ser candidatos a valores atípicos.

b. No, porque observaciones con valores idénticos de x tienen valores distintos de y.

c. No, porque al parecer los puntos no caen en absoluto cerca de una recta o una curva simple.

3. Sí. Sí.

5. b. Sí.

c. Parece haber una relación cuadrática aproximada (los puntos caen cerca de una parábola).

7. a. 5050 **b.** 1.3 **c.** 130 **d.** −130

9. a. 0.095 **b.** −0.475 **c.** 0.830, 1.305 **d.** 0.4207, 0.3446 **e.** 0.0036

11. a. −0.01, −0.10 **b.** 3.00, 2.50 **c.** 0.3627 **d.** 0.4641

13. a. Sí, porque $r^2 = 0.972$

15. a.

```
2 | 9
3 | 3 3 5 5 6 6 6 7 7 8 8 9
4 | 1 2 2 3 5 6 6 8 9
5 | 1
6 | 2 9
7 | 9
8 | 0
```

Valor característico en la parte baja de los 40, variabilidad razonable, sesgo positivo, dos posibles valores atípicos

b. No

c. $y = 3.2925 + 0.10748x = 7.59$. No; es peligroso extrapolar

d. 18.736, 71.605, 0.738, sí

17. a. Se estima que 0.144 es el cambio esperado en el contenido de calcio relacionado con un aumento de 1 mg/cm^2 en la cantidad de material disuelto. Más o menos 86% de la variación observada en el contenido de calcio se puede atribuir a la relación del modelo de regresión lineal simple entre el contenido y la cantidad de material disuelto.

b. 10.88 **c.** 1.46

19. a. $y = -45.5519 + 1.7114x$ **b.** 339.51 **c.** −85.57 **d.** Las $\hat{y}_i$ son 125.6, 168.4, 168.4, 211.1, 211.1, 296.7, 296.7, 382.3, 382.3, 467.9, 467.9, 553.4, 639.0, 639.0; una recta de 45° que pase por (0, 0).

21. b. $y = -2.182 + 0.660x$ c. 7.72 d. 7.72

23. a. 16 213.64; 16 205.45
b. 414 235.71; sí, porque $r^2 = 0.961$.

27. $\hat{\beta}_1 = \sum x_i Y_i / \sum x_i^2$

29.

Conjunto de datos	r^2	s
1	0.43	4.03
2	0.99	4.03
3	0.99	1.90

Más efectivo: conjunto 3

Menos efectivo: conjunto 1

31. a. 0.001017 b. (−0.00956, −0.00516)

33. a. (0.081, 0.133)
b. H_a: $\beta_1 > 0.1$, valor $P = 0.277$, no

35. a. (0.63, 2.44) es un IC de 95%.
b. Sí. $t \approx 3.6$, valor $P \approx 0.004$
c. No; extrapolación
d. (0.54, 2.82), no

37. a. 0.110, 0.000262 b. $t \approx 3.1$, valor $P \approx 0.014 < 0.10 = \alpha$, por lo que se debe rechazar H_0. Parece que los datos contradicen la opinión previa.

39. $f = 71.97$, $s_{\beta_1} = 0.004837$, $t = 8.48$, valor $P = 0.000$

43. $d = 1.20$, gl = 13 y $\beta \approx 0.1$.

45. a. (77.80, 78.38)
b. (76.90, 79.28), el mismo centro pero más amplio
c. Más amplio, puesto que 115 está lejos de $\bar{x}$
d. $t = -11$, valor $P = 0$

47. a. El IP de 95% es (20.21, 43.69), no
b. (28.53, 51.92), al menos 90%

49. (431.3, 628.5)

51. a. 0.40 está más próximo $\bar{x} = 0.7495$ b. (0.745, 0.875) c. (0.059, 0.523)

53. (a) más angosto que (b), (c) más angosto que (d), (a) más angosto que (c), (b) más angosto que (d)

55. a. $x_2 = x_3 = 12$, sin embargo $y_2 \neq y_3$. b. Sí
c. $y = -19.670 + 3.2847x$ d. (24.93, 67.12)

57. Sí, por ejemplo, 18 años es la edad mínima de elección, entonces para la mayoría de las personas $y \approx x - 18$.

59. a. 0.966
b. El porcentaje de peso de fibra seca en el primer espécimen tiende a ser mayor que en el segundo.
c. Sin cambio. d. 93.3%
e. $t = 14.9$, valor $P \approx 0$, por lo que sí parece haber esa relación.

61. a. $r = 0.748$, $t = 3.9$, valor $P = 0.001$. Con un valor de $\alpha = 0.05$ o 0.01, sí.
b. 0.560 (56%), igual

63. $r = 0.773$, sin embargo $t = 2.44 < 2.776$; por tanto, no se puede rechazar H_0: $\rho = 0$.

65. a. La gráfica de x es un poco curva, aunque no de manera preocupante si se considera el pequeño tamaño de muestra. La gráfica de y es bastante recta. b. $t = 6.3 \geq 3.355$, así que parece haber una relación lineal.

67. a. Se rechaza H_0
b. No. Valor $P = 0.00032 \Rightarrow z \approx 3.6 \Rightarrow r \approx 0.16$, lo que indica sólo una relación débil.
c. Sí, pero muy grande $n \Rightarrow \rho \approx 0.022$, con ninguna importancia práctica.

69. a. $t = -1.24 > -2.201$, de modo que no se puede rechazar H_0. b. 0.970

71. a. 0.507 b. 0.712 c. Valor $P = 0.0013 < 0.01 = \alpha$, así que se rechaza H_0: $\beta_1 = 0$ y se concluye que hay una relación lineal útil. d. Un intervalo de confianza de 95% es (1.056, 1.275). e. 1.0143, 0.2143

73. a. $y = 14.1904 - 0.14892x$ b. $t = -1.43$, así que no se rechaza H_0: $\beta_1 = -0.10$. c. No; $\sum(x - \bar{x})^2 = 143$ aquí, y 182 para los datos. d. Un intervalo de confianza de 95% para $\mu_{Y \cdot 28}$ es (9.599, 10.443).

75. a. Una relación lineal considerable
b. $y = -0.08259 + 0.044649x$
c. 98.3%
d. 0.7702, −0.0902 e. Sí; $t = 19.96$
f. (0.0394, 0.0499) g. (0.762, 0.858)

79. b. 0.573

85. $t = -1.14$, por lo que es factible que $\beta_1 = \gamma_1$.

Capítulo 13

1. a. 6.32, 8.37, 8.94, 8.37 y 6.32 b. 7.87, 8.49, 8.83, 8.94 y 2.83 c. Es probable que la desviación sea mucho menor para los valores de x del inciso (b).

3. a. Sí. b. −0.31, −0.31, 0.48, 1.23, −1.15, 0.35, −0.10, −1.39, 0.82, −0.16, 0.62, 0.09, 1.17, −1.50, 0.96, 0.02, 0.65, −2.16, −0.79, 1.74. Aquí e/e^* varía entre 0.57 y 0.65, así que e^* está próxima a e/s. c. No.

5. **a.** Alrededor de 98% de la variación observada en el espesor se explica mediante la relación.
b. Una relación no lineal.

7. **a.** No. **b.** Las e_i son -16.60, 9.70, 19.00, -0.70, 11.40; las e_i^* son -1.55, 0.68, 1.25, -0.05, -1.06; una función cuadrática.

9. Para el conjunto 1 es adecuada la regresión lineal simple. Una regresión cuadrática es razonable para el conjunto 2. En el conjunto 3, (13, 12.74) parece muy inconsistente con los demás datos. La pendiente estimada para el conjunto 4 depende mucho de la observación (19, 12.5) y no es convincente la evidencia para una relación lineal.

11. **c.** $V(\hat{Y}_i)$ aumenta, y $V(Y_i - \hat{Y}_i)$ disminuye.

13. t con $n - 2$ gl; 0.02

15. **a.** Un patrón curvado **b.** Un patrón lineal
c. $Y = \alpha x^{\beta} \cdot \epsilon$ **d.** Un IP de 95% (3.06, 6.50).
e. Un residuo estandarizado, que corresponde a la tercera observación, es un poco grande. Sólo hay dos residuos estandarizados positivos, pero otros dos son esencialmente 0. Los patrones en una gráfica residual estandarizada y en una de probabilidad normal son ligeramente aceptables.

