

Capítulo 11

Regresión lineal simple y correlación

11.1 Introducción a la regresión lineal

En la práctica a menudo se requiere resolver problemas que implican conjuntos de variables de las cuales se sabe que tienen alguna relación inherente entre sí. Por ejemplo, en una situación industrial quizá se sepa que el contenido de alquitrán en el flujo de salida de un proceso químico está relacionado con la temperatura en la entrada. Podría ser de interés desarrollar un método de pronóstico, es decir, un procedimiento que permita estimar el contenido de alquitrán para varios niveles de temperatura de entrada a partir de información experimental. Desde luego, es muy probable que para muchos ejemplos concretos en los que la temperatura de entrada sea la misma, por ejemplo 130°C, el contenido de alquitrán de salida no sea el mismo. Esto es muy similar a lo que ocurre cuando se estudian varios automóviles con un motor del mismo volumen; no todos tienen el mismo rendimiento de combustible. No todas las casas ubicadas en la misma zona del país, con la misma superficie de construcción, se venden al mismo precio. El contenido de alquitrán, el rendimiento del combustible (en millas por galón) y el precio de las casas (en miles de dólares) son **variables dependientes** naturales o respuestas en los tres escenarios. La temperatura en la entrada, el volumen del motor (pies cúbicos) y los metros cuadrados de superficie de construcción son, respectivamente, **variables independientes** naturales o **regresores**. Una forma razonable de relación entre la **respuesta Y** y el regresor x es la relación lineal,

$$Y = \beta_0 + \beta_1 x,$$

en la que, por supuesto, β_0 es la **intersección** y β_1 es la **pendiente**. Esta relación se ilustra en la figura 11.1.

Si la relación es exacta y no contiene ningún componente aleatorio o probabilístico, entonces se trata de una relación **determinista** entre dos variables científicas. Sin embargo, en los ejemplos que se mencionaron, así como en muchos otros fenómenos científicos y de ingeniería, la relación no es determinista, es decir, una x dada no siempre produce el mismo valor de Y . Como resultado, los problemas importantes en este caso son de naturaleza probabilística, toda vez que la relación anterior no puede considerarse exacta. El concepto de **análisis de regresión** se refiere a encontrar la mejor relación entre Y y x

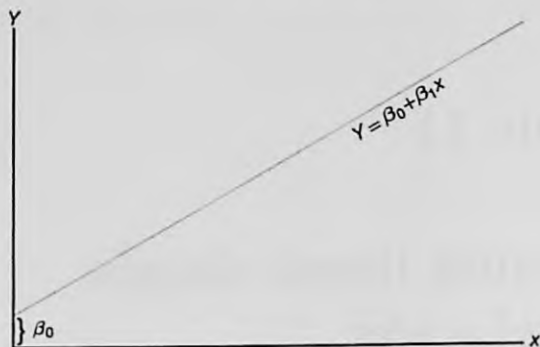


Figura 11.1: Una relación lineal; β_0 : intersección; β_1 : pendiente.

cuantificando la fuerza de esa relación, y empleando métodos que permitan predecir los valores de la respuesta dados los valores del regresor x .

En muchas aplicaciones habrá más de un regresor, es decir, más de una variable independiente que ayude a explicar a Y . Por ejemplo, si se tratara de explicar las razones para el precio de una casa, se esperaría que una de ellas fuera su antigüedad, en cuyo caso la estructura múltiple de la regresión se podría escribir como

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2,$$

donde Y es el precio, x_1 son los metros cuadrados y x_2 es la antigüedad de la casa en años. En el capítulo siguiente se estudiarán problemas con regresores múltiples. El análisis resultante se denomina **regresión múltiple**; en tanto que el análisis del caso con un solo regresor recibe el nombre de **regresión simple**. En un segundo ejemplo de la regresión múltiple, un ingeniero químico podría estar interesado en la cantidad de hidrógeno que se ha perdido en las muestras de un metal específico que se tiene almacenado. En este caso habría dos entradas, x_1 , el tiempo de almacenamiento en horas, y x_2 , la temperatura de almacenamiento en grados centígrados. De modo que la respuesta sería Y , la pérdida de hidrógeno en partes por millón.

En este capítulo estudiaremos el tema de la **regresión lineal simple**, que trata el caso de una sola variable regresora, en el que la relación entre x y y es lineal. Para el caso en el que hay más de una variable regresora el lector debe consultar el capítulo 12. Denotemos una muestra aleatoria de tamaño n mediante el conjunto $\{(x_i, y_i); i = 1, 2, \dots, n\}$. Si se tomaran muestras adicionales utilizando exactamente los mismos valores de x , se esperaría que los valores de y variarían. Así, el valor y_i en el par ordenado (x_i, y_i) es el valor de cierta variable aleatoria Y_i .

11.2 El modelo de regresión lineal simple (RLS)

Hemos limitado el uso del término *análisis de regresión* a los casos en los que las relaciones entre las variables no son deterministas, es decir, no son exactas. En otras palabras, debe existir un **componente aleatorio** en la ecuación que relaciona las variables. Este componente aleatorio toma en cuenta consideraciones que no son medibles o, de

hecho, que los científicos o los ingenieros no comprenden. En realidad, en la mayoría de aplicaciones de la regresión, la ecuación lineal, digamos, $Y = \beta_0 + \beta_1 x$ es una aproximación que representa de manera simplificada algo desconocido y mucho más complicado. Por ejemplo, en el caso que implica la respuesta $Y =$ contenido de alquitrán y $x =$ temperatura de entrada es probable que $Y = \beta_0 + \beta_1 x$ sea una aproximación razonable que podría funcionar dentro de un rango limitado de x . La mayoría de las veces los modelos que son simplificaciones de estructuras más complicadas y desconocidas son de naturaleza lineal, es decir, lineales en los **parámetros** β_0 y β_1 , o, en el caso del modelo que implica el precio, el tamaño y la antigüedad de la casa, lineal en los **parámetros** β_0 , β_1 y β_2 . Estas estructuras lineales son sencillas y de naturaleza empírica, por lo que se denominan **modelos empíricos**.

Un análisis de la relación entre x y Y requiere el planteamiento de un **modelo estadístico**. Con frecuencia un estadístico utiliza un modelo como representación de un **ideal** que, en esencia, define cómo percibimos que el sistema en cuestión generó los datos. El modelo debe incluir al conjunto $\{(x_i, y_i); i = 1, 2, \dots, n\}$ de datos que implica n pares de valores (x, y) . No debemos olvidar que el valor de y_i depende de x_i por medio de una estructura lineal que también incluye el componente aleatorio. La base para el uso de un modelo estadístico se relaciona con la manera en que la variable aleatoria Y cambia con x y el componente aleatorio. El modelo también incluye lo que se asume acerca de las propiedades estadísticas del componente aleatorio. A continuación se presenta el modelo estadístico para la regresión lineal simple. La respuesta Y se relaciona con la variable independiente x a través de la ecuación

Modelo de
regresión lineal
simple

$$Y = \beta_0 + \beta_1 x + \epsilon.$$

en la cual β_0 y β_1 son los parámetros desconocidos de la intersección y la pendiente, respectivamente, y ϵ es una variable aleatoria que se supone está distribuida con $E(\epsilon) = 0$ y $\text{Var}(\epsilon) = \sigma^2$. Es frecuente que a la cantidad σ^2 se le denomine varianza del error o varianza residual.

En el modelo anterior hay varias cuestiones evidentes. La cantidad Y es una variable aleatoria, ya que ϵ es aleatoria. El valor x de la variable regresora no es aleatorio y, de hecho, se mide con un error despreciable. La cantidad ϵ , que a menudo recibe el nombre de **error aleatorio** o **alteración aleatoria**, tiene varianza constante. Es común que a esta parte se le denomine **suposición de varianza homogénea**. La presencia de este error aleatorio ϵ evita que el modelo se convierta tan sólo en una ecuación determinista. Ahora, el hecho de que $E(\epsilon) = 0$ implica que para una x específica, los valores de y se distribuyen alrededor de la **recta verdadera** o **recta de regresión** de la población $y = \beta_0 + \beta_1 x$. Si se elige bien el modelo, es decir, si no hay otros regresores de importancia y la aproximación lineal es buena dentro de los rangos de los datos, entonces son razonables los errores positivos y negativos que rodean a la regresión verdadera. Debe recordarse que en la práctica β_0 y β_1 se desconocen y que deben estimarse a partir de los datos. Además, el modelo que se acaba de describir es de naturaleza conceptual. Como resultado, en la práctica nunca se observan los valores ϵ reales, por lo que nunca se puede trazar la verdadera recta de regresión, aunque suponemos que ahí está. Sólo es posible dibujar una recta estimada. En la figura 11.2 se ilustra la naturaleza de los datos (x, y) hipotéticos dispersos alrededor de la verdadera recta de regresión para un caso en que sólo se dispone de $n = 5$ observaciones. Debemos destacar que lo que observamos en la figura 11.2 no es la recta que utilizan el científico o ingeniero. En vez de esa recta, ¡lo

que describe la ilustración es el significado de las suposiciones! Ahora describiremos la regresión que el usuario tiene a su disposición.

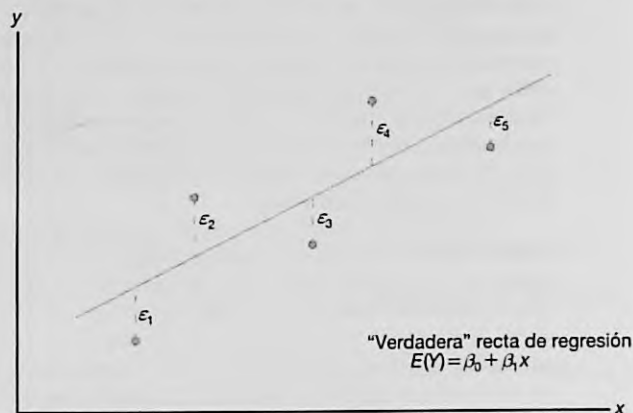


Figura 11.2: Datos (x, y) hipotéticos dispersos alrededor de la verdadera recta de regresión para $n = 5$.

La recta de regresión ajustada

Un aspecto importante del análisis de regresión es, en términos sencillos, estimar los parámetros β_0 y β_1 , es decir, estimar los llamados **coeficientes de regresión**. En la sección siguiente se estudiará el método para estimarlos. Suponga que denotamos los estimados b_0 para β_0 y b_1 para β_1 . Entonces, la recta de **regresión ajustada**, o estimada, es dada por

$$\hat{y} = b_0 + b_1 x,$$

donde \hat{y} es el valor pronosticado o ajustado. Es evidente que la recta ajustada es un estimado de la verdadera recta de regresión. Se espera que la recta ajustada esté más cerca de la verdadera línea de regresión cuando se dispone de una gran cantidad de datos. En el ejemplo siguiente se ilustra la recta ajustada para un estudio sobre contaminación en la vida real.

Uno de los problemas más desafiantes que enfrenta el campo del control de la contaminación del agua lo representa la industria de la peletería, ya que sus desechos son químicamente complejos; se caracterizan por valores elevados de la demanda de oxígeno químico, sólidos volátiles y otras medidas de contaminación. Considere los datos experimentales de la tabla 11.1, que se obtuvieron de 33 muestras de desechos tratados químicamente en un estudio realizado en Virginia Tech. Se registraron los valores de x , la reducción porcentual de los sólidos totales, y de y , el porcentaje de disminución de la demanda de oxígeno químico.

Los datos de la tabla 11.1 aparecen graficados en un **diagrama de dispersión** en la figura 11.3. Al inspeccionar dicho diagrama se observa que los puntos se acercan mucho a una línea recta, lo cual indica que la suposición de linealidad entre las dos variables parece ser razonable.

Tabla 11.1: Medidas de la reducción de los sólidos y de la demanda de oxígeno químico

Reducción de sólidos, x (%)	Reducción de la demanda de oxígeno, y (%)	Reducción de sólidos, x (%)	Reducción de la demanda de oxígeno, y (%)
3	5	36	34
7	11	37	36
11	21	38	38
15	16	39	37
18	16	39	36
27	28	39	45
29	27	40	39
30	25	41	41
30	35	42	40
31	30	42	44
31	40	43	37
32	32	44	44
33	34	45	46
33	32	46	46
34	34	47	49
36	37	50	51
36	38		

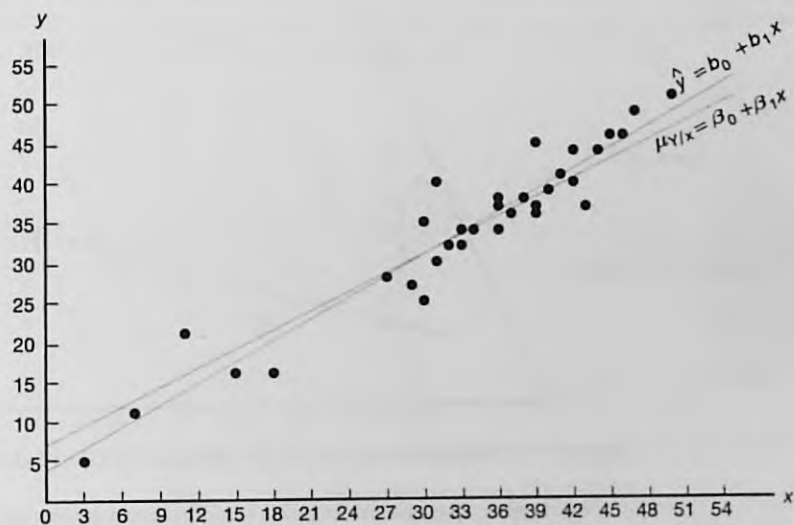


Figura 11.3: Diagrama de dispersión con rectas de regresión.

En el diagrama de dispersión de la figura 11.3 se ilustra la recta de regresión ajustada y una recta hipotética de regresión verdadera. Más adelante, en la sección 11.3, en la cual estudiaremos el método de estimación, revisaremos este ejemplo.

Otra mirada a las suposiciones del modelo

Resulta aleccionador repasar el modelo de regresión lineal simple que se presentó con anterioridad y analizar de forma gráfica la manera en que se relaciona con la denominada regresión verdadera. Daremos más detalles en la figura 11.2, cuando ilustremos no sólo el lugar en que los ϵ_i se localizan en la gráfica, sino también lo que implica la suposición de normalidad para los ϵ_i .

Suponga que tenemos una regresión lineal simple con $n = 6$, valores de x equidistantes y un valor único de y para cada x . Considere la gráfica de la figura 11.4, la cual debería proporcionar al lector una representación clara del modelo y de las suposiciones implicadas. La recta que aparece en la gráfica es la recta de regresión verdadera. Los puntos graficados (y, x) son puntos reales dispersos alrededor de la recta. Cada punto se ubica en su propia distribución normal, donde el centro de la distribución, es decir, la media de y , cae sobre la recta. Ciertamente esto es lo esperado, ya que $E(Y) = \beta_0 + \beta_1 x$. Como resultado, la verdadera recta de regresión **pasa a través de las medias de la respuesta** y las observaciones reales se encuentran sobre la distribución, alrededor de las medias. Observe también que todas las distribuciones tienen la misma varianza, que se denota con σ^2 . Desde luego, la desviación entre una y e individual y el punto sobre la recta será su valor individual ϵ . Esto queda claro porque

$$y_i - E(Y_i) = y_i - (\beta_0 + \beta_1 x_i) = \epsilon_i.$$

Así, con una x dada, tanto Y como el ϵ correspondiente tienen varianza σ^2 .

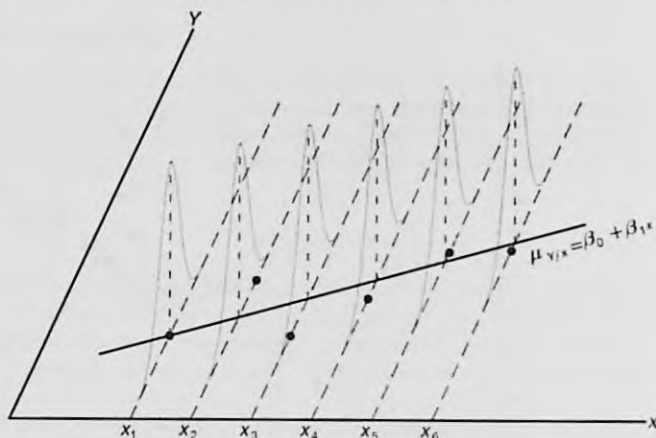


Figura 11.4: Observaciones individuales alrededor de la verdadera recta de regresión.

Note también que aquí escribimos la verdadera recta de regresión como $\mu_{y|x} = \beta_0 + \beta_1 x$ con el fin de reafirmar que la recta pasa a través de la media de la variable aleatoria Y .

11.3 Mínimos cuadrados y el modelo ajustado

En esta sección se estudia el método para ajustar una recta de regresión estimada a los datos, lo cual equivale a determinar los estimados b_0 para β_0 y b_1 para β_1 . Por supuesto,

esto permite el cálculo de los valores pronosticados a partir de la recta ajustada $\hat{y} = b_0 + b_1x$, y otros tipos de análisis y de información diagnóstica que determinarán la fuerza de la relación, así como la adecuación y el ajuste del modelo. Antes de analizar el método de estimación de los mínimos cuadrados es importante presentar el concepto de **residual**. En esencia, un residual es un error en el ajuste del modelo $\hat{y} = b_0 + b_1x$.

Residual: Error en el ajuste Dado un conjunto de datos de regresión $\{(x_i, y_i); i = 1, 2, \dots, n\}$ y un modelo ajustado $\hat{y}_i = b_0 + b_1x_i$, el i -ésimo residual e_i es dado por

$$e_i = y_i - \hat{y}_i, \quad i = 1, 2, \dots, n.$$

Es evidente que si un conjunto de n residuales es grande, entonces el ajuste del modelo no es bueno. Los residuales pequeños son indicadores de un ajuste adecuado. Otra relación interesante, y que a veces es útil, es la siguiente:

$$y_i = b_0 + b_1x_i + e_i.$$

El uso de la ecuación anterior debería aclarar la diferencia entre los residuales e_i y los errores del modelo conceptual ϵ_i . No debemos olvidar que, mientras que los ϵ_i no se observan, los e_i no sólo se observan sino que desempeñan un papel importante en el análisis total.

La figura 11.5 ilustra el ajuste de la recta a este conjunto de datos: a saber $\hat{y} = b_0 + b_1x$, y la recta que refleja el modelo $\mu_{y|x} = \beta_0 + \beta_1x$. Desde luego, β_0 y β_1 son parámetros desconocidos. La recta ajustada es un estimado de la recta que genera el modelo estadístico. Hay que tener presente que la recta $\mu_{y|x} = \beta_0 + \beta_1x$ es desconocida.

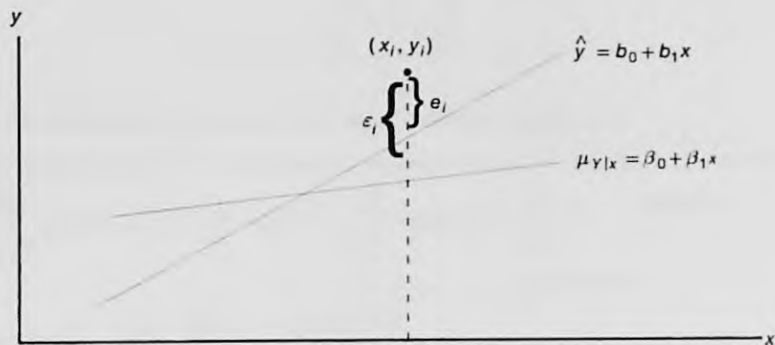


Figura 11.5: Comparación de ϵ_i con el residual e_i .

Método de mínimos cuadrados

Debemos calcular b_0 y b_1 , los estimados de β_0 y β_1 , de manera que la suma de los cuadrados de los residuales sea mínima. La suma residual de los cuadrados con frecuencia se denomina suma de los cuadrados del error respecto de la recta de regresión y se denota como *SCE*. Este procedimiento de minimización para estimar los parámetros

se denomina **método de mínimos cuadrados**. Por lo tanto, debemos calcular a y b para minimizar

$$SCE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2.$$

Al diferenciar la SCE con respecto a b_0 y b_1 , se obtiene

$$\frac{\partial(SCE)}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i), \quad \frac{\partial(SCE)}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) x_i.$$

Al igualar a cero las derivadas parciales y reacomodar los términos, obtenemos las ecuaciones siguientes (llamadas **ecuaciones normales**)

$$nb_0 + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i, \quad b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i,$$

que se resuelven simultáneamente para obtener fórmulas de cálculo para b_0 y b_1 .

Estimación de los coeficientes de regresión Dada la muestra $\{(x_i, y_i); i = 1, 2, \dots, n\}$, los estimados b_0 y b_1 de los mínimos cuadrados de los coeficientes de regresión β_0 y β_1 se calculan mediante las fórmulas

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} y$$

$$b_0 = \frac{\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i}{n} = \bar{y} - b_1 \bar{x}.$$

En el ejemplo siguiente se ilustra el cálculo de b_0 y b_1 usando los datos de la tabla 11.1.

Ejemplo 11.1: Estime la recta de regresión para los datos de contaminación de la tabla 11.1.

Solución:

$$\sum_{i=1}^{33} x_i = 1104, \quad \sum_{i=1}^{33} y_i = 1124, \quad \sum_{i=1}^{33} x_i y_i = 41,355, \quad \sum_{i=1}^{33} x_i^2 = 41,086$$

Por lo tanto,

$$b_1 = \frac{(33)(41,355) - (1104)(1124)}{(33)(41,086) - (1104)^2} = 0.903643 y$$

$$b_0 = \frac{1124 - (0.903643)(1104)}{33} = 3.829633.$$

Por consiguiente, la recta de regresión estimada es dada por

$$\hat{y} = 3.8296 + 0.9036x.$$

Si utilizáramos la recta de regresión del ejemplo 11.1, podríamos pronosticar una reducción de 31% en la demanda de oxígeno químico si los sólidos totales se redujeran

un 30%. La reducción de 31% en la demanda de oxígeno químico se puede interpretar como un estimado de la media de la población $\mu_{Y|30}$, o como un estimado de una observación nueva si la reducción de sólidos totales es de 30%. Sin embargo, dichas estimaciones están sujetas a error. Incluso si el experimento estuviera controlado para que la reducción de los sólidos totales fuera de 30%, es improbable que la reducción en la demanda de oxígeno químico que se midiera fuera exactamente igual a 31%. De hecho, los datos originales registrados en la tabla 11.1 indican que se registraron medidas de 25% y de 35% en la reducción de la demanda de oxígeno, cuando la disminución de los sólidos totales se mantuvo en 30%.

¿Qué es lo bueno de los mínimos cuadrados?

Debemos señalar que el criterio de los mínimos cuadrados está diseñado para brindar una recta ajustada que resulte en la “cercanía” entre la recta y los puntos graficados. Existen muchas formas de medir dicha cercanía. Por ejemplo, quizá deseáramos determinar los valores de b_0 y b_1 para los que se minimiza $\sum_{i=1}^n |y_i - \hat{y}_i|$ o para los que se minimiza $\sum_{i=1}^n |y_i - \hat{y}_i|^{1.5}$. Ambos métodos son viables y razonables. Observe que los dos, así como el procedimiento de mínimos cuadrados, obligan a que los residuales sean “pequeños” en cierto sentido. Debemos recordar que los residuales son el equivalente empírico de los valores de ϵ . La figura 11.6 ilustra un conjunto de residuales. Observe que la línea ajustada tiene valores predichos como puntos sobre la recta y, en consecuencia, los residuales son desviaciones verticales desde los puntos hasta la recta. Como resultado, el procedimiento de mínimos cuadrados genera una recta que **minimiza la suma de los cuadrados de las desviaciones verticales** desde los puntos hasta la recta.

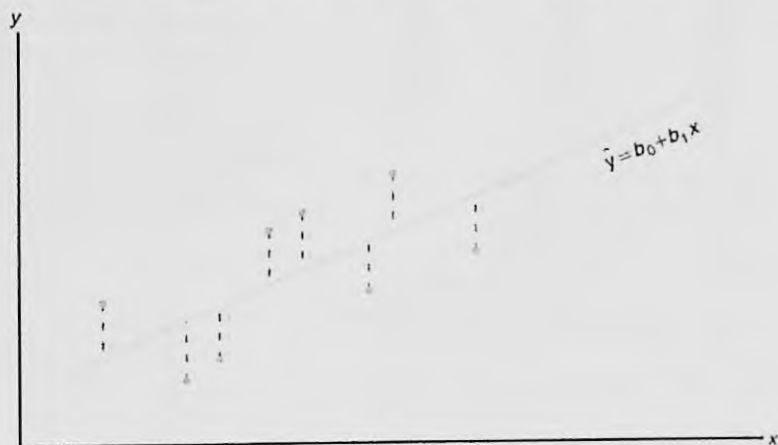


Figura 11.6: Los residuales como desviaciones verticales.

Ejercicios

11.1 Se realizó un estudio en Virginia Tech para determinar si ciertas medidas de la fuerza estática del brazo influyen en las características de "levantamiento dinámico" de un individuo. Veinticinco individuos se sometieron a pruebas de fuerza y luego se les pidió que hicieran una prueba de levantamiento de peso, en el que el peso se elevaba en forma dinámica por encima de la cabeza. A continuación se presentan los datos.

Individual	Fuerza del brazo, x	Levantamiento dinámico, y
1	17.3	71.7
2	19.3	48.3
3	19.5	88.3
4	19.7	75.0
5	22.9	91.7
6	23.1	100.0
7	26.4	73.3
8	26.8	65.0
9	27.6	75.0
10	28.1	88.3
11	28.2	68.3
12	28.7	96.7
13	29.0	76.7
14	29.6	78.3
15	29.9	60.0
16	29.9	71.7
17	30.3	85.0
18	31.3	85.0
19	36.0	88.3
20	39.5	100.0
21	40.4	100.0
22	44.3	100.0
23	44.6	91.7
24	50.4	100.0
25	55.9	71.7

- a) Estime los valores de β_0 y β_1 para la curva de regresión lineal $\mu_{y|x} = \beta_0 + \beta_1 x$.
 b) Calcule un estimado puntual de $\mu_{y|30}$.
 c) Grafique los residuales en comparación con las x (fuerza del brazo). Comente los resultados.

11.2 Las siguientes son las calificaciones de un grupo de 9 estudiantes en un informe de medio semestre (x) y en el examen final (y):

x	77	50	71	72	81	94	96	99	67
y	82	66	78	34	47	85	99	99	68

- a) Estime la recta de regresión lineal.
 b) Calcule la calificación final de un estudiante que obtuvo 85 de calificación en el informe de medio semestre.

11.3 Se registraron las cantidades de un compuesto químico y que se disuelve en 100 gramos de agua a distintas temperaturas x :

x ($^{\circ}\text{C}$)	y (gramos)		
0	8	6	8
15	12	10	14
30	25	21	24
45	31	33	28
60	44	39	42
75	48	51	44

- a) Calcule la ecuación de la recta de regresión.
 b) Grafique la recta en un diagrama de dispersión.
 c) Estime la cantidad de producto químico que se disolverá en 100 gramos de agua a 50°C .
- 11.4 Para fines de calibración se recabaron los siguientes datos, los cuales permitirían determinar la relación entre la presión y la lectura correspondiente en la escala.

Presión, x (lb/pulg ²)	Lectura en la escala, y
10	13
10	18
10	16
10	15
10	20
50	86
50	90
50	88
50	88
50	92

- a) Calcule la ecuación de la recta de regresión.
 b) En esta aplicación el propósito de la calibración es estimar la presión a partir de una lectura observada en la escala. Estime la presión para una lectura en la escala de 54, usando $\hat{x} = (54 - b_0)/b_1$.
- 11.5 Se realizó un estudio sobre la cantidad de azúcar convertida en cierto proceso a distintas temperaturas. Los datos se codificaron y registraron como sigue:

Temperatura, x	Azúcar convertida, y
1.0	8.1
1.1	7.8
1.2	8.5
1.3	9.8
1.4	9.5
1.5	8.9
1.6	8.6
1.7	10.2
1.8	9.3
1.9	9.2
2.0	10.5

- a) Estime la recta de regresión lineal.
 b) Calcule la cantidad media de azúcar convertida que se produce cuando se registra una temperatura codificada de 1.75.
 c) Grafique los residuales en comparación con la temperatura. Comente sus resultados.

11.6 En cierto tipo de espécimen de prueba metálico se sabe que la tensión normal sobre un espécimen se relaciona funcionalmente con la resistencia al corte. El siguiente es un conjunto de datos experimentales codificados para las dos variables:

Tensión normal, x	Resistencia al corte, y
26.8	26.5
25.4	27.3
28.9	24.2
23.6	27.1
27.7	23.6
23.9	25.9
24.7	26.3
28.1	22.5
26.9	21.7
27.4	21.4
22.6	25.8
25.6	24.9

- a) Estime la recta de regresión $\mu_{y|x} = \beta_0 + \beta_1 x$.
 b) Estime la resistencia al corte para una tensión normal de 24.5.

11.7 Los siguientes son algunos de los datos contenidos en un conjunto clásico denominado "datos piloto de graficación" que aparecen en *Fitting Equations to Data*, de Daniel y Wood, publicado en 1971. La respuesta y es el contenido de ácido del material determinado por análisis volumétrico; mientras que el regresor x es el contenido de ácido orgánico determinado por extracción y ponderación.

y	x	y	x
76	123	70	109
62	55	37	48
66	100	82	138
58	75	88	164
88	159	43	28

- a) Grafique los datos; ¿la regresión lineal simple parece un modelo adecuado?
 b) Haga un ajuste de regresión lineal simple; calcule la pendiente y la intersección.
 c) Grafique la recta de regresión en la gráfica del inciso a.

11.8 Se aplica un examen de colocación de matemáticas a todos los estudiantes de nuevo ingreso en una universidad pequeña. Se negará la inscripción al curso regular de matemáticas a los estudiantes que obtengan menos de 35 puntos y se les enviará a clases de regularización. Se registraron los resultados del examen de colocación y las calificaciones finales de 20 estudiantes que tomaron el curso regular:

- a) Elabore un diagrama de dispersión.
 b) Calcule la ecuación de la recta de regresión para predecir las calificaciones en el curso a partir de las del examen de colocación.
 c) Grafique la recta en el diagrama de dispersión.

d) Si la calificación aprobatoria mínima fuera 60 puntos, ¿qué calificación en el examen de colocación se debería usar en el futuro como criterio para negar a los estudiantes el derecho de admisión a ese curso?

Examen de colocación	Calificación en el curso
50	53
35	41
35	61
40	56
55	68
65	36
35	11
60	70
90	79
35	59
90	54
80	91
60	48
60	71
60	71
40	47
55	53
50	68
65	57
50	79

11.9 Un comerciante minorista realizó un estudio para determinar la relación que hay entre los gastos semanales de publicidad y las ventas.

Costos de publicidad (\$)	Ventas (\$)
40	385
20	400
25	395
20	365
30	475
50	440
40	490
20	420
50	560
40	525
25	480
50	510

- a) Elabore un diagrama de dispersión.
 b) Calcule la ecuación de la recta de regresión para pronosticar las ventas semanales a partir de los gastos de publicidad.
 c) Estime las ventas semanales si los costos de publicidad son de \$35.
 d) Grafique los residuales en comparación con los costos de publicidad. Comente sus resultados.

11.10 Los siguientes datos son los precios de venta z de cierta marca y modelo de automóvil usado con w años de antigüedad. Ajuste una curva de la forma $\mu_{z|w} = \gamma \delta^w$ mediante la ecuación de regresión muestral no lineal $\hat{z} = cd^w$ [Sugerencia: Escriba $\ln \hat{z} = \ln c + (\ln d)w = b_0 + b_1 w$].

w (años)	z (dólares)	w (años)	z (dólares)
1	6350	3	5395
2	5695	5	4985
2	5750	5	4895

11.11 La fuerza de impulso de un motor (y) es una función de la temperatura de escape (x) en $^{\circ}\text{F}$ cuando otras variables de importancia se mantienen constantes. Considere los siguientes datos.

y	x	y	x
4300	1760	4010	1665
4650	1652	3810	1550
3200	1485	4500	1700
3150	1390	3008	1270
4950	1820		

- Grafique los datos.
- Ajuste una recta de regresión simple a los datos y grafíquela a través de ellos.

11.12 Se realizó un estudio para analizar el efecto de la temperatura ambiente x sobre la energía eléctrica consumida por una planta química y . Otros factores se mantuvieron constantes y se recabaron los datos de una planta piloto experimental.

y (BTU)	x ($^{\circ}\text{F}$)	y (BTU)	x ($^{\circ}\text{F}$)
250	27	265	31
285	45	298	60
320	72	267	34
295	58	321	74

- Grafique los datos.
- Estime la pendiente y la intersección en un modelo de regresión lineal simple.
- Pronostique el consumo de energía para una temperatura ambiente de 65°F .

11.4 Propiedades de los estimadores de mínimos cuadrados

Además de los supuestos de que el término del error en el modelo

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

es una variable aleatoria con media igual a cero y varianza σ^2 constante, suponga que además damos por hecho que $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ son independientes de una corrida a otra del experimento, lo cual proporciona la base para calcular las medias y varianzas de los estimadores de β_0 y β_1 .

Es importante recordar que nuestros valores de b_0 y b_1 , basados en una muestra dada de n observaciones, sólo son estimaciones de los parámetros verdaderos β_0 y β_1 . Si el experimento se repitiera una y otra vez, usando en cada ocasión los mismos valores fijos de x , los estimados resultantes de β_0 y β_1 muy probablemente diferirían de un experimento a otro. Estos estimados distintos podrían ser considerados como valores adoptados por las variables aleatorias B_0 y B_1 ; en tanto que b_0 y b_1 son ejecuciones específicas.

Como los valores de x permanecen fijos, los valores de B_0 y B_1 dependen de las variaciones en los valores de y o, con más precisión, en los valores de las variables aleatorias

11.13 Un estudio sobre la cantidad de lluvia y la de contaminación del aire eliminada produjo los siguientes datos:

Cantidad de lluvia diaria, x (0.01 cm)	Partículas eliminadas, y ($\mu\text{g}/\text{m}^3$)
4.3	126
4.5	121
5.9	116
5.6	118
6.1	114
5.2	118
3.8	132
2.1	141
7.5	108

- Calcule la ecuación de la recta de regresión para predecir las partículas eliminadas de la cantidad de precipitación diaria.
- Estime la cantidad de partículas eliminadas si la precipitación diaria es $x = 4.8$ unidades.

11.14 Un profesor de la Escuela de Negocios de una universidad encuestó a una docena de colegas acerca del número de reuniones profesionales a que acudieron en los últimos cinco años (x) y el número de trabajos que enviaron a revistas especializadas (y) durante el mismo periodo. A continuación se presenta el resumen de los datos:

$$n = 12, \quad \bar{x} = 4, \quad \bar{y} = 12,$$

$$\sum_{i=1}^n x_i^2 = 232, \quad \sum_{i=1}^n x_i y_i = 318.$$

Ajuste un modelo de regresión lineal simple entre x y y calculando los estimados de la intersección y la pendiente. Comente si la asistencia a más reuniones profesionales da como resultado más publicaciones de artículos.

Y_1, Y_2, \dots, Y_n . Las suposiciones sobre la distribución implican que las Y_i , $i = 1, 2, \dots, n$ también están distribuidas de manera independiente, con media $\mu_{Y|x_i} = \beta_0 + \beta_1 x_i$ y varianzas σ^2 iguales, es decir,

$$\sigma_{Y|x_i}^2 = \sigma^2 \quad \text{para } i = 1, 2, \dots, n.$$

Media y varianza de los estimadores

En la exposición que sigue mostramos que el estimador B_1 es insesgado para β_1 , y se demuestran tanto las varianzas de B_0 como las de B_1 . Esto inicia una serie de procedimientos que conducen a la prueba de hipótesis y a la estimación de intervalos de confianza para la intersección y la pendiente.

Como el estimador

$$B_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

es de la forma $\sum_{i=1}^n c_i Y_i$,

$$c_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad i = 1, 2, \dots, n,$$

podemos concluir a partir del teorema 7.11 que B_1 tiene una distribución $n(\mu_{B_1}, \sigma_{B_1})$ con

$$\mu_{B_1} = \frac{\sum_{i=1}^n (x_i - \bar{x})(\beta_0 + \beta_1 x_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} = \beta_1 \quad \text{y} \quad \sigma_{B_1}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sigma_{Y_i}^2}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

También se puede demostrar (véase el ejercicio de repaso 11.60 de la página 438) que la variable aleatoria B_0 se distribuye normalmente con

$$\text{media } \mu_{B_0} = \beta_0 \quad \text{y} \quad \text{varianza } \sigma_{B_0}^2 = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \sigma^2.$$

A partir de estos resultados es evidente que los estimadores de mínimos cuadrados tanto para β_0 como para β_1 son insesgados.

Partición de la variabilidad total y estimación de σ^2

Para hacer inferencias sobre β_0 y β_1 es necesario llegar a una estimación del parámetro σ^2 que aparece en las dos fórmulas anteriores de la varianza de B_0 y B_1 . El parámetro σ^2 , el modelo de la varianza del error, refleja una variación aleatoria o una variación del

error experimental alrededor de la recta de regresión. En gran parte de lo que sigue se recomienda emplear la notación

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2, \quad S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2, \quad S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

De manera que la suma de los cuadrados del error se puede escribir como sigue:

$$\begin{aligned} SCE &= \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 = \sum_{i=1}^n [(y_i - \bar{y}) - b_1(x_i - \bar{x})]^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - 2b_1 \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) + b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= S_{yy} - 2b_1 S_{xy} + b_1^2 S_{xx} = S_{yy} - b_1 S_{xy}, \end{aligned}$$

que es el paso final que surge del hecho de que $b_1 = S_{xy}/S_{xx}$.

Teorema 11.1: Un estimador insesgado de σ^2 es

$$s^2 = \frac{SCE}{n-2} = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2}.$$

La prueba del teorema 11.1 se deja como ejercicio (véase el ejercicio de repaso 11.59).

El estimador de σ^2 como error cuadrado medio

Para darnos una idea del estimador de σ^2 deberíamos observar el resultado del teorema 11.1. El parámetro σ^2 mide la varianza o las desviaciones cuadradas entre los valores de Y y su media, dada por $\mu_{Y,x}$, es decir, las desviaciones cuadradas entre Y y $\beta_0 + \beta_1 x$. Por supuesto, $\beta_0 + \beta_1 x$ se estima por medio de $\hat{y} = b_0 + b_1 x$. Por consiguiente, tendría sentido que la varianza σ^2 se describa mejor como una desviación cuadrada de la observación típica y_i con respecto a la media estimada \hat{y}_i , que es el punto correspondiente sobre la recta ajustada. Entonces, los valores $(y_i - \hat{y}_i)$ revelan la varianza apropiada, de manera muy similar a como los valores $(y_i - \bar{y})^2$ miden la varianza cuando se realiza un muestreo en un escenario no relacionado con la regresión. En otras palabras, \bar{y} estima la media en la última situación sencilla, mientras que \hat{y}_i estima la media de y_i en una estructura de regresión. Ahora, ¿qué significa el divisor $n-2$? En las secciones que siguen observaremos que éstos son los grados de libertad asociados con el estimador s^2 de σ^2 . En tanto que en el escenario i.i.d. (independiente e idénticamente distribuidas), la normal estándar se resta un grado de libertad de n en el denominador, para lo cual una explicación razonable es que se estima un parámetro, que es la media μ por medio de, digamos, \bar{y} , pero en el problema de la regresión se estiman dos parámetros, que son β_0 y β_1 , por medio de b_0 y b_1 . Así, el parámetro importante σ^2 , que se estima mediante

$$s^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-2),$$

se denomina **error cuadrado medio**, que describe un tipo de media (división entre $n-2$) de los residuales cuadrados.

11.5 Inferencias sobre los coeficientes de regresión

Además de tan sólo estimar la relación lineal entre x y Y para fines de predicción, el experimentador podría estar interesado en hacer ciertas inferencias acerca de la pendiente y la intersección. Para dar ocasión a la prueba de hipótesis y a la construcción de intervalos de confianza para β_0 y β_1 , debemos estar dispuestos a hacer la suposición adicional de que cada ϵ_i , $i = 1, 2, \dots, n$, se distribuye de forma normal. Esta suposición implica que Y_1, Y_2, \dots, Y_n también están distribuidas normalmente, cada una con una distribución de probabilidad $n(y_i; \beta_0 + \beta_1 x_i, \sigma)$.

A partir de la sección 11.4 sabemos que B_1 tiene una distribución normal, y suponiendo normalidad, un resultado muy parecido al que se plantea en el teorema 8.4 nos permite concluir que $(n-2)S^2/\sigma^2$ es una variable chi cuadrada con $n-2$ grados de libertad, independiente de la variable aleatoria B_1 . Entonces, el teorema 8.5 garantiza que el estadístico

$$T = \frac{(B_1 - \beta_1)/(\sigma/\sqrt{S_{xx}})}{S/\sigma} = \frac{B_1 - \beta_1}{S/\sqrt{S_{xx}}}$$

tenga una distribución t con $n-2$ grados de libertad. Podemos utilizar el estadístico T para construir un intervalo de confianza del $100(1-\alpha)\%$ para el coeficiente β_1 .

Intervalo de confianza para β_1 Un intervalo de confianza de $100(1-\alpha)\%$ para el parámetro β_1 en la recta de regresión $\mu_{Y|x} = \beta_0 + \beta_1 x$ es

$$b_1 - t_{\alpha/2} \frac{s}{\sqrt{S_{xx}}} < \beta_1 < b_1 + t_{\alpha/2} \frac{s}{\sqrt{S_{xx}}},$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n-2$ grados de libertad.

Ejemplo 11.2: Calcule un intervalo de confianza de 95% para β_1 en la recta de regresión $\mu_{Y|x} = \beta_0 + \beta_1 x$, con base en los datos de contaminación de la tabla 11.1.

Solución: A partir de los resultados dados en el ejemplo 11.1, se determina que $S_{xx} = 4152.18$ y $S_{xy} = 3752.09$. Además, se observa que $S_{yy} = 3713.88$. Recuerde que $b_1 = 0.903643$. En consecuencia,

$$s^2 = \frac{S_{yy} - b_1 S_{xy}}{n-2} = \frac{3713.88 - (0.903643)(3752.09)}{31} = 10.4299.$$

Por lo tanto, al sacar la raíz cuadrada obtenemos $s = 3.2295$. Si usamos la tabla A.4 encontramos que $t_{0.025} \approx 2.045$ para 31 grados de libertad. Así, un intervalo de confianza de 95% para β_1 es

$$0.903643 - \frac{(2.045)(3.2295)}{\sqrt{4152.18}} < \beta_1 < 0.903643 + \frac{(2.045)(3.2295)}{\sqrt{4152.18}},$$

que se simplifica a

$$0.8012 < \beta_1 < 1.0061.$$



Prueba de hipótesis sobre la pendiente

Para probar la hipótesis nula H_0 de que $\beta_1 = \beta_{10}$, en comparación con una alternativa posible, utilizamos de nuevo la distribución t con $n - 2$ grados de libertad con el fin de establecer una región crítica y después basar nuestra decisión en el valor de

$$t = \frac{b_1 - \beta_{10}}{s/\sqrt{S_{xx}}}$$

El método se ilustra con el ejemplo siguiente.

Ejemplo 11.3: Utilice el valor estimado $b_1 = 0.903643$ del ejemplo 11.1 y pruebe la hipótesis de que $\beta_1 = 1.0$ en comparación con la alternativa de que $\beta_1 < 1.0$.

Solución: Las hipótesis son $H_0: \beta_1 = 1.0$ y $H_1: \beta_1 < 1.0$. Por lo tanto,

$$t = \frac{0.903643 - 1.0}{3.2295/\sqrt{4152.18}} = -1.92,$$

con $n - 2 = 31$ grados de libertad ($P \approx 0.03$).

Decisión: El valor t es significativo al nivel 0.03, lo cual sugiere evidencia sólida de que $\beta_1 < 1.0$. ▮

Una prueba t importante sobre la pendiente es la prueba de la hipótesis

$$H_0: \beta_1 = 0 \text{ en comparación con } H_1: \beta_1 \neq 0.$$

Cuando no se rechaza la hipótesis nula la conclusión es que no hay relación lineal significativa entre $E(y)$ y la variable independiente x . La gráfica de los datos del ejemplo 11.1 sugeriría que existe una relación lineal. Sin embargo, en ciertas aplicaciones en las que σ^2 es grande y, por ende, hay "ruido" considerable en los datos, una gráfica, aunque útil, quizá no produzca información clara para el investigador. El rechazo anterior de H_0 implica que hay una relación lineal significativa.

La figura 11.7 muestra la salida de resultados de MINITAB que presenta la prueba t para

$$H_0: \beta_1 = 0 \text{ en comparación con } H_1: \beta_1 \neq 0,$$

para los datos del ejemplo 11.1. Observe el coeficiente de regresión (Coef), el error estándar (EE Coef), el valor t (T) y el valor P (P). Se rechaza la hipótesis nula. Es claro que existe una relación lineal significativa entre la reducción de la demanda media del oxígeno químico y la reducción de los sólidos. Observe que el estadístico t se calcula como

$$t = \frac{\text{coeficiente}}{\text{error estándar}} = \frac{b_1}{s/\sqrt{S_{xx}}}$$

El no rechazo de $H_0: \beta_1 = 0$ sugiere que no hay una relación lineal entre Y y x . La figura 11.8 es una ilustración de la implicación de este resultado; podría significar que los cambios de x tienen poco efecto sobre los cambios de Y , como se ve en el inciso *a*. Sin embargo, también puede indicar que la relación verdadera es no lineal, como se aprecia en *b*.

Cuando se rechaza $H_0: \beta_1 = 0$ existe la implicación de que el término lineal en x que reside en el modelo explica una parte significativa de la variabilidad de Y . Las dos gráfi-

Regression Analysis: COD versus Per_Red
 The regression equation is COD = 3.83 + 0.904 Per_Red

Predictor	Coef	SE Coef	T	P
Constant	3.830	1.768	2.17	0.038
Per_Red	0.90364	0.05012	18.03	0.000

S = 3.22954 R-Sq = 91.3% R-Sq(adj) = 91.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	3390.6	3390.6	325.08	0.000
Residual Error	31	323.3	10.4		
Total	32	3713.9			

Figura 11.7: Salida de resultados de MINITAB para la prueba t de los datos del ejemplo 11.1.

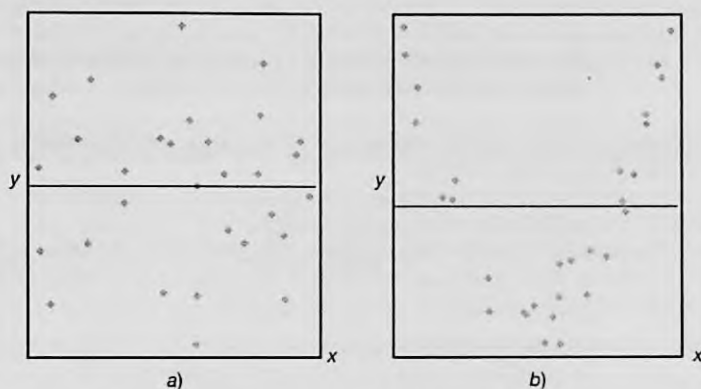


Figura 11.8: No se rechaza la hipótesis $H_0: \beta_1 = 0$.

cas que aparecen en la figura 11.9 ilustran los escenarios posibles. Como se muestra en el inciso a de la figura, el rechazo de H_0 sugiere que la relación en efecto es lineal. En el caso del inciso b, lo que se observa sugiere que, aunque el modelo contenga un efecto lineal, se podría obtener una mejor representación si se incluyera un término polinomial (tal vez cuadrático), es decir, términos que complementen el término lineal.

Inferencia estadística sobre la intersección

Los intervalos de confianza y la prueba de hipótesis del coeficiente β_0 se podrían establecer a partir del hecho de que B_0 también se distribuye de forma normal. No es difícil demostrar que

$$T = \frac{B_0 - \beta_0}{S \sqrt{\sum_{i=1}^n x_i^2 / (nS_{xx})}}$$

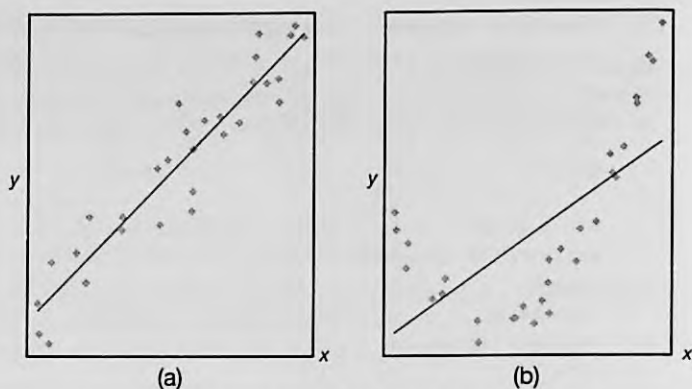


Figura 11.9: Se rechaza la hipótesis de que $H_0: \beta_1 = 0$.

tiene una distribución t con $n - 2$ grados de libertad, de manera que podemos construir un intervalo de confianza de $100(1 - \alpha)\%$ para α .

Intervalo de confianza para β_0 Un intervalo de confianza de $100(1 - \alpha)\%$ para el parámetro β_0 en la recta de regresión $\mu_{Y|x} = \beta_0 + \beta_1 x$ es

$$b_0 - t_{\alpha/2} \frac{s}{\sqrt{nS_{xx}}} \sqrt{\sum_{i=1}^n x_i^2} < \beta_0 < b_0 + t_{\alpha/2} \frac{s}{\sqrt{nS_{xx}}} \sqrt{\sum_{i=1}^n x_i^2},$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Ejemplo 11.4: Calcule un intervalo de confianza de 95% para β_0 en la recta de regresión $\mu_{Y|x} = \beta_0 + \beta_1 x$ con base en los datos de la tabla 11.1.

Solución: En los ejemplos 11.1 y 11.2 se encontró que

$$S_{xx} = 4152.18 \quad \text{y} \quad s = 3.2295.$$

Del ejemplo 11.1 se tiene que

$$\sum_{i=1}^n x_i^2 = 41,086 \quad \text{y} \quad b_0 = 3.829633.$$

Si usamos la tabla A.4, encontramos que $t_{0.025} \approx 2.045$ para 31 grados de libertad. Por lo tanto, un intervalo de confianza de 95% para β_0 es

$$3.829633 - \frac{(2.045)(3.2295) \sqrt{41,086}}{\sqrt{(33)(4152.18)}} < \beta_0 < 3.829633 + \frac{(2.045)(3.2295) \sqrt{41,086}}{\sqrt{(33)(4152.18)}},$$

que se simplifica a $0.2132 < \beta_0 < 7.4461$.

Para probar la hipótesis nula H_0 de que $\beta_0 = \beta_{00}$ en comparación con una alternativa posible utilizamos la distribución t con $n - 2$ grados de libertad para establecer una región crítica y, luego, basar nuestra decisión en el valor de

$$t = \frac{b_0 - \beta_{00}}{s \sqrt{\sum_{i=1}^n x_i^2 / (nS_{xx})}}$$

Ejemplo 11.5: Utilice el valor estimado de $b_0 = 3.829633$ del ejemplo 11.1 y, a un nivel de significancia de 0.05, pruebe la hipótesis de que $\beta_0 = 0$ en comparación con la alternativa de que $\beta_0 \neq 0$. Entonces

Solución: Las hipótesis son $H_0: \beta_0 = 0$ y $H_1: \beta_0 \neq 0$. Así que,

$$t = \frac{3.829633 - 0}{3.2295 \sqrt{41,086 / [(33)(4152.18)]}} = 2.17,$$

con 31 grados de libertad. Por lo tanto, $P = \text{valor } P \approx 0.038$ y concluimos que $\beta_0 \neq 0$. Observe que esto tan sólo es Coef/desviación estándar, como se aprecia en la salida de resultados de MINITAB en la figura 11.7. El SE Coef es el error estándar de la intersección estimada. ▮

Una medida de la calidad del ajuste: el coeficiente de determinación

Observe en la figura 11.7 que aparece un elemento denotado con R-Sq, cuyo valor es 91.3%. Esta cantidad, R^2 , se denomina **coeficiente de determinación** y es una medida de la **proporción de la variabilidad explicada por el modelo ajustado**. En la sección 11.8 se presentará el concepto del método del análisis de varianza para la prueba de hipótesis en la regresión. El enfoque del análisis de varianza utiliza la suma de los cuadrados del error $SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ y la **suma total de los cuadrados corregida STCC** $= \sum_{i=1}^n (y_i - \bar{y}_i)^2$. Esta última representa la variación en los valores de respuesta que *idealmente* serían explicados con el modelo. El valor de la SCE es la variación debida al error, o la **variación no explicada**. Resulta claro que si la $SCE = 0$, toda variación queda explicada. La cantidad que representa la variación explicada es $STCC - SCE$. R^2 es el

$$\text{Coeficiente de determinación: } R^2 = 1 - \frac{SCE}{STCC}.$$

Advierta que si el ajuste es perfecto, *todos los residuales son cero*, y así $R^2 = 1.0$. Pero si la SCE es tan sólo un poco menor que la STCC, $R^2 \approx 0$. Observe en la salida de resultados de la figura 11.7 que el coeficiente de determinación sugiere que el modelo ajustado a los datos explica el 91.3% de la variabilidad observada en la respuesta, la reducción en la demanda de oxígeno químico.

La figura 11.10 ofrece ejemplos de una gráfica con un buen ajuste ($R^2 \approx 1.0$) en *a*) y una gráfica con un ajuste deficiente ($R^2 \approx 0$) en *b*).

Errores en el uso de R^2

Los analistas citan con mucha frecuencia los valores de R^2 , quizá debido a su simplicidad. Sin embargo, hay errores en su interpretación. La confiabilidad de R^2 depende del

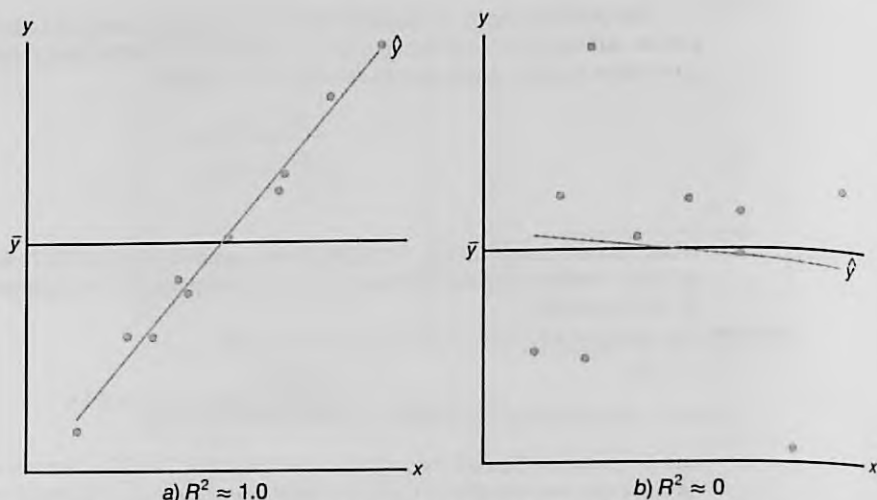


Figura 11.10: Gráficas que ilustran un ajuste muy bueno y otro deficiente.

tamaño del conjunto de los datos de la regresión y del tipo de aplicación. Resulta claro que $0 \leq R^2 \leq 1$, y el límite superior se logra cuando el ajuste a los datos es perfecto, es decir, cuando todos los residuales son cero. ¿Cuál es un valor aceptable de R^2 ? Se trata de una pregunta difícil de responder. Un químico encargado de establecer una calibración lineal de una pieza de equipo de alta precisión seguramente esperaría obtener un valor muy alto de R^2 (quizá superior a 0.99); mientras que un científico del comportamiento, que trabaja con datos en los que influye la variabilidad de la conducta humana, quizá se sentiría afortunado si obtuviera un valor de R^2 de hasta 0.70. Un individuo con experiencia en el ajuste de modelos tiene la sensibilidad para saber cuándo un valor es suficientemente grande dada la situación que está enfrentando. Es evidente que algunos fenómenos científicos se prestan más a un modelamiento más preciso que otros.

Es peligroso usar el criterio de R^2 para comparar *modelos en competencia* para el mismo conjunto de datos. Cuando se agregan términos adicionales al modelo, por ejemplo un regresor más, disminuye la *SCE*, lo que provoca que R^2 aumente (o al menos no disminuya). Esto implica que R^2 se puede volver artificialmente elevado por medio de la práctica inapropiada de *sobreajustar*, es decir, de incluir demasiados términos en el modelo. Por consiguiente, el incremento inevitable de R^2 que se logra al agregar términos adicionales no implica que éstos se necesitaban. En realidad, el modelo simple puede ser mejor para predecir los valores de la respuesta. En el capítulo 12, cuando se presente el concepto de los modelos que implican **más de un solo regresor**, se estudiará con detalle el papel del sobreajuste y su influencia sobre la capacidad de predicción. En este momento baste decir que *para seleccionar un modelo no se debe adoptar un proceso de selección que sólo incluya la consideración de R^2 .*

11.6 Predicción

Hay varias razones para construir un modelo de regresión lineal. Una de ellas es, desde luego, predecir valores de respuesta para uno o más valores de la variable independiente. En esta sección se centra el enfoque en los errores asociados con la predicción.

La ecuación $\hat{y} = b_0 + b_1x$ se puede utilizar para predecir o estimar la **respuesta media** $\mu_{Y|x_0}$ en $x = x_0$, donde x_0 no necesariamente es uno de los valores preestablecidos, o cuando $x = x_0$, se podría emplear para pronosticar un solo valor y_0 de la variable Y_0 . Se esperaría que el error de predicción fuera mayor para el caso de un solo valor pronosticado que para aquel en que se predice una media. Entonces, esto afectaría la anchura de los intervalos para los valores que se predicen.

Suponga que el experimentador desea construir un intervalo de confianza para $\mu_{Y|x_0}$. En tal caso debe usar el estimador puntual $\hat{Y}_0 = B_0 + B_1x_0$ para estimar $\mu_{Y|x_0} = \beta_0 + \beta_1x_0$. Se puede demostrar que la distribución muestral de \hat{Y}_0 es normal con media

$$\mu_{Y|x_0} = E(\hat{Y}_0) = E(B_0 + B_1x_0) = \beta_0 + \beta_1x_0 = \mu_{Y|x_0}$$

y varianza

$$\sigma_{\hat{Y}_0}^2 = \sigma_{B_0 + B_1x_0}^2 = \sigma_{\hat{Y} + B_1(x_0 - \bar{x})}^2 = \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right],$$

esta última surge del hecho de que $\text{Cov}(\bar{Y}_0, B_1) = 0$ (véase el ejercicio de repaso 11.61 de la página 438). Por consiguiente, ahora podemos construir un intervalo de confianza de $100(1 - \alpha)\%$ sobre la respuesta media $\mu_{Y|x_0}$ a partir del estadístico

$$T = \frac{\hat{Y}_0 - \mu_{Y|x_0}}{S \sqrt{1/n + (x_0 - \bar{x})^2/S_{xx}}},$$

que tiene una distribución t con $n - 2$ grados de libertad.

Intervalo de confianza para $\mu_{Y|x_0}$ Un intervalo de confianza de $100(1 - \alpha)\%$ para la respuesta media $\mu_{Y|x_0}$ es

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} < \mu_{Y|x_0} < \hat{y}_0 + t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}},$$

$t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Ejemplo 11.6: Con los datos de la tabla 11.1 construya límites de confianza de 95% para la respuesta media $\mu_{Y|x_0}$.

Solución: A partir de la ecuación de regresión encontramos que, para $x_0 = 20\%$ de reducción de sólidos, digamos,

$$\hat{y}_0 = 3.829633 + (0.903643)(20) = 21.9025.$$

Además, $\bar{x} = 33.4545$, $S_{xx} = 4152.18$, $s = 3.2295$ y $t_{0.025} \approx 2.045$ para 31 grados de libertad. Por lo tanto, un intervalo de confianza de 95% para $\mu_{Y|20}$ es

$$\begin{aligned} 21.9025 - (2.045)(3.2295) \sqrt{\frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}} &< \mu_{Y|20} \\ &< 21.9025 + (2.045)(3.2295) \sqrt{\frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}}, \end{aligned}$$

o simplemente, $20.1071 < \mu_{Y|20} < 23.6979$.

Si repetimos los cálculos anteriores para cada uno de los diferentes valores de x_0 , obtenemos los límites de confianza correspondientes para cada $\mu_{Y|x_0}$. En la figura 11.11 se presentan los datos de los puntos, la recta de regresión estimada y los límites de confianza superior e inferior sobre la media de $Y|x$.

En el ejemplo 11.6 tenemos 95% de confianza en que la reducción media poblacio-

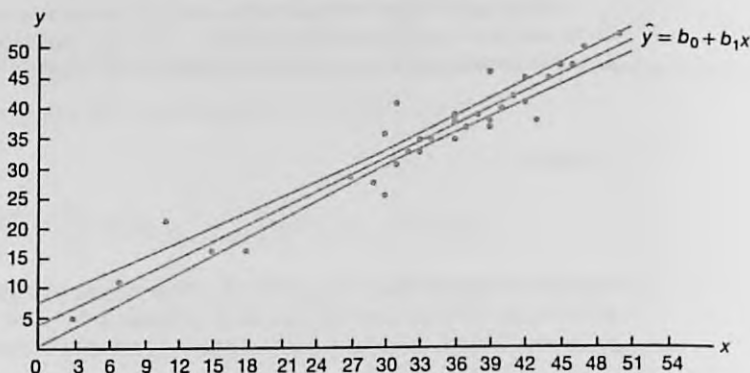


Figura 11.11: Límites de confianza para el valor medio de $Y|x$.

nal en la demanda de oxígeno químico estará entre el 20.1071% y 23.6979%, cuando la reducción de sólidos sea de 20%.

Predicción del intervalo

Otro tipo de intervalo que con frecuencia se malinterpreta y se confunde con aquel dado para $\mu_{Y|x}$ es el intervalo de la predicción para una respuesta futura observada. En realidad, en muchos casos el intervalo de la predicción es más relevante para el científico o el ingeniero que el intervalo de confianza sobre la media. En el ejemplo del contenido de alquitrán y la temperatura de entrada, mencionado en la sección 11.1, seguramente sería interesante no sólo estimar la media del contenido de alquitrán a una temperatura específica, sino también construir un intervalo que refleje el error en la predicción de una cantidad futura observada del contenido de alquitrán a la temperatura dada.

Para obtener un **intervalo de predicción** para cualquier valor único y_0 de la variable Y_0 es necesario estimar la varianza de las diferencias entre las ordenadas \hat{y}_0 , obtenidas de las rectas de regresión calculadas en el muestreo repetido cuando $x = x_0$, y la ordenada verdadera correspondiente y_0 . Podríamos considerar la diferencia $\hat{y}_0 - y_0$ como un valor de la variable aleatoria $\hat{Y}_0 - Y_0$, cuya distribución muestral se podría demostrar que es normal con media

$$\mu_{\hat{Y}_0 - Y_0} = E(\hat{Y}_0 - Y_0) = E[B_0 + B_1 x_0 - (\beta_0 + \beta_1 x_0 + \epsilon_0)] = 0$$

y varianza

$$\sigma_{\hat{Y}_0 - Y_0}^2 = \sigma_{B_0 + B_1 x_0 - \epsilon_0}^2 = \sigma_{\hat{Y}_0 + B_1(x_0 - \bar{x}) - \epsilon_0}^2 = \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]$$

Así, un intervalo de predicción de $100(1 - \alpha)\%$ para un solo valor pronosticado y_0 se puede construir a partir del estadístico

$$T = \frac{\hat{Y}_0 - Y_0}{S \sqrt{1 + 1/n + (x_0 - \bar{x})^2/S_{xx}}},$$

que tiene una distribución t con $n - 2$ grados de libertad.

Intervalo de predicción para y_0 Un intervalo de predicción de $100(1 - \alpha)\%$ para una sola respuesta y_0 es dado por

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} < y_0 < \hat{y}_0 + t_{\alpha/2} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}},$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Es claro que hay una diferencia entre el concepto de un intervalo de confianza y el del intervalo de predicción antes descrito. La interpretación del intervalo de confianza es idéntica a la que se describió para todos los intervalos de confianza sobre los parámetros de la población estudiados en el libro. De hecho, $\mu_{Y|x_0}$ es un parámetro de la población. Sin embargo, el intervalo de la predicción calculado representa un intervalo que tiene una probabilidad igual a $1 - \alpha$ de contener no un parámetro sino un valor futuro de y_0 de la variable aleatoria Y_0 .

Ejemplo 11.7: Con los datos de la tabla 11.1 construya un intervalo de predicción de 95% para y_0 cuando $x_0 = 20\%$.

Solución: Tenemos que $n = 33$, $x_0 = 20$, $\bar{x} = 33.4545$, $\hat{y}_0 = 21.9025$, $S_{xx} = 4152.18$, $s = 3.2295$, y $t_{0.025} \approx 2.045$ para 31 grados de libertad. Por lo tanto, un intervalo de predicción de 95% para y_0 es

$$\begin{aligned} 21.9025 - (2.045)(3.2295) \sqrt{1 + \frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}} &< y_0 \\ &< 21.9025 + (2.045)(3.2295) \sqrt{1 + \frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}}, \end{aligned}$$

que se simplifica como $15.0585 < y_0 < 28.7464$. ▮

En la figura 11.12 se presenta otra gráfica de los datos de reducción de la demanda de oxígeno químico, tanto con los intervalos de confianza de la respuesta media como con el intervalo de predicción sobre una respuesta individual. En el caso de la respuesta media la gráfica refleja un intervalo mucho más angosto alrededor de la recta de regresión.

Ejercicios

11.15 Remítase al ejercicio 11.1 de la página 398,

- evalúe s^2 ;
- pruebe la hipótesis de que $\beta_1 = 0$ en comparación con la alternativa de que $\beta_1 \neq 0$ a un nivel de significancia de 0.05, e interprete la decisión resultante.

11.16 Remítase al ejercicio 11.2 de la página 398,

- evalúe s^2 ;
- construya un intervalo de confianza de 95% para β_0 ;
- construya un intervalo de confianza de 95% para β_1 .

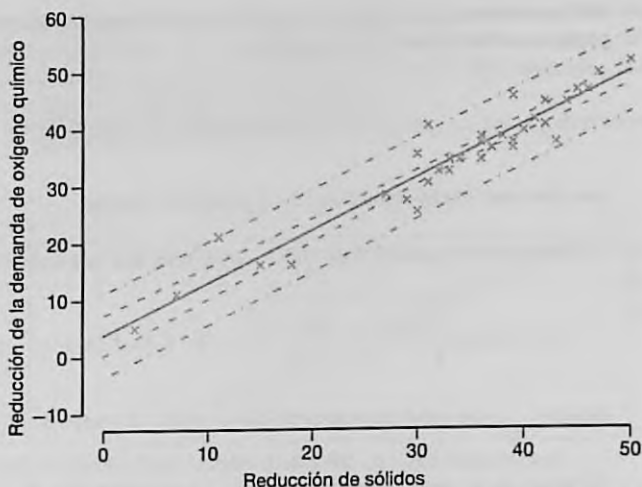


Figura 11.12: Intervalos de confianza y predicción para los datos de la reducción de la demanda de oxígeno químico; las bandas internas indican los límites de confianza para las respuestas medias y las externas señalan los límites de predicción para las respuestas futuras.

11.17 Remítase al ejercicio 11.5 de la página 398,

- evalúe s^2 ;
- construya un intervalo de confianza de 95% para β_0 ;
- construya un intervalo de confianza de 95% para β_1 .

11.18 Remítase al ejercicio 11.6 de la página 399,

- evalúe s^2 ;
- construya un intervalo de confianza de 99% para β_0 ;
- construya un intervalo de confianza de 99% para β_1 .

11.19 Remítase al ejercicio 11.3 de la página 398,

- evalúe s^2 ;
- construya un intervalo de confianza de 99% para β_0 ;
- construya un intervalo de confianza de 99% para β_1 .

11.20 Pruebe la hipótesis de que $\beta_0 = 10$ en el ejercicio 11.8 de la página 399, en comparación con la alternativa de que $\beta_0 < 10$. Utilice un nivel de significancia de 0.05.

11.21 Pruebe la hipótesis de que $\beta_1 = 6$ en el ejercicio 11.9 de la página 399, en comparación con la alternativa de que $\beta_1 < 6$. Utilice un nivel de significancia de 0.025.

11.22 Utilice el valor de s^2 que se obtuvo en el ejercicio 11.16a para construir un intervalo de confianza de 95% para $\mu_{Y|85}$ en el ejercicio 11.2 de la página 398.

11.23 Remítase al ejercicio 11.6 de la página 399 y utilice el valor de s^2 que se obtuvo en el ejercicio 11.18a para calcular

- un intervalo de confianza de 95% para la resistencia media al corte cuando $x = 24.5$;
- un intervalo de predicción de 95% para un solo valor pronosticado de la resistencia al corte cuando $x = 24.5$.

11.24 Utilice el valor de s^2 que se obtuvo en el ejercicio 11.17a) y grafique la regresión lineal y las bandas de confianza de 95% para la respuesta media $\mu_{Y|x}$ en el caso de los datos del ejercicio 11.5 de la página 398.

11.25 Utilice el valor de s^2 que se obtuvo en el ejercicio 11.17a) y construya un intervalo de confianza de 95% para la cantidad de azúcar convertida correspondiente a $x = 1.6$ en el ejercicio 11.5 de la página 398.

11.26 Remítase al ejercicio 11.3 de la página 398, y utilice el valor de s^2 que se obtuvo en el ejercicio 11.19a) para calcular

- un intervalo de confianza de 99% para la cantidad promedio del producto químico que se disolverá en 100 gramos de agua a 50°C ;

- b) un intervalo de predicción de 99% para la cantidad de producto químico que se disolverá en 100 gramos de agua a 50°C.

11.27 Considere la regresión de la distancia recorrida para ciertos automóviles, en millas por galón (mpg) y su peso en libras (wt). Los datos son de la revista *Consumer Reports* (abril de 1997). En la figura 11.13 se presenta una parte de la salida del SAS con los resultados del procedimiento.

- a) Estime la distancia recorrida para un vehículo que pesa 4000 libras.
 b) Suponga que los ingenieros de Honda afirman que, en promedio, el Civic (o cualquier otro modelo que pese 2440 libras) recorre más de 30 millas por galón (mpg). Con base en los resultados del análisis de regresión, ¿creería usted dicha afirmación? Explique su respuesta.
 c) Los ingenieros de diseño del Lexus ES300 consideraron que un rendimiento de 18 mpg sería el objetivo ideal para dicho modelo (o cualquier otro modelo que pese 3390 libras), aunque se espera que haya cierta variación. ¿Es probable que ese objetivo sea realista? Comente al respecto.

11.28 Existen aplicaciones importantes en las que, debido a restricciones científicas conocidas, la recta de regresión debe atravesar el origen, es decir, la intersección debe estar en el cero. En otras palabras, el modelo debe ser

$$Y_i = \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, n,$$

y tan sólo se requiere estimar un parámetro sencillo. Con frecuencia a este modelo se le denomina **modelo de regresión por el origen**.

- a) Demuestre que el estimador de mínimos cuadrados para la pendiente es

$$b_1 = \left(\sum_{i=1}^n x_i y_i \right) / \left(\sum_{i=1}^n x_i^2 \right).$$

- b) Demuestre que $\sigma_{b_1}^2 = \sigma^2 / \left(\sum_{i=1}^n x_i^2 \right)$.

- c) Demuestre que b_1 del inciso a es un estimador insesgado para β_1 . Es decir, demuestre que $E(b_1) = \beta_1$.

11.29 Dado el conjunto de datos

y	x
7	2
50	15
100	30
40	10
70	20

- a) Grafique los datos.
 b) Ajuste una recta de regresión por el origen.
 c) Grafique la recta de regresión sobre la gráfica de los datos.
 d) Calcule una fórmula general (en términos de y_i y la pendiente b_1) para el estimador de σ^2 .
 e) Calcule una fórmula para $\text{Var}(\hat{y}_i)$; $i = 1, 2, \dots, n$, aplicable a este caso.
 f) Grafique límites de confianza de 95% para la respuesta media alrededor de la recta de regresión.

11.30 Para los datos del ejercicio 11.29 calcule un intervalo de predicción de 95% en $x = 25$.

Root MSE	1.48794	R-Square	0.9509
Dependent Mean	21.50000	Adj R-Sq	0.9447

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	44.78018	1.92919	23.21	<.0001
WT	1	-0.00686	0.00055133	-12.44	<.0001

MODEL	WT	MPG	Predict	LMean	UMean	Lpred	Upred	Residual
GMC	4520	15	13.7720	11.9752	15.5688	9.8988	17.6451	1.22804
Geo	2065	29	30.6138	28.6063	32.6213	26.6385	34.5891	-1.61381
Honda	2440	31	28.0412	26.4143	29.6681	24.2439	31.8386	2.95877
Hyundai	2290	28	29.0703	27.2967	30.8438	25.2078	32.9327	-1.07026
Infiniti	3195	23	22.8618	21.7478	23.9758	19.2543	26.4693	0.13825
Isuzu	3480	21	20.9066	19.8160	21.9972	17.3062	24.5069	0.09341
Jeep	4090	15	16.7219	15.3213	18.1224	13.0158	20.4279	-1.72185
Land	4535	13	13.6691	11.8570	15.4811	9.7888	17.5493	-0.66905
Lexus	3390	22	21.5240	20.4390	22.6091	17.9253	25.1227	0.47599
Lincoln	3930	18	17.8195	16.5379	19.1011	14.1568	21.4822	0.18051

Figura 11.13: Salida de resultados del SAS para el ejercicio 11.27.

11.7 Selección de un modelo de regresión

Gran parte de lo que se ha presentado hasta ahora acerca de la regresión que involucra una sola variable independiente depende de la suposición de que el modelo elegido es correcto, la suposición de que $\mu_{y|x}$ se relaciona con x linealmente en los parámetros. Es cierto que no se esperaría que la predicción de la respuesta fuera buena si hubiera diversas variables independientes que no se tomaran en cuenta en el modelo, que afectarían la respuesta y variarían en el sistema. Además, la predicción seguramente sería inadecuada si la estructura verdadera que relaciona $\mu_{y|x}$ con x fuera extremadamente no lineal en el rango de las variables consideradas.

Es frecuente que se utilice el modelo de regresión lineal simple aun cuando se sepa que el modelo no es lineal o que se desconozca la estructura verdadera. Este método suele ser acertado, en particular cuando el rango de las x es estrecho. De esta manera, el modelo que se utiliza se vuelve una función de aproximación que se espera sea una representación adecuada del panorama verdadero en la región de interés. Sin embargo, hay que señalar el efecto que tendría un modelo inadecuado sobre los resultados presentados hasta este momento. Por ejemplo, si el modelo verdadero, desconocido para el experimentador, es lineal en más de una x , digamos,

$$\mu_{y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2,$$

entonces el estimado $b_1 = S_{xy}/S_{xx}$ de los mínimos cuadrados ordinarios que se calcula considerando tan sólo x_1 en el experimento es, en circunstancias generales, un estimado sesgado del coeficiente β_1 , donde el sesgo es una función del coeficiente adicional β_2 (véase el ejercicio de repaso 11.65 en la página 438). Asimismo, el estimado s^2 para σ^2 es sesgado debido a la variable adicional.

11.8 El método del análisis de varianza

Con frecuencia el problema de analizar la calidad de la recta de regresión estimada se maneja por medio del método del **análisis de varianza** (ANOVA), que es un procedimiento mediante el cual la variación total de la variable dependiente se subdivide en componentes significativos, que luego se observan y se tratan en forma sistemática. El análisis de varianza, que se estudia en el capítulo 13, es un recurso poderoso que se emplea en muchas situaciones.

Suponga que tenemos n puntos de datos experimentales en la forma usual (x_i, y_i) y que se estima la recta de regresión. En la sección 11.4 para la estimación de σ^2 se estableció la identidad

$$S_{yy} = b_1 S_{xy} + SCE.$$

Una formulación alternativa y quizá más informativa es la siguiente:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Logramos hacer una partición de la **suma total de los cuadrados corregida de y** en dos componentes que deberían proporcionar un significado particular para el experimentador. Esta partición se debería indicar en forma simbólica como

$$STCC = SCR + SCE.$$

El primer componente de la derecha, *SCR*, se denomina **suma de cuadrados de la regresión** y refleja la cantidad de variación de los valores y que se explica con el modelo, que en este caso es la línea recta postulada. El segundo componente es la ya conocida suma de cuadrados del error, que refleja la variación alrededor de la recta de regresión.

Suponga que nos interesa probar la hipótesis

$$H_0: \beta_1 = 0 \text{ en comparación con } H_1: \beta_1 \neq 0,$$

donde la hipótesis nula en esencia dice que el modelo es $\mu_{Y|x} = \beta_0$; es decir, la variación en los resultados *Y* debida a las fluctuaciones de probabilidad o aleatorias que son independientes de los valores de *x*. Esta condición se refleja en la figura 11.10b). En las condiciones de esta hipótesis nula se puede demostrar que *SCR*/ σ^2 , y *SCE*/ σ^2 son valores de variables cuadradas independientes con 1 y $n - 2$ grados de libertad, respectivamente y, usando el teorema 7.12, se sigue que *STCC*/ σ^2 también es un valor de una variable chi cuadrada con $n - 1$ grados de libertad. Para probar la hipótesis anterior calculamos

$$f = \frac{SCR / 1}{SCE / (n-2)} = \frac{SCR}{s^2}$$

y rechazamos H_0 al nivel de significancia α cuando $f > f_\alpha(1, n - 2)$.

Por lo general los cálculos se resumen mediante las medias de una **tabla de análisis de varianza**, como se indica en la tabla 11.2. Es costumbre referirse a las distintas sumas de los cuadrados divididos entre sus respectivos grados de libertad como **cuadrados medios**.

Tabla 11.2: Análisis de varianza para la prueba de $\beta_1 = 0$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	<i>f</i> calculada
Regresión	<i>SCR</i>	1	$\frac{SCR}{s^2}$	$\frac{SCR}{s^2}$
Error	<i>SCE</i>	$n - 2$	$s^2 = \frac{SCE}{n-2}$	
Total	<i>STCC</i>	$n - 1$		

Cuando se rechaza la hipótesis nula, es decir, cuando el estadístico *F* calculado excede al valor crítico $f_\alpha(1, n - 2)$, concluimos que **hay una cantidad significativa de variación en la respuesta justificada por el modelo postulada, que es la función de la línea recta**. Si el estadístico *F* está en la región de no rechazo, se concluye que los datos no reflejan evidencia suficiente para apoyar el modelo que se postula.

En la sección 11.5 se presentó un procedimiento donde se usa el estadístico

$$T = \frac{B_1 - \beta_{10}}{S / \sqrt{S_{xx}}}$$

para probar la hipótesis

$$H_0: \beta_1 = \beta_{10} \text{ contra } H_1: \beta_1 \neq \beta_{10},$$

donde *T* sigue la distribución *t* con $n - 2$ grados de libertad. La hipótesis se rechaza si $|t| > t_{\alpha/2}$ para un nivel de significancia α . Es interesante observar que en el caso especial en que probamos

$$H_0: \beta_1 = 0 \text{ en comparación con } H_1: \beta_1 \neq 0,$$

el valor del estadístico T se convierte en

$$t = \frac{b_1}{s/\sqrt{S_{xx}}},$$

y la hipótesis a considerar es idéntica a la que se prueba en la tabla 11.2. En otras palabras, la hipótesis nula establece que la variación en la respuesta se debe tan sólo al azar. El análisis de varianza utiliza la distribución F en vez de la distribución t . Para la alternativa bilateral ambos enfoques son idénticos. Esto se observa si se escribe

$$t^2 = \frac{b_1^2 S_{xx}}{s^2} = \frac{b_1 S_{xy}}{s^2} = \frac{SCR}{s^2},$$

que da como resultado un valor idéntico al valor f utilizado en el análisis de varianza. La relación fundamental entre la distribución t con v grados de libertad y la distribución F con 1 y v grados de libertad es

$$t^2 = f(1, v).$$

Desde luego, la prueba t permite probar en comparación con una alternativa unilateral, en tanto que la prueba F está restringida a una prueba en comparación con una alternativa bilateral.

Salida de resultados por computadora comentados para la regresión lineal simple

Considere nuevamente los datos de la tabla 11.1 sobre la reducción de la demanda de oxígeno químico. En las figuras 11.14 y 11.15 se presentan salidas de los resultados por computadora más completos. De nuevo se ilustran con el software *MINITAB*. La columna de la razón t indica pruebas para la hipótesis nula de valores de cero en el parámetro. El término "Fit" denota los valores \hat{y} , que con frecuencia se denominan **valores ajustados**. El término "SE Fit" se emplea para calcular los intervalos de confianza sobre la respuesta media. El elemento R^2 se calcula como $(SCR/STCC) \times 100$, y significa la proporción de variación en y explicada por la regresión de la línea recta. Asimismo, se incluyen los intervalos de confianza sobre la respuesta media y los intervalos de predicción sobre una observación nueva.

11.9 Prueba para la linealidad de la regresión: datos con observaciones repetidas

En ciertos tipos de situaciones experimentales el investigador tiene la capacidad de efectuar observaciones repetidas de la respuesta para cada valor de x . Aunque no es necesario tener dichas repeticiones para estimar β_0 y β_1 , las repeticiones permiten al experimentador obtener información cuantitativa acerca de lo apropiado que resulta el modelo. De hecho, si se generan observaciones repetidas, el investigador puede efectuar una prueba de significancia para determinar si el modelo es o no adecuado.

The regression equation is COD = 3.83 + 0.904 Per_Red						
Predictor	Coef	SE Coef	T	P		
Constant	3.830	1.768	2.17	0.038		
Per_Red	0.90364	0.05012	18.03	0.000		
S = 3.22954	R-Sq = 91.3%	R-Sq(adj) = 91.0%				
Analysis of Variance						
Source	DF	SS	MS	F	P	
Regression	1	3390.6	3390.6	325.08	0.000	
Residual Error	31	323.3	10.4			
Total	32	3713.9				

Obs	Per_Red	COD	Fit	SE Fit	Residual	St Resid
1	3.0	5.000	6.541	1.627	-1.541	-0.55
2	36.0	34.000	36.361	0.576	-2.361	-0.74
3	7.0	11.000	10.155	1.440	0.845	0.29
4	37.0	36.000	37.264	0.590	-1.264	-0.40
5	11.0	21.000	13.770	1.258	7.230	2.43
6	38.0	38.000	38.168	0.607	-0.168	-0.05
7	15.0	16.000	17.384	1.082	-1.384	-0.45
8	39.0	37.000	39.072	0.627	-2.072	-0.65
9	18.0	16.000	20.095	0.957	-4.095	-1.33
10	39.0	36.000	39.072	0.627	-3.072	-0.97
11	27.0	28.000	28.228	0.649	-0.228	-0.07
12	39.0	45.000	39.072	0.627	5.928	1.87
13	29.0	27.000	30.035	0.605	-3.035	-0.96
14	40.0	39.000	39.975	0.651	-0.975	-0.31
15	30.0	25.000	30.939	0.588	-5.939	-1.87
16	41.0	41.000	40.879	0.678	0.121	0.04
17	30.0	35.000	30.939	0.588	4.061	1.28
18	42.0	40.000	41.783	0.707	-1.783	-0.57
19	31.0	30.000	31.843	0.575	-1.843	-0.58
20	42.0	44.000	41.783	0.707	2.217	0.70
21	31.0	40.000	31.843	0.575	8.157	2.57
22	43.0	37.000	42.686	0.738	-5.686	-1.81
23	32.0	32.000	32.746	0.567	-0.746	-0.23
24	44.0	44.000	43.590	0.772	0.410	0.13
25	33.0	34.000	33.650	0.563	0.350	0.11
26	45.0	46.000	44.494	0.807	1.506	0.48
27	33.0	32.000	33.650	0.563	-1.650	-0.52
28	46.0	46.000	45.397	0.843	0.603	0.19
29	34.0	34.000	34.554	0.563	-0.554	-0.17
30	47.0	49.000	46.301	0.881	2.699	0.87
31	36.0	37.000	36.361	0.576	0.639	0.20
32	50.0	51.000	49.012	1.002	1.988	0.65
33	36.0	38.000	36.361	0.576	1.639	0.52

Figura 11.14: Salida de resultados de MINITAB de la regresión lineal simple para los datos de reducción de la demanda de oxígeno químico; parte I.

Seleccionemos una muestra aleatoria de n observaciones utilizando k valores distintos de x , por ejemplo, x_1, x_2, \dots, x_k , tales que la muestra contenga n_i valores observados de la variable aleatoria Y_1 correspondientes a los valores x_1 , con n_2 valores observados de Y_2 correspondientes a x_2, \dots, n_k valores observados de Y_k correspondientes a x_k . Necesariamente, $n = \sum_{i=1}^k n_i$.

Obs	Fit	SE Fit	95% CI	95% PI
1	6.541	1.627	(3.223, 9.858)	(-0.834, 13.916)
2	36.361	0.576	(35.185, 37.537)	(29.670, 43.052)
3	10.155	1.440	(7.218, 13.092)	(2.943, 17.367)
4	37.264	0.590	(36.062, 38.467)	(30.569, 43.960)
5	13.770	1.258	(11.204, 16.335)	(6.701, 20.838)
6	38.168	0.607	(36.931, 39.405)	(31.466, 44.870)
7	17.384	1.082	(15.177, 19.592)	(10.438, 24.331)
8	39.072	0.627	(37.793, 40.351)	(32.362, 45.781)
9	20.095	0.957	(18.143, 22.047)	(13.225, 26.965)
10	39.072	0.627	(37.793, 40.351)	(32.362, 45.781)
11	28.228	0.649	(26.905, 29.551)	(21.510, 34.946)
12	39.072	0.627	(37.793, 40.351)	(32.362, 45.781)
13	30.035	0.605	(28.802, 31.269)	(23.334, 36.737)
14	39.975	0.651	(38.648, 41.303)	(33.256, 46.694)
15	30.939	0.588	(29.739, 32.139)	(24.244, 37.634)
16	40.879	0.678	(39.497, 42.261)	(34.149, 47.609)
17	30.939	0.588	(29.739, 32.139)	(24.244, 37.634)
18	41.783	0.707	(40.341, 43.224)	(35.040, 48.525)
19	31.843	0.575	(30.669, 33.016)	(25.152, 38.533)
20	41.783	0.707	(40.341, 43.224)	(35.040, 48.525)
21	31.843	0.575	(30.669, 33.016)	(25.152, 38.533)
22	42.686	0.738	(41.181, 44.192)	(35.930, 49.443)
23	32.746	0.567	(31.590, 33.902)	(26.059, 39.434)
24	43.590	0.772	(42.016, 45.164)	(36.818, 50.362)
25	33.650	0.563	(32.502, 34.797)	(26.964, 40.336)
26	44.494	0.807	(42.848, 46.139)	(37.704, 51.283)
27	33.650	0.563	(32.502, 34.797)	(26.964, 40.336)
28	45.397	0.843	(43.677, 47.117)	(38.590, 52.205)
29	34.554	0.563	(33.406, 35.701)	(27.868, 41.239)
30	46.301	0.881	(44.503, 48.099)	(39.473, 53.128)
31	36.361	0.576	(35.185, 37.537)	(29.670, 43.052)
32	49.012	1.002	(46.969, 51.055)	(42.115, 55.908)
33	36.361	0.576	(35.185, 37.537)	(29.670, 43.052)

Figura 11.15: Salida de resultados de *MINITAB* de la regresión lineal simple para los datos de reducción de la demanda de oxígeno químico; parte II.