17. **a.** $\sum x_i' = 15.501$, $\sum y_i' = 13.352$, $\sum (x_i')^2 = 20.228$, $\sum x_i' y_i' = 18.109$, $\sum (y_i')^2 = 16.572$, $\hat{\beta}_1 = 1.254$, $\hat{\beta}_0 = -0.468$, $\hat{\alpha} = 0.626$, $\hat{\beta} = 1.254$ **c.** $t = -1.07$, así que no se rechaza H_0. **d.** H_0: $\beta = 1$, $t = -4.30$, así que se rechaza H_0.

19. **a.** No **b.** $Y' = \beta_0 + \beta_1 \cdot (1/t) + \epsilon'$, donde $Y' = \ln(Y)$, así que $Y = \alpha e^{\beta/t} \cdot \epsilon$. **c.** $\hat{\beta} = \hat{\beta}_1 = 3735.45$, $\hat{\beta}_0 = -10.2045$, $\hat{\alpha} = (3.70034) \cdot (10^{-5})$, $\hat{y}' = 6.7748$, $\hat{y} = 875.5$ **d.** SSE = 1.39587, SSPE = 1.36594 (usando valores transformados), $f = 0.33 < 8.68 = F_{0.01,1,15}$, así que no se rechaza H_0.

21. **a.** $\hat{\mu}_{Y \cdot x} = 18.14 - 1485/x$ **b.** $\hat{y} = 15.17$

23. Para el modelo exponencial, $V(Y \Omega x) = \alpha^2 e^{2\beta x} \sigma^2$, que no depende de x. Un resultado parecido se cumple para el modelo de potencia.

25. **a.** La estimación puntual de β_1 es $\hat{\beta}_1 = 0.1772$, y la relación de probabilidades estimada es 1.194. Se rechaza H_0: $\beta_1 = 0$ en favor de la conclusión de que al parecer la experiencia sí afecta la probabilidad de desempeñar con éxito la tarea.

27. **b.** 52.88, 0.12 **c.** 0.895 **d.** No
e. (48.54, 57.22) **f.** (42.85, 62.91)

29. **a.** SSE = 103.37, $s^2 = 51.69$ **b.** 0.961
c. $t = -3.83 > -4.303$, de modo que no se puede rechazar H_0: $\beta_2 = 0$. Los datos no contradicen mucho la inclusión del término cuadrático, quizá porque n es pequeña; H_0 se rechazaría a un nivel de significación de 0.10). **d.** β_1: $(-19.87, 35.99)$ β_2: $(-5.18, 1.50)$
e. (100.08, 129.34) **f.** 2.19

31. **a.** $13.636 + 11.406x - 1.7155x^2$
b. Sí. Sí, (6, 20) **c.** 2.040, 0.947. La prueba F de utilidad del modelo con MINITAB da como resultado $f = 35.9$, valor $P = 0.003$, lo que significa que el modelo es útil. **d.** Sí, sí **e.** (28.35, 35.28)

33. **a.** 0.9671, 0.9407
b. $0.0000492x^3 - 0.000446058x^2 + 0.007290688x + 0.96034944$ **c.** $t = 2 < 3.182 = t_{0.025,3}$, así que se debe eliminar el término cúbico. **d.** Son idénticas
e. 0.987, 0.994, sí

35. $\hat{y} = 7.6883e^{0.1799x - 0.0022x^2}$

37. **a.** 4.9 **b.** Cuando se mantiene fija la cantidad de entregas, el cambio promedio en tiempo de viaje relacionado con un aumento de una milla en la distancia recorrida es 0.060 h. Cuando se mantiene fija la distancia recorrida, el cambio promedio en el tiempo de viaje relacionado con una entrega extra es 0.900 h. **c.** 0.9861

39. **a.** 77.3 **b.** 40.4

41. $f = 24.4 > 5.12 = F_{0.001,6,30}$, por lo que el valor $P << 0.001$. Al parecer es útil el modelo elegido.

43. **a.** 48.31, 3.69 **b.** No. Si aumenta x_1 debe cambiar ya sea x_3 o x_2. **c.** Sí, puesto que $f = 18.924$, valor $P = 0.001$. **d.** Sí, con $\alpha = 0.01$, puesto que $t = 3.496$ y el valor $P = 0.003$.

45. **a.** $f = 87.6$, valor $P = 0$, al parecer hay una relación lineal útil entre y y por lo menos uno de los predictores.
b. 0.935 **c.** (9.095, 11.087)

47. **b.** Valor $P = 0.000$, por lo tanto se concluye que el modelo es útil. **c.** Valor $P = 0.034 \leq 0.05 = \alpha$, y, en consecuencia, se rechaza H_0: $\beta_3 = 0$; al parecer el porcentaje de basura proporciona información adicional útil.
d. (1479.8, 1531.1), precisión razonable **e.** Un IP de 95% es (1435.7, 1575.2).

49. **a.** 96.8303, -5.8303 **b.** $f = 14.9 \geq 8.02 = F_{0.05,2,9}$, por lo que se rechaza H_0 y se concluye que el modelo es útil. **c.** (78.28, 115.38) **d.** (38.50, 155.16) **e.** (46.91, 140.66) **f.** No. Valor $P = 0.208$, de modo que no se puede rechazar H_0: $\beta_1 = 0$.

51. **a.** No **b.** $f = 5.04 \geq 3.69 = F_{0.05,5,8}$. Al parecer hay una relación lineal útil. **c.** 6.16, 3.304, (16.67, 31.91)
d. $f = 3.44 < 4.07 = F_{0.05,3,8}$, así que no se puede rechazar H_0: $\beta_3 = \beta_4 = \beta_5 = 0$. Se pueden eliminar los términos cuadráticos.

55. **a.** La variable dependiente es $\ln(q)$, y los predictores son $x_1 = \ln(a)$ y $x_2 = \ln(b)$; $\hat{\beta} = \hat{\beta}_1 = 0.9450$, $\hat{\gamma} = \hat{\beta}_2 = 0.1815$, $\hat{\alpha} = 4.7836$, $\hat{q} = 18.27$.
b. Ahora se hace la regresión de $\ln(q)$ en función de $x_1 = a$ y $x_2 = b$.
c. (1.24, 5.78)

57.

k	R^2	R^2 ajustada	C_k
1	0.676	0.647	138.2
2	0.979	0.975	2.7
3	0.9819	0.976	3.2
4	0.9824		4

a. El modelo con $k = 2$ **b.** No

59. El modelo con predictores x_1, x_3 y x_5

61. No. Todos los valores de R^2 son mucho menores que 0.9.

63. Se debe investigar más el impacto de estas dos observaciones. No en absoluto. También se deben considerar la eliminación de la observación #6 seguida de una nueva regresión.

65. **a.** Las dos distribuciones tienen cantidades similares de variabilidad, ambas son razonablemente simétricas y no contienen valores atípicos. La diferencia principal es que la mediana de los valores con fisura es casi 840, mientras para los valores sin fisura es 480. Un IC de 95% t para la diferencia entre medias es (132, 557).

b. $r^2 = 0.577$ para el modelo de regresión lineal simple, valor P para la utilidad del modelo = 0, pero un residuo estandarizado es -4.11! Incluir un indicador para fisura-sin fisura no mejora el ajuste, ni tampoco incluir un indicador y un predictor de interacción.

67. **a.** Cuando se mantienen fijos género, peso y frecuencia cardiaca, se estima que el cambio promedio en VO_2 máx relacionado con un aumento de un minuto en el tiempo de caminata es -0.0996. **b.** Cuando se mantienen fijos peso, tiempo de caminata y frecuencia cardiaca, la estimación de la diferencia promedio entre VO_2 máx para hombres y mujeres es 0.6566. **c.** 3.669, -0.519 **d.** 0.706

e. $f = 9.0 \geq 4.89 = F_{0.01,4,15}$, así que al parecer hay una relación útil.

69. **a.** No. Hay curvatura considerable en el diagrama de dispersión. **b.** Con la regresión cúbica se obtiene $R^2 = 0.998$ y un IP de 95% de (261.98, 295.62), y al parecer el predictor cúbico es importante (valor (P = 0.001). Una regresión de y en función de $\ln(x)$ tiene r^2 = 0.991, pero hay un residuo estandarizado muy grande y la gráfica residual estandarizada no es satisfactoria.

71. **a.** $R^2 = 0.802, f = 21.03$, valor $P = 0.000$. El pH es un candidato a eliminación. Observe que hay un residuo estandarizado extremadamente grande.

b. $R^2 = 0.920$, ajustada $R^2 = 0.774, f = 6.29$, valor P = 0.002

c. $f = 1.08$, valor $P > 0.10$, no se rechaza H_0: $\beta_6 = \cdots = \beta_{20} = 0$. Al parecer no es útil el grupo de predictores de segundo orden.

d. $R^2 = 0.871, f = 28.50$, valor $P = 0.000$, y ahora son considerados importantes los seis predictores (el valor P máximo para cualquier relación t es 0.016); la importancia del pH^2 se enmascaró en el resto de (c). Observe que hay dos residuos estandarizados bastante grandes.

73. **a.** $f = 1783$, así que al parecer es útil el modelo.

b. $t = -48.1 \leq -6.689$, así que incluso al nivel 0.001 se debe conservar el predictor cuadrático.

c. No **d.** (21.07, 21.65) **e.** (20.67, 22.05)

75. **a.** $f = 30.8 \geq 9.55 = F_{0.01,2,7}$, de modo que al parecer es útil el modelo. **b.** $t = -7.69$ y el valor $P < 0.001$, así que se retiene el predictor cuadrático. **c.** (44.01, 47.91)

77. **a.** 231.75 **b.** 0.903 **c.** $f = 41.9$, lo que indica una relación útil. **d.** (220.9, 238.1)

79. Hay varias opciones rezonables en cada caso.

81. **a.** $f = 106$, valor $P \approx 0$ **b.** (0.014, 0.068)

c. $t = 5.9$, se rechaza H_0: $\beta_4 = 0$, parece importante el porcentaje de no blancos. **d.** 99.514, $y - \hat{y} = 3.486$

Capítulo 14

1. **a.** Se rechaza H_0. **b.** No se rechaza H_0.
c. No se rechaza H_0. **d.** No se rechaza H_0.

3. No, $\chi^2 = 1.57$ y valor $P > 0.10$, así que no se debe rechazar H_0.

5. $\chi^2 = 6.61 < 14.684 = \chi^2_{0.10,9}$, por lo tanto no se rechaza H_0.

7. $\chi^2 = 4.03$ y valor $P > 0.10$, así que no se rechaza H_0.

9. **a.** [0, 0.2231), [0.2231, 0.5108), [0.5108, 0.9163), [0.9163, 1.6094) y [1.6094, ∞) **b.** $\chi^2 = 1.25 < \chi^2_{\alpha,4}$ para cualquier α razonable, por lo que la distribución exponencial especificada es bastante factible.

11. **a.** $(-\infty, -0.97)$, $[-0.97, -0.43)$, $[-0.43, 0)$, $[0, 0.43)$, $[0.43, 0.97)$ y $[0.97, \infty)$ **b.** $(-\infty, 0.49806)$, [0.49806, 0.49914), 0[0.49914, 0.5), [0.5, 0.50086), [0.50086, 0.50194) y [0.50194, ∞) **c.** $\chi^2 = 5.53$, $\chi^2_{0.10,5} = 9.236$, así que el valor $P > 0.10$ y es factible la distribución normal especificada.

13. $\hat{p} = 0.0843$, $\chi^2 = 280.3 > \chi^2_{\alpha,1}$ para cualquier α tabulada, así que el modelo proporciona un ajuste deficiente.

15. La probabilidad es proporcional a $\theta^{233}(1 - \theta)^{367}$, de donde $\hat{\theta} = 0.3883$. Las cantidades esperadas estimadas son 21.00, 53.33, 50.78, 21.50 y 3.41. Al combinar las celdas 4 y 5, $\chi^2 = 1.62$, por lo tanto no se rechaza H_0.

17. $\hat{\lambda} = 3.167$, de lo cual $\chi^2 = 103.98 >> \chi^2_{\alpha,k-1} = \chi^2_{\alpha,7}$ para cualquier α tabulada, por lo que la distribución de Poisson proporciona un ajuste muy deficiente.

19. $\hat{\theta}_1 = (2n_1 + n_3 + n_5)/2n = 0.4275$, $\hat{\theta}_2 = 0.2750$, $\chi^2 = 29.1$, $\chi^2_{0.01,3} = 11.344$, por lo que se rechaza H_0.

21. Sí. No se puede rechazar la hipótesis nula de una distribución normal de la población.

23. Con MINITAB se obtiene $r = 0.967$, y como $c_{0.10} = 0.9707$ y $c_{0.05} = 0.9639$, $0.05 <$ valor $P < 0.10$. Con $\alpha = 0.05$, se juzga factible la normalidad.

25. $\chi^2 = 23.18 \geq 13.277 = \chi^2_{0.01,4}$, así que se rechaza H_0. Al parecer las proporciones son distintas.

27. Sí. $\chi^2 = 44.98$ y valor $P < 0.001$.

29. p_{ij} = proporción de la j-ésima combinación de sexo resultante del i-ésimo genotipo. $\chi^2 = 6.46$, así que el valor $P > 0.10$ y no se puede rechazar la hipótesis nula de homogeneidad.

31. Sí. $\chi^2 = 14.15$, por lo tanto $0.025 <$ valor $P < 0.03$ y se debe rechazar H_0 al nivel de significación 0.05.

35. N_{ij}/n, $n_k N_{ij}/n$, 24

37. $\chi^2 = 3.65 < 5.992 = \chi^2_{0.05,2}$, por lo tanto no se puede rechazar H_0.

39. Sí. $\chi^2 = 131$ y valor $P < 0.001$.

41. $\chi^2 = 22.4$ y valor $P < 0.001$, por lo que se rechaza la hipótesis nula de independencia.

43. Valor $P = 0$, por lo tanto se rechaza la hipótesis nula de homogeneidad.

47. **a.** Valor de la estadística de prueba = 19.2, valor $P = 0$
b. Evidencia de cuando mucho una relación débil; valor de la estadística de prueba = -2.13
c. Valor de la estadística de prueba = -0.98, valor $P > 0.10$
d. Valor de la estadística de prueba = 3.3, $0.01 <$ valor $P < 0.05$

Capítulo 15

1. $s_+ = 35$ y $14 < 35 < 64$, por lo que no se puede rechazar H_0.

3. $s_+ = 18 \leq 21$, así que se rechaza H_0.

5. Se rechaza H_0 si $s_+ \geq 64$ o $s_+ \leq 14$. Debido a que $s_+ = 72$, se rechaza H_0.

7. $s_+ = 442.5$, $z = 2.89 \geq 1.645$, así que se rechaza H_0.

9.

d	0	2	4	6	8	10	12	14	16	18	20
$p(d)$	$\frac{1}{24}$	$\frac{3}{24}$	$\frac{1}{24}$	$\frac{4}{24}$	$\frac{2}{24}$	$\frac{2}{24}$	$\frac{2}{24}$	$\frac{4}{24}$	$\frac{1}{24}$	$\frac{3}{24}$	$\frac{1}{24}$

11. $w = 37$ y $29 < 37 < 61$, por lo que no se puede rechazar H_0.

13. $z = 2.27 < 2.58$, de modo que no se puede rechazar H_0. Valor $P \approx 0.023$

15. $w = 39 \leq 41$, así que se rechaza H_0.

17. $(\bar{x}_{(5)}, \bar{x}_{(32)}) = (11.15, 23.80)$

19. $(-13.0, -6.0)$

21. $(d_{ij(5)}, d_{ij(21)}) = (16, 87)$

23. $k = 14.06 \geq 6.251$, así que se rechaza H_0.

25. $k = 9.23 \geq 5.992$, por lo que se rechaza H_0.

27. $f_r = 2.60 < 5.992$, por lo tanto no se rechaza H_0.

29. $f_r = 9.62 > 7.815 = \chi^2_{0.05,3}$, entonces se rechaza H_0.