Definimos

$$y_{ij} = \text{el } j\text{-ésimo valor de la variable aleatoria } Y_i,$$

$$y_i = T_i = \sum_{j=1}^{n_i} y_{ij},$$

$$\bar{y}_i = \frac{T_i}{n_i}.$$

Entonces, si se realizaron $n_4 = 3$ mediciones de Y que corresponden a $x = x_4$, estas observaciones se indicarían por medio de y_{41} , y_{42} y y_{43} . Por lo tanto,

$$T_4 = y_{41} + y_{42} + y_{43}.$$

El concepto de la falta de ajuste

La suma de cuadrados del error consta de dos partes: la cantidad debida a la variación entre los valores de Y dentro de valores dados de x , y un componente que normalmente

se denomina contribución a la **falta de ajuste**. El primer componente refleja tan sólo la variación aleatoria, o **error experimental puro**, en tanto que el segundo es una medida de la variación sistemática introducida por los términos de orden superior. En nuestro caso éstos son términos de x distintos de la contribución lineal o de primer orden. Observe que al elegir un modelo lineal en esencia asumimos que este segundo componente no existe y que, en consecuencia, la suma de cuadrados del error se debe por completo a errores aleatorios. Si éste fuera el caso, entonces $s^2 = SCE/(n-2)$ es un estimado insesgado de σ^2 . Sin embargo, si el modelo no se ajusta a los datos en forma apropiada, entonces la suma de cuadrados del error estará inflada y producirá un estimador sesgado de σ^2 . Ya sea que el modelo se ajuste o no a los datos, siempre que se tienen observaciones repetidas es posible obtener un estimador insesgado de σ^2 calculando

$$s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n_i - 1}, \quad i = 1, 2, \dots, k,$$

para cada uno de los k valores distintos de x y, después, agrupando estas varianzas, tenemos

$$s^2 = \frac{\sum_{i=1}^k (n_i - 1)s_i^2}{n - k} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n - k}.$$

El numerador de s^2 es una **medida del error experimental puro**. A continuación se presenta un procedimiento de cálculo para separar la suma de los cuadrados del error en los dos componentes que representan el error puro y la falta de ajuste:

Cálculo 1. Calcular la suma de los cuadrados del error puro

de la suma de los cuadrados de la falta de ajuste

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2.$$

Esta suma de cuadrados tiene $n - k$ grados de libertad asociados con ella, y el cuadrado medio resultante es el estimador insesgado s^2 de σ^2 .

2. Restar la suma de los cuadrados del error puro de la suma de los cuadrados del error, SCE, con lo que se obtiene la suma de los cuadrados debida a la falta de ajuste. Los grados de libertad de la falta de ajuste también se obtienen simplemente restando $(n-2) - (n-k) = k-2$.

Los cálculos necesarios para probar hipótesis en un problema de regresión con mediciones repetidas de la respuesta se pueden resumir como se muestra en la tabla 11.3.

Las figuras 11.16 y 11.17 ilustran los puntos muestrales para las situaciones del "modelo correcto" y del "modelo incorrecto". En la figura 11.16, donde $\mu_{y|x}$ cae sobre una línea recta, no hay falta de ajuste cuando se asume un modelo lineal, por lo que la variación muestral alrededor de la recta de regresión es un error puro que resulta de la variación que ocurre entre observaciones repetidas. En la figura 11.17, donde es evidente que $\mu_{y|x}$ no cae sobre una línea recta, la responsable de la mayor parte de la variación alrededor de la recta de regresión, además del error puro, es la falta de ajuste que resulta de seleccionar por error un modelo lineal.

Tabla 11.3: Análisis de varianza para la prueba de linealidad de la regresión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada
Regresión	SCR	1	SCR	$\frac{SCR}{s^2}$
Error	SCE	$n - 2$		
Falta de ajuste	$\left\{ \begin{array}{l} SCE - SCE \text{ (puro)} \\ SCE \text{ (puro)} \end{array} \right.$	$\left\{ \begin{array}{l} k - 2 \\ n - k \end{array} \right.$	$\frac{SCE - SCE \text{ (puro)}}{k - 2}$	$\frac{SCE - SCE \text{ (puro)}}{s^2(k - 2)}$
Error puro			$s^2 = \frac{SCE \text{ (puro)}}{n - k}$	
Total	$STCC$	$n - 1$		

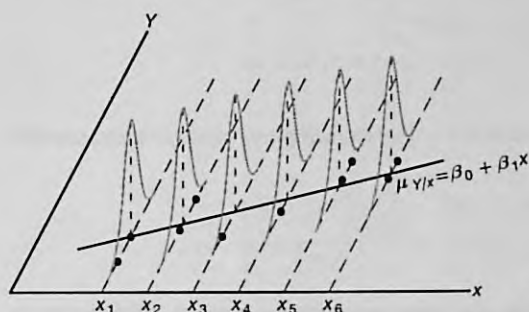


Figura 11.16: Modelo lineal correcto con componente sin falta de ajuste.

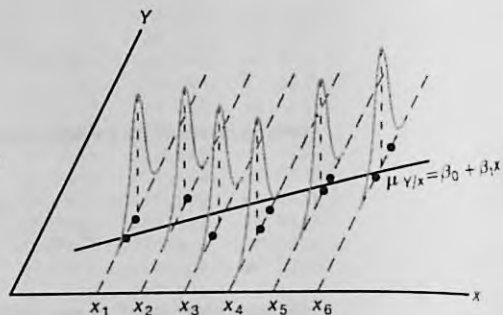


Figura 11.17: Modelo lineal incorrecto con componente de falta de ajuste.

¿Por qué es importante detectar la falta de ajuste?

El concepto de falta de ajuste es muy importante en las aplicaciones del análisis de regresión. De hecho, la necesidad de construir o diseñar un experimento que tome en cuenta la falta de ajuste se vuelve más crítica a medida que el problema y el mecanismo subyacente implicados se vuelven más complicados. Es cierto que no siempre se puede tener la certeza de que la estructura que se postula, en este caso el modelo de regresión lineal, sea una representación correcta o incluso adecuada. El ejemplo siguiente muestra la manera en que se parte la suma de cuadrados del error en los dos componentes que representan el error puro y la falta de ajuste. Lo adecuado del modelo se prueba al nivel de significancia α , comparando el cuadrado medio de la falta de ajuste dividido entre s^2 con $f_{\alpha}(k - 2, n - k)$.

Ejemplo 11.8: En la tabla 11.4 se presenta el registro de las observaciones del producto de una reacción química tomadas a distintas temperaturas. Calcule el modelo lineal $\mu_{y|x} = \beta_0 + \beta_1 x$ y pruebe la falta de ajuste.

Solución: Los resultados de los cálculos se presentan en la tabla 11.5.

Conclusión: La partición de la variación total de esta manera revela una variación significativa debida al modelo lineal y una cantidad insignificante de variación debida a la falta de ajuste. Por consiguiente, los datos experimentales no parecen sugerir la necesidad de considerar en el modelo términos superiores a los de primer orden y no se rechaza la hipótesis nula.

Tabla 11.4: Datos para el ejemplo 11.8

y (%)	x (°C)	y (%)	x (°C)
77.4	150	88.9	250
76.7	150	89.2	250
78.2	150	89.7	250
84.1	200	94.8	300
84.5	200	94.7	300
83.7	200	95.9	300

Tabla 11.5: Análisis de varianza de los datos de producto-temperatura

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada	Valores P
Regresión	509.2507	1	509.2507	1531.58	< 0.0001
Error	3.8660	10			
Falta de ajuste	1.2060	2	0.6030	1.81	0.2241
Error puro	2.6600	8	0.3325		
Total	513.1167	11			

Salida de resultados por computadora comentados para la prueba de falta de ajuste

En la figura 11.18 se presenta una salida de resultados por computadora para el análisis de los datos del ejemplo 11.8 con el programa SAS. Observe la "LOF" con 2 grados de libertad, que representa las contribuciones cuadrática y cúbica al modelo, y el valor P de 0.22, que sugiere que el modelo lineal (de primer orden) es adecuado.

Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	510.4566667	170.1522222	511.74	<.0001
Error	8	2.6600000	0.3325000		
Corrected Total	11	513.1166667			
R-Square		Coeff Var	Root MSE	yield Mean	
0.994816		0.666751	0.576628	86.48333	
Source	DF	Type III SS	Mean Square	F Value	Pr > F
temperature	1	509.2506667	509.2506667	1531.58	<.0001
LOF	2	1.2060000	0.6030000	1.81	0.2241

Figura 11.18: Salida de resultados del SAS que incluye el análisis de los datos del ejemplo 11.8.

Ejercicios

11.31 En el ejercicio 11.3 de la página 398 pruebe la linealidad de la regresión. Use un nivel de significancia de 0.05. Haga comentarios al respecto.

11.32 En el ejercicio 11.8 de la página 399 pruebe la linealidad de la regresión. Haga comentarios al respecto.

11.33 Suponga que tenemos una ecuación lineal que pasa por el origen $\mu_{Y|X} = \beta X$ (ejercicio 11.28).

a) Estime la regresión lineal que pasa por el origen para los siguientes datos:

x	0.5	1.5	3.2	4.2	5.1	6.5
y	1.3	3.4	6.7	8.0	10.0	13.2

- b) Suponga que se desconoce si la regresión verdadera debería pasar por el origen. Estime el modelo lineal $\mu_{y|x} = \beta_0 + \beta_1 x$ y pruebe la hipótesis de que $\beta_0 = 0$ a un nivel de significancia de 0.10, en comparación con la alternativa de que $\beta_0 \neq 0$.

11.34 En el ejercicio 11.5 de la página 398 utilice el método del análisis de varianza para probar la hipótesis de que $\beta_1 = 0$, en comparación con la hipótesis alternativa de que $\beta_1 \neq 0$, a un nivel de significancia de 0.05.

11.35 Los siguientes datos son el resultado de una investigación sobre el efecto de la temperatura de reacción x sobre la conversión porcentual de un proceso químico y . (Véase Myers, Montgomery y Anderson-Cook, 2009). Ajuste una regresión lineal simple y utilice pruebas de falta de ajuste para determinar si el modelo es adecuado. Analice los resultados.

Observación	Temperatura ($^{\circ}\text{C}$), x	Conversión (%), y
1	200	43
2	250	78
3	200	69
4	250	73
5	189.65	48
6	260.35	78
7	225	65
8	225	74
9	225	76
10	225	79
11	225	83
12	225	81

11.36 La ganancia de un transistor en un dispositivo de circuito integrado, entre el emisor y el colector (hFE), se relaciona con dos variables (Myers, Montgomery y Anderson-Cook, 2009) que se controlan en el proceso de deposición, controlado por el emisor en el tiempo (x_1 , en minutos) y la dosis del emisor (x_2 , en iones $\times 10^{14}$). Se observaron 14 muestras después de la deposición y los datos resultantes se presentan en la tabla siguiente. Consideraremos modelos de regresión lineal usando la ganancia como respuesta y el control del emisor en el tiempo o la dosis del emisor como la variable regresora.

Obs.	x_1 (tiempo de control, min)	x_2 (dosis, iones $\times 10^{14}$)	y (ganancia o hFE)
1	195	4.00	1004
2	255	4.00	1636
3	195	4.60	852
4	255	4.60	1506
5	255	4.20	1272
6	255	4.10	1270
7	255	4.60	1269
8	195	4.30	903
9	255	4.30	1555

10	255	4.00	1260
11	255	4.70	1146
12	255	4.30	1276
13	255	4.72	1225
14	340	4.30	1321

- a) Determine si el tiempo de control del emisor influye en la ganancia en una relación lineal. Es decir, pruebe $H_0: \beta_1 = 0$, donde β_1 es la pendiente de la variable regresora.
- b) Efectúe una prueba de falta de ajuste para determinar si la relación lineal es adecuada. Saque sus conclusiones.
- c) Determine si la dosis del emisor influye en la ganancia en una relación lineal. ¿Cuál variable regresora es el mejor predictor de la ganancia?

11.37 En los pesticidas se utilizan compuestos de organofosfatos (OF). Sin embargo, es importante estudiar el efecto que tienen sobre las especies expuestas a ellos. Como parte del estudio de laboratorio *Some Effects of Organophosphate Pesticides on Wildlife Species*, elaborado por el Departamento de Pesca y Vida Silvestre de Virginia Tech, se realizó un experimento en el cual se suministraron distintas dosis de un pesticida de OF específico a 5 grupos de 5 ratones (*peromyscus leucopus*). Los 25 ratones eran hembras de edad y condiciones similares. Un grupo no recibió el producto. La respuesta básica y consistió en medir la actividad cerebral. Se postuló que dicha actividad disminuiría con un incremento en la dosis de OF. A continuación se presentan los datos:

Animal	Dosis, x (mg/kg de peso corporal)	Actividad, y (moles/litro/min)
1	0.0	10.9
2	0.0	10.6
3	0.0	10.8
4	0.0	9.8
5	0.0	9.0
6	2.3	11.0
7	2.3	11.3
8	2.3	9.9
9	2.3	9.2
10	2.3	10.1
11	4.6	10.6
12	4.6	10.4
13	4.6	8.8
14	4.6	11.1
15	4.6	8.4
16	9.2	9.7
17	9.2	7.8
18	9.2	9.0
19	9.2	8.2
20	9.2	2.3
21	18.4	2.9
22	18.4	2.2
23	18.4	3.4
24	18.4	5.4
25	18.4	8.2

a) Con el modelo

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, 25,$$

calcule los estimados de los mínimos cuadrados de β_0 y β_1 .

b) Construya una tabla de análisis de varianza en la cual aparezcan por separado el error puro y el error por falta de ajuste. Determine si la falta de ajuste es significativa al nivel de 0.05. Interprete los resultados.

-272.0	0.860
-272.5	0.935
-272.6	0.961
-272.8	0.979
-272.9	0.990

- a) Ajuste un modelo de regresión lineal.
 b) ¿Parece que la proporción de impurezas que pasan a través del helio aumenta a medida que la temperatura se acerca a -273 grados centígrados?
 c) Calcule R^2 .
 d) Con base en la información anterior, ¿parece adecuado el modelo lineal? ¿Qué información adicional necesitaría usted para responder mejor a la pregunta?

11.38 Es frecuente que se utilice el tratamiento con calor para carburar partes metálicas como los engranes. El espesor de la capa carburada se considera una característica importante del engrane que contribuye a la confiabilidad general de la parte. Debido a la naturaleza crítica de esta característica, se realiza una prueba de laboratorio para cada lote del horno. La prueba es destructiva, ya que una parte real se corta en forma transversal y se sumerge en un producto químico durante cierto tiempo. Esta prueba requiere que se efectúe un análisis del carbono sobre la superficie, tanto de la parte superior del engrane (arriba de los dientes) como de su raíz (entre los dientes). Los datos siguientes son los resultados de la prueba de análisis de carbono en 19 partes.

Tiempo de inmersión		Tiempo de inmersión	
Grado	Grado	Grado	Grado
0.58	0.013	1.17	0.021
0.66	0.016	1.17	0.019
0.66	0.015	1.17	0.021
0.66	0.016	1.20	0.025
0.66	0.015	2.00	0.025
0.66	0.016	2.00	0.026
1.00	0.014	2.20	0.024
1.17	0.021	2.20	0.025
1.17	0.018	2.20	0.024
1.17	0.019		

- a) Ajuste una regresión lineal simple que relacione el grado del análisis de carbono y en comparación con el tiempo de inmersión. Pruebe $H_0: \beta_1 = 0$.
 b) Si se rechaza la hipótesis del inciso a, determine si el modelo lineal es adecuado.

11.39 Se desea obtener un modelo de regresión que relacione la temperatura con la proporción de impurezas de una sustancia que pasa a través de helio sólido. Se lista la temperatura en grados centígrados. A continuación se presentan los datos.

Temperatura ($^{\circ}$ C)	Proporción de impurezas
-260.5	0.425
-255.7	0.224
-264.6	0.453
-265.0	0.475
-270.0	0.705

11.40 Existe interés por estudiar el efecto que tiene el tamaño de la población de varias ciudades de Estados Unidos sobre las concentraciones de ozono. Los datos consisten en la población de 1999 en millones de habitantes y en la cantidad de ozono presente por hora en partes por mil millones (ppmm). Los datos son los siguientes:

Ozono (ppmm/hora), y	Población, x
126	0.6
135	4.9
124	0.2
128	0.5
130	1.1
128	0.1
126	1.1
128	2.3
128	0.6
129	2.3

- a) Ajuste un modelo de regresión lineal que relacione la concentración de ozono con la población. Pruebe $H_0: \beta_1 = 0$ usando el método ANOVA.
 b) Haga una prueba para la falta de ajuste. Con base en los resultados de la prueba, ¿es apropiado el modelo lineal?
 c) Pruebe la hipótesis del inciso a) utilizando el cuadrado medio del error puro en la prueba F. ¿Cambian los resultados? Comente las ventajas de cada prueba.

11.41 Evaluar la deposición del nitrógeno de la atmósfera es una tarea importante del National Atmospheric Deposition Program (NADP), que está asociado con muchas instituciones. Este programa está estudiando la deposición atmosférica y su efecto sobre los cultivos agrícolas, las aguas superficiales de los bosques y otros recursos. Los óxidos del nitrógeno pueden tener efectos sobre el ozono atmosférico y la cantidad de nitrógeno puro que se encuentra en el aire que respiramos. Los datos son los siguientes:

Año	Óxido de nitrógeno
1978	0.73
1979	2.55
1980	2.90
1981	3.83
1982	2.53
1983	2.77
1984	3.93
1985	2.03
1986	4.39
1987	3.04
1988	3.41
1989	5.07
1990	3.95
1991	3.14
1992	3.44
1993	3.63
1994	4.50
1995	3.95
1996	5.24
1997	3.30
1998	4.36
1999	3.33

- Grafique los datos.
- Ajuste un modelo de regresión lineal y calcule R^2 .
- ¿Qué puede decir acerca de la tendencia del óxido de nitrógeno con el paso del tiempo?

11.42 Para una variedad particular de planta los investigadores desean desarrollar una fórmula para predecir la cantidad de semillas (en gramos) como una función de la densidad de las plantas. Efectuaron un estudio con cuatro niveles del factor x , el número de plantas por parcela. Se utilizaron cuatro réplicas para cada nivel de x . A continuación se muestran los datos:

Plantas por parcela, x	Cantidad de semillas, y (gramos)			
	10	12.6	11.0	12.1
20	15.3	16.1	14.9	15.6
30	17.9	18.3	18.6	17.8
40	19.2	19.6	18.9	20.0

¿Es adecuado un modelo de regresión lineal simple para analizar este conjunto de datos?

11.10 Gráficas de datos y transformaciones

En este capítulo se estudia la construcción de modelos de regresión en los que hay una variable independiente o regresora. Además, se supone que durante la construcción del modelo tanto x como y entran en el modelo en *forma lineal*. Con frecuencia es aconsejable trabajar con un modelo alternativo en el que x o y (o ambas) intervengan en una forma no lineal. Se podría recomendar una **transformación** de los datos debido a consideraciones teóricas inherentes al estudio científico, o bien, una simple graficación de los datos podría sugerir la necesidad de *reexpresar* las variables en el modelo. La necesidad de llevar a cabo una transformación es muy fácil de diagnosticar en el caso de la regresión lineal simple, ya que las gráficas en dos dimensiones brindan un panorama verdadero de la manera en que las variables se comportan en el modelo.

Un modelo en el que x o y se transforman no debería considerarse como un *modelo de regresión no lineal*. Por lo general denominamos a un modelo de regresión como lineal cuando es **lineal en los parámetros**. En otras palabras, suponga que el aspecto de los datos u otra información científica sugiere que debe hacerse la **regresión de y^* en comparación con la de x^*** , donde cada una de ellas es una transformación de las variables naturales x y y . Entonces, el modelo de la forma

$$y_i^* = \beta_0 + \beta_1 x_i^* + \epsilon_i$$

es lineal porque lo es en los parámetros β_0 y β_1 . El material que se estudió en las secciones 11.2 a 11.9 permanece sin cambio, donde y_i^* y x_i^* reemplazan a y_i y x_i . Un ejemplo sencillo y útil es el modelo log-log:

$$\log y_i = \beta_0 + \beta_1 \log x_i + \epsilon_i.$$

Aunque este modelo es no lineal en x y y , sí lo es en los parámetros y por ello recibe el tratamiento de un modelo lineal. Por otro lado, un ejemplo de modelo verdaderamente no lineal es:

$$y_i = \beta_0 + \beta_1 x^{\beta_2} + \epsilon_i,$$

donde se debe estimar el parámetro β_2 así como β_0 y β_1 . El modelo es no lineal en β_2 .

Las transformaciones susceptibles de mejorar el ajuste y la capacidad de predicción de un modelo son muy numerosas. Para un análisis completo de las transformaciones el lector podría consultar a Myers (1990, véase la bibliografía). Decidimos incluir aquí algunas de ellas y mostrar la apariencia de las gráficas que sirven como herramientas diagnósticas. Considere la tabla 11.6, donde se presentan varias funciones que describen relaciones entre y y x que pueden producir una *regresión lineal* por medio de la transformación indicada. Además, en aras de que el análisis sea más exhaustivo, se presentan al lector las variables dependiente e independiente que se utilizan en la *regresión lineal simple* resultante. La figura 11.19 ilustra las funciones que se listan en la tabla 11.6, las cuales sirven como guía para el analista en la elección de una transformación a partir de la observación de la gráfica de y contra x .

Tabla 11.6: Algunas transformaciones útiles para linealizar

Forma funcional que relaciona y con x	Transformación propia	Forma de la regresión lineal simple
Exponencial: $y = \beta_0 e^{\beta_1 x}$	$y^* = \ln y$	Hacer la regresión de y^* contra x
Potencia: $y = \beta_0 x^{\beta_1}$	$y^* = \log y$; $x^* = \log x$	Hacer la regresión de y^* contra x^*
Recíproca: $y = \beta_0 + \beta_1 \left(\frac{1}{x}\right)$	$x^* = \frac{1}{x}$	Hacer la regresión de y contra x^*
Hiperbólica: $y = \frac{x}{\beta_0 + \beta_1 x}$	$y^* = \frac{1}{y}$; $x^* = \frac{1}{x}$	Hacer la regresión de y^* contra x^*

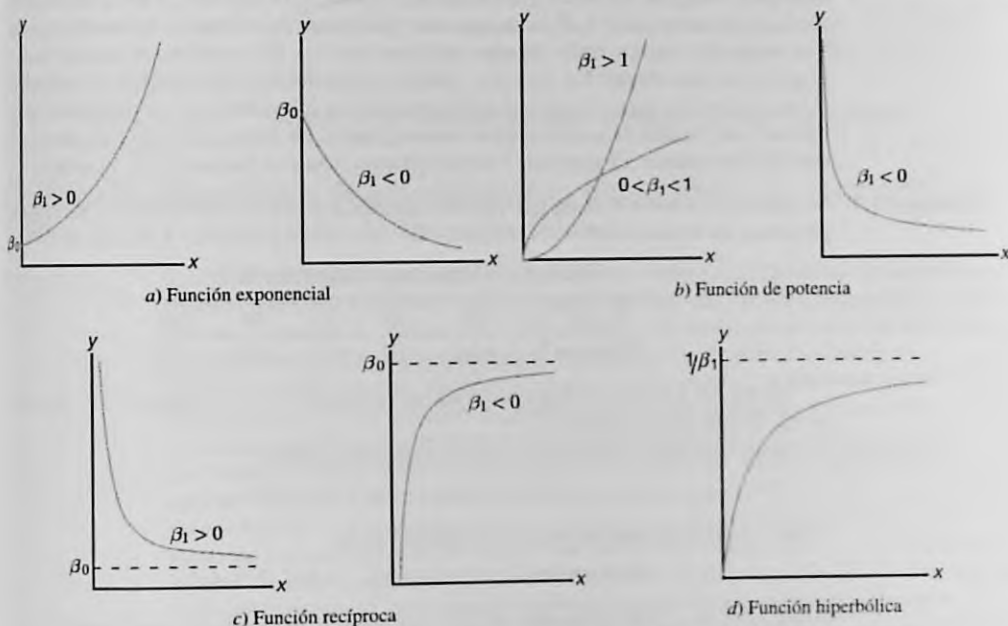


Figura 11.19: Diagramas que ilustran las funciones listadas en la tabla 11.6.

¿Cuáles son las implicaciones de un modelo transformado?

Lo que sigue intenta ser una ayuda para el analista cuando es evidente que una transformación producirá una mejoría. Sin embargo, antes de dar un ejemplo hay que mencionar dos puntos importantes. El primero tiene que ver con la escritura formal del modelo una vez que se hayan transformado los datos. Con mucha frecuencia el analista no piensa en esto y simplemente lleva a cabo la transformación sin preocuparse por la forma del modelo *antes ni después* de la transformación. El modelo exponencial sirve como una buena ilustración de esto. El modelo en las variables naturales (no transformadas) que produce un *modelo de error aditivo* en las variables transformadas es dado por

$$y_i = \beta_0 e^{\beta_1 x_i} + \epsilon_i,$$

que es un *modelo de error multiplicativo*. Al aplicar logaritmos es claro que se obtiene

$$\ln y_i = \ln \beta_0 + \beta_1 x_i + \ln \epsilon_i.$$

Como resultado, las suposiciones básicas se efectúan sobre $\ln \epsilon_i$. El propósito de esta presentación sólo es recordar al lector que no debemos considerar una transformación tan sólo como una manipulación algebraica a la cual se suma un error. Con frecuencia, un modelo en las variables transformadas que tiene una adecuada *estructura de error aditivo* es resultado de un modelo en las variables naturales con un tipo de estructura de error diferente.

El segundo aspecto importante se refiere a la noción de las medidas de mejoría. Las medidas evidentes de comparación son, por supuesto, el valor de R^2 y el cuadrado medio de los residuales s^2 . (En el capítulo 12 se estudian otras medidas de rendimiento que se usan para comparar modelos que compiten). Ahora, si la respuesta y no se transforma, entonces es claro que s^2 y R^2 se pueden usar para medir la utilidad de la transformación. Los residuales estarán en las mismas unidades para los dos modelos, el transformado y el que no se transformó. No obstante, cuando se transforma y los criterios de rendimiento para el modelo transformado deberían basarse en los valores de los residuales en las unidades de medida de la respuesta no transformada. De esta manera las comparaciones son más apropiadas. El siguiente ejemplo proporciona una ilustración de lo anterior.

Ejemplo 11.9: Se registra la presión P de un gas que corresponde a distintos volúmenes V y los datos se presentan en la tabla 11.7.

Tabla 11.7: Datos para el ejemplo 11.9

V (cm ³)	50	60	70	90	100
P (kg/cm ²)	64.7	51.3	40.5	25.9	7.8

La ley del gas ideal es dada por la forma funcional $PV^\gamma = C$, donde γ y C son constantes. Estime las constantes C y γ .

Solución: Se toman logaritmos naturales en ambos lados del modelo

$$P_i V_i^\gamma = C + \epsilon_i, \quad i = 1, 2, 3, 4, 5.$$

Como resultado, es posible escribir el modelo lineal

$$\ln P_i = \ln C - \gamma \ln V_i + \epsilon_i^*, \quad i = 1, 2, 3, 4, 5,$$

Donde $\epsilon_i^* = \ln \epsilon_i$. Los siguientes son los resultados de la regresión lineal simple:

$$\text{Intersección } \widehat{\ln C} = 14.7589, \quad \widehat{C} = 2,568,862.88, \quad \text{Pendiente: } \widehat{\gamma} = 2.65347221.$$

La siguiente tabla representa información tomada del análisis de regresión.

P_i	V_i	$\ln P_i$	$\ln V_i$	$\widehat{\ln P_i}$	$\widehat{P_i}$	$e_i = P_i - \widehat{P_i}$
64.7	50	4.16976	3.91202	4.37853	79.7	-15.0
51.3	60	3.93769	4.09434	3.89474	49.1	2.2
40.5	70	3.70130	4.24850	3.48571	32.6	7.9
25.9	90	3.25424	4.49981	2.81885	16.8	9.1
7.8	100	2.05412	4.60517	2.53921	12.7	-4.9

Resulta aleccionador graficar los datos y la ecuación de regresión. En la figura 11.20 se presenta una gráfica de los datos no transformados de presión y volumen; en tanto que la curva representa la ecuación de regresión.

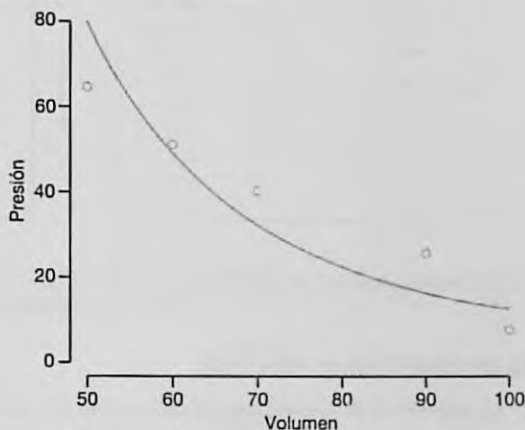


Figura 11.20: Datos de presión y volumen y la regresión ajustada.

Gráficas de diagnóstico de los residuales: detección gráfica de la transgresión de las suposiciones

Las gráficas de los datos brutos pueden ser muy útiles para determinar la naturaleza del modelo que debe ajustarse a ellos cuando sólo hay una variable independiente. En lo anterior tratamos de ilustrar esto. Sin embargo, la detección de la forma del modelo adecuado no es el único beneficio que se obtiene con la gráfica de diagnóstico. Como ocurre con gran parte del material asociado con las pruebas de hipótesis que se expone en el capítulo 10, los métodos de graficación ilustran y detectan la transgresión de las suposiciones. El lector debería recordar que muchos de los conceptos que se ilustran en este capítulo requieren suposiciones sobre los errores del modelo, las ϵ_i . De hecho, suponemos que las ϵ_i son variables aleatorias independientes $N(0, \sigma)$. Por supuesto, las ϵ_i no se observan. Sin embargo, las $e_i = y_i - \hat{y}_i$, los *residuales*, corresponden al error en el ajuste de la recta de regresión, por lo que sirven para imitar a las ϵ_i . Así, la apariencia general de estos residuales con frecuencia puede resaltar las dificultades. De manera ideal, por supuesto, la gráfica de los residuales es como la que se aprecia en la figura 11.21. Es decir, los residuales deberían demostrar en verdad fluctuaciones aleatorias alrededor del valor de cero.

Varianza no homogénea

Una suposición importante que se hace en el análisis de regresión es la varianza homogénea. A menudo las transgresiones se detectan mediante la apariencia de la gráfica de los residuales. Es común que en los datos científicos se incremente la varianza del error con el aumento de la variable regresora. Una varianza grande del error produce residuales grandes y, por ende, una gráfica de residuales como la que se presenta en la figura 11.22 es una señal de varianza no homogénea. En el capítulo 12, en el cual se expone la regresión lineal múltiple, se presenta un análisis más amplio acerca de las gráficas de los residuales e información acerca de los diferentes tipos de residuales.

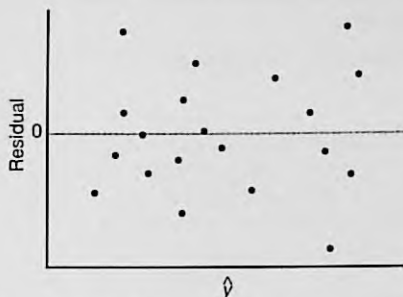


Figura 11.21: Gráfica ideal de los residuales.

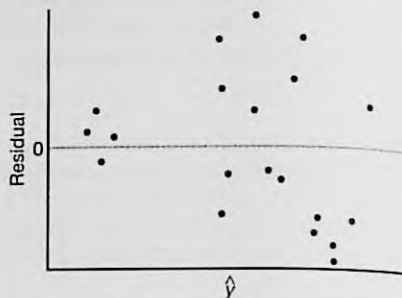


Figura 11.22: Gráfica de los residuales que ilustra una varianza heterogénea del error.

Gráfica de la probabilidad normal

La suposición de que los errores del modelo son normales se hace cuando el analista de los datos se ocupa de las pruebas de hipótesis o de la estimación de intervalos de confianza. De nuevo, los equivalentes numéricos de los ϵ_i , es decir, los residuales, son sujetos de diagnóstico mediante la graficación para detectar cualesquiera transgresiones extremas. En el capítulo 8 se presentaron las gráficas normales cuantil-cuantil y se analizaron en forma breve las de probabilidad normal. En el estudio de caso que se presenta en la siguiente sección se ilustran estas gráficas de residuales.

11.11 Estudio de caso de regresión lineal simple

En la fabricación de productos comerciales de madera es importante estimar la relación que hay entre la densidad de un producto de madera y su rigidez. Se está considerando un tipo relativamente nuevo de aglomerado que se puede formar con mucha mayor facilidad que el producto comercial ya aceptado. Es necesario saber a qué densidad su rigidez es comparable con la del producto comercial bien conocido y documentado. Terrance E. Connors realizó un estudio titulado *Investigation of Certain Mechanical Properties of a Wood-Foam Composite* (Tesis para el doctorado, Departamento de Bosques y Vida Silvestre, University of Massachusetts). Se produjeron 30 tableros de aglomerado con densidades que variaban aproximadamente de 8 a 26 libras por pie cúbico y se midió su rigidez en libras por pulgada cuadrada. En la tabla 11.8 se presentan los datos.

Es necesario que el analista de datos se concentre en un ajuste apropiado para los datos y que utilice los métodos de inferencia que se estudian en este capítulo. Tal vez lo más apropiado sea una prueba de hipótesis sobre la pendiente de la regresión, así como

la estimación de los intervalos de confianza o de predicción. Se comenzará presentando un simple diagrama de dispersión de los datos brutos con una regresión lineal simple sobrepuesta. En la figura 11.23 se observa dicha gráfica.

El ajuste de regresión lineal simple a los datos produce el modelo ajustado

$$\hat{y} = -25,433.739 + 3884.976x \quad (R^2 = 0.7975),$$

Tabla 11.8: Densidad y rigidez de 30 tableros de aglomerado

Densidad, x	Rigidez, y	Densidad, x	Rigidez, y
9.50	14,814.00	8.40	17,502.00
9.80	14,007.00	11.00	19,443.00
8.30	7573.00	9.90	14,191.00
8.60	9714.00	6.40	8076.00
7.00	5304.00	8.20	10,728.00
17.40	43,243.00	15.00	25,319.00
15.20	28,028.00	16.40	41,792.00
16.70	49,499.00	15.40	25,312.00
15.00	26,222.00	14.50	22,148.00
14.80	26,751.00	13.60	18,036.00
25.60	96,305.00	23.40	104,170.00
24.40	72,594.00	23.30	49,512.00
19.50	32,207.00	21.20	48,218.00
22.80	70,453.00	21.70	47,661.00
19.80	38,138.00	21.30	53,045.00

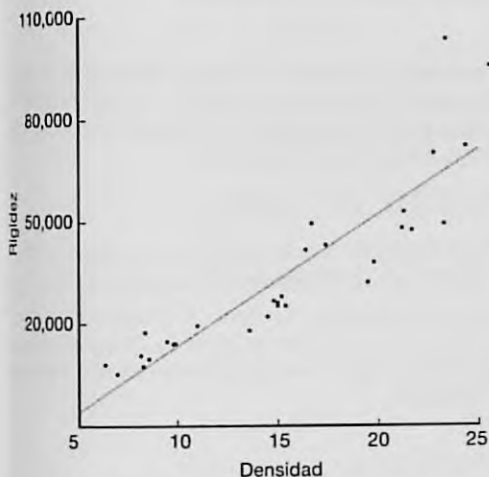


Figura 11.23: Diagrama de dispersión de los datos de densidad de la madera.

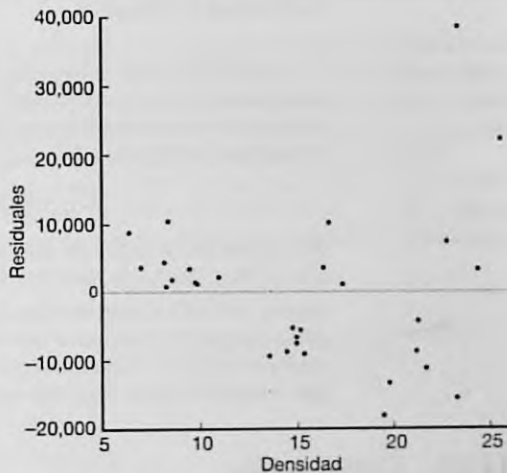


Figura 11.24: Gráfica de los residuales para los datos de densidad de la madera.

y se calcularon los residuales. En la figura 11.24 se presentan los residuales graficados contra las mediciones de la densidad. Difícilmente se trata de un conjunto de residuales ideal o satisfactorio, pues no muestran una distribución aleatoria alrededor del valor de cero. En realidad, los agrupamientos de valores positivos y negativos sugerirían que se debe investigar una tendencia curvilínea en los datos.

Para darnos una idea respecto a la suposición de error normal se dibujó una gráfica de probabilidad normal de los residuales. Es el tipo de gráfica que estudiamos en la sección 8.8, donde el eje horizontal representa la función de distribución normal empírica en una escala que produce una gráfica con línea recta cuando se grafica contra los residuales. En la figura 11.25 se presenta la gráfica de probabilidad normal de los residuales. Esta gráfica no refleja la apariencia de recta que a uno le gustaría ver, lo cual es otro síntoma de una selección errónea, quizá sobresimplificada, de un modelo de regresión.

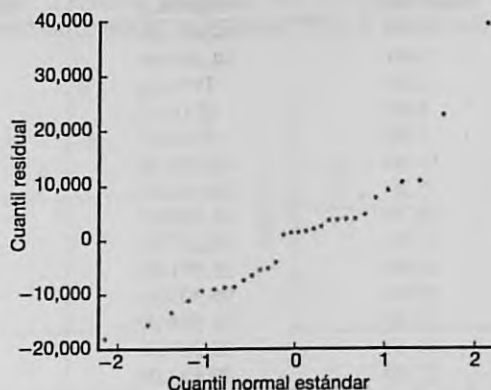


Figura 11.25: Gráfica de probabilidad normal de los residuales para los datos de densidad de la madera.

Los dos tipos de gráficas de residuales y, de hecho, el propio diagrama de dispersión, sugieren que sería adecuado un modelo algo más complicado. Una posibilidad es usar un modelo con transformación de logaritmos naturales. En otras palabras, hay que elegir hacer la regresión de $\ln y$ contra x . Esto produce la regresión

$$\widehat{\ln y} = 8.257 + 0.125x \quad (R^2 = 0.9016).$$

Para darse una idea de si el modelo transformado es más apropiado considere las figuras 11.26 y 11.27, que muestran las gráficas de los residuales de la rigidez [es decir, y -antilog ($\widehat{\ln y}$)] en comparación con las de la densidad. La figura 11.26 parece más cercana a un patrón aleatorio alrededor del cero, en tanto que la figura 11.27 con seguridad se acerca más a una línea recta. Esto, además de un valor de R^2 más elevado, sugeriría que el modelo transformado es más apropiado.

11.12 Correlación

Hasta este momento se ha supuesto que la variable regresora independiente x es una variable científica o física en lugar de una variable aleatoria. De hecho, en este contexto es frecuente que x se denomine **variable matemática**, la cual, en el proceso de muestreo, se mide con un error despreciable. En muchas aplicaciones de las técnicas de regresión es más realista suponer que tanto X como Y son variables aleatorias y que las mediciones $\{(x_i, y_i); i = 1, 2, \dots, n\}$ son observaciones de una población que tiene la función de

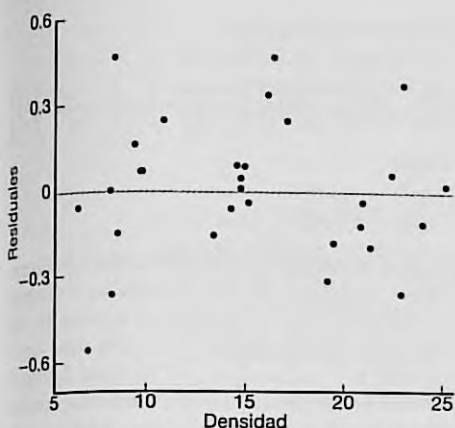


Figura 11.26: Gráfica de residuos donde se utiliza una transformación logarítmica para los datos de densidad de la madera.

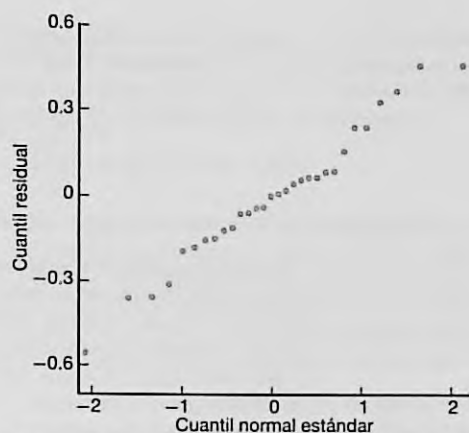


Figura 11.27: Gráfica de probabilidad normal de residuos en la cual se utiliza una transformación logarítmica para los datos de densidad de la madera.

densidad conjunta $f(x, y)$. Debemos considerar el problema de medir la relación entre las dos variables X y Y . Por ejemplo, si X y Y representaran la longitud y la circunferencia de una clase particular de hueso en el cuerpo de un adulto, podríamos realizar un estudio antropológico para determinar si los valores grandes de X se asocian con valores grandes de Y , y viceversa.

Por otro lado, si X representa la antigüedad de un automóvil usado y Y representa su precio de lista al menudeo, se esperaría que los valores grandes de X correspondan a valores pequeños de Y y que los valores pequeños de X correspondan a valores grandes de Y . El **análisis de correlación** intenta medir la fuerza de tales relaciones entre dos variables por medio de un solo número denominado **coeficiente de correlación**.

En teoría, con frecuencia se supone que la distribución condicional $f(y|x)$ de Y , para valores fijos de X , es normal con media $\mu_{y|x} = \beta_0 + \beta_1 x$ y varianza $\sigma_{y|x}^2 = \sigma^2$, y que, de igual manera, X se distribuye de forma normal con media μ y varianza σ_x^2 . Entonces, la densidad conjunta de X y Y es

$$\begin{aligned} f(x, y) &= n(y|x; \beta_0 + \beta_1 x, \sigma) n(x; \mu_x, \sigma_x) \\ &= \frac{1}{2\pi\sigma_x\sigma} \exp \left\{ -\frac{1}{2} \left[\left(\frac{y - \beta_0 - \beta_1 x}{\sigma} \right)^2 + \left(\frac{x - \mu_x}{\sigma_x} \right)^2 \right] \right\}, \end{aligned}$$

para $-\infty < x < \infty$ y $-\infty < y < \infty$.

Escribamos la variable aleatoria Y en la forma

$$Y = \beta_0 + \beta_1 X + \epsilon,$$

donde ahora X es una variable aleatoria independiente del error aleatorio ϵ . Como la media del error aleatorio ϵ es cero, se deduce que

$$\mu_Y = \beta_0 + \beta_1 \mu_X \quad \text{y} \quad \sigma_Y^2 = \sigma^2 + \beta_1^2 \sigma_X^2.$$

Al sustituir para α y σ^2 en la expresión anterior para $f(x, y)$, se obtiene la **distribución normal bivariada**

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}$$

para $-\infty < x < \infty$ y $-\infty < y < \infty$, donde

$$\rho^2 = 1 - \frac{\sigma^2}{\sigma_Y^2} = \beta_1^2 \frac{\sigma_X^2}{\sigma_Y^2}.$$

La constante ρ (ro) se denomina **coeficiente de correlación de la población** y desempeña un papel importante en muchos problemas de análisis de datos bivariados. Es importante que el lector entienda la interpretación física de este coeficiente de correlación, así como la diferencia entre correlación y regresión. El término *regresión* aún tiene algún significado aquí. De hecho, la línea recta dada por $\mu_{Y|x} = \beta_0 + \beta_1 x$ se sigue llamando recta de regresión, igual que antes, y los estimadores de β_0 y β_1 son idénticos a los que se presentaron en la sección 11.3. El valor de ρ es 0 cuando $\beta_1 = 0$, que resulta cuando en esencia no existe regresión lineal; es decir, cuando la recta de regresión es horizontal y cualquier conocimiento de X es inútil para predecir Y . Como $\sigma_Y^2 \geq \sigma^2$, se debe tener $\rho^2 \leq 1$ y, por lo tanto, $-1 \leq \rho \leq 1$. Los valores de $\rho \pm 1$ sólo ocurren cuando $\sigma^2 = 0$, en cuyo caso se tiene una relación lineal perfecta entre las dos variables. Así, un valor de ρ igual a $+1$ implica una relación lineal perfecta con pendiente positiva, en tanto que un valor de ρ igual a -1 resulta de una relación lineal perfecta con pendiente negativa. Entonces, se podría decir que los estimadores muestrales de ρ con magnitud cercana a la unidad implican una buena correlación o **asociación lineal** entre X y Y , mientras que valores cercanos a cero indican poca o ninguna correlación.

Para obtener un estimador muestral de ρ recordemos que en la sección 11.4 aprendimos que la suma de los cuadrados del error es

$$SCE = S_{yy} - b_1 S_{xy}.$$

Al dividir ambos lados de esta ecuación entre S_{yy} y reemplazar S_{xy} con $b_1 S_{xx}$, se obtiene la relación

$$b_1^2 \frac{S_{xx}}{S_{yy}} = 1 - \frac{SCE}{S_{yy}}.$$

El valor de $b_1^2 S_{xx}/S_{yy}$ es igual a cero cuando $b_1 = 0$, lo que ocurrirá cuando los puntos muestrales no tengan relación lineal. Como $S_{yy} \geq SCE$, se concluye que $b_1^2 S_{xx}/S_{yy}$ debe estar entre 0 y 1. En consecuencia, $b_1 \sqrt{S_{xx}/S_{yy}}$ debe variar entre -1 y $+1$, y los valores negativos corresponden a rectas con pendientes negativas, mientras que los valores positivos corresponden a rectas con pendientes positivas. Un valor de -1 o $+1$ sucederá cuando $SCE = 0$, pero éste es el caso en el que todos los puntos muestrales caen sobre una línea recta. Por lo tanto, una relación lineal perfecta se da en los datos muestrales cuando $b_1 \sqrt{S_{xx}/S_{yy}} = \pm 1$. Es claro que la cantidad $b_1 \sqrt{S_{xx}/S_{yy}}$, la cual se designará de aquí en adelante como r , se puede usar como un estimado del coeficiente de correlación ρ de la población. Se acostumbra hacer referencia al estimado r como **coeficiente de correlación producto-momento de Pearson**, o sólo como **coeficiente de correlación muestral**.

Coficiente de correlación La medida ρ de la asociación lineal entre dos variables X y Y se estima por medio del **coeficiente de correlación muestral** r , donde

$$r = b_1 \sqrt{\frac{S_{xx}}{S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}.$$

Hay que tener cuidado en la interpretación de valores de r entre -1 y $+1$. Por ejemplo, valores de r iguales a 0.3 y 0.6 significan sólo que hay dos correlaciones positivas, una un poco más fuerte que la otra. Sería un error concluir que $r = 0.6$ indica una relación lineal dos veces mejor que la del valor $r = 0.3$. Por otro lado, si escribimos

$$r^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}} = \frac{SCR}{S_{yy}},$$

entonces r^2 , que por lo general se denomina **coeficiente muestral de determinación**, representa la proporción de la variación de S_{yy} explicada por la regresión de Y sobre x , a saber, la SCR . Es decir, r^2 expresa la proporción de la variación total de los valores de la variable Y que son ocasionados o explicados por una relación lineal con los valores de la variable aleatoria X . Así, una correlación de 0.6 significa que 0.36 , o 36% , de la variación total de los valores de Y en la muestra se explica mediante la relación lineal con los valores de X .

Ejemplo 11.10: Es importante que los investigadores científicos del área de productos forestales sean capaces de estudiar la correlación entre la anatomía y las propiedades mecánicas de los árboles. Para el estudio *Quantitative Anatomical Characteristics of Plantation Grown Loblolly Pine (Pinus Taeda L.) and Cottonwood (Populus deltoides Bart. Ex Marsh.) and Their Relationships to Mechanical Properties*, realizado por el Departamento de Bosques y Productos Forestales de Virginia Tech, se seleccionaron al azar 29 pinos de Arkansas para investigarlos. En la tabla 11.9 se presentan los datos resultantes sobre la gravedad específica en gramos/cm^3 y el módulo de ruptura en kilopascales (kPa). Calcule e interprete el coeficiente de correlación muestral.

Tabla 11.9: Datos de 29 pinos de Arkansas para el ejemplo 11.10

Gravedad específica, x (g/cm^3)	Módulo de ruptura, y (kPa)	Gravedad específica, x (g/cm^3)	Módulo de ruptura, y (kPa)
0.414	29,186	0.581	85,156
0.383	29,266	0.557	69,571
0.399	26,215	0.550	84,160
0.402	30,162	0.531	73,466
0.442	38,867	0.550	78,610
0.422	37,831	0.556	67,657
0.466	44,576	0.523	74,017
0.500	46,097	0.602	87,291
0.514	59,698	0.569	86,836
0.530	67,705	0.544	82,540
0.569	66,088	0.557	81,699
0.558	78,486	0.530	82,096
0.577	89,869	0.547	75,657
0.572	77,369	0.585	80,490
0.548	67,095		

Solución: A partir de los datos se encuentra que

$$S_{xx} = 0.11273, \quad S_{yy} = 11,807,324,805, \quad S_{xy} = 34,422.27572.$$

Por lo tanto,

$$r = \frac{34,422.27572}{\sqrt{(0.11273)(11,807,324,805)}} = 0.9435.$$

Un coeficiente de correlación de 0.9435 indica una buena relación lineal entre X y Y . Como $r^2 = 0.8902$, se puede decir que aproximadamente 89% de la variación de los valores de Y es ocasionada por una relación lineal con X . \blacksquare

Una prueba de la hipótesis especial $\rho = 0$ en comparación con una alternativa apropiada es equivalente a probar $\beta_1 = 0$ para el modelo de regresión lineal simple y, por lo tanto, son aplicables los procedimientos de la sección 11.8, donde se usaba la distribución t con $n - 2$ grados de libertad o la distribución F con 1 y $n - 2$ grados de libertad. Sin embargo, si se desea evitar el procedimiento del análisis de varianza y tan sólo calcular el coeficiente de correlación muestral, se podría verificar (véase el ejercicio de repaso 11.66 en la página 438) que el valor t

$$t = \frac{b_1}{s/\sqrt{S_{xx}}}$$

también se puede escribir como

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}},$$

que, como antes, es un valor del estadístico T que tiene una distribución t con $n - 2$ grados de libertad.

Ejemplo 11.11: Para los datos del ejemplo 11.10 pruebe la hipótesis de que no existe asociación lineal entre las variables.

Solución: 1. $H_0: \rho = 0$.

2. $H_1: \rho \neq 0$.

3. $\alpha = 0.05$.

4. Región crítica: $t < -2.052$ o $t > 2.052$.

5. Cálculos: $t = \frac{0.9435 \sqrt{27}}{\sqrt{1-0.9435^2}} = 14.79$, $P < 0.0001$.

6. Decisión: Rechazar la hipótesis de que no existe asociación lineal. \blacksquare

A partir de la información muestral es fácil efectuar una prueba de la hipótesis más general de que $\rho = \rho_0$ en comparación con una hipótesis alternativa adecuada. Si X y Y siguen una distribución normal bivariada, la cantidad

$$\frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

es el valor de una variable aleatoria que sigue aproximadamente la distribución normal con media $\frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0}$ y varianza $1/(n-3)$. Entonces, el procedimiento de prueba consiste en calcular

$$z = \frac{\sqrt{n-3}}{2} \left[\ln \left(\frac{1+r}{1-r} \right) - \ln \left(\frac{1+\rho_0}{1-\rho_0} \right) \right] = \frac{\sqrt{n-3}}{2} \ln \left[\frac{(1+r)(1-\rho_0)}{(1-r)(1+\rho_0)} \right]$$

y compararlo con los puntos críticos de la distribución normal estándar.

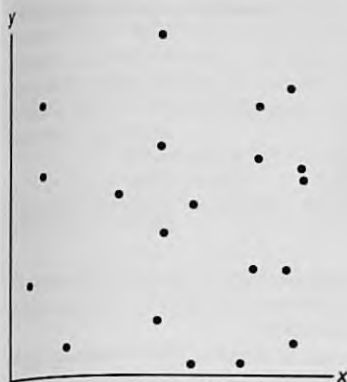
Ejemplo 11.12: Para los datos del ejemplo 11.10 pruebe la hipótesis nula de que $\rho = 0.9$ en comparación con la alternativa de que $\rho > 0.9$. Utilice un nivel de significancia de 0.05.

Solución: 1. $H_0: \rho = 0.9$.

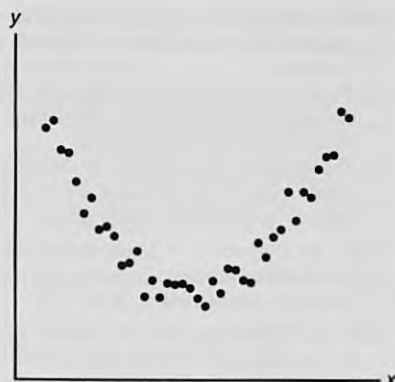
2. $H_1: \rho > 0.9$.

3. $\alpha = 0.05$.

4. Región crítica: $z > 1.645$.



a) Sin asociación



b) Relación causal

Figura 11.28: Diagrama de dispersión que muestra correlación de cero.

5. Cálculos:

$$z = \frac{\sqrt{26}}{2} \ln \left[\frac{(1 + 0.9435)(0.1)}{(1 - 0.9435)(1.9)} \right] = 1.51, \quad P = 0.0655.$$

6. Decisión: Existe con certeza alguna evidencia de que el coeficiente de correlación no excede a 0.9. ■

Debe precisarse que en los estudios de correlación, como en los problemas de regresión lineal, los resultados obtenidos sólo son tan buenos como el modelo que se adopte. En las técnicas de correlación estudiadas aquí se supone que las variables X y Y tienen una densidad normal bivariada, con el valor medio de Y para cada valor de x relacionado en forma lineal con x . Con frecuencia es útil elaborar una gráfica preliminar de los datos experimentales para observar qué tan adecuada es la suposición de linealidad. Un valor del coeficiente de correlación muestral cercano a cero resultará de datos que muestren un efecto estrictamente aleatorio, como los de la figura 11.28a, lo que implica que hay poca o ninguna relación causal. Es importante recordar que el coeficiente de correlación entre dos variables es una medida de su relación lineal, y que un valor de $r = 0$ implica *falta de linealidad* y *no falta de asociación*. Por lo tanto, si existiera una relación cuadrática fuerte entre X y Y , como la que se observa en la figura 11.28b, aún se podría obtener una correlación de cero que indicaría una relación no lineal.

Ejercicios

11.43 Calcule e interprete el coeficiente de correlación para las siguientes calificaciones de 6 estudiantes seleccionados al azar:

Calificación en matemáticas	70	92	80	74	65	83
Calificación en inglés	74	84	63	87	78	90

11.44 Remítase al ejercicio 11.1 de la página 398 y suponga que x y y son variables aleatorias con una distribución normal bivariada:

- Calcule r .
- Pruebe la hipótesis de que $\rho = 0$ en comparación con la alternativa de que $\rho \neq 0$ a un nivel de significancia de 0.05.

11.45 Remítase al ejercicio 11.13 de la página 400, suponga una distribución normal bivariada para x y y .

- Calcule r .
- Pruebe la hipótesis nula de que $\rho = -0.5$, en comparación con la alternativa de que $\rho < -0.5$, a un nivel de significancia de 0.025.
- Determine el porcentaje de la variación en la cantidad de partículas eliminadas que se debe a cambios en la cantidad de lluvia diaria.

11.46 En el ejercicio 11.43 pruebe la hipótesis de que $\rho = 0$ en comparación con la alternativa de que $\rho \neq 0$. Utilice un nivel de significancia de 0.05.

11.47 Los datos siguientes se obtuvieron en un estudio de la relación entre el peso y el tamaño del pecho de niños al momento de nacer.

Peso (kg)	Tamaño del pecho (cm)
2.75	29.5
2.15	26.3
4.41	32.2
5.52	36.5
3.21	27.2
4.32	27.7
2.31	28.3
4.30	30.3
3.71	28.7

- Calcule r .
- Pruebe la hipótesis nula de que $\rho = 0$ en comparación con la alternativa de que $\rho > 0$ a un nivel de significancia de 0.01.
- ¿Qué porcentaje de la variación del tamaño del pecho de los niños es explicado por la diferencia de peso?

Ejercicios de repaso

11.48 Remítase al ejercicio 11.8 de la página 399 y construya

- un intervalo de confianza de 95% para la calificación promedio en el curso de los estudiantes que obtuvieron 35 puntos en el examen de colocación;
- un intervalo de predicción de 95% para la calificación del curso de un estudiante que obtuvo 35 puntos en el examen de colocación.

11.49 El Centro de Consulta Estadística de Virginia Tech analizó datos sobre las marmotas normales para el Departamento de Veterinaria. Las variables de interés fueron el peso corporal en gramos y el peso del corazón en gramos. Se deseaba desarrollar una ecuación de regresión lineal con el fin de determinar si había una relación lineal significativa entre el peso del corazón y el peso total del cuerpo.

Peso corporal (gramos) Peso del corazón (gramos)

4050	11.2
2465	12.4
3120	10.5
5700	13.2
2595	9.8
3640	11.0
2050	10.8
4235	10.4
2935	12.2
4975	11.2
3690	10.8
2800	14.2
2775	12.2
2170	10.0
2370	12.3
2055	12.5
2025	11.8
2645	16.0
2675	13.8

Utilice el peso del corazón como la variable independiente, el peso del cuerpo como la dependiente y haga un ajuste de regresión lineal simple con los siguientes datos. Además, pruebe la hipótesis de que $H_0: \beta_1 = 0$ en comparación con $H_1: \beta_1 \neq 0$. Saque conclusiones.

11.50 A continuación se presentan las cantidades de sólidos eliminados de cierto material cuando se expone a periodos de secado de diferentes duraciones.

x (horas)	y (gramos)	
4.4	13.1	14.2
4.5	9.0	11.5
4.8	10.4	11.5
5.5	13.8	14.8
5.7	12.7	15.1
5.9	9.9	12.7
6.3	13.8	16.5
6.9	16.4	15.7
7.5	17.6	16.9
7.8	18.3	17.2

- Estime la recta de regresión lineal.
- Pruebe si es adecuado el modelo lineal a un nivel de significancia de 0.05.

11.51 Remítase al ejercicio 11.9 de la página 399 y construya

- un intervalo de confianza de 95% para las ventas semanales promedio cuando se gastan \$45 en publicidad.
- un intervalo de predicción de 95% para las ventas semanales cuando se gastan \$45 en publicidad.

11.52 Se diseñó un experimento para el Departamento de Ingeniería de Materiales de Virginia Tech con el fin de estudiar las propiedades de deterioro del nitrógeno con base en las mediciones de la presión de hidrógeno

electrolítico. Se utilizó una solución al 0.1 *N* NaOH y el material era cierto tipo de acero inoxidable. La densidad de corriente de carga catódica fue controlada y variada en cuatro niveles. Se observó la presión de hidrógeno efectiva como la respuesta. A continuación se presentan los datos.

Ensayo	Densidad de corriente de carga, x (mA/cm ²)	Presión de hidrógeno efectiva, y (atm)
1	0.5	86.1
2	0.5	92.1
3	0.5	64.7
4	0.5	74.7
5	1.5	223.6
6	1.5	202.1
7	1.5	132.9
8	2.5	413.5
9	2.5	231.5
10	2.5	466.7
11	2.5	365.3
12	3.5	493.7
13	3.5	382.3
14	3.5	447.2
15	3.5	563.8

- Efectúe un análisis de regresión lineal simple de y con x .
- Calcule la suma de cuadrados del error puro y haga una prueba para la falta de ajuste.
- ¿La información del inciso *b* indica la necesidad de un modelo en x más allá de una regresión de primer orden? Explique su respuesta.

11.53 Los datos siguientes representan la calificación en química de una muestra aleatoria de 12 estudiantes de nuevo ingreso a cierta universidad, así como sus calificaciones en una prueba de inteligencia aplicada mientras estudiaban el último año de preparatoria.

Estudiante	Calificación en la prueba, x	Calificación en química, y
1	65	85
2	50	74
3	55	76
4	65	90
5	55	85
6	70	87
7	65	94
8	70	98
9	55	81
10	70	91
11	50	76
12	55	74

- Calcule e interprete el coeficiente de correlación de la muestra.
- Establezca las suposiciones necesarias acerca de las variables aleatorias.

- Pruebe la hipótesis de que $\rho = 0.5$ en comparación con la alternativa de que $\rho > 0.5$. Use un valor P para las conclusiones.

11.54 La sección de negocios del *Washington Times* de marzo de 1997 listaba 21 diferentes computadoras e impresoras usadas, así como sus precios de lista. También se listaba la oferta promedio. En la figura 11.29 de la página 439 se presenta una parte de los resultados impresos por computadora del análisis de regresión usando el programa SAS.

- Explique la diferencia entre el intervalo de confianza sobre la media y el intervalo de predicción.
- Explique por qué los errores estándar de la predicción varían de una observación a otra.
- ¿Cuál observación tiene el menor error estándar de la predicción? Explique su respuesta.

11.55 Considere los datos de los vehículos de *Consumer Reports* que se incluyen en la figura 11.30 de la página 440. El peso se indica en toneladas, el rendimiento en millas por galón y también se incluye el coeficiente de manejo. Se ajustó un modelo de regresión que relaciona el peso x con el rendimiento y . En la figura 11.30 de la página 440 se observa una salida parcial del SAS con los resultados de dicho análisis de regresión, y en la figura 11.31 de la página 441 se incluye una gráfica de los residuales y el peso de cada vehículo.

- A partir del análisis y la gráfica de los residuales, ¿se podría concluir que cabría la posibilidad de encontrar un modelo mejorado si se usara una transformación? Explique su respuesta.
- Ajuste el modelo reemplazando el peso con el logaritmo del peso. Comente los resultados.
- Ajuste un modelo reemplazando mpg con los galones por cada 100 millas recorridas, como se reporta con frecuencia el rendimiento del combustible en otros países. ¿Cuál de los tres modelos es preferible? Explique su respuesta.

11.56 A continuación se presentan las observaciones registradas del producto de una reacción química tomadas a temperaturas diferentes:

x (°C)	y (%)	x (°C)	y (%)
150	75.4	150	77.7
150	81.2	200	84.4
200	85.5	200	85.7
250	89.0	250	89.4
250	90.5	300	94.8
300	96.7	300	95.3

- Grafique los datos.
- ¿La gráfica indica que la relación es lineal?
- Haga un análisis de regresión lineal simple y pruebe la falta de ajuste.

d) Saque conclusiones con base en el resultado del inciso c.

11.57 La prueba de acondicionamiento físico es un aspecto importante del entrenamiento atlético. Una medida común para determinar la aptitud cardiovascular es el volumen máximo de oxígeno que se inhala al realizar un ejercicio extenuante. Se realizó un estudio con 24 hombres de mediana edad para analizar cómo el tiempo que les tomaba correr una distancia de dos millas influía en el oxígeno que consumían, el cual se midió con métodos estándar de laboratorio mientras los sujetos se ejercitaban en una banda sin fin. El trabajo fue publicado en el artículo "Maximal Oxygen Intake Prediction in Young and Middle Aged Males", *Journal of Sports Medicine* 9, 1969, 17-22. A continuación se presentan los datos.

Sujeto	y, Volumen máximo de O ₂	x, Tiempo en segundos
1	42.33	918
2	53.10	805
3	42.08	892
4	50.06	962
5	42.45	968
6	42.46	907
7	47.82	770
8	49.92	743
9	36.23	1045
10	49.66	810
11	41.49	927
12	46.17	813
13	46.18	858
14	43.21	860
15	51.81	760
16	53.28	747
17	53.29	743
18	47.18	803
19	56.91	683
20	47.80	844
21	48.65	755
22	53.67	700
23	60.62	748
24	56.73	775

- a) Estime los parámetros en un modelo de regresión lineal simple.
 b) ¿El tiempo que toma correr dos millas influye de forma significativa en la cantidad máxima de oxígeno consumido? Utilice $H_0: \beta_0 = 0$ en comparación con $H_1: \beta_0 \neq 0$.
 c) Grafique los residuales en una gráfica en comparación con x y haga comentarios sobre qué tan apropiado es el modelo lineal simple.

11.58 Suponga que cierto científico postula el modelo

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, n,$$

y β_0 es un valor conocido no necesariamente igual a cero.

a) ¿Cuál es el estimador apropiado de mínimos cuadrados de β_1 ? Justifique su respuesta.

b) ¿Cuál es la varianza del estimador de la pendiente?

11.59 Para el modelo de regresión lineal simple demuestre que $E(s^2) = \sigma^2$.

11.60 Suponga que las ϵ_i son independientes y que se distribuyen normalmente con medias de cero y varianza común σ^2 , y demuestre que B_0 , el estimador de mínimos cuadrados de β_0 en $\mu_{Y|x} = \beta_0 + \beta_1 x$, se distribuye de manera normal con media β_0 y varianza

$$\sigma_{B_0}^2 = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \sigma^2.$$

11.61 Para un modelo de regresión lineal simple

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, n,$$

donde las ϵ_i son independientes y se distribuyen normalmente con medias de cero y varianzas iguales σ^2 , demuestre que

$$B_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

y tienen covarianza de cero.

11.62 Demuestre, en el caso de un ajuste de mínimos cuadrados al modelo de regresión lineal simple

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, n,$$

que $\sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n \epsilon_i = 0$.

11.63 Considere la situación del ejercicio de repaso 11.62 pero suponga que $n = 2$, es decir, que sólo disponemos de dos puntos de datos. Argumente que la recta de regresión de mínimos cuadrados tendrá como resultado $(y_1 - \hat{y}_1) = (y_2 - \hat{y}_2) = 0$. También demuestre que para este caso $R^2 = 1.0$.

11.64 En el ejercicio de repaso 11.62 se pidió al estudiante que demostrara que $\sum_{i=1}^n (y_i - \hat{y}_i) = 0$ para un modelo de regresión lineal simple estándar. ¿Se cumple también para un modelo con intersección en el origen? Demuestre su respuesta, ya sea afirmativa o negativa.

11.65 Suponga que un experimentador plantea un modelo como

$$Y_i = \beta_0 + \beta_1 x_{1i} + \epsilon_i, \quad i = 1, 2, \dots, n,$$

cundo en realidad una variable adicional, digamos x_2 también contribuye linealmente a la respuesta. Entonces, el verdadero modelo es dado por

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad i = 1, 2, \dots, n.$$

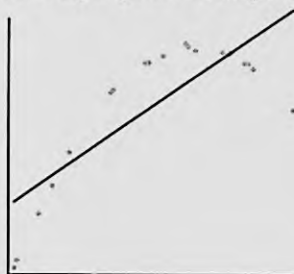
Calcule el valor esperado del estimador

$$B_1 = \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1) Y_i}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}$$

11.66 Demuestre los pasos necesarios para convertir

la ecuación $r = \frac{b_1}{s/\sqrt{S_x}}$ a la forma equivalente $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

11.67 Considere el siguiente grupo ficticio de datos, donde la línea que los atraviesa representa la recta de regresión lineal simple ajustada. Grafique los residuales.



11.68 Proyecto: Este proyecto se puede realizar en grupos o de manera individual. Cada grupo o persona debe encontrar un grupo de datos, preferiblemente de su campo de estudios, aunque también pueden ser de otro campo. Los datos se deben ajustar al esquema de regresión, con una variable de regresión x y una variable de respuesta y . Determine con cuidado cuál variable es x y cuál es y . Tal vez necesite consultar una revista científica de su campo si no cuenta con otros datos experimentales.

- Grafique y contra x . Comente sobre la relación que se observa en la gráfica.
- Diseñe un modelo de regresión adecuado a partir de los datos. Utilice una regresión lineal simple o ajuste un modelo polinomial a los datos. Comente acerca de medidas de calidad.
- Grafique los residuales como se indica en el texto. Verifique posibles violaciones de los supuestos. Muestre de forma gráfica una representación de los intervalos de confianza de una respuesta media graficada en comparación con x . Haga comentarios al respecto.

R-Square	Coeff Var	Root MSE	Price Mean					
0.967472	7.923338	70.83841	894.0476					
				Standard				
Parameter	Estimate	Error	t Value	Pr > t				
Intercept	59.93749137	38.34195754	1.56	0.1345				
Buyer	1.04731316	0.04405635	23.77	<.0001				
				PredictStd Err	Lower 95%	Upper 95%	Lower 95%	Upper 95%
product	Buyer Price	Value Predict	Mean	Mean	Predict	Predict		
IEM PS/1 486/66 420MB	325 375	400.3125.8906	346.12	454.50	242.46	558.17		
IEM ThinkPad500	450 625	531.2321.7232	485.76	576.70	376.15	686.31		
IEM Think-Dad755CX	1700 1850	1840.3742.7041	1750.99	1929.75	1667.25	2013.49		
AST Pentium90 540MB	800 875	897.7915.4590	865.43	930.14	746.03	1049.54		
Dell Pentium75 1GB	650 700	740.6916.7503	705.63	775.75	588.34	893.05		
Gateway486/75320MB	700 750	793.0616.0314	759.50	826.61	641.04	945.07		
Clone 586/1331GB	500 600	583.5920.2363	541.24	625.95	429.40	737.79		
CompaqContura4/25 120MB	450 600	531.2321.7232	485.76	576.70	376.15	686.31		
CompaqDeskproP90 1.2GB	800 850	897.7915.4590	865.43	930.14	746.03	1049.54		
MicronP75 810MB	800 675	897.7915.4590	865.43	930.14	746.03	1049.54		
MicronP100 1.2GB	900 975	1002.5216.1176	968.78	1036.25	850.46	1154.58		
Mac Quadra 840AV 500MB	450 575	531.2321.7232	485.76	576.70	376.15	686.31		
Mac Performer6116 700MB	700 775	793.0616.0314	759.50	826.61	641.04	945.07		
PowerBooks40c 320MB	1400 1500	1526.1830.7579	1461.80	1590.55	1364.54	1687.82		
PowerMac300 500MB	1350 1575	1473.8128.8747	1413.37	1534.25	1313.70	1633.92		
PowerMac 7500/1001GB	1150 1325	1264.3521.9454	1218.42	1310.28	1109.13	1419.57		
NEC Versa 486 340MB	800 900	897.7915.4590	865.43	930.14	746.03	1049.54		
Toshiba1960CS320MB	700 825	793.0616.0314	759.50	826.61	641.04	945.07		
Toshiba4800VCTS500MB	1000 1150	1107.2517.8715	1069.85	1144.66	954.34	1260.16		
HP Laser jet III	350 475	426.5025.0157	374.14	478.86	269.26	583.74		
Apple LaserWriter Pro 63	750 800	845.4215.5930	812.79	878.06	693.61	997.24		

Figura 11.29: Salida por computadora de los resultados del SAS que presenta el análisis parcial de datos del ejercicio de repaso 11.54.

Obs	Model	WT	MPG	DR_RATIO
1	BuickEstateWagon	4.360	16.9	2.73
2	FordCountrySquireWagon	4.054	15.5	2.26
3	ChevyMa libuWagon	3.605	19.2	2.56
4	ChryslerLeBaronWagon	3.940	18.5	2.45
5	Chevette	2.155	30.0	3.70
6	ToyotaCorona	2.560	27.5	3.05
7	Datsun510	2.300	27.2	3.54
8	DodgeOmni	2.230	30.9	3.37
9	Audi 5000	2.830	20.3	3.90
10	Volvo240 CL	3.140	17.0	3.50
11	Saab 99 GLE	2.795	21.6	3.77
12	Peugeot694 SL	3.410	16.2	3.58
13	BuickCenturySpecial	3.380	20.6	2.73
14	MercuryZephyr	3.070	20.8	3.08
15	DodgeAspen	3.620	18.6	2.71
16	AMC ConcordD/L	3.410	18.1	2.73
17	ChevyCapriceClassic	3.840	17.0	2.41
18	Ford LTP	3.725	17.6	2.26
19	MercuryGrandMarquis	3.955	16.5	2.26
20	Dodge St Regis	3.830	18.2	2.45
21	Ford Mustang4	2.585	26.5	3.08
22	Ford MustangGhia	2.910	21.9	3.08
23	Macda GLC	1.975	34.1	3.73
24	Dodge Colt	1.915	35.1	2.97
25	AMC Spirit	2.670	27.4	3.08
26	VW Scirocco	1.990	31.5	3.78
27	Honda AccordLX	2.135	29.5	3.05
28	Buick Skylark	2.570	28.4	2.53
29	Chevy Citation	2.595	28.8	2.69
30	Olds Omega	2.700	26.8	2.84
31	PontiacPhoenix	2.556	33.5	2.69
32	PlymouthHorizon	2.200	34.2	3.37
33	Datsun210	2.020	31.8	3.70
34	Fiat Strada	2.130	37.3	3.10
35	VW Dasher	2.190	30.5	3.70
36	Datsun810	2.815	22.0	3.70
37	BMW 320i	2.600	21.5	3.64
38	VW Rabbit	1.925	31.9	3.78
	R-Square	Coeff Var	Root MSE	MPG Mean
	0.817244	11.46010	2.837580	24.76053
			Standard	
Parameter	Estimate	Error	t Value	Pr > t
Intercept	48.67928080	1.94053995	25.09	<.0001
WT	-8.36243141	0.65908398	-12.69	<.0001

Figura 11.30: Salida de computadora de los resultados del SAS que muestra el análisis parcial de los datos del ejercicio de repaso 11.55.

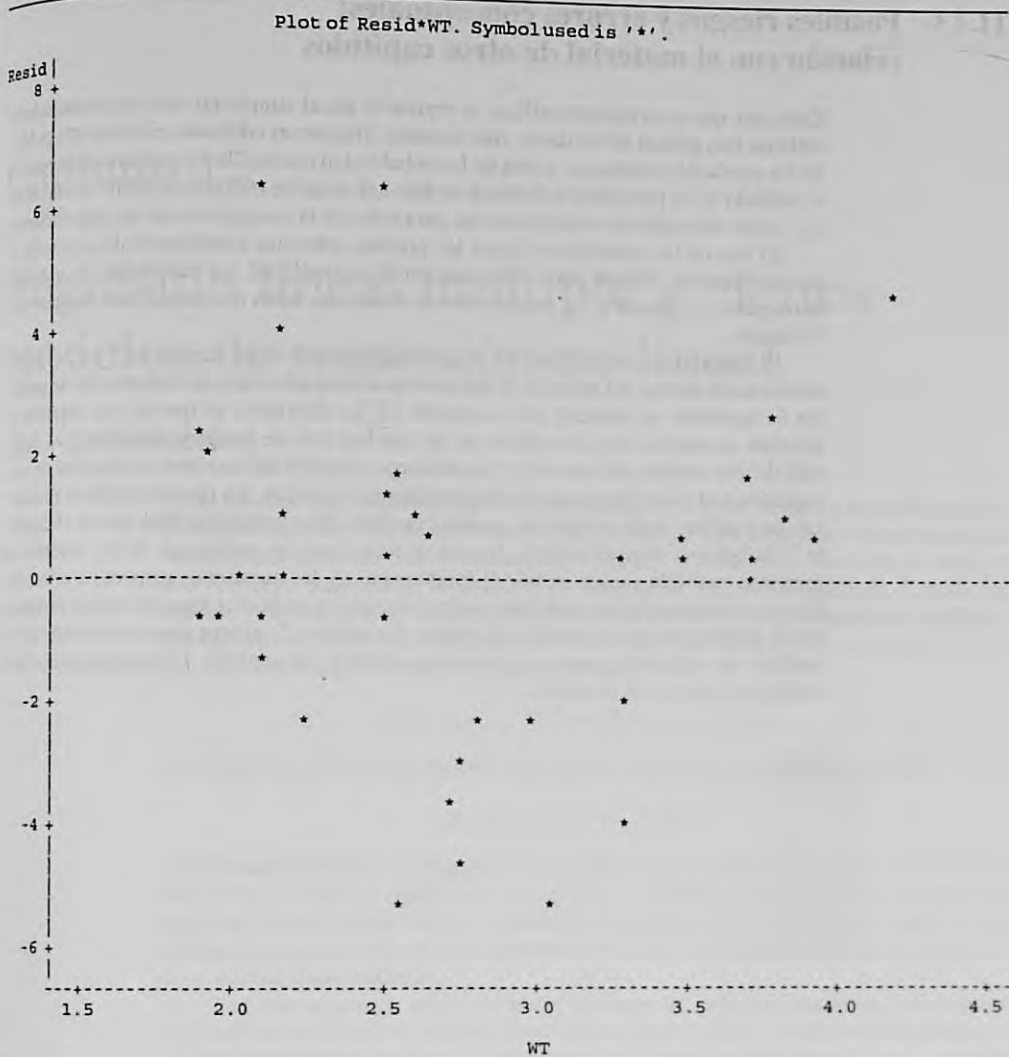


Figura 11.31: Salida de computadora de los resultados del SAS que muestra la gráfica de residuales del ejercicio de repaso 11.55.

11.13 Posibles riesgos y errores conceptuales; relación con el material de otros capítulos

Cada vez que se considere utilizar la regresión lineal simple no sólo es recomendable elaborar una gráfica de los datos, sino esencial. Siempre es edificante elaborar una gráfica de los residuales ordinarios y otra de la probabilidad normal de los mismos. Además, en el capítulo 12 se presentará e ilustrará un tipo adicional de residual en forma estandarizada. Todas esas gráficas están diseñadas para detectar la transgresión de las suposiciones.

El uso de los estadísticos t para las pruebas sobre los coeficientes de regresión es razonablemente robusto para la suposición de normalidad. La suposición de varianza homogénea es crucial y las gráficas de los residuales están diseñadas para detectar una violación.

El material de este capítulo se utiliza ampliamente en los capítulos 12 a 15. Toda la información acerca del método de los mínimos cuadrados para la elaboración de modelos de regresión se utilizará en el capítulo 12. La diferencia es que en ese capítulo se abordan las condiciones científicas en las que hay más de una sola variable x , es decir, más de una variable de regresión. Sin embargo, también utilizaremos el material de este capítulo en el que se exponen los diagnósticos de regresión, los tipos de gráficas residuales, las medidas de la calidad del modelo, etcétera. El estudiante notará que en el capítulo 12 habrá más complicaciones, lo cual se debe a que los problemas de los modelos de regresión múltiple suelen incluir el fundamento de las cuestiones respecto a cómo las diversas variables de regresión entran en el modelo, e incluso el tema de cuáles variables deben permanecer en el modelo. De hecho, el capítulo 15 incluye el uso constante de los modelos de regresión, pero en el resumen al final del capítulo 12 presentaremos una vista preliminar de la conexión.

Capítulo 12

Regresión lineal múltiple y ciertos modelos de regresión no lineal

12.1 Introducción

En la mayoría de los problemas de investigación en los que se aplica el análisis de regresión se necesita más de una variable independiente para el modelo de regresión. La complejidad de la mayoría de mecanismos científicos es tal que, con el fin de predecir una respuesta importante, se requiere un **modelo de regresión múltiple**. Cuando un modelo es lineal en los coeficientes se denomina **modelo de regresión lineal múltiple**. Para el caso de k variables independientes, el modelo que da x_1, x_2, \dots, x_k , la media de $Y|x_1, x_2, \dots, x_k$ es el modelo de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, \dots, x_k} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k,$$

y la respuesta estimada se obtiene a partir de la ecuación de regresión muestral

$$\hat{y} = b_0 + b_1 x_1 + \dots + b_k x_k,$$

donde cada coeficiente de regresión β_i se estima por medio de b_i , a partir de los datos muestrales, usando el método de los mínimos cuadrados. Como ocurre en el caso de una sola variable independiente, a menudo el modelo de regresión lineal múltiple es una representación adecuada de una estructura más complicada dentro de ciertos rangos de las variables independientes.

También se pueden aplicar técnicas similares de mínimos cuadrados para estimar los coeficientes cuando el modelo lineal incluye, por ejemplo, potencias y productos de las variables independientes. Un ejemplo de esto se presentaría cuando $k = 1$, en cuyo caso el experimentador podría pensar que las medias $\mu_{Y|x}$ no caen sobre una línea recta, sino que se describen de manera más adecuada mediante el **modelo de regresión polinomial**

$$\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_r x^r,$$

y la respuesta estimada se obtiene de la ecuación de regresión polinomial

$$\hat{y} = b_0 + b_1 x + b_2 x^2 + \dots + b_r x^r.$$

En ocasiones se genera confusión al decir que un modelo polinomial es un modelo lineal. Sin embargo, los estadísticos normalmente se refieren a un modelo lineal como aquel en el que los parámetros ocurren en forma lineal, independientemente de cómo las variables independientes entran en el modelo. Un ejemplo de modelo no lineal es la **relación exponencial**

$$\mu_{Y|x} = \alpha\beta^x,$$

que se estima mediante la ecuación de regresión

$$\hat{y} = ab^x.$$

En ciencias e ingeniería hay muchos fenómenos cuya naturaleza no es inherentemente lineal y, cuando se conoce su verdadera estructura, no hay duda de que habría que intentar ajustar el modelo real. Existe mucha literatura acerca de la estimación de modelos no lineales por medio de mínimos cuadrados. Los modelos no lineales que se analizan en este capítulo se relacionan con condiciones no ideales, en las cuales el analista está seguro de que la respuesta y , por lo tanto, el error de respuesta del modelo no se distribuyen normalmente sino que, más bien, tienen una distribución binomial o de Poisson. Estas situaciones ocurren a menudo en la práctica.

El estudiante que busque profundizar en la explicación de la regresión no lineal debe consultar la obra de Myers *Classical and Modern Regression with Applications* (1990; véase la bibliografía).

12.2 Estimación de los coeficientes

En esta sección se calculan los estimadores de mínimos cuadrados de los parámetros $\beta_0, \beta_1, \dots, \beta_k$ mediante el ajuste del modelo de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, \dots, x_k} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

a los puntos de los datos

$$\{(x_{1i}, x_{2i}, \dots, x_{ki}, y_i); \quad i = 1, 2, \dots, n \text{ y } n > k\},$$

donde y_i es la respuesta observada a los valores $x_{1i}, x_{2i}, \dots, x_{ki}$ de las k variables independientes x_1, x_2, \dots, x_k . Se supone que cada observación $(x_{1i}, x_{2i}, \dots, x_{ki}, y_i)$ satisface la siguiente ecuación:

Modelo de
regresión lineal
múltiple

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i$$

o bien,

$$y_i = \hat{y}_i + e_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki} + e_i,$$

donde ϵ_i y e_i son el error aleatorio y el residual, respectivamente, asociados con la respuesta y_i y con el valor ajustado \hat{y}_i .

Como en el caso de la regresión lineal simple, se supone que los ϵ_i son independientes y están distribuidos en forma idéntica con media cero y varianzas común σ^2 .

Si usamos el concepto de mínimos cuadrados para obtener los estimados b_0, b_1, \dots, b_k , minimizamos la expresión

$$SCE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{1i} - b_2 x_{2i} - \dots - b_k x_{ki})^2.$$

Si, a su vez, diferenciamos la SCE respecto a b_0, b_1, \dots, b_k e igualamos el resultado a cero, generamos el conjunto de $k + 1$ ecuaciones normales para la regresión lineal múltiple

Ecuaciones normales de estimación para la regresión lineal múltiple

$$\begin{array}{r}
 nb_0 + b_1 \sum_{i=1}^n x_{1i} + b_2 \sum_{i=1}^n x_{2i} + \cdots + b_k \sum_{i=1}^n x_{ki} = \sum_{i=1}^n y_i \\
 b_0 \sum_{i=1}^n x_{1i} + b_1 \sum_{i=1}^n x_{1i}^2 + b_2 \sum_{i=1}^n x_{1i}x_{2i} + \cdots + b_k \sum_{i=1}^n x_{1i}x_{ki} = \sum_{i=1}^n x_{1i}y_i \\
 \vdots \\
 b_0 \sum_{i=1}^n x_{ki} + b_1 \sum_{i=1}^n x_{ki}x_{1i} + b_2 \sum_{i=1}^n x_{ki}x_{2i} + \cdots + b_k \sum_{i=1}^n x_{ki}^2 = \sum_{i=1}^n x_{ki}y_i
 \end{array}$$

Estas ecuaciones se pueden resolver para $b_0, b_1, b_2, \dots, b_k$ utilizando cualquier método apropiado que permita resolver sistemas de ecuaciones lineales. Casi todos los programas estadísticos de cómputo se pueden utilizar para obtener soluciones numéricas de las ecuaciones anteriores.

Ejemplo 12.1: Se sometió a prueba un grupo de camiones ligeros con motores que utilizan diesel como combustible para saber si la humedad, la temperatura del aire y la presión barométrica influyen en la cantidad de óxido nitroso que emiten (en ppm). Las emisiones se midieron en distintos momentos y en diversas condiciones experimentales. Los datos se presentan en la tabla 12.1. El modelo es

$$\mu_{Y|x_1, x_2, x_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3,$$

o, en forma equivalente,

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i, \quad i = 1, 2, \dots, 20.$$

Ajuste este modelo de regresión lineal múltiple a los datos con los que cuenta y luego estime la cantidad de óxido nitroso que emiten los camiones en las siguientes condiciones: 50% de humedad, temperatura de 76°F y una presión barométrica de 29.30.

Tabla 12.1: Datos para el ejemplo 12.1

Óxido nitroso, y	Humedad, x_1	Temp., x_2	Presión, x_3	Óxido nitroso, y	Humedad, x_1	Temp., x_2	Presión, x_3
0.90	72.4	76.3	29.18	1.07	23.2	76.8	29.38
0.91	41.6	70.3	29.35	0.94	47.4	86.6	29.35
0.96	34.3	77.1	29.24	1.10	31.5	76.9	29.63
0.89	35.1	68.0	29.27	1.10	10.6	86.3	29.56
1.00	10.7	79.0	29.78	1.10	11.2	86.0	29.48
1.10	12.9	67.4	29.39	0.91	73.3	76.3	29.40
1.15	8.3	66.8	29.69	0.87	75.4	77.9	29.28
1.03	20.1	76.9	29.48	0.78	96.6	78.7	29.29
0.77	72.2	77.7	29.09	0.82	107.4	86.8	29.03
1.07	24.0	67.7	29.60	0.95	54.9	70.9	29.37

Fuente: Charles T. Hare, "Light-Duty Diesel Emission Correction Factors for Ambient Conditions", EPA-600/2-77-116. U. S. Environmental Protection Agency.

Solución: La solución del conjunto de las ecuaciones de estimación produce los estimadores únicos

$$b_0 = -3.507778, \quad b_1 = -0.002625, \quad b_2 = 0.000799, \quad b_3 = 0.154155.$$

Por lo tanto, la ecuación de regresión es

$$\hat{y} = -3.507778 - 0.002625x_1 + 0.000799x_2 + 0.154155x_3.$$

Para 50% de humedad, una temperatura de 76°F y una presión barométrica de 29.30, la cantidad estimada de óxido nitroso emitido es

$$\begin{aligned}\hat{y} &= -3.507778 - 0.002625(50.0) + 0.000799(76.0) + 0.1541553(29.30) \\ &= 0.9384 \text{ ppm.}\end{aligned}$$

Regresión polinomial

Ahora suponga que se desea ajustar la ecuación polinomial

$$\mu_{Y|X} = \beta_0 + \beta_1x + \beta_2x^2 + \cdots + \beta_r x^r$$

a los n pares de observaciones $\{(x_i, y_i); i = 1, 2, \dots, n\}$. Cada observación, y_i , satisface la ecuación

$$y_i = \beta_0 + \beta_1x_i + \beta_2x_i^2 + \cdots + \beta_r x_i^r + \epsilon_i$$

o bien,

$$y_i = \bar{y}_i + e_i = b_0 + b_1x_i + b_2x_i^2 + \cdots + b_r x_i^r + e_i,$$

donde r es el grado del polinomio y ϵ_i y e_i son, de nuevo, el error aleatorio y el residual asociados con la respuesta y_i y con el valor ajustado \bar{y}_i , respectivamente. Aquí el número de pares, n , debe ser al menos $r + 1$, que es el número de parámetros por estimar.

Observe que el modelo polinomial se puede considerar un caso especial del modelo de regresión lineal múltiple más general, donde establezcamos $x_1 = x$, $x_2 = x^2, \dots, x_r = x^r$. Las ecuaciones normales adoptan la misma forma que las que aparecen en la página 445. Luego se resuelven para $b_0, b_1, b_2, \dots, b_r$.

Ejemplo 12.2: Datos los datos

x	0	1	2	3	4	5	6	7	8	9
y	9.1	7.3	3.2	4.6	4.8	2.9	5.7	7.1	8.8	10.2

ajuste una curva de regresión de la forma $\mu_{Y|X} = \beta_0 + \beta_1x + \beta_2x^2$, luego, estime $\mu_{Y|X}$.

Solución: A partir de los datos se encuentra que

$$\begin{aligned}10b_0 + 45b_1 + 285b_2 &= 63.7, \\ 45b_0 + 285b_1 + 2025b_2 &= 307.3, \\ 285b_0 + 2025b_1 + 15,333b_2 &= 2153.3.\end{aligned}$$

Al resolver las ecuaciones normales se obtiene

$$\hat{b}_0 = 8.698, \quad \hat{b}_1 = -2.341, \quad \hat{b}_2 = 0.288.$$

Por lo tanto,

$$\hat{y} = 8.698 - 2.341x + 0.288x^2.$$

Cuando $x = 2$ el estimado de $\mu_{y|2}$ es

$$\hat{y} = 8.698 - (2.341)(2) + (0.288)(2^2) = 5.168. \quad \blacksquare$$

Ejemplo 12.3: Los datos de la tabla 12.2 representan el porcentaje de impurezas que resultaron de diversas temperaturas y del tiempo de esterilización durante una reacción asociada con la fabricación de cierta bebida. Estime los coeficientes de regresión en el modelo polinomial

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + \epsilon_i,$$

para $i = 1, 2, \dots, 18$.

Tabla 12.2: Datos para el ejemplo 12.3

Tiempo de esterilización, x_2 (min)	Temperatura, x_1 ($^{\circ}\text{C}$)		
	75	100	125
15	14.05	10.55	7.55
	14.93	9.48	6.59
20	16.56	13.63	9.23
	15.85	11.75	8.78
25	22.41	18.55	15.93
	21.66	17.98	16.44

Solución: Si usamos las ecuaciones normales, obtenemos

$$\begin{aligned} b_0 &= 56.4411, & b_1 &= -0.36190, & b_2 &= -2.75299, \\ b_{11} &= 0.00081, & b_{22} &= 0.08173, & b_{12} &= 0.00314, \end{aligned}$$

y nuestra ecuación de regresión estimada es

$$\hat{y} = 56.4411 - 0.36190x_1 - 2.75299x_2 + 0.00081x_1^2 + 0.08173x_2^2 + 0.00314x_1x_2. \quad \blacksquare$$

Muchos de los principios y procedimientos asociados con la estimación de funciones de regresión polinomiales caen en la categoría de **metodología de respuesta superficial**, que es un conjunto de técnicas que los científicos e ingenieros de muchos campos han utilizado con bastante éxito. Las x_i^2 se denominan **términos cuadráticos puros** y las $x_i x_j$ ($i \neq j$) se conocen como **términos de interacción**. Dichas técnicas a menudo se aplican a problemas tales como seleccionar un diseño experimental adecuado, en particular en casos en los que un número muy grande de variables entra en el modelo; y elegir condiciones óptimas de operación para x_1, x_2, \dots, x_k . Para profundizar en este tema se recomienda al lector consultar la obra de Myers, Montgomery y Anderson-Cook, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (2009; véase la bibliografía).

12.3 Modelo de regresión lineal en el que se utilizan matrices

Al ajustar un modelo de regresión lineal múltiple, en particular cuando contiene más de dos variables, tener conocimientos sobre la teoría de matrices facilita considerablemente el manejo de las matemáticas. Suponga que el experimentador tiene k variables

independientes x_1, x_2, \dots, x_k y n observaciones y_1, y_2, \dots, y_n , cada una de las cuales se puede expresar con la ecuación

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i.$$

Este modelo representa en esencia a n ecuaciones que describen cómo se generan los valores de la respuesta durante el proceso científico. Si usamos la notación de matrices, podemos escribir la ecuación siguiente

Modelo lineal

$$y = X\beta + \epsilon,$$

general donde

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}.$$

Después, el método de mínimos cuadrados para la estimación de β , que se estudió en la sección 12.2, implica calcular \mathbf{b} , para lo cual

$$SCE = (y - X\mathbf{b})'(y - X\mathbf{b})$$

se minimiza. Este proceso de minimización implica resolver para \mathbf{b} en la ecuación

$$\frac{\partial}{\partial \mathbf{b}}(SCE) = \mathbf{0}.$$

Aquí no presentaremos los detalles respecto a cómo se resuelven las ecuaciones anteriores. El resultado se reduce a la solución de \mathbf{b} en

$$(X'X)\mathbf{b} = X'y.$$

Observe la naturaleza de la matriz X . Además del elemento inicial, el i -ésimo renglón representa los valores de x que dan lugar a la respuesta y_i . Si escribimos

$$A = X'X = \begin{bmatrix} n & \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{2i} & \cdots & \sum_{i=1}^n x_{ki} \\ \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i}x_{2i} & \cdots & \sum_{i=1}^n x_{1i}x_{ki} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{i=1}^n x_{ki} & \sum_{i=1}^n x_{ki}x_{1i} & \sum_{i=1}^n x_{ki}x_{2i} & \cdots & \sum_{i=1}^n x_{ki}^2 \end{bmatrix}$$

y

$$g = X'y = \begin{bmatrix} g_0 = \sum_{i=1}^n y_i \\ g_1 = \sum_{i=1}^n x_{1i}y_i \\ \vdots \\ g_k = \sum_{i=1}^n x_{ki}y_i \end{bmatrix}$$

nos permite escribir las ecuaciones normales en la forma de matriz

$$A\mathbf{b} = \mathbf{g}.$$

Si la matriz A es no singular, la solución para los coeficientes de regresión se escribe como

$$\mathbf{b} = \mathbf{A}^{-1} \mathbf{g} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}.$$

De esta manera, obtenemos la ecuación de predicción o regresión resolviendo un conjunto de $k + 1$ ecuaciones con un número igual de incógnitas. Esto implica el invertir la matriz $\mathbf{X}'\mathbf{X}$ de orden $k + 1$ por $k + 1$. En la mayoría de libros que tratan sobre determinantes y matrices elementales se explican las técnicas para invertir matrices. Por supuesto, existen muchos paquetes de cómputo veloces para resolver problemas de regresión múltiple, los cuales no sólo proporcionan estimados de los coeficientes de regresión, sino que también ofrecen otra clase de información relevante para hacer inferencias acerca de la ecuación de regresión.

Ejemplo 12.4: Se midió el porcentaje de supervivencia de los espermatozoides de cierto tipo de semen animal, después de almacenarlo con distintas combinaciones de concentraciones de tres materiales que se emplean para incrementar la supervivencia. En la tabla 12.3 se presentan los datos. Obtenga el modelo de regresión lineal múltiple para los datos.

Tabla 12.3: Datos para el ejemplo 12.4

y (% de supervivencia)	x_1 (peso %)	x_2 (peso %)	x_3 (peso %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Solución: Las ecuaciones de estimación por mínimos cuadrados, $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y}$, son

$$\begin{bmatrix} 13.0 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.0780 \\ 81.82 & 360.6621 & 576.7264 & 728.3100 \\ 115.40 & 522.0780 & 728.3100 & 1035.9600 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 377.5 \\ 1877.567 \\ 2246.661 \\ 3337.780 \end{bmatrix}.$$

A partir de una salida de computadora se obtienen los elementos de la matriz inversa

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 8.0648 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0886 \end{bmatrix}.$$

y, luego, utilizando la relación $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, se obtienen los siguientes coeficientes de regresión estimados

$$b_0 = 39.1574, b_1 = 1.0161, b_2 = -1.8616, b_3 = -0.3433.$$

Entonces, la ecuación de regresión estimada es

$$\hat{y} = 39.1574 + 1.0161x_1 - 1.8616x_2 - 0.3433x_3.$$

Ejercicios

12.1 Se llevó a cabo un conjunto de ensayos experimentales con un horno para determinar una forma de predecir el tiempo de cocción, y , a diferentes niveles de ancho del horno, x_1 , y a diferentes temperaturas, x_2 . Se registraron los siguientes datos:

y	x_1	x_2
6.40	1.32	1.15
15.05	2.69	3.40
18.75	3.56	4.10
30.25	4.41	8.75
44.85	5.35	14.82
48.94	6.20	15.15
51.55	7.12	15.32
61.50	8.87	18.18
100.44	9.80	35.19
111.42	10.65	40.40

Estime la ecuación de regresión lineal múltiple

$$\mu_{Y|X_1, X_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$

12.2 En *Applied Spectroscopy* se estudiaron las propiedades de reflectancia infrarroja de un líquido viscoso que se utiliza como lubricante en la industria electrónica. El experimento que se diseñó consistió en medir el efecto de frecuencia de banda, x_1 , y el espesor de película, x_2 , sobre la densidad óptica, y , usando un espectrómetro infrarrojo Perkin-Elmer Modelo 621. (Fuente: Pacansky, J., England, C. D. y Wattman, R., 1986).

y	x_1	x_2
0.231	740	1.10
0.107	740	0.62
0.053	740	0.31
0.129	805	1.10
0.069	805	0.62
0.030	805	0.31
1.005	980	1.10
0.559	980	0.62
0.321	980	0.31
2.948	1235	1.10
1.633	1235	0.62
0.934	1235	0.31

Estime la ecuación de regresión lineal múltiple

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2.$$

12.3 En el ejercicio de repaso 11.53 de la página 437 suponga que también se proporciona el número de periodos de clase perdidos por los 12 estudiantes que toman el curso de química. A continuación se presentan los datos completos.

Estudiante	Calificación en química, y	Calificación en el examen, x_1	Clases perdidas, x_2
1	85	65	1
2	74	50	7
3	76	55	5
4	90	65	2
5	85	55	6
6	87	70	3
7	94	65	2
8	98	70	5
9	81	55	4
10	91	70	3
11	76	50	1
12	74	55	4

- Ajuste una ecuación de regresión lineal múltiple de la forma $\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i}$.
- Estime la calificación de química para un estudiante que en la prueba de inteligencia obtuvo 60 de calificación y perdió 4 clases.

12.4 Se realizó un experimento para determinar si era posible predecir el peso de un animal después de un periodo determinado con base en su peso inicial y la cantidad de alimento que consumía. Se registraron los siguientes datos, en kilogramos:

Peso final, y	Peso inicial, x_1	Peso del alimento, x_2
95	42	272
77	33	226
80	33	259
100	45	292
97	39	311
70	36	183
50	32	173
80	41	236
92	40	230
84	38	235

- Ajuste una ecuación de regresión múltiple de la forma $\mu_{Y|X_1, X_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.
- Prediga cuánto pesará un animal que comienza pesando 35 kilogramos después de consumir 250 kilogramos de alimento.

12.5 Se cree que la energía eléctrica que una planta química consume cada mes se relaciona con la temperatura ambiental promedio, x_1 , el número de días del mes, x_2 , la pureza promedio del producto, x_3 y las toneladas fabricadas del producto, x_4 . Se dispone de datos históricos del año anterior, los cuales se presentan en la siguiente tabla.

y	x ₁	x ₂	x ₃	x ₄
240	25	24	91	100
236	31	21	90	95
290	45	24	88	110
274	60	25	87	88
301	65	25	91	94
316	72	26	94	99
300	80	25	87	97
296	84	25	86	96
267	75	24	88	110
276	60	25	91	105
288	50	25	90	100
261	38	23	89	98

- a) Ajuste un modelo de regresión lineal múltiple usando el conjunto de datos anterior.
 b) Prediga el consumo de energía para un mes en que $x_1 = 75^\circ\text{F}$, $x_2 = 24$ días, $x_3 = 90\%$ y $x_4 = 98$ toneladas.

12.6 Se realizó un experimento sobre un modelo nuevo de una marca de automóvil específica para determinar la distancia de frenado a distintas velocidades. Se registraron los siguientes datos.

Velocidad, v (km/h)	35	50	65	80	95	110
Distancia de frenado, d (m)	16	26	41	62	88	119

- a) Ajuste una curva de regresión múltiple de la forma $\mu_{D|v} = \beta_0 + \beta_1 v + \beta_2 v^2$.
 b) Estime la distancia de frenado cuando el automóvil viaja a 70 kilómetros por hora.

12.7 Se realizó un experimento con el fin de determinar si el flujo sanguíneo cerebral de los seres humanos se podía predecir a partir de la tensión arterial del oxígeno (milímetros de mercurio). En el estudio participaron 15 pacientes y se reunieron los siguientes datos:

Flujo sanguíneo, y	Tensión arterial del oxígeno, x
84.33	603.40
87.80	582.50
82.20	556.20
78.21	594.60
78.44	558.90
80.01	575.20
83.53	580.10
79.46	451.20
75.22	404.00
76.58	484.00
77.90	452.40
78.80	448.40
80.67	334.80
86.60	320.30
78.20	350.30

Estime la ecuación de regresión cuadrática

$$\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2.$$

12.8 El siguiente es un conjunto de datos experimentales codificados acerca de la resistencia a la compresión de una aleación específica para distintos valores de la concentración de cierto aditivo:

Concentración, x	Resistencia a la compresión, y		
10.0	25.2	27.3	28.7
15.0	29.8	31.1	27.8
20.0	31.2	32.6	29.7
25.0	31.7	30.1	32.3
30.0	29.4	30.8	32.8

- a) Estime la ecuación de regresión cuadrática $\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2$.
 b) Pruebe la falta de ajuste del modelo.

12.9 a) Ajuste una ecuación de regresión múltiple de la forma $\mu_{Y|x} = \beta_0 + \beta_1 x_1 + \beta_2 x_2^2$ para los datos del ejemplo 11.8 de la página 420.

- b) Estime el producto de la reacción química para una temperatura de 225°C .

12.10 Para los datos siguientes

x	0	1	2	3	4	5	6
y	1	4	5	3	2	3	4

- a) Ajuste el modelo cúbico $\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$.
 b) Prediga el valor de Y cuando $x = 2$.

12.11 Se realizó un experimento para estudiar el tamaño de los calamares consumidos por tiburones y atunes. Las variables regresoras son características de la boca del calamar. Los datos del estudio son los siguientes:

x_1	x_2	x_3	x_4	x_5	y
1.31	1.07	0.44	0.75	0.35	1.95
1.55	1.49	0.53	0.90	0.47	2.90
0.99	0.84	0.34	0.57	0.32	0.72
0.99	0.83	0.34	0.54	0.27	0.81
1.01	0.90	0.36	0.64	0.30	1.09
1.09	0.93	0.42	0.61	0.31	1.22
1.08	0.90	0.40	0.51	0.31	1.02
1.27	1.08	0.44	0.77	0.34	1.93
0.99	0.85	0.36	0.56	0.29	0.64
1.34	1.13	0.45	0.77	0.37	2.08
1.30	1.10	0.45	0.76	0.38	1.98
1.33	1.10	0.48	0.77	0.38	1.90
1.86	1.47	0.60	1.01	0.65	8.56
1.58	1.34	0.52	0.95	0.50	4.49
1.97	1.59	0.67	1.20	0.59	8.49
1.80	1.56	0.66	1.02	0.59	6.17
1.75	1.58	0.63	1.09	0.59	7.54
1.72	1.43	0.64	1.02	0.63	6.36
1.68	1.57	0.72	0.96	0.68	7.63
1.75	1.59	0.68	1.08	0.62	7.78
2.19	1.86	0.75	1.24	0.72	10.15
1.73	1.67	0.64	1.14	0.55	6.88

En el estudio las variables regresoras y la respuesta considerada son

- x_1 = longitud del rostral, en pulgadas,
- x_2 = longitud de la aleta, en pulgadas,
- x_3 = longitud del rostral a la cola, en pulgadas,
- x_4 = longitud de la cola a la aleta, en pulgadas,
- x_5 = ancho, en pulgadas,
- y = peso, en libras.

Estime la ecuación de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, x_3, x_4, x_5} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5.$$

12.12 Los siguientes datos reflejan información obtenida en 17 hospitales de la marina estadounidense ubicados en diversos sitios del mundo. Los regresores son variables de la carga de trabajo, es decir, conceptos que dan como resultado la necesidad de personal en un hospital. A continuación se presenta una descripción breve de las variables:

- y = horas de trabajo mensuales,
- x_1 = carga diaria promedio de pacientes,
- x_2 = exposiciones de rayos X mensuales,
- x_3 = días-cama ocupados por mes,
- x_4 = población elegible en el área/1000,
- x_5 = duración promedio de la estancia de un paciente, en días.

Sitio	x_1	x_2	x_3	x_4	x_5	y
1	15.57	2463	472.92	18.0	4.45	566.52
2	44.02	2048	1339.75	9.5	6.92	696.82
3	20.42	3940	620.25	12.8	4.28	1033.15
4	18.74	6505	568.33	36.7	3.90	1003.62
5	49.20	5723	1497.60	35.7	5.50	1611.37
6	44.92	11,520	1365.83	24.0	4.60	1613.27
7	55.48	5779	1687.00	43.3	5.62	1854.17
8	59.28	5969	1639.92	46.7	5.15	2160.55
9	94.39	8461	2872.33	78.7	6.18	2305.58
10	128.02	20,106	3655.08	180.5	6.15	3503.93
11	96.00	13,313	2912.00	60.9	5.88	3571.59
12	131.42	10,771	3921.00	103.7	4.88	3741.40
13	127.21	15,543	3865.67	126.8	5.50	4026.52
14	252.90	36,194	7684.10	157.7	7.00	10,343.81
15	409.20	34,703	12,446.33	169.4	10.75	11,732.17
16	463.70	39,204	14,098.40	331.4	7.05	15,414.94
17	510.22	86,533	15,524.00	371.6	6.35	18,854.45

El objetivo es generar una ecuación empírica para estimar (o predecir) las necesidades de personal en los hospitales de la marina. Calcule la ecuación de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, x_3, x_4, x_5} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5.$$

12.13 Se llevó a cabo un estudio sobre un tipo de conexión para conocer la relación entre la cantidad de desgaste, y , para x_1 = viscosidad del aceite, y x_2 =

carga. Se obtuvieron los datos siguientes. (Tomado de *Response Surface Methodology*, Myers, Montgomery y Anderson-Cook, 2009).

y	x_1	x_2	y	x_1	x_2
193	1.6	851	230	15.5	816
172	22.0	1058	91	43.0	1201
113	33.0	1357	125	40.0	1115

a) Estime los parámetros desconocidos de la ecuación de regresión lineal múltiple

$$\mu_{Y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$

b) Prediga el desgaste cuando la viscosidad del aceite sea de 20 y la carga sea de 1200.

12.14 Once estudiantes normalistas participaron en un programa de evaluación diseñado para medir la eficacia de los maestros y determinar cuáles factores son importantes. La medición de la respuesta consistió en una evaluación cuantitativa del maestro. Las variables regresoras fueron las calificaciones de cuatro pruebas estandarizadas aplicadas a cada maestro. Los datos son los siguientes:

y	x_1	x_2	x_3	x_4
410	69	125	59.00	55.66
569	57	131	31.75	63.97
425	77	141	80.50	45.32
344	81	122	75.00	46.67
324	0	141	49.00	41.21
505	53	152	49.35	43.83
235	77	141	60.75	41.61
501	76	132	41.25	64.57
400	65	157	50.75	42.41
584	97	166	32.25	57.95
434	76	141	54.50	57.90

Estime la ecuación de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, x_3, x_4} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4.$$

12.15 Con el fin de determinar la relación entre la calificación de su desempeño laboral (y) y las calificaciones en cuatro exámenes, el departamento de personal de cierta empresa industrial realizó un estudio en el que participaron 12 sujetos. Los datos son los siguientes:

y	x_1	x_2	x_3	x_4
11.2	56.5	71.0	38.5	43.0
14.5	59.5	72.5	38.2	44.8
17.2	69.2	76.0	42.5	49.0
17.8	74.5	79.5	43.4	56.3
19.3	81.2	84.0	47.5	60.2
24.5	88.0	86.2	47.4	62.0
21.2	78.2	80.5	44.5	58.1
16.9	69.0	72.0	41.8	48.1
14.8	58.1	68.0	42.1	46.0
20.0	80.5	85.0	48.1	60.3
13.2	58.3	71.0	37.5	47.1
22.5	84.0	87.2	51.0	65.2

Estime los coeficientes de regresión del modelo

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4.$$

12.16 Un ingeniero de una empresa de semiconductores desea modelar la relación entre la ganancia o hFE de un dispositivo (y) y tres parámetros: RS del emisor (x_1), RS de la base (x_2) y RS del emisor a la base (x_3). A continuación se muestran los datos:

x_1 , RS del emisor	x_2 , RS de la base	x_3 , E-B-RS	y , hFE
14.62	226.0	7.000	128.40
15.63	220.0	3.375	52.62
14.62	217.4	6.375	113.90
15.00	220.0	6.000	98.01
14.50	226.5	7.625	139.90
15.25	224.1	6.000	102.60

(cont.)

x_1 , RS del emisor	x_2 , RS de la base	x_3 , E-B-RS	y , hFE
16.12	220.5	3.375	48.14
15.13	223.5	6.125	109.60
15.50	217.6	5.000	82.68
15.13	228.5	6.625	112.60
15.50	230.2	5.750	97.52
16.12	226.5	3.750	59.06
15.13	226.6	6.125	111.80
15.63	225.6	5.375	89.09
15.38	234.0	8.875	171.90
15.50	230.0	4.000	66.80
14.25	224.3	8.000	157.10
14.50	240.5	10.870	208.40
14.62	223.7	7.375	133.40

(Datos de Myers, Montgomery y Anderson-Cook, 2009).

- a) Ajuste una regresión lineal múltiple para los datos.
b) Prediga hFE cuando $x_1 = 14$, $x_2 = 220$ y $x_3 = 5$.

12.4 Propiedades de los estimadores de mínimos cuadrados

Las medias y varianzas de los estimadores b_0, b_1, \dots, b_k se obtienen con facilidad si se hacen ciertas suposiciones sobre los errores aleatorios $\epsilon_1, \epsilon_2, \dots, \epsilon_i$, que son idénticas a las que se hacen en el caso de la regresión lineal simple. Si suponemos que dichos errores son independientes, cada uno con media igual a cero y varianza σ^2 , entonces podemos demostrar que b_0, b_1, \dots, b_k son, respectivamente, estimadores no sesgados de los coeficientes de regresión $\beta_0, \beta_1, \dots, \beta_k$. Además, las varianzas de las b se obtienen por medio de los elementos del inverso de la matriz \mathbf{A} . Observe que los elementos fuera de la diagonal de $\mathbf{A} = \mathbf{X}'\mathbf{X}$ representan sumas de productos de los elementos en las columnas de \mathbf{X} ; mientras que los elementos en la diagonal de \mathbf{A} son las sumas de los cuadrados de los elementos en las columnas de \mathbf{X} . La matriz inversa, \mathbf{A}^{-1} , aparte del multiplicador σ^2 , representa la **matriz de varianza-covarianza** de los coeficientes de regresión estimados. Es decir, los elementos de la matriz $\mathbf{A}^{-1}\sigma^2$ muestran las varianzas de b_0, b_1, \dots, b_k en la diagonal principal y las covarianzas fuera de la diagonal. Por ejemplo, en un problema de regresión lineal múltiple con $k = 2$ se podría escribir

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} c_{00} & c_{01} & c_{02} \\ c_{10} & c_{11} & c_{12} \\ c_{20} & c_{21} & c_{22} \end{bmatrix}$$

con los elementos debajo de la diagonal principal determinados por la simetría de la matriz. Entonces, se escribe

$$\begin{aligned} \sigma_{b_i}^2 &= c_{ii} \sigma^2, & i &= 0, 1, 2, \\ \sigma_{b_i, b_j} &= \text{Cov}(b_i, b_j) = c_{ij} \sigma^2, & i &\neq j. \end{aligned}$$

Desde luego, los estimados de las varianzas y, por lo tanto, sus errores estándar, se obtienen reemplazando σ^2 con el estimado apropiado, el cual se obtuvo a partir de los datos experimentales. Un estimado no sesgado de σ^2 de nuevo se define en términos de

la suma de cuadrados del error, que se calcula utilizando la fórmula establecida en el teorema 12.1. En el teorema las suposiciones se basan en los ϵ_i descritos con anterioridad.

Teorema 12.1: Para la ecuación de regresión lineal

$$y = X\beta + \epsilon,$$

un estimador insesgado de σ^2 es dado por el error o media cuadrática residual

$$s^2 = \frac{SCE}{n - k - 1}, \quad \text{donde} \quad SCE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Podemos ver que, para el caso de la regresión lineal simple, el teorema 12.1 representa una generalización del teorema 11.1. La prueba se deja como ejercicio para el lector. Al igual que en el caso de la regresión lineal más simple, el estimado de s^2 es una medida de la variación de los errores de la predicción o residuales. En las secciones 12.10 y 12.11 se presentan otras inferencias importantes relacionadas con la ecuación ajustada de regresión, con base en los valores de los residuales individuales $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$.

La suma de cuadrados del error y de la regresión adoptan la misma forma y desempeñan el mismo papel que en el caso de la regresión lineal simple. De hecho, la identidad de la suma de cuadrados

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

se sigue cumpliendo, y se conserva la notación anterior, que es,

$$STCC = SCR + SCE,$$

con

$$STCC = \sum_{i=1}^n (y_i - \bar{y})^2 = \text{suma de cuadrados total}$$

y

$$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \text{suma de cuadrados de regresión}$$

Hay k grados de libertad asociados con la SCR , y, como siempre, la $STCC$ tiene $n - 1$ grados de libertad. Por lo tanto, después de restar, la SCE tiene $n - k - 1$ grados de libertad. Así, nuestro estimado de σ^2 de nuevo es dado por la suma de cuadrados del error dividida entre sus grados de libertad. Las tres sumas de cuadrados aparecen en la salida de resultados de la mayoría de los programas de cómputo de regresión múltiple. Observe que la condición $n > k$ en la sección 12.2 garantiza que los grados de libertad de la SCE no sean negativos.

Análisis de varianza en la regresión múltiple

La partición de la suma total de cuadrados en sus componentes, la suma de cuadrados de regresión y del error desempeña un papel importante. Puede efectuarse un **análisis de varianza** que arroje luz sobre la calidad de la ecuación de regresión. Una hipótesis que sirve para determinar si el modelo explica una cantidad significativa de variación, es la siguiente:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0.$$

El análisis de varianza implica una prueba F , mediante una tabla, como la siguiente:

Fuente	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Regresión	SCR	k	$CMR = \frac{SCR}{k}$	$f = \frac{CMR}{CME}$
Error	SCE	$n - (k + 1)$	$CME = \frac{SCE}{n - (k + 1)}$	
Total	$STCC$	$n - 1$		

Se trata de una **prueba de cola superior**. El rechazo de H_0 significa que la **ecuación de regresión difiere de una constante**. Es decir, al menos una variable regresora es importante. En las secciones que siguen se estudia más el uso del análisis de varianza.

Otra utilidad del cuadrado medio del error (o cuadrado medio residual) estriba en su uso para la prueba de hipótesis y la estimación de intervalos de confianza que se estudian en la sección 12.5. Además, el cuadrado medio del error desempeña un papel importante en las situaciones en las que el científico busca el mejor modelo entre un conjunto de ellos que están en competencia. Muchos criterios de construcción de modelos incluyen el estadístico s^2 . En la sección 12.11 se presentan criterios para comparar modelos en competencia.

12.5 Inferencias en la regresión lineal múltiple

El conocimiento de la distribución de los estimadores del coeficiente individual facilita al experimentador construir intervalos de confianza para los coeficientes y hacer pruebas de hipótesis acerca de ellos. Recuerde que en la sección 12.4 estudiamos que b_j ($j = 0, 1, 2, \dots, k$) se distribuyen de forma normal con media β_j y varianza $c_{jj}\sigma^2$. De esta manera, se puede utilizar el estadístico

$$t = \frac{b_j - \beta_{j0}}{s\sqrt{c_{jj}}}$$

con $n - k - 1$ grados de libertad para probar hipótesis y construir intervalos de confianza sobre β_j . Por ejemplo, si queremos probar

$$H_0: \beta_j = \beta_{j0},$$

$$H_1: \beta_j \neq \beta_{j0},$$

se calcula el estadístico t anterior y no se rechaza H_0 si $-t_{\alpha/2} < t < t_{\alpha/2}$, donde $t_{\alpha/2}$ tiene $n - k - 1$ grados de libertad.

Ejemplo 12.5: Para el modelo del ejemplo 12.4 pruebe la hipótesis de que $\beta_2 = -2.5$ en comparación con la alternativa de que $\beta_2 > -2.5$ a un nivel de significancia de 0.05.

Solución:

$$H_0: \beta_2 = -2.5,$$

$$H_1: \beta_2 > -2.5.$$

Cálculos:

$$t = \frac{b_2 - \beta_{20}}{s \sqrt{c_{22}}} = \frac{-1.8616 + 2.5}{2.073 \sqrt{0.0166}} = 2.390,$$

$$P = P(T > 2.390) = 0.04.$$

Decisión: Rechazar H_0 y concluir que $\beta_2 > -2.5$. ▮

Pruebas t individuales para la selección de variables

La prueba t que se utiliza con más frecuencia en la regresión múltiple es aquella que prueba la importancia de los coeficientes individuales, es decir, $H_0: \beta_j = 0$ en comparación con la hipótesis alternativa $H_1: \beta_j \neq 0$. Con frecuencia estas pruebas contribuyen a lo que se denomina **selección de variables**, con la cual el analista intenta llegar al modelo más útil, es decir, a la elección de cuál regresor utilizar. Aquí debemos destacar que, si se encuentra que un coeficiente es insignificante, es decir, si **no se rechaza** la hipótesis $H_0: \beta_j = 0$, la conclusión que se obtiene es que la **variable** es insignificante (explica una cantidad insignificante de la variación de y) **en la presencia de los demás regresores del modelo**. Más adelante se profundizará en este punto.

Inferencias sobre la respuesta media y la predicción

Una de las inferencias más útiles que se pueden hacer con respecto a la calidad de la respuesta predicha y_0 , correspondiente a los valores $x_{10}, x_{20}, \dots, x_{k0}$, es el intervalo de confianza sobre la respuesta media $\mu_1 | x_{10}, x_{20}, \dots, x_{k0}$. Estamos interesados en construir un intervalo de confianza sobre la respuesta media para el conjunto de condiciones determinadas por

$$\mathbf{x}'_0 = [1, x_{10}, x_{20}, \dots, x_{k0}].$$

Se aumentan en 1 las condiciones sobre las x para facilitar la notación de matrices. La normalidad en los ϵ_i producen normalidad en los b_j , y la media y la varianza siguen siendo las mismas, como se indica en la sección 12.4. Así es la covarianza entre b_i y b_j para $i \neq j$. De esta manera,

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_{j0}$$

también se distribuye normalmente y es, de hecho, un estimador no sesgado para la **respuesta media** sobre la que se intenta ligar un intervalo de confianza. La varianza de \hat{y}_0 , escrita con notación de matriz simplemente como función de σ^2 , $(\mathbf{X}'\mathbf{X})^{-1}$, y el vector de condiciones, \mathbf{x}'_0 es

$$\sigma_{\hat{y}_0}^2 = \sigma^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0.$$

Si esta expresión se extendiera para un caso dado, por ejemplo $k = 2$, ya vimos que explica de manera apropiada la varianza de b_j y la covarianza de b_i y b_j , para $i \neq j$. Después de sustituir σ^2 con s^2 , según se plantea en el teorema 12.1, el intervalo de confianza del $100(1 - \alpha)\%$ se puede construir sobre $\mu_{y|x} = x_{10}, x_{20}, \dots, x_{k0}$ a partir del estadístico

$$T = \frac{\hat{y}_0 - \mu_{Y|x_{10}, x_{20}, \dots, x_{k0}}}{s \sqrt{\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}},$$

que tiene una distribución t con $n - k - 1$ grados de libertad.

Intervalo de confianza para $\mu_{Y|x_{10}, x_{20}, \dots, x_{k0}}$ Un intervalo de confianza de $100(1 - \alpha)\%$ para la **respuesta media** $\mu_{Y|x_{10}, x_{20}, \dots, x_{k0}}$ es

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} < \mu_{Y|x_{10}, x_{20}, \dots, x_{k0}} < \hat{y}_0 + t_{\alpha/2} s \sqrt{\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0},$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - k - 1$ grados de libertad.

Es frecuente que a la cantidad $s \sqrt{\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}$ se le denomine **error estándar de la predicción** y aparece en la salida de resultados de muchos paquetes de cómputo para regresión.

Ejemplo 12.6: Con los datos del ejemplo 12.4 construya un intervalo de confianza de 95% para la respuesta media, cuando $x_1 = 3\%$, $x_2 = 8\%$ y $x_3 = 9\%$.

Solución: De la ecuación de regresión del ejemplo 12.4, el porcentaje estimado de supervivencia cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$, es:

$$\hat{y} = 39.1574 + (1.0161)(3) - (1.8616)(8) - (0.3433)(9) = 24.2232.$$

Y luego se determina que

$$\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0 = [1, 3, 8, 9] \begin{bmatrix} 8.0648 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0886 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 8 \\ 9 \end{bmatrix}$$

$$= 0.1267.$$

Si utilizamos el cuadrado medio del error, $s^2 = 4.298$ o $s = 2.073$, y la tabla A.4, observamos que $t_{0.025} = 2.262$ para 9 grados de libertad. Por lo tanto, un intervalo de confianza de 95% para el porcentaje medio de supervivencia para $x_1 = 3\%$, $x_2 = 8\%$ y $x_3 = 9\%$ es dado por

$$24.2232 - (2.262)(2.073)\sqrt{0.1267} < \mu_{Y|3,8,9} < 24.2232 + (2.262)(2.073)\sqrt{0.1267},$$

o simplemente $22.5541 < \mu_{Y|3,8,9} < 25.8923$. ▮

Como ocurre en el caso de la regresión lineal simple, necesitamos distinguir con claridad entre el intervalo de confianza sobre la respuesta media y el intervalo de predicción sobre una *respuesta observada*. Esta última proporciona un límite dentro del cual podemos decir que, con un grado preseleccionado de certidumbre, caerá una respuesta nueva observada.

Nuevamente se establece un intervalo de predicción para una sola respuesta predicha y_0 al considerar la diferencia $\hat{y}_0 - y_0$. Se puede demostrar que la distribución del muestreo es normal con media

$$\mu_{\hat{y}_0 - y_0} = 0$$

y varianza

$$\sigma_{\hat{y}_0 - y_0}^2 = \sigma^2 [1 + \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0].$$

Por consiguiente, se puede construir un intervalo de predicción del $100(1 - \alpha)\%$ para un solo valor de predicción y_0 a partir del estadístico

$$T = \frac{\hat{y}_0 - y_0}{s \sqrt{1 + \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}},$$

el cual tiene una distribución t con $n - k - 1$ grados de libertad.

Intervalo de predicción para y_0 Un intervalo de predicción del $100(1 - \alpha)\%$ para una sola respuesta y_0 es dado por

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{1 + \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0} < y_0 < \hat{y}_0 + t_{\alpha/2} s \sqrt{1 + \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0},$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - k - 1$ grados de libertad.

Ejemplo 12.7: Con los datos del ejemplo 12.4 construya un intervalo de predicción de 95% para una respuesta individual del porcentaje de supervivencia, cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$.

Solución: Si nos remitimos a los resultados del ejemplo 12.6, encontramos que el intervalo de predicción de 95% para la respuesta y_0 , cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$, es

$$24.2232 - (2.262)(2.073) \sqrt{1.1267} < y_0 < 24.2232 + (2.262)(2.073) \sqrt{1.1267},$$

que se reduce a $19.2459 < y_0 < 29.2005$. Observe que, como se esperaba, el intervalo de predicción es considerablemente más ancho que el intervalo de confianza para el porcentaje medio de supervivencia del ejemplo 12.6. \blacksquare

Salida de resultados comentado para los datos del ejemplo 12.4

La figura 12.1 muestra una salida de resultados por computadora con comentarios para el ajuste de regresión lineal múltiple de los datos del ejemplo 12.4. Se empleó el paquete SAS.

Observe los estimados de los parámetros del modelo, los errores estándar y los estadísticos t que aparecen en el listado. Los errores estándar se calcularon a partir de las raíces cuadradas de los elementos de la diagonal $(\mathbf{X}'\mathbf{X})^{-1} s^2$. En dicha ilustración la variable x_3 es insignificante en presencia de x_1 y x_2 con base en la prueba t y el valor P correspondiente de 0.5916. Los términos CLM y CLI son intervalos de confianza sobre la respuesta media y los límites de predicción sobre una observación individual, respectivamente. La prueba f en el análisis de varianza indica que se explica una cantidad significativa de variabilidad. Como ejemplo de las interpretaciones de CLM y CLI, considere la observación 10. Con una observación de 25.2000 y un valor predicho de 26.0676 tenemos 95% de confianza en que la respuesta media está entre 24.5024 y 27.6329, y en que una observación nueva caerá entre 21.1238 y 31.0114 con una probabilidad de 0.95. El valor R^2 de 0.9117 implica que el modelo explica el 91.17% de la variabilidad de la respuesta. En la sección 12.6 se analiza más a fondo R^2 .

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	399.45437	133.15146	30.98	<.0001
Error	9	38.67640	4.29738		
Corrected Total	12	438.13077			

Root MSE	2.07301	R-Square	0.9117
Dependent Mean	29.03846	Adj R-Sq	0.8823
Coeff Var	7.13885		

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	39.15735	5.88706	6.65	<.0001
x1	1	1.01610	0.19090	5.32	0.0005
x2	1	-1.86165	0.26733	-6.96	<.0001
x3	1	-0.34326	0.61705	-0.56	0.5916

Obs	Variable	Dependent Value	Predicted Mean	Std Error Predict	95% CL Mean	95% CL Predict	Residual	
1	25.5000	27.3514	1.4152	24.1500	30.5528	21.6734	33.0294	-1.8514
2	31.2000	32.2623	0.7846	30.4875	34.0371	27.2482	37.2764	-1.0623
3	25.9000	27.3495	1.3588	24.2757	30.4234	21.7425	32.9566	-1.4495
4	38.4000	38.3096	1.2818	35.4099	41.2093	32.7960	43.8232	0.0904
5	18.4000	15.5447	1.5789	11.9730	19.1165	9.6499	21.4395	2.8553
6	26.7000	26.1081	1.0358	23.7649	28.4512	20.8658	31.3503	0.5919
7	26.4000	28.2532	0.8094	26.4222	30.0841	23.2189	33.2874	-1.8532
8	25.9000	26.2219	0.9732	24.0204	28.4233	21.0414	31.4023	-0.3219
9	32.0000	32.0882	0.7828	30.3175	33.8589	27.0755	37.1008	-0.0882
10	25.2000	26.0676	0.6919	24.5024	27.6329	21.1238	31.0114	-0.8676
11	39.7000	37.2524	1.3070	34.2957	40.2090	31.7086	42.7961	2.4476
12	35.7000	32.4879	1.4648	29.1743	35.8015	26.7459	38.2300	3.2121
13	26.5000	28.2032	0.9841	25.9771	30.4294	23.0122	33.3943	-1.7032

Figura 12.1: Salida de resultados del SAS para los datos del ejemplo 12.4.

Más sobre el análisis de varianza en la regresión múltiple (opcional)

En la sección 12.4 se estudió brevemente la partición de la suma total de cuadrados $\sum_{i=1}^n (y_i - \bar{y})^2$ en sus dos componentes, el modelo de regresión y la suma de cuadrados del error (que se ilustran en la figura 12.1). El análisis de varianza conduce a la prueba de

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0.$$

El rechazo de la hipótesis nula implica una interpretación importante para el científico o el ingeniero. (A quienes les interese profundizar en el tema del uso de matrices les será útil estudiar el desarrollo de estas sumas de cuadrados que se usan en el ANOVA).

En primer lugar, de la sección 12.3 recuerde que \mathbf{b} , el vector de los estimadores de mínimos cuadrados, es dado por

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

Una partición de la **suma de cuadrados no corregida**,

$$y'y = \sum_{i=1}^n y_i^2$$

en dos componentes es dada por

$$\begin{aligned} y'y &= \mathbf{b}'\mathbf{X}'y + (y'y - \mathbf{b}'\mathbf{X}'y) \\ &= y'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y + [y'y - y'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y]. \end{aligned}$$

El segundo término (entre corchetes) en el lado derecho es tan sólo la suma de cuadrados del error $\sum_{i=1}^n (y_i - \hat{y}_i)^2$. El lector debería observar que una expresión alternativa para la suma de cuadrados del error es

$$SCE = y'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']y.$$

El término $y'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$ se denomina **suma de cuadrados de la regresión**. Sin embargo, no se trata de la expresión $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ que se usó para probar la "importancia" de los términos b_1, b_2, \dots, b_k , sino más bien de

$$y'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y = \sum_{i=1}^n \hat{y}_i^2,$$

que es la suma de cuadrados de la regresión no corregida para la media. Como tal, sólo se podría usar para probar si la ecuación de regresión difiere significativamente de cero, es decir,

$$H_0: \beta_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0.$$

En general, esto no es tan importante como probar

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0,$$

dado que esto plantea que la respuesta media es una constante, no necesariamente cero.

Grados de libertad

Así, la partición de las sumas de cuadrados y los grados de libertad se reduce a

Fuente	Suma de cuadrados	gl
Regresión	$\sum_{i=1}^n \hat{y}_i^2 = y'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$	$k + 1$
Error	$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = y'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}]y$	$n - (k + 1)$
Total	$\sum_{i=1}^n y_i^2 = y'y$	n

Hipótesis de interés

Desde luego, la hipótesis de interés para un ANOVA debe eliminar el papel de la intersección según se describió anteriormente. En términos estrictos, si $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$, entonces la recta de regresión estimada es simplemente $\hat{y}_i = \bar{y}$. Como resultado, en realidad se busca evidencia de que la ecuación de regresión “varíe a partir de una constante”. Así, la suma de cuadrados total y la suma de regresión deben corregirse para la media. Como resultado, tenemos

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

En notación de matrices esto es simplemente

$$\mathbf{y}'[\mathbf{I}_n - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y} = \mathbf{y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y} + \mathbf{y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}.$$

En esta expresión el $\mathbf{1}$ sólo es un vector de n unos. Como resultado, simplemente restamos

$$\mathbf{y}'\mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{y} = \frac{1}{n} \left(\sum_{i=1}^n y_i \right)^2$$

de $\mathbf{y}'\mathbf{y}$ y de $\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, es decir, corrigiendo la suma de cuadrados total y la de regresión para la media.

Por último, la partición apropiada de las sumas de cuadrados con grados de libertad es como sigue:

Fuente	Suma de cuadrados	gl
Regresión	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \mathbf{y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y}$	k
Error	$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \mathbf{y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}$	$n - (k + 1)$
Total	$\sum_{i=1}^n (y_i - \bar{y})^2 = \mathbf{y}'[\mathbf{I}_n - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y}$	$n - 1$

Ésta es la tabla ANOVA que aparece en la salida de resultados por computadora de la figura 12.1. Es frecuente denominar a la expresión $\mathbf{y}'[\mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y}$ como la **suma de cuadrados de la regresión asociada con la media**, y se le asigna 1 grado de libertad.

Ejercicios

12.17 Para los datos del ejercicio 12.2 de la página 450, estime σ^2 .

12.18 Para los datos del ejercicio 12.1 de la página 450, estime σ^2 .

12.19 Para los datos del ejercicio 12.5 de la página 450, estime σ^2 .

12.20 Obtenga estimados de las varianzas y la covarianza de los estimadores b_1 y b_2 , del ejercicio 12.2 de la página 450.

12.21 Remítase al ejercicio 12.5 de la página 450 y obtenga estimados de

a) $\sigma_{b_2}^2$;

b) $\text{Cov}(b_1, b_2)$.

12.22 Para el modelo del ejercicio 12.7 de la página 451, a un nivel de significancia de 0.05 pruebe la hipótesis de que $\beta_2 = 0$, en comparación con la hipótesis alternativa de que $\beta_2 \neq 0$.

12.23 Para el modelo del ejercicio 12.2 de la página 450 a un nivel de significancia de 0.05, pruebe la hipótesis de que $\beta_1 = 0$, en comparación con la hipótesis alternativa de que $\beta_1 \neq 0$.

12.24 Para el modelo del ejercicio 12.1 de la página 450 pruebe la hipótesis de que $\beta_1 = 2$, en comparación con la hipótesis alternativa de que $\beta_1 \neq 2$. Utilice un valor P en sus conclusiones.

12.25 Utilice los datos del ejercicio 12.2 de la página 450 y el estimado de σ^2 del ejercicio 12.17 para calcular intervalos de confianza de 95% para la respuesta predicha y la respuesta media cuando $x_1 = 900$ y $x_2 = 1.00$.

12.26 Para el ejercicio 12.8 de la página 451 construya un intervalo de confianza de 90% para la resistencia media a la compresión cuando la concentración es $x = 19.5$ y se utiliza un modelo cuadrático.

12.27 Utilice los datos del ejercicio 12.5 de la página 450 y el estimado de σ^2 del ejercicio 12.19 para calcular intervalos de confianza de 95% para la respuesta predicha y la respuesta media cuando $x_1 = 75$, $x_2 = 24$, $x_3 = 90$ y $x_4 = 98$.

12.28 Considere los siguientes datos del ejercicio 12.13 de la página 452.

y (desgaste)	x_1 (viscosidad del aceite)	x_2 (carga)
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

- a) Estime σ^2 usando regresión múltiple de y sobre x_1 y x_2 .
- b) Calcule valores predichos, un intervalo de confianza de 95% para el desgaste promedio y un intervalo de predicción de 95% para el desgaste observado si $x_1 = 20$ y $x_2 = 1000$.

12.29 Con los datos del ejercicio 12.28, y a un nivel de 0.05, pruebe:

- a) $H_0: \beta_1 = 0$ en comparación con $H_1: \beta_1 > 0$;
- b) $H_0: \beta_2 = 0$ en comparación con $H_1: \beta_2 > 0$;
- c) ¿Existe alguna razón para creer que habría que cambiar el modelo del ejercicio 12.28? Explique su respuesta.

12.30 Utilice los datos del ejercicio 12.16 de la página 453.

- a) Estime σ^2 usando la regresión múltiple de y sobre x_1 , x_2 y x_3 ;
- b) Calcule un intervalo de predicción de 95% para la ganancia observada con los tres regresores en $x_1 = 15.0$, $x_2 = 220.0$ y $x_3 = 6.0$.

12.6 Selección de un modelo ajustado mediante la prueba de hipótesis

En muchas situaciones de regresión los coeficientes individuales revisten importancia para el experimentador. Por ejemplo, en una aplicación de economía, β_1, β_2, \dots podrían tener un significado en particular, por lo que el economista tendría un interés especial en los intervalos de confianza y en las pruebas de hipótesis sobre dichos parámetros. Sin embargo, considere una situación de química industrial en la que el modelo propuesto supone que el producto de la reacción depende linealmente de la temperatura y concentración de la reacción de cierto catalizador. Es probable que se sepa que éste no es el verdadero modelo, sino una aproximación adecuada; de manera que el interés no estribaría en los parámetros individuales, sino en la capacidad de la función en su conjunto para predecir la respuesta verdadera en el rango de las variables consideradas. Por lo tanto, en esta situación, se pondría más énfasis en σ^2 , los intervalos de confianza de la respuesta media, y así sucesivamente, y disminuiría el interés en las inferencias sobre los parámetros individuales.

El experimentador que utiliza análisis de regresión también está interesado en eliminar variables cuando la situación impone que, además de llegar a una ecuación de pronóstico funcional, debe encontrar la "mejor regresión" que implique sólo variables que sean predictores útiles. Se dispone de varios programas de cómputo que llegan en secuencia a la denominada mejor ecuación de regresión, dependiendo de ciertos criterios. En la sección 12.9 profundizaremos en el estudio de esto.

Un criterio que suele utilizarse para ilustrar lo adecuado de un modelo ajustado de regresión es el coeficiente de determinación múltiple o R^2 .

Coefficiente de
determinación
múltiple o R^2

$$R^2 = \frac{SCR}{STCC} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SCE}{STCC}.$$

Advierta que esta descripción se parece a la que se hizo de R^2 en el capítulo 11. En este punto la explicación podría ser más clara, toda vez que ahora nos centramos en SCR como la **variabilidad explicada**. La cantidad R^2 tan sólo indica qué proporción de la variación total de la respuesta Y es explicada por el modelo ajustado. Con frecuencia los experimentadores reportan $R^2 \times 100\%$ e interpretan el resultado como el porcentaje de variación con respecto al modelo propuesto. La raíz cuadrada de R^2 se denomina **coeficiente de correlación múltiple** entre Y y el conjunto x_1, x_2, \dots, x_k . En el ejemplo 12.4 el valor de R^2 que indica la proporción de variación explicada por las tres variables independientes x_1, x_2 y x_3 es

$$R^2 = \frac{SCR}{STCC} = \frac{399.45}{438.13} = 0.9117,$$

lo cual significa que 91.17% de la variación del porcentaje de supervivencia queda explicada por el modelo de regresión lineal.

La suma de cuadrados de regresión se puede emplear para obtener algún indicio acerca de si el modelo representa o no una explicación adecuada de la verdadera situación. Podemos probar la hipótesis H_0 de que la **regresión no es significativa** con sólo plantear la razón

$$f = \frac{SCR/k}{SCE/(n-k-1)} = \frac{SCR/k}{s^2}$$

y rechazar H_0 al nivel de significancia α cuando $f > f_\alpha(k, n-k-1)$. Para los datos del ejemplo 12.4 se obtiene

$$f = \frac{399.45/3}{4.298} = 30.98.$$

De la salida de resultados por computadora que aparece en la figura 12.1, el valor P es menor que 0.0001. Esto no debe malinterpretarse. Aunque indica que la regresión explicada por el modelo es significativa, no descarta la posibilidad de que

1. El modelo de regresión lineal en este conjunto de x no sea el único que se puede usar para explicar los datos; de hecho, quizás haya otros modelos con transformaciones sobre las x que generen un valor mayor para el estadístico F .
2. El modelo podría ser más eficaz si se incluyeran otras variables, además de x_1, x_2 y x_3 , o quizá si se eliminaran una o más de las variables del modelo, por ejemplo x_3 , que tiene un valor $P = 0.5916$.

El lector debería recordar el análisis de la sección 11.5 sobre las desventajas de utilizar R^2 como criterio para comparar modelos en competencia. Es evidente que dichas desventajas son relevantes en la regresión lineal múltiple. De hecho, los riesgos de su empleo en la regresión múltiple son aún mayores debido a que es muy grande la tentación de hacer un sobreajuste. Hay que tener siempre presente que $R^2 \approx 1.0$ siempre puede

obtenerse a expensas de los grados de libertad del error cuando se emplea un exceso de términos en el modelo. Sin embargo, $R^2 = 1$, que describe un modelo con ajuste casi perfecto, no siempre genera un modelo que hace buenas predicciones.

El coeficiente de determinación ajustado (R^2_{ajus})

En el capítulo 11 se presentan varias figuras que muestran listados de resultados por computadora, tanto del SAS como de MINITAB, en las que aparece un estadístico llamado R^2 ajustado, o un coeficiente de determinación ajustado. R^2 ajustado es una variación de R^2 que proporciona un ajuste para los grados de libertad. El coeficiente de determinación, según se definió en la página 407, no puede disminuir a medida que se agregan términos al modelo. En otras palabras, R^2 no disminuye a medida que se reducen los grados de libertad del error $n - k - 1$, ya que este último resultado se produce por un incremento de k , el número de términos en el modelo. R^2 ajustado se calcula dividiendo la SCE y la STCC entre sus grados de libertad respectivos de la siguiente manera.

R^2 ajustado

$$R^2_{\text{ajus}} = 1 - \frac{SCE / (n - k - 1)}{STCC / (n - 1)}$$

Para ilustrar el uso de R^2_{ajus} se revisará el ejemplo 12.4.

¿Cómo la eliminación de x_3 afecta a R^2 y R^2_{ajus} ?

La prueba t (o la prueba F correspondiente) para x_3 sugiere que un modelo más sencillo que sólo implique x_1 y x_2 bien podría ser una mejoría. En otras palabras, el modelo completo con todos los regresores podría estar sobreajustado. Por supuesto que es interesante investigar R^2 y R^2_{ajus} tanto para el modelo completo (x_1 , x_2 y x_3) como para el modelo reducido (x_1 , x_2). A partir de la figura 12.1 ya sabemos que $R^2_{\text{compl}} = 0.9117$. La SCE para el modelo reducido es 40.01, por lo que $R^2_{\text{reduc}} = 1 - \frac{40.01}{438.13} = 0.9087$. De esta forma, con x_3 dentro del modelo se explica más variabilidad. No obstante, como ya se dijo, esto ocurriría aun si el modelo estuviera sobreajustado. Desde luego que R^2_{ajus} está diseñada para proporcionar un estadístico que castigue un modelo sobreajustado, de manera que podríamos esperar que se favorezca al modelo restringido. Entonces, para el modelo completo

$$R^2_{\text{ajus}} = 1 - \frac{38.6764/9}{438.1308/12} = 1 - \frac{4.2974}{36.5109} = 0.8823,$$

mientras que para el modelo reducido (eliminación de x_3)

$$R^2_{\text{ajus}} = 1 - \frac{40.01/10}{438.1308/12} = 1 - \frac{4.001}{36.5109} = 0.8904.$$

Así, R^2_{ajus} realmente favorece el modelo reducido y confirma la evidencia proporcionada por las pruebas t y F , sugiriendo que el modelo reducido es preferible sobre el que contiene los tres regresores. El lector quizás espere que otros estadísticos sugieran el rechazo del modelo sobreajustado. Véase el ejercicio 12.40 de la página 471.

Prueba sobre un coeficiente individual

Agregar cualquier variable sencilla a un sistema de regresión *incrementará la suma de cuadrados de regresión* y con ello *se reducirá la suma de cuadrados del error*. En consecuencia, se debe decidir si el incremento en la regresión es suficiente para garantizar el uso de la variable en el modelo. Como es de esperarse, el empleo de variables sin importancia reduciría la eficacia de la ecuación de predicción incrementando la varianza de la respuesta estimada. Profundizaremos más en este punto al considerar la importancia de x_3 en el ejemplo 12.4. Inicialmente podemos probar

$$H_0: \beta_3 = 0,$$

$$H_1: \beta_3 \neq 0$$

usando la distribución t con 9 grados de libertad. Se tiene

$$t = \frac{b_3 - 0}{s\sqrt{c_{33}}} = \frac{-0.3433}{2.073\sqrt{0.0886}} = -0.556,$$

que indica que β_3 no difiere en forma significativa de cero y, por lo tanto, bien podríamos sentir que se justifica eliminar x_3 del modelo. Suponga que se considera la regresión de Y sobre el conjunto (x_1, x_2) , las ecuaciones normales de mínimos cuadrados ahora se reducen a

$$\begin{bmatrix} 13.0 & 59.43 & 81.82 \\ 59.43 & 394.7255 & 360.6621 \\ 81.82 & 360.6621 & 576.7264 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 377.50 \\ 1877.5670 \\ 2246.6610 \end{bmatrix}.$$

Los coeficientes de regresión estimados para este modelo reducido son

$$b_0 = 36.094, \quad b_1 = 1.031, \quad b_2 = -1.870,$$

y la suma de cuadrados de regresión resultante, con 2 grados de libertad, es

$$R(\beta_1, \beta_2) = 398.12.$$

Aquí se utiliza la notación $R(\beta_1, \beta_2)$ para indicar la suma de cuadrados de regresión del modelo restringido, y no debe confundirse con la SCR , es decir, la suma de cuadrados de regresión del modelo original con 3 grados de libertad. Entonces, la nueva suma de cuadrados del error es

$$STCC - R(\beta_1, \beta_2) = 438.13 - 398.12 = 40.01,$$

y el cuadrado medio del error resultante, con 10 grados de libertad, es

$$s^2 = \frac{40.01}{10} = 4.001.$$

¿Una prueba t de una variable tiene una prueba equivalente F ?

En el ejemplo 12.4 la cantidad de variación en el porcentaje de supervivencia que se atribuye a x_3 , en presencia de las variables x_1 y x_2 , es

$$R(\beta_3 | \beta_1, \beta_2) = SCR - R(\beta_1, \beta_2) = 399.45 - 398.12 = 1.33,$$

que representa una pequeña proporción de toda la variación de la regresión. Esta cantidad de regresión agregada, como lo indica la prueba previa sobre β_3 , es estadísticamente insignificante. Una prueba equivalente implica la formación de la razón

$$f = \frac{R(\beta_3 | \beta_1, \beta_2)}{s^2} = \frac{1.33}{4.298} = 0.309,$$

que es un valor de la distribución F con 1 y 9 grados de libertad. Recuerde que la relación básica entre la distribución t con v grados de libertad y la distribución F con 1 y v grados de libertad es

$$t^2 = f(1, v),$$

y se observa que el valor f de 0.309 es en realidad el cuadrado del valor t de -0.56 .

Para generalizar los conceptos anteriores podemos evaluar el funcionamiento de una variable independiente x_i en el modelo general de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, \dots, x_k} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

observando la cantidad de regresión atribuida a x_i **sobre y por arriba de la atribuida a las demás variables**, es decir, la regresión sobre x_i *ajustada para las demás variables*. Por ejemplo, se dice que x_i se evalúa calculando

$$R(\beta_1 | \beta_2, \beta_3, \dots, \beta_k) = SCR - R(\beta_2, \beta_3, \dots, \beta_k),$$

donde $R(\beta_2, \beta_3, \dots, \beta_k)$ es la suma de cuadrados de regresión con $\beta_1 x_1$ eliminados del modelo. Para probar la hipótesis

$$H_0: \beta_1 = 0,$$

$$H_1: \beta_1 \neq 0,$$

se calcula

$$f = \frac{R(\beta_1 | \beta_2, \beta_3, \dots, \beta_k)}{s^2},$$

y se compara con $f_{\alpha}(1, n - k - 1)$.

Pruebas F parciales en subconjuntos de coeficientes

De manera similar, se puede hacer una prueba para la significancia de un *conjunto* de las variables. Por ejemplo, para investigar simultáneamente la importancia de incluir x_1 y x_2 en el modelo se prueba la hipótesis

$$H_0: \beta_1 = \beta_2 = 0,$$

$$H_1: \beta_1 \text{ y } \beta_2 \text{ no son ambas cero,}$$

calculando

$$f = \frac{[R(\beta_1, \beta_2 | \beta_3, \beta_4, \dots, \beta_k)]/2}{s^2} = \frac{[SCR - R(\beta_3, \beta_4, \dots, \beta_k)]/2}{s^2}$$

y comparando con $f_{\alpha}(2, n - k - 1)$. El número de grados de libertad asociados con el numerador, en este caso 2, es igual al número de variables en el conjunto que se investiga.

Suponga que se desea probar la hipótesis

$$\begin{aligned} H_0: \beta_2 = \beta_3 = 0, \\ H_1: \beta_2 \text{ y } \beta_3 \text{ no son ambas cero} \end{aligned}$$

para el ejemplo 12.4. Si desarrollamos el modelo de regresión

$$y = \beta_0 + \beta_1 x_1 + \epsilon,$$

podemos obtener $R(\beta_1) = SCR_{\text{reduc}} = 187.31179$. En la figura 12.1, de la página 459, tenemos $s^2 = 4.29738$ para el modelo completo. Por lo tanto, el valor de f para la prueba de hipótesis es

$$\begin{aligned} f &= \frac{R(\beta_2, \beta_3 \mid \beta_1)/2}{s^2} = \frac{[R(\beta_1, \beta_2, \beta_3) - R(\beta_1)]/2}{s^2} = \frac{[SCR_{\text{compl}} - SCR_{\text{reduc}}]/2}{s^2} \\ &= \frac{(399.45437 - 187.31179)/2}{4.29738} = 24.68278. \end{aligned}$$

Esto implica que β_2 y β_3 no son iguales a cero de forma simultánea. Se puede utilizar un programa de estadística como el SAS para obtener el resultado anterior de manera directa, con un valor P de 0.0002. Los lectores deben observar que en los resultados de los programas de estadística para computadora aparecen valores P asociados con cada coeficiente individual del modelo. La hipótesis nula para cada una es que el coeficiente es igual a cero. Sin embargo, debemos señalar que la insignificancia de cualquier coeficiente no implica necesariamente que no deba ser incluido en el modelo final; sólo sugiere que es insignificante ante la presencia de todas las otras variables en el problema. El estudio de caso que se incluye al final del capítulo ilustra más esta cuestión.

12.7 Caso especial de ortogonalidad (opcional)

Antes de nuestro desarrollo original del problema general de regresión lineal se planteó la suposición de que las variables independientes se miden sin error y que con frecuencia están bajo el control del experimentador. A menudo ocurren como resultado de un *experimento diseñado con gran detalle*. De hecho, se puede incrementar la eficacia de la ecuación de predicción resultante utilizando un plan de experimentación adecuado.

Suponga que nuevamente consideramos la matriz \mathbf{X} , tal como se definió en la sección 12.3. Podemos describirla como

$$\mathbf{X} = [\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k],$$

donde $\mathbf{1}$ representa una columna de unos y \mathbf{x}_j es un vector columna que representa los niveles de x_j . Si

$$\mathbf{x}_p' \mathbf{x}_q = 0, \quad \text{para } p \neq q,$$

se dice que las variables x_p y x_q son *ortogonales* entre sí. Hay ciertas ventajas evidentes en tener una situación completamente ortogonal, en la cual $\mathbf{x}_p' \mathbf{x}_q = 0$.

para toda posible p y q , $p \neq q$ y, además,

$$\sum_{i=1}^n x_{ji} = 0, \quad j = 1, 2, \dots, k.$$

La XX resultante es una matriz diagonal, y las ecuaciones normales de la sección 12.3 se reducen a

$$nb_0 = \sum_{i=1}^n y_i, \quad b_1 \sum_{i=1}^n x_{1i}^2 = \sum_{i=1}^n x_{1i} y_i, \dots, \quad b_k \sum_{i=1}^n x_{ki}^2 = \sum_{i=1}^n x_{ki} y_i.$$

Una ventaja importante es que es fácil hacer la partición de la *SCR* en **componentes de un solo grado de libertad**, cada uno de los cuales corresponde a la cantidad de variación de Y explicada por una variable controlada establecida. En la situación ortogonal se escribe

$$\begin{aligned} SCR &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n (b_0 + b_1 x_{1i} + \dots + b_k x_{ki} - b_0)^2 \\ &= b_1^2 \sum_{i=1}^n x_{1i}^2 + b_2^2 \sum_{i=1}^n x_{2i}^2 + \dots + b_k^2 \sum_{i=1}^n x_{ki}^2 \\ &= R(\beta_1) + R(\beta_2) + \dots + R(\beta_k). \end{aligned}$$

La cantidad $R(\beta_i)$ es la cantidad de la suma de cuadrados de regresión asociada con un modelo que implica una sola variable independiente x_i .

Para probar simultáneamente la significancia de un conjunto de m variables en una situación ortogonal, la suma de cuadrados de regresión se convierte en

$$R(\beta_1, \beta_2, \dots, \beta_m \mid \beta_{m+1}, \beta_{m+2}, \dots, \beta_k) = R(\beta_1) + R(\beta_2) + \dots + R(\beta_m),$$

y, por lo tanto,

$$R(\beta_1 \mid \beta_2, \beta_3, \dots, \beta_k) = R(\beta_1)$$

se simplifica cuando se evalúa una sola variable independiente. Por consiguiente, la contribución de una variable determinada o un conjunto de variables se encuentra, en esencia, *ignorando* las demás variables del modelo. Las evaluaciones independientes del beneficio de las variables individuales se llevan a cabo usando las técnicas de análisis de varianza, tal como se presentan en la tabla 12.4. La variación total en la respuesta está dividida en componentes de un solo grado de libertad más el término del error con $n - k - 1$ grados de libertad. Cada valor f calculado se utiliza para probar una de las hipótesis

$$\left. \begin{aligned} H_0: \beta_i &= 0 \\ H_1: \beta_i &\neq 0 \end{aligned} \right\} \quad i = 1, 2, \dots, k.$$

comparándolas con el punto crítico $f_{\alpha}(1, n - k - 1)$ o simplemente interpretando el valor P calculado a partir de la distribución f .

Tabla 12.4: Análisis de varianza para variables ortogonales

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada
β_1	$R(\beta_1) = b_1^2 \sum_{i=1}^n x_{1i}^2$	1	$R(\beta_1)$	$\frac{R(\beta_1)}{s^2}$
β_2	$R(\beta_2) = b_2^2 \sum_{i=1}^n x_{2i}^2$	1	$R(\beta_2)$	$\frac{R(\beta_2)}{s^2}$
\vdots	\vdots	\vdots	\vdots	\vdots
β_k	$R(\beta_k) = b_k^2 \sum_{i=1}^n x_{ki}^2$	1	$R(\beta_k)$	$\frac{R(\beta_k)}{s^2}$
Error	SCE	$n - k - 1$	$s^2 = \frac{SCE}{n - k - 1}$	
Total	$SST = S_{yy}$	$n - 1$		

Ejemplo 12.8: Suponga que un científico recaba datos experimentales sobre el radio de un grano propulsor, Y , en función de la temperatura del polvo, x_1 , la tasa de extrusión, x_2 y la temperatura del molde, x_3 . Ajuste un modelo de regresión lineal para predecir el radio del grano y determine la eficacia de cada variable que interviene en el modelo. Los datos se presentan en la tabla 12.5.

Tabla 12.5: Datos para el ejemplo 12.8

Radio del grano	Temperatura del polvo	Tasa de extrusión	Temperatura del molde
82	150 (-1)	12 (-1)	220 (-1)
93	190 (+1)	12 (-1)	220 (-1)
114	150 (-1)	24 (+1)	220 (-1)
124	150 (-1)	12 (-1)	250 (+1)
111	190 (+1)	24 (+1)	220 (-1)
129	190 (+1)	12 (-1)	250 (+1)
157	150 (-1)	24 (+1)	250 (+1)
164	190 (+1)	24 (+1)	250 (+1)

Solución: Observe que cada variable está controlada en dos niveles, y que el experimento está compuesto por las ocho combinaciones posibles. Por conveniencia, los datos de las variables independientes se codificaron mediante las siguientes fórmulas:

$$x_1 = \frac{\text{temperatura del polvo} - 170}{20},$$

$$x_2 = \frac{\text{tasa de extrusión} - 18}{6},$$

$$x_3 = \frac{\text{temperatura del molde} - 235}{15}.$$

Los niveles resultantes de x_1 , x_2 y x_3 toman los valores -1 y $+1$, tal como se indica en la tabla con los datos. Este diseño experimental en particular permite la ortogonalidad que

queremos ilustrar aquí. (En el capítulo 15 se analiza un tratamiento más completo de este tipo de diseño experimental). La matriz X es

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

y las condiciones de ortogonalidad se verifican con facilidad.

Ahora podemos calcular los coeficientes

$$b_0 = \frac{1}{8} \sum_{i=1}^8 y_i = 121.75, \quad b_1 = \frac{1}{8} \sum_{i=1}^8 x_{1i} y_i = \frac{20}{8} = 2.5,$$

$$b_2 = \frac{\sum_{i=1}^8 x_{2i} y_i}{8} = \frac{118}{8} = 14.75, \quad b_3 = \frac{\sum_{i=1}^8 x_{3i} y_i}{8} = \frac{174}{8} = 21.75,$$

de manera que, en términos de las variables codificadas, la ecuación de predicción es

$$\hat{y} = 121.75 + 2.5x_1 + 14.75x_2 + 21.75x_3.$$

El análisis de varianza de la tabla 12.6 presenta las contribuciones independientes a la SCR de cada variable. Cuando los resultados se comparan con $f_{0.05}(1,4)$, cuyo valor es 7.71, indican que x_1 no contribuye de manera significativa a un nivel de 0.05; mientras que las variables x_2 y x_3 sí son significativas. En este ejemplo el estimado para σ^2 es 23.1250. Igual que en el caso de una sola variable independiente, se debe señalar que este estimado no sólo contiene variación por el error experimental, a menos que el modelo postulado sea correcto. De otra manera, el estimado estará "contaminado" por la falta de ajuste, además del error puro, y la falta de ajuste sólo se puede separar si se obtienen múltiples observaciones experimentales para las distintas combinaciones (x_1, x_2, x_3) .

Tabla 12.6: Análisis de varianza para los datos del radio de los granos

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada	Valor P
β_1	$(2.5)^2(8) = 50.00$	1	50.00	2.16	0.2156
β_2	$(14.75)^2(8) = 1740.50$	1	1740.50	75.26	0.0010
β_3	$(21.75)^2(8) = 3784.50$	1	3784.50	163.65	0.0002
Error	92.50	4	23.13		
Total	5667.50	7			

Como x_1 no es significativa, simplemente se puede eliminar del modelo sin alterar los efectos de las otras variables. Observe que tanto x_2 como x_3 tienen un efecto positivo sobre el radio del grano, pero x_3 es el factor más importante debido a la pequeñez de su valor P .

Ejercicios

12.31 Calcule e interprete el coeficiente de determinación múltiple para las variables del ejercicio 12.1 de la página 450.

12.32 Pruebe si la regresión explicada por el modelo del ejercicio 12.1, que se encuentra en la página 450, es significativa a un nivel de significancia de 0.01.

12.33 Pruebe si la regresión explicada por el modelo del ejercicio 12.5, de la página 450, es significativa a un nivel de significancia de 0.01.

12.34 Para el modelo del ejercicio 12.5 de la página 450 pruebe la hipótesis

$$H_0: \beta_1 = \beta_2 = 0,$$

$$H_1: \beta_1 \text{ y } \beta_2 \text{ no son ambas cero.}$$

12.35 Repita el ejercicio 12.17 de la página 461 usando el estadístico F .

12.36 Se realizó un pequeño experimento para ajustar una ecuación de regresión múltiple que relaciona el producto, y , con la temperatura, x_1 , el tiempo de reacción, x_2 , y la concentración de uno de los reactantes, x_3 . Se eligieron dos niveles de cada variable y se registraron las siguientes mediciones correspondientes a las variables independientes codificadas:

y	x_1	x_2	x_3
7.6	-1	-1	-1
8.4	1	-1	-1
9.2	-1	1	-1
10.3	-1	-1	1
9.8	1	1	-1
11.1	1	-1	1
10.2	-1	1	1
12.6	1	1	1

a) Utilice las variables codificadas para estimar la ecuación de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, x_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

b) Divida la SCR , es decir, la suma de cuadrados de regresión, en tres componentes de un solo grado de libertad atribuibles a x_1 , x_2 y x_3 , respectivamente. Construya una tabla de análisis de varianza que indique las pruebas de significancia para cada variable.

12.37 Considere los datos de energía eléctrica del ejercicio 12.5 de la página 450. Pruebe $H_0: \beta_1 = \beta_2 = 0$ utilizando $R(\beta_1, \beta_2 | \beta_3, \beta_4)$. Proporcione un valor P y saque conclusiones.

12.38 Considere los datos del ejercicio 12.36. Calcule lo siguiente:

$$R(\beta_1 | \beta_0), \quad R(\beta_1 | \beta_0, \beta_2, \beta_3),$$

$$R(\beta_2 | \beta_0, \beta_1), \quad R(\beta_2 | \beta_0, \beta_1, \beta_3),$$

$$R(\beta_3 | \beta_0, \beta_1, \beta_2), \quad R(\beta_1, \beta_2 | \beta_3).$$

Haga comentarios al respecto.

12.39 Considere los datos del ejercicio 11.55 de la página 437. Ajuste un modelo de regresión utilizando el peso y el cociente de manejo como variables explicativas. Compare este modelo con el de la RLS (regresión lineal simple) utilizando sólo el peso. Utilice R^2 , R^2_{ajus} , y cualquier estadístico t (o F) que necesite para comparar la RLS con el modelo de regresión múltiple.

12.40 Considere el ejemplo 12.4. La figura 12.1 de la página 459 presenta una salida de resultados del SAS para un análisis del modelo que contiene las variables x_1 , x_2 y x_3 . Céntrese en el intervalo de confianza de la respuesta media μ_Y en las ubicaciones (x_1, x_2, x_3) que representan los 13 puntos de los datos. Considere el elemento en la salida de resultados indicado con C.V., que representa al coeficiente de variación, el cual se define como

$$C.V. = \frac{s}{\bar{y}} \cdot 100,$$

donde $s = \sqrt{s^2}$ es la raíz del cuadrado medio del error. El coeficiente de variación se utiliza con frecuencia como otro criterio para comparar modelos en competencia. Se trata de una cantidad sin escala que expresa al estimado de σ , es decir, s , como un porcentaje de la respuesta promedio \bar{y} . Al competir por el "mejor" modelo de un grupo de modelos en competencia se busca un modelo con un valor pequeño de C.V. Haga un análisis de regresión del conjunto de datos que se presenta en el ejemplo 12.4, pero elimine x_3 . Compare el modelo completo (x_1, x_2, x_3) con el restringido (x_1, x_2) y céntrate en dos criterios: i) C.V.; ii) la anchura de los intervalos de confianza sobre μ_Y . Para el segundo criterio usted quizá desearía usar la anchura promedio. Haga comentarios al respecto.

12.41 Considere el ejemplo 12.3 de la página 447. Compare los dos modelos en competencia

$$\text{Primer orden: } y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i,$$

$$\text{Segundo orden: } y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}$$

$$+ \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + \epsilon_i.$$

Utilice R^2_{ajus} para realizar la comparación. Pruebe $H_0: \beta_{11} = \beta_{22} = \beta_{12} = 0$. También utilice el C.V. que se mencionó en el ejercicio 12.40.

12.42 En el ejemplo 12.8 se trata el caso de eliminar del modelo x_1 , que representa la temperatura del polvo, ya que el valor P basado en la prueba F es 0.2156, en tanto que los valores P para x_2 y x_3 son casi cero.

- Reduzca el modelo eliminando x_1 , después genere un modelo completo y uno restringido (o reducido), y compárelos basándose en R^2_{ajus} .
- Compare los modelos completo y restringido usando intervalos de predicción de 95% de ancho sobre una nueva observación. El “mejor” de ambos modelos será aquel con intervalos de predicción más “estrechos”. Utilice el promedio del ancho de los intervalos de predicción.

12.43 Considere los datos del ejercicio 12.13 de la página 452. ¿La respuesta, o sea el uso, se puede explicar en forma adecuada mediante una sola variable (ya sea la viscosidad o la carga) con una RLS en vez de con la regresión completa con dos variables? Justifique su respuesta con pruebas de hipótesis, así como con la comparación de los tres modelos en competencia.

12.44 Para el conjunto de datos que se da en el ejercicio 12.16 de la página 453, ¿es posible explicar la respuesta en forma adecuada usando dos variables regresoras cualesquiera? Analice el problema.

12.8 Variables categóricas o indicadoras

Un caso especial de aplicación muy importante de la regresión lineal múltiple ocurre cuando una o más de las variables regresoras son **variables categóricas, indicadoras o ficticias**. Es probable que en un proceso químico el ingeniero desee modelar el producto del proceso en comparación con regresores tales como la temperatura del proceso y el tiempo de reacción. Sin embargo, hay interés por el uso de dos catalizadores diferentes y por incluir de algún modo el “catalizador” en el modelo. El efecto del catalizador no se puede medir sobre un continuo, de manera que es una variable categórica. Un analista podría desear modelar el precio de casas en comparación con regresores que incluyan los pies cuadrados de superficie habitable, x_1 , la superficie del terreno, x_2 , y la antigüedad de la vivienda, x_3 . Estos regresores son de naturaleza claramente continua. Sin embargo, es evidente que el costo de las casas podría variar en forma sustancial de una zona del país a otra. Si reuniéramos datos acerca de casas en el este, el medio oeste, en el sur y en el oeste, tendríamos una variable indicadora con **cuatro categorías**. En el ejemplo del proceso químico, si utilizáramos dos catalizadores tendríamos una variable indicadora con dos categorías. En un ejemplo biomédico, donde se compara un medicamento con un placebo, a todos los sujetos se les evalúa con varias mediciones continuas, como su edad, presión sanguínea, etcétera, al igual que el género, que por supuesto es una variable categórica con dos categorías. De esta manera, además de las variables continuas existen dos variables indicadoras, el tratamiento con dos categorías (medicamento activo y placebo) y el género con dos categorías (hombre y mujer).

Modelo con variables categóricas

Para ilustrar la forma en que las variables indicadoras participan en el modelo utilizaremos el ejemplo del proceso químico. Suponga que y = producto, x_1 = temperatura y x_2 = tiempo de reacción. Ahora denotaremos con z la variable indicadora. Sea $z = 0$ para el catalizador 1 y $z = 1$ para el catalizador 2. La asignación del indicador (0, 1) al catalizador es arbitraria. Como resultado, el modelo se convierte en

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 z_i + \epsilon_i, \quad i = 1, 2, \dots, n.$$

Tres categorías

Continuamos aplicando la estimación de los coeficientes con el método de los mínimos cuadrados. En el caso de tener tres niveles o categorías de una sola variable indicadora.

el modelo incluirá dos regresores, digamos z_1 y z_2 , donde la asignación (0, 1) es como sigue:

$$\begin{bmatrix} z_1 & z_2 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

donde $\mathbf{0}$ y $\mathbf{1}$ son vectores de ceros y unos, respectivamente. En otras palabras, si hay ℓ categorías, el modelo incluye $\ell - 1$ términos reales.

Puede ser aleccionador observar la representación gráfica del modelo con 3 categorías. En aras de la simplicidad, se considerará una sola variable continua x . Como resultado, el modelo quedará representado como

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_{1i} + \beta_3 z_{2i} + \epsilon_i.$$

Así, la figura 12.2 refleja la naturaleza del modelo. Las siguientes son expresiones del modelo para las tres categorías.

$$\begin{aligned} E(Y) &= (\beta_0 + \beta_2) + \beta_1 x, & \text{categoría 1,} \\ E(Y) &= (\beta_0 + \beta_3) + \beta_1 x, & \text{categoría 2,} \\ E(Y) &= \beta_0 + \beta_1 x, & \text{categoría 3.} \end{aligned}$$

Como resultado, el modelo que incluye variables categóricas en esencia implica un **cambio en la intersección** a medida que se pasa de una categoría a otra. Desde luego, aquí se asume que los **coeficientes de las variables continuas son los mismos entre las categorías**.

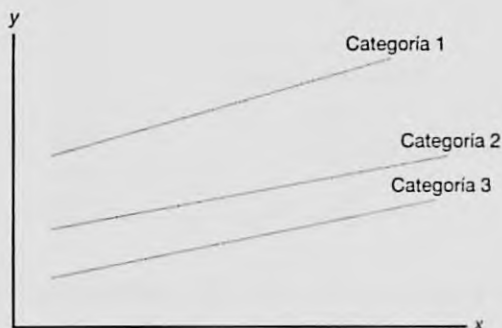


Figura 12.2: Caso de tres categorías.

Ejemplo 12.9: Considere los datos de la tabla 12.7. La respuesta y es la cantidad de sólidos en suspensión en un sistema de limpieza de carbón. La variable x es el pH del sistema y se utilizan tres polímeros diferentes. Así, “polímero” es categórico con tres categorías, de manera que produce dos términos en el modelo, el cual queda como

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_{1i} + \beta_3 z_{2i} + \epsilon_i, \quad i = 1, 2, \dots, 18.$$

Luego, tenemos

$$z_1 = \begin{cases} 1, & \text{para el polímero 1,} \\ 0, & \text{en cualquier caso,} \end{cases} \quad \text{y} \quad z_2 = \begin{cases} 1, & \text{para el polímero 2,} \\ 0, & \text{en cualquier caso.} \end{cases}$$

A partir del análisis de la figura 12.3 se sacan las siguientes conclusiones. El coeficiente b_1 para el pH es el estimado de la **pendiente común** que se asume en el análisis de regresión. Todos los términos del modelo son estadísticamente significativos. Así, el pH y la naturaleza del polímero tienen un efecto sobre la cantidad de limpieza. Los signos y las magnitudes de los coeficientes de z_1 y z_2 indican que el polímero más eficaz para la limpieza es el polímero 1 (produce más sólidos en suspensión), seguido por el polímero 2, y que el menos eficaz es el polímero 3. J

Tabla 12.7: Datos para el ejemplo 12.9

x (pH)	y (cantidad de sólidos en suspensión)	Polímero
6.5	292	1
6.9	329	1
7.8	352	1
8.4	378	1
8.8	392	1
9.2	410	1
6.7	198	2
6.9	227	2
7.5	277	2
7.9	297	2
8.7	364	2
9.2	375	2
6.5	167	3
7.0	225	3
7.2	247	3
7.6	268	3
8.7	288	3
9.2	342	3

La pendiente puede variar con las categorías indicadoras

En el análisis efectuado hasta el momento se ha supuesto que los términos de las variables indicadoras entran al modelo en forma aditiva, lo cual sugiere que las pendientes, como las que se aprecian en la figura 12.2, son constantes en todas las categorías. Es evidente que éste no siempre será el caso. Existe la posibilidad de que las pendientes varíen y realmente se ponga a prueba esta condición de **paralelismo** al incluir términos de producto o **interacción** entre los términos indicadores y las variables continuas. Por ejemplo, suponga que se eligen un modelo con un regresor continuo y una variable indicadora con dos niveles. El modelo entonces quedaría como sigue

$$y = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 xz + \epsilon.$$

		Sum of			
Source	DF	Squares	Mean Square	F Value	Pr > F
Mode 1	3	80181.73127	26727.24376	73.68	<.0001
Error	14	5078.71318	362.76523		
Corrected Total	17	85260.44444			

R-Square	Coeff Var	Root MSE	y Mean	
0.940433	6.316049	19.04640	301.5556	

Parameter	Estimate	Error	t Value	Pr > t
Intercept	-161.8973333	37.43315576	-4.32	0.0007
x	54.2940260	4.75541126	11.42	<.0001
z1	89.9980606	11.05228237	8.14	<.0001
z2	27.1656970	11.01042883	2.47	0.0271

Figura 12.3: Salida de resultados del SAS para el ejemplo 12.9.

Este modelo sugiere que para la categoría 1 ($z = 1$),

$$E(y) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)x,$$

mientras que para la categoría 2 ($z = 0$),

$$E(y) = \beta_0 + \beta_1 x.$$

Por consiguiente, se permite que varíen la intersección y las pendientes para las dos categorías. En la figura 12.4 se presentan las rectas de regresión con pendientes variables para las dos categorías.

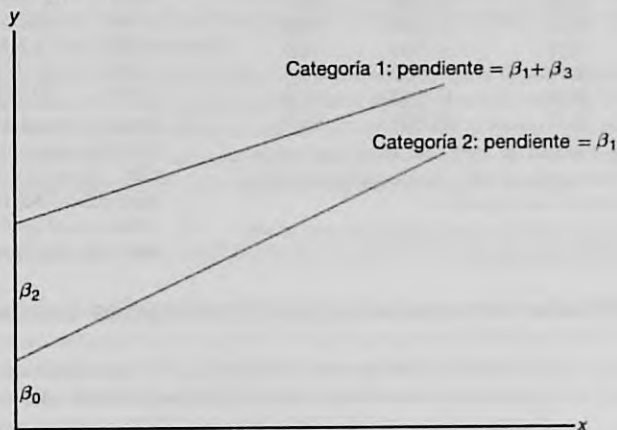


Figura 12.4: Falta de paralelismo en las variables categóricas.

En este caso β_0 , β_1 y β_2 son positivas, mientras que β_3 es negativa con $|\beta_3| < \beta_1$. Por supuesto, si el coeficiente de interacción β_3 es insignificante, regresamos al modelo común de la pendiente.

Ejercicios

12.45 Se realizó un estudio para evaluar el combustible que se ahorra al conducir un automóvil sedán de cuatro puertas en vez de una camioneta o un SUV (vehículo deportivo utilitario). Las variables continuas son la lectura del odómetro y el octanaje de la gasolina empleada. La variable de respuesta se da en millas por galón. Los datos se presentan a continuación.

MP	Tipo de automóvil	Odómetro	Octanaje
34.5	sedan	75,000	87.5
33.3	sedan	60,000	87.5
30.4	sedan	88,000	78.0
32.8	sedan	15,000	78.0
35.0	sedan	25,000	90.0
29.0	sedan	35,000	78.0
32.5	sedan	102,000	90.0
29.6	sedan	98,000	87.5
16.8	van	56,000	87.5
19.2	van	72,000	90.0
22.6	van	14,500	87.5
24.4	van	22,000	90.0
20.7	van	66,500	78.0
25.1	van	35,000	90.0
18.8	van	97,500	87.5
15.8	van	65,500	78.0
17.4	van	42,000	78.0
15.6	SUV	65,000	78.0
17.3	SUV	55,500	87.5
20.8	SUV	26,500	87.5
22.2	SUV	11,500	90.0
16.5	SUV	38,000	78.0
21.3	SUV	77,500	90.0
20.7	SUV	19,500	78.0
24.1	SUV	87,000	90.0

- a) Ajuste un modelo de regresión lineal que incluya dos variables indicadoras. Utilice (0, 0) para denotar al sedán de cuatro puertas.
- b) ¿Qué tipo de vehículo parece tener un mayor rendimiento del combustible?

c) Analice la diferencia entre una camioneta y un SUV en términos del rendimiento del combustible.

12.46 Se efectuó un estudio para determinar si el género del titular de la tarjeta de crédito era un factor importante en la generación de utilidades para cierta empresa de tarjetas de crédito. Las variables consideradas fueron el ingreso, el número de miembros de la familia y el género del titular de la tarjeta. Los datos son los siguientes:

Utilidad	Ingreso	Género	Miembros de la familia
157	45,000	M	1
-181	55,000	M	2
-253	45,800	M	4
158	38,000	M	3
75	75,000	M	4
202	99,750	M	4
-451	28,000	M	1
146	39,000	M	2
89	54,350	M	1
-357	32,500	M	1
522	36,750	F	1
78	42,500	F	3
5	34,250	F	2
-177	36,750	F	3
123	24,500	F	2
251	27,500	F	1
-56	18,000	F	1
453	24,500	F	1
288	88,750	F	1
-104	19,750	F	2

- a) Ajuste un modelo de regresión lineal usando las variables disponibles. Con base en el modelo ajustado, ¿la empresa preferiría clientes del género masculino o del femenino?
- b) ¿Diría usted que el ingreso fue un factor importante para explicar la variabilidad de la utilidad?

12.9 Métodos secuenciales para la selección del modelo

En ocasiones las pruebas de significancia estudiadas en la sección 12.6 son muy adecuadas para determinar cuáles variables se deben usar en el modelo final de regresión. Dichas pruebas sin duda son eficaces si el experimento se puede planear y las variables son ortogonales entre sí. Incluso si las variables no son ortogonales, las pruebas t individuales se pueden usar en muchos problemas en donde se investigan pocas variables. Sin embargo, existen muchos problemas en los que es necesario utilizar técnicas más elaboradas para seleccionar las variables, en particular si el experimento exhibe una desviación sustancial de la ortogonalidad. Los coeficientes de correlación de la muestra r_{x_i, x_j} proporcionan medidas útiles de **multicolinealidad** (dependencia lineal) entre las

variables independientes. Como sólo estamos interesados en la dependencia lineal entre variables independientes, no nos confundiremos si eliminamos las x de la notación y sólo escribimos $r_{x_i, x_j} = r_{ij}$, donde

$$r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii} S_{jj}}}$$

Observe que, en sentido estricto, las r_{ij} no proporcionan estimados verdaderos de los coeficientes de correlación de la población, ya que las x en realidad no son variables aleatorias en el contexto que se estudia aquí. Así, el término *correlación*, aunque estándar, quizá sea inadecuado.

Cuando uno o más de esos coeficientes de correlación muestral se desvíe de manera sustancial de cero, suele ser muy difícil encontrar el subconjunto de variables más eficaz para incluirlo en la ecuación de predicción. De hecho, en ciertos problemas la multicolinealidad es tan extrema que no es posible encontrar un predictor adecuado, a menos que se investiguen todos los subconjuntos posibles de variables. En la bibliografía se citan los análisis informativos de Hocking (1976) para la selección de modelos de regresión. En la obra de Myers (1990), también citado, se estudian procedimientos para detectar la multicolinealidad.