31. $(-5.9, -3.8)$

33. **a.** 0.021 **b.** Con $c = 14$ se obtiene $\alpha = 0.058$; $y = 12$, por lo que no se puede rechazar H_0.

35. $w' = 26 < 27$, por lo tanto no se rechaza H_0.

Capítulo 16

1. Todos los puntos de la gráfica caen entre los límites de control.

3. 0.9802, 0.9512, 53

5. **a.** 0.0301 **b.** 0.2236 **c.** 0.6808

7. LCI = 12.20, LCS = 13.70. No.

9. LCI = 94.91, LCS = 98.17. Al parecer hay un problema en el día 22.

11. **a.** 200 **b.** 4.78 **c.** 384.62 (mayor), 6.30 (menor)

13. LCI = 12.37, LCS = 13.53

15. **a.** LCI = 0, LCS = 6.48
b. LCI = 0.48, LCS = 6.60

17. LCI = 0.045, LCS = 2.484. Sí, puesto que todos los puntos están dentro de los límites de control.

19. **a.** LCI = 0.105, LCS = 0.357
b. Sí, puesto que 0.39 > LCS.

21. $\bar{p} > 3/53$

23. LCI = 0, LCS = 10.1

25. Cuando el área = 0.6, LCI = 0 y LCS = 14.6; cuando el área = 0.8, LCI = 0 y LCS = 13.4; cuando el área = 1.0, LCI = 0 y LCS = 12.6.

27.

l:	1	2	3	4	5	6	7	8
d_l:	0	0.001	0.017	0	0	0.010	0	0
e_l:	0	0	0	0.038	0	0	0	0.054
l:	9	10	11	12	13	14	15	
d_l:	0	0.024	0.003	0	0	0	0.005	
e_l:	0	0	0	0.015	0	0	0	

No hay señales fuera de control.

29. $n = 5$, $h = 0.00626$

31. Las probabilidades hipergeométricas (determinadas con una calculadora HP21S) son 0.9919, 0.9317, 0.8182, 0.6775, 0.5343, 0.4047, 0.2964, 0.2110, 0.1464 y 0.0994, mientras que las probabilidades binomiales correspondientes son 0.9862, 0.9216, 0.8108, 0.6767, 0.5405, 0.4162, 0.3108, 0.2260, 0.1605 y 0.1117. La aproximación es satisfactoria.

33. 0.9206, 0.6767, 0.4198, 0.2321, 0.1183; es preferible el plan con $n = 100$, $c = 2$.

35. 0.9981, 0.5968 y 0.0688

37. **a.** 0.010, 0.018, 0.024, 0.027, 0.027, 0.025, 0.022, 0.018, 0.014, 0.011 **b.** 0.0477, 0.0274 **c.** 77.3, 202.1, 418.6, 679.9, 945.1, 1188.8, 1393.6, 1559.3, 1686.1, 1781.6

39. Gráfica $\bar{x}$ basada en desviaciones estándar muestrales: LCI = 402.42, LCS = 442.20. $\bar{x}$ basada en recorridos muestrales: LCI = 402.36, LCS = 442.26. Gráfica S: LCI = 0.55, LCS = 30.37. Gráfica R: LCI = 0, LCS = 82.75.

41. Gráfica S: LCI = 0, LCS = 2.3020; como $s_{21} = 2.931 >$ LCS, parece que el proceso está fuera de control en ese momento. Debido a que se identifica una causa asignable, después de la eliminación se recalculan los límites: para una gráfica S, LCI = 0 y LCS = 2.0529; para una gráfica $\bar{x}$, LCI = 48.583 y LCS = 51.707. Todos los puntos en ambas gráficas quedan entre los límites de control.

43. $\bar{\bar{x}} = 430.65$, $s = 24.2905$; para una gráfica S, LCS = 62.43 cuando $n = 3$ y LCS = 55.11 cuando $n = 4$; para una gráfica $\bar{x}$, LCI = 383.16 y LCS = 478.14 cuando $n = 3$, y LCI = 391.09 y LCS = 470.21 cuando $n = 4$.

Índice

Adecuación del modelo en la regresión, 556-561, 602-603
  regresión, 500, 577, 587
Alias, 470
Análisis de datos categóricos, 632-666
Análisis de varianza, 410
  de un solo factor (unidireccional), 410-440
  de varios factores, 441-495
  en la regresión, 527
  sin distribución, 689-692
Análisis residual, 556-561
ANOVA de dos factores, 442-465
ANOVA de tres factores, 465-475
ANOVA de un solo factor, 410
ANOVA de varios factores, 441-495
ANOVA unidireccional, 410
Asociación, 224
Atributos, 713
Axiomas de probabilidad, 58

Bloque principal, 484
Bloques, 449, 452, 481-482

Calidad promedio de salida, 732
Causa asignable, 698
Causalidad, 224
Censo, 3
Censura, 36, 262
Centrado de valores de x, 582-583
Coeficiente ajustado de determinación múltiple, 580, 587, 617
Coeficiente de correlación
  intervalo de confianza. *Véase* Intervalo de confianza
  múltiple, 594
  estimación puntual. *Véase* Estimación puntual
  poblacional, 542
  muestral, 539
  prueba de hipótesis. *Véase* Prueba de hipótesis
  variables aleatorias, 222
Coeficiente de correlación múltiple, 594
Coeficiente de determinación múltiple, 580, 594
  ajustado, 580, 594
Coeficiente de determinación, 514
Coeficiente de variación, 48
Coeficientes de autocorrelación, 548
Coincidencias, 94
Combinación lineal, 244
  de variables normales, 245
  valor esperado de, 244
  varianza de, 244
Combinación, 71
Comparaciones múltiples, 422, 433-434, 448, 460
Complemento de un suceso, 55
Componente de varianza, 453
Computadoras, 4
Conexión de continuidad, 169
Confusión, 482-483
Contraste, 440, 477
Covarianza, 220
Cuadrado grecolatino, 494
Cuadrado latino, 470
Cuadrado medio, 415, 419
Cuadrados medios esperados, 421, 429, 437, 453, 459, 462, 465, 467, 474
Cuarta dispersión, 41
Cuartil, 32
Cuarto de replicación, 485
Cuarto, inferior, 41
Cuarto medio, 50
Cuarto, superior, 41
Cuentas de celda esperada estimada, 644, 656, 659
  esperada, 635
  observada, 635, 655

Curva característica de operación, 728
Curva de densidad, 147
Curva de Gompertz, 576

Datos por pares, 381, 387-388, 672
Datos, 3
  bivariados, 3
  censurados, 36, 262
  cualitativos (categóricos), 22, 33
  multivariados, 4, 22
  univariados, 3
Deming, W. E., 7
Desigualdad de Chebyshev, 119
Desigualdad de Jensen, 203
Desviación aleatoria, 500
Desviación estándar
  estimación puntual. *Véase* Estimación puntual
  intervalo de confianza. *Véase* Intervalo de confianza
  poblacional, 38
  muestral, 37
  prueba de hipótesis. *Véase* Prueba de hipótesis
  variable aleatoria continua, 158
  variable aleatoria discreta, 116
Desviaciones de la media, 36
Determinación del tamaño de la muestra
  intervalo de confianza, 287, 296, 369
  prueba de hipótesis, 330, 335, 341, 366, 395
Diagrama comparativo de tallo y hojas, 24
Diagrama de árbol, 68-69
Diagrama de caja comparativo, 43
Diagrama de caja, 41-42
  comparativo, 43
Diagrama de dispersión, 498
  smoother, 570
Diagrama de Pareto, 27
Diagrama de puntos y dígitos, 14
Diagrama de puntos, 14
Diagrama de tallo y hoja, 11
  comparativo, 24
Diagrama de Venn, 56
Diagramas residuales, 558, 595
Diseño completo, 470
Diseño de medidas repetidas, 451
Diseño incompleto, 470
Disjuntos por pares, 56
Distribución beta, 185
Distribución binomial negativa, 132
  estimación puntual, 271
Distribución binomial, 123-126
  aproximación a la hipergeométrica, 131
  aproximación de la distribución normal a la, 169
  aproximación de Poisson a la, 136
  tablas, 1124, 736-738
Distribución chi-cuadrada, 179, 308, 635
  tabla, 309, 636, 745, 755-756
  valor crítico, 309, 635
Distribución de Bernoulli, 105
Distribución de Cauchy, 262
Distribución de Erlang, 180
Distribución de frecuencia, 15
Distribución de intervalo estudentizado, 422
  tabla, 422, 754
Distribución de muestreo, 225
  de una media muestral, 237-239
Distribución de Pareto, 151
Distribución de Poisson, 135
  aproximación a la binomial, 136
  aproximación normal a la, 243
  en gráficas de control, 715
  estimación puntual, 267, 273
  intervalo de confianza, 299
  prueba de hipótesis, 357, 647
  relación con la distribución exponencial, 178
  tabla, 136, 738-739
Distribución de probabilidad (*Véase también* cada una de las distribuciones)
  Bernoulli, 105
  beta, 185
  binomial, 123-126
  Cauchy, 262
  condicional, 215
  continua, 147
  chi-cuadrada, 179
  de valores extremos, 195, 201
  de una función, 202
  discreta, 101-102
  Erlang, 180
  exponencial, 177
  F, 399, 416
  familia, 104
  gamma, 175
  geométrica, 133
  normal bivariada, 543
Distribución de probabilidad conjunta, 206, 208, 212
Distribución de probabilidad, (*continuación*)
  binomial negativa, 132
  continua conjunta, 208, 212
  discreta conjunta, 206, 212
  de combinación lineal, 244-245
  de la media muestral, 237-239
  de la muestra total, 237-239
  de Pareto, 151
  de Poisson, 135
  de Rayleigh, 151
  de Weibull, 181
  exponencial desplazada, 278