El usuario de la regresión lineal múltiple busca lograr uno de tres objetivos:

1. Obtener estimados de coeficientes individuales en un modelo completo.
2. Estudiar variables para determinar cuáles tienen un efecto significativo sobre la respuesta.
3. Calcular la ecuación de predicción más eficaz.

En 1) se sabe de antemano que todas las variables deben incluirse en el modelo. En 2) la predicción es secundaria; mientras que en 3) los coeficientes de regresión individuales no son tan importantes como la calidad de la respuesta estimada \hat{y} . Para cada una de las situaciones anteriores la multicolinealidad en el experimento puede tener un efecto profundo sobre el éxito de la regresión.

En esta sección se estudian algunos procedimientos secuenciales estándar para seleccionar variables, los cuales se basan en la idea de que una sola variable o un conjunto de ellas no debería aparecer en la ecuación de estimación, a menos que origine un incremento significativo en la suma de cuadrados de regresión o, en forma equivalente, un incremento significativo de R^2 , el coeficiente de determinación múltiple.

Ilustración de la selección de las variables en presencia de colinealidad

Ejemplo 12.10: Considere los datos de la tabla 12.8, que muestra mediciones de 9 bebés. El objetivo del experimento era calcular una ecuación de estimación apropiada que relacionara la talla del bebé con todas las variables independientes o un subconjunto de ellas. Los coeficientes de correlación muestral, que indican la dependencia lineal entre las variables independientes, se incluyen en la matriz simétrica

x_1	x_2	x_3	x_4
1.0000	0.9523	0.5340	0.3900
0.9523	1.0000	0.2626	0.1549
0.5340	0.2626	1.0000	0.7847
0.3900	0.1549	0.7847	1.0000

Tabla 12.8: Datos relacionados con la talla de bebés*

Talla del bebé, y (cm)	Edad, x_1 (días)	Talla al nacer, x_2 (cm)	Peso al nacer, x_3 (kg)	Tamaño del pecho al nacer, x_4 (cm)
57.5	78	48.2	2.75	29.5
52.8	69	45.5	2.15	26.3
61.3	77	46.3	4.41	32.2
67.0	88	49.0	5.52	36.5
53.5	67	43.0	3.21	27.2
62.7	80	48.0	4.32	27.7
56.2	74	48.0	2.31	28.3
68.5	94	53.0	4.30	30.3
69.2	102	58.0	3.71	28.7

*Datos analizados por el Statistical Consulting Center, Virginia Tech, Blacksburg, Virginia.

Observe que parece haber una cantidad apreciable de multicolinealidad. Se utilizó la técnica de mínimos cuadrados descrita en la sección 12.2 y se usó el modelo completo para ajustar la ecuación de regresión estimada, la cual quedó como sigue:

$$\hat{y} = 7.1475 + 0.1000x_1 + 0.7264x_2 + 3.0758x_3 - 0.0300x_4.$$

El valor de s^2 con 4 grados de libertad es 0.7414, y se encontró que el valor del coeficiente de determinación para este modelo es 0.9908. En la tabla 12.9 se proporciona la suma de cuadrados de regresión que mide la variación atribuida a cada variable individual en presencia de las demás, así como los valores t correspondientes.

Tabla 12.9: Valores t para los datos de regresión de la tabla 12.8

Variable x_1	Variable x_2	Variable x_3	Variable x_4
$R(\beta_1 \beta_2, \beta_3, \beta_4)$	$R(\beta_2 \beta_1, \beta_3, \beta_4)$	$R(\beta_3 \beta_1, \beta_2, \beta_4)$	$R(\beta_4 \beta_1, \beta_2, \beta_3)$
= 0.0644	= 0.6334	= 6.2523	= 0.0241
$t = 0.2947$	$t = 0.9243$	$t = 2.9040$	$t = -0.1805$

Una región crítica de dos colas, con 4 grados de libertad y un nivel de significancia de 0.05, es dada por $|t| > 2.776$. De los cuatro valores t calculados sólo la variable x_3 parece ser significativa. Sin embargo, recuerde que aunque el estadístico t descrito en la sección 12.6 mide el beneficio que aporta una variable ajustada a todas las demás, no detecta la importancia potencial de una variable en combinación con un subconjunto de variables. Por ejemplo, considere el modelo sólo con las variables x_2 y x_3 en la ecuación. El análisis de los datos proporciona la función de regresión

$$\hat{y} = 2.1833 + 0.9576x_2 + 3.3253x_3,$$

con $R^2 = 0.9905$, que por supuesto no es una reducción sustancial de $R^2 = 0.9907$ para el modelo completo. Sin embargo, a menos que las características del desempeño de esta combinación particular hayan sido observadas, no estaríamos conscientes de su potencial predictivo. Esto, desde luego, apoya una metodología que observe *todas las regresiones posibles*, o un procedimiento secuencial sistemático diseñado para probar subconjuntos diferentes.

Regresión por etapas

Un procedimiento estándar para buscar el “subconjunto óptimo” de variables ante la ausencia de ortogonalidad es una técnica denominada **regresión por etapas**, que se basa en el procedimiento de introducir en forma secuencial las variables al modelo, una por una. Dado un tamaño α predeterminado, la descripción de la rutina por etapas se entenderá mejor si primero se describen los métodos de **selección hacia delante** y **eliminación hacia atrás**.

La **selección hacia delante** se basa en el concepto de que las variables deben insertarse una por una hasta obtener una ecuación de regresión satisfactoria. El procedimiento es como sigue:

PASO 1. Elija la variable que proporcione la mayor suma de cuadrados de regresión cuando se ejecute la regresión lineal simple con y o, en forma equivalente, aquella que proporcione el mayor valor de R^2 . Esta variable inicial se llamará x_1 . Si x_1 es insignificante, el procedimiento se suspende.

PASO 2. Seleccione la variable que al ser integrada al modelo proporciona el mayor incremento de R^2 , en presencia de x_1 , sobre la R^2 encontrada en el paso 1. Ésta, por supuesto, es la variable x_2 para la que

$$R(\beta_j | \beta_1) = R(\beta_1, \beta_j) - R(\beta_1)$$

es más grande. Dicha variable se llamará x_2 . Luego se ajusta el modelo de regresión con x_1 y x_2 , y se observa R^2 . Si x_2 es insignificante, el procedimiento se suspende.

PASO 3. Elija la variable x_j que proporciona el valor más grande de

$$R(\beta_j | \beta_1, \beta_2) = R(\beta_1, \beta_2, \beta_j) - R(\beta_1, \beta_2),$$

otra vez da como resultado el incremento mayor de R^2 sobre el que se obtuvo en el paso 2. A esta variable se le denomina x_3 , y ahora se tiene un modelo de regresión que incluye x_1 , x_2 y x_3 . Si x_3 es insignificante, el procedimiento se suspende.

Este proceso continúa hasta que la variable más reciente incluida ya no produce un incremento significativo en la regresión explicada. Tal incremento se puede determinar en cada paso utilizando adecuadamente una prueba F o una prueba t parciales. Por ejemplo, en el paso 2 el valor

$$f = \frac{R(\beta_2 | \beta_1)}{s^2}$$

se determina para probar la pertinencia de x_2 en el modelo. Aquí, el valor de s^2 es el cuadrado medio del error para el modelo que contiene las variables x_1 y x_2 . De manera similar, en el paso 3 la razón

$$f = \frac{R(\beta_3 | \beta_1, \beta_2)}{s^2}$$

prueba la pertinencia de x_3 en el modelo. Sin embargo, ahora el valor de s^2 es el cuadrado medio del error para el modelo que contiene las tres variables x_1 , x_2 y x_3 . Si en el paso 2, $f < f_\alpha(1, n - 3)$ para un nivel de significancia preseleccionado, x_2 no está incluida y el

proceso finaliza, lo que da como resultado una ecuación lineal simple que relaciona y y x_1 . Sin embargo, si $f > f_{\alpha}(1, n - 3)$, se avanza al paso 3. De nuevo, si en el paso 3, $f < f_{\alpha}(1, n - 4)$, entonces x_3 no se incluye y el proceso termina con la ecuación de la regresión apropiada que contiene las variables x_1 y x_2 .

La **eliminación hacia atrás** implica los mismos conceptos que la selección hacia delante, excepto que se comienza con todas las variables en el modelo. Por ejemplo, suponga que hay cinco variables en consideración. Los pasos son:

PASO 1. Ajuste una ecuación de regresión con las cinco variables incluidas en el modelo. Elija la variable que proporcione el valor más pequeño de la suma de cuadrados de regresión **ajustada para las demás**. Suponga que dicha variable es x_2 . Elimine x_2 del modelo si

$$f = \frac{R(\beta_2 \mid \beta_1, \beta_3, \beta_4, \beta_5)}{s^2}$$

es insignificante.

PASO 2. Ajuste una ecuación de regresión utilizando las variables restantes x_1, x_3, x_4 y x_5 , y repita el paso 1. Suponga que esta vez elige la variable x_5 . Nuevamente, si

$$f = \frac{R(\beta_5 \mid \beta_1, \beta_3, \beta_4)}{s^2}$$

es insignificante, se retira del modelo la variable x_5 . En cada paso la s^2 que se usa en la prueba F es el cuadrado medio del error para el modelo de regresión en esa etapa.

Este proceso se repite hasta que en algún paso la variable con la suma de cuadrados de regresión ajustada más pequeña produce un valor f significativo a un nivel de significancia predeterminado.

La **regresión por etapas** se lleva a cabo con una modificación ligera pero importante del procedimiento de selección hacia delante. La modificación requiere efectuar más pruebas en cada etapa para garantizar la eficacia continuada de las variables que se hubieran incluido en el modelo durante alguna etapa anterior. Esto representa una mejoría sobre la selección hacia delante, ya que es muy posible que una variable que haya entrado a la ecuación de regresión en una etapa temprana resulte poco importante o redundante debido a las relaciones que existen entre ella y las otras variables que se incluyeron en etapas posteriores. Por lo tanto, en la etapa en que se incluyó una variable nueva a la ecuación de regresión mediante un incremento significativo de R^2 , según lo determina la prueba F , todas las variables que ya estén en el modelo se someten a pruebas F (o bien, a pruebas t) a la luz de esta nueva variable, y si no muestran un valor f significativo, se eliminan. El procedimiento continúa hasta que se alcance una etapa donde ya no sea posible insertar ni eliminar variables adicionales. Este procedimiento por etapas se ilustra con el siguiente ejemplo.

Ejemplo 12.11: Utilice las técnicas de regresión por etapas y calcule un modelo de regresión lineal ajustado para predecir la talla de los bebés cuyos datos se presentan en la tabla 12.8.

Solución: **PASO 1.** Se considera cada variable por separado y se ajustan cuatro ecuaciones individuales de regresión lineal simple. Se calculan las siguientes sumas de cua-

datos de regresión pertinentes:

$$\begin{aligned} R(\beta_1) &= 288.1468, & R(\beta_2) &= 215.3013, \\ R(\beta_3) &= 186.1065, & R(\beta_4) &= 100.8594. \end{aligned}$$

Es evidente que la variable x_1 proporciona la suma de cuadrados de regresión más elevada. El cuadrado medio del error para la ecuación que implica sólo x_1 es $s^2 = 4.7276$, y como

$$f = \frac{R(\beta_1)}{s^2} = \frac{288.1468}{4.7276} = 60.9500,$$

que excede a $f_{0.05}(1, 7) = 5.59$, la variable x_1 es significativa y se introduce al modelo.

PASO 2. En esta etapa se ajustan tres ecuaciones de regresión y todas incluyen a x_1 . Los resultados importantes para las combinaciones (x_1, x_2) , (x_1, x_3) y (x_1, x_4) son

$$R(\beta_2|\beta_1) = 23.8703, \quad R(\beta_3|\beta_1) = 29.3086, \quad R(\beta_4|\beta_1) = 13.8178.$$

La variable x_3 muestra la mayor suma de cuadrados de regresión en presencia de x_1 . La regresión que implica x_1 y x_3 proporciona un valor nuevo de $s^2 = 0.6307$, y como

$$f = \frac{R(\beta_3|\beta_1)}{s^2} = \frac{29.3086}{0.6307} = 46.47,$$

que excede a $f_{0.05}(1, 6) = 5.99$, la variable x_3 es significativa y se incluye en el modelo junto con x_1 . Ahora debemos someter a x_1 a una prueba de significancia en presencia de x_3 . Encontramos que $R(\beta_1|\beta_3) = 131.349$, en consecuencia,

$$f = \frac{R(\beta_1|\beta_3)}{s^2} = \frac{131.349}{0.6307} = 208.26,$$

que es muy significativa. Por lo tanto, se mantiene x_1 junto con x_3 .

PASO 3. Con x_1 y x_3 incluidas en el modelo, ahora se requiere $R(\beta_2|\beta_1, \beta_3)$ y $R(\beta_4|\beta_1, \beta_3)$ para determinar cuál de las dos variables restantes, si es que acaso se puede incluir alguna, se debe incluir en esta etapa. Del análisis de regresión, usando x_2 junto con x_1 y x_3 , se observa que $R(\beta_2|\beta_1, \beta_3) = 0.7948$, y cuando x_4 se utiliza con x_1 y x_3 se obtiene $R(\beta_4|\beta_1, \beta_3) = 0.1855$. El valor de s^2 es 0.5979 para la combinación (x_1, x_2, x_3) , y 0.7198 para la combinación (x_1, x_3, x_4) . Como ningún valor f es significativo al nivel $\alpha = 0.05$, el modelo final de regresión sólo incluye las variables x_1 y x_3 . Se encuentra que la ecuación de estimación es

$$\hat{y} = 20.1084 + 0.4136x_1 + 2.0253x_3,$$

y el coeficiente de determinación para este modelo es $R^2 = 0.9882$.

Aunque (x_1, x_3) es la combinación elegida mediante la regresión por etapas, no es necesariamente la combinación de dos variables que proporciona el valor más grande de R^2 . De hecho, ya observamos que la combinación (x_2, x_3) da un valor de $R^2 = 0.9905$. Desde luego, el procedimiento por etapas nunca tomó en cuenta dicha combinación. Se podría plantear un argumento racional de que en realidad hay una diferencia despreciable en el desempeño entre esas dos ecuaciones de estimación, al menos en términos del

porcentaje de variación explicado. Sin embargo, es interesante observar que el procedimiento de eliminación hacia atrás proporciona la combinación (x_2, x_3) en la ecuación final (véase el ejercicio 12.49 en la página 494). ▮

Resumen

La función principal de cada uno de los procedimientos explicados en esta sección consiste en exponer las variables a una metodología sistemática, diseñada para garantizar la inclusión final de las mejores combinaciones de las mismas. Es evidente que no es seguro que esto pase en todos los problemas y, por supuesto, es posible que la multicolinealidad sea tan extensa que no haya más alternativa que apoyarse en procedimientos de estimación diferentes de los mínimos cuadrados. Tales procedimientos de estimación se estudian en Myers (1990), listado en la bibliografía.

Los procedimientos secuenciales que se estudian aquí son tres de los muchos métodos de ese tipo que aparecen en la literatura y que están incluidos en diversos paquetes de regresión por computadora. Estos métodos fueron diseñados para ser eficientes en cuanto al cálculo pero, por supuesto, no proporcionan resultados para todos los subconjuntos posibles de variables. Debido a esto los procedimientos son más eficaces para conjuntos de datos que incluyen un **número grande de variables**. En el caso de los problemas de regresión que implican un número relativamente pequeño de variables, los paquetes modernos de cómputo para la regresión permiten el cálculo y resumen la información cuantitativa de todos los modelos para cada subconjunto posible de variables. En la sección 12.11 se proporcionan ilustraciones.

Elección de valores P

Como es de esperarse, la elección del modelo final con estos procedimientos podría depender en gran medida del valor P que se seleccione. Además, un procedimiento es más exitoso cuando es forzado a probar una gran cantidad de variables posibles. Por esta razón, cualquier procedimiento hacia delante es más útil cuando se utiliza un valor P relativamente grande. A esto se debe que algunos programas de cómputo empleen un valor P predeterminado de 0.50.

12.10 Estudio de los residuales y violación de las suposiciones (verificación del modelo)

Anteriormente en este capítulo se sugirió que los residuales, o errores en el ajuste de regresión, con frecuencia proporcionan información que puede ser muy valiosa para el analista de datos. Los $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$, que son el equivalente numérico de los ϵ_i , los errores del modelo, a menudo revelan la posible violación de las suposiciones o la presencia de datos de puntos "sospechosos". Suponga que el vector \mathbf{x}_i denota los valores de las variables regresoras que corresponden al i -ésimo punto de los datos, complementado por un 1 en la posición inicial. Es decir,

$$\mathbf{x}'_i = [1, x_{1i}, x_{2i}, \dots, x_{ki}]$$

Considere la cantidad

$$h_{ii} = \mathbf{x}'_i (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_i, \quad i = 1, 2, \dots, n.$$

El lector debería notar que en la sección 12.5 se utilizó h_{ii} para calcular los intervalos de confianza de la respuesta media. Además de σ^2 , h_{ii} representa la varianza del valor ajustado \hat{y}_i . Los valores h_{ii} son los elementos de la diagonal de la matriz "SOMBRERO"

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

que desempeña un papel importante en cualquier estudio de residuales y en otros aspectos modernos del análisis de regresión (véase Myers, 1990, citado en la bibliografía). El término *matriz SOMBRERO* se deriva del hecho de que \mathbf{H} genera las "y sombrero", o los valores ajustados cuando se multiplica por el vector \mathbf{y} de respuestas observadas. Es decir, $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$, por lo tanto,

$$\hat{\mathbf{y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y}$$

donde \hat{y}_i es el vector cuyo i -ésimo elemento es \hat{y}_i .

Si se hacen las suposiciones acostumbradas de que los e_i son independientes y están distribuidos normalmente, con media cero y varianza σ^2 , las propiedades estadísticas de los residuales se establecen con facilidad. Entonces,

$$E(e_i) = E(y_i - \hat{y}_i) = 0 \quad \text{y} \quad \sigma_{e_i}^2 = (1 - h_{ii})\sigma^2$$

para $i = 1, 2, \dots, n$. (Para mayores detalles véase Myers, 1990). Es posible demostrar que los valores de la diagonal de la matriz SOMBRERO están acotados de acuerdo con la desigualdad

$$\frac{1}{n} \leq h_{ii} \leq 1.$$

Además, $\sum_{i=1}^n h_{ii} = k + 1$, el número de parámetros de la regresión. Como resultado, cualquier punto de los datos cuyo elemento diagonal SOMBRERO sea grande, es decir, esté muy por encima del valor promedio de $(k + 1)/n$, está en una posición dentro del conjunto de datos donde la varianza de \hat{y}_i es relativamente grande y la varianza de un residuo es relativamente pequeña. Como resultado, el analista de datos puede tener una idea de qué tan grande puede ser un residuo antes de que su desviación de cero se pueda atribuir a algo distinto del azar. Muchos de los paquetes comerciales para computadora que permiten calcular la regresión producen el conjunto de **residuales estudentizados**.

Residuo
estudentizado

$$r_i = \frac{e_i}{s\sqrt{1 - h_{ii}}}, \quad i = 1, 2, \dots, n$$

Aquí, cada residuo se **dividió entre una estimación de su desviación estándar** creando un estadístico *tipo t* diseñado para dar al analista una cantidad sin escala que proporcione información sobre el *tamaño* del residual. Además, a menudo los paquetes de cómputo comunes proporcionan valores de otro conjunto de residuales tipo estudentizados denominados **valores R de Student**.

Residual R de
Student

$$t_i = \frac{e_i}{s_{-i}\sqrt{1 - h_{ii}}}, \quad i = 1, 2, \dots, n.$$

donde s_{-i} es un estimador de la desviación estándar del error calculado con el i -ésimo punto de los datos eliminado.

Hay tres tipos de transgresiones de las suposiciones fáciles de detectar mediante el uso de los residuales o *gráficas de residuales*. Aunque las gráficas de los residuales brutos, los e_i , ayudan a esto, con frecuencia es más informativo graficar los residuales estudentizados. Las tres transgresiones son las siguientes:

1. Presencia de valores extremos
2. Varianza heterogénea del error
3. Especificación inadecuada del modelo

En el caso 1 elegimos definir un **valor extremo** como un punto de los datos que se desvía de la suposición común de que $E(\epsilon_i) = 0$ para un valor específico de i . Si hay una razón para creer que un punto de un dato específico es un valor extremo que ejerce una gran influencia sobre el modelo ajustado, r_i o t_i , esto podría estar informando algo. Es de esperarse que los valores R de Student sean más sensibles a los valores extremos que los valores r_i .

En realidad, en el caso de que $E(\epsilon_i) = 0$, t_i es un valor de una variable aleatoria que sigue una distribución t con $n - 1 - (k + 1) = n - k - 2$ grados de libertad. Por consiguiente, es posible utilizar una prueba t de dos colas para proporcionar información con el fin de detectar si el punto i -ésimo es o no un valor extremo.

Aunque el estadístico R de Student t_i produce una prueba t exacta para detectar un valor extremo en una ubicación específica, la distribución t no se aplicaría para probar simultáneamente varios valores extremos en todas las ubicaciones. Como resultado, los residuales estudentizados o valores R de Student se deberían usar estrictamente como herramientas de diagnóstico *sin* un mecanismo de prueba de hipótesis formal. La implicación es que dichos estadísticos resaltan puntos de los datos en los que el error del ajuste es mayor de lo esperado por el azar. Los valores R de Student de gran magnitud sugieren la necesidad de “verificar” los datos con todos los recursos disponibles. La práctica de eliminar observaciones de conjuntos de datos de la regresión no debería llevarse a cabo de forma indiscriminada. (Para más información sobre el uso de los diagnósticos sobre valores extremos véase Myers, 1990, en la bibliografía).

Ilustración de la detección de valores extremos

Estudio de caso 12.1: Método para capturar saltamontes. En un experimento biológico, que fue efectuado en el Departamento de Entomología de Virginia Tech, se hicieron n ensayos experimentales con dos métodos diferentes para capturar saltamontes. Los métodos consistieron en la captura por caída de la red y la captura por barrido de la red. El número promedio de saltamontes atrapados con cada método se registró en un conjunto de cuadrantes del campo en una fecha determinada. También se registró una variable regresora adicional, la altura promedio de las plantas en los cuadrantes. Los datos experimentales aparecen en la tabla 12.10.

El objetivo consiste en estimar cuántos saltamontes se capturan empleando sólo el método del barrido de la red, que es menos costoso. Hay cierta preocupación por la validez del cuarto punto de los datos. La captura observada utilizando el método de caída de la red que se reportó parece inusualmente alta, dadas las demás condiciones, de hecho se pensó que la cifra podía ser errónea. Ajuste un modelo del tipo

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$$

para los 17 puntos de los datos y estudie los residuales para determinar si el punto 4 es un valor extremo.

Tabla 12.10: Conjunto de datos para el estudio de caso 12.1

Observación	Captura por caída de la red, y	Captura por barrido de la red, x_1	Altura de las plantas, x_2 (cm)
1	18.0000	4.15476	52.705
2	8.8750	2.02381	42.069
3	2.0000	0.15909	34.766
4	20.0000	2.32812	27.622
5	2.3750	0.25521	45.879
6	2.7500	0.57292	97.472
7	3.3333	0.70139	102.062
8	1.0000	0.13542	97.790
9	1.3333	0.12121	88.265
10	1.7500	0.10937	58.737
11	4.1250	0.56250	42.386
12	12.8750	2.45312	31.274
13	5.3750	0.45312	31.750
14	28.0000	6.68750	35.401
15	4.7500	0.86979	64.516
16	1.7500	0.14583	25.241
17	0.1333	0.01562	36.354

Solución: Un paquete de cómputo generó el modelo de regresión ajustado

$$\hat{y} = 3.6870 + 4.1050x_1 - 0.0367x_2$$

junto con los estadísticos $R^2 = 0.9244$ y $s^2 = 5.580$. También se obtuvieron los residuales y otra información de diagnóstico que fueron registrados en la tabla 12.11.

Como se esperaba, el residual en la cuarta ubicación parece inusualmente grande, a saber, 7.769. La cuestión fundamental aquí es si este residual es más grande de lo que se esperaría debido al azar. El error estándar del residual para el punto 4 es 2.209. El valor R de Student t_4 que se obtuvo es 9.9315. Al considerarlo como el valor de una variable aleatoria que tiene una distribución t con 13 grados de libertad, se concluiría sin duda que el residuo de la cuarta observación se estima algo mayor que 0, y que la medición del presunto error es apoyada por el estudio de los residuales. Observe que ningún otro valor de los residuales proporciona un valor R de Student que sea motivo de alarma. ■

Gráfica de los residuales para el estudio de caso-12.1

En el capítulo 11 estudiamos con cierto detalle la utilidad de graficar los residuos en el análisis de regresión. Es frecuente que con base en dichas gráficas se detecte la violación de las suposiciones del modelo. En la regresión múltiple en ocasiones es útil graficar la probabilidad normal de los residuales o los residuales en comparación con \hat{y} . Sin embargo, a menudo es preferible graficar los residuales estudentizados.

Recuerde que la preferencia por los residuales estudentizados sobre los residuales ordinarios para propósitos de graficación se debe a que, como la varianza de i -ésimo

Tabla 12.11: Información sobre los residuales para el conjunto de datos del estudio de caso 12.1

Obs.	y_i	\hat{y}_i	$y_i - \hat{y}_i$	h_{ii}	$s\sqrt{1-h_{ii}}$	r_i	t_i
1	18.000	18.809	-0.809	0.2291	2.074	-0.390	-0.3780
2	8.875	10.452	-1.577	0.0766	2.270	-0.695	-0.6812
3	2.000	3.065	-1.065	0.1364	2.195	-0.485	-0.4715
4	20.000	12.231	7.769	0.1256	2.209	3.517	9.9315
5	2.375	3.052	-0.677	0.0931	2.250	-0.301	-0.2909
6	2.750	2.464	0.286	0.2276	2.076	0.138	0.1329
7	3.333	2.823	0.510	0.2669	2.023	0.252	0.2437
8	1.000	0.656	0.344	0.2318	2.071	0.166	0.1601
9	1.333	0.947	0.386	0.1691	2.153	0.179	0.1729
10	1.750	1.982	-0.232	0.0852	2.260	-0.103	-0.0989
11	4.125	4.442	-0.317	0.0884	2.255	-0.140	-0.1353
12	12.875	12.610	0.265	0.1152	2.222	0.119	0.1149
13	5.375	4.383	0.992	0.1339	2.199	0.451	0.4382
14	28.000	29.841	-1.841	0.6233	1.450	-1.270	-1.3005
15	4.750	4.891	-0.141	0.0699	2.278	-0.062	-0.0598
16	1.750	3.360	-1.610	0.1891	2.127	-0.757	-0.7447
17	0.133	2.418	-2.285	0.1386	2.193	-1.042	-1.0454

residuo depende del i -ésimo elemento en la diagonal SOMBRERO, las varianzas de los residuos diferirán si hay dispersión en las diagonales SOMBRERO. Así, es probable que la apariencia de una gráfica de residuales sugiera heterogeneidad debido a que los propios residuales no se comportan, en general, de manera ideal. El propósito de utilizar residuales estudentizados es proporcionar un tipo de *estandarización*. Es evidente que si se conociera σ , en condiciones ideales, es decir, en las que el modelo fuera correcto y la varianza homogénea, se tendría

$$E\left(\frac{e_i}{\sigma\sqrt{1-h_{ii}}}\right) = 0 \quad \text{y} \quad \text{Var}\left(\frac{e_i}{\sigma\sqrt{1-h_{ii}}}\right) = 1.$$

De manera que los residuales estudentizados producen un conjunto de estadísticos que en condiciones ideales se comportan en forma estándar. La figura 12.5 presenta una gráfica con los valores **R de Student** para los datos de los saltamontes del estudio de caso 12.1. Advierta que el valor para la observación 4 se destaca de los demás. La gráfica **R de Student** se generó con el programa *SAS*. La gráfica presenta los residuales en comparación con los valores \hat{y} .

Verificación de la normalidad

El lector debe recordar, de acuerdo con lo que se estudió en el capítulo 11, la importancia de verificar la normalidad utilizando una gráfica de probabilidad normal. La misma recomendación es válida para el caso de la regresión lineal múltiple. Las gráficas de probabilidad normal se pueden generar utilizando software estándar para regresión. Sin embargo, como ya se indicó, éstas pueden ser más eficaces si se usan residuales estudentizados o valores **R de Student** en vez de residuales comunes.

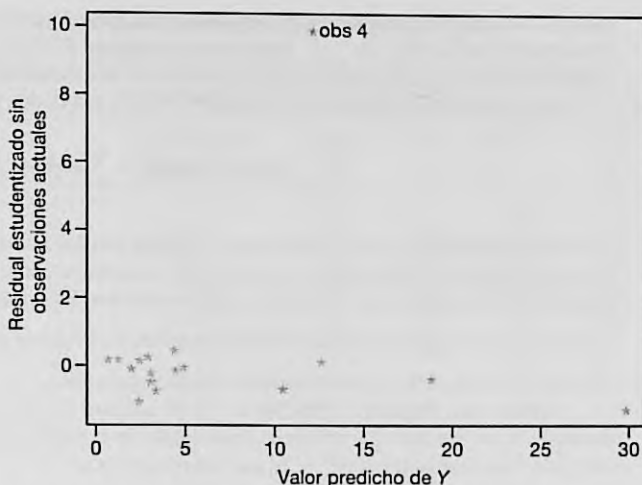


Figura 12.5: Valores R de Student graficados en comparación con los valores predichos para los datos de los saltamontes del estudio de caso 12.1.

12.11 Validación cruzada, C_p , y otros criterios para la selección del modelo

Para muchos problemas de regresión el experimentador debe elegir entre varios modelos alternativos o formas de modelo que se desarrollan a partir del mismo conjunto de datos. Con mucha frecuencia se requiere el modelo que predice o estima mejor la respuesta media. El experimentador debe tomar en cuenta los tamaños relativos de los valores de s^2 para los posibles modelos y, sin duda, la naturaleza general de los intervalos de confianza sobre la respuesta media. También se debe considerar lo bien que el modelo predice los valores de la respuesta que **no se hayan utilizado para construir los posibles modelos**. Los modelos deben estar sujetos a **validación cruzada**. Entonces, lo que se requiere son los errores de la validación cruzada en lugar de los errores del ajuste. Estos errores en la predicción son los **residuales PRESS**.

$$\delta_i = y_i - \hat{y}_{i,-i}, \quad i = 1, 2, \dots, n,$$

donde $\hat{y}_{i,-i}$ es la predicción del i -ésimo punto de los datos por medio de un modelo que no utiliza el i -ésimo punto en el cálculo de los coeficientes. Estos residuales PRESS se calculan mediante la fórmula

$$\delta_i = \frac{e_i}{1 - h_{ii}}, \quad i = 1, 2, \dots, n.$$

(La derivación se encuentra en Myers, 1990).

Uso del estadístico PRESS

La motivación para utilizar PRESS y la utilidad de los residuales PRESS es muy fácil de entender. El propósito de extraer o *separar* puntos de datos, uno a la vez, consiste en

permitir el empleo de metodologías separadas para ajustar y evaluar un modelo específico. Para evaluar un modelo la “-i” indica que el residual PRESS proporciona un error de predicción donde la observación que se predice es *independiente del ajuste del modelo*.

Los criterios que utilizan los residuales PRESS son dados por

$$\sum_{i=1}^n |\delta_i| \quad \text{y} \quad \text{PRESS} = \sum_{i=1}^n \delta_i^2.$$

(El término **PRESS** es un acrónimo que se forma con las iniciales de los términos de la frase en inglés *prediction sum of squares*, que se traduce como **suma de cuadrados de predicción**). Se sugiere que se utilicen ambos criterios. Es posible que PRESS sea dominado por uno o algunos residuales PRESS grandes. Es evidente que el criterio sobre $\sum_{i=1}^n |\delta_i|$ es menos sensible a un número pequeño de valores grandes.

Además del estadístico PRESS en sí, el analista puede simplemente calcular un estadístico similar a R^2 que refleje el desempeño de la predicción. Con frecuencia a este estadístico se le denomina R^2_{pred} y se calcula como sigue:

R^2 de predicción Dado un modelo ajustado con valor específico para PRESS, R^2_{pred} es dado por

$$R^2_{\text{pred}} = 1 - \frac{\text{PRESS}}{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

Observe que R^2_{pred} es tan sólo el estadístico común R^2 donde la *SCE* fue reemplazada por el estadístico PRESS.

En el siguiente estudio de caso se proporciona un ejemplo en el que se ajustan muchos posibles modelos a un conjunto de datos y se elige el mejor de ellos. No se emplean los procedimientos secuenciales descritos en la sección 12.9. En vez de eso se ilustra el papel que desempeñan los residuales PRESS y otros valores estadísticos cuando se trata de seleccionar la mejor ecuación de regresión.

Estudio de caso 12.2: Patada de fútbol. La fuerza de las piernas es un requisito necesario para que un pateador tenga éxito en el fútbol americano. Una medida de la calidad de una buena patada es el “tiempo de vuelo” del balón, es decir, el tiempo que el balón se mantiene en el aire antes de ser atrapado por el regresador de patadas. Para determinar cuáles factores de la fuerza de las piernas influyen en el tiempo de vuelo y desarrollar un modelo empírico para predecir esta respuesta, el Departamento de Salud, Educación Física y Recreación de Virginia Tech llevó a cabo un estudio sobre *La relación entre variables seleccionadas de desempeño físico y la capacidad de despejes en el fútbol*. Se eligieron 13 pateadores para el experimento y cada uno pateó 10 veces el balón. En la tabla 12.12 aparece el registro del tiempo de vuelo promedio, junto con las medidas de fuerza usadas en el análisis.

Cada variable regresora se define como sigue:

1. **FPD**, fuerza de la pierna derecha (libras)
2. **FPI**, fuerza de la pierna izquierda (libras)
3. **FTD**, flexibilidad muscular del tendón derecho (grados)
4. **FTI**, flexibilidad muscular del tendón izquierdo (grados)

5. Potencia, fuerza general de las piernas (pie-libras)

Determine el modelo más adecuado para predecir el tiempo de vuelo.

Tabla 12.12: Datos para el estudio de caso 12.2.

Pateador	Tiempo de vuelo, y (seg)	FPD, x_1	FPI, x_2	FTD, x_3	FTI, x_4	Potencia, x_5
1	4.75	170	170	106	106	240.57
2	4.07	140	130	92	93	195.49
3	4.04	180	170	93	78	152.99
4	4.18	160	160	103	93	197.09
5	4.35	170	150	104	93	266.56
6	4.16	150	150	101	87	260.56
7	4.43	170	180	108	106	219.25
8	3.20	110	110	86	92	132.68
9	3.02	120	110	90	86	130.24
10	3.64	130	120	85	80	205.88
11	3.68	120	140	89	83	153.92
12	3.60	140	130	92	94	154.64
13	3.85	160	150	95	95	240.57

Solución: Al buscar el mejor modelo posible para predecir el tiempo de vuelo se obtuvo la información de la tabla 12.13 utilizando un paquete de cómputo para regresión. Los modelos están clasificados en orden ascendente con respecto a los valores del estadístico PRESS. Esta presentación brinda información suficiente acerca de todos los modelos posibles con el fin de permitir que el usuario elimine algunos de ellos. Al parecer, el mejor modelo para predecir el tiempo de vuelo para los pateadores es el que contiene a x_2 y x_5 (FPI y potencia), denotadas por x_2, x_5 . Asimismo, observe que todos los modelos con valores bajos de PRESS, de s^2 , de $\sum_{i=1}^n |\delta_i|$, y con valores altos de R^2 , contienen esas dos variables. Para obtener información de los residuales de la regresión ajustada

$$\hat{y}_i = b_0 + b_2 x_{2i} + b_5 x_{5i},$$

se generaron los residuales y los residuales PRESS. El modelo de predicción real (véase el ejercicio 12.47 de la página 494) es dado por

$$\hat{y} = 1.10765 + 0.01370x_2 + 0.00429x_5.$$

En la tabla 12.14 se listan los residuales, los valores de la diagonal testada y los valores PRESS.

Observe el ajuste relativamente bueno de los modelos de regresión con dos variables para los datos. Los residuales PRESS reflejan la capacidad de la ecuación de regresión para predecir el tiempo de vuelo si se hicieran predicciones independientes. Por ejemplo, para el pateador número 4 el tiempo de vuelo de 4.180 tendría un error de predicción de 0.039 si se construyera el modelo usando a los 12 pateadores restantes. Para este modelo el error promedio de la predicción, o error de validación cruzada, es

$$\frac{1}{13} \sum_{i=1}^n |\delta_i| = 0.1489 \text{ segundos.}$$

Tabla 12.13: Comparación de diferentes modelos de regresión

Modelo	s^2	$\sum \delta_i $	PRESS	R^2
x_2x_5	0.036907	1.93583	0.54683	0.871300
$x_1x_2x_5$	0.041001	2.06489	0.58998	0.871321
$x_2x_4x_5$	0.037708	2.18797	0.59915	0.881658
$x_2x_3x_5$	0.039636	2.09553	0.66182	0.875606
$x_1x_2x_4x_5$	0.042265	2.42194	0.67840	0.882093
$x_1x_2x_3x_5$	0.044578	2.26283	0.70958	0.875642
$x_2x_3x_4x_5$	0.042421	2.55789	0.86236	0.881658
$x_1x_3x_5$	0.053664	2.65276	0.87325	0.831580
$x_1x_4x_5$	0.056279	2.75390	0.89551	0.823375
x_1x_5	0.059621	2.99434	0.97483	0.792094
x_2x_3	0.056153	2.95310	0.98815	0.804187
x_1x_3	0.059400	3.01436	0.99697	0.792864
$x_1x_2x_3x_4x_5$	0.048302	2.87302	1.00920	0.882096
x_2	0.066894	3.22319	1.04564	0.743404
x_3x_5	0.065678	3.09474	1.05708	0.770971
x_1x_2	0.068402	3.09047	1.09726	0.761474
x_3	0.074518	3.06754	1.13555	0.714161
$x_1x_3x_4$	0.065414	3.36304	1.15043	0.794705
$x_2x_3x_4$	0.062082	3.32392	1.17491	0.805163
x_2x_4	0.063744	3.59101	1.18531	0.777716
$x_1x_2x_3$	0.059670	3.41287	1.26558	0.812730
x_3x_4	0.080605	3.28004	1.28314	0.718921
x_1x_4	0.069965	3.64415	1.30194	0.756023
x_1	0.080208	3.31562	1.30275	0.692334
$x_1x_3x_4x_5$	0.059169	3.37362	1.36867	0.834936
$x_1x_2x_4$	0.064143	3.89402	1.39834	0.798692
$x_3x_4x_5$	0.072505	3.49695	1.42036	0.772450
$x_1x_2x_3x_4$	0.066088	3.95854	1.52344	0.815633
x_5	0.111779	4.17839	1.72511	0.571234
x_4x_5	0.105648	4.12729	1.87734	0.631593
x_4	0.186708	4.88870	2.82207	0.283819

que es pequeño comparado con el tiempo de vuelo promedio para los 13 pateadores. ▮

En la sección 12.9 indicamos que a menudo es aconsejable utilizar todos los subconjuntos posibles de regresión cuando se busca el mejor modelo. La mayoría de los programas comerciales de cómputo para estadística contienen una rutina de *todas las regresiones posibles*. Tales algoritmos calculan diversos criterios para todos los subconjuntos de términos del modelo. Es evidente que criterios como R^2 , s^2 y PRESS son razonables para elegir entre subconjuntos de candidatos. Otro estadístico muy popular y útil, en particular para las ciencias físicas e ingeniería, es el estadístico C_p , que se describe a continuación.

Tabla 12.14: Residuales PRESS

Pateador	y_i	\hat{y}_i	$e_i = y_i - \hat{y}_i$	h_{ii}	δ_i
1	4.750	4.470	0.280	0.198	0.349
2	4.070	3.728	0.342	0.118	0.388
3	4.040	4.094	-0.054	0.444	-0.097
4	4.180	4.146	0.034	0.132	0.039
5	4.350	4.307	0.043	0.286	0.060
6	4.160	4.281	-0.121	0.250	-0.161
7	4.430	4.515	-0.085	0.298	-0.121
8	3.200	3.184	0.016	0.294	0.023
9	3.020	3.174	-0.154	0.301	-0.220
10	3.640	3.636	0.004	0.231	0.005
11	3.680	3.687	-0.007	0.152	-0.008
12	3.600	3.553	0.047	0.142	0.055
13	3.850	4.196	-0.346	0.154	-0.409

El estadístico C_p

Muy a menudo la selección del modelo más adecuado implica tomar en cuenta muchas cosas. Evidentemente el número de términos del modelo es importante; el tema de la parsimonia no debe ignorarse. Por otro lado, el analista no debe sentirse satisfecho con un modelo demasiado simple hasta el punto de una simplificación excesiva. En este sentido, un estadístico único que implica un compromiso aceptable es C_p . (Véase Mallows, 1973, en la bibliografía).

El estadístico C_p apela de forma muy adecuada al sentido común y se desarrolla tomando en cuenta el equilibrio apropiado entre el sesgo excesivo en que se incurre cuando se subajusta, es decir, cuando se eligen muy pocos términos para el modelo; y la varianza excesiva de la predicción que se genera cuando se sobreajusta, o sea cuando hay redundancias en el modelo. El estadístico C_p es una función simple del número total de parámetros en el posible modelo y la media cuadrada del error s^2 .

Aquí no presentaremos el desarrollo completo del estadístico C_p . (Para mayores detalles se recomienda consultar a Myers, 1990, listado en la bibliografía). El C_p para un subconjunto particular de modelos es un *estimado* de lo siguiente:

$$\Gamma_{(p)} = \frac{1}{\sigma^2} \sum_{i=1}^n \text{Var}(\hat{y}_i) + \frac{1}{\sigma^2} \sum_{i=1}^n (\text{Sesgo } \hat{y}_i)^2.$$

Se descubre que, bajo las suposiciones estándar de los mínimos cuadrados que se indicaron con anterioridad en este capítulo, y asumiendo que el modelo "verdadero" es aquel que contiene todas las posibles variables,

$$\frac{1}{\sigma^2} \sum_{i=1}^n \text{Var}(\hat{y}_i) = p \quad (\text{número de parámetros en el posible modelo})$$

(véase el ejercicio de repaso 12.63) y un estimado no sesgado de

$$\frac{1}{\sigma^2} \sum_{i=1}^n (\text{Sesgo } \hat{y}_i)^2 \text{ es dado por } \frac{1}{\sigma^2} \sum_{i=1}^n (\widehat{\text{Sesgo}} \hat{y}_i)^2 = \frac{(s^2 - \sigma^2)(n - p)}{\sigma^2}.$$

En las ecuaciones anteriores s^2 es el cuadrado medio del error para el posible modelo y σ^2 es la varianza del error de la población. Así, si asumimos que se dispone de algún estimado $\hat{\sigma}^2$ para σ^2 , entonces C_p es dado por la siguiente ecuación:

Estadístico C_p

$$C_p = p + \frac{(s^2 - \hat{\sigma}^2)(n - p)}{\hat{\sigma}^2},$$

donde p es el número de parámetros en el modelo, s^2 es el cuadrado medio del error para el modelo candidato y $\hat{\sigma}^2$ es un estimador de σ^2 .

Es evidente que el científico debería adoptar modelos con valores pequeños de C_p . El lector observará que, a diferencia del estadístico PRESS, C_p carece de una escala. Además, se puede obtener cierta información acerca de qué tan adecuado es un posible modelo observando su valor de C_p . Por ejemplo, $C_p > p$ indica que un modelo está sesgado debido a que está subajustado, mientras que $C_p \approx p$ indica un modelo razonable.

Con frecuencia hay confusión respecto a la precedencia de $\hat{\sigma}^2$ en la fórmula para C_p . Es evidente que el científico o ingeniero no tienen acceso a la cantidad σ^2 de la población. En aplicaciones donde se dispone de corridas repetidas, digamos en situaciones de diseño experimental, se dispone de un estimado de σ^2 independiente del modelo (véase los capítulos 11 y 15). Sin embargo, la mayoría de paquetes de cómputo utilizan $\hat{\sigma}^2$ como el *cuadrado medio del error del modelo más completo*. Evidentemente, si éste no es un buen estimado, la parte de sesgo del estadístico C_p puede ser negativa. Por consiguiente, C_p puede ser menor que p .

Ejemplo 12.12: Considere el conjunto de datos de la tabla 12.15, los cuales reflejan el interés de un fabricante de grava asfáltica en la relación que existe entre las ventas durante un año específico y los factores que influyen en ellas. (Los datos fueron tomados de Kutner *et al.*, 2004, véase la bibliografía).

En los subconjuntos de modelos posibles, hay tres que revisten interés especial. Estos tres son los de x_2x_3 , $x_1x_2x_3$ y $x_1x_2x_3x_4$. A continuación se presenta la información pertinente para comparar los tres modelos. Para ayudar a la toma de decisiones incluimos los estadísticos PRESS de los tres modelos.

Modelo	R^2	R^2_{pred}	s^2	PRESS	C_p
x_2x_3	0.9940	0.9913	44.5552	782.1896	11.4013
$x_1x_2x_3$	0.9970	0.9928	24.7956	643.3578	3.4075
$x_1x_2x_3x_4$	0.9971	0.9917	26.2073	741.7557	5.0

A partir de la información de la tabla parece claro que el modelo $x_1x_2x_3$ es mejor que los otros dos. Observe que para el modelo completo $C_p = 5.0$. Esto ocurre porque la *parte de sesgo* es igual a cero y $\hat{\sigma}^2 = 26.2073$ es el cuadrado medio del error del modelo completo.

La figura 12.6 es una salida de resultados de la función PROC REG del SAS, la cual muestra información sobre todas las regresiones posibles. A partir de ella es posible hacer comparaciones de otros modelos con (x_1, x_2, x_3) . Observe que (x_1, x_2, x_3) parece muy bueno en comparación con todos los demás modelos.

Como verificación final del modelo (x_1, x_2, x_3) , la figura 12.7 presenta una gráfica de probabilidad normal de los residuales del modelo.

Tabla 12.15: Datos para el ejemplo 12.12

Distrito	Cuentas promocionales, x_1	Cuentas activas, x_2	Marcas en competencia, x_3	Potencial, x_4	Ventas, y (miles)
1	5.5	31	10	8	\$ 79.3
2	2.5	55	8	6	200.1
3	8.0	67	12	9	163.2
4	3.0	50	7	16	200.1
5	3.0	38	8	15	146.0
6	2.9	71	12	17	177.7
7	8.0	30	12	8	30.9
8	9.0	56	5	10	291.9
9	4.0	42	8	4	160.0
10	6.5	73	5	16	339.4
11	5.5	60	11	7	159.6
12	5.0	44	12	12	86.3
13	6.0	50	6	6	237.5
14	5.0	39	10	4	107.2
15	3.5	55	10	4	155.0

Dependent Variable: sales						
Number in Model	C(p)	R-Square	Adjusted R-Square	MSE	Variables in Model	
3	3.4075	0.9970	0.9961	24.79560	x1	x2 x3
4	5.0000	0.9971	0.9959	26.20728	x1	x2 x3 x4
2	11.4013	0.9940	0.9930	44.55518	x2	x3
3	13.3770	0.9940	0.9924	48.54787	x2	x3 x4
3	1053.643	0.6896	0.6049	2526.96144	x1	x3 x4
2	1082.670	0.6805	0.6273	2384.14286	x3	x4
2	1215.316	0.6417	0.5820	2673.83349	x1	x3
1	1228.460	0.6373	0.6094	2498.68333	x3	
3	1653.770	0.5140	0.3814	3956.75275	x1	x2 x4
2	1668.699	0.5090	0.4272	3663.99357	x1	x2
2	1685.024	0.5042	0.4216	3699.64814	x2	x4
1	1693.971	0.5010	0.4626	3437.12846	x2	
2	3014.641	0.1151	-.0324	6603.45109	x1	x4
1	3088.650	0.0928	0.0231	6248.72283	x4	
1	3364.884	0.0120	-.0640	6805.59568	x1	

Figura 12.6: Salida de resultados del SAS de todos los subconjuntos posibles sobre los datos de las ventas para el ejemplo 12.12.

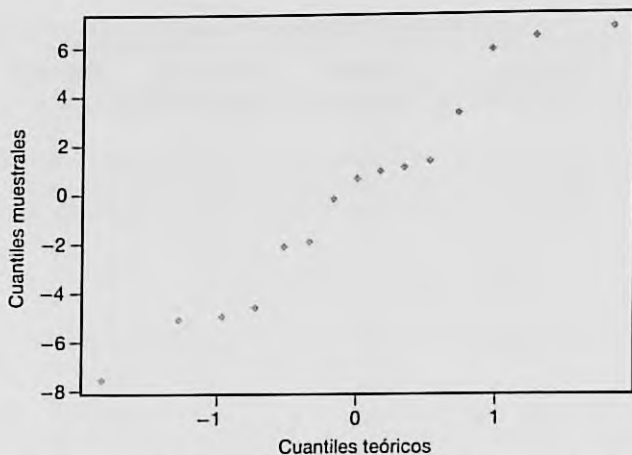


Figura 12.7: Gráfica de la probabilidad normal de los residuales, utilizando el modelo x_1, x_2, x_3 , para el ejemplo 12.12.

Ejercicios

12.47 Considere los datos sobre el “tiempo de vuelo” de los pateadores que se presentaron en el estudio de caso 12.2 y, utilizando sólo las variables x_2 y x_3 ,

- Verifique la ecuación de regresión que se presenta en la página 489.
- Prediga el tiempo de vuelo para un pateador con FPI = 180 libras y potencia = 260 pie-libras.
- Construya un intervalo de confianza de 95% para el tiempo de vuelo promedio de un pateador con FPI = 180 libras y potencia = 260 pies-libras.

12.48 Para los datos del ejercicio 12.15 de la página 452 utilice las técnicas de

- selección hacia delante* a un nivel de significancia de 0.05 para elegir un modelo de regresión lineal;
- eliminación hacia atrás* a un nivel de significancia de 0.05 para seleccionar un modelo de regresión lineal;
- regresión por etapas* a un nivel de significancia de 0.05 para escoger un modelo de regresión lineal.

12.49 Emplee las técnicas de *eliminación hacia atrás* con $\alpha = 0.05$ para elegir una ecuación de predicción para los datos de la tabla 12.8.

12.50 Para los datos de los pateadores del estudio de caso 12.2 también se registró una respuesta adicional, la “distancia de la patada”. Los siguientes son los valores de distancia promedio para cada uno de los 13 pateadores:

- Utilice los datos de distancia en lugar de los de tiempo de vuelo para estimar un modelo de regresión lineal múltiple del tipo

$$\mu_Y | x_1, x_2, x_3, x_4, x_5$$

$$= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5$$

que permita predecir la distancia de la patada.

- Utilice la regresión por etapas a un nivel de significancia de 0.10 para seleccionar una combinación de variables.
- Genere valores para s^2 , R^2 , PRESS y $\sum_{i=1}^n |\delta_i|$ para el

conjunto completo de 31 modelos. Utilice esta información para determinar la mejor combinación de variables para predecir la distancia de la patada.

- Para el modelo final que seleccione, grafique los residuales estandarizados en comparación con Y y elabore una gráfica de probabilidad normal de los residuales ordinarios. Haga comentarios al respecto.

Pateador Distancia, y (pies)

1	162.50
2	144.00
3	147.50
4	163.50
5	192.00
6	171.75
7	162.00
8	104.93
9	105.67
10	117.59
11	140.25
12	150.17
13	165.16

12.51 El siguiente es un conjunto de datos para y , la cantidad de dinero (en miles de dólares) aportado a la asociación de exalumnos del Virginia Tech por la generación de 1960; y para x , el número de años que han transcurrido desde la graduación:

y	x	y	x
812.52	1	2755.00	11
822.50	2	4390.50	12
1211.50	3	5581.50	13
1348.00	4	5548.00	14
1301.00	8	6086.00	15
2567.50	9	5764.00	16
2526.50	10	8903.00	17

a) Ajuste un modelo de regresión del tipo

$$\mu_{Y|x} = \beta_0 + \beta_1 x.$$

b) Ajuste un modelo cuadrático del tipo

$$\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_{11} x^2.$$

c) Determine cuál de los modelos de los incisos a) o b) es preferible. Utilice s^2 , R^2 y los residuales PRESS para sustentar su decisión.

12.52 Para el modelo del ejercicio 12.50a) pruebe la hipótesis

$$H_0: \beta_4 = 0,$$

$$H_1: \beta_4 \neq 0.$$

Utilice un valor P para su conclusión.

12.53 Para el modelo cuadrático del ejercicio 12.51b) proporcione estimados de las varianzas y las covarianzas de los estimados de β_1 y β_{11} .

12.54 Un cliente del Departamento de Ingeniería Mecánica se acercó al Centro de Consulta de Virginia Tech para que lo ayudaran a analizar un experimento sobre motores con turbina de gas. Se midieron varias salidas del voltaje de los motores con distintas combinaciones de velocidad de las aspas y de la extensión de los sensores. Los datos son los siguientes:

y (voltios)	Velocidad, x_1 (pulg/seg)	Extensión, x_2 (pulg)
1.95	6336	0.000
2.50	7099	0.000
2.93	8026	0.000
1.69	6230	0.000
1.23	5369	0.000
3.13	8343	0.000
1.55	6522	0.006
1.94	7310	0.006
2.18	7974	0.006
2.70	8501	0.006
1.32	6646	0.012
1.60	7384	0.012
1.89	8000	0.012
2.15	8545	0.012
1.09	6755	0.018
1.26	7362	0.018
1.57	7934	0.018
1.92	8554	0.018

a) Ajuste una regresión lineal múltiple a los datos.

b) Calcule las pruebas t sobre los coeficientes. Proporcione valores P .

c) Haga comentarios sobre la calidad del modelo ajustado.

12.55 La blancura del rayón es un factor importante para los científicos que estudian la calidad de las telas. La blancura se ve afectada por la calidad de la pulpa y otras variables de procesamiento. Algunas de las variables son la temperatura del baño con ácido, $^{\circ}\text{C}$ (x_1); la concentración del ácido en cascada, % (x_2); la temperatura del agua, $^{\circ}\text{C}$ (x_3); la concentración del sulfuro, % (x_4); la cantidad del blanqueador de cloro, lb/min (x_5) y la temperatura de terminado de la tela, $^{\circ}\text{C}$ (x_6). A continuación se proporciona un conjunto de datos de especímenes de rayón. La respuesta, y , es la medida de la blancura.

y	x_1	x_2	x_3	x_4	x_5	x_6
88.7	43	0.211	85	0.243	0.606	48
89.3	42	0.604	89	0.237	0.600	55
75.5	47	0.450	87	0.198	0.527	61
92.1	46	0.641	90	0.194	0.500	65
83.4	52	0.370	93	0.198	0.485	54
44.8	50	0.526	85	0.221	0.533	60
50.9	43	0.486	83	0.203	0.510	57
78.0	49	0.504	93	0.279	0.489	49
86.8	51	0.609	90	0.220	0.462	64
47.3	51	0.702	86	0.198	0.478	63
53.7	48	0.397	92	0.231	0.411	61
92.0	46	0.488	88	0.211	0.387	88
87.9	43	0.525	85	0.199	0.437	63
90.3	45	0.486	84	0.189	0.499	58
94.2	53	0.527	87	0.245	0.530	65
89.5	47	0.601	95	0.208	0.500	67

a) Utilice los criterios CME , c_p y PRESS para obtener el "mejor" modelo de todos los subconjuntos de los modelos.

b) Grafique los residuales estandarizados en comparación con Y y dibuje una gráfica de probabilidad normal de los residuales para el "mejor" modelo. Comente sus resultados.

12.56 En un esfuerzo para modelar las remuneraciones de los ejecutivos en el año 1979 se seleccionaron 33 empresas y se recabaron datos acerca de las remuneraciones, las ventas, las utilidades y el empleo. Se reunieron los siguientes datos para el año 1979.

Empresa	Remuneraciones, y (miles)	Ventas, x_1 , (millones)	Utilidades, x_2 , (millones)	Empleo, x_3
1	\$450	\$4600.6	\$128.1	48,000
2	387	9255.4	783.9	55,900
3	368	1526.2	136.0	13,783
4	277	1683.2	179.0	27,765
5	676	2752.8	231.5	34,000
6	454	2205.8	329.5	26,500
7	507	2384.6	381.8	30,800
8	496	2746.0	237.9	41,000
9	487	1434.0	222.3	25,900

(cont.)

Empresa	Remuneraciones, y (miles)	Ventas, x_1 (millones)	Utilidades, x_2 (millones)	Empleo, x_3
10	\$383	\$470.6	\$63.7	8600
11	311	1508.0	149.5	21,075
12	271	464.4	30.0	6874
13	524	9329.3	577.3	39,000
14	498	2377.5	250.7	34,300
15	343	1174.3	82.6	19,405
16	354	409.3	61.5	3586
17	324	724.7	90.8	3905
18	225	578.9	63.3	4139
19	254	966.8	42.8	6255
20	208	591.0	48.5	10,605
21	518	4933.1	310.6	65,392
22	406	7613.2	491.6	89,400
23	332	3457.4	228.0	55,200
24	340	545.3	54.6	7800
25	698	22,862.8	3011.3	337,119
26	306	2361.0	203.0	52,000
27	613	2614.1	201.0	50,500
28	302	1013.2	121.3	18,625
29	540	4560.3	194.6	97,937
30	293	855.7	63.4	12,300
31	528	4211.6	352.1	71,800
32	456	5440.4	655.2	87,700
33	417	1229.9	97.5	14,600

Considere el modelo

$$y_i = \beta_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + \beta_3 \ln x_{3i} + \epsilon_i, \quad i = 1, 2, \dots, 33.$$

- Ajuste la regresión con el modelo anterior.
- ¿Un modelo con un subconjunto de variables es preferible al modelo completo?

12.57 La resistencia a la tracción de una unión de alambre es una característica importante. La siguiente tabla brinda información sobre la resistencia a la tracción, y , la altura del molde, x_1 , la altura del perno, x_2 , la altura del lazo, x_3 , la longitud del alambre, x_4 , el ancho de la unión sobre el molde, x_5 y el ancho del molde sobre el perno, x_6 . (Datos tomados de Myers, Montgomery y Anderson-Cook, 2009).

- Ajuste un modelo de regresión usando todas las variables independientes.
- Utilice la regresión por etapas a un nivel de significancia de entrada de 0.25 y un nivel de significancia de eliminación de 0.05. Proporcione el modelo final.
- Utilice todos los modelos de regresión posibles y

calcule R^2 , C_p , s^2 y R^2 ajustada para todos los modelos.

- Proporcione el modelo final.
- Para el modelo del inciso *d*) grafique los residuos estudentizados (o la R de Student) y haga comentarios al respecto.

y	x_1	x_2	x_3	x_4	x_5	x_6
8.0	5.2	19.6	29.6	94.9	2.1	2.3
8.3	5.2	19.8	32.4	89.7	2.1	1.8
8.5	5.8	19.6	31.0	96.2	2.0	2.0
8.8	6.4	19.4	32.4	95.6	2.2	2.1
9.0	5.8	18.6	28.6	86.5	2.0	1.8
9.3	5.2	18.8	30.6	84.5	2.1	2.1
9.3	5.6	20.4	32.4	88.8	2.2	1.9
9.5	6.0	19.0	32.6	85.7	2.1	1.9
9.8	5.2	20.8	32.2	93.6	2.3	2.1
10.0	5.8	19.9	31.8	86.0	2.1	1.8
10.3	6.4	18.0	32.6	87.1	2.0	1.6
10.5	6.0	20.6	33.4	93.1	2.1	2.1
10.8	6.2	20.2	31.8	83.4	2.2	2.1
11.0	6.2	20.2	32.4	94.5	2.1	1.9
11.3	6.2	19.2	31.4	83.4	1.9	1.8
11.5	5.6	17.0	33.2	85.2	2.1	2.1
11.8	6.0	19.8	35.4	84.1	2.0	1.8
12.3	5.8	18.8	34.0	86.9	2.1	1.8
12.5	5.6	18.6	34.2	83.0	1.9	2.0

12.58 Para el ejercicio 12.57 pruebe $H_0: \beta_1 = \beta_6 = 0$. Proporcione valores P y comente al respecto.

12.59 En el ejercicio 12.28 de la página 462 se tienen los siguientes datos sobre el desgaste de un cojinete:

y (desgaste)	x_1 (viscosidad del aceite)	x_2 (carga)
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

a) Puede considerar el siguiente modelo para describir los datos:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} + \epsilon_i,$$

para $i = 1, 2, \dots, 6$. El término $x_1 x_2$ es una "interacción". Ajuste este modelo y estime los parámetros.

b) Utilice los modelos (x_1) , (x_1, x_2) , (x_2) , $(x_1, x_2, x_1 x_2)$ y calcule PRESS, C_p , y s^2 para determinar el "mejor" modelo.

12.12 Modelos especiales no lineales para condiciones no ideales

En gran parte del material anterior de este capítulo y en el del capítulo 11 nos hemos beneficiado mucho de la suposición de que los errores del modelo, los ϵ_i , son normales.

con media igual a cero y varianza constante σ^2 . Sin embargo, en la vida real hay muchas situaciones en las cuales es evidente que la respuesta no es normal. Por ejemplo, existe una gran cantidad de aplicaciones en las que la **respuesta es binaria** (0 o 1), por lo que su naturaleza es de Bernoulli. En las ciencias sociales un problema podría ser el de desarrollar un modelo que prediga si un individuo representa riesgos para un crédito (0 o 1), en función de ciertos regresores socioeconómicos, como sus ingresos, edad, género y nivel de escolaridad. En una prueba biomédica para un fármaco a menudo se observa si el paciente responde o no de manera favorable a éste, en tanto que los regresores podrían incluir la dosis y factores biológicos como la edad, el peso y la presión sanguínea. Nuevamente la respuesta es de naturaleza binaria. También abundan las aplicaciones en las áreas de manufactura en que ciertos factores controlables influyen en el hecho de que un artículo fabricado esté o no defectuoso.

Un segundo tipo de aplicación que no es normal y del que haremos una mención breve tiene que ver con el **conteo de datos**. Aquí a menudo es conveniente suponer una respuesta de Poisson. En aplicaciones biomédicas la respuesta que se modela en comparación con las dosis de medicamentos podría ser el número de colonias de células cancerosas. En la industria textil una respuesta razonable que se modela en comparación con ciertas variables de los procesos es el número de imperfecciones por yarda de tela.

Varianza no homogénea

El lector debería notar la comparación de la situación ideal, es decir, la respuesta normal, con la de la respuesta de Bernoulli (o binomial) o la de Poisson. Nos hemos acostumbrado al hecho de que el caso normal es muy especial debido a que la varianza es **independiente de la media**. Resulta claro que éste no es el caso para la respuesta de Bernoulli ni la de Poisson. Por ejemplo, si la respuesta es 0 o 1, lo cual sugiere una respuesta de Bernoulli, entonces el modelo adopta la forma

$$p = f(\mathbf{x}, \beta),$$

donde p es la **probabilidad de un éxito** (por ejemplo, la respuesta = 1). El parámetro p desempeña el papel de $\mu_{Y|x}$ en el caso normal. Sin embargo, la varianza de Bernoulli es $p(1 - p)$ que, desde luego, también es una función del regresor \mathbf{x} . Como resultado, la varianza no es constante. Esto descarta el uso de los mínimos cuadrados estándar que hemos utilizado en nuestro trabajo de regresión lineal hasta este momento. Lo mismo se aplica para el caso de Poisson, ya que el modelo adopta la forma

$$\lambda = f(\mathbf{x}, \beta),$$

con $\text{Var}(y) = \mu_y = \lambda$, que varía con \mathbf{x} .

Respuesta binaria (regresión logística)

El enfoque más popular para modelar respuestas binarias es la técnica llamada **regresión logística**, la cual se emplea mucho en las ciencias biológicas, en la investigación biomédica y en la ingeniería. De hecho, se observa que incluso en las ciencias sociales abundan las respuestas binarias. La distribución básica para la respuesta es la de Bernoulli o la binomial. La primera se encuentra en estudios observacionales donde no hay corridas repetidas en cada nivel de regresor; mientras que la segunda será el caso en que se utilice un diseño experimental. Por ejemplo, en un ensayo clínico en el cual se evalúa un fármaco nuevo, el objetivo podría ser el de determinar la dosis del medicamento que es

eficaz. Así, en el experimento se utilizarán ciertas dosis y para cada una de ellas se emplearán a varios sujetos, un caso al que se le denomina **caso agrupado**.

¿Cuál es el modelo para la regresión logística?

En el caso de respuestas binarias la respuesta media es una probabilidad. En la ilustración anterior del ensayo clínico podríamos decir que deseamos estimar la probabilidad de que el paciente responda en forma adecuada al fármaco, $P(\text{éxito})$. Entonces, el modelo se escribe en términos de una probabilidad. Dados los regresores x , la función logística es dada por

$$p = \frac{1}{1 + e^{-x'\beta}}$$

La porción $x'\beta$ se llama **predicador lineal** y, en el caso de un solo regresor x , se puede escribir $x'\beta = \beta_0 + \beta_1 x$. Desde luego, no descartamos la inclusión de regresores múltiples y de términos polinomiales en el llamado predicador lineal. En el caso agrupado el modelo implica el modelado de la media de una binomial en vez de una de Bernoulli, por lo que la media es dada por

$$np = \frac{n}{1 + e^{-x'\beta}}$$

Características de la función logística

Una gráfica de la función logística revela mucho sobre sus características y del porqué se utiliza para este tipo de problema. En primer lugar, la función es no lineal. Además, la gráfica de la figura 12.8 revela la forma de S con la función que tiende a la asíntota en $p = 1.0$. En este caso, $\beta_1 > 0$. Así, nunca se experimentaría una probabilidad estimada mayor que 1.0.

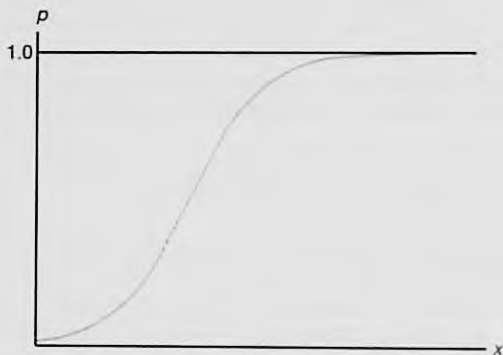


Figura 12.8: La función logística.

Los coeficientes de regresión en el predicador lineal se estiman con el método de probabilidad máxima, tal como se describió en el capítulo 9. La solución de las ecuaciones

de probabilidad requiere una metodología iterativa que no se describe aquí. Sin embargo, presentaremos un ejemplo y analizaremos la salida de resultados por computadora y las conclusiones.

Ejemplo 12.13: El conjunto de datos de la tabla 12.16 se utilizará con el fin de ilustrar el uso de la regresión logística para analizar un ensayo biológico cuantal de agente único en un experimento de toxicidad. Los resultados muestran el efecto de diferentes dosis de nicotina en la mosca común de la fruta.

Tabla 12.16: Conjunto de datos para el ejemplo 12.13

x Concentración (gramos/100 cc)	n_i Número de insectos	y Número de muertes	Porcentaje de muertes
0.10	47	8	17.0
0.15	53	14	26.4
0.20	55	24	43.6
0.30	52	32	61.5
0.50	46	38	82.6
0.70	54	50	92.6
0.95	52	50	96.2

El propósito del experimento era el de obtener un modelo adecuado que relacionara la probabilidad de “muerte” con la concentración. Además, el analista buscaba la denominada **dosis eficaz** (DE), es decir, la concentración de nicotina que da como resultado cierta probabilidad. La DE_{50} tiene interés particular, ya que es la concentración que produce una probabilidad de 0.5 de que el “insecto muera”.

Este ejemplo es agrupado, por lo que el modelo es dado por

$$E(Y_i) = n_i p_i = \frac{n_i}{1 + e^{-(\beta_0 + \beta_1 x_i)}}$$

Los estimados de β_0 y β_1 , y sus errores estándar, se calculan usando el método de probabilidad máxima. Las pruebas de los coeficientes individuales se calculan utilizando el estadístico χ^2 en lugar del estadístico t , puesto que no hay una varianza común σ^2 . El estadístico χ^2 se obtiene a partir de (coef/valor estándar)².

Por consiguiente, obtenemos la siguiente salida de resultados por computadora de la función PROC LOGIST del SAS.

Análisis de los estimados de los parámetros					
	gl	Estimado	Error estándar	Chi cuadrada	Valor P
β_0	1	-1.7361	0.2420	51.4482	< 0.0001
β_1	1	6.2954	0.7422	71.9399	< 0.0001

Ambos coeficientes difieren significativamente de cero. Por consiguiente, el modelo ajustado que se emplea para predecir la probabilidad de “muerte” es dado por

$$\hat{p} = \frac{1}{1 + e^{-(-1.7361 + 6.2954 x)}}$$

Estimado de la dosis eficaz

El estimado de la DE_{50} para el ejemplo 12.13 se calcula de manera muy sencilla a partir de los estimados b_0 para β_0 y b_1 para β_1 . A partir de la función logística se observa que

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x.$$

Como resultado, para $p = 0.5$ se calcula un estimado de x a partir de

$$b_0 + b_1 x = 0.$$

Así, DE_{50} es dada por

$$x = -\left(\frac{b_0}{b_1}\right) = 0.276 \text{ gramos/100 cc.}$$

Concepto de razón de probabilidad

Otra forma de inferencia que se lleva a cabo de manera adecuada usando la regresión logística se deriva del uso de la razón de probabilidad, la cual está diseñada para determinar cómo se incrementa la **probabilidad de éxitos**, $\frac{p}{1-p}$, a medida que ocurren ciertos cambios en los valores del regresor. Por ejemplo, en el caso del ejemplo 12.13, quizá se deseara saber cómo aumentarían las probabilidades si la dosis se incrementara en, digamos, 0.2 gramos/100 cc.

Definición 12.1: En la regresión logística una **razón de probabilidad** es la razón de la probabilidad de éxito en la condición 2 con respecto a la de la condición 1 en los regresores, es decir,

$$\frac{[p/(1-p)]_2}{[p/(1-p)]_1}.$$

Esto permite que el analista tenga una idea de la utilidad de cambiar el regresor en cierto número de unidades. Ahora, como $\left(\frac{p}{1-p}\right) = e^{\beta_0 + \beta_1 x}$, para el ejemplo 12.13 la razón que refleja el incremento de las probabilidades de éxito cuando aumenta la dosis de nicotina en 0.2 gramos/100 cc es dada por

$$e^{0.2b_1} = e^{(0.2)(6.2954)} = 3.522.$$

La implicación de una razón de probabilidad de 3.522 es que la probabilidad de éxito aumenta en un factor de 3.522 cuando la dosis de nicotina aumenta en 0.2 gramos/100 cc.

Ejercicios

12.60 A partir de un conjunto de datos de respuestas a la dosis de estreptomina un investigador desea desarrollar una relación entre la proporción de linfoblastos muestreados que contienen aberraciones y la dosis del medicamento. Se aplicaron cinco niveles de dosis a los conejos que se emplearon para el experimento. Los datos son los siguientes (véase Myers, 1990, listado en la bibliografía):

Dosis (mg/kg)	Número de linfoblastos	Número de aberraciones
0	600	15
30	500	96
60	600	187
75	300	100
90	300	145

- a) Ajuste una regresión logística al conjunto de datos, y así estime β_0 y β_1 en el modelo

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

donde n es el número de linfoblastos, x es la dosis y p la probabilidad de una aberración.

- b) Muestre los resultados de pruebas χ^2 que revelen la significancia de los coeficientes de regresión β_0 y β_1 .
- c) Estime la DE_{50} e interprétela.

12.61 En un experimento para estudiar el efecto de la carga, x , en lb/pulgadas², sobre la probabilidad de falla de especímenes de cierto tipo de tela, varios especímenes se expusieron a cargas de entre 5 lb/pulg² a 90 lb/pulg². Se observaron los números de "fallas". Los datos son los siguientes:

Carga	Número de especímenes	Número de fallas
5	600	13
35	500	95
70	600	189
80	300	95
90	300	130

- a) Utilice regresión logística para ajustar el modelo

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

donde p es la probabilidad de falla y x es la carga.

- b) Emplee el concepto de razón de probabilidad para determinar el incremento de la probabilidad de falla que resulta de aumentar la carga en 20 lb/pulg².

Ejercicios de repaso

12.62 En el Departamento de Pesca y Vida Silvestre de Virginia Tech se realizó un experimento para estudiar el efecto de las características de la corriente sobre la biomasa de los peces. Las variables regresoras son las siguientes: profundidad promedio (de 50 células), x_1 ; área de la cubierta en la corriente, es decir, riberas socavadas, troncos, cantos rodados, etc., x_2 ; porcentaje de cubierta de material translúcido (promedio de 12), x_3 ; y un área ≥ 25 centímetros de profundidad, x_4 . La respuesta es y , la biomasa de los peces. Los datos son los siguientes:

Obs.	y	x_1	x_2	x_3	x_4
1	100	14.3	15.0	12.2	48.0
2	388	19.1	29.4	26.0	152.2
3	755	54.6	58.0	24.2	469.7
4	1288	28.8	42.6	26.1	485.9
5	230	16.1	15.9	31.6	87.6
6	0	10.0	56.4	23.3	6.9
7	551	28.5	95.1	13.0	192.9
8	345	13.8	60.6	7.5	105.8
9	0	10.7	35.2	40.3	0.0
10	348	25.9	52.0	40.3	116.6

- a) Ajuste una regresión lineal múltiple que incluya las cuatro variables regresoras.
- b) Utilice C_r , R^2 y s^2 para determinar el mejor subconjunto de variables. Calcule dichos estadísticos para todos los subconjuntos posibles.
- c) Compare lo adecuado de los modelos de los incisos a) y b) para efectos de predecir la biomasa de los peces.

12.63 Demuestre que, en un conjunto de datos de regresión lineal múltiple,

$$\sum_{i=1}^n h_{ii} = p.$$

12.64 Se efectuó un experimento sencillo para ajustar una ecuación de regresión múltiple que relaciona al producto, y , con la temperatura, x_1 , el tiempo de reacción, x_2 , y la concentración de uno de los reactivos, x_3 . Se eligieron dos niveles de cada variable y se hicieron las siguientes mediciones correspondientes a las variables independientes definidas:

y	x_1	x_2	x_3
7.6	-1	-1	-1
5.5	1	-1	-1
9.2	-1	1	-1
10.3	-1	-1	1
11.6	1	1	-1
11.1	1	-1	1
10.2	-1	1	1
14.0	1	1	1

- a) Utilice las variables codificadas y estime la ecuación de regresión lineal múltiple

$$\mu_{Y | x_1, x_2, x_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

- b) Separe la SCR, es decir, la suma de cuadrados de regresión, en tres componentes con un grado de libertad, atribuibles a x_1 , x_2 y x_3 , respectivamente. Construya una tabla de análisis de varianza donde se indiquen pruebas de significancia sobre cada variable. Comente los resultados.

12.65 En un experimento de ingeniería química relacionado con la transferencia de calor en una capa de fluido superficial se recabaron datos sobre las cuatro variables regresoras siguientes: la tasa de flujo del gas fluido en lb/hr (x_1), la tasa de flujo del gas flotante en lb/hr (x_2), la abertura de la boquilla de entrada del gas flotante en milímetros (x_3) y la temperatura de entrada del gas flotante en °F (x_4). Las respuestas medidas son la eficacia de la transferencia de calor (y_1) y la eficacia térmica (y_2). Los datos son los siguientes:

Obs.	y_1	y_2	x_1	x_2	x_3	x_4
1	41.852	38.75	69.69	170.83	45	219.74
2	155.329	51.87	113.46	230.06	25	181.22
3	99.628	53.79	113.54	228.19	65	179.06
4	49.409	53.84	118.75	117.73	65	281.30
5	72.958	49.17	119.72	117.69	25	282.20
6	107.702	47.61	168.38	173.46	45	216.14
7	97.239	64.19	169.85	169.85	45	223.88
8	105.856	52.73	169.85	170.86	45	222.80
9	99.348	51.00	170.89	173.92	80	218.84
10	111.907	47.37	171.31	173.34	25	218.12
11	100.008	43.18	171.43	171.43	45	219.20
12	175.380	71.23	171.59	263.49	45	168.62
13	117.800	49.30	171.63	171.63	45	217.58
14	217.409	50.87	171.93	170.91	10	219.92
15	41.725	54.44	173.92	71.73	45	296.60
16	151.139	47.93	221.44	217.39	65	189.14
17	220.630	42.91	222.74	221.73	25	186.08
18	131.666	66.60	228.90	114.40	25	285.80
19	80.537	64.94	231.19	113.52	65	286.34
20	152.966	43.18	236.84	167.77	45	221.72

Considere el modelo para predecir la respuesta del coeficiente de transferencia de calor

$$y_i = \beta_0 + \sum_{j=1}^4 \beta_j x_{ij} + \sum_{i=1}^4 \beta_{ij} x_{ij}^2 + \sum_{j \neq l} \sum_{i=1}^4 \beta_{jl} x_{ij} x_{li} + \epsilon_i, \quad i = 1, 2, \dots, 20.$$

- Calcule PRESS y $\sum_{i=1}^n |y_i - \hat{y}_i|$ para ajustar el modelo anterior con los mínimos cuadrados de regresión.
- Ajuste un modelo de segundo orden con x_4 eliminada por completo, es decir, elimine todos los términos que impliquen x_4 . Calcule los criterios de predicción para el modelo reducido. Comente qué tan adecuada es x_4 para predecir el coeficiente de transferencia de calor.
- Repita los incisos a) y b) para la eficacia térmica.

12.66 En la fisiología del deporte una medición objetiva de la condición física es el consumo de oxígeno en volumen por unidad de peso corporal por unidad de tiempo. Se estudiaron 31 individuos en un experimento con el fin de modelar el consumo de oxígeno en comparación con la edad en años, x_1 , el peso en kilogramos, x_2 , el tiempo para correr 1 1/2 millas, x_3 , las pulsaciones en reposo, x_4 , las pulsaciones al final de la carrera, x_5 , y las pulsaciones máximas durante la carrera, x_6 .

- Realice una regresión por etapas a un nivel de significancia de 0.25 en la entrada. Proporcione el modelo final.
- Estudie todos los subconjuntos posibles usando s^2 , C_p , R^2 y R_{ajust}^2 . Tome una decisión y determine el modelo final.

ID	y	x_1	x_2	x_3	x_4	x_5	x_6
1	44.609	44	89.47	11.37	62	178	182
2	45.313	40	75.07	10.07	62	185	185
3	54.297	44	85.84	8.65	45	156	168
4	59.571	42	68.15	8.17	40	166	172
5	49.874	38	89.02	9.22	55	178	180
6	44.811	47	77.45	11.63	58	176	176
7	45.681	40	75.98	11.95	70	176	180
8	49.091	43	81.19	10.85	64	162	170
9	39.442	44	81.42	13.08	63	174	176
10	60.055	38	81.87	8.63	48	170	186
11	50.541	44	73.03	10.13	45	168	168
12	37.388	45	87.66	14.03	56	186	192
13	44.754	45	66.45	11.12	51	176	176
14	47.273	47	79.15	10.60	47	162	164
15	51.855	54	83.12	10.33	50	166	170
16	49.156	49	81.42	8.95	44	180	185
17	40.836	51	69.63	10.95	57	168	172
18	46.672	51	77.91	10.00	48	162	168
19	46.774	48	91.63	10.25	48	162	164
20	50.388	49	73.37	10.08	76	168	168
21	39.407	57	73.37	12.63	58	174	176
22	46.080	54	79.38	11.17	62	156	165
23	45.441	52	76.32	9.63	48	164	166
24	54.625	50	70.87	8.92	48	146	155
25	45.118	51	67.25	11.08	48	172	172
26	39.203	54	91.63	12.88	44	168	172
27	45.790	51	73.71	10.47	59	186	188
28	50.545	57	59.08	9.93	49	148	155
29	48.673	49	76.32	9.40	56	186	188
30	47.920	48	61.24	11.50	52	170	176
31	47.467	52	82.78	10.50	53	170	172

12.67 Considere los datos del ejercicio de repaso 12.64. Suponga que le interesa agregar algunos términos de "interacción". En específico, considere el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_{12} x_{1i} x_{2i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i} + \beta_{123} x_{1i} x_{2i} x_{3i} + \epsilon_i$$

- ¿Aún se tiene ortogonalidad? Comente al respecto.
- Con el modelo ajustado del inciso a), ¿puede usted encontrar intervalos de predicción y de confianza sobre la respuesta media? Explique su respuesta.
- Considere un modelo en el que se eliminó $\beta_{123} x_1 x_2 x_3$. Para determinar si son necesarias las interacciones (como un todo), pruebe

$$H_0: \beta_{12} = \beta_{13} = \beta_{23} = 0.$$

Proporcione el valor P y saque conclusiones.

12.68 Para extraer petróleo crudo se utiliza una técnica de inyección de dióxido de carbono (CO_2). El flujo de CO_2 envuelve el petróleo y lo desplaza. En un experimento se introducen tubos de flujo en muestras de cavidades de petróleo que contienen una cantidad conocida del mismo. Se utilizan tres valores diferentes de

presión de flujo y tres valores diferentes de ángulos de inyección, las cavidades de petróleo se inyectan con CO_2 y se registra el porcentaje de petróleo desplazado. Considere el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + \epsilon_i$$

Ajuste el modelo anterior a los datos y sugiera cualquier modificación al modelo que considere necesaria.

Presión lb/pulg ² , x_1	Ángulo de inyección, x_2	Recuperación de petróleo, (%), y
1000	0	60.58
1000	15	72.72
1000	30	79.99
1500	0	66.83
1500	15	80.78
1500	30	89.78
2000	0	69.18
2000	15	80.31
2000	30	91.99

Fuente: Wang, G. C. "Microscopic Investigations of CO_2 Flooding Process", *Journal of Petroleum Technology*, vol. 4, núm. 8, agosto de 1982.

12.69 Un artículo del *Journal of Pharmaceutical Sciences* (vol. 80, 1991) presenta datos de la solubilidad de una fracción molar de un soluto a temperatura constante. También se midió la dispersión, x_1 , y los parámetros de solubilidad del enlace bipolar y de hidrógeno, x_2 y x_3 . En la tabla siguiente se presenta una parte de los datos. En el modelo, y es el logaritmo negativo de la fracción molar. Ajuste el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i$$

para $i = 1, 2, \dots, 20$.

Obs.	y	x_1	x_2	x_3
1	0.2220	7.3	0.0	0.0
2	0.3950	8.7	0.0	0.3
3	0.4220	8.8	0.7	1.0
4	0.4370	8.1	4.0	0.2
5	0.4280	9.0	0.5	1.0
6	0.4670	8.7	1.5	2.8
7	0.4440	9.3	2.1	1.0
8	0.3780	7.6	5.1	3.4
9	0.4940	10.0	0.0	0.3
10	0.4560	8.4	3.7	4.1
11	0.4520	9.3	3.6	2.0
12	0.1120	7.7	2.8	7.1
13	0.4320	9.8	4.2	2.0
14	0.1010	7.3	2.5	6.8
15	0.2320	8.5	2.0	6.6
16	0.3060	9.5	2.5	5.0
17	0.0923	7.4	2.8	7.8
18	0.1160	7.8	2.8	7.7
19	0.0764	7.7	3.0	8.0
20	0.4390	10.3	1.7	4.2

- Pruebe $H_0: \beta_1 = \beta_2 = \beta_3 = 0$.
- Grafique los residuales estudentizados en comparación con x_1 , x_2 y x_3 (tres gráficas). Haga comentarios al respecto.
- Considere dos modelos adicionales que compitan con el modelo anterior:

Modelo 2: Agregue x_1^2, x_2^2, x_3^2 .

Modelo 3: Agregue $x_1^2, x_2^2, x_3^2, x_1 x_2, x_1 x_3, x_2 x_3$.

Utilice PRESS y C_p con estos tres modelos para saber cuál de los tres es el mejor.

12.70 Se realizó un estudio para determinar si los cambios en el estilo de vida podrían sustituir la medicación para reducir la presión sanguínea de los individuos hipertensos. Los factores considerados fueron una dieta saludable con un programa de ejercicios, la dosis común de medicamentos para la hipertensión y ningún tratamiento. También se calculó el índice de masa corporal (IMC) previo al tratamiento, debido a que se sabe que éste afecta la presión sanguínea. La respuesta considerada en este estudio fue el cambio en la presión sanguínea. La variable "grupo" tenía los siguientes niveles.

1 = Dieta saludable y programa de ejercicios

2 = Medicación

3 = Sin tratamiento

- Ajuste un modelo adecuado utilizando los datos anteriores. ¿Parece que el ejercicio y la dieta se pueden utilizar en forma eficaz para disminuir la presión sanguínea? Explique su respuesta a partir de los resultados.
- ¿El ejercicio y la dieta son una alternativa eficaz a la medicación?

(Sugerencia: Para responder a estas preguntas quizás usted desee construir el modelo en más de una forma).

Cambio en la presión sanguínea	Grupo	IMC
-32	1	27.3
-21	1	22.1
-26	1	26.1
-16	1	27.8
-11	2	19.2
-19	2	26.1
-23	2	28.6
-5	2	23.0
-6	3	28.1
5	3	25.3
-11	3	26.7
14	3	22.3

12.71 Demuestre que al elegir el llamado mejor modelo del subconjunto de entre una serie de posibles modelos, elegir el modelo con la menor s^2 equivale a escoger el modelo con el R_{ajust}^2 más pequeño.

12.72 Estudio de caso: Considere el conjunto de datos para el ejercicio 12.12 de la página 452 (datos de un hospital) que se repite a continuación.

Sitio	x_1	x_2	x_3	x_4	x_5	y
1	15.57	2463	472.92	18.0	4.45	566.52
2	44.02	2048	1339.75	9.5	6.92	696.82
3	20.42	3940	620.25	12.8	4.28	1033.15
4	18.74	6505	568.33	36.7	3.90	1003.62
5	49.20	5723	1497.60	35.7	5.50	1611.37
6	44.92	11,520	1365.83	24.0	4.60	1613.27
7	55.48	5779	1687.00	43.3	5.62	1854.17
8	59.28	5969	1639.92	46.7	5.15	2160.55
9	94.39	8461	2872.33	78.7	6.18	2305.58
10	128.02	20,106	3655.08	180.5	6.15	3503.93
11	96.00	13,313	2912.00	60.9	5.88	3571.59
12	131.42	10,771	3921.00	103.7	4.88	3741.40
13	127.21	15,543	3865.67	126.8	5.50	4026.52
14	252.90	36,194	7684.10	157.7	7.00	10,343.81
15	409.20	34,703	12,446.33	169.4	10.75	11,732.17
16	463.70	39,204	14,098.40	331.4	7.05	15,414.94
17	510.22	86,533	15,524.00	371.6	6.35	18,854.45

- a) Los listados de resultados de la función PROC REG del SAS que se presentan en las figuras 12.9 y 12.10 proporcionan una cantidad considerable de información. El propósito es detectar los valores extremos y , a final de cuentas, determinar cuáles términos del modelo deben utilizarse en la versión final.
- b) A menudo ocurre que el papel que desempeña una sola variable regresora no es evidente cuando se estudia en presencia de otras variables; esto se debe a la multicolinealidad. Con esto presente haga comentarios sobre la importancia de x_2 y x_3 en el modelo completo en comparación con su importancia en un modelo en el cual éstas son las únicas variables.
- c) Comente acerca de qué otros análisis se tendrían que hacer.
- d) Elabore análisis apropiados y escriba sus conclusiones respecto al modelo final.

Dependent Variable: y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	490177488	98035498	237.79	<.0001
Error	11	4535052	412277		
Corrected Total	16	494712540			
Root MSE		642.08838	R-Square	0.9908	
Dependent Mean		4978.48000	Adj R-Sq	0.9867	
Coeff Var		12.89728			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	1962.94816	1071.36170	1.83	0.0941
x_1	Average Daily Patient Load	1	-15.85167	97.65299	-0.16	0.8740
x_2	Monthly X-Ray Exposure	1	0.05593	0.02126	2.63	0.0234
x_3	Monthly Occupied Bed Days	1	1.58962	3.09208	0.51	0.6174
x_4	Eligible Population in the Area/100	1	-4.21867	7.17656	-0.59	0.5685
x_5	Average Length of Patients Stay in Days	1	-394.31412	209.63954	-1.88	0.0867

Figura 12.9: Salida de resultados del SAS para el ejercicio de repaso 12.72; parte I.

Obs	Dependent	Predicted	Std Error		95% CL Mean		95% CL Predict	
	Variable	Value	Mean	Predict	Mean	Mean	Predict	Predict
1	566.5200	775.0251	241.2323	244.0765	1306	-734.6494	2285	
2	696.8200	740.6702	331.1402	11.8355	1470	-849.4275	2331	
3	1033	1104	278.5116	490.9234	1717	-436.5244	2644	
4	1604	1240	268.1298	650.3459	1831	-291.0028	2772	
5	1611	1564	211.2372	1099	2029	76.6816	3052	
6	1613	2151	279.9293	1535	2767	609.5796	3693	
7	1854	1690	218.9976	1208	2172	196.5345	3183	
8	2161	1736	468.9903	703.9948	2768	-13.8306	3486	
9	2306	2737	290.4749	2098	3376	1186	4288	
10	3504	3682	585.2517	2394	4970	1770	5594	
11	3572	3239	189.0989	2823	3655	1766	4713	
12	3741	4353	328.8507	3630	5077	2766	5941	
13	4027	4257	314.0481	3566	4948	2684	5830	
14	10344	8768	252.2617	8213	9323	7249	10286	
15	11732	12237	573.9168	10974	13500	10342	14133	
16	15415	15038	585.7046	13749	16328	13126	16951	
17	18854	19321	599.9780	18000	20641	17387	21255	

Obs	Residual	Std Error	Student					
		Residual	Residual	-2	-1	0	1	2
1	-208.5051	595.0	-0.350					
2	-43.8502	550.1	-0.0797					
3	-70.7734	578.5	-0.122					
4	363.1244	583.4	0.622				*	
5	46.9483	606.3	0.0774					
6	-538.0017	577.9	-0.931		*			
7	164.4696	603.6	0.272					
8	424.3145	438.5	0.968				*	
9	-431.4090	572.6	-0.753		*			
10	-177.9234	264.1	-0.674		*			
11	332.6011	613.6	0.542				*	
12	-611.9330	551.5	-1.110		**			
13	-230.5684	560.0	-0.412					
14	1576	590.5	2.669				*****	
15	-504.8574	287.9	-1.753		***			
16	376.5491	263.1	1.431				**	
17	-466.2470	228.7	-2.039		****			

Figura 12.10: Salida de resultados del SAS para el ejercicio de repaso 12.72; parte II.

12.13 Posibles riesgos y errores conceptuales; relación con el material de otros capítulos

En este capítulo estudiamos varios procedimientos para usarlos en un “intento” por encontrar el mejor modelo. Sin embargo, una de las confusiones más importantes en el trabajo de los científicos e ingenieros novatos es que existe un **modelo lineal verdadero**, y que es posible encontrarlo. En la mayoría de fenómenos de la ciencia las relaciones entre las variables científicas son de naturaleza no lineal y se desconoce el modelo verdadero. Los modelos estadísticos lineales son **aproximaciones empíricas**.

En ocasiones, la decisión sobre cuál modelo adoptar depende de la información que se necesita obtener de éste. ¿Se usará para realizar predicciones? ¿Para explicar el papel que desempeña cada regresor? Esta “decisión” podría ser difícil ante la presencia de colinealidad. Es un hecho que para muchos problemas de regresión hay modelos múltiples con un desempeño muy similar. Para mayores detalles véase la referencia de Myers (1990).

Uno de los abusos más nocivos del material de este capítulo consiste en dar demasiada importancia a R^2 en la selección del llamado mejor modelo. Es importante recordar que para cualquier conjunto de datos se puede obtener una R^2 tan grande como se desee, dentro de la restricción de que $0 \leq R^2 \leq 1$. **Prestar demasiada atención a R^2 con frecuencia conduce a un sobreajuste.**

En este capítulo se dio mucha importancia a la detección de los valores extremos. Un clásico y grave abuso de la estadística radica en la decisión relacionada con la detección de los valores extremos. Esperamos que quede claro que el analista no debería por ningún motivo detectar los valores extremos, eliminarlos del conjunto de datos, ajustar un modelo nuevo, informar sobre los valores extremos, y así sucesivamente. Se trata de un procedimiento tentador y desastroso para llegar a un modelo que se ajuste bien a los datos, el cual conlleva a un ejemplo de **cómo mentir con estadísticos**. Si se detecta un valor extremo, lo correcto es revisar la historia de los datos en busca de posibles errores de captura o de procedimiento antes de eliminarlos del conjunto de datos. Se debe recordar que, por definición, un valor extremo es aquel para el cual el modelo no se ajusta bien. El problema podría no estar en los datos sino en la selección del modelo. Cambiar el modelo quizás haría que el punto no se detecte como un valor extremo.

Existen muchos tipos de respuestas que ocurren de forma natural en la práctica, pero que no se pueden utilizar en un análisis de mínimos cuadrados estándar porque sus supuestos de mínimos cuadrados clásicos no se cumplen. Los supuestos que suelen fallar son los de los errores normales y de la varianza homogénea. Por ejemplo, si la respuesta es una proporción, digamos la proporción de artículos defectuosos, la distribución de las respuestas se relaciona con la distribución binomial. Una segunda respuesta que ocurre con frecuencia en la práctica es la del conteo de Poisson. Evidentemente, la distribución no es normal, y la varianza de la respuesta, que es igual a la media de Poisson, varía de una observación a otra. Para conocer más detalles sobre estas condiciones poco ideales véase Myers y colaboradores (2008), citado en la bibliografía.

Capítulo 13

Experimentos con un solo factor: generales

13.1 Técnica del análisis de varianza

En el material sobre estimación y prueba de hipótesis que se cubrió en los capítulos 9 y 10 en cada caso nos limitamos a considerar sólo dos parámetros de la población. Ése fue el caso, por ejemplo, en la prueba de la igualdad de dos medias de la población, en la cual se usaron muestras independientes de poblaciones normales con varianza común pero desconocida, y en donde se necesitaba obtener un estimado agrupado de σ^2 .

El material que se refiere a las inferencias de dos muestras representa un caso especial de lo que se denomina *problema de un solo factor*. Por ejemplo, en el ejercicio 10.35 de la página 357 se midió el tiempo de supervivencia para dos muestras de ratones, en donde una muestra recibió un tratamiento de suero contra la leucemia y la otra no lo recibió. En este caso decimos que hay *un factor*, es decir, un *tratamiento*, y el factor se halla en *dos niveles*. Si en el proceso de muestreo se utilizaran varios tratamientos en competencia, se necesitarían más muestras de ratones. En ese caso el problema implicaría un factor con más de dos niveles, por lo tanto, con más de dos muestras.

En el problema de $k > 2$ muestras se supone que hay k muestras provenientes de k poblaciones. Un procedimiento muy común que se utiliza cuando se prueban medias de la población se denomina **análisis de varianza**, o ANOVA.

Si el lector ha estudiado el material acerca de la teoría de la regresión, el análisis de varianza no será, por supuesto, una técnica nueva para él. Utilizamos el método del análisis de varianza para partir la suma total de cuadrados en dos partes, una parte debida a la regresión y otra debida al error.

Suponga que en un experimento industrial a un ingeniero le interesa la forma en que la absorción media de humedad del concreto varía para 5 agregados de concreto diferentes. Las muestras se exponen a la humedad durante 48 horas y se decide que para cada agregado deben probarse 6 muestras, lo que hace que se requiera probar un total de 30 muestras. En la tabla 13.1 se presentan los datos registrados.

El modelo que se considera para esta situación es el siguiente. Se tomaron 6 observaciones de cada una de las 5 poblaciones, con medias $\mu_1, \mu_2, \dots, \mu_5$, respectivamente. Deseamos probar

$$H_0: \mu_1 = \mu_2 = \dots = \mu_5,$$

$$H_1: \text{Al menos dos de las medias no son iguales.}$$

Tabla 13.1: Absorción de humedad en agregados para concreto

Agregado:	1	2	3	4	5	
	551	595	639	417	563	
	457	580	615	449	631	
	450	508	511	517	522	
	731	583	573	438	613	
	499	633	648	415	656	
	632	517	677	555	679	
Total	3320	3416	3663	2791	3664	16,854
Media	553.33	569.33	610.50	465.17	610.67	561.80

Además, estamos interesados en realizar comparaciones individuales entre estas 5 medias de la población.

Dos fuentes de variabilidad en los datos

En el procedimiento del análisis de varianza se supone que cualquier variación que exista entre los promedios de los agregados se atribuye a 1) la variación en la absorción entre observaciones *dentro* de los tipos de agregados, y 2) la variación *entre* los tipos de agregados, es decir, a las diferencias en la composición química de los agregados. Por supuesto, la **variación dentro de los agregados** se debe a varias causas. Quizá las condiciones de temperatura y humedad no se mantuvieron constantes durante el experimento. Es posible que haya habido cierta cantidad de heterogeneidad en los lotes de materias primas que se usaron. En todo caso debe considerarse la variación dentro de la muestra como una **variación aleatoria o al azar**. Parte del objetivo del análisis de varianza consiste en determinar si las diferencias entre las 5 medias muestrales son lo que se esperaría debido sólo a la variación aleatoria o si, más bien, se trata de una variación más allá de los simples efectos del azar, como las diferencias en la composición química de los agregados.

En esta etapa surgen muchas preguntas acerca del problema anterior. Por ejemplo, ¿cuántas muestras deben probarse para cada agregado? Ésta es una pregunta que desafía continuamente al analista. Además, ¿qué pasaría si la variación dentro de la muestra fuera tan grande que al procedimiento estadístico le resultara difícil detectar las diferencias sistemáticas? ¿Es posible controlar de manera sistemática fuentes externas de variación y así eliminarlas de la parte que llamamos variación aleatoria? En las secciones siguientes intentaremos responder éstas y otras preguntas.

13.2 La estrategia del diseño de experimentos

En los capítulos 9 y 10 se estudiaron los conceptos de la estimación y la prueba de hipótesis para el caso de dos muestras, bajo la importante perspectiva de la manera en que se realiza el experimento. Esto forma parte de la categoría amplia de los diseños experimentales. Por ejemplo, para la **prueba *t* agrupada** que se estudió en el capítulo 10, se supone que los niveles de los factores (los tratamientos en el ejemplo de los ratones) se asignan al azar a las unidades experimentales (los ratones). En los capítulos 9 y 10 analizamos el

concepto de unidades experimentales y lo ilustramos por medio de varios ejemplos. En pocas palabras, las unidades experimentales son las unidades (ratones, pacientes, especímenes de concreto, tiempo) que **proporcionan la heterogeneidad que conduce al error experimental** en una investigación científica. La asignación aleatoria elimina el sesgo que podría originarse con una asignación sistemática. El objetivo consiste en distribuir en forma uniforme entre los niveles de los factores los riesgos que introduce la heterogeneidad de las unidades experimentales. Una asignación al azar simula mejor las condiciones que se asumen en el modelo. En la sección 13.7 analizamos la **formación de bloques** en los experimentos. En los capítulos 9 y 10 se presentó el concepto, cuando se efectuaron comparaciones entre las medias usando el **emparejamiento**, es decir, la división de las unidades experimentales en pares homogéneos denominados **bloques**. Entonces, los niveles de los factores o tratamientos se asignan al azar dentro de los bloques. El propósito de la formación de bloques es reducir el error experimental efectivo. En este capítulo se extiende de manera natural el emparejamiento a bloques de tamaño mayor, con el análisis de varianza como la herramienta analítica principal.

13.3 Análisis de varianza de un factor: diseño completamente aleatorizado (ANOVA de un factor)

De k poblaciones se seleccionan muestras aleatorias de tamaño n . Las k poblaciones diferentes se clasifican con base en un criterio único, como tratamientos o grupos distintos. En la actualidad el término **tratamiento** se utiliza por lo general para designar las diversas clasificaciones, ya sean diferentes agregados, analistas, fertilizadores o regiones del país.

Suposiciones e hipótesis del ANOVA de un solo factor

Se supone que las k poblaciones son independientes y que están distribuidas en forma normal con medias $\mu_1, \mu_2, \dots, \mu_k$, y varianza común σ^2 . Como se indicó en la sección 13.2, estas suposiciones son más aceptables mediante la aleatoriedad. Se desean obtener métodos adecuados para probar las hipótesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k,$$

H_1 : Al menos dos de las medias no son iguales.

Sea que y_{ij} denote la j -ésima observación del i -ésimo tratamiento, y el acomodo de los datos es el que se observa en la tabla 13.2. Aquí, Y_i es el total de todas las observaciones de la muestra, del i -ésimo tratamiento, \bar{y}_i es la media de todas las observaciones en la muestra del i -ésimo tratamiento, $Y_{..}$ es el total de todas las nk observaciones, y $\bar{y}_{..}$ es la media de todas las nk observaciones.

Modelo de ANOVA para un solo factor

Cada observación puede escribirse en la forma

$$Y_{ij} = \mu_i + \epsilon_{ij},$$

donde ϵ_{ij} mide la desviación que tiene la observación j -ésima de la i -ésima muestra, con respecto de la media del tratamiento correspondiente. El término ϵ_{ij} representa el error aleatorio y desempeña el mismo papel que los términos del error en los modelos de

Tabla 13.2: k muestras aleatorias

Tratamiento:	1	2	...	i	...	k	
	y_{11}	y_{21}	...	y_{i1}	...	y_{k1}	
	y_{12}	y_{22}	...	y_{i2}	...	y_{k2}	
	\vdots	\vdots		\vdots		\vdots	
	y_{1n}	y_{2n}	...	y_{in}	...	y_{kn}	
Total	Y_1	Y_2	...	Y_i	...	Y_k	$Y_{..}$
Media	\bar{y}_1	\bar{y}_2	...	\bar{y}_i	...	\bar{y}_k	$\bar{y}_{..}$

regresión. Una forma alternativa y preferible de esta ecuación se obtiene sustituyendo $\mu_i = \mu + \alpha_i$, sujeta a la restricción $\sum_{i=1}^k \alpha_i = 0$. Por lo tanto, se escribe

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

donde μ tan sólo es la **media general** de todas las μ_i , es decir,

$$\mu = \frac{1}{k} \sum_{i=1}^k \mu_i,$$

y α_i se denomina el **efecto** del i -ésimo tratamiento.

La hipótesis nula de que k medias de la población son iguales, en comparación con la alternativa de que al menos dos de las medias son distintas, ahora se puede reemplazar por las hipótesis equivalentes.

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0,$$

$$H_1: \text{Al menos una de las } \alpha_i \text{ no es igual a cero.}$$

Resolución de la variabilidad total en componentes

Nuestra prueba se basará en una comparación de dos estimados independientes de la varianza poblacional común σ^2 . Dichos estimadores se obtendrán haciendo la partición de la variabilidad total de nuestros datos, denotados mediante la sumatoria doble

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2,$$

en dos componentes.

Teorema 13.1: Identidad de la suma de cuadrados

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_i - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$$

En lo que sigue, será conveniente identificar los términos de la identidad de la suma de cuadrados con la siguiente notación:

Tres medidas importantes de variabilidad

$$STC = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \text{suma total de cuadrados,}$$

$$SCT = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 = \text{suma de los cuadrados del tratamiento,}$$

$$SCE = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 = \text{suma de los cuadrados del error.}$$

Entonces, la identidad de la suma de los cuadrados se puede representar simbólicamente con la ecuación

$$STC = SCT + SCE$$

La identidad anterior expresa cómo las variaciones entre los tratamientos y dentro de los tratamientos contribuyen a la suma total de cuadrados. Sin embargo, se puede obtener mucha información si se investiga el **valor esperado tanto de SCT como de SCE**. Eventualmente calcularemos estimados de la varianza que determinan la razón que

Teorema 13.2:

$$E(SCT) = (k-1)\sigma^2 + n \sum_{i=1}^k \alpha_i^2$$

La prueba del teorema se deja como ejercicio para el lector (véase el ejercicio 13.53 de la página 556).

Si H_0 es verdadera, un estimado de σ^2 basado en $k-1$ grados de libertad es dado por la expresión:

Media cuadrática del tratamiento

$$s_1^2 = \frac{SCT}{k-1}$$

Si H_0 es verdadera y por ello cada α_i en el teorema 13.2 es igual a cero, se observa que

$$E\left(\frac{SCT}{k-1}\right) = \sigma^2,$$

y s_1^2 es un estimado no sesgado de σ^2 . Sin embargo, si H_1 es verdadera, se tiene que

$$E\left(\frac{SCT}{k-1}\right) = \sigma^2 + \frac{n}{k-1} \sum_{i=1}^k \alpha_i^2,$$

y s_1^2 estima a σ^2 más un término adicional, que mide la variación debida a los efectos sistemáticos.

Otro estimado independiente de σ^2 , basado en $k(n-1)$ grados de libertad, es la fórmula familiar:

Cuadrado medio del error

$$s^2 = \frac{SCE}{k(n-1)}$$

Resulta aleccionador puntualizar la importancia de los valores esperados de los cuadrados medios a los que recién nos referimos. En la sección siguiente se estudia el empleo de una razón F con el cuadrado medio del tratamiento en el numerador. Se observa que cuando H_1 es verdadera, la presencia de la condición $E(s_1^2) > E(s^2)$ sugiere que la razón F se utiliza en el contexto de una **prueba unilateral de cola superior**. Es decir, cuando H_1 es verdadera se esperaría que el numerador s_1^2 fuera mayor que el denominador.

Uso de la prueba F en el ANOVA

El estimado s^2 es no sesgado, independientemente de la veracidad o falsedad de la hipótesis nula (véase el ejercicio de repaso 13.52 de la página 556). Es importante señalar que la identidad de la suma de cuadrados ha hecho la partición no sólo de la variabilidad total de los datos, sino también del número total de grados de libertad. Es decir,

$$nk - 1 = k - 1 + k(n - 1).$$

Razón F para probar la igualdad de las medias

Cuando H_0 es verdadera, la razón $f = s_1^2/s^2$ es un valor de la variable aleatoria F , que tiene la distribución F con $k - 1$ y $k(n - 1)$ grados de libertad (véase el teorema 8.8). Como s_1^2 sobrestima a σ^2 cuando H_0 es falsa, se tiene una prueba de una cola con la región crítica localizada por completo en la cola derecha de la distribución.

A un nivel de significancia de α se rechaza la hipótesis nula H_0 cuando

$$f > f_{\alpha}[k - 1, k(n - 1)].$$

Otro método, el del valor P , sugiere que la evidencia a favor o en contra de H_0 es

$$P = P\{f[k - 1, k(n - 1)] > f\}.$$

Los cálculos para un problema de análisis de varianza por lo general se resumen en forma tabular, como se observa en la tabla 13.3.

Tabla 13.3: Análisis de varianza para el ANOVA de un solo factor

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada
Tratamientos	SCT	$k - 1$	$s_1^2 = \frac{SCT}{k - 1}$	$\frac{s_1^2}{s^2}$
Error	SCE	$k(n - 1)$	$s^2 = \frac{SCE}{k(n - 1)}$	
Total	STC	$kn - 1$		

Ejemplo 13.1: Pruebe la hipótesis de que $\mu_1 = \mu_2 = \dots = \mu_k$ a un nivel de significancia de 0.05 para los datos de la tabla 13.1 sobre la absorción de humedad por varios tipos de agregados para cemento.

Solución: Las hipótesis son

$$H_0: \mu_1 = \mu_2 = \dots = \mu_5,$$

H_1 : Al menos dos de las medias no son iguales.

$$\alpha = 0.05$$

Región crítica: $f > 2.76$ con $v_1 = 4$ y $v_2 = 25$ grados de libertad. Los cálculos de la suma de cuadrados proporcionan

$$STC = 209,377, \quad SCT = 85,356,$$

$$SCE = 209,377 - 85,356 = 124,021.$$

En la figura 13.1 se presentan estos resultados y el resto de los cálculos del procedimiento ANOVA del programa SAS.

The GLM Procedure					
Dependent Variable: moisture					
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	4	85356.4667	21339.1167	4.30	0.0088
Error	25	124020.3333	4960.8133		
Corrected Total	29	209376.8000			
R-Square	Coeff Var	Root MSE	moisture Mean		
0.407669	12.53703	70.43304	561.8000		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
aggregate	4	85356.46667	21339.11667	4.30	0.0088

Figura 13.1: Salida de resultados del programa SAS para el procedimiento de análisis de varianza.

Decisión: Rechazar H_0 y concluir que los agregados no tienen la misma media de absorción. El valor P para $f = 4.30$ es 0.0088, que es menor que 0.05.

Además del ANOVA, se construyeron gráficas de caja para cada agregado, las cuales se presentan en la figura 13.2. Al observar las gráficas vemos que es evidente que no todos los agregados tienen la misma absorción. De hecho, parece que el agregado 4 destaca del resto. En el ejercicio 13.21 de la página 531 se incluye un análisis más formal que revela este resultado.

Durante el trabajo experimental es frecuente que se pierdan algunas de las observaciones deseadas. Los animales experimentales mueren, el material experimental se daña o los seres humanos abandonan el estudio. El análisis anterior para el mismo tamaño de la muestra aún es válido si modificamos ligeramente las fórmulas de la suma de cuadrados. Ahora suponemos que las k muestras aleatorias son de tamaño n_1, n_2, \dots, n_k , respectivamente.

Suma de cuadrados:
tamaños desiguales
de las muestras

$$STC = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2, \quad SCT = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y}_{..})^2, \quad SCE = STC - SCT$$

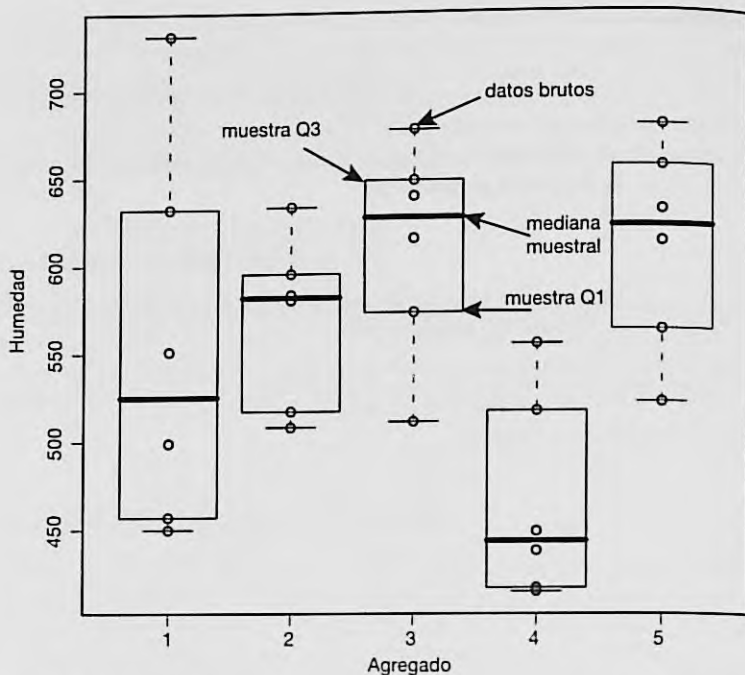


Figura 13.2: Gráficas de caja para la absorción de la humedad en agregados de concreto.

Después se hace la partición de los grados de libertad, como antes: $N - 1$ para STC , $k - 1$ para SCT y $N - 1 - (k - 1) = N - k$ para SCE , donde $N = \sum_{i=1}^k n_i$.

Ejemplo 13.2: Parte de un estudio realizado en Virginia Tech se diseñó para medir los niveles de actividad de la fosfatasa alcalina sérica (en unidades de Bessey-Lowry) en niños con trastornos convulsivos que recibían terapia de anticonvulsivantes bajo el cuidado de un médico privado. Se reclutaron 45 sujetos para el estudio y se clasificaron en cuatro grupos de medicamentos:

- G-1: Control (no recibieron anticonvulsivantes ni tenían historia de trastornos convulsivos)
- G-2: Fenobarbital
- G-3: Carbamazepina
- G-4: Otros anticonvulsivantes

De las muestras de sangre tomadas a cada sujeto se determinó el nivel de actividad de la fosfatasa alcalina sérica y se registró tal como se observa en la tabla 13.4. Pruebe la hipótesis de que, a un nivel de significancia de 0.05, el nivel promedio de actividad de la fosfatasa alcalina sérica es el mismo para los cuatro grupos de medicamentos.

Tabla 13.4: Nivel de actividad de la fosfatasa alcalina sérica

G-1		G-2	G-3	G-4
49.20	97.50	97.07	62.10	110.60
44.54	105.00	73.40	94.95	57.10
45.80	58.05	68.50	142.50	117.60
95.84	86.60	91.85	53.00	77.71
30.10	58.35	106.60	175.00	150.00
36.50	72.80	0.57	79.50	82.90
82.30	116.70	0.79	29.50	111.50
87.85	45.15	0.77	78.40	
105.00	70.35	0.81	127.50	
95.22	77.40			

Solución: A un nivel de significancia de 0.05, las hipótesis son

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4,$$

H_1 : Al menos dos de las medias no son iguales.

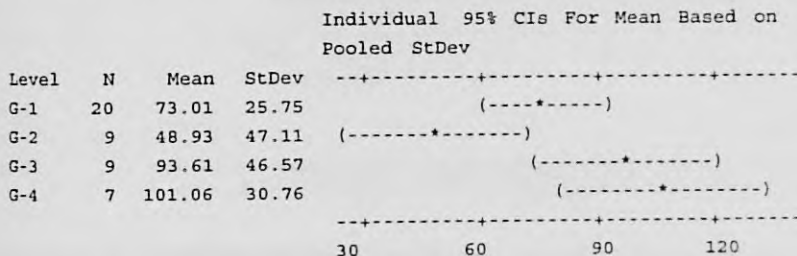
Región crítica: $f > 2.836$, al interpolar los valores de la tabla A.6.

Cálculos: $Y_1 = 1460.25$, $Y_2 = 440.36$, $Y_3 = 842.45$, $Y_4 = 707.41$ y $Y = 3450.47$. El análisis de varianza se incluye en la salida de resultados de MINITAB que se presenta en la figura 13.3.

One-way ANOVA: G-1, G-2, G-3, G-4

Source	DF	SS	MS	F	P
Factor	3	13939	4646	3.57	0.022
Error	41	53376	1302		
Total	44	67315			

S = 36.08 R-Sq = 20.71% R-Sq(adj) = 14.90%



Pooled StDev = 36.08

Figura 13.3: Análisis de MINITAB para los datos de la tabla 13.4.

Decisión: Rechazar H_0 y concluir que los niveles de actividad promedio de la fosfatasa alcalina sérica para los cuatro grupos de medicamentos no son los mismos. El valor calculado de P es 0.022. \blacksquare

Para concluir nuestro estudio del análisis de varianza para la clasificación de un solo factor mencionaremos las ventajas de elegir muestras del mismo tamaño en vez de otras de tamaños distintos. La primera ventaja es que la razón f no es sensible a pequeñas desviaciones de la suposición de varianzas iguales para las k poblaciones cuando las muestras son del mismo tamaño. La segunda consiste en que muestras del mismo tamaño minimizan la probabilidad de cometer un error tipo II.

13.4 Pruebas de la igualdad de varias varianzas

Aunque la razón f que se obtiene con el procedimiento del análisis de varianza no es sensible a las desviaciones de la suposición de varianzas iguales para las k poblaciones normales cuando las muestras son de igual tamaño, debe tenerse precaución y efectuar una prueba preliminar sobre la homogeneidad de las varianzas. En el caso de muestras de tamaños distintos, salta a la vista que es aconsejable realizar una prueba como ésta, si existe duda razonable acerca de la homogeneidad de las varianzas de la población. Por lo tanto, suponga que se desea probar la hipótesis nula

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

en comparación con la alternativa

$$H_1: \text{No todas las varianzas son iguales.}$$

La prueba que usaremos, denominada **prueba de Bartlett**, se basa en un estadístico cuya distribución muestral proporciona valores críticos exactos cuando los tamaños de las muestras son iguales. Dichos valores críticos para tamaños de las muestras iguales también se pueden utilizar para obtener aproximaciones muy exactas de los valores críticos para tamaños muestrales distintos.

En primer lugar calculamos las k varianzas muestrales $s_1^2, s_2^2, \dots, s_k^2$ a partir de muestras de tamaño n_1, n_2, \dots, n_k , con $\sum_{i=1}^k n_i = N$. En segundo lugar combinamos las varianzas muestrales para obtener la estimación agrupada

$$s_p^2 = \frac{1}{N - k} \sum_{i=1}^k (n_i - 1) s_i^2.$$

Ahora,

$$b = \frac{[(s_1^2)^{n_1-1} (s_2^2)^{n_2-1} \dots (s_k^2)^{n_k-1}]^{1/(N-k)}}{s_p^2}$$

es un valor de una variable aleatoria B que tiene la **distribución de Bartlett**. Para el caso especial en que $n_1 = n_2 = \dots = n_k = n$, se rechaza H_0 a un nivel de significancia α si

$$b < b_k(\alpha; n).$$

donde $b_k(\alpha; n)$ es el valor crítico que deja una área de tamaño α en el extremo izquierdo de la distribución de Bartlett. En la tabla A.10 se incluyen los valores críticos, $b_k(\alpha; n)$, para $\alpha = 0.01$ y 0.05 ; $k = 2, 3, \dots, 10$; y valores seleccionados de n , desde 3 hasta 100.

Cuando los tamaños de las muestras son distintos, se rechaza la hipótesis nula al nivel de significancia α si

$$b < b_k(\alpha; n_1, n_2, \dots, n_k),$$

donde

$$b_k(\alpha; n_1, n_2, \dots, n_k) \approx \frac{n_1 b_k(\alpha; n_1) + n_2 b_k(\alpha; n_2) + \dots + n_k b_k(\alpha; n_k)}{N}.$$

Igual que antes, todas las $b_k(\alpha; n_i)$ para los tamaños muestrales n_1, n_2, \dots, n_k se obtienen de la tabla A.10.

Ejemplo 13.3: Utilice la prueba de Bartlett a un nivel de significancia de 0.01 para probar la hipótesis de que las varianzas de la población de los cuatro grupos de medicamentos del ejemplo 13.2 son iguales.

Solución: Tenemos la hipótesis

$$\begin{aligned} H_0: \sigma_1^2 &= \sigma_2^2 = \sigma_3^2 = \sigma_4^2, \\ H_1: &\text{Las varianzas no son iguales,} \\ &\text{con } \alpha = 0.01. \end{aligned}$$

Región crítica: Si nos remitimos al ejemplo 13.2, tenemos que $n_1 = 20$, $n_2 = 9$, $n_3 = 9$, $n_4 = 7$, $N = 45$ y $k = 4$. Por lo tanto, se rechaza cuando

$$\begin{aligned} b &< b_4(0.01; 20, 9, 9, 7) \\ &\approx \frac{(20)(0.8586) + (9)(0.6892) + (9)(0.6892) + (7)(0.6045)}{45} \\ &= 0.7513. \end{aligned}$$

Cálculos: El primero se obtiene

$$s_1^2 = 662.862, \quad s_2^2 = 2219.781, \quad s_3^2 = 2168.434, \quad s_4^2 = 946.032,$$

y después

$$\begin{aligned} s_p^2 &= \frac{(19)(662.862) + (8)(2219.781) + (8)(2168.434) + (6)(946.032)}{41} \\ &= 1301.861. \end{aligned}$$

Ahora,

$$b = \frac{[(662.862)^{19} (2219.781)^8 (2168.434)^8 (946.032)^6]^{1/41}}{1301.861} = 0.8557.$$

Decisión: no rechazar la hipótesis y concluir que las varianzas de la población de los cuatro grupos de medicamentos no son significativamente distintas. ▀

Aunque la prueba de Bartlett se utiliza con mayor frecuencia para probar la homogeneidad de varianzas, se dispone de otros métodos. Un método creado por Cochran proporciona un procedimiento de cálculo sencillo, aunque está limitado a situaciones en

que los tamaños muestrales son iguales. La **prueba de Cochran** es especialmente útil para detectar si alguna de las varianzas es mucho mayor que las demás. El estadístico que se emplea es:

$$G = \frac{S_i^2 \text{ más grande}}{\sum_{i=1}^k S_i^2},$$

y se rechaza la hipótesis de igualdad de varianzas si $g > g_\alpha$, donde el valor de g_α se obtiene de la tabla A.11.

Para ilustrar la prueba de Cochran nos remitiremos otra vez a los datos de la tabla 13.1 sobre la absorción de humedad de los agregados para concreto. ¿Se justificó la suposición de varianzas iguales al realizar el análisis de varianza en el ejemplo 13.1? Se encontró que

$$s_1^2 = 12,134, s_2^2 = 2303, s_3^2 = 3594, s_4^2 = 3319, s_5^2 = 3455.$$

Por lo tanto,

$$g = \frac{12,134}{24,805} = 0.4892,$$

que no excede el valor de la tabla $g_{0.05} = 0.5065$. En consecuencia, se concluye que es razonable la suposición de que las varianzas son iguales.

Ejercicios

13.1 Se están considerando seis máquinas diferentes para la fabricación de sellos de goma y se están comparando con respecto a la resistencia a la tensión del producto. Se utiliza una muestra aleatoria de cuatro sellos hechos con cada máquina para determinar si la resistencia media a la tensión varía de una máquina a otra. A continuación se presentan las medidas de la resistencia a la tensión en kilogramos por centímetro cuadrado $\times 10^{-1}$:

Máquina					
1	2	3	4	5	6
17.5	16.4	20.3	14.6	17.5	18.3
16.9	19.2	15.7	16.7	19.2	16.2
15.8	17.7	17.8	20.8	16.5	17.5
18.6	15.4	18.9	18.9	20.5	20.1

Realice el análisis de varianza a un nivel de significancia de 0.05 e indique si la resistencia promedio a la tensión de las seis máquinas difiere o no de manera significativa.

13.2 Los datos que se presentan en la siguiente tabla representan el número de horas de alivio proporcionadas por cinco marcas diferentes de tabletas para el dolor de cabeza administradas a 25 sujetos que tenían fiebre de 38°C o más. Realice el análisis de varianza y, a un nivel de significancia de 0.05, pruebe la hipótesis de que las cinco marcas proporcionan el mismo número medio de horas de alivio. Analice los resultados.

Tabletas				
A	B	C	D	E
5.2	9.1	3.2	2.4	7.1
4.7	7.1	5.8	3.4	6.6
8.1	8.2	2.2	4.1	9.3
6.2	6.0	3.1	1.0	4.2
3.0	9.1	7.2	4.0	7.6

13.3 En el artículo "Shelf-Space Strategy in Retailing", que se publicó en *Proceedings: Southern Marketing Association*, se investigó el efecto que tenía la altura de los anaqueles en los supermercados sobre las ventas de alimento enlatado para perro. Durante un periodo de 8 días se llevó a cabo un experimento en un supermercado pequeño acerca de las ventas de una marca de alimento para perro conocida como Arf y se utilizaron tres niveles de altura de anaquel: a las rodillas, a la cintura y a los ojos. Cada día se cambió al azar tres veces la altura del anaquel en la que estaba dicho alimento. Las secciones restantes de la góndola que contenía la marca dada se llenaban con una mezcla de marcas de comida canina, las cuales resultaban tanto familiares como desconocidas para los consumidores de esa área geográfica específica. Se presentan las ventas diarias, expresadas en cientos de dólares, del alimento Arf para las tres alturas del anaquel. Con base en los datos, ¿existe una diferencia significativa en el promedio de ventas diarias de dicho alimento, con base en la altura del anaquel? Utilice un nivel de significancia de 0.01.

Altura de anaque!

Nivel de las rodillas	Nivel de la cintura	Nivel de los ojos
77	88	85
82	94	85
86	93	87
78	90	81
81	91	80
86	94	79
77	90	87
81	87	93

13.4 La inmovilización de los venados silvestres de cola blanca usando tranquilizantes da a los investigadores la oportunidad de estudiarlos de cerca y obtener información fisiológica valiosa. En el estudio denominado *Influence of Physical Restraint and Restraint Facilitating Drugs on Blood Measurements of White-Tailed Deer and Other Selected Mammals*, realizado en Virginia Tech, los biólogos de la vida silvestre probaron el tiempo del "derribo" (el periodo transcurrido entre la inyección y la inmovilización) de tres sustancias inmovilizadoras distintas. En este caso la inmovilización se define como el punto en que el animal ya no tiene control muscular suficiente para permanecer de pie. Se asignaron 30 venados machos de cola blanca al azar a cada uno de tres tratamientos. El grupo A recibió 5 miligramos de cloruro de succinilcolina líquida (SCC); al grupo B se le suministraron 8 miligramos de SCC en polvo; y al grupo C, 200 miligramos de hidrocloreto de fenciclidina. A continuación se presentan los tiempos de derribo, en minutos. Haga un análisis de varianza a un nivel de significancia de 0.01 y determine si el tiempo promedio de derribo es o no igual para las tres sustancias.

	Grupo		
	A	B	C
11	10	4	
5	7	4	
14	16	6	
7	7	3	
10	7	5	
7	5	6	
23	10	8	
4	10	3	
11	6	7	
11	12	3	

13.5 La enzima mitocondrial transhidrogenasa NADPH:NAD, de la tenia de la rata común (*Hymenolepis diminuta*) cataliza el hidrógeno en la transferencia de NADPH a NAD, lo que produce NADH. Se sabe que esta enzima desempeña un papel vital en el metabolismo anaerobio de la tenia, y recientemente se planteó la hipótesis de que podría servir como una bomba de intercambio de protones, es decir, para transferir protones a través de la membrana mitocondrial. Un estudio sobre el *Effect*

of Various Substrate Concentrations on the Conformational Variation of the NADPH:NAD Transhydrogenase of Hymenolepis diminuta llevado a cabo por la Bowling Green State University, se diseñó para evaluar la capacidad de dicha enzima para sufrir cambios en su conformación o su forma. Podría considerarse que los cambios en la actividad específica de la enzima ocasionados por las variaciones en la concentración de NADP sustentan la teoría del cambio de conformación. La enzima en cuestión se localiza en la membrana interior de las mitocondrias de la tenia. Se homogeneizaron las tenias y se aisló la enzima mediante una serie de centrifugaciones. Después se agregaron diferentes concentraciones de NADP a la solución de enzima aislada y la mezcla se incubó durante tres minutos en un baño de agua a 56°C. Luego, se analizó la enzima con un espectrómetro de rayo dual y se calcularon los resultados que se presentan a continuación, en términos de la actividad específica de la enzima, en nanomoles por minuto por miligramo de proteína. Pruebe la hipótesis de que la actividad específica promedio es la misma para las cuatro concentraciones, a un nivel de significancia de 0.01.

Concentración de NADP (nm)				
0	80	160	360	
11.01	11.38	11.02	6.04	10.31
12.09	10.67	10.67	8.65	8.30
10.55	12.33	11.50	7.76	9.48
11.26	10.08	10.31	10.13	8.89
			9.36	

13.6 Un estudio midió la tasa de sorción (ya sea absorción o adsorción) de tres tipos diferentes de solventes químicos orgánicos. Estos solventes se utilizan para limpiar partes industriales metálicas, y son desechos potencialmente riesgosos. Se probaron muestras independientes de solventes de cada tipo y se registraron sus tasas de sorción como un porcentaje molar. (Véase McClave, Dietrich y Sincich, 1997).

Aromáticos		Cloroalcalinos		Ésteres		
1.06	0.95	1.58	1.12	0.29	0.43	0.06
0.79	0.65	1.45	0.91	0.06	0.51	0.09
0.82	1.15	0.57	0.83	0.44	0.10	0.17
0.89	1.12	1.16	0.43	0.55	0.53	0.17
1.05				0.61	0.34	0.60

¿Existe una diferencia significativa en la tasa promedio de sorción de los tres solventes? Utilice un valor *P* para sus conclusiones. ¿Qué solvente usaría?

13.7 Se ha demostrado que el fertilizante fosfato amoniacal de magnesio, $MgNH_4PO_4$, es un proveedor eficaz de los nutrientes necesarios para el crecimiento de las plantas. Los compuestos que suministra son muy solubles en agua, lo cual permite su aplicación directa sobre la superficie del suelo o que se mezcle con el sustrato de crecimiento durante el proceso de encapsu-

lamiento. Se efectuó un estudio denominado *Effect of Magnesium Ammonium Phosphate on Height of Chrysanthemums* en George Mason University para determinar el nivel óptimo posible de la fertilización con base en la mejoría de la respuesta de crecimiento vertical del crisantemo. Se dividieron 40 semillas de crisantemo en 4 grupos de diez plantas cada uno. Se sembró cada una en una maceta similar que contenía un medio uniforme de crecimiento. Se agregó a cada grupo de plantas una concentración cada vez mayor de $MgNH_4PO_4$, medido en gramos por bushel. Los cuatro grupos de plantas se cultivaron durante cuatro semanas en condiciones uniformes en un invernadero. A continuación se presentan los tratamientos y los cambios respectivos de sus alturas, medidas en centímetros:

Tratamiento							
50 g/bu		100 g/bu		200 g/bu		400 g/bu	
13.2	12.4	16.0	12.6	7.8	14.4	21.0	14.8
12.8	17.2	14.8	13.0	20.0	15.8	19.1	15.8
13.0	14.0	14.0	23.6	17.0	27.0	18.0	26.0
14.2	21.6	14.0	17.0	19.6	18.0	21.1	22.0
15.0	20.0	22.2	24.4	20.2	23.2	25.0	18.2

A un nivel de significancia de 0.05, ¿podría concluirse que concentraciones diferentes de $MgNH_4PO_4$ afectan la altura promedio que alcanzan los crisantemos? ¿Qué cantidad del fertilizante parece ser la mejor?

13.8 Para el conjunto de datos del ejercicio 13.7 use la prueba de Bartlett para probar si las varianzas son iguales. Utilice $\alpha = 0.05$.

13.9 Utilice la prueba de Bartlett a un nivel de significancia de 0.01 para probar la homogeneidad de las varianzas en el ejercicio 13.5 de la página 519.

13.10 Utilice la prueba de Cochran a un nivel de significancia de 0.01 para probar la homogeneidad de las varianzas en el ejercicio 13.4 de la página 519.

13.11 Utilice la prueba de Bartlett a un nivel de significancia de 0.05 para probar la homogeneidad de las varianzas en el ejercicio 13.6 de la página 519.

13.5 Comparaciones de un grado de libertad

El análisis de varianza en la clasificación de un solo factor, o experimento de un solo factor, como se le denomina con frecuencia, tan sólo indica si puede rechazarse o no la hipótesis de medias de tratamientos iguales. Por lo general, el experimentador preferiría efectuar un análisis más profundo. Como ilustración, en el ejemplo 13.1, mediante el rechazo de la hipótesis nula, concluimos que las medias no son iguales, pero aún no sabemos en dónde residen las diferencias entre los agregados. Es probable que el ingeniero intuya de antemano que los agregados 1 y 2 deberían poseer propiedades similares de absorción, al igual que los agregados 3 y 5. Sin embargo, sería interesante estudiar las diferencias entre los dos grupos. Así, parece apropiado probar las hipótesis

$$H_0: \mu_1 + \mu_2 - \mu_3 - \mu_5 = 0,$$

$$H_1: \mu_1 + \mu_2 - \mu_3 - \mu_5 \neq 0.$$

Se observa que la hipótesis es una función lineal de las medias de la población, en las cuales los coeficientes suman cero.

Definición 13.1: Cualquier función lineal de la forma

$$\omega = \sum_{i=1}^k c_i \mu_i,$$

donde $\sum_{i=1}^k c_i = 0$ se llama **comparación** o **contraste** en las medias de los tratamientos.

Con frecuencia el experimentador puede hacer comparaciones múltiples probando la significancia de los contrastes de las medias de los tratamientos, es decir, probando una hipótesis del siguiente tipo:

Hipótesis para un
contraste

$$H_0: \sum_{i=1}^k c_i \mu_i = 0,$$

$$H_1: \sum_{i=1}^k c_i \mu_i \neq 0,$$

donde $\sum_{i=1}^k c_i = 0$.

La prueba se efectúa calculando primero un contraste similar de las medias de las muestras,

$$w = \sum_{i=1}^k c_i \bar{y}_i.$$

Como $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_k$ son variables aleatorias independientes que tienen distribuciones normales con medias $\mu_1, \mu_2, \dots, \mu_k$ y varianzas $\sigma_1^2/n_1, \sigma_2^2/n_2, \dots, \sigma_k^2/n_k$, respectivamente, el teorema 7.11 nos garantiza que w es un valor de la variable aleatoria normal W con

$$\text{media } \mu_W = \sum_{i=1}^k c_i \mu_i \text{ y varianza } \sigma_W^2 = \sigma^2 \sum_{i=1}^k \frac{c_i^2}{n_i}.$$

Por lo tanto, cuando H_0 es verdadera, $\mu_W = 0$ y, según el ejemplo 7.5, el estadístico

$$\frac{W^2}{\sigma_W^2} = \frac{\left(\sum_{i=1}^k c_i \bar{Y}_i \right)^2}{\sigma^2 \sum_{i=1}^k (c_i^2/n_i)}$$

se distribuye como una variable aleatoria chi cuadrada con 1 grado de libertad.

Estadístico de prueba para probar un contraste Nuestra hipótesis se prueba a un nivel de significancia α calculando

$$f = \frac{\left(\sum_{i=1}^k c_i \bar{y}_i \right)^2}{s^2 \sum_{i=1}^k (c_i^2/n_i)} = \frac{\left[\sum_{i=1}^k (c_i Y_i / n_i) \right]^2}{s^2 \sum_{i=1}^k (c_i^2/n_i)} = \frac{SC_W}{s^2}.$$

Aquí f es un valor de la variable aleatoria F que tiene distribución F con 1 y $N - k$ grados de libertad.

Cuando los tamaños de las muestras son iguales a n ,

$$SC_W = \frac{\left(\sum_{i=1}^k c_i Y_i \right)^2}{n \sum_{i=1}^k c_i^2}.$$

La cantidad SC_W , que se denomina **suma de cuadrados de los contrastes**, indica la parte de la SCT que se explica por el contraste en cuestión.

Esta suma de cuadrados se empleará para probar la hipótesis de que

$$\sum_{i=1}^k c_i \mu_i = 0.$$

Con frecuencia es de interés probar contrastes múltiples, en particular contrastes que son linealmente independientes u ortogonales. Como resultado, se vuelve necesaria la siguiente definición:

Definición 13.2: Se dice que los dos contrastes

$$\omega_1 = \sum_{i=1}^k b_i \mu_i \quad \text{y} \quad \omega_2 = \sum_{i=1}^k c_i \mu_i$$

son **ortogonales**, si $\sum_{i=1}^k b_i c_i / n_i = 0$, o bien, cuando las n_i son iguales a n , si

$$\sum_{i=1}^k b_i c_i = 0.$$

Si ω_1 y ω_2 son ortogonales, entonces las cantidades SC_{ω_1} y SC_{ω_2} son componentes de SCT , cada una con un solo grado de libertad. La suma de cuadrados de los tratamientos con $k-1$ grados de libertad se puede dividir en, a lo sumo, $k-1$ sumas de cuadrados de contrastes independientes con un solo grado de libertad que satisfacen la identidad

$$SCT = SC_{\omega_1} + SC_{\omega_2} + \dots + SC_{\omega_{k-1}},$$

si los contrastes son ortogonales entre sí.

Ejemplo 13.4: Remítase al ejemplo 13.1 y calcule la suma de cuadrados de los contrastes que corresponden a los contrastes ortogonales

$$\omega_1 = \mu_1 + \mu_2 - \mu_3 - \mu_5, \quad \omega_2 = \mu_1 + \mu_2 + \mu_3 - 4\mu_4 + \mu_5,$$

y efectúe las pruebas de significancia adecuadas. En este caso es de interés *a priori* comparar los dos grupos (1, 2) y (3, 5). Un contraste importante e independiente consiste en realizar la comparación entre el conjunto de agregados (1, 2, 3, 5) y el agregado 4.

Solución: Es evidente que los dos contrastes son ortogonales, puesto que

$$(1)(1) + (1)(1) + (-1)(1) + (0)(-4) + (-1)(1) = 0.$$

El segundo contraste indica una comparación entre los agregados (1, 2, 3 y 5) y el agregado 4. Podemos escribir dos contrastes adicionales ortogonales a los dos primeros, es decir:

$$\omega_3 = \mu_1 - \mu_2 \text{ (agregado 1 contra agregado 2),}$$

$$\omega_4 = \mu_3 - \mu_5 \text{ (agregado 3 contra agregado 5).}$$

De los datos de la tabla 13.1, se tiene que

$$SC_{w_1} = \frac{(3320 + 3416 - 3663 - 3664)^2}{6[(1)^2 + (1)^2 + (-1)^2 + (-1)^2]} = 14,553,$$

$$SC_{w_2} = \frac{[3320 + 3416 + 3663 + 3664 - 4(2791)]^2}{6[(1)^2 + (1)^2 + (1)^2 + (1)^2 + (-4)^2]} = 70,035.$$

En la tabla 13.5 se presenta un análisis de varianza más extenso. Se observa que las dos sumas de cuadrados de los contrastes explican casi toda la suma de cuadrados de los agregados. Existe una diferencia significativa entre las propiedades de absorción de los agregados, y el contraste ω_1 es significativo marginalmente. Sin embargo, el valor f de 14.12 para ω_2 es muy significativo, y se rechaza la hipótesis

$$H_0: \mu_1 + \mu_2 + \mu_3 + \mu_5 = 4\mu_4$$

Tabla 13.5: Análisis de varianza usando contrastes ortogonales

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Agregados	85,356	4	21,339	4.30
(1, 2) vs. (3,5)	{14,553	{ 1	{14,553	2.93
(1, 2, 3, 5) vs. 4	{70,035	{ 1	{70,035	14.12
Error	124,021	25	4961	
Total	209,377	29		

Los contrastes ortogonales permiten al profesional dividir la variación del tratamiento en componentes independientes. Por lo general el experimentador tiene interés en hacer ciertos contrastes. Eso ocurrió en nuestro ejemplo, donde las consideraciones *a priori* sugerían que los agregados (1, 2) y (3, 5) constituirían grupos distintos con propiedades diferentes de absorción, un planteamiento que no obtuvo mucho respaldo con la prueba de significancia. Sin embargo, la segunda comparación apoyó la conclusión de que el agregado 4 parecía “destacar” de los demás. En este caso no fue necesaria la partición completa de *SCT*, dado que dos de las cuatro comparaciones independientes posibles explicaban la mayor parte de la variación en los tratamientos.

En la figura 13.4 se presenta un procedimiento GLM del programa SAS, que proporciona un conjunto completo de contrastes ortogonales. Observe que la suma de cuadrados de los cuatro contrastes se agrega a la suma de cuadrados de los agregados. Asimismo, note que los últimos dos contrastes (1 contra 2, 3 contra 5) revelan comparaciones insignificantes. ▣

13.6 Comparaciones múltiples

El análisis de varianza es un procedimiento poderoso para probar la homogeneidad de un conjunto de medias. No obstante, si se rechazara la hipótesis nula y se aceptara la alternativa que se planteó (que no todas las medias son iguales), aún no se sabría cuáles de las medias de la población son iguales y cuáles son diferentes.

The GLM Procedure					
Dependent Variable: moisture					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	85356.4667	21339.1167	4.30	0.0088
Error	25	124020.3333	4960.8133		
Corrected Total	29	209376.8000			
	R-Square	Coeff Var	Root MSE	moisture Mean	
	0.407669	12.53703	70.43304	561.8000	
Source	DF	Type I SS	Mean Square	F Value	Pr > F
aggregate	4	85356.46667	21339.11667	4.30	0.0088
Source	DF	Type III SS	Mean Square	F Value	Pr > F
aggregate	4	85356.46667	21339.11667	4.30	0.0088
Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
(1,2,3,5) vs. 4	1	70035.00833	70035.00833	14.12	0.0009
(1,2) vs. (3,5)	1	14553.37500	14553.37500	2.93	0.0991
1 vs. 2	1	768.00000	768.00000	0.15	0.6973
3 vs. 5	1	0.08333	0.08333	0.00	0.9968

Figura 13.4: Un conjunto de contrastes ortogonales.

A menudo es de interés efectuar varias **comparaciones por pares** (quizá todas las que sean posibles) entre los tratamientos. En realidad, una comparación por pares se puede ver como un contraste simple, es decir, una prueba de

$$H_0: \mu_i - \mu_j = 0,$$

$$H_1: \mu_i - \mu_j \neq 0,$$

para toda $i \neq j$. Hacer todas las comparaciones posibles por pares entre las medias puede ser muy benéfico cuando no se conocen *a priori* contrastes complejos particulares. Por ejemplo, suponga que se desea probar las hipótesis siguientes, con los datos de los agregados de la tabla 13.1:

$$H_0: \mu_1 - \mu_5 = 0,$$

$$H_1: \mu_1 - \mu_5 \neq 0.$$

La prueba se desarrolla usando una F , una t , o el método de los intervalos de confianza. Si se usa la t , se tiene que

$$t = \frac{\bar{y}_1 - \bar{y}_5}{s\sqrt{2/n}},$$

donde s es la raíz cuadrada del cuadrado medio del error y $n = 6$ es el tamaño de la muestra por tratamiento. En este caso,

$$t = \frac{553.33 - 610.67}{\sqrt{4961}\sqrt{1/3}} = -1.41.$$

El valor P para la prueba t con 25 grados de libertad es 0.17. Así que no hay evidencia suficiente para rechazar H_0 .

Relación entre T y F

Anteriormente se expuso el uso de una prueba t agrupada, junto con los lineamientos que se estudiaron en el capítulo 10. El estimado agrupado se tomó del cuadrado medio del error con el fin de aprovechar los grados de libertad que están agrupados en las cinco muestras. Además, probamos un contraste. El lector debería observar que si el valor t se eleva al cuadrado, el resultado tiene exactamente la misma forma que el valor de f para una prueba del contraste, analizada en la sección anterior. En efecto,

$$f = \frac{(\bar{y}_1 - \bar{y}_5)^2}{s^2(1/6 + 1/6)} = \frac{(553.33 - 610.67)^2}{4961(1/3)} = 1.988,$$

que es, por supuesto, t^2 .

Método del intervalo de confianza para una comparación por pares

Es fácil resolver el mismo problema de una comparación por pares (o un contraste) usando el método del intervalo de confianza. Es claro que, si se calcula un intervalo de confianza del $100(1 - \alpha)\%$ sobre $\mu_1 - \mu_5$, se tiene que

$$\bar{y}_1 - \bar{y}_5 \pm t_{\alpha/2} s \sqrt{\frac{2}{6}},$$

donde $t_{\alpha/2}$ es el punto superior de $100(1 - \alpha/2)\%$ de una distribución t con 25 grados de libertad (grados de libertad que provienen de s^2). Esta conexión inmediata entre las pruebas de hipótesis y los intervalos de confianza debería ser evidente a partir de los análisis de los capítulos 9 y 10. La prueba de un contraste simple $\mu_1 - \mu_5$ no implica más que observar si el intervalo de confianza anterior cubre o no al cero. Al sustituir los números se tiene que el intervalo de confianza de 95%:

$$(553.33 - 610.67) \pm 2.060 \sqrt{4961} \sqrt{\frac{1}{3}} = -57.34 \pm 83.77.$$

Por consiguiente, como el intervalo de confianza cubre al cero, el contraste no es significativo. En otras palabras, no hay diferencia significativa entre las medias de los agregados 1 y 5.

Tasa de error por experimento

Se presentan muchas dificultades cuando el analista intenta hacer muchas o todas las comparaciones por pares posibles. Para el caso de k medias habrá, desde luego, $r = k(k - 1)/2$ comparaciones por pares posibles. Si se suponen comparaciones independientes, la tasa de error por experimento o tasa de error por familia, es decir, la probabilidad de un falso rechazo de al menos una de las hipótesis, es dada por $1 - (1 - \alpha)^r$, donde α es la probabilidad seleccionada del error tipo I para una comparación específica. Es claro que esta medida del error tipo I por experimento sería bastante grande. Por

ejemplo, aun si sólo hubiera 6 comparaciones, digamos, en el caso de 4 medias, y $\alpha = 0.05$, la tasa de experimento-juicio sería

$$1 - (0.95)^6 \approx 0.26.$$

Cuando se prueban muchas comparaciones por pares, por lo general existe la necesidad de hacer el contraste efectivo sobre una sola comparación más conservadora. Es decir, usando el método del intervalo de confianza, los intervalos de confianza serían mucho más anchos que $\pm t_{\alpha/2} s \sqrt{2/n}$ que se emplea para el caso de una sola comparación.

Prueba de Tukey

Hay varios métodos estándar para realizar comparaciones por pares que den credibilidad a la tasa del error tipo I. Aquí se analizarán e ilustrarán dos de ellos. El primero, denominado **procedimiento de Tukey**, permite la formación de intervalos de confianza del $100(1 - \alpha)\%$ simultáneos para todas las comparaciones por pares. El método se basa en la distribución del rango *estudentizado*. El punto apropiado del percentil es una función de α , k y $v =$ grados de libertad para s^2 . En la tabla A.12 se presenta una lista de puntos porcentuales superiores adecuados para $\alpha = 0.05$. El método de Tukey de comparaciones por pares implica encontrar una diferencia significativa entre las medias i y j ($i \neq j$) si $|\bar{y}_i - \bar{y}_j|$ excede a $q(\alpha, k, v) \sqrt{\frac{s^2}{n}}$.

El procedimiento de Tukey se ilustra con facilidad. Considere un ejemplo hipotético en el que se tienen 6 tratamientos en un diseño completamente aleatorizado de un solo factor, en el que se hacen 5 observaciones por tratamiento. Suponga que el cuadrado medio del error tomado de la tabla del análisis de varianza es $s^2 = 2.45$ (24 grados de libertad). Las medias muestrales están en orden ascendente,

$$\begin{array}{cccccc} \bar{y}_2 & \bar{y}_5 & \bar{y}_1 & \bar{y}_3 & \bar{y}_6 & \bar{y}_4 \\ 14.50 & 16.75 & 19.84 & 21.12 & 22.90 & 23.20. \end{array}$$

Con $\alpha = 0.05$, el valor de $q(0.05, 6, 24)$ es 4.37. Así, todas las diferencias absolutas se comparan con

$$4.37 \sqrt{\frac{2.45}{5}} = 3.059.$$

Como resultado, las siguientes representan medias que, usando el procedimiento de Tukey, se encuentra que son significativamente diferentes:

$$\begin{array}{cccccc} 4 \text{ y } 1, & 4 \text{ y } 5, & 4 \text{ y } 2, & 6 \text{ y } 1, & 6 \text{ y } 5, \\ 6 \text{ y } 2, & 3 \text{ y } 5, & 3 \text{ y } 2, & 1 \text{ y } 5, & 1 \text{ y } 2. \end{array}$$

¿De dónde proviene el nivel α en la prueba de Tukey?

Se mencionó brevemente el concepto de **intervalos de confianza simultáneos** que se emplean para el procedimiento de Tukey. El lector obtendrá una perspectiva útil del concepto de comparaciones múltiples, si comprende el significado de los intervalos de confianza simultáneos.

En el capítulo 9 vimos que, si se calcula un intervalo de confianza de 95% para, digamos, una media μ , entonces la probabilidad de que el intervalo cubra la media verdadera μ es 0.95.

Sin embargo, como vimos antes, para el caso de comparaciones múltiples la probabilidad efectiva de interés está ligada con la tasa de error por experimento, y debe hacerse énfasis en que los intervalos de confianza del tipo $\bar{y}_i - \bar{y}_j \pm q(\alpha, k, v)s\sqrt{1/n}$ no son independientes, ya que todos implican a s y muchos utilizan los mismos promedios, las \bar{y}_i . A pesar de tales dificultades, si se utiliza la $q(0.05, k, v)$, el nivel de confianza simultáneo está controlado en un 95%. Lo mismo es cierto para $q(0.01, k, v)$, es decir, el nivel de confianza está controlado en un 99%. En el caso de $\alpha = 0.05$, hay una probabilidad de 0.05 de que se encuentre falsamente que al menos un par de mediciones son diferentes (falso rechazo de al menos una hipótesis nula). En el caso de $\alpha = 0.01$, la probabilidad correspondiente será 0.01.

Prueba de Duncan

El segundo procedimiento que se estudiará se denomina **procedimiento de Duncan o prueba de Duncan de rango múltiple**. Este procedimiento también se basa en el concepto general del rango estudentizado. El rango de cualquier subconjunto de p medias muestrales debe exceder cierto valor antes de que se encuentre que cualquiera de las p medias es diferente. Este valor recibe el nombre de **rango de menor significancia** para las p medias, y se denota por R_p , donde

$$R_p = r_p \sqrt{\frac{s^2}{n}}$$

Los valores de la cantidad r_p , llamados **rango estudentizado de menor significancia**, dependen del nivel de significancia deseado y del número de grados de libertad del cuadrado medio del error. Estos valores se obtienen de la tabla A.13 para $p = 2, 3, \dots, 10$ medias.

Para ilustrar el procedimiento de prueba de rango múltiple, consideremos el ejemplo hipotético en el cual se comparan 6 tratamientos con 5 observaciones por tratamiento. Se trata del mismo ejemplo que se empleó para ilustrar la prueba de Tukey. Se obtiene R_p multiplicando cada r_p por 0.70. Los resultados de estos cálculos se resumen como sigue:

p	2	3	4	5	6
r_p	2.919	3.066	3.160	3.226	3.276
R_p	2.043	2.146	2.212	2.258	2.293

Si se comparan estos rangos de menor significancia con las diferencias en medias ordenadas, se llega a las conclusiones siguientes:

1. Como $\bar{y}_4 - \bar{y}_2 = 8.70 > R_6 = 2.293$, se concluye que μ_4 y μ_2 son significativamente distintas.
2. Si se comparan $\bar{y}_4 - \bar{y}_5$ y $\bar{y}_6 - \bar{y}_2$ con R_3 , se concluye que μ_4 es significativamente mayor que μ_5 y que μ_6 es significativamente mayor que μ_2 .
3. Si se comparan $\bar{y}_4 - \bar{y}_1$, $\bar{y}_6 - \bar{y}_5$ y $\bar{y}_3 - \bar{y}_2$ con R_4 , se concluye que cada diferencia es significativa.
4. Si se comparan $\bar{y}_4 - \bar{y}_3$, $\bar{y}_6 - \bar{y}_1$, $\bar{y}_3 - \bar{y}_5$ y $\bar{y}_1 - \bar{y}_2$ con R_5 , se encuentra que todas las diferencias son significativas excepto para $\mu_4 - \mu_3$. Por lo tanto, μ_4 , μ_5 y μ_6 constituyen un subconjunto de medias homogéneas.
5. Si se comparan $\bar{y}_3 - \bar{y}_1$, $\bar{y}_1 - \bar{y}_5$ y $\bar{y}_5 - \bar{y}_2$ con R_2 , se concluye que sólo μ_1 y μ_1 no son significativamente distintas.

Se acostumbra resumir las conclusiones anteriores con el dibujo de una línea debajo de cualquier subconjunto de medias adyacentes que no sean significativamente diferentes. Así, tenemos

\bar{y}_2	\bar{y}_5	\bar{y}_1	\bar{y}_3	\bar{y}_6	\bar{y}_4
14.50	16.75	19.84	21.12	22.90	23.20

Es evidente que en este caso los resultados con los procedimientos de Tukey y Duncan son muy similares. El procedimiento de Tukey no detectó ninguna diferencia entre 2 y 5; mientras que el de Duncan sí lo hizo.

Prueba de Dunnett: comparación de tratamientos con un control

En muchos problemas científicos y de ingeniería no nos interesa hacer inferencias acerca de todas las comparaciones posibles entre las medias de los tratamientos del tipo $\mu_i - \mu_j$. En vez de ello es frecuente que el experimento dicte la necesidad de comparar simultáneamente cada *tratamiento* con un *control*. Un procedimiento de prueba desarrollado por C. W. Dunnett determina diferencias significativas entre cada media de tratamiento y el control, con un solo nivel conjunto de significancia α . Para ilustrar el procedimiento de Dunnett, se considerarán los datos experimentales de la tabla 13.6 para la clasificación de un solo factor, donde se estudió el efecto de tres catalizadores sobre el producto de una reacción. Como control se emplea un cuarto tratamiento en el que no se aplica un catalizador.

Tabla 13.6: Producto de una reacción

Control	Catalizador 1	Catalizador 2	Catalizador 3
50.7	54.1	52.7	51.2
51.5	53.8	53.9	50.8
49.2	53.1	57.0	49.7
53.1	52.5	54.1	48.0
52.7	54.0	52.5	47.2
$\bar{y}_0 = 51.44$	$\bar{y}_1 = 53.50$	$\bar{y}_2 = 54.04$	$\bar{y}_3 = 49.38$

En general, se desea probar las k hipótesis

$$\left. \begin{array}{l} H_0: \mu_0 = \mu_i \\ H_1: \mu_0 \neq \mu_i \end{array} \right\} \quad i = 1, 2, \dots, k,$$

donde μ_0 representa el producto medio para la población de medidas en que se utiliza el control. Como se mencionó en la sección 13.3, se espera que las suposiciones habituales del análisis de varianza sigan siendo válidas. Para probar la hipótesis nula especificada con H_0 en comparación con alternativas bilaterales para una situación experimental donde existen k tratamientos, sin incluir el control, y n observaciones por tratamiento, primero calculamos los valores

$$d_i = \frac{\bar{y}_i - \bar{y}_0}{\sqrt{2s^2/n}}, \quad i = 1, 2, \dots, k,$$

Como antes, la varianza muestral s^2 se obtiene a partir del cuadrado medio del error en el análisis de varianza. Ahora bien, la región crítica para rechazar H_0 a un nivel de significancia α , se establece con la desigualdad

$$|d_i| > d_{\alpha/2}(k, \nu),$$

donde ν es el número de grados de libertad para el cuadrado medio del error. Los valores de la cantidad $d_{\alpha/2}(k, \nu)$ para una prueba de dos colas se incluyen en la tabla A.14 para $\alpha = 0.05$ y $\alpha = 0.01$, para diversos valores de k y ν .

Ejemplo 13.5: Para los datos de la tabla 13.6, pruebe la hipótesis que compara cada catalizador con el control, usando alternativas bilaterales. Como nivel de significancia conjunto elija $\alpha = 0.05$.

Solución: El cuadrado medio del error con 16 grados de libertad se obtiene de la tabla de análisis de varianza, usando todos los tratamientos $k + 1$. El cuadrado medio del error es dado por

$$s^2 = \frac{36.812}{16} = 2.30075 \text{ y } \sqrt{\frac{2s^2}{n}} = \sqrt{\frac{(2)(2.30075)}{5}} = 0.9593.$$

Entonces,

$$d_1 = \frac{53.50 - 51.44}{0.9593} = 2.147, \quad d_2 = \frac{54.04 - 51.44}{0.9593} = 2.710,$$

$$d_3 = \frac{49.38 - 51.44}{0.9593} = -2.147.$$

De la tabla A.14 el valor crítico para $\alpha = 0.05$ resulta ser $d_{0.025}(3, 16) = 2.59$. Como $|d_1| < 2.59$ y $|d_3| < 2.59$, se concluye que tan sólo la producción media para el catalizador 2 es significativamente diferente de la respuesta media de la reacción utilizando el control. \blacksquare

Muchas aplicaciones prácticas imponen la necesidad de una prueba de una cola para comparar los tratamientos con un control. En efecto, cuando un farmacólogo está interesado en el efecto de varias dosis de un medicamento sobre el nivel del colesterol, y su control consiste en una dosis de cero, sería interesante determinar si cada dosis produce una reducción significativamente mayor que la del control. En la tabla A.15 se presentan los valores críticos de $d_{\alpha}(k, \nu)$ para alternativas unilaterales.

Ejercicios

13.12 Considere los datos del ejercicio de repaso 13.45 de la página 555. Efectúe pruebas de significancia sobre los siguientes contrastes:

- B contra A, C y D;
- C contra A y D;
- A contra D.

13.13 El propósito del estudio *The Incorporation of a Chelating Agent into a Flame Retardant Finish of a Cotton Flannelette and the Evaluation of Selected Fabric Properties*, llevado a cabo en Virginia Tech, fue evaluar el uso de un agente quelante como parte del acabado retardante del fuego de la franela de algodón, determinando sus efectos en la inflamabilidad después de lavar la tela en condiciones específicas. Se pre-

pararon dos baños, uno con celulosa de carboximetilo y otro sin ella. Se lavaron 12 piezas de tela 5 veces en el baño I, y otras 12 piezas se lavaron 10 veces en el baño I. Esto se repitió con 24 piezas adicionales de tela en el baño II. Después de los lavados se midieron las longitudes quemadas de la tela, así como los tiempos de combustión. Por conveniencia, definamos los siguientes tratamientos:

Tratamiento 1: 5 lavados en el baño I.

Tratamiento 2: 5 lavados en el baño II.

Tratamiento 3: 10 lavados en el baño I.

Tratamiento 4: 10 lavados en el baño II.

Los registros del tiempo de combustión, en segundos, son los siguientes:

Tratamiento			
1	2	3	4
13.7	6.2	27.2	18.2
23.0	5.4	16.8	8.8
15.7	5.0	12.9	14.5
25.5	4.4	14.9	14.7
15.8	5.0	17.1	17.1
14.8	3.3	13.0	13.9
14.0	16.0	10.8	10.6
29.4	2.5	13.5	5.8
9.7	1.6	25.5	7.3
14.0	3.9	14.2	17.7
12.3	2.5	27.4	18.3
12.3	7.1	11.5	9.9

- a) Efectúe un análisis de varianza a un nivel de significancia de 0.01, y determine si hay diferencias significativas entre las medias de los tratamientos.
- b) Use contrastes de un solo grado de libertad con $\alpha = 0.01$ para comparar el tiempo medio de combustión del tratamiento 1 en comparación con el tratamiento 2, y también del tratamiento 3 en comparación con el 4.

13.14 El Departamento de Alimentación y Nutrición Humana de Virginia Tech realizó el estudio *Loss of Nitrogen Through Sweat by Preadolescent Boys Consuming Three Levels of Dietary Protein* para determinar la pérdida de nitrógeno por transpiración con varios niveles dietéticos de proteínas. En el experimento participaron 12 hombres preadolescentes cuyas edades iban de 7 años 8 meses a 9 años 8 meses, y a quienes de les calificó de clínicamente saludables. Cada muchacho estuvo sujeto a una de tres dietas controladas en las cuales consumía 29, 54 u 84 gramos de proteínas por día. Los siguientes datos representan la pérdida de nitrógeno corporal por transpiración, en miligramos, recabados durante los dos días últimos del periodo de experimentación:

Nivel de proteínas		
29 gramos	54 gramos	84 gramos
190	318	390
266	295	321
270	271	396
	438	399
	402	

- a) Realice un análisis de varianza a un nivel de significancia de 0.05, para demostrar que las pérdidas medias de nitrógeno por transpiración son diferentes con los tres niveles de proteínas.
- b) Utilice una prueba de Tukey para determinar cuáles niveles de proteínas difieren significativamente entre sí en la pérdida media de nitrógeno.

13.15 Utilice la prueba de Tukey a un nivel de significancia de 0.05, para analizar las medias de las 5 mar-

cas diferentes de tabletas para el dolor de cabeza del ejercicio 13.2 de la página 518.

13.16 Se realizó una investigación para determinar la fuente de reducción en el rendimiento de cierto producto químico. Se sabía que la pérdida en el rendimiento ocurría en el licor madre, es decir, el material eliminado en la etapa de filtración. Se pensaba que mezclas distintas del material original podrían ocasionar reducciones diferentes del rendimiento en la etapa de licor madre. A continuación se presentan los resultados de la reducción porcentual para tres lotes de cada una de cuatro mezclas seleccionadas con anterioridad.

Mezcla			
1	2	3	4
25.6	25.2	20.8	31.6
24.3	28.6	26.7	29.8
27.9	24.7	22.2	34.3

- a) Haga el análisis de varianza al nivel de significancia $\alpha = 0.05$.
- b) Utilice la prueba de Duncan de rango múltiple para determinar cuáles mezclas difieren.
- c) Resuelva el inciso b usando la prueba de Tukey.

13.17 En el estudio, denominado *An Evaluation of the Removal Method for Estimating Benthic Populations and Diversity*, realizado por Virginia Tech en el río Jackson, se emplearon 5 procedimientos distintos de muestreo para determinar los conteos de especies. Se seleccionaron 20 muestras al azar y los 5 procedimientos de muestreo se repitieron 4 veces. Se registraron los siguientes conteos de especies:

Disminución	Procedimiento de muestreo			
	De Hess modificado	Remoción del sustrato de Surber	Kicknet	Kicknet
85	75	31	43	17
55	45	20	21	10
40	35	9	15	8
77	67	37	27	15

- a) ¿Hay alguna diferencia significativa en el conteo promedio de especies para los distintos procedimientos de muestreo? Use un valor P en su conclusión.
- b) Emplee una prueba de Tukey con $\alpha = 0.05$ para determinar cuáles procedimientos de muestreo difieren.

13.18 Los siguientes datos son valores de presión (psi) en un resorte de torsión para valores distintos del ángulo entre las vueltas del resorte en posición libre.

Ángulo					
67	71	75	79	83	
83	84	86	87	89	90
85	85	87	87	90	92
	85	88	88	90	
	86	88	88	91	
	86	88	89		
	87	90			

Calcule un análisis de varianza de un solo factor para este experimento y plantee sus conclusiones acerca del efecto que tiene el ángulo sobre la presión en el resorte. (Tomado de C. R. Hicks, *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart y Winston, Nueva York, 1973).

13.19 Se sospecha que la temperatura del ambiente en que se activan las baterías afecta su vida. Se probaron 30 baterías homogéneas, seis a cada una de cinco temperaturas, y los datos se presentan a continuación (vida activada en segundos). Analice e interprete los datos. (Tomado de C. R. Hicks, *Fundamental Concepts in Design of Experiments*, Holt, Rinehart y Winston, Nueva York, 1973.)

	Temperatura(°C)				
	0	25	50	75	100
	55	60	70	72	65
	55	61	72	72	66
	57	60	72	72	60
	54	60	68	70	64
	54	60	77	68	65
	56	60	77	69	65

13.20 La tabla siguiente (tomada de A. Hald, *Statistical Theory with Engineering Applications*, John Wiley & Sons, Nueva York, 1952) proporciona las resistencias a la tensión (en desviaciones desde 340) para conductores extraídos de nueve cables que deben usarse para una red de alto voltaje. Cada cable está constituido por 12 conductores. Se desea saber si las resistencias medias de los conductores en los nueve cables son las mismas. Si los cables son diferentes, ¿cuáles son los que difieren? Utilice un valor P en su análisis de varianza.

Cable	Resistencia a la tensión											
1	5	-13	-5	-2	-10	-6	-5	0	-3	2	-7	-5
2	-11	-13	-8	8	-3	-12	-12	-10	5	-6	-12	-10
3	0	-10	-15	-12	-2	-8	-5	0	-4	-1	-5	-11
4	-12	4	2	10	-5	-8	-12	0	-5	-3	-3	0
5	7	1	5	0	10	6	5	2	0	-1	-10	-2
6	1	0	-5	-4	-1	0	2	5	1	-2	6	7
7	-1	0	2	1	-4	2	7	5	1	0	-4	2
8	-1	0	7	5	10	8	1	2	-3	6	0	5
9	2	6	7	8	15	11	-7	7	10	7	8	1

13.21 La salida de resultados que se presenta en la figura 13.5 de la página 532 proporciona información sobre la prueba de Duncan para los datos de los agregados del ejemplo 13.1 obtenidos con la función PROC GLM del programa SAS. Saque conclusiones sobre las comparaciones por pares usando los resultados de la prueba de Duncan.

13.22 Realice la prueba de Duncan para comparaciones por pares con los datos del ejercicio 13.6 de la página 519. Comente los resultados.

13.23 En un experimento biológico se emplearon 4 concentraciones de cierto producto químico para mejorar el crecimiento de cierto tipo de planta con el paso del tiempo. Se utilizaron cinco plantas con cada concentración y se midió su crecimiento, en centímetros. Se obtuvieron los siguientes datos y también se aplicó un control (ausencia de producto químico)

Control	Concentración			
	1	2	3	4
6.8	8.2	7.7	6.9	5.9
7.3	8.7	8.4	5.8	6.1
6.3	9.4	8.6	7.2	6.9
6.9	9.2	8.1	6.8	5.7
7.1	8.6	8.0	7.4	6.1

Utilice una prueba bilateral de Dunnett a un nivel de significancia de 0.05 para comparar de manera simultánea las concentraciones con el control.

13.24 La estructura financiera de una empresa consiste en la forma en que sus activos se dividen en capital y deuda, y el apalancamiento financiero se refiere al porcentaje de activos financiados con endeudamiento. En el artículo *The Effect of Financial Leverage on Return*, Tai Ma, de Virginia Tech, afirma que es posible utilizar el apalancamiento financiero para incrementar la tasa de rendimiento sobre el capital. Dicho de otra manera, los accionistas pueden recibir rendimientos más elevados sobre el capital propio con la misma cantidad de inversión si usan apalancamiento financiero. Los siguientes datos muestran las tasas de rendimiento sobre el capital utilizando 3 niveles distintos de apalancamiento financiero, así como un nivel de control (deuda igual a cero) para 24 empresas seleccionadas al azar.

Control	Apalancamiento financiero		
	Bajo	Medio	Alto
2.1	6.2	9.6	10.3
5.6	4.0	8.0	6.9
3.0	8.4	5.5	7.8
7.8	2.8	12.6	5.8
5.2	4.2	7.0	7.2
2.6	5.0	7.8	12.0

Fuente: Standard & Poor's *Machinery Industry Survey*, 1975.

- Haga el análisis de varianza a un nivel de significancia de 0.05.
- Use una prueba de Dunnett a un nivel de significancia de 0.01, para determinar si las tasas medias de rendimiento sobre el capital son más elevadas con los niveles bajo, medio y alto de apalancamiento financiero que con el nivel de control.

The GLM Procedure				
Duncan's Multiple Range Test for moisture				
NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.				
Alpha	0.05			
Error Degrees of Freedom	25			
Error Mean Square	4960.813			
Number of Means	2	3	4	5
Critical Range	83.75	87.97	90.69	92.61
Means with the same letter are not significantly different.				
Duncan Grouping	Mean	N	aggregate	
A	610.67	6	5	
A				
A	610.50	6	3	
A				
A	569.33	6	2	
A				
A	553.33	6	1	
B	465.17	6	4	

Figura 13.5: Salida de resultados del SAS para el ejercicio 13.21.

13.7 Comparación de un conjunto de tratamientos en bloques

En la sección 13.2 estudiamos la idea de la formación de bloques, es decir, de aislar conjuntos de unidades experimentales que son razonablemente homogéneas y asignarles tratamientos de forma aleatoria. Ésta es una extensión del concepto de "formar pares" que se analizó en los capítulos 9 y 10, y se hace para reducir el error experimental, ya que las unidades en un bloque tienen más características comunes que las unidades localizadas en diferentes bloques.

El lector no debería considerar los bloques como un segundo factor, aunque ésa sea una forma tentadora de visualizar el diseño. De hecho, el factor principal (los tratamientos) aún lleva el peso mayor del experimento. Las unidades experimentales siguen siendo la fuente del error, igual que en el diseño completamente aleatorizado. Con la formación de bloques simplemente tratamos a dichas unidades de manera más sistemática. De ese modo, se dice que la aleatoriedad tiene restricciones. Antes de iniciar el estudio de la formación de bloques revisaremos dos ejemplos de un **diseño completamente aleatorizado**. El primer ejemplo es un experimento químico diseñado para determinar si hay una diferencia en la reacción media producida por cuatro catalizadores. Las muestras de los materiales que tienen que probarse se extraen de los mismos lotes de materias primas, a la vez que se mantienen constantes otras condiciones como la temperatura y concentración de los reactivos. En este caso, la hora del día en que se efectúan las corridas experimentales podría representar las unidades experimentales, y si el experimentador considera que es posible que haya un ligero efecto del tiempo, aleatorizaría la asignación

de los catalizadores a las corridas para contrarrestar la posible tendencia. Como un segundo ejemplo de dicho diseño, considere un experimento para comparar cuatro métodos para medir una propiedad física en particular de un fluido. Suponga que el proceso de muestreo es destructivo, es decir, que una vez que se ha medido una muestra de la sustancia usando un método, ya no puede medirse nuevamente con ningún otro. Si se decide hacer cinco mediciones con cada método, entonces se seleccionan al azar 20 muestras del material de un lote grande y se utilizan en el experimento para comparar los cuatro métodos de medición. Las unidades experimentales son las muestras seleccionadas al azar. Cualquier variación de una muestra a otra aparecerá en la variación del error, según se mida con s^2 en el análisis.

¿Cuál es el propósito de formar bloques?

Si la variación debida a la heterogeneidad en las unidades experimentales es tan grande que la sensibilidad para detectar diferencias de tratamiento se reduce debido a un valor aumentado de s^2 , un plan mejor sería “bloquear” la variación debida a esas unidades y, por consiguiente, reducir la variación ajena a la que es explicada por bloques más pequeños o más homogéneos. Por ejemplo, suponga que en el ejemplo anterior de los catalizadores se supiera *a priori* que existe en definitiva un efecto significativo diario sobre el producto, y que es posible medir el producto para cuatro catalizadores en un día específico. En lugar de asignar los 4 catalizadores a las 20 corridas de prueba completamente al azar, se eligen, por ejemplo, 5 días y se prueba cada uno de los cuatro catalizadores cada día, asignándolos al azar a las corridas dentro de los días. De esta manera se elimina la variación diaria del análisis y, en consecuencia, el error experimental, que aún incluye cualquier tendencia temporal *dentro de los días*, representa con más precisión la variación aleatoria. A cada día se le denomina **bloque**.

El más directo de los diseños aleatorizados de bloques es aquel en el cual se asigna al azar un tratamiento a la vez a cada bloque. A un plan experimental como éste se le denomina **diseño de bloques completos aleatorizados (BCA)** y cada bloque constituye una sola réplica de los tratamientos.

13.8 Diseños de bloques completos aleatorizados

Un plan clásico para el diseño de bloques completos aleatorizados (BCA) usando tres mediciones en cuatro bloques es el siguiente:

Bloque 1	Bloque 2	Bloque 3	Bloque 4
t_2	t_1	t_3	t_2
t_1	t_3	t_2	t_1
t_3	t_2	t_1	t_3

Las t denotan la asignación de cada uno de 3 tratamientos a los bloques. Desde luego, la asignación verdadera de los tratamientos a las unidades dentro de los bloques se hace al azar. Una vez que ha finalizado el experimento, los datos se pueden registrar como en el siguiente arreglo de 3×4 :

Tratamiento	Bloque: 1	2	3	4
1	y_{11}	y_{12}	y_{13}	y_{14}
2	y_{21}	y_{22}	y_{23}	y_{24}
3	y_{31}	y_{32}	y_{33}	y_{34}

donde y_{11} representa la respuesta que se obtiene al utilizar el tratamiento 1 en el bloque 1, y_{12} es la respuesta que se obtiene al utilizar el tratamiento 1 en el bloque 2, ..., y y_{34} es la respuesta que se obtiene al utilizar el tratamiento 3 en el bloque 4.

Ahora vamos a generalizar y a considerar el caso de k tratamientos asignados a b bloques. Los datos se pueden resumir tal como se observa en el arreglo rectangular de $k \times b$ de la tabla 13.7. Se supondrá que las y_{ij} , $i = 1, 2, \dots, k$ y $j = 1, 2, \dots, b$, son valores de variables aleatorias independientes que tienen distribuciones normales con media μ_{ij} y varianzas común σ^2 .

Tabla 13.7: Arreglo de $k \times b$ para el diseño de BCA

Tratamiento	Bloque						Total	Media
	1	2	...	j	...	b		
1	y_{11}	y_{12}	...	y_{1j}	...	y_{1b}	$T_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2b}	$T_{2.}$	$\bar{y}_{2.}$
⋮	⋮	⋮	...	⋮	...	⋮	⋮	⋮
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ib}	$T_{i.}$	$\bar{y}_{i.}$
⋮	⋮	⋮	...	⋮	...	⋮	⋮	⋮
k	y_{k1}	y_{k2}	...	y_{kj}	...	y_{kb}	$T_{k.}$	$\bar{y}_{k.}$
Total	$T_{.1}$	$T_{.2}$...	$T_{.j}$...	$T_{.b}$	$T_{..}$	
Media	$\bar{y}_{.1}$	$\bar{y}_{.2}$...	$\bar{y}_{.j}$...	$\bar{y}_{.b}$		$\bar{y}_{..}$

Sea μ_i el promedio (en lugar del total) de las b medias de la población para el i -ésimo tratamiento. Es decir,

$$\mu_i = \frac{1}{b} \sum_{j=1}^b \mu_{ij}, \text{ para } i = 1, \dots, k.$$

De manera similar, el promedio de las medias de la población para el j -ésimo bloque, μ_j , es definido por

$$\mu_j = \frac{1}{k} \sum_{i=1}^k \mu_{ij}, \text{ para } j = 1, \dots, b$$

y el promedio de las bk medias de la población, μ , es definido por

$$\mu = \frac{1}{bk} \sum_{i=1}^k \sum_{j=1}^b \mu_{ij}.$$

Para determinar si parte de la variación de nuestras observaciones se debe a diferencias entre los tratamientos, se considera la siguiente prueba:

Hipótesis de
medias iguales
de los tratamientos

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k = \mu,$$

$$H_1: \text{No todas las } \mu_i \text{ son iguales.}$$

Modelo para el diseño BCA

Cada observación se puede escribir en la forma siguiente:

$$y_{ij} = \mu_{ij} + \epsilon_{ij},$$

donde ϵ_{ij} mide la desviación del valor observado y_{ij} de la media de la población μ_{ij} . La forma preferida de esta ecuación se obtiene sustituyendo

$$\mu_{ij} = \mu + \alpha_i + \beta_j,$$

donde α_i es, como antes, el efecto del i -ésimo tratamiento, y β_j es el efecto del j -ésimo bloque. Se supone que el tratamiento y los efectos de los bloques son aditivos. Por lo tanto, se puede escribir

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}.$$

Observe que el modelo se parece al de clasificación de un solo factor; la diferencia esencial es la introducción del efecto de bloque β_j . El concepto básico se parece mucho al de la clasificación de un solo factor, excepto que en el análisis debe tomarse en cuenta el efecto adicional debido a los bloques, ya que ahora la variación se controla de manera sistemática *en dos direcciones*. Si ahora imponemos las restricciones de que

$$\sum_{i=1}^k \alpha_i = 0 \quad \text{y} \quad \sum_{j=1}^b \beta_j = 0,$$

entonces

$$\mu_i = \frac{1}{b} \sum_{j=1}^b (\mu + \alpha_i + \beta_j) = \mu + \alpha_i, \text{ para } i = 1, \dots, k,$$

y

$$\mu_j = \frac{1}{k} \sum_{i=1}^k (\mu + \alpha_i + \beta_j) = \mu + \beta_j, \text{ para } j = 1, \dots, b.$$

La hipótesis nula de que las k medias de los tratamientos μ_i son iguales y, por lo tanto, iguales a μ , ahora es **equivalente a probar la hipótesis:**

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0,$$

$$H_1: \text{Al menos una de las } \alpha_i \text{ no es igual a cero.}$$

Cada una de las pruebas de tratamientos se basará en una comparación de los estimados independientes de la varianza común poblacional σ^2 . Esos estimados se

obtendrán separando la suma total de cuadrados de los datos en tres componentes mediante la siguiente identidad:

Teorema 13.3: Identidad de la suma de cuadrados

$$\sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2 = b \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 + k \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

La demostración se deja como ejercicio para el lector.