hiperexponencial, 201
hipergeométrica, 129-131, 217
intervalo estudentizado, 422
lognormal, 184
multinomial, 213
normal, 161
normal estándar, 162
simétrica, 156
t, 300
uniforme, 148
Distribución de Rayleigh, 151, 268
Distribución de valores extremos, 195, 201
Distribución de Weibull, 181
estimación puntual, 274
gráfica de probabilidad, 196
Distribución exponencial (continuación)
gráfica de probabilidad, 199
prueba de hipótesis, 358
relación con la distribución de Poisson, 178
Distribución exponencial, 177, 278
estimación puntual, 262, 265, 272, 278
intervalo de confianza, 288-289
Distribución *F*, 399, 416
tabla, 399, 416, 748-753
Distribución gamma, 175
estándar, 175
estimación puntual, 270
Distribución geométrica, 133
Distribución hiperexponencial, 201
Distribución hipergeométrica, 129-131
alternativa, 316
aproximación binomial a la, 132
compuesta, 643
Hipótesis, 316
multivariada, 217
nula, 316
simple, 642
Distribución lognormal, 184
de un producto, 242
Distribución multinomial, 213
Distribución normal, 161
aproximación a la binomial, 169
aproximación a distribuciones discretas, 168
aproximación a la distribución gamma, 243
aproximación a la distribución de Poisson, 243
bivariada, 543
estándar, 162
estimación puntual, 261, 273
intervalo de confianza, 286, 302, 310
intervalo de predicción, 304
intervalo de tolerancia, 305
gráfica de probabilidad, 192, 487
percentil, 163, 167
prueba chi-cuadrada, 649
prueba basada en la gráfica de probabilidad, 651
prueba de hipótesis, 328, 333, 358, 362-363
tabla de la distribución normal estándar, 162, 740-741
Distribución normal bivariada, 543
Distribución normal estándar, 162
tabla, 162
Distribución sesgada, 193-194
Distribución simétrica, 156
Distribución *t*, 300
tabla, 301, 743, 746-747
Distribución uniforme, 148
estimación puntual, 276
intervalo de confianza, 313
Distribución *z*, 162
Distribución, probabilidad. *Véase* Distribución de probabilidad
Duración de corrida promedio, 705, 724

Ecuación de modelo ANOVA, 428-429, 435, 445, 452, 453, 457, 461, 466, 471, 476
Ecuaciones normales, 507, 578, 592
Efectos
aleatorio, 435
fijo, 435
mixto, 452-453
principal, 457
Efectos aleatorios, 435, 452
Efectos fijos, 435
Efectos mixtos, 452-453
Eficiencia relativa asintótica, 675, 682
Eficiencia, relativa asintótica, 675, 682
Eliminación hacia atrás, 619
Empates, 673, 681
Ensayo, 120
Error aleatorio, 500
Error cuadrático medio, 279, 616
Error de tipo I, 319
Error tipo II, 319
prueba *t*, 335-336, 377-378, 530, 762
prueba *F* ANOVA, 430
prueba sin distribución, 675, 682
prueba *z*, 330, 333, 341, 365, 394
Error estándar, 263
Error total esperado normalizado de la estimación, 617
Errores en la prueba de hipótesis
tipo I, 319
tipo II, 319
Escala de densidad, 19
Espacio muestral, 53
Estadístico de prueba, 318
Estadística descriptiva, 4
Estadística inferencial, 5

Estadístico, 227
prueba estadística, 318
Estadística, 3-4
descriptiva, 4
inferencial, 5
Estandarización, 165
variables, 586, 614
Estimación de máxima verosimilitud, 272-277
y pruebas de bondad de ajuste, 643
Estimación *M*, 277
Estimación puntual
combinación lineal de medias, 426
coeficiente de correlación, 542
desviación estándar, 256, 259, 275
distribución binomial negativa, 271
distribución de Cauchy, 262
diferencia entre medias, 361
distribución de Poisson, 267, 273
distribución de Rayleigh, 268
diferencia entre proporciones, 391
distribución exponencial, 262, 265, 272
distribución gamma, 270
distribución lognormal, 267
distribución normal, 255, 261, 273
distribución simétrica, 259
distribución uniforme, 276
distribución de Weibull, 274
en la regresión, 512
estimación agrupada, 377
efectos del ANOVA, 446, 465, 467
estimador de Hodges-Lehmann, 279
estimador *M*, 277
insesgada, 257
insesgada de varianza mínima, 260
método de mínimos cuadrados, 506-507
método de máxima verosimilitud, 272
media, 255, 259, 261
media de la variable dependiente en la regresión, 509
método de momentos, 270
ordenada al origen de la regresión, 507
parámetros de la regresión, 507, 592
pendiente de la recta de regresión, 507
principio de invarianza, 275
proporción, 255, 257
robusta, 262
tamaño de la población, 276
técnica de respuesta aleatorizada, 269
varianza, 256, 259, 377
Estimación puntual, 254
Estimación sesgada, 257, 261
Estimación, intervalo. *Véase* Intervalo de confianza
Estimación, puntual. *Véase* Estimación puntual
Estimador combinado, 377
Estimador consistente, 278
Estimador de Hodges-Lehmann, 279
Estimador de momento, 270
Estimador insesgado de varianza mínima, 260
Estimador insesgado, 257
Estimador puntual, 254
consistente, 278
error estándar, 263
error estándar estimado, 263
inicialización (bootstrap), 264
insesgado, 257
insesgado de varianza mínima, 260
sesgado, 257
máxima verosimilitud, 272
método de momentos, 270
robusto, 262
Estimador robusto, 262
Estimador, puntual. *Véase* Estimador puntual
Estudio analítico, 7
Estudio enumerativo, 7
Estudio observacional, 364
Estudio retrospectivo, 364
Evento, 54
complemento, 55
compuesto, 54
dependiente, 86
disjunto, 56
exhaustivo, 81
independiente, 86, 89
intersección, 55, 56
mutuamente excluyente, 56
simple, 54
unión, 55, 56
Evento simple, 54
Eventos exhaustivos, 81
Eventos mutuamente excluyentes, 56
Experimento binomial, 120
Experimento controlado aleatorizado, 365
Experimento de captura-recaptura, 276
Experimento de simulación, 231
Experimento doble ciego, 395
Experimento factorial, 476
Experimento multinomial, 213, 634
Experimento por bloques aleatorizado, 449
Experimento, 9, 53
binomial, 120
controlado aleatorizado, 365
de bloques aleatorizado, 449
doble ciego, 395
Extrapolación, peligro de, 509
Extremos densos, 193
Extremos ligeros, 193

Factor de corrección de población finita, 131
Factor de correlación, 418
Factor, 410
Familia de distribuciones de probabilidad, 104
Frecuencia acumulada, 27
Frecuencia relativa limitante, 59
Frecuencia relativa y probabilidad, 59-60
Frecuencia relativa, 15
Frecuencia, 15
  acumulada, 27
  relativa, 15
Frecuencia, relativa limitante, 60
Función convexa, 203
Función de densidad de probabilidad marginal, 209, 218
Función de probabilidad, 147
  condicional, 215
  conjunta, 208, 212
  marginal, 209
Función de densidad de probabilidad, 208, 212, 218
Función de densidad, probabilidad, 147
Función de distribución acumulada, 106, 152
Función de distribución. *Véase* Función de distribución acumulada
Función de probabilidad, 272
Función de proporción de fracasos, 203
Función de variables aleatorias
  distribución de, 202
  valor esperado de, 114-115, 157, 219, 252
  varianza de, 117-118, 158
Función escalón, 106
Función gamma incompleta, 176, 742
Función gamma, 174
  incompleta, 176
  tabla de, 176
Función intrínsecamente lineal, 564-565
Función masa de probabilidad, 102
  condicional, 215
  conjunta, 206, 212
  marginal, 207
Función masa de probabilidad conjunta, 206, 212
Función masa de probabilidad marginal, 207

Function logit, 572

Galton, Francis, 516
Grados de libertad
  ANOVA de un solo factor, 419
  ANOVA de varios factores, 446, 458, 467, 471
  con dos muestras, 373
  con una sola muestra, 39, 300
  distribución chi-cuadrada, 179
  distribución de intervalo estudentizado, 422
  distribución *F*, 399, 416
  distribución *t*, 300
  prueba de homogeneidad, 657
  prueba de independencia, 660
  pruebas de bondad del ajuste, 636, 644, 646
  regresión, 512, 580, 594
Gráfica de control promedio móvil ponderada exponencialmente, 734
Gráfica de línea, 103
Gráfica de probabilidad, 188, 190
Gráfica de series de tiempo, 14
Gráfica residual ajustada, 595
Gráfica seminormal, 198
Gráficas de control, 667
  características de desempeño, 704-706
  CUSUM, 718-726
  datos de atributos, 713, 715, 717, 734
  localización, 699, 702, 703
  promedio móvil ponderado exponencialmente, 734
  robustas, 706-707
  reglas complementarias, 706
  variación 709, 711, 712
Gran media, 414
Gran total, 418

Hipótesis alternativa, 316
Hipótesis compuesta, 643
Hipótesis estadística, 316
Hipótesis nula, 316
Hipótesis simple, 642
Histograma bimodal, 21
Histograma de probabilidad, 104
Histograma multimodal, 21
Histograma sesgado, 21
Histograma simétrico, 21
Histograma unimodal, 21
Histograma, 15
  bimodal, 21
  con sesgo negativo, 21
  con sesgo positivo, 21
  multimodal, 21
  para datos continuos, 18-19
  para datos discretos, 16
  probabilidad, 104
  simétrico, 21
  unimodal, 21
Homogeneidad, prueba para, 657

Independencia mutua, 89
Independencia por pares, 91
Independencia, prueba para, 660
Indicador variable, 590
Inicialización intervalo de confianza, 289-290, 306
  estimación del error estándar, 264

Inspección de rectificación, 732
Interacción generalizada, 483
Interacción, 589
  de dos factores, 459-457, 466
  de tres factores, 466
  generalizada, 483
Interpretación objetiva de probabilidad, 60
Intersección de sucesos, 55, 56
Intervalo de confianza. *Véase* Intervalo de confianza
de predicción. *Véase* Intervalo de predicción
aleatorio, 284
Intervalo aleatorio, 284
Intervalo de clase (clase), 17
Intervalo de confianza de la suma de rangos de Wilcoxon, 687, 761
Intervalo de confianza de los rangos con signo de Wilcoxon, 686, 760
Intervalo de confianza de los signos, 694
Intervalo de confianza de muestra grande, 292, 294, 368
Intervalo de confianza *t*, 302, 373, 377, 385, 426
Intervalo de confianza *z*, 286, 292, 294, 368, 396
Intervalo de confianza, 281
  coeficiente de correlación, 546
  coeficiente de regresión, 581, 598
  combinación lineal de medias, 426
  con muestras grandes, 292, 294, 368
  contraste, 440
  desviación estándar, 310
  determinación del tamaño de muestra, 287, 296, 369
  diferencia entre medias, 368, 373, 377, 385
  diferencia entre proporciones, 396
  diferencia media con datos agrupados, 385
  distribución de Poisson, 299
  distribución exponencial, 288-289
  distribución uniforme, 313
  inicialización, 289-290
  límite, 288
  media, 286, 292, 302, 313
  media de la variable independiente en la regresión, 532, 582, 598
  mediana, 313, 694
  obtención de, 288
  ordenada al origen de la recta de regresión, 552
  pendiente de la recta de regresión, 523
  proporción, 295
  rangos con signo de Wilcoxon, 686
  relación de desviaciones estándar, 402
  relación de proporciones, 398
  relación de varianzas, 402
  signo, 694
  simultáneo, 422, 434, 440
  sin distribución, 684-688
  suma de rangos de Wilcoxon, 687
  unilateral, 297
  varianza, 310
Intervalo de predicción, 304, 535, 582, 598
Intervalo de tolerancia, 305
Intervalo intercuartílico, 706
Intervalo, 36
Intervalos de confianza de Scheffé, 440
Intervalos de confianza simultáneos
  ANOVA, 422, 434, 440
  regresión, 534
Intervalos de confianza simultáneos de Bonferroni, 534

K-tupla, 69
Ley de probabilidad total, 81
Leyes de DeMorgan, 58
Límite de confianza inferior, 297, 302
Límite de confianza superior, 297, 302
Límite de confianza, 297, 302
Límite en el error de estimación, 264
Límites de control, 697
  recálculo, 704
LOWESS, 570

Marco de muestreo, 7
Marco, muestreo, 7
Máscara V, 718
Media recortada, 32
Media(s)
  distribución de muestra, 237-239
  estimación puntual. *Véase* Estimación puntual
  gran, 414
  intervalo de confianza. *Véase* Intervalo de confianza
  variable aleatoria continua. *Véase* Valor esperado
  variable aleatoria discreta. *Véase* Valor esperado
Mediana
  distribución continua, 156
  estimación puntual. *Véase* Estimación puntual
  intervalo de confianza. *Véase* Intervalo de confianza
  poblacional, 31
  muestral, 30
  prueba de hipótesis. *Véase* Prueba de hipótesis
Medidas resistentes, 40
Método de comparaciones múltiples de Tukey, 423, 434
Método de Yates, 481
Método *T*. *Véase* Método de comparaciones múltiples de Tukey
Métodos de Taguchi, 697
Métodos no paramétricos. *Véase* Métodos sin distribución
Métodos sin distribución, 667
Mínimos cuadrados ponderados, 560
Mínimos cuadrados, principio de, 506, 577
MINITAB, 4
Moda, 49, 143, 201

Modelo aditivo, 444
Modelo cuadrático completo, 588
Modelo de regresión intrínsecamente lineal, 566
Modelo de regresión múltiple de primer orden, 588
Modelo de regresión múltiple de segundo orden, 588
Modelo de segundo orden completo, 588
Modelo probabilístico lineal, 500
Modelo probabilístico, 497
Momento poblacional, 269
  muestral, 269
Momentos, método de, 270
Muestra, 3
  de conveniencia, 8
  aleatoria simple, 8, 228
  estratificada, 8
Muestra aleatoria simple, 8
Muestra aleatoria, 228
Muestreo de aceptación, 727-733
  planes de muestreo doble, 730-733
  planes de muestreo simple, 727-730
Muestreo estratificado, 8
Muestreo restringido, 731
Multicolinealidad, 622-623

Nivel de calidad aceptable, 729
Nivel de confianza simultáneo, 426
Nivel de confianza, 281, 285, 286, 426
Nivel de predicción, 304
Nivel de significación observado, 346
Nivel de significación, 323
Niveles, 411
Nomograma de Kemp, 724-725
Notación factorial, 71
Número total inspeccionado promedio, 732