La identidad de la suma de cuadrados podría presentarse simbólicamente mediante la ecuación

$$STC = SCT + SCB + SCE,$$

donde

$$STC = \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2 \quad = \text{suma total de cuadrados,}$$

$$SCT = b \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 \quad = \text{suma de los cuadrados de los tratamientos,}$$

$$SCB = k \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{..})^2 \quad = \text{suma de los cuadrados de los bloques,}$$

$$SCE = \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 = \text{suma de los cuadrados del error.}$$

Si se sigue el procedimiento descrito en el teorema 13.2, donde se interpretó a las sumas de cuadrados como funciones de las variables aleatorias independientes, $Y_{11}, Y_{12}, \dots, Y_{1b}, Y_{21}, \dots, Y_{kb}$, se puede demostrar que los valores esperados de las sumas de los cuadrados de los tratamientos, los bloques y los errores son dados por

$$E(SCT) = (k-1)\sigma^2 + b \sum_{i=1}^k \alpha_i^2, \quad E(SCB) = (b-1)\sigma^2 + k \sum_{j=1}^b \beta_j^2,$$

$$E(SCE) = (b-1)(k-1)\sigma^2.$$

Como en el caso del problema de un solo factor, tenemos que el cuadrado medio del tratamiento es

$$s_1^2 = \frac{SCT}{k-1}.$$

Si los efectos del tratamiento $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, entonces s_1^2 es un estimado insesgado de σ^2 . Sin embargo, si los efectos de los tratamientos no son todos iguales a cero, se tiene que:

Media cuadrada
esperada
del tratamiento

$$E\left(\frac{SCT}{k-1}\right) = \sigma^2 + \frac{b}{k-1} \sum_{i=1}^k \alpha_i^2$$

En este caso s_1^2 sobrestima σ^2 . Un segundo estimado de σ^2 , basado en $b-1$ grados de libertad, es

$$s_2^2 = \frac{SCB}{b-1}.$$

El estimado s_2^2 es un estimado no sesgado de σ^2 si los efectos de los bloques $\beta_1 = \beta_2 = \dots = \beta_b = 0$. Si los efectos de los bloques no son iguales a cero, entonces,

$$E\left(\frac{SCB}{b-1}\right) = \sigma^2 + \frac{k}{b-1} \sum_{j=1}^b \beta_j^2,$$

y s_2^2 sobrestimaré a σ^2 . Un tercer estimado de σ^2 , basado en $(k-1)(b-1)$ grados de libertad e independiente de s_1^2 y s_2^2 , es

$$s^2 = \frac{SCE}{(k-1)(b-1)},$$

que es no sesgado independientemente de la veracidad o falsedad de cualquier hipótesis nula.

Para probar la hipótesis nula de que los efectos de los tratamientos son iguales a cero, se calcula la razón $f_1 = s_1^2/s^2$, que es un valor de la variable aleatoria F_1 , que tiene una distribución F con $k-1$ y $(k-1)(b-1)$ grados de libertad, cuando la hipótesis nula es verdadera. La hipótesis nula se rechaza al nivel de significancia α cuando

$$f_1 > f_\alpha[k-1, (k-1)(b-1)].$$

En la práctica, primero calculamos STC , SCT y SCB , y después, utilizando la identidad de la suma de cuadrados, obtenemos SCE mediante una resta. Los grados de libertad asociados con SCE por lo general también se obtienen por sustracción; es decir,

$$(k-1)(b-1) = kb - 1 - (k-1) - (b-1).$$

Los cálculos necesarios para un problema de análisis de varianza para un diseño de bloques completos aleatorizados se puede resumir como se observa en la tabla 13.8.

Ejemplo 13.6: Se consideran cuatro máquinas diferentes, M_1 , M_2 , M_3 y M_4 , para ensamblar un producto específico. Se decidió que para comparar las máquinas se usarían 6 operadores distintos en un experimento de bloques aleatorizados. Las máquinas se asignaron al azar a cada operador. La operación de las máquinas requiere destreza física, y se anticipó que habría una diferencia en la velocidad con que los operadores trabajaban con las máquinas. En la tabla 13.9 se observan los tiempos (en segundos) requeridos para ensamblar el producto.

A un nivel de significancia de 0.05, pruebe la hipótesis H_0 de que las máquinas se desempeñan con el mismo índice de velocidad promedio.

Tabla 13.8: Análisis de varianza para el diseño de bloques completos aleatorizados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Tratamientos	SCT	$k - 1$	$s_1^2 = \frac{SCT}{k - 1}$	$f_1 = \frac{s_1^2}{s^2}$
Bloques	SCB	$b - 1$	$s_2^2 = \frac{SCB}{b - 1}$	
Error	SCE	$(k - 1)(b - 1)$	$s^2 = \frac{SCE}{(k - 1)(b - 1)}$	
Total	STC	$kb - 1$		

Tabla 13.9: Tiempo para ensamblar el producto, en segundos

Máquina	Operador						Total
	1	2	3	4	5	6	
1	42.5	39.3	39.6	39.9	42.9	43.6	247.8
2	39.8	40.1	40.5	42.3	42.5	43.1	248.3
3	40.2	40.5	41.3	43.4	44.9	45.1	255.4
4	41.3	42.2	43.5	44.2	45.9	42.3	259.4
Total	163.8	162.1	164.9	169.8	176.2	174.1	1010.9

Solución: Las hipótesis son

H_0 : $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ (los efectos de las máquinas son iguales a cero),

H_1 : Al menos una de las α_j no es igual a cero.

Para producir el análisis de varianza que aparece en la tabla 13.10 se emplean las fórmulas de la suma de cuadrados que se presentan en la página 536 y los grados de libertad. El valor $f = 3.34$ es significativo con $P = 0.048$. Si se emplea $\alpha = 0.05$ como al menos una aproximación burda, se concluye que las máquinas no se desempeñan con el mismo índice de velocidad media. J

Tabla 13.10: Análisis de varianza para los datos de la tabla 13.9

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Máquinas	15.93	3	5.31	3.34
Operadores	42.09	5	8.42	
Error	23.84	15	1.59	
Total	81.86	23		

Comentarios adicionales acerca de la formación de bloques

En el capítulo 10 presentamos un procedimiento para comparar medias cuando las observaciones estaban *ordenadas por pares*. El procedimiento implicaba “restar” el efecto

debido a la paridad homogénea para así trabajar con las diferencias. Este es un caso especial de diseño de bloques completos aleatorizados con $k = 2$ tratamientos. Las n unidades homogéneas a las que se asignaron los tratamientos adoptan el papel de bloques.

Si hay heterogeneidad en las unidades experimentales, el experimentador no debería confundirse y pensar que siempre es ventajoso reducir el error experimental mediante el uso de pequeños bloques homogéneos. De hecho podría haber casos en los que no es deseable formar bloques. El propósito de reducir la varianza del error es incrementar la *sensibilidad* de la prueba para detectar diferencias en las medias de los tratamientos. Esto se refleja en la potencia del procedimiento de prueba. (En la sección 13.11 se analiza con mayor detalle la potencia del procedimiento de prueba del análisis de varianza). La potencia para detectar ciertas diferencias entre las medias de los tratamientos se incrementa con una disminución de la varianza del error. Sin embargo, la potencia también se ve afectada por los grados de libertad con los que se estima la varianza, y la formación de bloques reduce los grados de libertad que están disponibles desde $k(b-1)$ para la clasificación de un solo factor, hasta $(k-1)(b-1)$. De modo que se podría perder potencia con la formación de bloques si no hay una reducción significativa de la varianza del error.

Interacción entre bloques y tratamientos

Otra suposición importante que está implícita en la escritura del modelo para un diseño de bloques completos aleatorizados es que los efectos de los bloques y del tratamiento son aditivos. Esto equivale a decir que

$$\mu_{ij} - \mu_{ij'} = \mu_{i'j} - \mu_{i'j'} \quad \text{o bien} \quad \mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'},$$

para cada valor de i, i', j y j' . Es decir, la diferencia entre las medias de la población para los bloques j y j' es la misma para cada tratamiento, y la diferencia entre las medias de la población para los tratamientos i e i' es la misma para cada bloque. Las líneas paralelas de la figura 13.6a ilustran un conjunto de respuestas medias para las cuales los efectos de los tratamientos y los bloques son aditivos, mientras que las líneas que se intersecan en la figura 13.6b exhiben una situación en la que se dice que los efectos de los tratamientos y de los bloques **interactúan**. Con respecto al ejemplo 13.6, si el operador 3 es en promedio 0.5 segundos más rápido que el operador 2 cuando utiliza la máquina 1, entonces el operador 3 será 0.5 segundos más rápido, en promedio, que el operador 2 cuando se empleen las máquinas 2, 3 o 4. En muchos experimentos no se cumple la suposición de aditividad y el análisis descrito en esta sección llevaría a conclusiones erróneas. Por ejemplo, suponga que el operador 3 es 0.5 segundos más rápido, en promedio, que el operador 2 si emplea la máquina 1, pero que es 0.2 segundos más lento, en promedio, que el operador 2 si utiliza la máquina 2. En ese caso los operadores y las máquinas estarían interactuando.

Una inspección de la tabla 13.9 sugiere la posible presencia de interacción. Esta aparente interacción podría ser real o podría deberse al error experimental. El análisis del ejemplo 13.6 se basó en la suposición de que la aparente interacción se debe por completo al error experimental. Si la variabilidad total de nuestros datos se debiera en parte al efecto de la interacción, esa fuente de variación seguiría formando parte de la suma de cuadrados del error, **provocando que el cuadrado medio del error sobrestime a σ^2** , incrementando así la probabilidad de cometer un error tipo II. De hecho, hemos supuesto un modelo incorrecto. Si permitimos que $(\alpha\beta)_{ij}$ denote el efecto de la interacción del

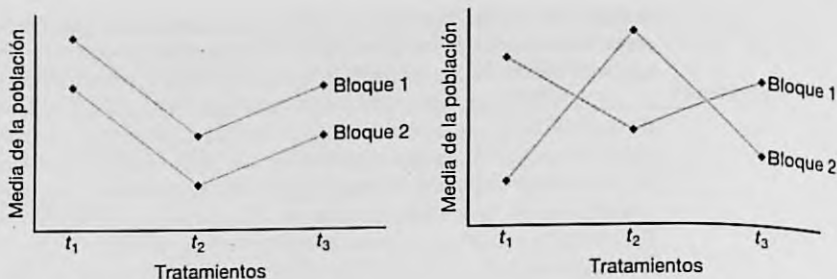


Figura 13.6: Medias de la población para a) resultados aditivos y b) efectos de interacción.

i -ésimo tratamiento y el j -ésimo bloque, podríamos escribir un modelo más adecuado con la forma siguiente:

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij},$$

al que se impondrían las restricciones adicionales

$$\sum_{i=1}^k (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = 0, \text{ para } i = 1, \dots, k \text{ y } j = 1, \dots, b.$$

Ahora es fácil comprobar que

$$E \left[\frac{SCE}{(b-1)(k-1)} \right] = \sigma^2 + \frac{1}{(b-1)(k-1)} \sum_{i=1}^k \sum_{j=1}^b (\alpha\beta)_{ij}^2.$$

Así, el cuadrado medio del error es considerado un **estimado sesgado de σ^2** cuando se ha ignorado la interacción existente. En este momento parecería necesario utilizar un procedimiento para detectar la interacción en aquellos casos en que se sospecha que exista. Tal procedimiento requiere que se disponga de un estimado no sesgado e independiente de σ^2 . Por desgracia, el diseño de bloques aleatorizados no se presta a una prueba de este tipo, a menos que se modifique el diseño inicial del experimento. En el capítulo 14 se estudia este tema de manera detallada.

13.9 Métodos gráficos y verificación del modelo

En varios capítulos de este libro se hace referencia a procedimientos gráficos para mostrar datos y resultados analíticos. En los primeros capítulos se usaron gráficas de tallo y hojas y de caja y extensión como auxiliares visuales para resumir muestras. En el capítulo 10 se emplearon diagnósticos similares para entender mejor los datos de dos problemas de muestreo. En el capítulo 11 se introdujo el concepto de gráfica de residuales para detectar violaciones de las suposiciones estándar. En los últimos años gran parte de la atención dedicada al análisis de datos se ha centrado en los **métodos gráficos**. Al igual que en la regresión, el análisis de varianza se presta a la elaboración de gráficas que

ayudan a resumir los datos y a detectar violaciones. Por ejemplo, una gráfica sencilla de las observaciones brutas alrededor de la media de cada tratamiento proporciona al analista una noción de la variabilidad entre las medias muestrales y dentro de las muestras. La figura 13.7 ilustra una de tales gráficas para los datos de agregados que se presentan en la tabla 13.1. A partir de la apariencia de la gráfica se obtiene incluso una idea de cuáles agregados (si los hubiera) destacan de los demás. Es evidente que el agregado 4 resalta del resto, y que los agregados 3 y 5 forman un grupo homogéneo, así como los agregados 1 y 2.

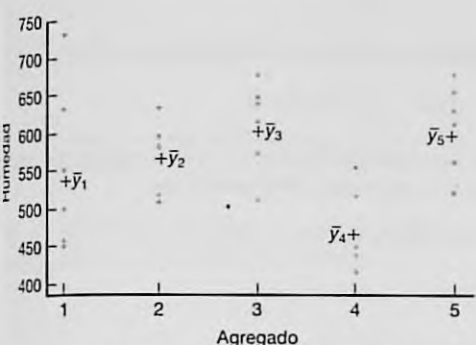


Figura 13.7: Gráfica de los datos alrededor de la media para los datos de los agregados de la tabla 13.1.

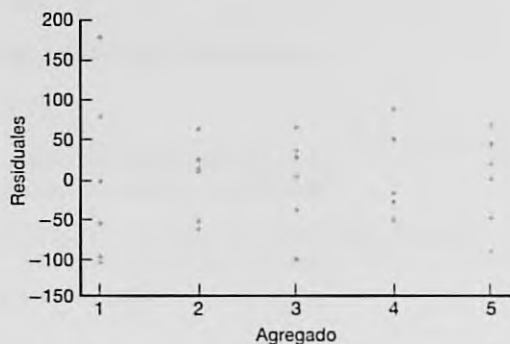


Figura 13.8: Gráfica de los residuales para cinco agregados con los datos de la tabla 13.1.

Como ocurre en el caso de la regresión, los residuales son útiles en el análisis de varianza para dar un diagnóstico sobre la detección de violaciones de los supuestos. Para formar los residuales sólo necesitamos considerar el modelo del problema de un solo factor, que es

$$y_{ij} = \mu_i + \epsilon_{ij}.$$

Es fácil determinar que el estimado de μ_i es \bar{y}_i . Por lo tanto, el ij -ésimo residual es $\bar{y}_i - y_{ij}$, lo cual se extiende fácilmente al modelo de bloques completos aleatorizados. Sería aconsejador graficar los residuos para cada agregado con el fin de obtener cierta información sobre la suposición de varianzas homogéneas. Esta gráfica se muestra en la figura 13.8.

Las tendencias en gráficas como éstas podrían revelar dificultades en ciertas situaciones, especialmente cuando la violación de una suposición en particular se manifiesta en la gráfica. En el caso de la figura 13.8, los residuales parecen indicar que las varianzas *dentro de los tratamientos* son razonablemente homogéneas, excepto la del agregado 1. Hay cierta evidencia gráfica de que la varianza del agregado 1 es más grande que la del resto.

¿Qué es un residual para un diseño de BCA?

El diseño de bloques completos aleatorizados es otra situación experimental en la cual una gráfica permite que el analista se sienta cómodo con una "imagen ideal" o que tal

vez detecte dificultades. Recuerde que el modelo para el diseño de bloques completos aleatorizados es

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, b,$$

con las restricciones impuestas

$$\sum_{i=1}^k \alpha_i = 0, \quad \sum_{j=1}^b \beta_j = 0.$$

Para determinar qué es lo que en realidad constituye un residual considere que

$$\alpha_i = \mu_i - \mu, \quad \beta_j = \mu_j - \mu$$

y que μ se estima por medio de $\bar{y}_{..}$, μ_i se estima por medio de $\bar{y}_{i.}$ y μ_j se estima por medio de $\bar{y}_{.j}$. Como resultado, el *valor ajustado* o pronosticado \bar{y}_{ij} es dado por

$$\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j = \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}$$

y, entonces, el residual en la observación (i, j) es dado por

$$y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$$

Observe que \bar{y}_{ij} , el valor ajustado, es un estimado de la media μ_{ij} . Esto es congruente con la partición de la variabilidad dada en el teorema 13.3, en la que la suma de los cuadrados del error es

$$SCE = \sum_i^k \sum_j^b (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2.$$

Las técnicas visuales para el diseño de bloques completos aleatorizados requieren que se grafiquen los residuos por separado para cada tratamiento y cada bloque. Si la suposición de varianza homogénea se cumple, el analista debería esperar una variabilidad aproximadamente igual. El lector seguramente recordará que en el capítulo 12 se estudiaron gráficas de los residuales con el objetivo de detectar si el modelo era inadecuado. En el caso del diseño de bloques completos aleatorizados, una grave falla del modelo podría estar relacionada con la suposición de aditividad (lo cual significa que no hay interacción). Si no hay interacción debe surgir un patrón aleatorio.

Considere los datos del ejemplo 13.6, donde los tratamientos son cuatro máquinas y los bloques son seis operadores. Las figuras 13.9 y 13.10 incluyen las gráficas de los residuales para tratamientos separados y bloques separados. La figura 13.11 presenta una gráfica de los residuales contra los valores ajustados. La figura 13.9 revela que quizá la varianza del error no sea la misma para todas las máquinas, y lo mismo podría ocurrir con la varianza del error para cada uno de los seis operadores. Sin embargo, al parecer dos residuales inusualmente grandes son los que provocan la aparente dificultad. La figura 13.11 es una gráfica de residuales que revela evidencia razonable de un comportamiento aleatorio. Sin embargo, sobresalen los dos residuales grandes ya detectados.

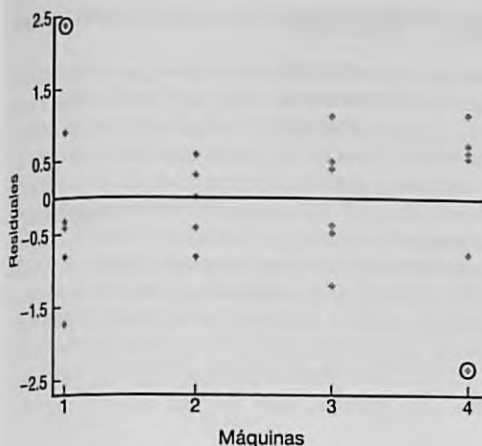


Figura 13.9: Gráfica de residuos para las cuatro máquinas de los datos del ejemplo 13.6.

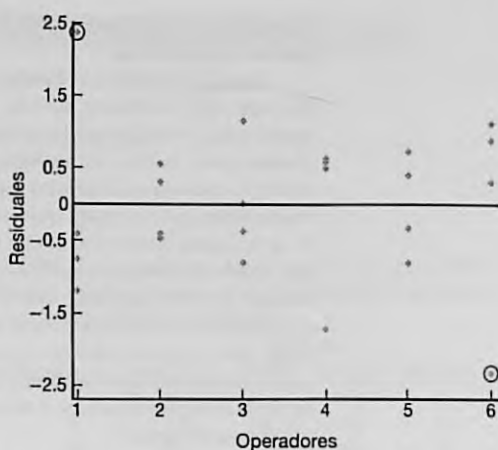


Figura 13.10: Gráfica de residuos para los seis operadores de los datos del ejemplo 13.6.

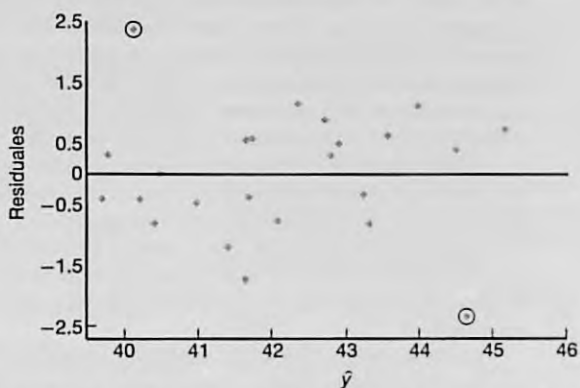


Figura 13.11: Residuales graficados contra los valores ajustados para los datos del ejemplo 13.6.

13.10 Transformaciones de datos en el análisis de varianza

En el capítulo 11 se puso mucha atención a la transformación de la respuesta y en situaciones para las que se ajustaba un modelo de regresión lineal a un conjunto de datos. Es evidente que se aplican los mismos conceptos a la regresión lineal múltiple, aunque esto no se analizó en el capítulo 12. En el estudio de los modelos de regresión se hizo énfasis en las transformaciones de y que producirían un modelo que se ajustara mejor a los datos que uno en el que la y ingresara de forma lineal. Por ejemplo, si la estructura del "tiempo" es de naturaleza exponencial, entonces una transformación logarítmica de

y linealiza la estructura y, por lo tanto, se anticipa más éxito cuando se utiliza la respuesta transformada.

Aunque el propósito fundamental de la transformación de datos que se ha analizado hasta este momento ha sido mejorar el ajuste del modelo, hay otras razones para transformar o reexpresar la respuesta y , y muchas de ellas se relacionan con las suposiciones que se hacen, por ejemplo, las suposiciones de las cuales depende la validez del análisis. Una suposición muy importante en el análisis de varianza es la de la varianza homogénea que se estudió antes en la sección 13.4. Se supone una **varianza común** σ^2 . Si la varianza difiere mucho de un tratamiento a otro, y se realiza el ANOVA estándar que se estudia en este capítulo (y en otros posteriores), los resultados serían muy deficientes. En otras palabras, el análisis de varianza no es **robusto** respecto a la suposición de varianza homogénea. Como se ha dicho hasta el momento, se trata del motivo principal para la graficación de los residuales que estudiamos en la sección anterior y que ilustramos en las figuras 13.9, 13.10 y 13.11. Esas gráficas permiten detectar problemas debidos a una varianza no homogénea. Sin embargo, ¿qué hay que hacer al respecto? ¿Cómo se corrigen?

¿De dónde proviene la varianza no homogénea?

Con frecuencia, aunque no siempre, la varianza no homogénea en el ANOVA existe debido a la distribución de las respuestas. Ahora, por supuesto, se supone la normalidad de la respuesta, pero hay ciertas situaciones en las que se necesitan pruebas de las medias aunque la distribución de la respuesta sea una de las distribuciones no normales que se estudiaron en los capítulos 5 y 6, es decir, la distribución de Poisson, la logarítmica normal, la exponencial y la gamma. Realmente existen problemas del tipo del ANOVA con datos de conteo, duración antes de la falla, etcétera.

En los capítulos 5 y 6 se demostró que, además del caso de la normal, la varianza de una distribución con frecuencia será función de la media, es decir, $\sigma_i^2 = g(\mu_i)$. Por ejemplo, en el caso de la distribución de Poisson, $\text{Var}(Y_i) = \mu_i = \sigma_i^2$, lo que significa que la varianza es igual a la media. En el caso de la distribución exponencial, $\text{Var}(Y_i) = \sigma_i^2 = \mu_i^2$, o sea que la varianza es igual al cuadrado de la media. Para el caso de la logarítmica normal, una transformación logarítmica produce una distribución normal con varianza constante σ^2 .

Los mismos conceptos que usamos en el capítulo 4 para determinar la varianza de una función no lineal pueden ayudarnos a determinar la naturaleza de la *transformación estabilizadora de la varianza* $g(y)$. Recuerde la expansión de las series de Taylor de primer orden de $g(y)$ alrededor de $y_i = \mu_i$, donde $g'(\mu_i) = \left[\frac{\partial g(y_i)}{\partial y_i} \right]_{y_i = \mu_i}$. La función de transformación $g(y)$ debe ser independiente de μ para que baste como la transformación estabilizadora de la varianza. De lo anterior

$$\text{Var}[g(y_i)] \approx [g'(\mu_i)]^2 \sigma_i^2.$$

Como resultado, $g(y)$ debe ser tal que $g'(\mu_i) \propto \frac{1}{\sigma}$. Así, si se sospecha que la respuesta tiene una distribución de Poisson, $\sigma_i = \mu_i^{1/2}$, de modo que $g'(\mu_i) \propto \frac{1}{\mu_i^{1/2}}$. Entonces, la transformación estabilizadora de la varianza es $g(y_i) = y_i^{1/2}$. A partir de esta ilustración y manipulaciones similares para las distribuciones exponencial y gamma, se obtiene lo siguiente.

Distribución	Transformaciones estabilizadoras de la varianza
Poisson	$g(y) = y^{1/2}$
Exponencial	$g(y) = \ln y$
Gamma	$g(y) = \ln y$

Ejercicios

13.25 Se utilizaron cuatro clases de fertilizante f_1, f_2, f_3, f_4 para estudiar la cosecha de frijol. El suelo se dividió en 3 bloques, cada uno de los cuales contiene 4 parcelas homogéneas. A continuación se presentan las cosechas en kilogramos por parcela, así como los tratamientos correspondientes:

Bloque 1	Bloque 2	Bloque 3
$f_1 = 42.7$	$f_3 = 50.9$	$f_4 = 51.1$
$f_3 = 48.5$	$f_1 = 50.0$	$f_2 = 46.3$
$f_4 = 32.8$	$f_2 = 38.0$	$f_1 = 51.9$
$f_2 = 39.3$	$f_4 = 40.2$	$f_3 = 53.5$

Realice un análisis de varianza a un nivel de significancia de 0.05 utilizando el modelo de bloques completos aleatorizados.

13.26 Se compararon las cosechas de tres variedades de papas. El experimento se efectuó asignando cada variedad de manera aleatoria a 3 parcelas del mismo tamaño, en 4 lugares diferentes. Se registraron las siguientes cosechas para las variedades A, B y C, en 100 kilogramos por parcela:

Lugar 1	Lugar 2	Lugar 3	Lugar 4
B: 13	C: 21	C: 9	A: 11
A: 18	A: 20	B: 12	C: 10
C: 12	B: 23	A: 14	B: 17

Realice un análisis de varianza de bloques completos aleatorizados con el objetivo de probar la hipótesis de que no hay diferencia en la capacidad de rendimiento de las 3 variedades de papas. Utilice un nivel de significancia de 0.05 y saque conclusiones.

13.27 Los siguientes datos son los porcentajes de aditivos extranjeros, medidos por 5 analistas, de 3 marcas distintas de mermelada de fresa, A, B y C.

Analista 1	Analista 2	Analista 3	Analista 4	Analista 5
B: 2.7	C: 7.5	B: 2.8	A: 1.7	C: 8.1
C: 3.6	A: 1.6	A: 2.7	B: 1.9	A: 2.0
A: 3.8	B: 5.2	C: 6.4	C: 2.6	B: 4.8

A un nivel de significancia de 0.05, realice un análisis de varianza de bloques completos aleatorizados para probar la hipótesis de que el porcentaje de aditivos extranjeros es el mismo para las tres marcas de mermelada. ¿Cuál de ellas parece tener menos aditivos?

13.28 Los siguientes datos representan las calificaciones finales obtenidas por 5 estudiantes en matemáticas, inglés, francés y biología:

Estudiante	Materia			
	Matemáticas	Inglés	Francés	Biología
1	68	57	73	61
2	83	94	91	86
3	72	81	63	59
4	55	73	77	66
5	92	68	75	87

Pruebe la hipótesis de que los cursos tienen la misma dificultad. Use un valor P en sus conclusiones y analice sus hallazgos.

13.29 En el estudio *The Periphyton of the South River, Virginia: Mercury Concentration, Productivity, and Autotrophic Index Studies*, efectuado por el Departamento de Ciencias e Ingeniería Ambientales de Virginia Tech, se midió la concentración total de mercurio en sólidos totales de perifitón en seis estaciones distintas durante seis días diferentes. Determine si el contenido medio de mercurio difiere significativamente entre las estaciones utilizando los siguientes datos. Use un valor P y analice sus hallazgos.

Fecha	Estación					
	CA	CB	E1	E2	E3	E4
8 de abril	0.45	3.24	1.33	2.04	3.93	5.93
23 de junio	0.10	0.10	0.99	4.31	9.92	6.49
1 de julio	0.25	0.25	1.65	3.13	7.39	4.43
8 de julio	0.09	0.06	0.92	3.66	7.88	6.24
15 de julio	0.15	0.16	2.17	3.50	8.82	5.39
23 de julio	0.17	0.39	4.30	2.91	5.50	4.29

13.30 Una planta de energía nuclear produce una gran cantidad de calor que generalmente se descarga en los sistemas de agua. Ese calor eleva la temperatura del líquido, lo cual da como resultado una mayor concentración de clorofila a que, a su vez, alarga la temporada de crecimiento. Para estudiar este efecto se tomaron muestras de agua mensualmente en 3 estaciones, durante un periodo de 12 meses. La estación A es la que se ubica más cerca de una descarga potencial de agua caliente, la estación C es la más lejana de la descarga y la estación B se encuentra entre las estaciones A y C. Se registraron las siguientes concentraciones de clorofila a .

Mes	Estación		
	A	B	C
Enero	9.867	3.723	4.410
Febrero	14.035	8.416	11.100
Marzo	10.700	20.723	4.470
Abril	13.853	9.168	8.010
Mayo	7.067	4.778	34.080
Junio	11.670	9.145	8.990
Julio	7.357	8.463	3.350
Agosto	3.358	4.086	4.500
Septiembre	4.210	4.233	6.830
Octubre	3.630	2.320	5.800
Noviembre	2.953	3.843	3.480
Diciembre	2.640	3.610	3.020

Realice un análisis de varianza y, a un nivel de significancia de 0.05, pruebe la hipótesis de que no hay diferencia en las concentraciones medias de clorofila *a* en las 3 estaciones.

13.31 En un estudio realizado por el Departamento de Salud y Educación Física de Virginia Tech, se asignaron 3 dietas durante 3 días a 6 sujetos utilizando un diseño de bloques completos aleatorizados. Los sujetos, que desempeñan el papel de bloques, recibieron las siguientes 3 dietas en orden aleatorio:

- Dieta 1: grasas mixtas y carbohidratos,
- Dieta 2: alta en grasas,
- Dieta 3: alta en carbohidratos.

Al terminar el periodo de tres días se puso a cada sujeto en una banda caminadora y se midió el tiempo, en segundos, que transcurría hasta que se sentían exhaustos. Efectúe un análisis de varianza separando la dieta, los sujetos y la suma de cuadrados del error. Utilice un valor *P* para determinar si existen diferencias significativas entre las dietas. Los datos registrados son los siguientes:

Dieta	Sujeto					
	1	2	3	4	5	6
1	84	35	91	57	56	45
2	91	48	71	45	61	61
3	122	53	110	71	91	122

13.32 El personal forestal utiliza arsénico orgánico como arboricida. La cantidad de arsénico que absorbe el cuerpo cuando se expone a este producto constituye un grave problema de salud. Es importante que la cantidad de exposición se determine rápido, de manera que pueda retirarse del trabajo a los empleados con niveles elevados de arsénico. En un experimento descrito en el artículo "A Rapid Method for the Determination of Arsenic Concentrations in Urine at Field Locations", publicado en el *American Industrial Hygiene Association Journal* (Vol. 37, 1976), especímenes de orina de 4 personas del servicio forestal fueron divididos por igual en

tres muestras para que pudiera analizarse el contenido de arsénico en la orina de cada individuo en un laboratorio universitario: las muestras eran analizadas por un químico con un sistema portátil, así como también por un empleado forestal que había recibido una capacitación breve. Se registraron los siguientes niveles de arsénico, en partes por millón:

Individuo	Analista		
	Empleado	Químico	Laboratorio
1	0.05	0.05	0.04
2	0.05	0.05	0.04
3	0.04	0.04	0.03
4	0.15	0.17	0.10

Realice un análisis de varianza y, a un nivel de significancia de 0.05, pruebe la hipótesis de que no hay diferencia en los niveles de arsénico con los tres métodos de análisis.

13.33 Los científicos del Departamento de Patología Vegetal de Virginia Tech realizaron un experimento en el que se aplicaron 5 tratamientos diferentes en 6 lugares distintos de un huerto de manzanas para determinar si había diferencias significativas en el crecimiento entre los tratamientos. Los tratamientos 1 a 4 representan distintos herbicidas y el 5 es un control. El periodo de crecimiento fue de mayo a noviembre de 1982, y los datos de crecimiento nuevo, medido en centímetros, para muestras seleccionadas de los 6 lugares en el huerto, son los siguientes:

Tratamiento	Ubicaciones					
	1	2	3	4	5	6
1	455	72	61	215	695	501
2	622	82	444	170	437	134
3	695	56	50	443	701	373
4	607	650	493	257	490	262
5	388	263	185	103	518	622

Lleve a cabo un análisis de varianza, separando el tratamiento, el lugar y la suma de cuadrados del error. Determine si hay diferencias significativas entre las medias de los tratamientos. Mencione un valor *P*.

13.34 En el artículo "Self-Control and Therapist Control in the Behavioral Treatment of Overweight Women", publicado en *Behavioral Research and Therapy* (Vol. 10, 1972), se estudiaron dos tratamientos de reducción y otro de control para observar sus efectos en el cambio del peso en mujeres obesas. Los dos tratamientos reductores involucrados fueron un programa autodirigido de reducción de peso y otro controlado por un terapeuta. Se asignó a cada uno de 10 sujetos a uno de los 3 programas de tratamiento en orden aleatorio y se midió la pérdida de peso. Se registraron los siguientes cambios en el peso:

Sujeto	Tratamiento		
	Control	Autodirigido	Con terapeuta
1	1.00	-2.25	-10.50
2	3.75	-6.00	-13.50
3	0.00	-2.00	0.75
4	-0.25	-1.50	-4.50
5	-2.25	-3.25	-6.00
6	-1.00	-1.50	4.00
7	-1.00	-10.75	-12.25
8	3.75	-0.75	-2.75
9	1.50	0.00	-6.75
10	0.50	-3.75	-7.00

Realice un análisis de varianza y, a un nivel de significancia de 0.01, pruebe la hipótesis de que no hay diferencia en las pérdidas de peso promedio para los 3 tratamientos. ¿Cuál tratamiento fue el mejor?

13.35 En el libro *Design of Experiments for the Quality Improvement*, publicado por la Japanese Standards Association (1989) se reportó un estudio sobre la cantidad de tinta que se requiere para obtener el mejor color para cierto tipo de tela. En dos plantas diferentes se administraron tres cantidades de tinta: $\frac{1}{3}$ del porcentaje de wof, es decir, $\frac{1}{3}$ del porcentaje del peso de la tela, 1% de wof y 3% de wof. Después se observó cuatro veces la densidad del color de la tela para cada nivel de tinta aplicada en cada planta.

	Cantidad de tinta					
	1/3%		1%		3%	
Planta 1	5.2	6.0	12.3	10.5	22.4	17.8
	5.9	5.9	12.4	10.9	22.5	18.4
Planta 2	6.5	5.5	14.5	11.8	29.0	23.2
	6.4	5.9	16.0	13.6	29.7	24.0

A un nivel de significancia de 0.05, realice un análisis de varianza para probar la hipótesis de que no hay dife-

rencia en la densidad de color de la tela con los tres niveles de tinta. Considere a las plantas como bloques.

13.36 Se realizó un experimento con el fin de comparar tres tipos de materiales para recubrir alambres de cobre. El propósito del recubrimiento consiste en eliminar los "defectos" del alambre. A cada recubrimiento se le asignaron al azar 10 especímenes distintos, de 5 milímetros de longitud, para que los fuera aplicado. Después se sometió a los 30 especímenes a cierto tipo de desgaste abrasivo. Al final se midió el número de defectos en cada uno y se obtuvieron los siguientes resultados:

	Material										
	1			2			3				
6	8	4	5	3	3	5	4	12	8	7	14
7	7	9	6	2	4	4	5	18	6	7	18
7	8			4	3			8	5		

Suponga que se acepta que se puede aplicar un proceso de Poisson, por lo que el modelo es $Y_{ij} = \mu_i + \epsilon_{ij}$, donde μ_i es la media de la distribución de Poisson y $\sigma_{ij}^2 = \mu_i$.

- Haga una transformación apropiada de los datos y un análisis de varianza.
- Determine si hay evidencia suficiente para preferir un material de recubrimiento sobre los demás. Muestre cualesquiera hallazgos que sugieran una conclusión.
- Haga una gráfica de residuales y coméntela.
- Mencione el propósito de la transformación de los datos.
- ¿Qué otra suposición se hace en este caso, que quizá la transformación no cumpla por completo?
- Comente el inciso e después de elaborar una gráfica de probabilidad normal sobre los residuales.

13.11 Modelos de efectos aleatorios

A lo largo de este capítulo estudiamos los procedimientos del análisis de varianza en los que el objetivo principal es estudiar el efecto sobre ciertas respuestas de ciertos tratamientos fijos o predeterminados. Los experimentos en los que los tratamientos o los niveles de tratamiento son preseleccionados por el experimentador, y no elegidos al azar, se denominan **experimentos de efectos fijos**. Para el modelo de efectos fijos sólo se hacen inferencias acerca de los tratamientos específicos que se utilizaron en el experimento.

Con frecuencia es importante que el experimentador sea capaz de hacer inferencias acerca de una población de tratamientos a través de un experimento en el que los tratamientos empleados se elijan al azar de entre la población. Por ejemplo, un biólogo podría estar interesado en saber si hay o no una varianza significativa en alguna característica fisiológica debida a un tipo de animal. Los tipos de animales que en realidad se usan en el experimento se eligen al azar y representan los efectos del tratamiento. Un químico podría estar interesado en estudiar el efecto de los laboratorios sobre el análisis químico de

una sustancia; no le interesa un laboratorio en particular, sino una población grande de laboratorios. Así, podría seleccionar al azar un grupo de laboratorios y asignar muestras a cada uno para su análisis. Entonces, la inferencia estadística implicaría 1) probar si los laboratorios contribuyen o no a una varianza diferente de cero en los resultados de los análisis, y 2) estimar la varianza debida a los laboratorios y a la varianza dentro de los mismos.

Modelo y suposiciones para el modelo de efectos aleatorios

El **modelo de efectos aleatorios** de un solo factor se escribe como el modelo de efectos fijos, pero sus términos tienen significados diferentes. La respuesta $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ es ahora un valor de la variable aleatoria

$$Y_{ij} = \mu + A_i + \epsilon_{ij}, \text{ con } i = 1, 2, \dots, k \text{ y } j = 1, 2, \dots, n,$$

donde las A_i tienen distribución normal e independiente con media igual a cero y varianza σ_α^2 , y son independientes de las ϵ_{ij} . Al igual que para el modelo de efectos fijos, las ϵ_{ij} también tienen distribución normal e independiente con media igual a cero y varianza σ^2 . Observe que para un experimento de efectos aleatorios, ya no se aplica la restricción de que $\sum_{i=1}^k \alpha_i = 0$.

Teorema 13.4: Para el modelo del análisis de varianza de efectos aleatorios de un solo factor.

$$E(SCT) = (k-1)\sigma^2 + n(k-1)\sigma_\alpha^2 \quad \text{y} \quad E(SCE) = k(n-1)\sigma^2.$$

La tabla 13.11 presenta los cuadrados medios esperados tanto para un experimento de efectos fijos como para uno de efectos aleatorios. Los cálculos para un experimento de efectos aleatorios se realizan exactamente de la misma forma que los del experimento de efectos fijos. Es decir, la suma de cuadrados, los grados de libertad y las columnas de los cuadrados medios en la tabla del análisis de varianza son iguales para ambos modelos.

Tabla 13.11: Cuadrados medios esperados para el experimento de un solo factor

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrados medios esperados	
			Efectos fijos	Efectos aleatorios
Tratamientos	$k-1$	s_1^2	$\sigma^2 + \frac{n}{k-1} \sum_i \alpha_i^2$	$\sigma^2 + n\sigma_\alpha^2$
Error	$k(n-1)$	s^2	σ^2	σ^2
Total	$nk-1$			

Para el modelo de efectos aleatorios la hipótesis de que todos los efectos del tratamiento son iguales a cero se escribe como sigue:

Hipótesis para un experimento de efectos aleatorios

$$H_0: \sigma_\alpha^2 = 0,$$

$$H_1: \sigma_\alpha^2 \neq 0.$$

Esta hipótesis afirma que los diferentes tratamientos no contribuyen en absoluto a la variabilidad de la respuesta. De la tabla 13.11 es evidente que tanto s_1^2 como s^2 son estimados de σ^2 cuando H_0 es verdadera, y que la razón

$$f = \frac{s_1^2}{s^2}$$

es un valor de la variable aleatoria F que tiene la distribución F con $k-1$ y $k(n-1)$ grados de libertad. La hipótesis nula se rechaza a un nivel de significancia α cuando

$$f > f_\alpha[k-1, k(n-1)].$$

En muchos estudios científicos y de ingeniería el interés no se centra en la prueba F . El científico sabe que el efecto aleatorio, en efecto, es significativo. Lo más importante es la estimación de los diversos componentes de la varianza. Esto produce un sentido de *jerarquía* en términos de cuáles factores producen la mayor variabilidad y en qué cantidad. En este contexto podría ser interesante cuantificar cuánto más grande es el *componente de la varianza de un solo factor* que el producido por el azar (variación aleatoria).

Estimación de los componentes de la varianza

La tabla 13.11 también se utiliza para estimar los **componentes de la varianza** σ^2 y σ_α^2 . Como s_1^2 estima $\sigma^2 + n\sigma_\alpha^2$ y s^2 estima σ^2 ,

$$\hat{\sigma}^2 = s^2, \quad \hat{\sigma}_\alpha^2 = \frac{s_1^2 - s^2}{n}.$$

Ejemplo 13.7: Los datos de la tabla 13.12 representan observaciones codificadas sobre el producto de un proceso químico en el que se utilizan 5 lotes de materia prima seleccionados al azar. Demuestre que el componente de la varianza del lote es significativamente mayor que cero y obtenga su estimado.

Tabla 13.12: Datos para el ejemplo 13.7

Lote	1	2	3	4	5	
	9.7	10.4	15.9	8.6	9.7	
	5.6	9.6	14.4	11.1	12.8	
	8.4	7.3	8.3	10.7	8.7	
	7.9	6.8	12.8	7.6	13.4	
	8.2	8.8	7.9	6.4	8.3	
	7.7	9.2	11.6	5.9	11.7	
	8.1	7.6	9.8	8.1	10.7	
Total	55.6	59.7	80.7	58.4	75.3	329.7

Solución: La suma total de cuadrados, la del lote y la suma de cuadrados del error son, respectivamente,

$$STC = 194.64, \quad SCT = 72.60 \text{ y } SCE = 194.64 - 72.60 = 122.04.$$

En la tabla 13.13 se presentan estos resultados con el resto de los cálculos.

Tabla 13.13: Análisis de la varianza para el ejemplo 13.7

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Lotes	72.60	4	18.15	4.46
Error	122.04	30	4.07	
Total	194.64	34		

La razón f es significativa al nivel $\alpha = 0.05$, lo que indica que se rechaza la hipótesis de un componente del lote igual a cero. Una estimación del componente de la varianza del lote es

$$\hat{\sigma}_\alpha^2 = \frac{18.15 - 4.07}{7} = 2.01.$$

Observe que mientras que el **componente de la varianza del lote** es significativamente diferente de cero, cuando se compara contra el estimado de σ^2 , es decir, $\hat{\sigma}^2 = CME = 4.07$, pareciera que el componente de varianza del lote no es considerablemente grande. ▮

Si el resultado que se obtiene con la fórmula para $\hat{\sigma}_\alpha^2$ es negativo, o sea, que s_i^2 es menor que s^2 , entonces a $\hat{\sigma}_\alpha^2$ se le asigna un valor de cero. Éste es un estimador sesgado. Para tener un mejor estimador de σ_α^2 , suele emplearse un método llamado **verosimilitud restringida** (o **residual**) **máxima** (**REML** por sus siglas en inglés) (véase Harville, 1977, en la bibliografía). Este tipo de estimador se puede encontrar en muchos paquetes estadísticos para computadora. Los detalles de dicho procedimiento rebasan el alcance de este libro.

Diseño de bloques aleatorizados con bloques aleatorios

En un experimento de bloques completos aleatorizados, donde los bloques representan días, es concebible que el experimentador quiera que los resultados se apliquen no sólo a los días reales utilizados en el análisis, sino a cada día del año. Entonces, seleccionaría al azar los días en que se haría el experimento, así como los tratamientos y el modelo de efectos aleatorios a utilizar.

$$Y_{ij} = \mu + A_i + B_j + \epsilon_{ij}, \text{ para } i = 1, 2, \dots, k \text{ y } j = 1, 2, \dots, b,$$

donde las A_i , B_j y ϵ_{ij} son variables aleatorias independientes con medias igual a cero y varianzas σ_α^2 , σ_β^2 y σ^2 , respectivamente. Se obtienen los cuadrados medios esperados para un diseño de bloques completos aleatorizados de efectos aleatorios usando el mismo procedimiento que se usó en el problema de un solo factor; en la tabla 13.14 se presentan junto con los de un experimento de efectos fijos.

Nuevamente, los cálculos para las sumas de cuadrados y grados de libertad individuales son idénticos a los del modelo de efectos fijos. Las hipótesis

$$H_0: \sigma_\alpha^2 = 0,$$

$$H_1: \sigma_\alpha^2 \neq 0$$

se obtienen calculando

$$f = \frac{s_1^2}{s^2}$$

Tabla 13.14: Cuadrados medios esperados para un diseño de bloques completos aleatorizados

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrados medios esperados	
			Efectos fijos	Efectos aleatorios
Tratamientos	$k - 1$	s_1^2	$\sigma^2 + \frac{b}{k-1} \sum_i \alpha_i^2$	$\sigma^2 + b\sigma_\alpha^2$
Bloques	$b - 1$	s_2^2	$\sigma^2 + \frac{k}{b-1} \sum_j \beta_j^2$	$\sigma^2 + k\sigma_\beta^2$
Error	$(k-1)(b-1)$	s^2	σ^2	σ^2
Total	$kb - 1$			

y H_0 se rechaza cuando $f > f_\alpha[k-1, (b-1)(k-1)]$.

Los estimados no sesgados de los componentes de la varianza son

$$\hat{\sigma}^2 = s^2, \quad \hat{\sigma}_\alpha^2 = \frac{s_1^2 - s^2}{b}, \quad \hat{\sigma}_\beta^2 = \frac{s_2^2 - s^2}{k}.$$

Las pruebas de las hipótesis referentes a los diversos componentes de la varianza se realizan calculando las razones de los cuadrados medios adecuados, tal como se indica en la tabla 13.14, y comparándolos con los valores f correspondientes de la tabla A.6.

13.12 Estudio de caso

Estudio de caso 13.1: Análisis químico. Se pidió al personal del Departamento de Química de Virginia Tech que analizara un conjunto de datos que se obtuvo para comparar 4 métodos distintos de análisis del aluminio en cierta mezcla deflagradora sólida. Para considerar una amplia gama de laboratorios de análisis se utilizaron 5 de ellos en el experimento. Se seleccionaron esos laboratorios porque suelen realizar esa clase de análisis. Se asignaron al azar 20 muestras de material deflagrador que contenían 2.70% de aluminio, cuatro a cada laboratorio, y se dieron instrucciones acerca de cómo efectuar los análisis químicos utilizando los cuatro métodos. Los datos que se obtuvieron son los siguientes:

Método	Laboratorio					Media
	1	2	3	4	5	
A	2.67	2.69	2.62	2.66	2.70	2.668
B	2.71	2.74	2.69	2.70	2.77	2.722
C	2.76	2.76	2.70	2.76	2.81	2.758
D	2.65	2.69	2.60	2.64	2.73	2.662

Los laboratorios no se consideran efectos aleatorios, ya que no fueron seleccionados al azar de entre una población más grande de ellos. Los datos se analizaron como un diseño de bloques completos aleatorizados. Se dibujaron gráficas de los datos para determinar si era apropiado un modelo aditivo del tipo:

$$y_{ij} = \mu + m_i + l_j + \epsilon_{ij}$$

en otras palabras, un modelo con efectos aditivos. El bloque aleatorizado no es adecuado cuando existe interacción entre los laboratorios y los métodos. Considere la gráfica de la figura 13.12. Aunque es un poco difícil de interpretar porque cada punto representa una sola observación, parece que no hay interacción evidente entre los métodos y los laboratorios.

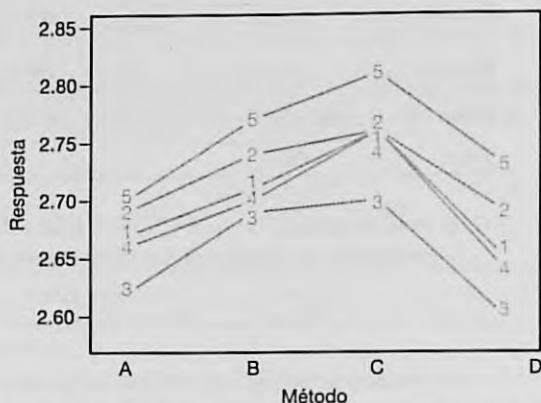


Figura 13.12: Gráfica de interacción para los datos del estudio de caso 13.1.

Gráficas de residuales

Las gráficas de residuales se usaron como indicaciones de diagnóstico con respecto a la suposición de una varianza homogénea. La figura 13.13 presenta una gráfica de residuales contra los métodos de análisis. La variabilidad descrita en los residuales parece ser bastante homogénea. Para completar, en la figura 13.14 se presenta una gráfica de probabilidad normal de los residuales.

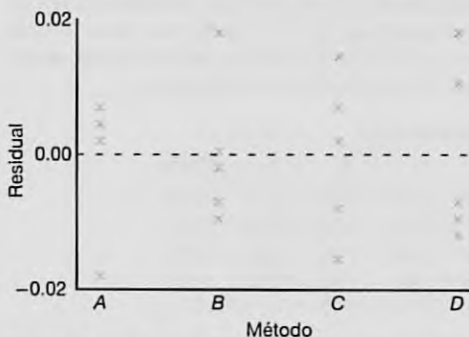


Figura 13.13: Gráfica de residuales en comparación con el método para los datos del estudio de caso 13.1.

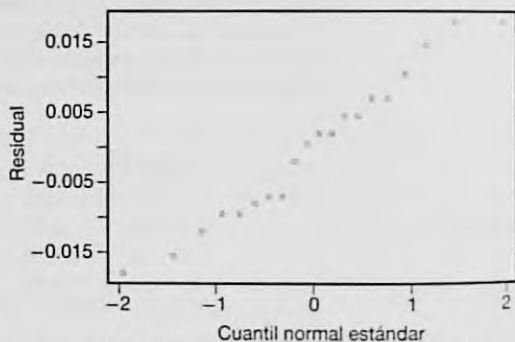


Figura 13.14: Gráfica de probabilidad normal de residuales para los datos del estudio de caso 13.1.

Las gráficas de residuales no muestran problemas con la suposición de errores normales ni con la de varianza homogénea. Para hacer el análisis de varianza se utilizó la

función PROC GLM del programa SAS. En la figura 13.15 se incluye una salida de resultados por computadora con comentarios.

Los valores f y P calculados sí indican una diferencia significativa entre los métodos de análisis. A este análisis le puede seguir un análisis de comparación múltiple para determinar en dónde están las diferencias entre los métodos.

Ejercicios

13.37 Al probar muestras de sangre de un paciente para detectar anticuerpos del VIH un espectrómetro determina la densidad óptica de cada muestra. La densidad óptica se mide como la absorbencia de la luz de cierta longitud de onda. La muestra de sangre es positiva si excede a cierto valor límite que se determina con muestras de control para esa corrida. A los investigadores les interesa comparar la variabilidad del laboratorio para los valores de control positivo. Los datos representan valores de control positivo para 10 corridas distintas en cuatro laboratorios seleccionados al azar.

Corrida	Laboratorio			
	1	2	3	4
1	0.888	1.065	1.325	1.232
2	0.983	1.226	1.069	1.127
3	1.047	1.332	1.219	1.051
4	1.087	0.958	0.958	0.897
5	1.125	0.816	0.819	1.222
6	0.997	1.015	1.140	1.125
7	1.025	1.071	1.222	0.990
8	0.969	0.905	0.995	0.875
9	0.898	1.140	0.928	0.930
10	1.018	1.051	1.322	0.775

- Escriba un modelo adecuado para este experimento.
- Estime el componente de varianza del laboratorio y la varianza dentro de los laboratorios.

13.38 Se efectúa un experimento en el que se compararán 4 tratamientos en 5 bloques. Los datos son los siguientes:

Tratamiento	Bloque				
	1	2	3	4	5
1	12.8	10.6	11.7	10.7	11.0
2	11.7	14.2	11.8	9.9	13.8
3	11.5	14.7	13.6	10.7	15.9
4	12.6	16.5	15.4	9.6	17.1

- Suponga que se trata de un modelo de efectos aleatorios y pruebe la hipótesis de que no hay diferencia entre las medias de los tratamientos, a un nivel de significancia de 0.05.
- Calcule estimados de los componentes de la varianza del tratamiento y del bloque.

13.39 Los siguientes datos muestran el efecto de cuatro operadores, elegidos al azar, sobre la producción de una máquina específica:

	Operador			
	1	2	3	4
	175.4	168.5	170.1	175.2
	171.7	162.7	173.4	175.7
	173.0	165.0	175.7	180.1
	170.5	164.1	170.7	183.7

- Realice un análisis de varianza de efectos aleatorios a un nivel de significancia de 0.05.
- Calcule un estimado del componente de la varianza del operador y del componente de la varianza del error experimental.

13.40 De cinco "vaciados" de metales se tomaron cinco muestras del núcleo y en cada una se analizó la cantidad de un elemento traza. Los siguientes son los datos de los 5 vaciados seleccionados al azar:

Núcleo	Vaciado				
	1	2	3	4	5
1	0.98	0.85	1.12	1.21	1.00
2	1.02	0.92	1.68	1.19	1.21
3	1.57	1.16	0.99	1.32	0.93
4	1.25	1.43	1.26	1.08	0.86
5	1.16	0.99	1.05	0.94	1.41

- La intención es que los vaciados sean idénticos. Por lo tanto, pruebe que el componente de la varianza del "vaciado" es igual a cero. Saque conclusiones.
- Realice un ANOVA completo y obtenga un estimado de la varianza dentro del vaciado.

13.41 Una empresa textil produce cierta tela en un número grande de telares. Los gerentes quieren que los telares sean homogéneos para que la tela que producen tenga una resistencia uniforme. Se sospecha que puede haber una variación significativa entre la resistencia de los telares. Considere los siguientes datos para 4 telares seleccionados al azar. Cada observación es una determinación de la resistencia de la tela expresada en libras por pulgada cuadrada.

	Telar			
	1	2	3	4
	99	97	94	93
	97	96	95	94
	97	92	90	90
	96	98	92	92

- Escriba un modelo para el experimento.
- ¿El componente de la varianza del telar difiere significativamente de cero?
- Haga comentarios sobre la sospecha de los gerentes.

The GLM Procedure					
Class Level Information					
Class	Levels	Values			
Method	4	A	B	C	D
Lab	5	1	2	3	4 5
Number of Observations Read					20
Number of Observations Used					20
Dependent Variable: Response					
Sum of					
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	7	0.05340500	0.00762929	42.19	<.0001
Error	12	0.00217000	0.00018083		
Corrected Total	19	0.05557500			
R-Square Coeff Var Root MSE Response Mean					
0.960954	0.497592	0.013447	2.702500		
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Method	3	0.03145500	0.01048500	57.98	<.0001
Lab	4	0.02195000	0.00548750	30.35	<.0001
Observation	Observed	Predicted	Residual		
1	2.67000000	2.66300000	0.00700000		
2	2.71000000	2.71700000	-0.00700000		
3	2.76000000	2.75300000	0.00700000		
4	2.65000000	2.65700000	-0.00700000		
5	2.69000000	2.68550000	0.00450000		
6	2.74000000	2.73950000	0.00050000		
7	2.76000000	2.77550000	-0.01550000		
8	2.69000000	2.67950000	0.01050000		
9	2.62000000	2.61800000	0.00200000		
10	2.69000000	2.67200000	0.01800000		
11	2.70000000	2.70800000	-0.00800000		
12	2.60000000	2.61200000	-0.01200000		
13	2.66000000	2.65550000	0.00450000		
14	2.70000000	2.70950000	-0.00950000		
15	2.76000000	2.74550000	0.01450000		
16	2.64000000	2.64950000	-0.00950000		
17	2.70000000	2.71800000	-0.01800000		
18	2.77000000	2.77200000	-0.00200000		
19	2.81000000	2.80800000	0.00200000		
20	2.73000000	2.71200000	0.01800000		

Figura 13.15: Salida de resultados por computadora del SAS para los datos del estudio de caso 13.1.

Ejercicios de repaso

13.42 El Centro de Consultoría en Estadística de Virginia Tech, junto con el Departamento de Silvicultura, llevaron a cabo un análisis. Se aplicó cierto tratamiento a tres cepas de árbol. Se empleó el producto químico Garlon con el fin de regenerar las raíces de las cepas. Se usó un aerosol con cuatro niveles de concentración de Garlon. Después de cierto tiempo, se observó la altura de los retoños. Realice un análisis de varianza de un solo factor con los siguientes datos. Haga pruebas para saber si la concentración de Garlon tiene un efecto significativo sobre la altura de los retoños. Emplee $\alpha = 0.05$.

Nivel de Garlon							
1	2	3	4	1	2	3	4
2.87	2.31	3.27	2.66	2.39	1.91	3.05	0.91
3.91	2.04	3.15	2.00	2.89	1.89	2.43	0.01

13.43 Considere los datos de los agregados del ejemplo 13.1. Efectúe una prueba de Bartlett a un nivel de significancia $\alpha = 0.1$ para determinar si hay heterogeneidad en la varianza entre los agregados.

13.44 En un proceso químico se utilizaron 3 catalizadores y también se incluyó un control (no catalizador). Se tienen los datos siguientes de la producción del proceso:

Control	Catalizador		
	1	2	3
74.5	77.5	81.5	78.1
76.1	82.0	82.3	80.2
75.9	80.6	81.4	81.5
78.1	84.9	79.5	83.0
76.2	81.0	83.0	82.1

Use una prueba de Dunnett a un nivel de significancia $\alpha = 0.01$ para determinar si se obtuvo una producción significativamente más alta con los catalizadores que sin ellos.

13.45 Se emplean cuatro laboratorios para efectuar análisis químicos. Se envían muestras del mismo material a los laboratorios para que, como parte del estudio, las analicen para determinar si dan o no, en promedio, los mismos resultados. Los resultados analíticos de los cuatro laboratorios son los siguientes:

Laboratorio	A	B	C	D
	58.7	62.7	55.9	60.7
61.4	64.5	56.1	60.3	
60.9	63.1	57.3	60.9	
59.1	59.2	55.2	61.4	
58.2	60.3	58.1	62.3	

a) Utilice una prueba de Bartlett para demostrar que las varianzas dentro de los laboratorios no difieren de manera significativa a un nivel de significancia $\alpha = 0.05$.

b) Realice el análisis de varianza y saque conclusiones acerca de los laboratorios.

c) Dibuje una gráfica de probabilidad normal de residuales.

13.46 Se diseñó un experimento para el personal del Departamento de Ciencia Animal de Virginia Tech, con el propósito de estudiar el tratamiento con urea y amoníaco acuoso de la espiga del trigo. El propósito era mejorar el valor nutricional para las ovejas macho. Los tratamientos dietéticos son: control, urea en la alimentación, espiga tratada con amoníaco, espiga tratada con urea. En el experimento se emplearon 24 ovejas y se separaron de acuerdo con su peso relativo. En cada grupo homogéneo había cuatro ovejas (según el peso) y cada una recibió una de las cuatro dietas en orden aleatorio. Se midió el porcentaje de materia seca digerida de las 24 ovejas. Los siguientes son los datos:

Dieta	Grupo por peso (bloque)					
	1	2	3	4	5	6
Control	32.68	36.22	36.36	40.95	34.99	33.89
Urea en la alimentación	35.90	38.73	37.55	34.64	37.36	34.35
Tratada con amoníaco	49.43	53.50	52.86	45.00	47.20	49.76
Tratada con urea	46.58	42.82	45.41	45.08	43.81	47.40

a) Use un análisis de bloques completos aleatorizados para probar las diferencias entre las dietas. Emplee $\alpha = 0.05$.

b) Utilice la prueba de Dunnett para comparar las tres dietas con el control. Utilice $\alpha = 0.05$.

c) Dibuje una gráfica de probabilidad normal de los residuales.

13.47 En un estudio realizado por el personal del Departamento de Bioquímica de Virginia Tech se dieron tres dietas a un grupo de ratas con el objetivo de estudiar el efecto de cada una sobre el zinc dietético residual en el torrente sanguíneo. Se asignaron al azar cinco ratas preñadas a cada grupo dietético, y cada una recibió la dieta en el día 22 del embarazo. Se midió la cantidad de zinc en partes por millón. Los datos son los que siguen:

1	0.50	0.42	0.65	0.47	0.44
Dieta 2	0.42	0.40	0.73	0.47	0.69
3	1.06	0.82	0.72	0.72	0.82

Determine si hay una diferencia significativa en el zinc dietético residual entre las tres dietas. Use $\alpha = 0.05$. Lleve a cabo un ANOVA de un solo factor.

13.48 Se realizó un experimento para comparar tres tipos de pintura para buscar evidencia de diferencias en su calidad de desgaste. Las pinturas se expusieron a acciones abrasivas y se registró el tiempo, en horas, que tardaba en observarse la abrasión. Se usaron seis especímenes para cada tipo de pintura. Los datos son los siguientes:

Tipo de pintura								
1			2			3		
158	97	282	515	264	544	317	662	213
315	220	115	525	330	525	536	175	614

- Realice un análisis de varianza para determinar si la evidencia sugiere que la calidad del desgaste de las tres pinturas es diferente. Utilice un valor P en sus conclusiones.
- Si se encuentran diferencias significativas, diga cuáles son. ¿Hay alguna pintura que destaque? Analice sus hallazgos.
- Haga todos los análisis gráficos que necesite para determinar si son válidas las suposiciones que se hicieron en el inciso a . Analice sus hallazgos.
- Suponga que se determina que los datos para cada tratamiento tienen una distribución exponencial. ¿Sugiere esto un análisis alternativo? Si fuera así, hágalo y presente sus hallazgos.

13.49 Una empresa que troquea juntas de hojas de caucho, plástico y corcho desea comparar el número medio de juntas producidas por hora para los tres tipos de material. Se eligieron al azar dos máquinas troqueladoras como bloques. Los datos representan el número de juntas (en miles) producidas por hora. En la figura 13.16 de la página 557 se observa la salida de resultados del análisis.

Máquina	Corcho		Material Caucho			Plástico			
	A	4.31	4.27	4.40	3.36	3.42	3.48	4.01	3.94
B	3.94	3.81	3.99	3.91	3.80	3.85	3.48	3.53	3.42

- ¿Por qué se eligieron las máquinas troqueladoras como bloques?
- Grafique las seis medias para las combinaciones de máquinas y materiales.
- ¿Hay un material que sea mejor?
- ¿Existe interacción entre los tratamientos y los bloques? Si es así, diga si la interacción ocasiona alguna dificultad seria para llegar a una conclusión adecuada. Explique su respuesta.

13.50 Se hizo un estudio para comparar el rendimiento de tres marcas de gasolina competidoras. Se seleccionaron al azar cuatro modelos de automóvil de tamaño variable. A continuación se presentan los datos, en millas por galón. El orden de prueba es aleatorio para cada modelo.

Modelo	Marca de gasolina		
	A	B	C
A	32.4	35.6	38.7
B	28.8	28.6	29.9
C	36.5	37.6	39.1
D	34.4	36.2	37.9

- Analice la necesidad de utilizar más de un solo modelo de automóvil.
- Considere el ANOVA de la salida de resultados del SAS en la figura 13.17. ¿Es importante la marca de la gasolina?
- ¿Qué marca de gasolina seleccionaría usted? Consulte el resultado de la prueba de Duncan.

13.51 Se utilizaron cuatro localidades diferentes del noreste para hacer mediciones de ozono, en partes por millón. Se recolectaron las cantidades de ozono en cinco muestras de cada localidad.

Localidad			
1	2	3	4
0.09	0.15	0.10	0.10
0.10	0.12	0.13	0.07
0.08	0.17	0.08	0.05
0.08	0.18	0.08	0.08
0.11	0.14	0.09	0.09

- ¿Hay información suficiente que sugiera que existen diferencias en los niveles medios de ozono entre las diferentes localidades? Gúfese usando un valor P .
- Si se encuentran diferencias significativas en el inciso a , determine su naturaleza. Emplee cualesquiera métodos que haya aprendido.

13.52 Demuestre que el cuadrado medio del error

$$s^2 = \frac{SCE}{k(n-1)}$$

para el análisis de varianza en la clasificación de un factor es un estimado no sesgado de σ^2 .

13.53 Demuestre el teorema 13.2.

13.54 Demuestre que la fórmula para calcular la SCB, en el análisis de varianza del diseño de bloques completos aleatorizados, es equivalente al término correspondiente en la identidad del teorema 13.3.

13.55 Para el diseño de bloques aleatorizados con k tratamientos y b bloques, demuestre que

$$E(SCB) = (b-1)\sigma^2 + k \sum_{j=1}^k \beta_j^2$$

The GLM Procedure

Dependent Variable: gasket

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1.68122778	0.33624556	76.52	<.0001
Error	12	0.05273333	0.00439444		
Corrected Total	17	1.73396111			
R-Square	Coeff Var	Root MSE	gasket Mean		
0.969588	1.734095	0.066291	3.822778		

Source	DF	Type III SS	Mean Square	F Value	Pr > F
material	2	0.81194444	0.40597222	92.38	<.0001
machine	1	0.10125000	0.10125000	23.04	0.0004
material*machine	2	0.76803333	0.38401667	87.39	<.0001

Level of material	Level of machine	N	Mean	Std Dev
cork	A	3	4.32666667	0.06658328
cork	B	3	3.91333333	0.09291573
plastic	A	3	3.94666667	0.06027714
plastic	B	3	3.47666667	0.05507571
rubber	A	3	3.42000000	0.06000000
rubber	B	3	3.85333333	0.05507571

Level of material	N	Mean	Std Dev
cork	6	4.12000000	0.23765521
plastic	6	3.71166667	0.26255793
rubber	6	3.63666667	0.24287171

Level of machine	N	Mean	Std Dev
A	9	3.89777778	0.39798800
B	9	3.74777778	0.21376259

Figura 13.16: Salida de resultados por computadora del SAS para el ejercicio de repaso 13.49.

The GLM Procedure

Dependent Variable: MPG

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	153.2508333	30.6501667	24.66	0.0006
Error	6	7.4583333	1.2430556		
Corrected Total	11	160.7091667			

R-Square	Coeff Var	Root MSE	MPG Mean
0.953591	3.218448	1.114924	34.64167

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Model	3	130.3491667	43.4497222	34.95	0.0003
Brand	2	22.9016667	11.4508333	9.21	0.0148

Duncan's Multiple Range Test for MPG

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	6
Error Mean Square	1.243056

Number of Means	2	3
Critical Range	1.929	1.999

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	Brand
A	36.4000	4	C
A			
B	34.5000	4	B
B			
B	33.0250	4	A

Figura 13.17 Salida de resultados por computadora del SAS para el ejercicio de repaso 13.50.

13.56 Proyecto de grupo: Resulta de interés determinar qué tipo de pelota deportiva se puede lanzar a la mayor distancia. La competencia incluye una pelota de tenis, una de beisbol y una de softbol. Divida el grupo en equipos de cinco estudiantes. Cada equipo debe diseñar y realizar un experimento separado, también debe analizar los datos de su propio experimento. Los cinco miembros del equipo lanzarán cada pelota (después de calentar el brazo el tiempo adecuado). La respuesta experimental será la distancia (en pies) que se lanza la pelota. Los datos de cada equipo incluirán 15 observaciones. Aspectos importantes:

a) No se trata de una competencia entre equipos. La competencia es entre los tres tipos de pelotas. Se esperaría que las conclusiones de cada equipo sean similares.

- En cada equipo debe haber hombres y mujeres.
- El diseño experimental de cada equipo deberá ser un diseño de bloques completos aleatorizados. Los cinco individuos que lanzan la pelota son los bloques.
- Asegúrese de incorporar la aleatorización adecuada para realizar el experimento.
- Los resultados deberán contener una descripción del experimento con una tabla de ANOVA que incluya un valor P y las conclusiones apropiadas. Se usarán técnicas gráficas y comparaciones múltiples en caso de ser necesarias. Saquen conclusiones prácticas con respecto a las diferencias entre los tipos de pelotas. Sean meticulosos.

13.13 Posibles riesgos y errores conceptuales; relación con el material de otros capítulos

Al igual que otros procedimientos estudiados en capítulos anteriores, el análisis de varianza es razonablemente robusto con respecto a la suposición de normalidad, pero no lo es tanto en cuanto a la suposición de varianza homogénea. También observamos que la prueba de Bartlett para varianzas iguales es sumamente débil en relación con la normalidad.

Este capítulo es sumamente importante, ya que se trata de un punto “de inicio” para temas importantes, como el diseño de experimentos y el análisis de varianza. En el capítulo 14 se tratan los mismos temas, pero en los casos de extensiones a más de un factor y el análisis más complicado por la interpretación de la interacción entre factores. Hay ocasiones en que el papel de la interacción en un experimento científico es más importante que el papel de los factores principales (efectos principales). Ante la presencia de interacciones se hace un énfasis aún mayor en las técnicas gráficas. En los capítulos 14 y 15 será necesario proporcionar más detalles acerca del proceso de aleatorización, ya que el número de combinaciones de factores puede ser muy grande.

Capítulo 14

Experimentos factoriales (dos o más factores)

14.1 Introducción

Considere una situación en la que haya interés por estudiar el efecto de **dos factores**, *A* y *B*, sobre alguna respuesta. Por ejemplo, en un experimento químico nos gustaría variar en forma simultánea la presión de reacción y el tiempo de reacción, y estudiar el efecto que cada uno tiene sobre el producto. En un experimento biológico resulta de interés estudiar el efecto que tienen el tiempo de secado y la temperatura sobre la cantidad de sólidos (porcentaje por peso) que queda en muestras de levadura. Igual que en el capítulo 13, el término **factor** se utiliza en un sentido general para denotar cualquier característica del experimento que pueda variar de un ensayo a otro, como la temperatura, el tiempo o la presión. Los **niveles** de un factor se definen como los valores reales que se utilizan en el experimento.

Para cada uno de estos casos es importante determinar no sólo si cada uno de los 2 factores influye en la respuesta, sino también si hay una interacción significativa entre ellos. En lo que se refiere a la terminología, el experimento descrito aquí es de 2 factores, y el diseño experimental podría ser uno completamente aleatorizado, en el que las distintas combinaciones de tratamiento se asignan al azar a todas las unidades experimentales, o bien, un diseño de bloques completos aleatorizados, donde las combinaciones de factores se asignan al azar dentro de los bloques. En el ejemplo de la levadura, si se empleara un diseño completamente aleatorizado, las distintas combinaciones de tratamientos de temperatura y tiempo de secado se asignarían al azar a las muestras de levadura.

Muchos de los conceptos que se estudiaron en el capítulo 13 se extienden a 2 y 3 factores en este capítulo. El objetivo principal de este material es el uso del diseño completamente aleatorizado con un *experimento factorial*. Un experimento factorial con 2 factores implica ensayos experimentales (o uno solo) con todas las combinaciones de factores. Por ejemplo, en el caso de la temperatura y tiempo de secado con, digamos, 3 niveles de cada uno y $n = 2$ corridas por cada una de las 9 combinaciones, tendríamos un *experimento factorial de 2 factores en un diseño completamente aleatorizado*. Ninguno de ellos es un factor de bloqueo; nos interesa la manera en que cada uno influye en el porcentaje de sólidos en las muestras, y si interactúan o no. El biólogo dispondría de 18 muestras físicas de material que constituyen unidades experimentales. Luego, éstas se asignarían al azar a las 18 combinaciones (9 combinaciones de tratamiento, cada una de ellas por duplicado).

Antes de entrar en detalles analíticos, sumas de cuadrados y demás, sería interesante que el lector observe la clara conexión que existe entre lo que hemos descrito y la situación con el problema de un solo factor. Considere el experimento de la levadura. La explicación de los grados de libertad ayuda a que el lector o el analista visualicen la extensión. En un inicio, las 9 combinaciones de tratamientos deberían considerarse como si representaran un factor con 9 niveles (8 grados de libertad). Así, un vistazo inicial a los grados de libertad arroja lo siguiente:

Combinaciones de tratamiento	8
Error	9
Total	17

Efectos principales e interacción

En realidad el experimento se podría analizar como se describe en la tabla anterior. Sin embargo, es probable que la prueba F para las combinaciones no dé al analista la información que desea, es decir, el papel de la temperatura y del tiempo de secado. Tres tiempos de secado tienen asociados 2 grados de libertad, y a 3 temperaturas se asocian también 2 grados de libertad. Los factores principales, la temperatura y el tiempo de secado reciben el nombre de **efectos principales**, los cuales representan 4 de los 8 grados de libertad para las *combinaciones de factores*. Los 4 grados de libertad adicionales se asocian con la *interacción* entre los 2 factores. Como resultado, el análisis incluye

Combinaciones	8
Temperatura	2
Tiempo de secado	2
Interacción	4
Error	9
Total	17

En el capítulo 13 vimos que en un análisis de varianza los factores pueden considerarse fijos o aleatorios, dependiendo del tipo de inferencia deseada y de la manera en que se eligieron los niveles. Aquí debemos considerar los efectos fijos, los efectos aleatorios e incluso los casos en que los efectos son mixtos. Conforme avancemos en estos temas pondremos mayor atención a los cuadrados medios esperados. En la siguiente sección nos centraremos en el concepto de interacción.

14.2 Interacción en el experimento de dos factores

En el modelo de bloques aleatorizados que se estudió previamente se supuso que en cada bloque se toma una observación de cada tratamiento. Si la suposición del modelo es correcta, es decir, si los bloques y los tratamientos son los únicos efectos reales y no hay interacción, el valor esperado del cuadrado medio del error es la varianza del error experimental σ^2 . Sin embargo, suponga que existe interacción entre los tratamientos y los bloques, como lo indica el modelo

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$$

de la sección 13.8. El valor esperado del cuadrado medio del error entonces es dado por

$$E \left[\frac{SCE}{(b-1)(k-1)} \right] = \sigma^2 + \frac{1}{(b-1)(k-1)} \sum_{i=1}^k \sum_{j=1}^b (\alpha\beta)_{ij}^2.$$

Los efectos del tratamiento y los bloques no aparecen en el cuadrado medio del error esperado, pero los efectos de la interacción sí. Entonces, si en el modelo hay interacción, el cuadrado medio del error refleja variación debida al error experimental más una contribución de la interacción y, para este plan experimental, no hay forma de separarlos.

La interacción y la interpretación de los efectos principales

Desde el punto de vista del experimentador, parecería necesario llegar a una prueba significativa sobre la existencia de una interacción, al separar la variación del error verdadero de aquel que se debe a la interacción. Los efectos principales, A y B , adoptan un significado distinto en presencia de la interacción. En el ejemplo biológico anterior el efecto que tiene el tiempo de secado sobre la cantidad de sólidos que quedan en la levadura muy bien podría depender de la temperatura a la que se expusieron las muestras. En general, podrían existir situaciones experimentales en las que el factor A tuviera un efecto positivo sobre la respuesta en un nivel del factor B ; en tanto que con un nivel distinto de B el efecto de A sería negativo. Aquí se usa el término **efecto positivo** para indicar que el producto o la respuesta se incrementan conforme los niveles de un factor dado aumentan de acuerdo con cierto orden definido. En el mismo sentido, un **efecto negativo** corresponde a una disminución de la respuesta al aumentar los niveles del factor.

Considere, por ejemplo, los siguientes datos de temperatura (factor A con niveles t_1 , t_2 y t_3 (en orden creciente) y tiempo de secado d_1 , d_2 y d_3 (también en orden creciente). La respuesta es el porcentaje de sólidos. Estos datos son completamente hipotéticos y se dan para ilustrar un aspecto.

A	B			Total
	d_1	d_2	d_3	
t_1	4.4	8.8	5.2	18.4
t_2	7.5	8.5	2.4	18.4
t_3	9.7	7.9	0.8	18.4
Total	21.6	25.2	8.4	55.2

Es evidente que el efecto de la temperatura sobre el porcentaje de sólidos es positivo para el tiempo breve de secado d_1 , pero negativo para el tiempo prolongado d_3 . Esta **interacción clara** entre la temperatura y el tiempo de secado es evidentemente interesante para el biólogo; sin embargo, con base en los totales de las respuestas para las temperaturas t_1 , t_2 y t_3 , la suma de cuadrados de la temperatura, SCT , producirá un valor de 0. Entonces, se dice que la presencia de la interacción **enmascara** el efecto de la temperatura. Por ello, si se considera el efecto medio de la temperatura, promediado para el tiempo de secado, **no existe efecto alguno**. Entonces, esto define el efecto principal. Pero, por supuesto, es probable que esto no sea pertinente para el biólogo.

Antes de sacar cualquier conclusión final a partir de las pruebas de significancia sobre los efectos principales y los efectos de la interacción, el **experimentador debería observar primero si la prueba para la interacción es o no significativa**. Si la

interacción no es significativa, entonces los resultados de las pruebas sobre los efectos principales son importantes. No obstante, si la interacción debe ser significativa, entonces solamente son importantes aquellas pruebas sobre los efectos principales que resultan significativas. En presencia de una interacción, los efectos principales no significativos bien podrían ser resultado de enmascaramiento e indicar la necesidad de observar la influencia de cada factor a niveles fijos del otro.

Representación gráfica de la interacción

La presencia de interacción, así como su impacto científico, se puede interpretar adecuadamente usando **gráficas de interacción**. Las gráficas proporcionan una clara imagen de la tendencia de los datos para mostrar el efecto que tiene el cambio de un factor conforme se pasa de un nivel a otro del segundo factor. La figura 14.1 ilustra la fuerte interacción entre la temperatura y el tiempo de secado. La interacción se revela en las líneas no paralelas.

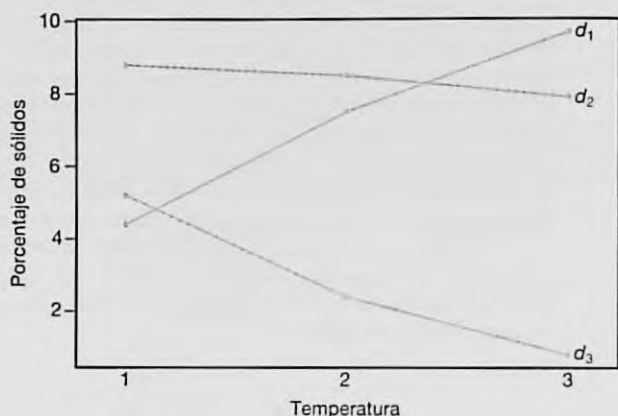


Figura 14.1: Gráfica de la interacción para los datos de temperatura y tiempo de secado.

El efecto relativamente fuerte de la temperatura sobre el porcentaje de sólidos en el tiempo de secado más breve se refleja en la marcada pendiente de d_1 . En el tiempo de secado medio, d_2 , la temperatura tiene muy poco efecto, mientras que en el tiempo de secado prolongado d_3 la pendiente negativa indica un efecto negativo de la temperatura. Las gráficas de interacción como ésta le permiten al científico hacer una interpretación rápida y significativa de la interacción que existe. Debe ser evidente que el paralelismo en las gráficas indica la ausencia de interacción.

Necesidad de observaciones múltiples

En el experimento de 2 factores, la interacción y el error experimental sólo se separan si se hacen observaciones múltiples con las distintas combinaciones de tratamiento. Para máxima eficiencia debe haber el mismo número n de observaciones para cada combinación. Éstas deben ser verdaderas réplicas, no sólo medidas repetidas. Por ejemplo, en

el caso de la levadura, si para cada combinación de temperatura y tiempo de secado se toman $n = 2$ observaciones, debería haber dos muestras separadas y no sólo mediciones repetidas en la misma muestra. Esto permite que la variabilidad debida a las unidades experimentales aparezca en el "error", de manera que la variación no es un simple error de medición.

14.3 Análisis de varianza de dos factores

Al presentar las fórmulas generales para el análisis de varianza de un experimento de 2 factores utilizando observaciones repetidas en un diseño completamente aleatorizado, debe considerarse el caso de n réplicas de las combinaciones del tratamiento, determinadas por a niveles del factor A y b niveles del factor B . Las observaciones se podrían clasificar usando un arreglo rectangular, donde los renglones representan los niveles del factor A y las columnas representan los niveles del factor B . Cada combinación de tratamiento define una celda del arreglo. Así, se tienen ab celdas, cada una de las cuales contiene n observaciones. Se denota con y_{ijk} la k -ésima observación tomada en el i -ésimo nivel del factor A y el j -ésimo nivel del factor B . En la tabla 14.1 se muestran las abn observaciones.

Tabla 14.1: Experimento de dos factores con n réplicas

A	B				Total	Media
	1	2	...	b		
1	y_{111}	y_{121}	...	y_{1b1}	$Y_{1..}$	$\bar{y}_{1..}$
	y_{112}	y_{122}	...	y_{1b2}		
	\vdots	\vdots		\vdots		
	y_{11n}	y_{12n}	...	y_{1bn}		
2	y_{211}	y_{221}	...	y_{2b1}	$Y_{2..}$	$\bar{y}_{2..}$
	y_{212}	y_{222}	...	y_{2b2}		
	\vdots	\vdots		\vdots		
	y_{21n}	y_{22n}	...	y_{2bn}		
\vdots	\vdots		\vdots	\vdots	\vdots	
a	y_{a11}	y_{a21}	...	y_{ab1}	$Y_{a..}$	$\bar{y}_{a..}$
	y_{a12}	y_{a22}	...	y_{ab2}		
	\vdots	\vdots		\vdots		
	y_{a1n}	y_{a2n}	...	y_{abn}		
Total	$Y_{.1}$	$Y_{.2}$...	$Y_{.b}$	$Y_{...}$	
Media	$\bar{y}_{.1}$	$\bar{y}_{.2}$...	$\bar{y}_{.b}$		$\bar{y}_{...}$

Las observaciones en la celda (ij) -ésima constituyen una muestra aleatoria de tamaño n de una población que se supone tiene distribución normal con media μ_{ij} y varianza σ^2 . Se supone que todas las ab poblaciones tienen la misma varianza σ^2 . Se

definen los siguientes símbolos útiles, algunos de los cuales se utilizan en la tabla 14.1:

- Y_{ij} = suma de las observaciones en la (ij) -ésima celda,
- $Y_{i.}$ = suma de las observaciones para el i -ésimo nivel del factor A ,
- $Y_{.j}$ = suma de las observaciones para el j -ésimo nivel del factor B ,
- $Y_{..}$ = suma de todas las abn observaciones,
- \bar{y}_{ij} = media de las observaciones en la (ij) -ésima celda,
- $\bar{y}_{i.}$ = media de las observaciones para el i -ésimo nivel del factor A ,
- $\bar{y}_{.j}$ = media de las observaciones para el j -ésimo nivel del factor B ,
- $\bar{y}_{..}$ = media de todas las abn observaciones.

A diferencia de la situación para un solo factor, que se cubrió con amplitud en el capítulo 13, en éste supondremos que las **poblaciones**, de las que se toman n observaciones independientes con distribución idéntica, son **combinaciones** de los factores. Asimismo, se supondrá siempre que de cada combinación de factores se toma un número igual (n) de observaciones. En los casos en que los tamaños de las muestras por combinación son desiguales, los cálculos son más complicados, aunque los conceptos son transferibles.

Modelo e hipótesis para el problema de dos factores

Cada observación de la tabla 14.1 se puede escribir en la siguiente forma:

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk},$$

donde ϵ_{ijk} mide las desviaciones de los valores y_{ijk} observados en la (ij) -ésima celda a partir de la media de la población μ_{ij} . Si $(\alpha\beta)_{ij}$ denota el efecto de la interacción del i -ésimo nivel del factor A y el j -ésimo nivel del factor B , α_i el efecto del i -ésimo nivel del factor A , β_j el efecto del j -ésimo nivel del factor B , y μ la media conjunta, escribimos

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij},$$

y, entonces,

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$

a las que se imponen las restricciones

$$\sum_{i=1}^a \alpha_i = 0, \quad \sum_{j=1}^b \beta_j = 0, \quad \sum_{i=1}^a (\alpha\beta)_{ij} = 0, \quad \sum_{j=1}^b (\alpha\beta)_{ij} = 0.$$

Las 3 hipótesis por probar son las siguientes:

1. H_0' : $\alpha_1 = \alpha_2 = \dots = \alpha_a = 0$,
 H_1' : Al menos una de las α_i no es igual a 0.
2. H_0'' : $\beta_1 = \beta_2 = \dots = \beta_b = 0$,
 H_1'' : Al menos una de las β_j no es igual a 0.

$$3. H_0''': (\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{ab} = 0,$$

H_1''' : Al menos una de las $(\alpha\beta)_{ij}$ no es igual a 0.

Se alerta al lector acerca del problema del enmascaramiento de los efectos principales cuando la interacción contribuye de manera importante en el modelo. Se recomienda considerar primero el resultado de la prueba de interacción y, luego, la interpretación de la prueba del efecto principal; la naturaleza de la conclusión científica depende de si se encontró interacción. Si ésta se descarta, entonces se pueden probar las hipótesis 1 y 2 y la interpretación es muy sencilla. Sin embargo, si se descubre que hay interacción, la interpretación puede ser más complicada, como se vio al analizar el tiempo de secado y la temperatura en la sección previa. La estructura de las pruebas de hipótesis 1, 2 y 3 se estudiará en las secciones siguientes. En el análisis del ejemplo 14.1 se incluirá la interpretación de los resultados.

Las pruebas de las hipótesis anteriores se basarán en la comparación de estimados independientes de σ^2 , obtenidos al separar la suma de cuadrados total de los datos en 4 componentes mediante la siguiente identidad.

Partición de la variabilidad en el caso de dos factores

Teorema 14.1: Identidad de la suma de cuadrados

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 &= bn \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{...})^2 + an \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{...})^2 \\ &+ n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij})^2 \end{aligned}$$

Simbólicamente, la identidad de la suma de cuadrados se escribe como

$$SCT = SCA + SCB + SC(AB) + SCE$$

donde a SCA y SCB se les denomina la suma de cuadrados para los efectos principales A y B , respectivamente, $SC(AB)$ recibe el nombre de suma de cuadrados de la interacción para A y B , y SCE es la suma de cuadrados del error. La partición de los grados de libertad se efectúa de acuerdo con la identidad

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1).$$

Formación de los cuadrados medios

Si dividimos cada una de las sumas de cuadrados en el lado derecho de la identidad de la suma de cuadrados entre su número correspondiente de grados de libertad, obtenemos los cuatro estadísticos

$$S_1^2 = \frac{SCA}{a - 1}, \quad S_2^2 = \frac{SCB}{b - 1}, \quad S_3^2 = \frac{SC(AB)}{(a - 1)(b - 1)}, \quad S^2 = \frac{SCE}{ab(n - 1)}.$$

Todos estos estimados de la varianza son estimados independientes de σ^2 , siempre que no haya efectos α_i , β_j ni, por supuesto, $(\alpha\beta)_{ij}$. Si las sumas de cuadrados se interpretan

como funciones de las variables aleatorias independientes $y_{111}, y_{112}, \dots, y_{abn}$, no es difícil comprobar que

$$E(S_1^2) = E\left[\frac{SCA}{a-1}\right] = \sigma^2 + \frac{nb}{a-1} \sum_{i=1}^a \alpha_i^2,$$

$$E(S_2^2) = E\left[\frac{SCB}{b-1}\right] = \sigma^2 + \frac{na}{b-1} \sum_{j=1}^b \beta_j^2,$$

$$E(S_3^2) = E\left[\frac{SC(AB)}{(a-1)(b-1)}\right] = \sigma^2 + \frac{n}{(a-1)(b-1)} \sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2,$$

$$E(S^2) = E\left[\frac{SCE}{ab(n-1)}\right] = \sigma^2,$$

a partir de lo cual se observa de inmediato que los 4 estimados de σ^2 son no sesgados cuando $H_0', H_0'',$ y H_0''' son verdaderas.