Observación preponderante, 620-622
Orden estándar, 478

Parámetro
  de escala, 195
  de forma, 196
  distribución de probabilidad, 104
  no centralidad, 430, 438
  ubicación, 195
Parámetro de escala, 195
Parámetro de forma, 196
Parámetro de localización, 195
Parámetro de no centralidad, 430, 438
Parámetro de pendiente en la regresión, 501
Pares correlacionados, 387
Pares de alias, 485
Pares ordenados, 67
Percentil
  distribución continua, 155
  distribución normal, 163, 167
  muestral, 32, 189
Permutación, 70
Población conceptual, 7
Población hipotética, 7
Porcentaje de recorte, 32
Porcentaje de tolerancia de artículos defectuosos en el lote, 729
Potencia, 336
Principio de relación de probabilidad, 354-355
Probabilidad a priori, 81
Probabilidad condicional, 76-77
  distribución, 215
Probabilidad a posteriori, 81
Probabilidad subjetiva, 61
Probabilidad total, ley de, 81
Probabilidad, 6, 52, 58
  a priori, 81
  a posteriori, 81
  axiomas, 58
  propiedad de complemento, 61
  condicional, 76-77
  histograma, 104
  interpretación, 59
  propiedad de intersección, 78-79
  ley de probabilidad total, 81
  propiedad de multiplicación, 78-79, 87, 89
  propiedad de unión, 62, 63
Problema del cumpleaños, 94
Procedimiento de prueba, 318
Procedimientos CUSUM, 718-726
Procedimientos *t* combinados, 377
Proceso de nacimiento, 279
Proceso, 7
Propiedad de falta de memoria, 178
Proporción de errores por experimento, 426
Proporción(es)
  estimación puntual. *Véase* Estimación puntual
  intervalo de confianza. *Véase* Intervalo de confianza
  muestral, 33
  poblacional, 34
  prueba de hipótesis. *Véase* Prueba de hipótesis
Prueba chi-cuadrada de bondad de ajuste, 636, 644, 648
Prueba de dos extremos, 327
Prueba de extremo inferior, 320
Prueba de extremo superior, 319
Prueba *F*
  ANOVA, 417, 447, 459, 462, 468
  igualdad de dos varianzas, 400
  regresión, 527, 596, 601
Prueba de falta de ajuste, 564
Prueba de Fisher-Irwin, 397

Prueba de Friedman, 691
Prueba de grupo, 141
Prueba de hipótesis con muestras grandes, 331, 339, 367
Prueba de hipótesis, 316
  alta de ajuste en la regresión, 564
  ANOVA de un solo factor, 417, 432
  ANOVA de dos factores, 447, 459, 462, 465
  ANOVA de tres factores, 468
  Ansari-Bradley, 695
  bondad de ajuste, 636, 644, 648
  coeficiente de correlación, 544, 545
  coeficiente de regresión, 581, 598
  con muestras grandes, 331, 339, 367
  cuadrado latino, 471-472
  de dos extremos, 327
  de extremo inferior, 320
  de extremo superior, 319
  desviación estándar, 358
  determinación del tamaño de muestra, 330, 336, 341, 366, 395
  diferencia entre medias, 362-363, 367, 373, 377, 383
  diferencia entre proporciones, 393
  diferencia media con datos por pares, 383
  distribución de Poisson, 357, 647
  distribución exponencial, 358
  distribución normal, 328, 333, 358, 362-363, 649, 651
  e intervalos de confianza, 684
  errores, 319
  estadístico de prueba, 318
  homogeneidad de varias poblaciones, 657
  igualdad de pendientes, 553
  igualdad de varianzas, 400
  independencia de factores categóricos, 660
  media, 328, 331, 333
  media de variable dependiente en la regresión, 534
  mediana, 694
  modelo reducido contra completo, 601
  nivel de significación, 323
  pasos, 328
  pendiente de la recta de regresión, 526
  percentil normal, 358
  potencia de una prueba, 336
  probabilidades de error tipo II, 330, 334-336, 341, 365, 377, 394, 430-432
  procedimiento de prueba, 318
  proporción, 340, 342
  prueba de Fisher-Irwin, 397
  prueba de Friedman, 691
  Prueba *F* ANOVA, 417, 432
  prueba de Kruskal-Wallis, 689-690
  prueba de la gráfica de probabilidad para normalidad, 651
  prueba de McNemar, 408
  región de rechazo, 318
  prueba de relación de verosimilitudes, 354-355
  prueba de Wilcoxon de la suma de rangos, 679, 681
  prueba de Wilcoxon de los rangos con signo, 671, 674
  prueba del signo, 694
  Siegel-Tukey, 695
  sin distribución, 355, 668-682, 689-692
  tendencia, 677
  utilidad del modelo, 526, 596
  Valor *P*, 346, 347
  varianza, 358
Prueba de hipótesis. *Véase* Prueba de hipótesis
Prueba de Kruskal-Wallis, 689-690
Prueba de la suma de rangos de Wilcoxon, 679, 681, 759
Prueba de los rangos con signo de Wilcoxon, 671, 674, 758
Prueba de los signos, 694
Prueba de Mann-Whitney, 677
Prueba de McNemar, 408
Prueba de nivel a, 323
Prueba de normalidad de Ryan-Joiner, 651, 757
Prueba de un extremo. *Véase* Prueba de extremo superior; prueba de extremo inferior
Prueba de utilidad del modelo
  regresión múltiple, 596
  regresión lineal simple, 526
Prueba *t*, 333, 373, 377, 383
Prueba *z*, 328, 331, 340, 362-363, 367, 393
Pruebas de bondad de ajuste, 634-651

Rangos con signo, 669
Rangos, 669
Recorrido medio, 50
Recta de mínimos cuadrados, 506
Recta de regresión poblacional, 500
Recta de regresión verdadera, 500
Región de rechazo, 318
  de dos extremos, 327
  de extremo inferior, 320
  de extremo superior, 319
Regla de multiplicación para probabilidades, 78-79, 87, 89
Regla del producto
  k-tuplas, 69
  pares ordenados, 67-68
Regla empírica, 167
Reglas de conteo, 67-72
Regresión cuadrática, 577
Regresión cúbica, 577
Regresión de función de potencia, 566
Regresión exponencial, 566
Regresión lineal simple, 500-536
Regresión logística, 573
Regresión MAD, 561
Regresión múltiple, 587-632
  ecuación de modelo, 587
  estandarización de variables, 614
  estimación puntual, 592

idoneidad del modelo, 602-603
intervalo de confianza, 598
intervalo de predicción, 598
modelos, 588-592
multicolinealidad, 622-623
observación preponderante, 620-622
prueba de hipótesis, 596, 598, 601
prueba de utilidad del modelo, 596
selección de variable, 615-620
transformaciones, 612-614
Regresión no lineal, 564-574
Regresión polinomial, 576-583
centrado de valores de x, 582-583
estimación puntual, 577-578
intervalo de confianza, 581-582
intervalo de predicción, 582
modelo, 577
prueba de hipótesis, 581
variable independiente estandarizada, 586
Regresión por pasos, 619-620
Regresión robusta, 561
Regresión, 497
análisis residual, 556-561
coeficientes, 587, 595
cuadrática, 577
cúbica, 577
de potencia, 566
efecto, 516
exponencial, 566
función, 587
idoneidad del modelo, 556-561
intrínsecamente lineal, 566
lineal simple, 500-536
logística, 572
LOWESS, 570
multicolinealidad, 622-623
múltiple, 587-632
no lineal, 564-574
origen de la terminología, 516
observación preponderante, 620-622
polinomial, 576-583
por el origen, 519
recta de, 500, 506
robusta, 561
selección de variable, 615-620
suma de cuadrados, 516
transformaciones, 564-570, 612-614
y ANOVA, 527
Relación de probabilidades, 573
Relación determinística, 496
Relación lineal, medida de, 223
Relación t, 526
Replicación fraccional, 484-488
Residuo estandarizado, 557
Residuo, 510, 580, 594
estandarizado, 557
Respuesta aleatorizada, 269
Resultados equiprobables, 64
Resumen de cinco números, 41