Para probar la hipótesis H_0' , de que los efectos de los factores A son todos iguales a cero, se calcula la siguiente razón:

Prueba F para
el factor A

$$f_1 = \frac{s_1^2}{s^2},$$

que es un valor de la variable aleatoria F_1 , el cual tiene la distribución F con $a-1$ y $ab(n-1)$ grados de libertad cuando H_0' es verdadera. La hipótesis nula se rechaza al nivel de significancia α cuando $f_1 > f_{\alpha}[a-1, ab(n-1)]$.

De manera similar, para probar la hipótesis H_0'' , de que todos los efectos del factor B son iguales a cero, se calcula la razón:

Prueba F para
el factor B

$$f_2 = \frac{s_2^2}{s^2},$$

que es un valor de la variable aleatoria F_2 que tiene la distribución F con $b-1$ y $ab(n-1)$ grados de libertad cuando H_0'' es verdadera. Esta hipótesis se rechaza al nivel de significancia α cuando $f_2 > f_{\alpha}[b-1, ab(n-1)]$.

Por último, para probar la hipótesis H_0''' , de que todos los efectos de interacción son iguales a 0, se calcula la razón siguiente:

Prueba F para
la interacción

$$f_3 = \frac{s_3^2}{s^2},$$

que es un valor de la variable aleatoria F_3 , el cual tiene la distribución F con $(a-1)(b-1)$ y $ab(n-1)$ grados de libertad cuando H_0''' es verdadera. Concluimos que, a un nivel de significancia α , hay interacción cuando $f_3 > f_{\alpha}[(a-1)(b-1), ab(n-1)]$.

Como se indicó en la sección 14.2, se recomienda interpretar la prueba para la interacción antes de tratar de hacer inferencias sobre los efectos principales. Si la interacción no es significativa, entonces hay evidencia de que las pruebas sobre los efectos principales son interpretables. El rechazo de la hipótesis 1 de la página 566 implica que las medias de la respuesta en los niveles del factor A difieren significativamente, mientras que

el rechazo de la hipótesis 2 implica una condición similar para las medias en los niveles del factor B . Sin embargo, una interacción significativa podría muy bien implicar que los datos se deberían analizar de una manera un poco diferente, quizá observando el efecto del factor A en niveles fijos del factor B , y así sucesivamente.

Los cálculos en un problema de análisis de varianza para un experimento de 2 factores con n réplicas suelen resumirse como se ilustra en la tabla 14.2.

Tabla 14.2: Análisis de varianza para el experimento de 2 factores con n réplicas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Efecto principal				
A	SCA	$a - 1$	$s_1^2 = \frac{SCA}{a-1}$	$f_1 = \frac{s_1^2}{s^2}$
B	SCB	$b - 1$	$s_2^2 = \frac{SCB}{b-1}$	$f_2 = \frac{s_2^2}{s^2}$
Interacciones de 2 factores				
AB	$SC(AB)$	$(a - 1)(b - 1)$	$s_3^2 = \frac{SC(AB)}{(a-1)(b-1)}$	$f_3 = \frac{s_3^2}{s^2}$
Error	SCE	$ab(n - 1)$	$s^2 = \frac{SCE}{ab(n-1)}$	
Total	STC	$abn - 1$		

Ejemplo 14.1: En un experimento realizado para determinar cuál de 3 sistemas de misiles distintos es preferible, se midió la tasa de combustión del propulsor para 24 arranques estáticos. Se emplearon 4 tipos de combustible diferentes y el experimento generó observaciones duplicadas de las tasas de combustión para cada combinación de los tratamientos.

Los datos, ya codificados, se presentan en la tabla 14.3. Pruebe las siguientes hipótesis: $a) H_0'$: no hay diferencia en las tasas medias de combustión del propulsor cuando se emplean diferentes sistemas de misiles, $b) H_0''$: no existe diferencia en las tasas medias de combustión de los 4 tipos de propulsor, $c) H_0'''$: no hay interacción entre los distintos sistemas de misiles y los diferentes tipos de propulsor.

Tabla 14.3: Tasas de combustión del propulsor

Sistema de misiles	Tipo de propulsor			
	b_1	b_2	b_3	b_4
a_1	34.0	30.1	29.8	29.0
	32.7	32.8	26.7	28.9
a_2	32.0	30.2	28.7	27.6
	33.2	29.8	28.1	27.8
a_3	28.4	27.3	29.7	28.8
	29.3	28.9	27.3	29.1

- Solución:** 1. $a) H_0'$: $\alpha_1 = \alpha_2 = \alpha_3 = 0$.
 $b) H_0''$: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$.
 $c) H_0'''$: $(\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{34} = 0$.

2. a) H'_1 : Al menos una de las α_i no es igual a 0.
 b) H''_1 : Al menos una de las β_j no es igual a 0.
 c) H'''_1 : Al menos una de las $(\alpha\beta)_{ij}$ no es igual a 0.

Se utiliza la fórmula de la suma de cuadrados que se describió en el teorema 14.1. En la tabla 14.4 se presenta el análisis de varianza.

Tabla 14.4: Análisis de varianza para los datos de la tabla 14.3

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Sistema de misiles	14.52	2	7.26	5.84
Tipo de propulsor	40.08	3	13.36	10.75
Interacción	22.16	6	3.69	2.97
Error	14.91	12	1.24	
Total	91.68	23		

Se remite al lector al procedimiento de modelos lineales generales (GLM, por sus siglas en inglés) del SAS para el análisis de los datos de la tasa de combustión de la figura 14.2. Observe la forma en que al principio se prueba el "modelo" (11 grados de libertad), y por separado se prueban el sistema, el tipo y el sistema por tipo de interacción. La prueba f en el modelo ($P = 0.0030$) prueba la acumulación de los 2 efectos principales y la interacción.

- a) Rechace H'_0 y concluya que los distintos sistemas de misiles resultan en diferentes tasas medias de combustión del propulsor. El valor P es de aproximadamente 0.0169.
 b) Rechace H''_0 y concluya que las tasas medias de combustión del propulsor no son las mismas para los 4 tipos de propulsores. El valor P es de aproximadamente 0.0010.
 c) La interacción es apenas insignificante al nivel 0.05, pero el valor P de aproximadamente 0.0513 indicaría que la interacción debe tomarse en serio.

En este momento debemos hacer algún tipo de interpretación de la interacción. Debe destacarse que la significancia estadística de un efecto principal tan sólo implica que las *medias marginales son significativamente diferentes*. Sin embargo, considere la tabla de promedios de 2 factores de la tabla 14.5.

Tabla 14.5: Interpretación de la interacción

	b_1	b_2	b_3	b_4	Promedio
a_1	33.35	31.45	28.25	28.95	30.50
a_2	32.60	30.00	28.40	27.70	29.68
a_3	28.85	28.10	28.50	28.95	28.60
Promedio	31.60	29.85	28.38	28.53	

Es evidente que hay más información importante en el cuerpo de la tabla, tendencias que son inconsistentes con la tendencia que describe los promedios marginales. La tabla 14.5 sugiere con certeza que el efecto del tipo de propulsor depende del sistema que se utiliza. Por ejemplo, para el sistema 3, el efecto del tipo de propulsor no parece ser

The GLM Procedure					
Dependent Variable: rate					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	76.76833333	6.97893939	5.62	0.0030
Error	12	14.91000000	1.24250000		
Corrected Total	23	91.67833333			
R-Square	Coeff Var	Root MSE	rate Mean		
0.837366	3.766854	1.114675	29.59167		
Source	DF	Type III SS	Mean Square	F Value	Pr > F
system	2	14.52333333	7.26166667	5.84	0.0169
type	3	40.08166667	13.36055556	10.75	0.0010
system*type	6	22.16333333	3.69388889	2.97	0.0512

Figura 14.2: Salida de resultados del SAS para el análisis de los datos de la tasa de combustión del propulsor de la tabla 14.3.

importante, aunque tiene un efecto grande si se utiliza el sistema 1 o el 2. Esto explica la interacción "significativa" entre esos 2 factores. Más adelante se revelará más información acerca de esta interacción. ▮

Ejemplo 14.2: Remítase al ejemplo 14.1 y elija 2 contrastes ortogonales para dividir la suma de cuadrados del sistema de misiles en componentes con un solo grado de libertad, los cuales utilizará para comparar los sistemas 1 y 2 con el 3, y el sistema 1 contra el sistema 2.

Solución: El contraste para comparar los sistemas 1 y 2 con el 3 es

$$w_1 = \mu_1 + \mu_2 - 2\mu_3.$$

Un segundo contraste, ortogonal a w_1 , para comparar el sistema 1 con el 2, es dado por $w_2 = \mu_1 - \mu_2$. Las sumas de cuadrados con un solo grado de libertad son

$$SC_{w_1} = \frac{[244.0 + 237.4 - (2)(228.8)]^2}{(8)[(1)^2 + (1)^2 + (-2)^2]} = 11.80$$

y

$$SC_{w_2} = \frac{(244.0 - 237.4)^2}{(8)[(1)^2 + (-1)^2]} = 2.72.$$

Observe que $SC_{w_1} + SC_{w_2} = SCA$, como se esperaba. Los valores f calculados correspondientes a w_1 y w_2 son, respectivamente,

$$f_1 = \frac{11.80}{1.24} = 9.5 \quad \text{y} \quad f_2 = \frac{2.72}{1.24} = 2.2.$$

Al comparar con el valor crítico $f_{0.05}(1, 12) = 4.75$, se encuentra que f_1 es significativo. De hecho, el valor P es menor que 0.01. Así, el primer contraste indica que se rechaza la hipótesis

$$H_0: \frac{1}{2}(\mu_1 + \mu_2) = \mu_3.$$

Como $f_2 < 4.75$, las tasas medias de combustión del primer y segundo sistemas no son significativamente diferentes. J

Impacto de la interacción significativa en el ejemplo 14.1

Si la hipótesis de que no hay interacción en el ejemplo 14.1 es verdadera, podríamos hacer las comparaciones *generales* del ejemplo 14.2 relacionado con los sistemas de misiles, en lugar de comparaciones separadas para cada propulsor. De manera similar, se podrían realizar comparaciones generales entre los propulsores, en vez de comparar por separado cada sistema de misiles. Por ejemplo, se podrían comparar los propulsores 1 y 2 con el 3 y 4, y también el 1 contra el 2. Las razones f resultantes, cada una con 1 y 12 grados de libertad, resultan ser de 24.81 y 7.39, respectivamente, y ambas son muy significativas al nivel 0.05.

Por los promedios de los propulsores, parece haber evidencia de que el 1 ofrece la tasa media de combustión más alta. Un experimentador prudente sería cauteloso al sacar conclusiones generales en un problema como éste, donde la razón f de la interacción está apenas por debajo del valor crítico de 0.05. Por ejemplo, la evidencia general, 31.60 contra 29.85 sobre el promedio para los 2 propulsores, indica con claridad que el 1 es superior al 2, en términos de una mayor tasa de combustión. Sin embargo, si nos restringimos al sistema 3, donde tenemos un promedio de 28.85 para el propulsor 1 en oposición a 28.10 para el propulsor 2, parece haber una diferencia mínima o incluso ninguna entre estos 2 propulsores. De hecho, parece que hay una estabilización de las tasas de combustión para los distintos propulsores si se opera con el sistema 3. Es claro que existe evidencia general que indica que el sistema 1 ofrece una tasa de combustión más alta que el sistema 3, pero parece que esta conclusión no se sostiene si nos restringimos al propulsor 4.

Para recabar evidencias concluyentes de que la interacción está *produciendo dificultades considerables en la obtención de conclusiones generales sobre los efectos principales*, el analista puede hacer una prueba t sencilla utilizando las tasas de combustión promedio del sistema 3. Considere una comparación del propulsor 1 contra el 2 usando únicamente el sistema 3. Se toma prestado un estimado de σ^2 del análisis general, es decir, se utiliza $s^2 = 1.24$ con 12 grados de libertad, y se obtiene

$$|t| = \frac{0.75}{\sqrt{2s^2/n}} = \frac{0.75}{\sqrt{1.24}} = 0.67,$$

que no se acerca a ser significativa. Esta ilustración sugiere que, en presencia de interacción, debería tenerse cautela con la interpretación estricta de los efectos principales.

Análisis gráfico para el problema de dos factores del ejemplo 14.1

Muchos de los mismos tipos de ilustraciones gráficas que se sugirió emplear en los problemas de un factor también se aplican en el caso de 2 factores. Las gráficas en 2 dimensiones de las medias de las celdas o de las medias de las combinaciones de tratamientos ofrecen información sobre la presencia de interacciones entre los 2 factores.

Además, una gráfica de los residuales contra los valores ajustados bien podría indicar si se cumple o no la suposición de la varianza homogénea. Por supuesto, es frecuente que una violación de la suposición de varianza homogénea implique un aumento en la varianza del error conforme *los números de la respuesta se vuelven más grandes*. Como resultado, esta gráfica podría resaltar la violación.

La figura 14.3 presenta la gráfica de las medias de las celdas para el caso del propulsor de los sistemas de misiles del ejemplo 14.1. Observe gráficamente (en este caso) cuánta falta de paralelismo hay. Note el aplanamiento de la parte de la figura que indica el efecto del propulsor para el sistema 3. Esto ilustra la interacción entre los factores. La figura 14.4 muestra la gráfica de los residuales contra los valores ajustados para los mismos datos. Al parecer no hay dificultades con la suposición de la varianza homogénea.

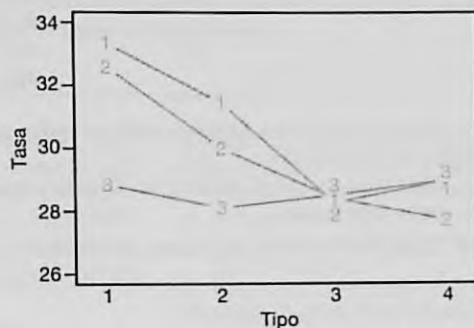


Figura 14.3: Gráfica de las medias de las celdas para los datos del ejemplo 14.1. Los números representan los sistemas de misiles.

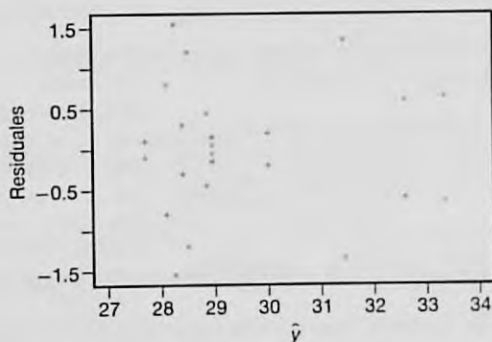


Figura 14.4: Gráfica de los residuales de los datos del ejemplo 14.1.

Ejemplo 14.3: Un ingeniero eléctrico investiga un proceso de grabado con plasma que se emplea en la fabricación de semiconductores. Es de interés estudiar los efectos de 2 factores, la cantidad de flujo (A) del gas C_2F_6 y la potencia aplicada al cátodo (B). La respuesta es la velocidad de grabado. Cada factor se aplica a 3 niveles y se hacen 2 corridas experimentales sobre la velocidad de grabado para cada una de las 9 combinaciones. El plan representa un diseño completamente aleatorizado. En la tabla 14.6 se presentan los datos. La velocidad de grabado se expresa en A°/min .

Tabla 14.6: Datos para el ejemplo 14.3

Cantidad de flujo del C_2F_6	Potencia suministrada		
	1	2	3
1	288	488	670
	360	465	720
2	385	482	692
	411	521	724
3	488	595	761
	462	612	801

Los niveles de los factores están en orden ascendente, donde el nivel 1 es el más bajo y el 3 el más alto.

a) Elabore una tabla de análisis de varianza y saque conclusiones; empiece con la prueba de interacción.

b) Haga pruebas sobre los efectos principales y saque conclusiones.

Solución: En la figura 14.5 se muestra una salida de resultados por computadora del SAS. De ese listado se concluye lo siguiente.

```

The GLM Procedure
Dependent Variable: etchrate

              Sum of
Source          DF          Squares    Mean Square    F Value    Pr > F
Model              8    379508.7778    47438.5972     61.00    <.0001
Error              9     6999.5000     777.7222
Corrected Total   17    386508.2778

R-Square          Coeff Var          Root MSE    etchrate Mean
0.981890          5.057714          27.88767     551.3889

Source          DF    Type III SS    Mean Square    F Value    Pr > F
c2f6              2    46343.1111    23171.5556     29.79    0.0001
power            2    330003.4444    165001.7222    212.16    <.0001
c2f6*power        4     3162.2222     790.5556       1.02    0.4485

```

Figura 14.5: Una salida de resultados por computadora del SAS para el ejemplo 14.3.

a) El valor P para la prueba de interacción es 0.4485. Se concluye que la interacción no es significativa.

b) Existe una diferencia significativa en la velocidad media de grabado para los 3 niveles de la velocidad de flujo del C_2F_6 . Una prueba de Duncan muestra que la velocidad media de grabado para el nivel 3 es significativamente mayor que para el nivel 2, y la

velocidad para el nivel 2 es significativamente mayor que para el nivel 1. Véase la figura 14.6a.

Existe una diferencia significativa en la velocidad media de grabado basada en el nivel de potencia al cátodo. Una prueba de Duncan revela que la velocidad de grabado para el nivel 3 es significativamente más alta que para el 2, y que la velocidad para el nivel 2 es significativamente más alta que para el 1. Véase la figura 14.6b.

Duncan Grouping	Mean	N	c2f6	Duncan Grouping	Mean	N	power
A	619.83	6	3	A	728.00	6	3
B	535.83	6	2	B	527.17	6	2
C	498.50	6	1	C	399.00	6	1
(a)				(b)			

Figura 14.6: Una salida de resultados por computadora del SAS para el ejemplo 14.3. a) Prueba de Duncan de la cantidad de flujo del gas; b) Prueba de Duncan de la potencia.

Ejercicios

14.1 Se realizó un experimento para estudiar los efectos de la temperatura y el tipo de horno sobre la vida de un componente en particular. En el experimento se utilizaron 4 tipos de horno y 3 niveles de temperatura. Se asignaron 24 piezas al azar, 2 para cada combinación de tratamientos y se registraron los siguientes resultados.

Temperatura (°F)	Horno			
	O ₁	O ₂	O ₃	O ₄
500	227	214	225	260
	221	259	236	229
550	187	181	232	246
	208	179	198	273
600	174	198	178	206
	202	194	213	219

A un nivel de significancia de 0.05 pruebe las hipótesis de que

- las diferentes temperaturas no tienen efecto en la vida del componente;
- los diferentes hornos no tienen efecto en la vida del componente;
- no hay interacción entre el tipo de horno y la temperatura.

14.2 El Departamento de Nutrición Humana y Alimentos de la Virginia Tech realizó un estudio titulado *Vitamin C Retention in Reconstituted Frozen Orange Juice* sobre la estabilidad de la vitamina C en el concentrado de jugo de naranja congelado reconstituido y almacenado en un refrigerador durante un periodo de hasta una semana. Se probaron 3 tipos de concentrado de jugo de naranja congelado en 3 periodos distintos, los cuales se refieren al número de días transcurridos desde que se mezcló el jugo hasta que se probó. Se

registraron los resultados en miligramos de ácido ascórbico por litro. Utilice un nivel de significancia de 0.05 para probar las hipótesis de que

- no hay diferencias en el contenido de ácido ascórbico entre las diferentes marcas de concentrado de jugo de naranja;
- no existen diferencias en el contenido de ácido ascórbico para distintos periodos;
- no hay interacción entre las marcas de concentrado de jugo de naranja y el número de días transcurridos desde que el jugo se mezcló hasta que se probó.

Marca	Tiempo (días)					
	0		3		7	
Richfood	52.6	54.2	49.4	49.2	42.7	48.8
	49.8	46.5	42.8	53.2	40.4	47.6
Sealed-Sweet	56.0	48.0	48.8	44.0	49.2	44.0
	49.6	48.4	44.0	42.4	42.0	43.2
Minute Maid	52.5	52.0	48.0	47.0	48.5	43.3
	51.8	53.6	48.2	49.6	45.2	47.6

14.3 Se estudió el desempeño de 3 cepas de ratas en una prueba de laberintos en 2 condiciones ambientales diferentes. Se registraron las puntuaciones de error de las 48 ratas:

Ambiente	Cepa					
	Brillante		Mezclada		Torpe	
Libre	28	12	33	83	101	94
	22	23	36	14	33	56
	25	10	41	76	122	83
	36	86	22	58	35	23
Restringido	72	32	60	89	136	120
	48	93	35	126	38	153
	25	31	83	110	64	128
	91	19	99	118	87	140

Utilice un nivel de significancia de 0.01 para probar las hipótesis de que

- no hay diferencia en las puntuaciones de error para ambientes diferentes;
- no existe diferencia en las puntuaciones de error para cepas diferentes;
- no hay interacción entre los ambientes y las cepas de las ratas.

14.4 La fatiga por corrosión de los metales se define como la acción simultánea de tensión cíclica y ataque químico sobre una estructura metálica. Una técnica muy utilizada para minimizar el daño de la fatiga por corrosión en el aluminio requiere la aplicación de un recubrimiento protector. En un estudio efectuado por el Departamento de Ingeniería Mecánica de Virginia Tech se utilizaron 3 niveles diferentes de humedad:

Bajo: 20 a 25% de humedad relativa

Medio: 55 a 60% de humedad relativa

Alto: 86 a 91% de humedad relativa

y 3 tipos de recubrimiento:

No revestido: Sin recubrimiento

Anodizado: Recubrimiento de óxido anódico por ácido sulfúrico

Conversión: Recubrimiento por conversión química de cromato.

Los datos de fatiga por corrosión, expresados en miles de ciclos hasta que se presenta la falla, se registraron como sigue:

Recubrimiento	Humedad relativa					
	Baja		Media		Alta	
No revestido	361	469	314	522	1344	1216
	466	937	244	739	1027	1097
	1069	1357	261	134	1011	1011
Anodizado	114	1032	322	471	78	466
	1236	92	306	130	387	107
	533	211	68	398	130	327
Conversión	130	1482	252	874	586	524
	841	529	105	755	402	751
	1595	754	847	573	846	529

- Lleve a cabo un análisis de varianza con $\alpha = 0.05$ para probar si existen efectos principales y efectos de interacción.
- Utilice la prueba de Duncan de rango múltiple a un nivel de significancia de 0.05 para determinar cuáles niveles de humedad relativa dan como resultado daños distintos de fatiga por corrosión.

14.5 Para determinar cuáles músculos necesitan someterse a un programa de acondicionamiento para mejorar el rendimiento individual en el servicio tendido que se usa en el tenis, el Departamento de Salud, Educación Física y Recreación de Virginia Tech realizó

un estudio de 5 músculos diferentes:

- 1: deltoides anterior
- 2: pectoral mayor
- 3: deltoides posterior
- 4: deltoides medio
- 5: tríceps

los cuales se probaron en cada uno de 3 sujetos; el experimento se efectuó 3 veces para cada combinación de tratamiento. Los datos electromiográficos que se registraron durante el servicio se presentan a continuación.

Sujeto	Músculo				
	1	2	3	4	5
1	32	5	58	10	19
	59	1.5	61	10	20
	38	2	66	14	23
2	63	10	64	45	43
	60	9	78	61	61
	50	7	78	71	42
3	43	41	26	63	61
	54	43	29	46	85
	47	42	23	55	95

Utilice un nivel de significancia de 0.01 para probar las hipótesis de que

- diferentes sujetos tienen medidas electromiográficas iguales;
- los diferentes músculos no tienen un efecto en las medidas electromiográficas;
- no hay interacción entre los sujetos y los tipos de músculos.

14.6 Se realizó un experimento para determinar si los aditivos incrementan la adherencia de productos de caucho. Se elaboraron 16 productos con el aditivo nuevo y otros 16 sin dicho aditivo. Se registró la siguiente adherencia.

	Temperatura (°C)			
	50	60	70	80
Sin el aditivo	2.3	3.4	3.8	3.9
	2.9	3.7	3.9	3.2
	3.1	3.6	4.1	3.0
	3.2	3.2	3.8	2.7
Con el aditivo	4.3	3.8	3.9	3.5
	3.9	3.8	4.0	3.6
	3.9	3.9	3.7	3.8
	4.2	3.5	3.6	3.9

Haga un análisis de varianza para probar la existencia de efectos principales y de interacción significativos.

14.7 Se sabe que la velocidad de extracción de cierto polímero depende de la temperatura de reacción y de la cantidad de catalizador empleada. Se hizo un experimento en 4 niveles de temperatura y 5 niveles de catalizador, y se registró la velocidad de extracción en la siguiente tabla:

Cantidad de catalizador

	0.5%	0.6%	0.7%	0.8%	0.9%
50°C	38	45	57	59	57
	41	47	59	61	58
60°C	44	56	70	73	61
	43	57	69	72	58
70°C	44	56	70	73	61
	47	60	67	61	59
80°C	49	62	70	62	53
	47	65	55	69	58

Realice un análisis de varianza. Pruebe si hay efectos principales y de interacción significativos.

14.8 En Myers, Montgomery y Anderson-Cook (2009) se estudia un escenario donde se describe un proceso de laminado por prensado. La respuesta es el espesor del material. Los factores que podrían afectar el espesor incluyen la cantidad de níquel (A) y el pH (B). Se diseñó un experimento con 2 factores. El plan es un diseño completamente aleatorizado en el que las prensas individuales se asignan al azar a las combinaciones de factores. En el experimento se utilizan 3 niveles de pH y 2 niveles de contenido de níquel. Los espesores, en $\text{cm} \times 10^{-3}$, son los siguientes:

Contenido de níquel (gramos)	pH		
	6	5.5	6
18	250	211	221
	195	172	150
	188	165	170
10	115	88	69
	165	112	101
	142	108	72

- Elabore la tabla del análisis de varianza con pruebas para los efectos principales y de interacción. Incluya valores P .
- Saque conclusiones para ingeniería. ¿Qué aprendió del análisis de estos datos?
- Elabore una gráfica que ilustre la presencia o ausencia de interacción.

14.9 Un ingeniero está interesado en los efectos de la velocidad de corte y la geometría de la herramienta sobre las horas de vida de una máquina-herramienta. Se utilizan 2 velocidades de corte y 2 geometrías distintas. Se llevan a cabo 3 pruebas experimentales con cada una de las 4 combinaciones. Los datos son los siguientes:

Geometría de la herramienta	Velocidad de corte			
	Baja		Alta	
1	22	28	20	34
2	18	15	16	11

- Calcule la tabla del análisis de varianza con pruebas sobre los efectos principales y de interacción.
- Haga comentarios sobre el efecto que tiene la interacción sobre la prueba de la velocidad de corte.

- Efectúe pruebas secundarias que permitan al ingeniero conocer el verdadero impacto de la velocidad de corte.
- Construya una gráfica que ilustre el efecto de interacción.

14.10 En un experimento se estudiaron 2 factores de un proceso de manufactura de un circuito integrado. El propósito del experimento es conocer su efecto sobre la resistividad de las obleas de silicio. Los factores son la dosis del implante (2 niveles) y la posición de la caldera (3 niveles). El experimento es costoso, por lo que sólo se hizo una corrida con cada combinación. Los datos son los siguientes.

Dosis	Posición			
	1	15.5	14.8	21.3
2	27.2	24.9	26.1	

Se supone que no hay interacción entre esos 2 factores.

- Escriba el modelo y explique sus términos.
- Elabore la tabla de análisis de varianza.
- Explique los 2 grados de libertad del "error".
- Use una prueba de Tukey para hacer pruebas de comparaciones múltiples sobre la posición de la caldera. Explique qué es lo que muestran los resultados.

14.11 Se realizó un estudio para determinar la influencia de 2 factores, el método de análisis y el laboratorio que hace el análisis, sobre el nivel de contenido de azufre del carbón. Se asignaron al azar 28 especímenes de carbón a 14 combinaciones de factores, la estructura de las unidades experimentales representada por las combinaciones de 7 laboratorios y 2 métodos de análisis con 2 especímenes por combinación de factores. Los datos, expresados en porcentaje de azufre, son los siguientes.

Laboratorio	Método			
	1		2	
1	0.109	0.105	0.105	0.108
2	0.129	0.122	0.127	0.124
3	0.115	0.112	0.109	0.111
4	0.108	0.108	0.117	0.118
5	0.097	0.096	0.110	0.097
6	0.114	0.119	0.116	0.122
7	0.155	0.145	0.164	0.160

(Los datos se tomaron de G. Taguchi, "Signal to Noise Ratio and Its Applications to Testing Material", *Reports of Statistical Application Research*, Union of Japanese Scientists and Engineers, Vol. 18, Núm. 4, 1971).

- Haga un análisis de varianza y exprese los resultados en la tabla correspondiente.
- ¿Es significativa la interacción? Si lo es, analice lo que significa para el científico. Utilice un valor P en sus conclusiones.

- c) ¿Son estadísticamente significativos los efectos principales individuales, el laboratorio y el método de análisis? Analice la información y lo que aprendió y base su respuesta en el contexto de cualquier interacción significativa.
- d) Dibuje una gráfica de interacción que ilustre el efecto de la interacción.
- e) Efectúe una prueba para comparar los métodos 1 y 2 en el laboratorio 1, y haga lo mismo para el laboratorio 7. Comente lo que revelan esos resultados.

14.12 En un experimento efectuado en el departamento de Ingeniería Civil de Virginia Tech se observó el crecimiento que cierto tipo de alga tenía en el agua, en función del tiempo y la dosis de cobre que se agregaba al líquido. Los datos se presentan a continuación. La respuesta se expresa en unidades de algas.

Cobre	Tiempo en días		
	5	12	18
1	0.30	0.37	0.25
	0.34	0.36	0.23
	0.32	0.35	0.24
2	0.24	0.30	0.27
	0.23	0.32	0.25
	0.22	0.31	0.25
3	0.20	0.30	0.27
	0.28	0.31	0.29
	0.24	0.30	0.25

- a) Haga un análisis de varianza y elabore la tabla correspondiente.
- b) Comente acerca de si los datos son suficientes para mostrar un efecto del tiempo sobre la concentración de algas.
- c) Haga lo mismo para el contenido de cobre. ¿El nivel de contenido de cobre tiene algún efecto sobre la concentración de algas?
- d) Comente los resultados de la prueba de interacción. ¿Cómo influye el tiempo sobre el efecto del contenido de cobre?

14.13 En *Classical and Modern Regression with Applications* (Duxbury Classic Series, 2a. ed., 1990), de Myers, se describe un experimento en el que la Agencia de Protección Ambiental busca determinar el efecto de 2 métodos de tratamiento de aguas sobre la absorción del magnesio. Se miden los niveles de magnesio, en gramos por centímetro cúbico (cc) y se incorporan 2 niveles diferentes de tiempo al experimento. Los datos son los siguientes:

Tiempo (horas)	Tratamiento	
	1	2
1	2.19 2.15 2.16	2.03 2.01 2.04
2	2.01 2.03 2.04	1.88 1.86 1.91

- a) Dibuje una gráfica de la interacción. ¿Cuál es su impresión?

- b) Efectúe un análisis de varianza y presente pruebas para los efectos principales y de interacción.
- c) Mencione los hallazgos científicos acerca de cómo influyen el tiempo y el tratamiento en la absorción del magnesio.
- d) Ajuste el modelo de regresión adecuado usando el tratamiento como variable categórica. Incluya la interacción en el modelo.
- e) ¿La interacción es significativa en el modelo de regresión?

14.14 Considere los datos del ejercicio 14.12 y responda las siguientes preguntas.

- a) Ambos factores, el cobre y el tiempo, son cuantitativos. Como resultado, podría ser de interés un modelo de regresión. Describa cuál sería un modelo adecuado si se usa x_1 = contenido de cobre y x_2 = tiempo. Ajuste el modelo a los datos mostrando los coeficientes de regresión y haga una prueba t para cada uno.
- b) Ajuste el modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \epsilon,$$

y compárelo con el que eligió en el inciso a. ¿Cuál es más apropiado? Como criterio utilice R_{ajust}^2 .

14.15 El propósito del estudio *The Incorporation of a Chelating Agent into a Flame Retardant Finish of a Cotton Flannelette and the Evaluation of Selected Fabric Properties*, llevado a cabo en Virginia Tech, fue evaluar el uso de un agente quelante como parte del acabado retardante del fuego de la franela de algodón, determinando sus efectos en la inflamabilidad después de lavar la tela en condiciones específicas. Se utilizaron 2 tratamientos con 2 niveles; se prepararon 2 baños, uno con celulosa de carboximetilo (baño I) y otro sin ella (baño II). La mitad de la tela se lavó 5 veces y la otra mitad se lavó 10 veces. Hubo 12 pedazos de tela en cada combinación de baño/número de lavados. Después de los lavados se midieron las longitudes quemadas de la tela, así como los tiempos de combustión. Se registraron los siguientes tiempos de combustión (en segundos):

Lavados	Baño I			Baño II		
	5	13.7	23.0	15.7	6.2	5.4
	25.5	15.8	14.8	4.4	5.0	3.3
	14.0	29.4	9.7	16.0	2.5	1.6
	14.0	12.3	12.3	3.9	2.5	7.1
10	27.2	16.8	12.9	18.2	8.8	14.5
	14.9	17.1	13.0	14.7	17.1	13.9
	10.8	13.5	25.5	10.6	5.8	7.3
	14.2	27.4	11.5	17.7	18.3	9.9

- a) Realice un análisis de varianza. ¿Existe un término de interacción significativo?
- b) ¿Se encontraron diferencias en los efectos principales? Analice la información.

14.4 Experimentos de tres factores

En esta sección consideramos un experimento con 3 factores, A , B y C , en los niveles a , b y c , respectivamente, en un diseño experimental completamente aleatorizado. Suponga de nuevo que se tienen n observaciones para cada una de las abc combinaciones de tratamientos. Debemos proceder a realizar las pruebas de significancia para los 3 efectos principales y las interacciones implicadas. Se espera que el lector podrá utilizar después esta descripción para generalizar el análisis a $k > 3$ factores.

Modelo para el experimento de tres factores

El modelo para el experimento de 3 factores es

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$$

$i = 1, 2, \dots, a$; $j = 1, 2, \dots, b$; $k = 1, 2, \dots, c$; y $l = 1, 2, \dots, n$, donde α_i , β_j y γ_k son los efectos principales y $(\alpha\beta)_{ij}$, $(\alpha\gamma)_{ik}$ y $(\beta\gamma)_{jk}$ son los efectos de la interacción de 2 factores que tienen la misma interpretación que en el experimento con 2 factores.

El término $(\alpha\beta\gamma)_{ijk}$ se denomina **efecto de interacción de 3 factores**, y representa la no aditividad de las $(\alpha\beta)_{ij}$ sobre los diferentes niveles del factor C . Igual que antes, la suma de todos los efectos principales es igual a 0, y la suma sobre cualesquiera de los subíndices de los efectos de la interacción entre 2 y 3 factores es igual a 0. En muchas situaciones experimentales estas interacciones de orden superior son insignificantes y sus cuadrados medios sólo reflejan variación aleatoria; pero se debe describir el análisis en su forma más general.

Nuevamente, para realizar pruebas válidas de significancia debe suponerse que los errores son valores de variables aleatorias independientes y con distribución normal, cada una con media igual a 0 y varianza común σ^2 .

La filosofía general respecto al análisis es la misma que la que se estudió para los experimentos de 1 y 2 factores. La suma de cuadrados se divide en 8 términos, donde cada uno representa una fuente de variación de los que se obtienen estimados independientes de σ^2 cuando todos los efectos principales y de la interacción son iguales a 0. Si los efectos de cualquier factor dado o interacción no son iguales a 0, entonces el cuadrado medio estimará la varianza del error más un componente debido al efecto sistemático en cuestión.

Suma de cuadrados para un experimento de tres factores

$$SCA = bcn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 \quad SC(AB) = cn \sum_i \sum_j (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$

$$SCB = acn \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{...})^2 \quad SC(AC) = bn \sum_i \sum_k (\bar{y}_{ik.} - \bar{y}_{i..} - \bar{y}_{.k.} + \bar{y}_{...})^2$$

$$SCC = abn \sum_{k=1}^c (\bar{y}_{..k} - \bar{y}_{...})^2 \quad SC(BC) = an \sum_j \sum_k (\bar{y}_{.jk} - \bar{y}_{.j.} - \bar{y}_{.k.} + \bar{y}_{...})^2$$

$$SC(ABC) = n \sum_i \sum_j \sum_k (\bar{y}_{ijk} - \bar{y}_{ij.} - \bar{y}_{i.k} - \bar{y}_{.jk} + \bar{y}_{i..} + \bar{y}_{.j.} + \bar{y}_{.k.} - \bar{y}_{...})^2$$

$$STC = \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{...})^2 \quad SCE = \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{ijk})^2$$

Aunque en esta sección hacemos énfasis en la interpretación de una salida de resultados por computadora con comentarios, en vez de preocuparnos por cálculos laboriosos de sumas de cuadrados, ofrecemos lo siguiente como la suma de cuadrados para los 3 efectos principales y las interacciones. Observe la evidente extensión del problema de 2 factores a uno de 3.

Los promedios en las fórmulas se definen como sigue:

$\bar{y}_{..}$ = promedio de todas las *abcn* observaciones,

$\bar{y}_{i..}$ = promedio de las observaciones para el *i*-ésimo nivel del factor A,

$\bar{y}_{.j.}$ = promedio de las observaciones para el *j*-ésimo nivel del factor B,

$\bar{y}_{..k}$ = promedio de las observaciones para el *k*-ésimo nivel del factor C,

$\bar{y}_{ij.}$ = promedio de las observaciones para el *i*-ésimo nivel de A y el *j*-ésimo nivel de B,

$\bar{y}_{i.k}$ = promedio de las observaciones para el *i*-ésimo nivel de A y el *k*-ésimo nivel de C,

$\bar{y}_{.jk}$ = promedio de las observaciones para el *j*-ésimo nivel de B y el *k*-ésimo nivel de C,

\bar{y}_{ijk} = promedio de las observaciones para la (*ijk*)-ésima combinación de tratamientos.

Los cálculos en una tabla de análisis de varianza para un problema de 3 factores con *n* réplicas de corridas para cada combinación de factores se resumen en la tabla 14.7.

Tabla 14.7: ANOVA para el experimento de 3 factores con *n* réplicas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	<i>f</i> calculada
Efecto principal:				
A	SCA	<i>a</i> - 1	s_1^2	$f_1 = \frac{s_1^2}{s^2}$
B	SCB	<i>b</i> - 1	s_2^2	$f_2 = \frac{s_2^2}{s^2}$
C	SCC	<i>c</i> - 1	s_3^2	$f_3 = \frac{s_3^2}{s^2}$
Interacción de 2 factores:				
AB	SC(AB)	(<i>a</i> - 1)(<i>b</i> - 1)	s_4^2	$f_4 = \frac{s_4^2}{s^2}$
AC	SC(AC)	(<i>a</i> - 1)(<i>c</i> - 1)	s_5^2	$f_5 = \frac{s_5^2}{s^2}$
BC	SC(BC)	(<i>b</i> - 1)(<i>c</i> - 1)	s_6^2	$f_6 = \frac{s_6^2}{s^2}$
Interacción de 3 factores:				
ABC	SC(ABC)	(<i>a</i> - 1)(<i>b</i> - 1)(<i>c</i> - 1)	s_7^2	$f_7 = \frac{s_7^2}{s^2}$
Error	SCE	<i>abcn</i> - 1	s^2	
Total	STC	<i>abcn</i> - 1		

Para el experimento de 3 factores con una sola corrida experimental por combinación se podría utilizar el análisis de la tabla 14.7 con *n* = 1 y usando la suma de cuadrados de la interacción ABC para SCE. En este caso suponemos que los efectos de la interacción $(\alpha\beta\gamma)_{ijk}$ son todos iguales a cero, de modo que

$$E \left[\frac{SC(ABC)}{(a-1)(b-1)(c-1)} \right] = \sigma^2 + \frac{n}{(a-1)(b-1)(c-1)} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c (\alpha\beta\gamma)_{ijk}^2 = \sigma^2.$$

Es decir, $SC(ABC)$ representa la variación que sólo se debe al error experimental. Por lo tanto, su cuadrado medio proporciona un estimado no sesgado de la varianza del error. Con $n = 1$ y $SCE = SC(ABC)$, la suma de cuadrados del error se obtiene restando la suma de cuadrados de los efectos principales y las interacciones de 2 factores a la suma de cuadrados total.

Ejemplo 14.4: En la producción de un material en particular hay 3 variables de interés: A , el efecto del operador (3 operadores); B , el catalizador utilizado en el experimento (3 catalizadores); y C , el tiempo de lavado del producto después del proceso de enfriamiento (15 minutos y 20 minutos). Se realizaron 3 corridas con cada combinación de factores. Se consideró que debían estudiarse todas las interacciones entre los factores. En la tabla 14.8 se presentan los productos codificados. Realice un análisis de varianza para probar si existen efectos significativos.

Tabla 14.8: Datos para el ejemplo 14.4

Operador, A	Tiempo de lavado, C					
	15 minutos			20 minutos		
	Catalizador, B			Catalizador, B		
	1	2	3	1	2	3
1	10.7	10.3	11.2	10.9	10.5	12.2
	10.8	10.2	11.6	12.1	11.1	11.7
	11.3	10.5	12.0	11.5	10.3	11.0
2	11.4	10.2	10.7	9.8	12.6	10.8
	11.8	10.9	10.5	11.3	7.5	10.2
	11.5	10.5	10.2	10.9	9.9	11.5
3	13.6	12.0	11.1	10.7	10.2	11.9
	14.1	11.6	11.0	11.7	11.5	11.6
	14.5	11.5	11.5	12.7	10.9	12.2

Solución: La tabla 14.9 muestra el análisis de varianza de los datos. Ninguna de las interacciones muestra un efecto significativo a un nivel $\alpha = 0.05$. Sin embargo, el valor P para BC es 0.0610, de modo que no debe ignorarse. Los efectos del operador y el catalizador son significativos, en tanto que el del tiempo de lavado no lo es. ■

Impacto de la interacción BC

Se deben analizar otros aspectos del ejemplo 14.4, en particular acerca del manejo del efecto que la interacción entre el catalizador y el tiempo de lavado tienen sobre la prueba del efecto principal del tiempo de lavado (factor C). Recuerde el análisis de la sección 14.2. Se proporcionaron ejemplos de la manera en que la presencia de la interacción podría cambiar la interpretación que se da a los efectos principales. En el ejemplo 14.4 la interacción BC es significativa aproximadamente al nivel 0.06. No obstante, suponga que se obtiene una tabla de medias de 2 factores como la 14.10.

Queda claro por qué el tiempo de lavado no fue significativo. Un analista poco cuidadoso se quedaría con la impresión de que el tiempo de lavado podría eliminarse de cualquier estudio futuro en el que se mida el producto. Sin embargo, es notorio cómo cambia el efecto del tiempo de lavado de uno negativo para el primer catalizador, a lo

Tabla 14.9: ANOVA para un experimento de 3 factores en un diseño completamente aleatorizado

Fuente	gl	Suma de cuadrados	Cuadrado medio	Valor F	Valor P
A	2	13.98	6.99	11.64	0.0001
B	2	10.18	5.09	8.48	0.0010
AB	4	4.77	1.19	1.99	0.1172
C	1	1.19	1.19	1.97	0.1686
AC	2	2.91	1.46	2.43	0.1027
BC	2	3.63	1.82	3.03	0.0610
ABC	4	4.91	1.23	2.04	0.1089
Error	36	21.61	0.60		
Total	53	63.19			

Tabla 14.10: Tabla de medias de 2 factores para el ejemplo 14.4

Catalizador, B	Tiempo de lavado, C	
	15 min	20 min
1	12.19	11.29
2	10.86	10.50
3	11.09	11.46
Medias	11.38	11.08

que parece ser un efecto positivo para el tercer catalizador. Si sólo nos concentramos en los datos para el catalizador 1, una comparación simple entre las medias de los 2 tiempos de lavado produciría un estadístico t sencillo:

$$t = \frac{12.19 - 11.29}{\sqrt{0.6(2/9)}} = 2.5,$$

que es significativo a un nivel menor que 0.02. Así, bien podría ignorarse un importante efecto negativo del tiempo de lavado para el catalizador 1 si el analista hace la interpretación general incorrecta de la razón F insignificante del tiempo de lavado.

Agrupamiento en modelos multifactoriales

El modelo de 3 factores y su análisis se describió de la manera más general mediante la inclusión en el modelo de todas las interacciones posibles. Por supuesto, hay muchas situaciones en las que *a priori* se sabe que el modelo no debería contener ciertas interacciones. Así, es posible aprovechar este conocimiento al combinar o agrupar las sumas de cuadrados correspondientes a interacciones despreciables con la suma de cuadrados del error para formar un nuevo estimador de σ^2 con un número más grande de grados de libertad. Por ejemplo, en un experimento de metalurgia diseñado para estudiar el efecto de 3 variables importantes del proceso sobre el espesor de película, suponga que se sabe que el factor A , la concentración de ácido, no interactúa con los factores B y C . Las

Tabla 14.11: ANOVA sin interacción del factor A

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Efecto principal:				
A	SCA	$a - 1$	s_1^2	$f_1 = \frac{s_1^2}{s^2}$
B	SCB	$b - 1$	s_2^2	$f_2 = \frac{s_2^2}{s^2}$
C	SCC	$c - 1$	s_3^2	$f_3 = \frac{s_3^2}{s^2}$
Interacción de 2 factores:				
BC	$SC(BC)$	$(b - 1)(c - 1)$	s_4^2	$f_4 = \frac{s_4^2}{s^2}$
Error	SCE	Resta	s^2	
Total	STC	$abcn - 1$		

sumas de cuadrados SCA , SCB , SCC y $SC(BC)$ se calculan usando los métodos descritos en un apartado anterior de esta sección. Todos los cuadrados medios de los efectos restantes ahora estimarán de manera independiente la varianza del error σ^2 . Por lo tanto, formamos el nuevo **cuadrado medio del error agrupando** $SC(AB)$, $SC(AC)$, $SC(ABC)$ y SCE junto con los grados de libertad correspondientes. El denominador resultante de las pruebas de significancia es, entonces, el cuadrado medio del error dado por

$$s^2 = \frac{SC(AB) + SC(AC) + SC(ABC) + SCE}{(a - 1)(b - 1) + (a - 1)(c - 1) + (a - 1)(b - 1)(c - 1) + abc(n - 1)}.$$

Por supuesto, con una resta se obtienen la suma de cuadrados agrupada y los grados de libertad agrupados, una vez que se calcula la STC y las sumas de cuadrados para los efectos existentes. La tabla del análisis de varianza adoptaría así la forma de la tabla 14.11.

Experimentos factoriales en bloques

En este capítulo se ha supuesto que el diseño experimental utilizado es un diseño completamente aleatorizado. Al interpretar los niveles del factor A en la tabla 14.11 como **bloques diferentes** se tiene el procedimiento del análisis de varianza para un experimento de 2 factores en un diseño de bloques aleatorizados. Por ejemplo, si se interpretan los operadores del ejemplo 14.4 como bloques, y se supone que no hay interacción entre los bloques y los otros 2 factores, el análisis de varianza adopta la forma de la tabla 14.12, en vez de la de la tabla 14.9. El lector puede verificar que el cuadrado medio del error también es

$$s^2 = \frac{4.77 + 2.91 + 4.91 + 21.61}{4 + 2 + 4 + 36} = 0.74,$$

lo que demuestra el agrupamiento de las sumas de cuadrados para los efectos de la interacción inexistente. Observe que el factor B , el catalizador, tiene un efecto significativo sobre el producto.

Tabla 14.12: ANOVA para un experimento de 2 factores en un diseño de bloques aleatorizados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada	Valor P
Bloques	13.98	2	6.99		
Efecto principal:					
<i>B</i>	10.18	2	5.09	6.88	0.0024
<i>C</i>	1.18	1	1.18	1.59	0.2130
Interacción de 2 factores					
<i>BC</i>	3.64	2	1.82	2.46	0.0966
Error	34.21	46	0.74		
Total	63.19	53			

Ejemplo 14.5: Se realizó un experimento para determinar los efectos de la temperatura, la presión y la intensidad de agitación sobre la tasa de filtración del producto. Esto se hizo en una planta piloto. El experimento se corrió en 2 niveles de cada factor. Además, se decidió que debían utilizarse 2 lotes de materia prima, los cuales fueron tratados como bloques. Se hicieron 8 corridas experimentales en orden aleatorio para cada lote de materia prima. Se piensa que todas las interacciones de los 2 factores podrían ser de interés. No se supone que haya interacciones con los lotes. Los datos aparecen en la tabla 14.13. Las letras "B" y "A" implican niveles bajo y alto, respectivamente. La tasa de filtración se expresa en galones por hora.

- Elabore la tabla ANOVA completa. Agrupe todas las "interacciones" con los bloques dentro del error.
- ¿Cuáles interacciones parecen ser significativas?
- Construya gráficas que revelen las interacciones significativas e interprételas. Explique el significado de la gráfica para el ingeniero.

Tabla 14.13: Datos para el ejemplo 14.5

Lote 1					
Temp.	Tasa de agitación baja		Temp.	Tasa de agitación alta	
	Presión B	Presión A		Presión B	Presión A
B	43	49	B	44	47
A	64	68	A	97	102
Lote 2					
Temp.	Tasa de agitación baja		Temp.	Tasa de agitación alta	
	Presión B	Presión A		Presión B	Presión A
B	49	57	B	51	55
A	70	76	A	103	106

- Solución:** a) En la figura 14.7 se presenta una salida de resultados impresos por computadora del SAS.
- b) Como se aprecia en la figura 14.7, la interacción de la temperatura con la tasa de agitación (*strate*) parece ser muy significativa. Asimismo, la interacción de la presión con la tasa de agitación también parece ser significativa. A propósito, si se hicieran más agrupamientos al combinar las interacciones insignificantes con el error, las conclusiones serían las mismas y el valor *P* para la interacción de la presión con la tasa de agitación se volvería más fuerte, a saber, 0.0517.
- c) Como se aprecia en la figura 14.7, los efectos principales tanto de la tasa de agitación como de la temperatura son muy significativos. Un vistazo a la gráfica de interacción de la figura 14.8a revela que el efecto de la tasa de agitación depende del nivel de la temperatura. Con la temperatura baja el efecto de la tasa de agitación es despreciable, mientras que con la temperatura alta la tasa de agitación tiene un efecto positivo fuerte sobre la tasa media de filtración. En la figura 14.8b la interacción entre la presión y la tasa de agitación, aunque no de manera tan pronunciada como la de la figura 14.8a, todavía muestra una ligera inconsistencia del efecto de la tasa de agitación a través de la presión. ▮

Source	DF	Type III SS	Mean Square	F Value	Pr > F
batch	1	175.562500	175.562500	177.14	<.0001
pressure	1	95.062500	95.062500	95.92	<.0001
temp	1	5292.562500	5292.562500	5340.24	<.0001
pressure*temp	1	0.562500	0.562500	0.57	0.4758
strate	1	1040.062500	1040.062500	1049.43	<.0001
pressure*strate	1	5.062500	5.062500	5.11	0.0583
temp*strate	1	1072.562500	1072.562500	1082.23	<.0001
pressure*temp*strate	1	1.562500	1.562500	1.58	0.2495
Error	7	6.937500	0.991071		
Corrected Total	15	7689.937500			

Figura 14.7: ANOVA para el ejemplo 14.5, interacción del lote agrupado con el error.

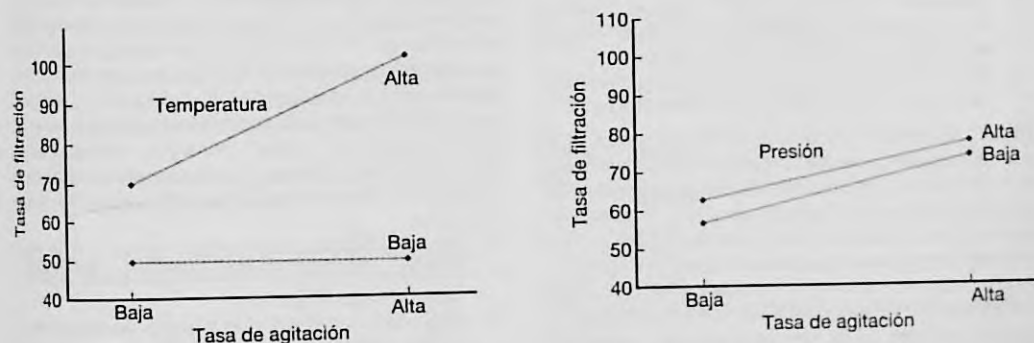


Figura 14.8: Gráficas de interacción para el ejemplo 14.5.

Ejercicios

14.16 Considere una situación experimental que implique los factores A , B y C , en la que se supone un modelo de efectos fijos de 3 factores de la forma $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijk}$. Se considera que todas las demás interacciones no existen o son despreciables. Los datos se presentan en seguida.

	B_1			B_2		
	C_1	C_2	C_3	C_1	C_2	C_3
A_1	4.0	3.4	3.9	4.4	3.1	3.1
	4.9	4.1	4.3	3.4	3.5	3.7
A_2	3.6	2.8	3.1	2.7	2.9	3.7
	3.9	3.2	3.5	3.0	3.2	4.2
A_3	4.8	3.3	3.6	3.6	2.9	2.9
	3.7	3.8	4.2	3.8	3.3	3.5
A_4	3.6	3.2	3.2	2.2	2.9	3.6
	3.9	2.8	3.4	3.5	3.2	4.3

- Haga una prueba de significancia sobre la interacción BC al nivel $\alpha = 0.05$.
- Desarrolle pruebas de significancia sobre los efectos principales A , B y C usando un cuadrado medio del error agrupado, con un nivel $\alpha = 0.05$.

14.17 Los siguientes datos son medidas de un experimento donde se usaron 3 factores, A , B y C , todos de efectos fijos.

	C_1			C_2			C_3		
	B_1	B_2	B_3	B_1	B_2	B_3	B_1	B_2	B_3
A_1	15.0	14.8	15.9	16.8	14.2	13.2	15.8	15.5	19.2
	18.5	13.6	14.8	15.4	12.9	11.6	14.3	13.7	13.5
	22.1	12.2	13.6	14.3	13.0	10.1	13.0	12.6	11.1
A_2	11.3	17.2	16.1	18.9	15.4	12.4	12.7	17.3	7.8
	14.6	15.5	14.7	17.3	17.0	13.6	14.2	15.8	11.5
	18.2	14.2	13.4	16.1	18.6	15.2	15.9	14.6	12.2

- Haga pruebas de significancia sobre todas las interacciones a un nivel $\alpha = 0.05$.
- Realice pruebas de significancia sobre los efectos principales a un nivel $\alpha = 0.05$.
- Dé una explicación de la forma en que una interacción significativa enmascara el efecto del factor C .

14.18 El método de fluorescencia por rayos X es una herramienta analítica importante para determinar la concentración de material en los propulsores sólidos para misiles. En el artículo *An X-ray Fluorescence Method for Analyzing Polybutadiene Acrylic Acid (PBAA) Propellants* (Quarterly Report, RK-TR-62-1, Army Ordnance Missile Command, 1962), se afirma que el proceso de mezcla del propulsor y el tiempo de análisis influyen en la homogeneidad del material y, por lo tanto, en la precisión de las mediciones de la

intensidad de los rayos X. Se hizo un experimento utilizando 3 factores: A , las condiciones de mezcla (4 niveles); B , el tiempo de análisis (2 niveles); y C , el método de carga del propulsor en los recipientes para muestras (temperatura elevada y de la habitación). Se obtuvieron los datos siguientes, que representan el porcentaje de peso del perclorato de amoniaco en un propulsor dado.

	Método de carga, C			
	Caliente		Temp. de la hab.	
	B_1	B_2	B_1	B_2
A				
1	38.62	38.45	39.82	39.82
	37.20	38.64	39.15	40.26
	38.02	38.75	39.78	39.72
2	37.67	37.81	39.53	39.56
	37.57	37.75	39.76	39.25
	37.85	37.91	39.90	39.04
3	37.51	37.21	39.34	39.74
	37.74	37.42	39.60	39.49
	37.58	37.79	39.62	39.45
4	37.52	37.60	40.09	39.36
	37.15	37.55	39.63	39.38
	37.51	37.91	39.67	39.00

- Realice un análisis de varianza con $\alpha = 0.01$ para probar la existencia de efectos principales y de interacción significativos.
- Analice la influencia de los 3 factores sobre el porcentaje de peso del perclorato de amoniaco. Incluya en su análisis el papel que desempeña cualquier interacción significativa.

14.19 La fatiga por corrosión de los metales se ha definido como la acción simultánea de tensión cíclica y ataque químico sobre una estructura metálica. En el estudio *Effect of Humidity and Several Surface Coatings on the Fatigue Life of 2024-T351 Aluminum Alloy*, realizado por el Departamento de Ingeniería Mecánica de Virginia Tech, se utilizó una técnica que requería la aplicación de un recubrimiento protector de cromato para minimizar el daño de la fatiga por corrosión en el aluminio. En la investigación se emplearon 3 factores con 5 réplicas para cada combinación de tratamientos: recubrimiento, en 2 niveles; humedad y esfuerzo constante, ambos en 3 niveles. A continuación se presentan los datos de fatiga expresados en miles de ciclos antes de la falla.

- Realice un análisis de varianza con $\alpha = 0.01$ para probar la existencia de efectos principales y de interacción significativos.
- Haga una recomendación para las combinaciones de los 3 factores que producirían poco daño por fatiga.

Recubrimiento	Humedad	Esfuerzo cortante (psi)		
		13,000	17,000	20,000
Sin recubrimiento	Bajo (20-25% RH)	4580	5252	361
		10,126	897	466
		1341	1465	1069
		6414	2694	469
	Medio (50-60% RH)	2858	799	314
		8829	3471	244
		10,914	685	261
		4067	810	522
	Alto (86-91% RH)	2595	3409	739
		6489	1862	1344
		5248	2710	1027
		6816	2632	663
Cromado	Bajo (20-25% RH)	5860	2131	1216
		5901	2470	1097
		5395	4035	130
		2768	2022	841
	Medio (50-60% RH)	1821	914	1595
		3604	2036	1482
		4106	3524	529
		4833	1847	252
	Alto (86-91% RH)	7414	1684	105
		10,022	3042	847
		7463	4482	874
		21,906	996	755
	3287	1319	586	
	5200	929	402	
	5493	1263	846	
	4145	2236	524	
	3336	1392	751	

Dentista	Método	Tipo	
		Lámina dorada	Goldent
3	1	715	724
	2	792	715
	3	762	606
4	1	673	946
	2	657	743
	3	690	245
5	1	634	715
	2	649	724
	3	724	627

14.21 Las copiatoras electrónicas funcionan adhiriendo tinta negra al papel mediante electricidad estática. La etapa final del proceso de copiado comprende el calentamiento y adhesión de la tinta sobre el papel. La potencia de la adhesión durante este proceso final determina la calidad de la copia. Se plantea que la temperatura, el estado superficial de la adhesión en el rodillo y la dureza del rodillo de la prensa influyen en la potencia de adhesión de la copiadora. Se hizo un experimento con tratamientos, que consistían en una combinación de estos 3 factores en cada uno de 3 niveles. Los datos siguientes muestran la potencia de la adhesión para cada combinación de tratamientos. Lleve a cabo un análisis de varianza con $\alpha = 0.05$ para probar si hay efectos principales y de interacción significativos.

Temp.	Estado superficial de la adhesión en el rodillo	Dureza del rodillo de la prensa					
		20		40		60	
baja	Suave	0.52	0.44	0.54	0.52	0.60	0.55
		0.57	0.53	0.65	0.56	0.78	0.68
	Medio	0.64	0.59	0.79	0.73	0.49	0.48
		0.58	0.64	0.79	0.78	0.74	0.50
	Duro	0.67	0.77	0.58	0.68	0.55	0.65
		0.74	0.65	0.57	0.59	0.57	0.58
media	Suave	0.46	0.40	0.31	0.49	0.56	0.42
		0.58	0.37	0.48	0.66	0.49	0.49
	Medio	0.60	0.43	0.66	0.57	0.64	0.54
		0.62	0.61	0.72	0.56	0.74	0.56
	Duro	0.53	0.65	0.53	0.45	0.56	0.66
		0.66	0.56	0.59	0.47	0.71	0.67
alta	Suave	0.52	0.44	0.54	0.52	0.65	0.49
		0.57	0.53	0.65	0.56	0.65	0.52
	Medio	0.53	0.65	0.53	0.45	0.49	0.48
		0.66	0.56	0.59	0.47	0.74	0.50
	Duro	0.43	0.43	0.48	0.31	0.55	0.65
		0.47	0.44	0.43	0.27	0.57	0.58

14.22 Considere el conjunto de datos del ejercicio 14.21.

- Construya una gráfica de la interacción para cualquier interacción de 2 factores que sea significativa.
- Dibuje una gráfica de probabilidad normal de residuales y coméntela.

Dentista	Método	Tipo	
		Lámina dorada	Goldent
1	1	792	824
	2	772	772
	3	782	803
2	1	803	803
	2	752	772
	3	715	707

(cont.)

14.23 Considere combinaciones de 3 factores en el retiro de la suciedad de cargas estándar de lavandería. El primer factor es la marca del detergente: X, Y o Z. El segundo factor es el tipo de detergente: líquido o en polvo. El tercer factor es la temperatura del agua, caliente o tibia. El experimento se replicó 3 veces. La respuesta es el porcentaje de suciedad eliminada. Los datos son los siguientes:

Marca	Tipo	Temperatura			
X	En polvo	Caliente	85	88	80
		Tibia	82	83	85
	Líquido	Caliente	78	75	72
		Tibia	75	75	73
Y	En polvo	Caliente	90	92	92
		Tibia	88	86	88
	Líquido	Caliente	78	76	70
		Tibia	76	77	76
Z	En polvo	Caliente	85	87	88
		Tibia	76	74	78
	Líquido	Caliente	60	70	68
		Tibia	55	57	54

- ¿Existen efectos de la interacción significativos a un nivel $\alpha = 0.05$?
- ¿Hay diferencias significativas entre las tres marcas de detergente?
- ¿Cuál combinación de factores preferiría utilizar?

14.24 Un científico recaba datos experimentales sobre el radio de un grano de combustible propulsor, y, en función de la temperatura del polvo, la tasa de extrusión y la temperatura del molde. Los resultados de los 3 factores del experimento son los siguientes:

Tasa	Temp. del polvo			
	150		190	
	Temp. del molde 220	Temp. del molde 250	Temp. del molde 220	Temp. del molde 250
12	82	124	88	129
24	114	157	121	164

No se dispone de recursos para hacer experimentos repetidos con las 8 combinaciones de factores. Se cree

14.5 Experimentos factoriales para efectos aleatorios y modelos mixtos

En un experimento de 2 factores con efectos aleatorios se tiene el modelo

$$Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \epsilon_{ijk}$$

para $i = 1, 2, \dots, a$; $j = 1, 2, \dots, b$; y $k = 1, 2, \dots, n$, donde A , B , $(AB)_{ij}$ y ϵ_{ijk} son variables aleatorias independientes con medias igual a 0 y varianzas σ_a^2 , σ_b^2 , σ_{ab}^2 y σ^2 , respectivamente. Las sumas de cuadrados para experimentos de efectos aleatorios se calculan

que la tasa de extrusión no interactúa con la temperatura del molde, y que la interacción entre los 3 factores es despreciable. Así, esas 2 interacciones pueden agruparse para producir un término de "error" con 2 grados de libertad.

- Haga un análisis de varianza que incluya los 3 efectos principales e interacciones de 2 factores. Determine cuáles efectos influyen en el radio del grano de combustible.
- Construya gráficas de interacción para la temperatura del polvo usando la temperatura del molde y la del polvo mediante las interacciones de la tasa de extrusión.
- Comente acerca de la consistencia de la apariencia de las gráficas de interacción y las pruebas sobre las 2 interacciones en el ANOVA.

14.25 En el libro *Design of Experiments for Quality Improvement*, publicado por la Japanese Standards Association (1989), se reporta un estudio sobre la extracción de polietileno por medio de un solvente, y la manera en que la cantidad de gel (proporción) se ve influida por 3 factores: el tipo de solvente, la temperatura de extracción y el tiempo de extracción. Se diseñó un experimento factorial y se obtuvieron los datos siguientes, expresados en proporción de gel.

Temp. del solvente		Tiempo					
		4		8		16	
Etanol	120	94.0	94.0	93.8	94.2	91.1	90.5
	80	95.3	95.1	94.9	95.3	92.5	92.4
Tolueno	120	94.6	94.5	93.6	94.1	91.1	91.0
	80	95.4	95.4	95.6	96.0	92.1	92.1

- Haga un análisis de varianza y determine cuáles factores e interacciones influyen en la proporción de gel.
- Construya una gráfica de la interacción entre cualesquiera 2 factores que sea significativa. Además, explique qué conclusión se podría extraer de la presencia de la interacción.
- Haga una gráfica de probabilidad normal de los residuales y comente.

exactamente de la misma forma que en los experimentos de efectos fijos. Ahora se tiene interés en probar hipótesis con la forma

$$\begin{aligned} H_0': \sigma_\alpha^2 &= 0, & H_0'': \sigma_\beta^2 &= 0, & H_0''': \sigma_{\alpha\beta}^2 &= 0, \\ H_1': \sigma_\alpha^2 &\neq 0, & H_1'': \sigma_\beta^2 &\neq 0, & H_1''': \sigma_{\alpha\beta}^2 &\neq 0, \end{aligned}$$

donde el denominador en la razón f no es necesariamente el cuadrado medio del error. El denominador apropiado se determina examinando los valores esperados de los distintos cuadrados medios, los cuales se presentan en la tabla 14.14.

Tabla 14.14: Cuadrados medios esperados para un experimento de efectos aleatorios de 2 factores

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
A	$a - 1$	s_1^2	$\sigma^2 + n\sigma_{\alpha\beta}^2 + bn\sigma_\alpha^2$
B	$b - 1$	s_2^2	$\sigma^2 + n\sigma_{\alpha\beta}^2 + an\sigma_\beta^2$
AB	$(a - 1)(b - 1)$	s_3^2	$\sigma^2 + n\sigma_{\alpha\beta}^2$
Error	$ab(n - 1)$	s^2	σ^2
Total	$abn - 1$		

En la tabla 14.14 se observa que H_0' y H_0'' se prueban usando s_3^2 en el denominador de la razón f ; mientras que H_0''' se prueba con s^2 en el denominador. Los estimados no sesgados de los componentes de la varianza son

$$\hat{\sigma}^2 = s^2, \quad \hat{\sigma}_{\alpha\beta}^2 = \frac{s_3^2 - s^2}{n}, \quad \hat{\sigma}_\alpha^2 = \frac{s_1^2 - s_3^2}{bn}, \quad \hat{\sigma}_\beta^2 = \frac{s_2^2 - s_3^2}{an}.$$

Tabla 14.15: Cuadrados medios esperados para un experimento de efectos aleatorios de 3 factores

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
A	$a - 1$	s_1^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + cn\sigma_{\alpha\beta}^2 + bn\sigma_{\alpha\gamma}^2 + bcn\sigma_\alpha^2$
B	$b - 1$	s_2^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + cn\sigma_{\alpha\beta}^2 + an\sigma_{\beta\gamma}^2 + acn\sigma_\beta^2$
C	$c - 1$	s_3^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + bn\sigma_{\alpha\gamma}^2 + an\sigma_{\beta\gamma}^2 + abn\sigma_\gamma^2$
AB	$(a - 1)(b - 1)$	s_4^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + cn\sigma_{\alpha\beta}^2$
AC	$(a - 1)(c - 1)$	s_5^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + bn\sigma_{\alpha\gamma}^2$
BC	$(b - 1)(c - 1)$	s_6^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + an\sigma_{\beta\gamma}^2$
ABC	$(a - 1)(b - 1)(c - 1)$	s_7^2	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2$
Error	$abc(n - 1)$	s^2	σ^2
Total	$abcn - 1$		

En la tabla 14.15 se presentan los cuadrados medios esperados para el experimento de 3 factores con efectos aleatorios en un diseño completamente aleatorizado. A partir de los cuadrados medios esperados de la tabla 14.15 es evidente que se pueden formar

razones f adecuadas para probar todos los componentes de la varianza de la interacción de 2 y 3 factores. Sin embargo, para probar una hipótesis de la forma

$$H_0: \sigma_\alpha^2 = 0,$$

$$H_1: \sigma_\alpha^2 \neq 0,$$

parece que no hay razón f apropiada, a menos que se encontrara que uno o más de los componentes de la varianza de interacción de 2 factores no es significativo. Por ejemplo, suponga que se hubiera comparado s_5^2 (cuadrado medio AC) con s_7^2 (cuadrado medio ABC) y se encontrara que $\sigma_{\alpha\gamma}^2$ es despreciable. Entonces podría argumentarse que el término $\sigma_{\alpha\gamma}^2$ debería eliminarse de todos los cuadrados medios esperados de la tabla 14.15; entonces, la razón s_5^2/s_4^2 ofrece una prueba de la significancia del componente σ_α^2 de la varianza. Por lo tanto, si se prueba la hipótesis concerniente a los componentes de la varianza de los efectos principales, es necesario investigar primero la significancia de los componentes de la interacción de 2 factores. Cuando se encuentra que ciertos componentes de la varianza de la interacción de 2 factores son significativos, por lo que deben permanecer como parte del cuadrado medio esperado, se utiliza una prueba aproximada derivada por Satterthwaite (1946; véase la bibliografía).

Ejemplo 14.6: En un estudio realizado para determinar cuáles son las fuentes importantes de la variación en un proceso industrial, se toman 3 mediciones del producto para 3 operadores elegidos al azar, y se eligen en forma aleatoria 4 lotes de materia prima. Se decidió que debe hacerse una prueba estadística a un nivel de significancia de 0.05 para determinar si los componentes de la varianza debidos a los lotes, los operadores y la interacción son significativos. Además, tienen que calcularse los estimados de los componentes de la varianza. En la tabla 14.16 se presentan los datos con la respuesta expresada en porcentaje por peso:

Tabla 14.16: Datos para el ejemplo 14.6

Operador	Lote			
	1	2	3	4
1	66.9	68.3	69.0	69.3
	68.1	67.4	69.8	70.9
	67.2	67.7	67.5	71.4
2	66.3	68.1	69.7	69.4
	65.4	66.9	68.8	69.6
	65.8	67.6	69.2	70.0
3	65.6	66.0	67.1	67.9
	66.3	66.9	66.2	68.4
	65.2	67.3	67.4	68.7

Solución: Las sumas de cuadrados se calculan de la forma acostumbrada y se obtienen los siguientes resultados:

$$STC \text{ (total)} = 84.5564, \quad SCE \text{ (error)} = 10.6733,$$

$$SCA \text{ (operadores)} = 18.2106, \quad SCB \text{ (lotes)} = 50.1564,$$

$$SC(AB) \text{ (interacción)} = 5.5161.$$

Se realizaron todos los demás cálculos y se presentan en la tabla 14.17. Como

$$f_{0.05}(2, 6) = 5.14, \quad f_{0.05}(3, 6) = 4.76, \quad \text{y} \quad f_{0.05}(6, 24) = 2.51,$$

se descubre que los componentes de la varianza de los operadores y el lote son significativos. Aunque la varianza de la interacción no es significativa a un nivel $\alpha = 0.05$, el valor P es de 0.095. Los estimados de los componentes de la varianza del efecto principal son

$$\hat{\sigma}_\alpha^2 = \frac{9.1053 - 0.9194}{12} = 0.68, \quad \hat{\sigma}_\beta^2 = \frac{16.7188 - 0.9194}{9} = 1.76.$$

Tabla 14.17: Análisis de varianza para el ejemplo 14.6

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Operadores	18.2106	2	9.1053	9.90
Lotes	50.1564	3	16.7188	18.18
Interacción	5.5161	6	0.9194	2.07
Error	10.6733	24	0.4447	
Total	84.5564	35		

Experimento del modelo mixto

Hay situaciones en que el experimento dicta la suposición de un **modelo mixto**, es decir, una mezcla de efectos aleatorios y fijos. Por ejemplo, para el caso de 2 factores se tiene que

$$Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \epsilon_{ijk},$$

para $i = 1, 2, \dots, a; j = 1, 2, \dots, b; k = 1, 2, \dots, n$. Las A_i pueden ser variables aleatorias independientes de ϵ_{ijk} , y las B_j pueden ser efectos fijos. La naturaleza mixta del modelo requiere que los términos de la interacción sean variables aleatorias. Como resultado, las hipótesis relevantes adoptan la forma

$$\begin{aligned} H'_0: \sigma_\alpha^2 = 0, \quad H''_0: B_1 = B_2 = \dots = B_b = 0, \quad H'''_0: \sigma_{\alpha\beta}^2 = 0, \\ H'_1: \sigma_\alpha^2 \neq 0, \quad H''_1: \text{Al menos una de las } B_j \text{ no es igual a } 0, \quad H'''_1: \sigma_{\alpha\beta}^2 \neq 0. \end{aligned}$$

Otra vez, los cálculos de la suma de cuadrados son idénticos a los de las situaciones de efectos fijos y aleatorios, y la prueba F es determinada por los cuadrados medios esperados. La tabla 14.18 proporciona los cuadrados medios esperados para el problema del modelo mixto de 2 factores.

Tabla 14.18: Cuadrados medios esperados para el experimento del modelo mixto de 2 factores

Factor	Cuadrado medio esperado
A (aleatorios)	$\sigma^2 + bn\sigma_\alpha^2$
B (fijos)	$\sigma^2 + n\sigma_{\alpha\beta}^2 + \frac{an}{b-1} \sum_j B_j^2$
AB (aleatorios)	$\sigma^2 + n\sigma_{\alpha\beta}^2$
Error	σ^2

A partir de la naturaleza de los cuadrados medios esperados queda claro que la prueba sobre el efecto aleatorio emplea el cuadrado medio del error s^2 como denominador, mientras que la prueba sobre el efecto fijo utiliza el cuadrado medio de interacción. Suponga que ahora se consideran 3 factores. En este caso, por supuesto, debe tomarse en cuenta la situación en que un factor es fijo y la situación en que 2 factores son fijos. La tabla 14.19 cubre ambas situaciones.

Tabla 14.19: Cuadrados medios esperados para experimentos factoriales de modelo mixto de 3 factores

	A aleatoria	A aleatoria, B aleatoria
A	$\sigma^2 + bcn\sigma_\alpha^2$	$\sigma^2 + cn\sigma_{\alpha\beta}^2 + bcn\sigma_\alpha^2$
B	$\sigma^2 + cn\sigma_{\alpha\beta}^2 + acn \sum_{j=1}^b \frac{B_j^2}{b-1}$	$\sigma^2 + cn\sigma_{\alpha\beta}^2 + acn\sigma_\beta^2$
C	$\sigma^2 + bn\sigma_{\alpha\gamma}^2 + abn \sum_{k=1}^c \frac{C_k^2}{c-1}$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + an\sigma_{\beta\gamma}^2 + bn\sigma_{\alpha\gamma}^2 + abn \sum_{k=1}^c \frac{C_k^2}{c-1}$
AB	$\sigma^2 + cn\sigma_{\alpha\beta}^2$	$\sigma^2 + cn\sigma_{\alpha\beta}^2$
AC	$\sigma^2 + bn\sigma_{\alpha\gamma}^2$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + bn\sigma_{\alpha\gamma}^2$
BC	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + an \sum_j \sum_k \frac{(BC)_{jk}^2}{(b-1)(c-1)}$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2 + an\sigma_{\beta\gamma}^2$
AB	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2$
Error	σ^2	σ^2

Observe que en el caso de A aleatoria todos los efectos tienen pruebas f apropiadas. No obstante, para A y B aleatorias, el efecto principal C debe probarse utilizando un procedimiento tipo Satterthwaite, similar al que se emplea en el experimento de efectos aleatorios.

Ejercicios

14.26 Suponga un experimento de efectos aleatorios para el ejercicio 14.2 de la página 575 y estime los componentes de la varianza para las marcas de concentrado de jugo de naranja, para el número de días transcurridos a partir del día en que se mezcló el jugo hasta el día en que se hizo la prueba, y para el error experimental.

14.27 Para estimar los diversos componentes de la variabilidad en un proceso de filtración el porcentaje de material que se pierde en el licor madre se mide en 12 condiciones experimentales, con 3 corridas en cada condición. Se seleccionan al azar 3 filtros y 4 operadores para usarlos en el experimento.

a) Pruebe la hipótesis de que no hay un componente de interacción de la varianza entre los filtros y los operadores a un nivel de significancia $\alpha = 0.05$.

b) Pruebe la hipótesis de que los operadores y los filtros no tienen ningún efecto sobre la variabilidad del proceso de filtración a un nivel de significancia $\alpha = 0.05$.

c) Estime los componentes de la varianza que se deben a los filtros, a los operadores y al error experimental.

Filtro	Operador			
	1	2	3	4
1	16.2	15.9	15.6	14.9
	16.8	15.1	15.9	15.2
	17.1	14.5	16.1	14.9
2	16.6	16.0	16.1	15.4
	16.9	16.3	16.0	14.6
	16.8	16.5	17.2	15.9
3	16.7	16.5	16.4	16.1
	16.9	16.9	17.4	15.4
	17.1	16.8	16.9	15.6

14.28 Un contratista de la defensa está interesado en estudiar un proceso de inspección para detectar la falla o la fatiga de partes de recambio. Se utilizan 3 niveles de inspección que ejecutan 3 inspectores elegidos al azar. Se emplean 5 lotes para cada combinación en el estudio. Los niveles de los factores están en los datos.

La respuesta se expresa en fallas por cada 1000 piezas.

- a) Escriba un modelo adecuado, con suposiciones.
 b) Utilice análisis de varianza para probar las hipótesis apropiadas para los inspectores, el nivel de inspección y la interacción.

Nivel de inspección

Inspector	Inspección militar completa		Inspección militar reducida		Comercial	
A	7.50	7.42	7.08	6.17	6.15	5.52
	5.85	5.89	5.65	5.30	5.48	5.48
	5.35		5.02		5.98	
B	7.58	6.52	7.68	5.86	6.17	6.20
	6.54	5.64	5.28	5.38	5.44	5.75
	5.12		4.87		5.68	
C	7.70	6.82	7.19	6.19	6.21	5.66
	6.42	5.39	5.85	5.35	5.36	5.90
	5.35		5.01		6.12	

14.29 Considere el análisis de varianza siguiente para un experimento de efectos aleatorios:

Fuente de variación	Grados de libertad	Cuadrado medio
A	3	140
B	1	480
C	2	325
AB	3	15
AC	6	24
BC	2	18
ABC	6	2
Error	24	5
Total	47	

Pruebe si existen componentes significativos de la varianza entre todos los efectos principales y los efectos de interacción a un nivel de significancia de 0.01,

- a) utilice un estimado agrupado del error cuando esto sea apropiado;
 b) sin agrupar las sumas de los cuadrados de efectos insignificantes.

14.30 A un gerente de una planta le gustaría demostrar que la producción de una fábrica de lana de su planta no depende del operador de la máquina ni de la hora del día, y que es consistentemente elevada. Se eligen al azar 4 operadores y 3 horas del día para el estudio. Se mide el producto en yardas por minuto y se toman muestras 3 días elegidos al azar.

- a) Escriba el modelo apropiado.
 b) Evalúe los componentes de la varianza para el operador y la hora.
 c) Saque sus conclusiones.

Hora	Operador			
	1	2	3	4
1	9.5	9.8	9.8	10.0
	9.8	10.1	10.3	9.7
	10.0	9.6	9.7	10.2
2	10.2	10.1	10.2	10.3
	9.9	9.8	9.8	10.1
	9.5	9.7	9.7	9.9
3	10.5	10.4	9.9	10.0
	10.2	10.2	10.3	10.1
	9.3	9.8	10.2	9.7

14.31 Un fabricante de pintura de látex para interiores (marca A) quisiera demostrar que su pintura es más robusta para el material donde se aplica, que la de sus 2 competidores más cercanos. La respuesta es el tiempo, en años, hasta que comienza a picarse. El estudio incluye las 3 marcas de pintura y 3 materiales seleccionados al azar. Para cada combinación se utilizan 2 piezas.

Material	Marca de pintura					
	A		B		C	
A	5.50	5.15	4.75	4.60	5.10	5.20
B	5.60	5.55	5.50	5.60	5.40	5.50
C	5.40	5.48	5.05	4.95	4.50	4.55

- a) ¿Cómo se le llama a este tipo de modelo?
 b) Analice los datos usando el modelo apropiado.
 c) ¿Los datos apoyan la afirmación del fabricante de la marca A?

14.32 Un ingeniero de procesos desea determinar si el ajuste de potencia de las máquinas que se usan para llenar ciertos tipos de cajas de cereal tienen un efecto significativo sobre el peso real del producto. El estudio consta de 3 tipos de cereal elaborados por la empresa, elegidos al azar, y 3 flujos fijos de energía. Para cada combinación se mide el peso de 4 cajas de cereal diferentes seleccionadas al azar. El peso deseado es de 400 gramos. A continuación se presentan los datos.

Ajuste de potencia	Tipo del cereal					
	1		2		3	
Bajo	395	390	392	392	402	405
	401	400	394	401	399	399
Actual	396	399	390	392	404	403
	400	402	395	502	400	399
Alto	410	408	404	406	415	412
	408	407	401	400	413	415

- a) Proporcione el modelo adecuado y liste las suposiciones que se hacen.
 b) ¿Hay un efecto significativo debido al ajuste de potencia?
 c) ¿Existe un componente de la varianza significativo debido al tipo de cereal?

Ejercicios de repaso

14.33 El Centro de Consulta Estadística de Virginia Tech participó en el análisis de un conjunto de datos tomados por el personal del Departamento de Nutrición Humana y Alimentos, al cual le interesaba estudiar los efectos del tipo de harina y el porcentaje de edulcorante sobre ciertos atributos físicos de un tipo de pastel. Se usó harina multiusos y para pasteles, y el porcentaje de edulcorante varió en 4 niveles. Los siguientes datos presentan información acerca de la gravedad específica de las muestras de pastel. Se prepararon 3 pasteles con cada una de las 8 combinaciones de factores.

Concentración de edulcorante	Harina					
	Multiusos			Para pasteles		
0	0.90	0.87	0.90	0.91	0.90	0.80
50	0.86	0.89	0.91	0.88	0.82	0.83
75	0.93	0.88	0.87	0.86	0.85	0.80
100	0.79	0.82	0.80	0.86	0.85	0.85

- a) Realice un análisis de varianza con 2 factores. Pruebe si existen diferencias entre los tipos de harina. Pruebe si hay diferencias entre las concentraciones de edulcorante.
- b) Analice el efecto de la interacción, si lo hubiera. Proporcione valores P para todas las pruebas.

14.34 Se llevó a cabo un experimento en el Departamento de Ciencias de Alimentos de Virginia Tech. El objetivo fue caracterizar la textura de cierto tipo de pescado de la familia de los arenques. También se estudió el efecto de los tipos de salsa empleada para preparar el pescado. La respuesta en el experimento era un "valor de textura", medido con una máquina que rebanaba el producto de los peces. Los siguientes datos son los valores de textura:

Tipo de salsa	Tipo de pescado					
	Sábalo sin curar		Sábalo curado		Arenque	
Crema ácida	27.6	57.4	64.0	66.9	107.0	83.9
	47.8	71.1	66.5	66.8	110.4	93.4
	53.8		53.8		83.1	
Salsa envinada	49.8	31.0	48.3	62.2	88.0	95.2
	11.8	35.1	54.6	43.6	108.2	86.7
	16.1		41.8		105.2	

- a) Haga un análisis de varianza. Determine si hay o no interacción entre el tipo de salsa y el tipo de pescado.
- b) Con base en los resultados del inciso a) y en pruebas F de los efectos principales, determine si hay una diferencia significativa en la textura debido a los tipos de salsa, y determine si existe una diferencia significativa entre los tipos de pescado.

14.35 Se hizo un estudio para determinar si las condiciones de humedad afectan la fuerza que se requiere

para separar piezas de plástico engomadas. Se probaron 3 tipos de plástico con 4 niveles de humedad. Los resultados, en kilogramos, son los siguientes:

Tipo de plástico	Humedad			
	30%	50%	70%	90%
A	39.0	33.1	33.8	33.0
	42.8	37.8	30.7	32.9
B	36.9	27.2	29.7	28.5
	41.0	26.8	29.1	27.9
C	27.4	29.2	26.7	30.9
	30.3	29.9	32.0	31.5

- a) Suponga un experimento de efectos fijos, realice un análisis de varianza y pruebe la hipótesis de que no hay interacción entre la humedad y el tipo de plástico a un nivel de significancia de 0.05.
- b) Utilice sólo los plásticos A y B y el valor de s^2 del inciso a) y vuelva a probar la presencia de interacción a un nivel de significancia de 0.05.

14.36 Personal del Departamento de Ingeniería de Materiales de Virginia Tech llevó a cabo un experimento para estudiar los efectos de los factores ambientales sobre la estabilidad de cierto tipo de aleación cobre-níquel. La respuesta básica fue la vida de fatiga del material. Los factores son el nivel de esfuerzo y el ambiente. Los datos son los siguientes:

Ambiente	Nivel de esfuerzo		
	Bajo	Medio	Alto
Hidrógeno seco	11.08	13.12	14.18
	10.98	13.04	14.90
	11.24	13.37	15.10
Humedad elevada (95%)	10.75	12.73	14.15
	10.52	12.87	14.42
	10.43	12.95	14.25

- a) Haga un análisis de varianza para probar la interacción entre los factores. Use $\alpha = 0.05$.
- b) Con base en el inciso a) efectúe un análisis sobre los 2 efectos principales y saque sus conclusiones. Utilice el método del valor P para sus conclusiones.

14.37 En el experimento del ejercicio de repaso 14.33 también se utilizó el volumen del pastel como respuesta. Las unidades en que se expresa son pulgadas cúbicas. Pruebe la interacción entre los factores y analice los efectos principales. Suponga que los 2 factores son efectos fijos

Concentración de edulcorante	Harina					
	Multiusos			Para pasteles		
0	4.48	3.98	4.42	4.12	4.92	5.10
50	3.68	5.04	3.72	5.00	4.26	4.34
75	3.92	3.82	4.06	4.82	4.34	4.40
100	3.26	3.80	3.40	4.32	4.18	4.30

14.38 Una válvula de control necesita ser muy sensible al voltaje de entrada para así generar un voltaje de salida adecuado. Un ingeniero gira las perillas de control para cambiar el voltaje de entrada. En el libro *SN-Ratio for the Quality Evaluation*, publicado por la Japanese Standards Association (1988), se describe un estudio sobre la forma en que esos 3 factores (posición relativa de las perillas de control, rango de control de las perillas y voltaje de entrada) influyen en la sensibilidad de una válvula de control. A continuación se presentan los factores y sus niveles. Los datos se refieren a la sensibilidad de una válvula de control.

Factor A: posición relativa de las perillas de control:
centro -0.5, centro y centro + 0.5

Factor B: rango de control de las perillas:
2, 4.5 y 7 (mm)

Factor C: voltaje de entrada: 100, 120 y 150 (V)

A	B	C		
		C ₁	C ₂	C ₃
A ₁	B ₁	151 135	151 135	151 138
A ₁	B ₂	178 171	180 173	181 174
A ₁	B ₃	204 190	205 190	206 192
A ₂	B ₁	156 148	158 149	158 150
A ₂	B ₂	183 168	183 170	183 172
A ₂	B ₃	210 204	211 203	213 204
A ₃	B ₁	161 145	162 148	163 148
A ₃	B ₂	189 182	191 184	192 183
A ₃	B ₃	215 202	216 203	217 205

Realice un análisis de varianza con $\alpha = 0.05$ para probar la existencia de efectos principales y de interacción significativos. Saque sus conclusiones.

14.39 En el ejercicio 14.25 de la página 588 se describe un experimento que implica la extracción de polietileno a través de un solvente.