S-PLUS, 4
SAS, 4
Selección de variable, 615-620
Selección hacia delante, 619
Semirreplicación, 484
Series de tiempo, 50, 548
Sesgo, 257
Sesgo negativo, 21
Sesgo positivo, 21
Significación
práctica, 354
estadística, 353
Significación estadística, 353
Significación práctica, 354
Significativo, 347
Simetría en el conjunto de datos, 51
Sistema K de n, 140
Suavizamiento exponencial, 51
Subrayado en comparaciones múltiples, 422-424
Suceso compuesto, 54
Sucesos dependientes, 86
Sucesos disjuntos, 56
Sucesos independientes, 86
Suma de cuadrados
error. *Véase* Suma de cuadrados del error
interacción. *Véase* Suma de cuadrados de la interacción
total. *Véase* Suma total de cuadrados
tratamiento. *Véase* Suma de cuadrados del tratamiento
Suma de cuadrados de la interacción, 458, 467
Suma de cuadrados de tratamiento, 418, 432
Suma de cuadrados residual, 512
Suma de errores de cuadrados, 418, 432, 512, 580
Suma total de cuadrados, 418, 432, 514

Tabla ANOVA, 419
Tabla de contingencia, 655
Tabla de probabilidad conjunta, 206
Tabla t, 301
Tabla z, 162
Tallos repetidos, 14
Tasa de error por familia, 426
Teorema de Bayes, 81-82
Teorema del binomio, 217
Teorema del límite central, 239-242
Total, muestra, 237
Transformación de Fisher, 545
Transformación para estabilizar la varianza, 434
Transformaciones

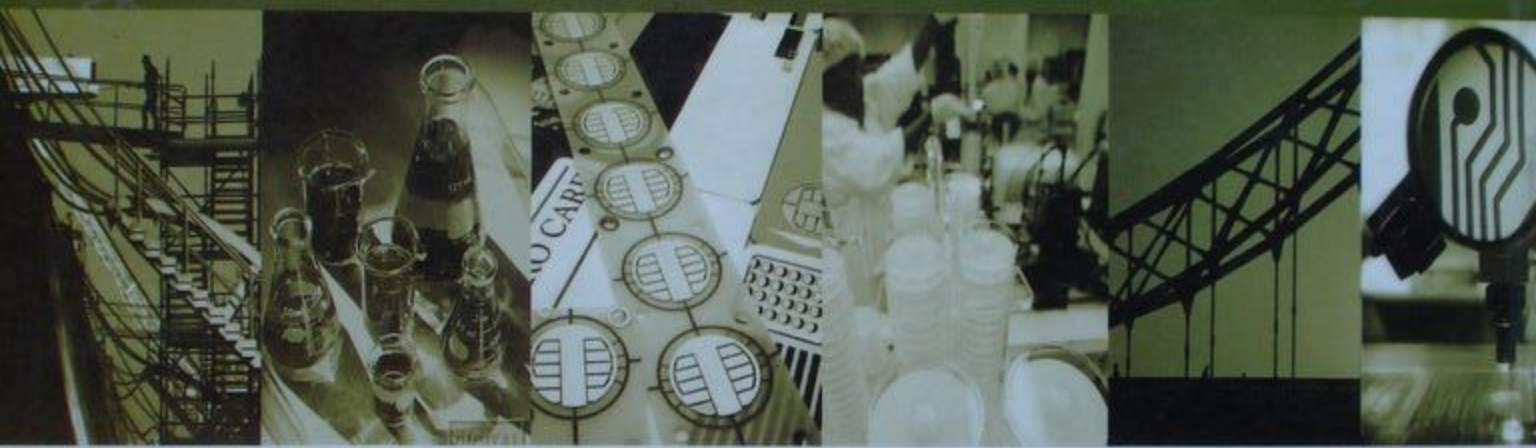
de datos, 26
en ANOVA, 434
en gráficas de control, 717
en la regresión, 564-570, 612-614
Tratamiento, 441

Unión de sucesos, 55, 56

Valor ajustado, 510, 594
Valor atípico extremo, 42
Valor atípico moderado, 42
Valor atípico, 30, 42
extremo, 42
moderado, 42
Valor crítico chi-cuadrado, 309
F, 399
intervalo estudentizado, 422
t, 301
tolerancia, 305
z, 164-165
Valor crítico de tolerancia, 305, 744
Valor crítico *t*, 301
Valor crítico *z*, 164-165
Valor esperado
combinación lineal, 244
de una diferencia, 245
función de variable aleatoria, 114-115, 157, 219, 252
media muestral, 237
reglas, 115-116
total de la muestra, 237
variable aleatoria continua, 157
variable aleatoria discreta, 112
Valor medio, 112, 157, 219
Valor nulo, 317
Valor *P*, 346, 347
prueba chi-cuadrada, 638
Prueba *F*, 401
prueba *t*, 349-350
prueba *z*, 348
Valor predicho, 510, 594
Variable aleatoria binomial, 122
Variable aleatoria de Bernoulli, 99
Variable aleatoria estandarizada, 165
Variable continua, 15
Variable de respuesta, 498
Variable dependiente, 498
Variable discreta, 15
Variable explicativa, 498
Variable ficticia, 590
Variable independiente, 498
Variable predictora, 498
Variable(s) aleatoria(s) continua(s), 100, 146
distribución de probabilidad, 147
función de distribución acumulada, 152
valor esperado, 157
varianza, 158
Variable(s) aleatoria(s) discreta(s), 100
distribución de probabilidad, 101-102
función de distribución acumulada, 106
valor esperado, 112
varianza, 116
Variable(s) aleatoria(s), 98
de Bernoulli, 99
binomial, 122
continua, 100, 146
dependiente, 211
discreta, 100
valor esperado, 112
geométrica, 133
hipergeométrica, 129
independiente, 211, 214
valor medio, 112, 157
binomial negativa, 132
desviación estándar, 116, 158
no correlacionada, 223
varianza, 116, 158
Variable, 3
aleatoria, 98
continua, 15
dependiente, 498
de respuesta, 498
discreta, 15
explicativa, 498
independiente, 498
predictora, 498
Variables aleatorias dependientes, 211
Variables aleatorias independientes, 211, 214
Variables aleatorias no correlacionadas, 223
Variación, coeficiente de, 48
Varianza(s)
combinación lineal, 244
de una diferencia, 245
estimador combinado, 377
estimación puntual. *Véase* Estimación puntual
función de variables aleatorias, 117-118, 158
intervalo de confianza. *Véase* Intervalo de confianza
media muestral, 237
muestral, 37
poblacional, 38
prueba de hipótesis. *Véase* Prueba de hipótesis
reglas, 40, 118
total de la muestra, 237
variable aleatoria continua, 158
variable aleatoria discreta, 116

Libro que proporciona un panorama amplio de los modelos de probabilidad y los métodos estadísticos. Ahora, en la sexta edición, el autor presenta numerosos cambios sustanciales a favor de la actualización del libro, así como para mejorar los resultados del aprendizaje en el alumno. Además, conserva las características que han hecho de este título uno de los más eficaces para la enseñanza de esta disciplina:

- Numerosos ejemplos basados en casos reales y con los datos auténticos.
- Las actividades a desarrollar tienen diferentes niveles de dificultad.
- El autor hace hincapié en que el estudiante aprenda a interpretar y no únicamente mecanice.



THOMSON

MÉXICO Y AMÉRICA CENTRAL
Tel. 52(55) 1500-6000
Fax 52(55) 5281-2656
editor@thomsonlearning.com.mx
México, D.F., MÉXICO

AMÉRICA DEL SUR
thomson@thomsonlearning.com.ar
Buenos Aires, ARGENTINA

EL CARIBE
Tel. (787) 641-11-12
Fax (787) 641-11-19
San Juan, PUERTO RICO

PACTO ANDINO
Tel. (571) 340-9470
Fax (571) 340-9475
cliente@thomsonlearning.com.co
Bogotá, COLOMBIA

ESPAÑA
Tel. (3491) 446-3350
Fax (3491) 445-6218
clientes@paraninfo.es
Madrid, ESPAÑA

ISBN 970-686-457-1
9 789706 864574