Solvente	Tiempo					
	4		8		16	
Etanol	120	94.0 94.0	93.8 94.2	91.1 90.5		
	80	95.3 95.1	94.9 95.3	92.5 92.4		
Tolueno	120	94.6 94.5	93.6 94.1	91.1 91.0		
	80	95.4 95.4	95.6 96.0	92.1 92.1		

- Haga una clase diferente de análisis de los datos. Ajuste un modelo adecuado de regresión con una variable categórica del solvente, un término de temperatura, un término de tiempo, una interacción de la temperatura y el tiempo, una interacción del solvente y la temperatura y una interacción del solvente y el tiempo. Realice pruebas *t* para todos los coeficientes y describa sus hallazgos.
- ¿Sus resultados sugieren que el etanol y el tolueno requieren modelos diferentes, o son equivalentes aparte de las intersecciones? Explique su respuesta.
- ¿Encontró alguna conclusión que contradiga las conclusiones que sacó de la solución del ejercicio 14.25? Explique su respuesta.

14.40 En el libro *SN-Ratio for the Quality Evaluation*, publicado por la Japanese Standards Association (1988), se describe un estudio acerca de cómo la presión del aire de los neumáticos afecta la maniobrabilidad de un automóvil. Se compararon 3 presiones distintas de aire en los neumáticos sobre 3 superficies diferentes de manejo. Las 3 presiones del aire fueron: los neumáticos tanto del lado izquierdo como del derecho inflados a 6 kgf/cm², los neumáticos del lado izquierdo inflados a 6 kgf/cm² y los del lado derecho inflados a 3 kgf/cm², y los neumáticos de ambos lados inflados a 3 kgf/cm². Las tres superficies de manejo fueron asfalto, asfalto seco y cemento seco. Se observó 2 veces el radio de giro de un vehículo de prueba para cada nivel de presión de los neumáticos sobre cada una de las 3 superficies de manejo.

Superficie de manejo	Presión del aire de los neumáticos					
	1		2		3	
Asfalto	44.0	25.5	34.2	37.2	27.4	42.8
Asfalto seco	31.9	33.7	31.8	27.6	43.7	38.2
Cemento seco	27.3	39.5	46.6	28.1	35.5	34.6

Realice un análisis de varianza con los datos anteriores. Haga comentarios acerca de la interpretación de los efectos principales y de interacción.

14.41 El fabricante de cierta marca de café secado por congelación espera reducir el tiempo del proceso sin arriesgar la integridad del producto. El ingeniero de procesos desea usar 3 temperaturas para la cámara de secado y 4 tiempos de secado. El tiempo de secado actual es de 3 horas a una temperatura de -15°C. La respuesta del sabor es un promedio de las calificaciones de 4 jueces profesionales. La calificación está en una escala de 1 a 10, donde 10 es la mejor. En la tabla que sigue se presentan los datos.

Tiempo	Temperatura					
	-20°C		-15°C		-10°C	
1 hr	9.60	9.63	9.55	9.50	9.40	9.43
1.5 hr	9.75	9.73	9.60	9.61	9.55	9.48
2 hr	9.82	9.93	9.81	9.78	9.50	9.52
3 hr	9.78	9.81	9.80	9.75	9.55	9.58

- ¿Qué tipo de modelo se debe utilizar? Plantee las suposiciones.
- Analice los datos en forma apropiada.
- Redacte un breve informe para el vicepresidente encargado y hágale una recomendación para la elaboración futura de este producto.

14.42 Para garantizar el número de cajeros necesarios durante las horas pico de operación, un banco urbano recabó datos. Se estudiaron 4 cajeros durante 3 horarios "ocupados", 1) entre semana, de 10:00 a 11:00 A.M., 2) por las tardes entre semana, entre las 2:00 y las 3:00 P.M., y 3) las mañanas de los sábados, entre 11:00 y las 12:00. Un analista eligió al azar 4 horarios

dentro de cada uno de los 3 periodos, para cada una de las 4 posiciones de los cajeros durante varios meses y se observó el número de clientes atendidos. Los datos son los siguientes:

Cajero	Periodo													
	1			2			3							
1	18	24	17	22	25	29	23	32	25	32	29	30	21	34
2	16	11	19	14	23	32	25	17	27	29	18	16		
3	12	19	11	22	27	33	27	24	25	20	29	15		
4	11	9	13	8	10	7	19	8	11	9	17	9		

Se supone que el número de clientes atendidos es una variable aleatoria de Poisson.

a) Comente sobre el riesgo de llevar a cabo un análisis

- de varianza estándar con los datos anteriores. ¿Qué suposiciones, si las hubiera, se violarían?
- b) Elabore una tabla de ANOVA estándar que incluya pruebas F de los efectos principales y las interacciones. Si las interacciones y los efectos principales resultan significativos, establezca las conclusiones científicas. ¿Qué aprendimos? Asegúrese de interpretar cualquier interacción significativa. Utilice su propio juicio respecto a los valores P .
- c) Vuelva a hacer el análisis completo usando una transformación apropiada de la respuesta. ¿Encontró alguna diferencia en los resultados? Haga comentarios al respecto.

14.6 Posibles riesgos y errores conceptuales; relación con el material de otros capítulos

Uno de los temas más susceptibles de confusión en el análisis de experimentos factoriales radica en la interpretación de los efectos principales ante la presencia de interacción. La existencia de un valor P relativamente grande para un efecto principal, cuando es clara la presencia de interacciones, podría tentar al analista a concluir que “no existe efecto principal significativo”. Sin embargo, debe entenderse que si un efecto principal está implicado en una interacción significativa, entonces el efecto principal **está influyendo en la respuesta**. La naturaleza del efecto es inconsistente a través de los niveles de otros efectos. La naturaleza del papel que desempeña el efecto principal se deduce de las **gráficas de interacción**.

Debido a lo que se expresa en el párrafo anterior, hay un gran peligro de usar la estadística de manera equivocada cuando se emplea una prueba de comparación múltiple sobre los efectos principales ante la presencia clara de interacción entre los factores.

Debe tenerse precaución en el análisis de un experimento factorial cuando se supone un diseño completamente aleatorizado y en realidad no se hizo tal aleatorización. Por ejemplo, es común que se encuentren factores que son **muy difíciles de cambiar**. Como resultado, podría ser necesario mantener sin cambio los niveles de factores durante largos periodos a lo largo de todo el experimento. El ejemplo más común es el factor temperatura. Subirla o bajarla en un esquema aleatorio es un plan costoso y la mayoría de los experimentadores evitarán hacerlo. Los diseños experimentales con **restricciones en la aleatorización** son muy comunes y reciben el nombre de **diseños de gráficas separadas**. Esos diseños rebasan el alcance de este libro, pero en Montgomery (2008a) se encuentra su presentación.

Muchos de los conceptos que se analizaron en este capítulo se utilizarán en el capítulo 15, por ejemplo, la importancia de la aleatorización y el papel que desempeña la interacción en la interpretación de los resultados. Sin embargo, en el capítulo 15 se cubren 2 áreas que representan una expansión de los principios que se estudiaron en este capítulo y en el capítulo 13. En el capítulo 15 la solución de problemas con el uso de experimentos factoriales se realiza por medio del análisis de regresión, ya que se supone que la mayoría de los factores son cuantitativos y que se miden en un continuo, como la temperatura y el tiempo. Se derivan ecuaciones de predicción a partir de los datos del experimento diseñado y se utilizan para la mejora de procesos o incluso para su optimización. Además, se estudia el tema de los factoriales fraccionarios, en los que sólo una parte o fracción de todo el experimento factorial se aplica debido al costo excesivo que implica la realización de todo el experimento.

Capítulo 15

Experimentos factoriales 2^k y fracciones

15.1 Introducción

Ya se han expuesto ciertos conceptos del diseño experimental. El plan de muestreo para la prueba t simple sobre la media de una población normal y el análisis de varianza implican la asignación aleatoria de los tratamientos preseleccionados a las unidades experimentales. El diseño de bloques aleatorizados, en el que los tratamientos se asignan a las unidades dentro de bloques relativamente homogéneos implica una aleatorización restringida.

En este capítulo se presta atención especial a los diseños experimentales en los que el plan experimental requiere estudiar el efecto sobre una respuesta de k factores, cada uno en dos niveles. A éstos se les conoce como **experimentos factoriales 2^k** . Es frecuente que los niveles se denoten por “alto” y “bajo”, aunque esa notación podría ser arbitraria en el caso de variables cualitativas. El diseño factorial completo requiere que cada nivel de cada factor ocurra con cada nivel de cada uno de los demás factores, lo que da un total de 2^k **combinaciones de tratamientos**.

Filtrado de factores y experimentación secuencial

A menudo, cuando se realizan experimentos, ya sea en una investigación o a un nivel de desarrollo, un diseño experimental bien planeado corresponde a una **etapa** de lo que en realidad es el **plan secuencial** de la experimentación. Lo más frecuente al comienzo de un estudio es que los científicos e ingenieros no estén conscientes de cuáles factores son importantes ni de cuáles son los rangos apropiados para los factores potenciales sobre los que deben realizar la experimentación. Por ejemplo, en el libro *Response Surface Methodology*, Myers, Montgomery y Anderson-Cook (2009) dan un ejemplo de una investigación realizada en una planta piloto, la cual incluye un experimento en el que cuatro factores, temperatura, presión, concentración de formaldehído y tasa de agitación, se varían para establecer su influencia sobre la respuesta, es decir, la tasa de filtración de cierto producto químico. Incluso al nivel de planta piloto los científicos no están seguros respecto a si deben incluir los 4 factores en el modelo. Además, el objetivo final consiste en determinar la configuración adecuada de los factores contribuyentes que maximice la tasa de filtración. Por lo tanto, es necesario determinar

la **región apropiada de experimentación**. Estas preguntas sólo pueden responderse si todo el plan experimental se realiza en forma secuencial. Muchos procesos experimentales son planes que implican un *aprendizaje iterativo*, el tipo de aprendizaje consistente con el método científico, en el que la palabra *iterativo* implica experimentación por etapas.

Por lo común la primera etapa del plan secuencial ideal es variable o de **filtrado de factores**, un procedimiento que implica un diseño experimental de bajo costo en el que se utilizan **factores candidatos**. Esto es especialmente importante cuando el plan requiere un sistema complejo, como un proceso de manufactura. La información obtenida a partir de los resultados de un *diseño de filtrado* se emplea para diseñar uno o más experimentos posteriores, en los que se realizan ajustes de los factores importantes, los cuales proporcionan mejoras en el sistema o en el proceso.

Los experimentos factoriales 2^k y fracciones de 2^k son poderosas herramientas que constituyen diseños de filtrado ideales; son sencillos y prácticos, y atraen por intuición. Muchos de los conceptos generales que se estudian en el capítulo 14 siguen siendo válidos. Sin embargo, hay métodos gráficos que brindan información útil para el análisis de los diseños de 2 niveles.

Diseños de selección para cantidades grandes de factores

Cuando k es pequeña, digamos $k = 2$ o incluso $k = 3$, es evidente la utilidad del factorial 2^k para el filtrado de factores. Tanto el análisis de varianza como el de regresión, que se estudiaron e ilustraron en los capítulos 12, 13 y 14, continúan siendo herramientas útiles. Además, los enfoques gráficos también pueden ser de ayuda.

Si k es grande, por ejemplo 6, 7 u 8, el número de combinaciones de factores y, por lo tanto, de corridas experimentales necesarias para el factorial 2^k con frecuencia se vuelve prohibitivo. Por ejemplo, suponga que hay interés en realizar un diseño de selección que involucre $k = 8$ factores. Podría desearse obtener información acerca de todos los $k = 8$ efectos principales, así como de las $\frac{k(k-1)}{2} = 28$ interacciones de dos factores. Sin embargo, incluso $2^8 = 256$ corridas parecería que hace al estudio demasiado grande y excesivo para estudiar $28 + 8 = 36$ efectos. No obstante, como se verá en secciones posteriores, cuando k es grande es posible obtener gran cantidad de información de manera eficaz usando sólo una fracción del experimento factorial 2^k completo. Esta clase de diseños constituye la clase de *diseños factoriales fraccionarios*. La meta consiste en recuperar información de alta calidad acerca de los efectos principales y las interacciones interesantes, aun cuando el tamaño del diseño se reduzca en forma considerable.

15.2 El factorial 2^k : cálculo de efectos y análisis de varianza

Considere inicialmente un factorial 2^2 con factores A y B , y n *observaciones experimentales por combinación de factores*. Es útil emplear los símbolos (1), a , b y ab para denotar los puntos del diseño, donde la presencia de una letra minúscula implica que el factor (A o B) está en el *nivel alto*. Así, la ausencia de la minúscula implica que el factor está en el *nivel bajo*. Por lo que ab es el punto de diseño (+, +), a es (+, -), b es (-, +) y (1) es (-, -). Asimismo existen situaciones en las que la notación también se aplica

para los datos de respuesta en el punto de diseño en cuestión. Como introducción al cálculo de **efectos importantes** que ayuden a determinar la influencia de los factores y **sumas de cuadrados** que están incorporados en los cálculos del análisis de varianza se presenta la tabla 15.1.

Tabla 15.1: Un experimento factorial 2²

		A		Media
B	{	b	ab	$\frac{b+ab}{2n}$
		(1)	a	$\frac{(1)+a}{2n}$
Media		$\frac{(1)+b}{2n}$	$\frac{a+ab}{2n}$	

En esta tabla, (1), a, b y ab representan totales de los n valores de la respuesta en los puntos de diseño individuales. La simplicidad del factorial 2² reside en el hecho de que, aparte del error experimental, el analista obtiene la información importante en componentes con un solo grado de libertad, uno para cada uno de los dos efectos principales A y B, y un grado de libertad para la interacción AB. La información que se recupera sobre todos estos aspectos adopta la forma de tres **contrastos**. Se definirán los siguientes contrastes entre los totales de los tratamientos:

$$\text{contraste } A = ab + a - b - (1),$$

$$\text{contraste } B = ab - a + b - (1),$$

$$\text{contraste } AB = ab - a - b + (1).$$

Los tres **efectos** del experimento implican estos contrastes y apelan al sentido común y a la intuición. Los dos efectos principales calculados tienen la forma

$$\text{efecto} = \bar{y}_H - \bar{y}_L,$$

donde \bar{y}_H y \bar{y}_L son las respuestas promedio en el nivel alto o “+” y en el nivel bajo o “-”, respectivamente. Como resultado,

Cálculo de los
efectos
principales

y

$$A = \frac{ab + a - b - (1)}{2n} = \frac{\text{contraste } A}{2n}$$

$$B = \frac{ab - a + b - (1)}{2n} = \frac{\text{contraste } B}{2n}.$$

La cantidad A es considerada la *diferencia entre la respuesta media en los niveles alto y bajo del factor A*. De hecho, A se denomina **efecto principal** del factor A. En forma similar, B es el efecto principal del factor B. Al inspeccionar la diferencia entre $ab - b$ y $a - (1)$ o entre $ab - a$ y $b - (1)$ en la tabla 15.1, se observa una aparente interacción en los datos. Si, por ejemplo,

$$ab - a \approx b - (1) \quad \text{o bien} \quad ab - a - b + (1) \approx 0,$$

una recta que conecta las respuestas para cada nivel del factor A en el nivel alto del factor B será aproximadamente paralela a una recta que conecte la respuesta para cada nivel del factor A en el nivel bajo del factor B . Las rectas no paralelas de la figura 15.1 sugieren la presencia de interacción. Para probar que esta interacción aparente es significativa se construye un tercer contraste en los totales del tratamiento, ortogonal a los contrastes del efecto principal, al cual se denomina **efecto de interacción**. La construcción del tercer contraste mencionado se realiza evaluando

Efecto de interacción

$$AB = \frac{ab - a - b + (1)}{2n} = \frac{\text{contraste } AB}{2n}.$$

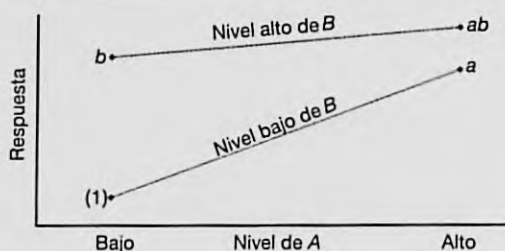


Figura 15.1: Respuesta que sugiere una interacción aparente.

Ejemplo 15.1: Considere los datos de las tablas 15.2 y 15.3 con $n = 1$ para un experimento factorial 2^2 .

Tabla 15.2: Factorial 2^2 sin interacción

A	B	
	-	+
+	50	70
-	80	100

Tabla 15.3: Factorial 2^2 con interacción

A	B	
	-	+
+	50	70
-	80	40

Los números en las celdas de las tablas 15.2 y 15.3 ilustran con claridad la manera en que los contrastes y el cálculo resultante de los dos efectos principales y de las conclusiones resultantes pueden estar muy influidos por la presencia de interacción. En la tabla 15.2 el efecto de A es -30 tanto en el nivel bajo como en el nivel alto del factor B , y el efecto de B es 20 en los niveles bajo y alto del factor A . Esta "consistencia del efecto" (no hay interacción) puede ser información muy importante para el analista. Los efectos principales son

$$A = \frac{70 + 50}{2} - \frac{100 + 80}{2} = 60 - 90 = -30,$$

$$B = \frac{100 + 70}{2} - \frac{80 + 50}{2} = 85 - 65 = 20,$$

mientras que el efecto de la interacción es

$$AB = \frac{100 + 50}{2} - \frac{80 + 70}{2} = 75 - 75 = 0.$$

Por otro lado, en la tabla 15.3 el efecto A es nuevamente -30 al nivel bajo de B , pero $+30$ al nivel alto de B . Esta “inconsistencia del efecto” (interacción) también está presente para B en todos los niveles de A . En estos casos los efectos principales podrían carecer de significado y, de hecho, prestarse mucho a la confusión. Por ejemplo, el efecto de A es

$$A = \frac{50 + 70}{2} - \frac{80 + 40}{2} = 0,$$

ya que hay un “enmascaramiento” completo del efecto conforme se promedia sobre los niveles de B . La fuerte interacción se ilustra con el efecto calculado

$$AB = \frac{70 + 80}{2} - \frac{50 + 40}{2} = 30.$$

Aquí es conveniente ilustrar los escenarios de las tablas 15.2 y 15.3 con las gráficas de interacción. Observe el paralelismo en la gráfica de la figura 15.2 y la interacción aparente en la figura 15.3.

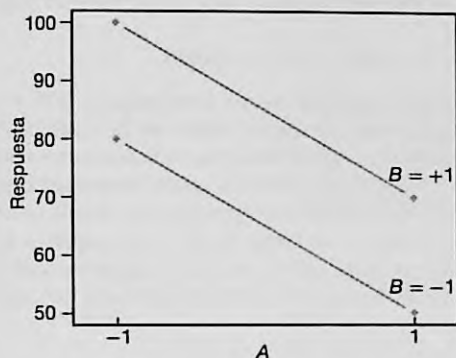


Figura 15.2: Gráfica de interacción para los datos de la tabla 15.2.

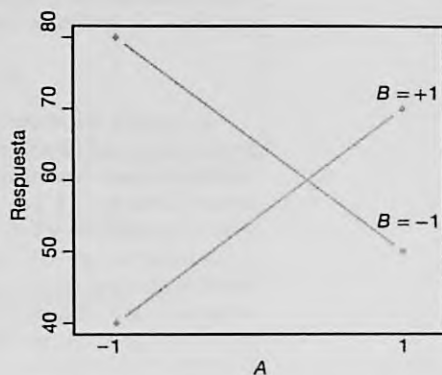


Figura 15.3: Gráfica de interacción para los datos de la tabla 15.3.

Cálculo de las sumas de cuadrados

Se aprovecha el hecho de que en el factorial 2^k, o para el caso en el experimento factorial 2^k general, cada efecto principal y efecto de interacción tiene asociado **un solo grado de libertad**. Por lo tanto, es posible escribir contrastes ortogonales 2^k - 1 con un solo grado de libertad en las combinaciones de tratamientos, donde cada uno es responsable de la variación debida a cierto efecto principal o interacción. Así, con base en las suposiciones usuales de independencia y normalidad en el modelo experimental, se hacen pruebas para determinar si el contraste refleja variación sistemática, o bien, sólo variaciones probabilísticas o aleatorias. Las sumas de cuadrados para cada contraste se calculan siguiendo los procedimientos que se estudiaron en la sección 13.5. Si se escribe

$$Y_{1..} = b + (1), \quad Y_{2..} = ab + a, \quad c_1 = -1 \quad \text{y} \quad c_2 = 1,$$

donde $Y_{1..}$ y $Y_{2..}$ constituyen el total de $2n$ observaciones, se tiene

$$SCA = SC_{w_A} = \frac{\left(\sum_{i=1}^2 c_i Y_{i..}\right)^2}{2n \sum_{i=1}^2 c_i^2} = \frac{[ab + a - b - (1)]^2}{2^2 n} = \frac{(\text{contraste } A)^2}{2^2 n},$$

con 1 grado de libertad. De forma similar, se encuentra que

$$SCB = \frac{[ab + b - a - (1)]^2}{2^2 n} = \frac{(\text{contraste } B)^2}{2^2 n}$$

y

$$SC(AB) = \frac{[ab + (1) - a - b]^2}{2^2 n} = \frac{(\text{contraste } AB)^2}{2^2 n}.$$

Cada contraste tiene 1 grado de libertad, mientras que las sumas de cuadrados del error, con $2^2(n-1)$ grados de libertad, se obtienen mediante una resta a partir de la fórmula

$$SCE = SCT - SCA - SCB - SC(AB).$$

Al calcular las sumas de cuadrados para los efectos principales A y B , y el efecto de interacción AB , es conveniente presentar las salidas totales de las combinaciones de tratamiento junto con los signos algebraicos apropiados para cada contraste, como se observa en la tabla 15.4. Los efectos principales se obtienen como comparaciones simples entre los niveles alto y bajo. Por lo tanto, se asigna un signo positivo para la combinación de tratamientos que esté en el nivel alto de un factor dado, y uno negativo a la combinación de tratamientos del nivel bajo. Los signos positivo y negativo para el efecto de interacción se obtienen multiplicando los signos correspondientes de los contrastes de los factores de la interacción.

Tabla 15.4: Signos para los contrastes en un experimento factorial 2²

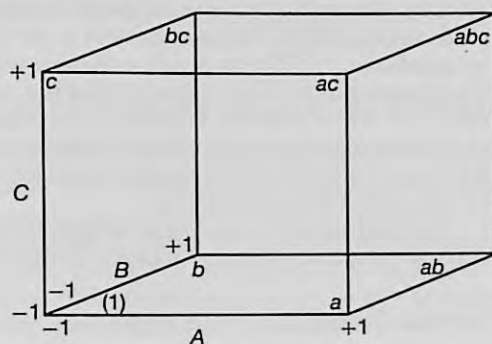
Combinación de tratamientos	Efecto factorial		
	A	B	AB
(1)	-	-	+
a	+	-	-
b	-	+	-
ab	+	+	+

El factorial 2³

Ahora consideremos un experimento en el que intervienen tres factores, A , B y C , cada uno con niveles -1 y $+1$. Se trata de un experimento factorial 2³ que proporciona ocho combinaciones de tratamientos (1), a , b , c , ab , ac , bc y abc . En la tabla 15.5 se presentan las combinaciones de tratamientos y los signos algebraicos apropiados para cada contraste que se usan en el cálculo de las sumas de los cuadrados para los efectos principales y los efectos de interacción.

Tabla 15.5: Signos de los contrastes en un experimento factorial 2³

Combinación de tratamiento	Efecto factorial (simbólico)						
	A	B	C	AB	AC	BC	ABC
(1)	-	-	-	+	+	+	-
a	+	-	-	-	-	+	+
b	-	+	-	-	+	-	+
c	-	-	+	+	-	-	+
ab	+	+	-	+	-	-	-
ac	+	-	+	-	+	-	-
bc	-	+	+	-	-	+	-
abc	+	+	+	+	+	+	+

Figura 15.4: Vista geométrica de 2³.

Es útil analizar e ilustrar la geometría del factorial 2³ del mismo modo que se hizo para el factorial 2² en la figura 15.1. Para el 2³ los **ocho puntos de diseño** representan los vértices de un cubo, como se observa en la figura 15.4.

Las columnas de la tabla 15.5 representan los signos que se utilizan para los contrastes, así como los cálculos de siete efectos y las sumas de cuadrados correspondientes. Estas columnas son análogas a las que se observan en la tabla 15.4 para el caso de 2². Como son ocho puntos de diseño hay siete efectos disponibles. Por ejemplo,

$$A = \frac{a + ab + ac + abc - (1) - b - c - bc}{4n},$$

$$AB = \frac{(1) + c + ab + abc - a - b - ac - bc}{4n},$$

y así sucesivamente. Las sumas de cuadrados son dadas por

$$SC(\text{efecto}) = \frac{(\text{contraste})^2}{2^3 n}.$$

Al observar la tabla 15.5 se revela que para el experimento 2³ todos los contrastes

entre los siete son mutuamente ortogonales y, por lo tanto, los siete efectos se evalúan en forma independiente.

Efectos y sumas de cuadrados para el 2^k

Para un experimento factorial 2^k las sumas de cuadrados de un solo grado de libertad para los efectos principales y los efectos de interacción se obtienen elevando al cuadrado los contrastes apropiados en los totales del tratamiento y dividiendo entre $2^k n$, donde n es el número de réplicas de las combinaciones del tratamiento.

Como antes, un efecto siempre se calcula restando la respuesta promedio en el nivel "bajo", de la respuesta promedio en el nivel "alto". Quedan muy claros los niveles alto y bajo para los efectos principales. Los niveles alto y bajo simbólicos para las interacciones son evidentes a partir de la información de la tabla 15.5.

La propiedad de ortogonalidad tiene la misma importancia aquí que en el material sobre las comparaciones que se estudió en el capítulo 13. La ortogonalidad de los contrastes implica que los efectos estimados y, por lo tanto, las sumas de cuadrados, son independientes. Esta independencia se ilustra con claridad en el experimento factorial 2^3 si las respuestas, con el factor A en su nivel alto, se incrementan en una cantidad x en la tabla 15.5. Sólo el contraste A conduce a una suma de cuadrados más grande, ya que el efecto x se cancela cuando se forman los seis contrastes restantes como resultado de los dos signos positivos y los dos negativos asociados con las combinaciones de tratamientos en los que A está en el nivel alto.

La ortogonalidad produce otras ventajas, las cuales se abordarán cuando se estudie el experimento factorial 2^k en situaciones de regresión.

15.3 Experimento factorial 2^k sin réplicas

El factorial completo 2^k con frecuencia requiere mucha experimentación, en particular cuando k es grande. Como resultado, a menudo no es posible replicar cada combinación de factores. Si en el modelo del experimento se incluyen todos los efectos, con todas las interacciones, no se permite ningún grado de libertad para el error. A menudo, cuando k es grande, el analista de datos *agrupará* las sumas de los cuadrados y los grados de libertad correspondientes para las interacciones de orden superior que se sabe, o se supone, son despreciables. Esto producirá pruebas F para los efectos principales e interacciones de orden inferior.

Graficación de diagnóstico con experimentos factoriales 2^k sin réplicas

Las gráficas de probabilidad normal constituyen una metodología muy útil para determinar la importancia relativa de los efectos en un experimento con factores de dos niveles razonablemente grandes cuando no hay réplica. Este tipo de gráfica de diagnóstico puede ser útil sobre todo cuando el analista de datos duda en agrupar interacciones de orden superior por temor de agrupar en el "error" algunos efectos verdaderamente reales y no sólo aleatorios. El lector debe recordar que todos los efectos que no son reales, es decir, que son *estimados de cero* independientes, siguen una distribución normal con media cercana a cero y varianza constante. Por ejemplo, en un experimento factorial 2^4 se debe recordar que todos los efectos, teniendo en cuenta que $n = 1$, son de la forma

$$AB = \frac{\text{contraste}}{8} = \bar{y}_H - \bar{y}_L,$$

donde \bar{y}_H es el promedio de ocho corridas experimentales independientes en el nivel alto, o "+", y \bar{y}_L es el promedio de ocho corridas independientes en el nivel bajo, o "-". Así, la varianza de cada contraste es $\text{Var}(\bar{y}_H - \bar{y}_L) = \sigma^2/4$. Para cualesquiera efectos reales $E(\bar{y}_H - \bar{y}_L) \neq 0$. Así, la gráfica de probabilidad normal debería revelar efectos "significativos" como aquellos que caen fuera de la línea recta que describe realizaciones de variables aleatorias normales independientes distribuidas de forma idéntica.

La gráfica de probabilidad puede adoptar una de muchas formas. Se recomienda al lector que consulte el capítulo 8, en el que se presentaron dichas gráficas por primera vez. Se puede usar la gráfica cuantil-cuantil, normal y empírica. También es posible utilizar el procedimiento de graficación que emplea el papel de probabilidad normal. Además, existen otros tipos de gráficas de probabilidad normal para el diagnóstico. En resumen, las gráficas de efectos para el diagnóstico son como sigue.

Gráficas de efectos de probabilidad para experimentos factoriales 2^k sin réplica

1. Calcular los efectos como

$$\text{efecto} = \frac{\text{contraste}}{2^{k-1}}$$

2. Construir una gráfica de probabilidad normal de todos los efectos.

3. Los efectos que caigan fuera de la línea recta deben considerarse reales.

A continuación se hacen más comentarios respecto de las gráficas de probabilidad normal de los efectos. En primer lugar, el analista podría sentirse frustrado si utiliza las gráficas con un experimento pequeño. Por otro lado, la graficación puede proporcionar resultados satisfactorios cuando hay *dispersión de efectos*, muchos efectos que no son verdaderamente reales. Esta dispersión será evidente en experimentos grandes, en los que es poco probable que las interacciones de orden superior sean reales.

Estudio de caso 15.1:

Moldeado por inyección. Muchas empresas fabricantes de Estados Unidos y otros países utilizan partes moldeadas como componentes de un proceso. Un problema grande que enfrentan con frecuencia es el rebasamiento. A menudo, un molde troquelado de una parte se construye con un tamaño más grande que el nominal para permitir que se contraiga. En la siguiente situación experimental se produce un molde nuevo para el cual es importante encontrar las especificaciones adecuadas del proceso para minimizar la contracción. En el siguiente experimento los valores de la respuesta son desviaciones de los nominales, es decir, contracciones. Los factores y niveles son los siguientes:

	Niveles codificados	
	-1	+1
A. Velocidad de inyección (pies/seg)	1.0	2.0
B. Temperatura de moldeado (°C)	100	150
C. Presión de moldeado (psi)	500	1000
D. Contrapresión (psi)	75	120

El propósito del experimento fue determinar cuáles efectos (principales y de interacción) influyen en la contracción. El experimento se consideró un filtrado preliminar a partir del cual se determinaron los factores para un análisis más completo. Asimismo, se espera obtener información respecto a cómo los factores importantes repercuten en la contracción. En la tabla 15.6 se presentan los datos de un experimento factorial 2^k sin réplica.

Tabla 15.6: Datos para el estudio de caso 15.1

Combinación de factores	Respuesta (cm × 10 ⁴)	Combinación de factores	Respuesta (cm × 10 ⁴)
(1)	72.68	<i>d</i>	73.52
<i>a</i>	71.74	<i>ad</i>	75.97
<i>b</i>	76.09	<i>bd</i>	74.28
<i>ab</i>	93.19	<i>abd</i>	92.87
<i>c</i>	71.25	<i>cd</i>	79.34
<i>ac</i>	70.59	<i>acd</i>	75.12
<i>bc</i>	70.92	<i>bcd</i>	79.67
<i>abc</i>	104.96	<i>abcd</i>	97.80

Inicialmente se calcularon los efectos y se plasmaron en una gráfica de probabilidad normal. Los efectos calculados son los siguientes:

$$\begin{aligned}
 A &= 10.5613, & BD &= -2.2787, & B &= 12.4463, \\
 C &= 2.4138, & D &= 2.1438, & AB &= 11.4038, \\
 AC &= 1.2613, & AD &= -1.8238, & BC &= 1.8163, \\
 CD &= 1.4088, & ABC &= 2.8588, & ABD &= -1.7813, \\
 ACD &= -3.0438, & BCD &= -0.4788, & ABCD &= -1.3063.
 \end{aligned}$$

En la figura 15.5 se observa la gráfica cuantil-cuantil normal, la cual parece implicar que los efectos *A*, *B* y *AB* son importantes. Los signos de los efectos importantes indican las conclusiones preliminares.

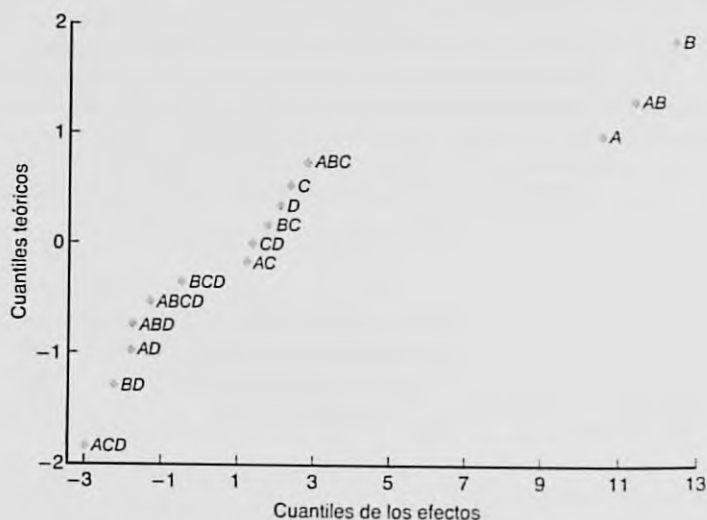


Figura 15.5: Gráfica cuantil-cuantil normal de los efectos para el estudio de caso del ejemplo 15.1.

1. Un incremento en la velocidad de inyección de 1.0 a 2.0 aumenta la contracción.
2. Un aumento en la temperatura de moldeo de 100°C a 150°C incrementa la contracción.
3. Hay una interacción entre la velocidad de inyección y la temperatura del moldeo; aunque ambos efectos principales son importantes es crucial entender el efecto de la interacción de los dos factores. ▬

Interpretación de la interacción de dos factores

Como se esperaría, una tabla de medias de dos factores facilita la interpretación de la interacción *AB*. Considere la situación de dos factores de la tabla 15.7.

Tabla 15.7: Ilustración de una interacción de dos factores

A (velocidad)	B (temperatura)	
	100	150
2	73.355	97.205
1	74.1975	75.240

Observe que la media muestral grande a velocidad y temperatura elevadas creó la interacción significativa. La **contracción se incrementa en forma no aditiva**. La temperatura del moldeo parece tener un efecto positivo a pesar del nivel de velocidad. Sin embargo, el efecto es el mayor a velocidad elevada. El efecto de la velocidad es muy ligero a temperaturas bajas, pero es claramente positivo a una temperatura elevada de moldeo. Para controlar la contracción a bajo nivel *debería evitarse el uso simultáneo de una alta velocidad de inyección y una temperatura de moldeo elevada*. Todos estos resultados se ilustran en forma gráfica en la figura 15.6.

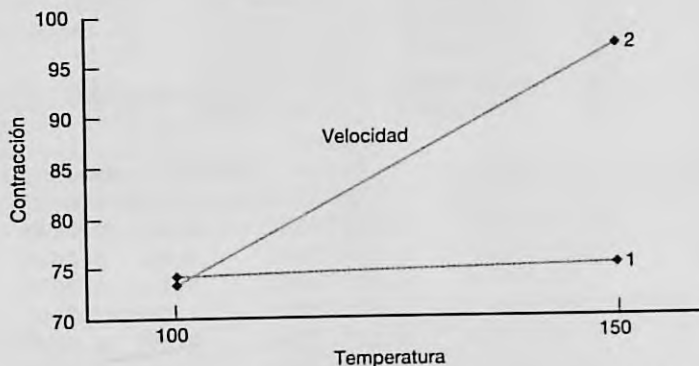


Figura 15.6: Gráfica de la interacción para el estudio de caso 15.1.

Análisis con el cuadrado medio del error agrupado: salida de resultados por computadora comentada

Puede ser de interés observar un análisis de varianza de los datos del moldeado por inyección con interacciones de orden superior agrupadas para formar un cuadrado medio del error. Las interacciones de órdenes tres y cuatro están agrupadas. En la figura 15.7 se observa una salida de resultados por computadora de la función PROC GLM del SAS. El análisis de varianza revela, en esencia, la misma conclusión que la gráfica de probabilidad normal.

Las pruebas y los valores *P* que se observan en la figura 15.7 requieren una interpretación. Un valor *P* significativo sugiere que el efecto difiere de cero en forma significativa. Las pruebas sobre los efectos principales (que en presencia de las interacciones se pueden considerar como los efectos promediados sobre los niveles de los demás factores) indican la significancia de los efectos *A* y *B*. Los signos de los efectos también son importantes. Un aumento en el nivel de bajo a alto en *A*, la velocidad de inyección, ocasiona

The GLM Procedure						
Dependent Variable: y						
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	
Model	10	1689.237462	168.923746	9.37	0.0117	
Error	5	90.180831	18.036166			
Corrected Total	15	1779.418294				
R-Square		Coeff Var	Root MSE	y Mean		
0.949320		5.308667	4.246901	79.99938		
Source	DF	Type III SS	Mean Square	F Value	Pr > F	
A	1	446.1600062	446.1600062	24.74	0.0042	
B	1	619.6365563	619.6365563	34.36	0.0020	
C	1	23.3047563	23.3047563	1.29	0.3072	
D	1	18.3826563	18.3826563	1.02	0.3590	
A*B	1	520.1820562	520.1820562	28.84	0.0030	
A*C	1	6.3630063	6.3630063	0.35	0.5784	
A*D	1	13.3042562	13.3042562	0.74	0.4297	
B*C	1	13.1950562	13.1950562	0.73	0.4314	
B*D	1	20.7708062	20.7708062	1.15	0.3322	
C*D	1	7.9383063	7.9383063	0.44	0.5364	
Parameter	Estimate	Standard Error	t Value	Pr > t		
Intercept	79.99937500	1.06172520	75.35	<.0001		
A	5.28062500	1.06172520	4.97	0.0042		
B	6.22312500	1.06172520	5.86	0.0020		
C	1.20687500	1.06172520	1.14	0.3072		
D	1.07187500	1.06172520	1.01	0.3590		
A*B	5.70187500	1.06172520	5.37	0.0030		
A*C	0.63062500	1.06172520	0.59	0.5784		
A*D	-0.91187500	1.06172520	-0.86	0.4297		
B*C	0.90812500	1.06172520	0.86	0.4314		
B*D	-1.13937500	1.06172520	-1.07	0.3322		
C*D	0.70437500	1.06172520	0.66	0.5364		

Figura 15.7: Salida de resultados por computadora del SAS para los datos del estudio de caso 15.1.

un incremento en la contracción. Lo mismo es verdad para *B*. Sin embargo, debido a la interacción significativa *AB*, las interpretaciones del efecto principal podrían considerarse como tendencias en todos los niveles de los demás factores. El impacto de la interacción *AB* significativa se entiende mejor si se emplea una tabla de medias de dos factores.

Ejercicios

15.1 Los siguientes datos se obtuvieron de un experimento factorial 2^3 que se replicó tres veces. Utilice el método del contraste para evaluar las sumas de cuadrados de todos los efectos factoriales. Saque sus conclusiones.

Combinación de tratamientos	Réplica		
	1	2	3
(1)	12	19	10
<i>a</i>	15	20	16
<i>b</i>	24	16	17
<i>ab</i>	23	17	27
<i>c</i>	17	25	21
<i>ac</i>	16	19	19
<i>bc</i>	24	23	29
<i>abc</i>	28	25	20

15.2 En un experimento efectuado por el Departamento de Ingeniería de Minas de Virginia Tech con el fin de estudiar un sistema de filtrado particular para carbón se agregó un coagulante a la solución contenida en un tanque con carbón y sedimentos, que luego se puso en un sistema de recirculación para purificar el carbón. En el proceso experimental se variaron tres factores:

Factor A: porcentaje de sólidos que circularon inicialmente en el sobreflujo

Factor B: tasa de flujo del polímero

Factor C: pH del tanque

La cantidad de sólidos en el flujo inferior del sistema de purificación determina qué tan puro ha quedado el carbón. Se emplearon dos niveles de cada factor y se hicieron dos corridas experimentales para cada una de las $2^3 = 8$ combinaciones. En la siguiente tabla se especifican las mediciones de respuesta en porcentajes de sólidos por peso:

Combinación de tratamientos	Respuesta	
	Réplica 1	Réplica 2
(1)	4.65	5.81
<i>a</i>	21.42	21.35
<i>b</i>	12.66	12.56
<i>ab</i>	18.27	16.62
<i>c</i>	7.93	7.88
<i>ac</i>	13.18	12.87
<i>bc</i>	6.51	6.26
<i>abc</i>	18.23	17.83

Suponga que todas las interacciones son potencialmente importantes y con base en esto haga un análisis completo de los datos. Use valores *P* en la conclusión.

15.3 En un experimento metalúrgico se desea probar el efecto de cuatro factores y sus interacciones sobre la concentración (porcentaje por peso) de cierto compuesto particular de fósforo en el material de fundición. Las variables son *A*, porcentaje de fósforo en la refinación; *B*, porcentaje del material vuelto a fundir; *C*, tiempo de flujo; y *D*, tiempo de espera. Se varían los cuatro factores en un experimento factorial 2^4 , con dos fundiciones tomadas de cada combinación de factores. Las 32 fundiciones se hicieron en orden aleatorio. Los datos se muestran en la siguiente tabla, y en la figura 15.8 de la página 610 se incluye la tabla del ANOVA. Analice los efectos de los factores y sus interacciones sobre la concentración del compuesto de fósforo.

Combinación de tratamientos	Peso		
	% de compuesto	Réplica 1	Réplica 2
(1)	30.3	28.6	58.9
<i>a</i>	28.5	31.4	59.9
<i>b</i>	24.5	25.6	50.1
<i>ab</i>	25.9	27.2	53.1
<i>c</i>	24.8	23.4	48.2
<i>ac</i>	26.9	23.8	50.7
<i>bc</i>	24.8	27.8	52.6
<i>abc</i>	22.2	24.9	47.1
<i>d</i>	31.7	33.5	65.2
<i>ad</i>	24.6	26.2	50.8
<i>bd</i>	27.6	30.6	58.2
<i>abd</i>	26.3	27.8	54.1
<i>cd</i>	29.9	27.7	57.6
<i>acd</i>	26.8	24.2	51.0
<i>bcd</i>	26.4	24.9	51.3
<i>abcd</i>	26.9	29.3	56.2
Total	428.1	436.9	865.0

15.4 Se realizó un experimento preliminar para estudiar los efectos de cuatro factores y sus interacciones sobre la producción de la operación de cierta máquina. Se realizan dos corridas de cada una de las combinaciones de tratamientos para obtener una medida del error experimental puro. Se emplean dos niveles de cada factor y se obtienen los datos que se observan en la siguiente página. Pruebe todos los efectos principales y las interacciones a un nivel de significancia de 0.05. Saque sus conclusiones.

Fuente de variación	Efectos	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada	Valor P
Efecto principal:						
A	-1.2000	11.52	1	11.52	4.68	0.0459
B	-1.2250	12.01	1	12.01	4.88	0.0421
C	-2.2250	39.61	1	39.61	16.10	0.0010
D	1.4875	17.70	1	17.70	7.20	0.0163
Interacción de dos factores:						
AB	0.9875	7.80	1	7.80	3.17	0.0939
AC	0.6125	3.00	1	3.00	1.22	0.2857
AD	-1.3250	14.05	1	14.05	5.71	0.0295
BC	1.1875	11.28	1	11.28	4.59	0.0480
BD	0.6250	3.13	1	3.13	1.27	0.2763
CD	0.7000	3.92	1	3.92	1.59	0.2249
Interacción de tres factores:						
ABC	-0.5500	2.42	1	2.42	0.98	0.3360
ABD	1.7375	24.15	1	24.15	9.82	0.0064
ACD	1.4875	17.70	1	17.70	7.20	0.0163
BCD	-0.8625	5.95	1	5.95	2.42	0.1394
Interacción de cuatro factores:						
ABCD	0.7000	3.92	1	3.92	1.59	0.2249
Error		39.36	16	2.46		
Total		217.51	31			

Figura 15.8: Tabla ANOVA para el ejercicio 15.3.

Combinación de tratamientos	Réplica 1	Réplica 2	Observación	Estado físico	Tiempo de mezclado	Vel. de las aspas	Condición de nitrógeno	Aluminio
(1)	7.9	9.6	1	1	1	2	2	16.3
a	9.1	10.2	2	1	2	2	2	16.0
b	8.6	5.8	3	1	1	1	1	16.2
c	10.4	12.0	4	1	2	1	2	16.1
d	7.1	8.3	5	1	1	1	2	16.0
ab	11.1	12.3	6	1	2	1	1	16.0
ac	16.4	15.5	7	1	2	2	1	15.5
ad	7.1	8.7	8	1	1	2	1	15.9
bc	12.6	15.2	9	2	1	2	2	16.7
bd	4.7	5.8	10	2	2	2	2	16.1
cd	7.4	10.9	11	2	1	1	1	16.3
abc	21.9	21.9	12	2	2	1	2	15.8
abd	9.8	7.8	13	2	1	1	2	15.9
acd	13.8	11.2	14	2	2	1	1	15.9
bcd	10.2	11.1	15	2	2	2	1	15.6
abcd	12.8	14.3	16	2	1	2	1	15.8

15.5 En el estudio *An X-Ray Fluorescence Method for Analyzing Polybutadiene-Acrylic Acid (PBAA) Propellants* (Quarterly Reports, RK-TR-62-1, Army Ordnance Missile Command) se realizó un experimento para determinar si existe o no una diferencia significativa en la cantidad de aluminio obtenido en

un análisis con ciertos niveles de ciertas variables de procesamiento. A continuación se presentan los datos.

Observación	Estado físico	Tiempo de mezclado	Vel. de las aspas	Condición de nitrógeno	Aluminio
1	1	1	2	2	16.3
2	1	2	2	2	16.0
3	1	1	1	1	16.2
4	1	2	1	2	16.1
5	1	1	1	2	16.0
6	1	2	1	1	16.0
7	1	2	2	1	15.5
8	1	1	2	1	15.9
9	2	1	2	2	16.7
10	2	2	2	2	16.1
11	2	1	1	1	16.3
12	2	2	1	2	15.8
13	2	1	1	2	15.9
14	2	2	1	1	15.9
15	2	2	2	1	15.6
16	2	1	2	1	15.8

Las variables para los datos son:

- A: tiempo de mezcla
 - nivel 1: 2 horas
 - nivel 2: 4 horas

- B:** Velocidad de las aspas
 nivel 1: 36 rpm
 nivel 2: 78 rpm
- C:** Condición de nitrógeno que pasa por el propulsor
 nivel 1: seco
 nivel 2: 72% de humedad relativa
- D:** estado físico del propulsor
 nivel 1: no refinado
 nivel 2: refinado

Analice los datos suponiendo que todas las interacciones de tres y cuatro factores son despreciables. Utilice un nivel de significancia de 0.05. Escriba un breve informe que resuma sus hallazgos.

15.6 Es importante estudiar el efecto de la concentración del reactivo y la tasa de alimentación de la viscosidad del producto de cierto proceso químico. La concentración del reactivo será el factor *A* a los niveles 15% y 25%. La tasa de alimentación será el factor *B* a niveles de 20 lb/h y 30 lb/h. El experimento implica 2 corridas experimentales en cada una de las cuatro combinaciones (*L* = bajo y *H* = alto). Las lecturas de la viscosidad son las siguientes.

H	132	149
	137	152
B		
L	145	154
	147	150
	L	H
	A	

- a) Suponga un modelo que contiene dos efectos principales y una interacción y calcule los tres efectos. ¿Tiene usted alguna interpretación en este momento?
- b) Realice un análisis de varianza y haga pruebas de interacción. Saque conclusiones.
- c) Realice pruebas para los efectos principales y saque conclusiones finales acerca de la importancia de todos estos efectos.

15.7 Considere el ejercicio 15.3. Al investigador no sólo le interesa saber que las interacciones *AD*, *BC* y quizá *AB* son importantes, sino también su significado científico. Dibuje gráficas de interacción bidimensional para las tres e interpréte las.

15.8 Considere nuevamente el ejercicio 15.3. Es frecuente que las interacciones de tres factores no sean significativas y, aun si lo fueran, serían difíciles de in-

terpretar. La interacción *ABD* parece ser importante. Para hacer cierta interpretación dibuje dos gráficas de la interacción *AD*, una para *B* = -1 y otra para *B* = +1. A partir de la apariencia de éstas interprete la interacción *ABD*.

15.9 Considere el ejercicio 15.6. Utilice una escala de +1 y -1, para "alto" y "bajo", respectivamente, y calcule una regresión lineal múltiple con el modelo

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} + \epsilon_i$$

con x_{1i} = concentración del reactivo (-1, +1) y x_{2i} = tasa de alimentación (-1, +1).

- a) Calcule los coeficientes de regresión.
- b) ¿Cómo se relacionan los coeficientes b_1 , b_2 y b_{12} con los efectos que encontró en el ejercicio 15.6a)?
- c) En su análisis de regresión haga pruebas *t* sobre b_1 , b_2 y b_{12} . ¿Cómo se relacionan estos resultados de la prueba con los del ejercicio 15.6b) y c)?

15.10 Considere el ejercicio 15.5. Calcule los 15 efectos y haga gráficas de probabilidad normal de los efectos.

- a) ¿Parece válida la suposición de que las interacciones de tres y cuatro factores son despreciables?
- b) ¿Los resultados de las gráficas del efecto son consistentes con lo que usted planteó sobre la importancia de los efectos principales y las interacciones de dos factores en su informe de resumen?

15.11 En Myers, Montgomery y Anderson-Cook (2009) se analiza un conjunto de datos para el que un ingeniero empleó un factorial 2^3 con el fin de estudiar los efectos de la velocidad de corte (*A*), la geometría de la herramienta (*B*) y el ángulo de corte (*C*) sobre la vida (en horas) de una máquina. Se eligen dos niveles de cada factor y se hacen pruebas dobles en cada punto del diseño en un orden aleatorio. A continuación se presentan los datos.

	<i>A</i>	<i>B</i>	<i>C</i>	Vida
(1)	-	-	-	22, 31
<i>a</i>	+	-	-	32, 43
<i>b</i>	-	+	-	35, 34
<i>ab</i>	+	+	-	35, 47
<i>c</i>	-	-	+	44, 45
<i>ac</i>	+	-	+	40, 37
<i>bc</i>	-	+	+	60, 50
<i>abc</i>	+	+	+	39, 41

- a) Calcule los siete efectos. Con base en su magnitud, ¿cuál parece ser importante?
- b) Haga un análisis de varianza y observe los valores *P*.
- c) ¿Coinciden los resultados de los incisos *a* y *b*?

d) El ingeniero confía en que debe haber una interacción entre la velocidad y el ángulo de corte. Si esta interacción es significativa, dibuje una gráfica de la interacción y analice su significado desde el punto de vista de la ingeniería.

15.12 Considere el ejercicio 15.11 y suponga que hubo cierta dificultad experimental para hacer las corridas; que en realidad se tuvo que suspender todo el experimento después de sólo cuatro corridas. Como resultado, el experimento abreviado es dado por

	Vida
a	43
b	35
c	44
abc	39

Con sólo estas corridas los signos para los contrastes son dados por

	A	B	C	AB	AC	BC	ABC
a	+	-	-	-	-	+	+
b	-	+	-	-	+	-	+
c	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+

Comente y determine si los contrastes son o no ortogonales. ¿Cuáles lo son y cuáles no? ¿Los efectos principales son ortogonales entre sí? En ese experimento abreviado (denominado *factorial fraccionario*) ¿es posible estudiar las interacciones de los efectos principales en forma independiente? ¿Se trataría de un experimento útil si estuviéramos convencidos de que las interacciones son despreciables? Explique su respuesta.

15.4 Experimentos factoriales en un ajuste de regresión

Hasta ahora hemos limitado el análisis de los datos para un factorial 2^k al método del análisis de varianza. La única referencia a un análisis alternativo se hizo en el ejercicio 15.9 de la página 611. De hecho, este ejercicio introduce gran parte del material que da origen a la presente sección. Hay situaciones en las que el ajuste de un modelo es importante y en la que es posible controlar los factores que se estudian. Por ejemplo, un biólogo podría querer estudiar el crecimiento de cierto tipo de alga en el agua, en cuyo caso sería muy útil un modelo que relacionara las unidades de algas como una función de la cantidad de cierto contaminante, y, digamos, del tiempo. Así, el estudio involucra un experimento factorial en un ambiente de laboratorio en el que los factores son la concentración del contaminante y el tiempo. Como se verá más adelante en esta sección, es posible ajustar un modelo más preciso si los factores están controlados en un arreglo factorial, para el que con frecuencia es útil elegir un factorial 2^k. En muchos procesos biológicos y químicos los niveles de las variables regresoras pueden y deberían controlarse.

Hay que recordar que el modelo de regresión empleado en el capítulo 12 se puede escribir con notación de matriz de la siguiente manera

$$y = X\beta + \epsilon.$$

La matriz X se denomina **matriz del modelo**. Suponga, por ejemplo, que se utiliza un experimento factorial 2³ con las variables

Temperatura:	150°C	200°C
Humedad:	15%	20%
Presión (psi):	1000	1500

Los niveles familiares +1 y -1 se generan a través del siguiente centrado y escalado a unidades de diseño:

$$x_1 = \frac{\text{temperatura} - 175}{25}, \quad x_2 = \frac{\text{humedad} - 17.5}{2.5}, \quad x_3 = \frac{\text{presión} - 1250}{250}.$$

Como resultado, la matriz X se vuelve

$$X = \begin{matrix} & x_1 & x_2 & x_3 & \text{Identificación del diseño} \\ \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} & & & & \begin{matrix} (1) \\ a \\ b \\ c \\ ab \\ ac \\ bc \\ abc \end{matrix} \end{matrix}$$

Ahora se observa que los contrastes ilustrados y analizados en la sección 15.2 están relacionados directamente con los coeficientes de regresión. Observe que todas las columnas de la matriz X en el ejemplo 2³ son *ortogonales*. Como resultado, el cálculo de los coeficientes de regresión que se describió en la sección 12.3 se convierte en

$$b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = (X'X)^{-1}X'y = \left(\frac{1}{8}I\right)X'y \\ = \frac{1}{8} \begin{bmatrix} a + ab + ac + abc + (1) + b + c + bc \\ a + ab + ac + abc - (1) - b - c - bc \\ b + ab + bc + abc - (1) - a - c - ac \\ c + ac + bc + abc - (1) - a - b - ab \end{bmatrix},$$

donde a , ab , etc., son medidas de la respuesta.

Ahora se observa que el concepto de *principales efectos calculados* que se enfatiza a lo largo de todo este capítulo con diseños factoriales 2^k , se relaciona con los coeficientes de un modelo de regresión ajustado cuando los factores son cuantitativos. De hecho, para un 2^k con, digamos, n corridas experimentales por punto del diseño, las relaciones entre los efectos y los coeficientes de regresión son como sigue:

$$\text{Efecto} = \frac{\text{contraste}}{2^{k-1}(n)} \\ \text{Coeficiente de regresión} = \frac{\text{contraste}}{2^k(n)} = \frac{\text{efecto}}{2}.$$

Esta relación debería tener sentido para el lector, ya que un coeficiente de regresión b_j es una tasa promedio del cambio en la respuesta *por cambio de unidad* en x_j . Por supuesto, cuando se va de -1 a $+1$ en x_j (de bajo a alto), la variable de diseño cambia en 2 unidades.

Ejemplo 15.2: Considere un experimento donde un ingeniero desea ajustar una regresión lineal del producto y contra el tiempo de retención x_1 y el tiempo de flexión x_2 en cierto sistema químico. Todos los demás factores se mantienen fijos. Los datos en las unidades naturales se incluyen en la tabla 15.8. Estime el modelo de regresión lineal múltiple.

Solución: El modelo de regresión ajustado es

$$\hat{y} = b_0 + b_1x_1 + b_2x_2.$$

Tabla 15.8: Datos para el ejemplo 15.2

Tiempo de retención (hr)	Tiempo de flexión (hr)	Producto (%)
0.5	0.10	28
0.8	0.10	39
0.5	0.20	32
0.8	0.20	46

Las unidades de diseño son

$$x_1 = \frac{\text{tiempo de retención} - 0.65}{0.15}, \quad x_2 = \frac{\text{tiempo de flexión} - 0.15}{0.05}$$

y la matriz \mathbf{X} es

$$\begin{bmatrix} 1 & x_1 & x_2 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

con los coeficientes de regresión

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} (1) + a + b + ab \\ \frac{4}{a + ab - (1) - b} \\ \frac{4}{b + ab - (1) - a} \\ 4 \end{bmatrix} = \begin{bmatrix} 36.25 \\ 6.25 \\ 2.75 \end{bmatrix}$$

Así, la ecuación de regresión de mínimos cuadrados es

$$\hat{y} = 36.25 + 6.25x_1 + 2.75x_2.$$

Este ejemplo ilustra el uso del experimento factorial de dos niveles en un ajuste de regresión. Las cuatro corridas experimentales en el diseño 2² se usaron para obtener una ecuación de regresión, con la interpretación evidente de los coeficientes de regresión. El valor $b_1 = 6.25$ representa el incremento estimado en la respuesta (porcentaje de producción) por cambio en la *unidad de diseño* (0.15 horas) en el tiempo de retención. El valor $b_2 = 2.75$ representa una tasa de cambio similar para el tiempo de flexión. ▮

Interacción en el modelo de regresión

Los contrastes de interacción que se estudiaron en la sección 15.2 tienen interpretaciones definidas en el contexto de la regresión. De hecho, las interacciones se explican en los modelos de regresión en términos de producto. Esto se ilustra en el ejemplo 15.2, en donde el modelo con interacción es

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$$

con b_0 , b_1 y b_2 , como antes, y

$$b_{12} = \frac{ab + (1) - a - b}{4} = \frac{46 + 28 - 39 - 32}{4} = 0.75.$$

Así, la ecuación de regresión que expresa dos *efectos principales lineales* e interacción, es

$$\hat{y} = 36.25 + 6.25x_1 + 2.75x_2 + 0.75x_1x_2.$$

El contexto de la regresión proporciona un marco de referencia mediante el cual el lector debería entender mejor la ventaja de la ortogonalidad de que goza el factorial 2^k . En la sección 15.2 se analizaron las ventajas de la ortogonalidad desde el punto de vista del *análisis de varianza* de los datos en un experimento factorial 2^k . Se señaló que la ortogonalidad entre los efectos conduce a la independencia entre las sumas de cuadrados. Desde luego, la presencia de variables de regresión no descarta el uso del análisis de varianza. De hecho, las pruebas *f* se llevan a cabo tal como se describió en la sección 15.2. No obstante, se debe hacer una distinción. En el caso del ANOVA las hipótesis surgen de medias poblacionales, mientras que en el caso de la regresión las hipótesis implican coeficientes de regresión.

Por ejemplo, considere el diseño experimental del ejercicio 15.2 de la página 609. Cada factor es continuo. Suponga que los niveles son

$A(x_1)$:	20%	50%
$B(x_2)$:	5 lb/sec	10 lb/sec
$C(x_3)$:	5	5.5

y que se tiene, para los niveles de diseño,

$$x_1 = \frac{\% \text{ sólidos} - 30}{10}, \quad x_2 = \frac{\text{tasa de flujo} - 7.5}{2.5}, \quad x_3 = \frac{\text{pH} - 5.25}{0.25}.$$

Suponga que es de interés ajustar un modelo de regresión múltiple, en el cual se considerarán todos los coeficientes lineales y las interacciones disponibles. Además, el ingeniero desea obtener información acerca de cuáles niveles del factor *maximizarán* la purificación, es decir, maximizar la respuesta. Este problema es el tema del estudio de caso 15.2.

Estudio de caso 15.2: Experimento de purificación del carbón:¹ La figura 15.9 representa una salida de resultados comentados del análisis de regresión del modelo ajustado

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3,$$

donde x_1 , x_2 y x_3 representan al porcentaje de sólidos, la tasa de flujo y el pH del sistema, respectivamente. Se utilizó la función PROC REG del sistema de cómputo SAS.

Observe los estimados del parámetro, el error estándar y los valores *P* en la salida de resultados por computadora. Los estimados del parámetro representan los coeficientes del modelo. Todos ellos son significativos, excepto el término x_2x_3 (interacción *BC*). También observe que los residuales, los intervalos de confianza y los intervalos de predicción aparecen como se presentaron en el material sobre regresión de los capítulos 11 y 12.

El lector puede usar los valores de los coeficientes del modelo y los valores pronosticados en la salida de resultados por computadora para asegurarse de que la combinación de los factores dé como resultado la **mayor eficiencia de pureza**. El factor *A* (porcentaje de sólidos circulados) tiene un coeficiente positivo alto, lo cual sugiere un valor elevado para el porcentaje de sólidos. Además, se sugiere un valor bajo para el factor *C* (pH del tanque). Aunque el coeficiente del efecto principal *B* (tasa de flujo

¹Véase el ejercicio 15.2.

Dependent Variable: Y									
Analysis of Variance									
Sum of Mean									
Source	DF	Squares	Square	F Value	Pr > F				
Model	7	490.23499	70.03357	254.43	<.0001				
Error	8	2.20205	0.27526						
Corrected Total	15	492.43704							
Root MSE	0.52465	R-Square	0.9955						
Dependent Mean	12.75188	Adj R-Sq	0.9916						
Coeff Var	4.11429								
Parameter Estimates									
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t				
Intercept	1	12.75188	0.13116	97.22	<.0001				
A	1	4.71938	0.13116	35.98	<.0001				
B	1	0.86563	0.13116	6.60	0.0002				
C	1	-1.41563	0.13116	-10.79	<.0001				
AB	1	-0.59938	0.13116	-4.57	0.0018				
AC	1	-0.52813	0.13116	-4.03	0.0038				
BC	1	0.00562	0.13116	0.04	0.9668				
ABC	1	2.23063	0.13116	17.01	<.0001				
Dependent Predicted Std Error									
Obs	Variable	Value	Mean	Predict	9 5% CL Mean	95% CL	Predict	Residual	
1	4.6500	5.2300	0.3710	4.3745	6.0855	3.7483	6.7117	-0.5800	
2	21.4200	21.3850	0.3710	20.5295	22.2405	19.9033	22.8667	0.0350	
3	12.6600	12.6100	0.3710	11.7545	13.4655	11.1283	14.0917	0.0500	
4	18.2700	17.4450	0.3710	16.5895	18.3005	15.9633	18.9267	0.8250	
5	7.9300	7.9050	0.3710	7.0495	8.7605	6.4233	9.3867	0.0250	
6	13.1800	13.0250	0.3710	12.1695	13.8805	11.5433	14.5067	0.1550	
7	6.5100	6.3850	0.3710	5.5295	7.2405	4.9033	7.8667	0.1250	
8	18.2300	18.0300	0.3710	17.1745	18.8855	16.5483	19.5117	0.2000	
9	5.8100	5.2300	0.3710	4.3745	6.0855	3.7483	6.7117	0.5800	
10	21.3500	21.3850	0.3710	20.5295	22.2405	19.9033	22.8667	-0.0350	
11	12.5600	12.6100	0.3710	11.7545	13.4655	11.1283	14.0917	-0.0500	
12	16.6200	17.4450	0.3710	16.5895	18.3005	15.9633	18.9267	-0.8250	
13	7.8800	7.9050	0.3710	7.0495	8.7605	6.4233	9.3867	-0.0250	
14	12.8700	13.0250	0.3710	12.1695	13.8805	11.5433	14.5067	-0.1550	
15	6.2600	6.3850	0.3710	5.5295	7.2405	4.9033	7.8667	-0.1250	
16	17.8300	18.0300	0.3710	17.1745	18.8855	16.5483	19.5117	-0.2000	

Figura 15.9: Lista de resultados del SAS para los datos del estudio de caso 15.2.

del polímero) es positivo, el coeficiente positivo elevado de $x_1 x_2 x_3$ (ABC) sugiere que la tasa de flujo debería estar en el nivel bajo para aumentar la eficiencia. De hecho, el modelo de regresión generado en la salida de resultados por computadora del SAS sugiere que la combinación de factores que podrían producir resultados óptimos, o quizá sugerir experimentos futuros, es dada por

A: nivel alto

B: nivel bajo

C: nivel bajo

15.5 El diseño ortogonal

En situaciones experimentales en las que es apropiado ajustar modelos que son lineales en las variables de diseño y que posiblemente impliquen interacciones o términos de producto, el *diseño ortogonal* de dos niveles, o arreglo ortogonal, plantea algunas ventajas. Por diseño ortogonal nos referimos a uno en el que hay ortogonalidad entre las columnas de la matriz X . Considere la matriz X para el factorial 2^2 del ejemplo 15.2. Observe que las tres columnas son mutuamente ortogonales. La matriz X del factorial 2^3 también contiene columnas ortogonales. El factorial 2^3 con interacciones produciría una matriz X del tipo

$$X = \begin{bmatrix} & x_1 & x_2 & x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 & x_1 x_2 x_3 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

La descripción de los grados de libertad es

Fuente	g.l.	
Regresión	3	
Falta de ajuste	4	$(x_1 x_2, x_1 x_3, x_2 x_3, x_1 x_2 x_3)$
Error (puro)	8	
Total	15	

Los ocho grados de libertad para el error puro se obtienen a partir de las *corridas duplicadas* en cada punto del diseño. Los grados de libertad de la falta de ajuste podrían considerarse como la diferencia entre el número de puntos de diseño distintos y el número total de términos en el modelo; en este caso hay ocho puntos y cuatro términos en el modelo.

Error estándar de los coeficientes y pruebas T

En las secciones anteriores vimos cómo el diseñador de un experimento puede aprovechar el concepto de ortogonalidad para diseñar un experimento de regresión con coeficientes que obtienen una varianza mínima sobre la base del costo. Debemos ser capaces de utilizar el material sobre la regresión que se expuso en la sección 12.4 para calcular estimados de las varianzas de los coeficientes y, con ello, los errores estándar. También resulta de interés observar la relación entre el estadístico t de un coeficiente y el estadístico F descrito e ilustrado en capítulos anteriores.

En la sección 12.4 vimos que las varianzas y las covarianzas de los coeficientes aparecen en A^{-1} , o, en términos de la notación actual, la *matriz de varianza-covarianza* de coeficientes es

$$\sigma^2 A^{-1} = \sigma^2 (X'X)^{-1}.$$

En el caso del experimento factorial 2^k las columnas de X son mutuamente ortogonales.

lo que impone una estructura muy especial. En general, para 2^k se puede escribir

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_k & x_1 x_2 & \cdots \\ \pm 1 & \pm 1 & \pm 1 & \cdots & \pm 1 & \pm 1 & \cdots \end{bmatrix},$$

donde cada columna contiene 2^k o $2^k n$ entradas, donde n es el número de réplicas de las corridas en cada punto del diseño. Así, la formación de $\mathbf{X}'\mathbf{X}$ lleva a

$$\mathbf{X}'\mathbf{X} = 2^k n \mathbf{I}_p,$$

donde \mathbf{I} es la matriz de identidad de la dimensión p , el número de parámetros del modelo.

Ejemplo 15.3: Considere un diseño factorial 2^3 con corridas por duplicado que se ajusta al modelo

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3.$$

Proporcione expresiones para los errores estándar de los estimados de mínimos cuadrados de $b_0, b_1, b_2, b_3, b_{12}, b_{13}$ y b_{23} .

Solución:

$$\mathbf{X} = \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 \end{matrix} \\ \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

considerando cada unidad como *repetida*, es decir, considerando que cada observación está duplicada. Como resultado,

$$\mathbf{X}'\mathbf{X} = 16\mathbf{I}_7.$$

Por consiguiente,

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{16} \mathbf{I}_7.$$

A partir de lo anterior debe quedar claro que las varianzas de todos los coeficientes para un factorial 2^k con n corridas en cada punto de diseño son

$$\text{Var}(b_j) = \frac{\sigma^2}{2^k n},$$

y, desde luego, todas las covarianzas son iguales a cero. Como resultado, los errores estándar de los coeficientes se calculan como

$$s_{b_j} = s \sqrt{\frac{1}{2^k n}},$$

donde s se calcula por medio de la raíz cuadrada del cuadrado medio del error que se espera obtener a partir de una réplica adecuada. Así, en nuestro caso con 2^3 ,

$$s_{b_j} = s \left(\frac{1}{4} \right).$$

Ejemplo 15.4: Considere el experimento metalúrgico del ejercicio 15.3 de la página 609. Suponga que el modelo ajustado es

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4.$$

¿Cuáles son los errores estándar de los coeficientes de regresión de los mínimos cuadrados?

Solución: Los errores estándar de todos los coeficientes para el factorial 2^4 son iguales, y son

$$s_{b_j} = s \sqrt{\frac{1}{2^k n}},$$

que en este ejemplo es

$$s_{b_j} = s \sqrt{\frac{1}{(16)(2)}}.$$

En este caso el cuadrado medio del error puro es dado por $s^2 = 2.46$ (16 grados de libertad). Entonces,

$$s_{b_j} = 0.28.$$

Los errores estándar de los coeficientes se usan para construir estadísticos t de todos los coeficientes. Estos valores t se relacionan con los estadísticos F del análisis de varianza. Ya se demostró que un estadístico F sobre un coeficiente, usando el factorial 2^k , es

$$f = \frac{(\text{contraste})^2}{(2^k n) s^2}.$$

Ésta es la forma del estadístico F de la página 610 para el experimento metalúrgico (ejercicio 15.3). Es fácil comprobar que si se escribe

$$t = \frac{b_j}{s_{b_j}}, \quad \text{donde} \quad b_j = \frac{\text{contraste}}{2^k n},$$

entonces

$$t^2 = \frac{(\text{contraste})^2}{s^2 2^k n} = f. \quad \blacksquare$$

Como resultado, se mantiene la relación acostumbrada entre los estadísticos t sobre los coeficientes y los valores F . Como era de esperarse, la única diferencia entre utilizar t y F para evaluar la significancia radica en el hecho de que el estadístico t indica el signo o la dirección del efecto del coeficiente.

Parecería que el plan del factorial 2^k se adapta a muchas situaciones prácticas a las cuales se ajustan modelos de regresión. Puede incluir términos lineales y de interacción, lo que proporciona estimados óptimos de todos los coeficientes (desde un punto de vista de la varianza). Sin embargo, cuando k es grande, el número de puntos del diseño requiriendo es muy grande. A menudo es posible utilizar partes del diseño total y aun así conservar la ortogonalidad, con todas sus ventajas. En la sección 15.6 se estudian esos diseños.

Una mirada más cercana a la propiedad de ortogonalidad del factorial 2^k

Ya vimos que para el caso del factorial 2^k toda la información que obtiene el analista sobre los efectos y las interacciones principales aparece en forma de contrastes. Estas " $2^k - 1$ piezas de información" conllevan un solo grado de libertad cada una y son independientes entre sí. En un análisis de varianza se manifiestan como *efectos*; mientras que si se construye un modelo de regresión, los efectos que resultan son coeficientes de regresión, aparte de un factor de 2. Con cada forma de análisis es posible hacer pruebas de significancia y la prueba t para un efecto dado es la misma en términos numéricos que para el coeficiente de regresión correspondiente. En el caso del ANOVA son importantes la selección de las variables y la interpretación científica de las interacciones; en tanto que en el caso de un análisis de regresión se usa un modelo para predecir la respuesta y/o determinar cuáles combinaciones de factores o niveles son las óptimas, por ejemplo, maximizar la producción o la eficiencia de la purificación, como en el estudio de caso 15.2.

Resulta que la propiedad de ortogonalidad es importante, ya sea que se trate de un ANOVA o de una regresión. La ortogonalidad entre las columnas de X , la matriz del modelo en, digamos, el ejemplo 15.3, ofrece condiciones especiales que tienen un impacto importante sobre los **efectos de la varianza** o los **coeficientes de regresión**. De hecho, ya es evidente que el diseño ortogonal da como resultado la igualdad de varianza para todos los efectos o coeficientes. Es así como, para propósitos de estimación o de prueba, la precisión es la misma para todos los coeficientes, los efectos principales o las interacciones. Además, si el modelo de regresión sólo contiene términos lineales, por lo cual sólo los efectos principales son de interés, las condiciones siguientes dan como resultado la minimización de las varianzas de todos los efectos, o, en forma correspondiente, de los coeficientes de regresión de primer orden.

Condiciones para varianzas mínimas de los coeficientes	Si el modelo de regresión contiene términos no mayores de primer orden, y si los rangos de las variables son dados por $x_j \in [-1, +1]$ para $j = 1, 2, \dots, k$, entonces $\text{Var}(b_j)/\sigma^2$, para $j = 1, 2, \dots, k$, se minimiza si el diseño es ortogonal y todos los niveles x_i del diseño son ± 1 para $i = 1, 2, \dots, k$.
--------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Así, en términos de los coeficientes del modelo o los efectos principales, la ortogonalidad en el 2^k es una propiedad muy deseable.

Otro método para lograr una mejor comprensión del "balance" proporcionado por el factorial 2^3 consiste en observar la situación mediante una gráfica. En la figura 15.10 se aprecia cada uno de los contrastes ortogonales y, por lo tanto, mutuamente independientes. En las gráficas se comparan los planos de los cuadrados cuyos vértices contienen las respuestas etiquetadas con "+" con las que tienen el signo "-". Las que aparecen en el inciso *a* presentan contrastes para efectos principales y deberían ser evidentes para el lector. Las del inciso *b* presentan los planos determinados por los vértices "+" y "-" para los tres contrastes de interacción de dos factores. En el inciso *c* se aprecia la representación geométrica de los contrastes para la interacción de tres factores (*ABC*).

Corridas centrales con diseños factoriales 2^k

En la situación en que se aplica el diseño 2^k con variables **continuas** de diseño y se busca ajustar un modelo de regresión lineal, el uso de réplicas de corridas en el **diseño central** puede ser sumamente útil. De hecho, además de las ventajas que se analizarán a continuación, la mayoría de los científicos e ingenieros considerarían que las corridas

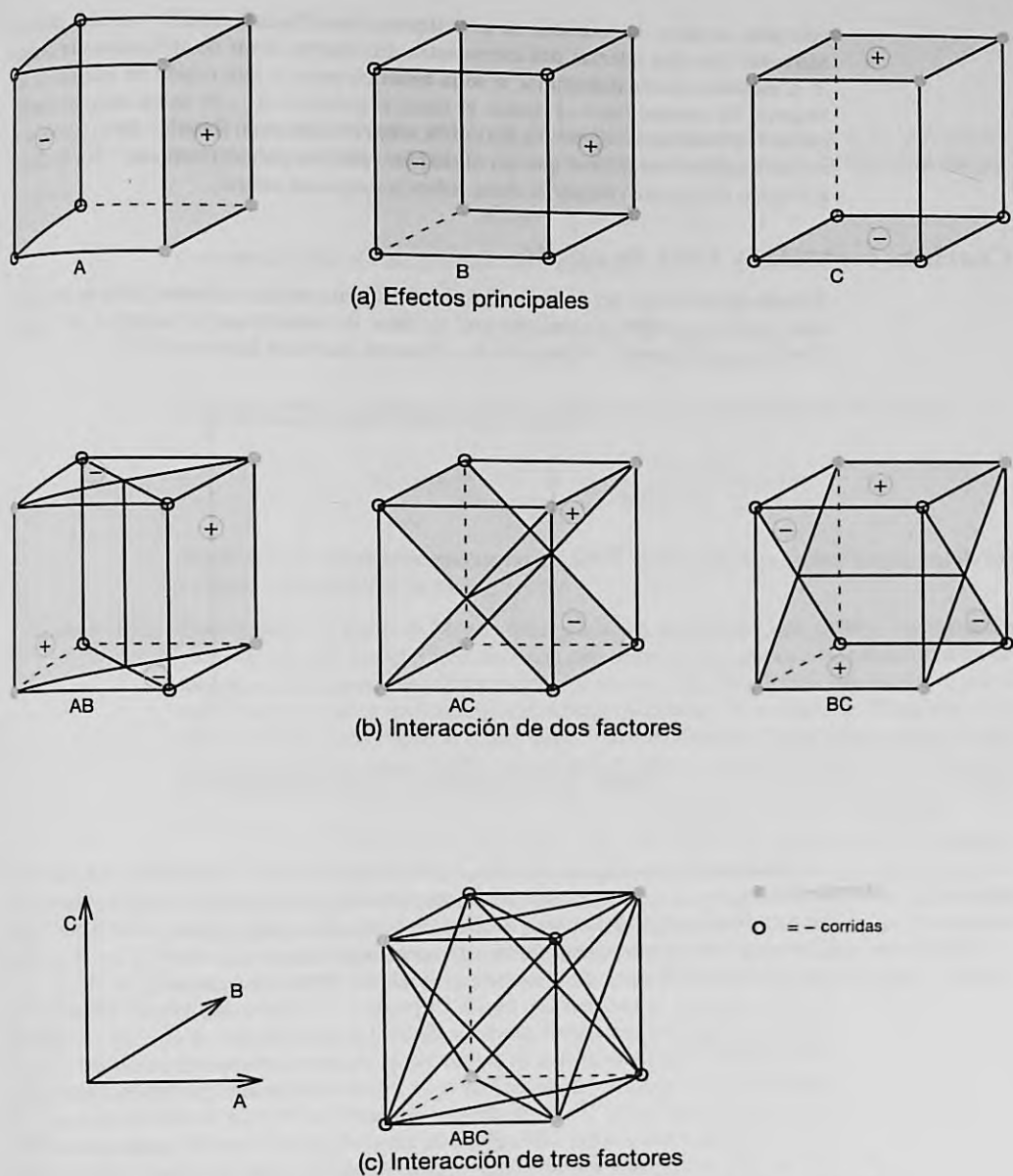


Figura 15.10: Presentación geométrica de los contrastes para el diseño factorial 2^3 .

centrales, es decir, las corridas en $x_i = 0$ para $i = 1, 2, \dots, k$, no sólo son una práctica razonable sino que además son interesantes. En muchas áreas de aplicación del diseño 2^k el científico desea determinar si sería benéfico pasar a otra región de interés en los factores. En muchos casos el centro, es decir, el punto $(0, 0, \dots, 0)$ en los factores codificados, con frecuencia representa las condiciones de operación actuales del proceso, o al menos aquellas condiciones que se consideran "óptimas para el momento". Por lo tanto, a menudo el científico requerirá datos sobre la respuesta central.

Corridas centrales y falta de ajuste

Además del atractivo del aumento del diseño 2^k con corridas centrales, otra de sus ventajas consiste en que se relaciona con la clase de modelo que se ajusta a los datos. Considere, por ejemplo, el caso con $k = 2$ que se ilustra en la figura 15.11.

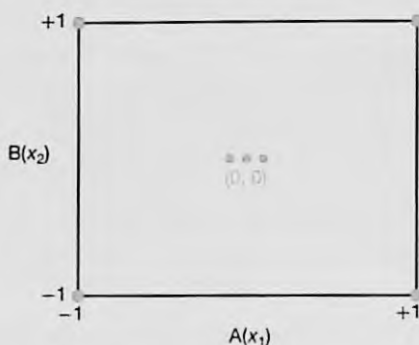


Figura 15.11: Un diseño 2^2 con corridas centrales.

Queda claro que *sin las corridas centrales* los términos del modelo son la intersección, x_1, x_2, x_1x_2 . Esto explica los cuatro grados de libertad del modelo producidos por los cuatro puntos del diseño, además de cualquier réplica. Como cada factor tiene información de respuesta disponible *sólo en dos ubicaciones* $\{-1, +1\}$, no es posible incluir términos "puros" de curvatura de segundo orden en el modelo, es decir, x_1^2 o x_2^2 . Sin embargo, la información en $(0, 0)$ produce un grado de libertad adicional del modelo. Si bien este importante grado de libertad no permite que ni x_1^2 ni x_2^2 se empleen en el modelo, sí permite probar la significancia de una combinación lineal de x_1^2 y x_2^2 . Entonces, para n_c corridas centrales, hay $n_c - 1$ grados de libertad disponibles para réplicas o para el error "puro". Esto permite un estimado de σ^2 para probar los términos del modelo y la significancia del único grado de libertad para la **falta de ajuste cuadrático**. El concepto aquí es muy similar al que se describió en el material sobre la falta de ajuste del capítulo 11.

Para entender por completo cómo funciona la prueba de falta de ajuste suponga que para $k = 2$ el **modelo verdadero** contiene todo el complemento de segundo orden de los términos, incluyendo a x_1^2 y x_2^2 . En otras palabras,

$$E(Y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2.$$

Ahora, considere el contraste

$$\bar{y}_f = \bar{y}_0$$

donde \bar{y}_f es la respuesta promedio de las ubicaciones factoriales y \bar{y}_0 es la respuesta promedio en el punto central. Es fácil demostrar (véase el ejercicio de repaso 15.46) que

$$E(\bar{y}_f - \bar{y}_0) = \beta_{11} + \beta_{22},$$

y, en efecto, para el caso general con k factores,

$$E(\bar{y}_f - \bar{y}_0) = \sum_{i=1}^k \beta_{ii}.$$

Como resultado, la prueba de falta de ajuste es una prueba t simple (o $F = t^2$) con

$$t_{n_c-1} = \frac{\bar{y}_f - \bar{y}_0}{s_{\bar{y}_f - \bar{y}_0}} = \frac{\bar{y}_f - \bar{y}_0}{\sqrt{CME(1/n_f + 1/n_c)}},$$

donde n_f es el número de puntos factoriales y CME sólo es la varianza muestral de los valores de la respuesta en $(0, 0, \dots, 0)$.

Ejemplo 15.5: Este ejemplo se tomó de Myers, Montgomery y Anderson-Cook (2009). Un ingeniero químico trata de modelar la conversión porcentual en un proceso. Hay dos variables de interés, el tiempo de reacción y la temperatura de reacción. En un intento por llegar al modelo apropiado se realiza un experimento preliminar en un factorial 2^2 usando la región actual de interés en el tiempo y temperatura de reacción. Se hicieron corridas únicas en cada uno de los cuatro puntos factoriales, y cinco corridas en el centro del diseño con el fin de poder realizar una prueba de falta de ajuste para la curvatura. En la figura 15.12 se presenta la región del diseño y las corridas experimentales sobre el producto.

Las lecturas del tiempo y la temperatura en el centro son, desde luego, 35 minutos y 145°C . Los estimados de los efectos principales y el coeficiente de interacción único se calculan mediante contrastes, igual que antes. Las corridas en el centro **no intervienen en el cálculo de b_1 , b_2 y b_{12}** . Esto debería ser razonable para el lector. La intersección es sólo \bar{y} para todo el experimento. Este valor es $\bar{y} = 40.4444$. Los errores estándar se calculan usando los elementos de la diagonal de $(\mathbf{X}'\mathbf{X})^{-1}$, como ya se expuso. Para este caso,

$$\mathbf{X} = \begin{bmatrix} & x_1 & x_2 & x_1 x_2 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

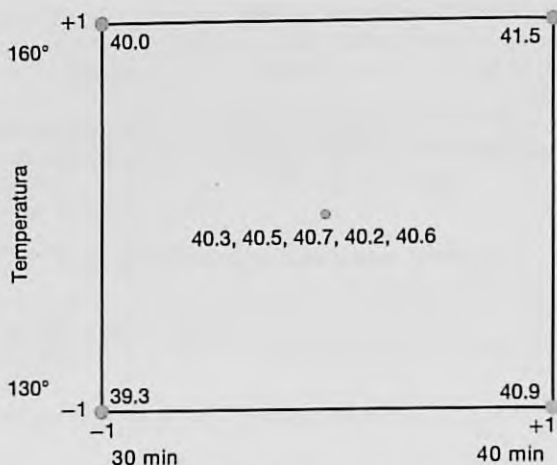


Figura 15.12: Factorial 2^2 con 5 corridas en el centro.

Después de hacer los cálculos se obtiene

$$\begin{aligned}
 b_0 &= 40.4444, & b_1 &= 0.7750, & b_2 &= 0.3250, & b_{12} &= -0.0250, \\
 s_{b_0} &= 0.06231, & s_{b_1} &= 0.09347, & s_{b_2} &= 0.09347, & s_{b_{12}} &= 0.09347, \\
 t_{b_0} &= 649.07, & t_{b_1} &= 8.29, & t_{b_2} &= 3.48, & t_{b_{12}} &= -0.27 \quad (P = 0.800).
 \end{aligned}$$

El contraste $\bar{y}_f - \bar{y}_0 = 40.425 - 40.46 = -0.035$ y el estadístico t que prueba la curvatura son dados por

$$t = \frac{40.425 - 40.46}{\sqrt{0.0430(1/4 + 1/5)}} = 0.251 \quad (P = 0.814).$$

Como resultado, parece que el modelo apropiado debería contener sólo términos de primer orden (además de la intersección). ▮

Una mirada intuitiva a la prueba de curvatura

Si se considera el caso sencillo con una sola variable de diseño con corridas en -1 y $+1$ debe quedar claro que la respuesta promedio en -1 y $+1$ debe estar cerca de la respuesta en 0 , el centro, si el modelo es de primer orden. Cualquier desviación sugeriría, con seguridad, curvatura. Esto se puede extender fácilmente a dos variables. Considere la figura 15.13.

La figura muestra el plano sobre y que pasa a través de los valores de y de los puntos factoriales. Éste es el plano que representaría el ajuste perfecto para el modelo que contiene x_1 , x_2 y x_1x_2 . Si el modelo no contiene curvatura cuadrática, es decir, $\beta_{11} = \beta_{22} = 0$, se esperaría que la respuesta en $(0, 0)$ esté en el plano o cerca del mismo. Si la respuesta estuviera lejos del plano, como ocurre en la figura 15.13, entonces se podría ver en forma gráfica que la curvatura cuadrática está presente.

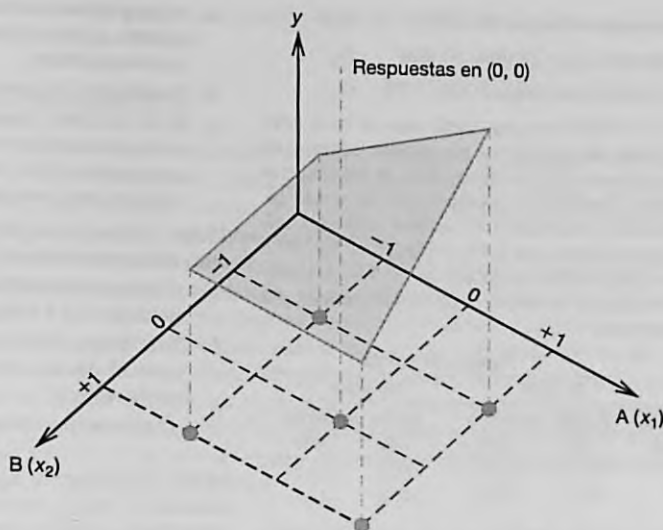


Figura 15.13: Factorial 2^2 con corridas en $(0, 0)$.

Ejercicios

15.13 Considere un experimento 2^5 donde se realizan corridas experimentales sobre 4 máquinas diferentes. Use las máquinas como bloques y suponga que todos los efectos principales y las interacciones de dos factores son importantes.

- ¿Cuáles corridas se harían sobre cada una de las 4 máquinas?
- ¿Cuáles efectos se confunden con los bloques?

15.14 En un experimento descrito en Myers, Montgomery y Anderson-Cook (2009) se buscan las condiciones óptimas para almacenar semen de bovinos con el fin de obtener la supervivencia máxima. Las variables son el porcentaje de citrato de sodio, el porcentaje de glicerol y el tiempo de equilibrio en horas. La respuesta es el porcentaje de supervivencia de los espermatozoides móviles. Los niveles naturales se encuentran en la referencia mencionada. A continuación se presentan los datos con los niveles codificados para la parte factorial del diseño y las corridas centrales.

- Ajuste un modelo de regresión lineal con los datos y determine cuáles términos lineales y de interacción son significativos. Suponga que la interacción $x_1x_2x_3$ es despreciable.
- Pruebe la falta de ajuste cuadrático y comente la respuesta.

x_1 , % de citrato de sodio	x_2 , % de glicerol	x_3 Tiempo de equilibrio	% Supervivencia
-1	-1	-1	57
1	-1	-1	40
-1	1	1	19
1	1	1	40
-1	-1	-1	54
1	-1	-1	41
-1	1	1	21
1	1	1	43
0	0	0	63
0	0	0	61

15.15 Los productores de petróleo están interesados en aleaciones de níquel que sean fuertes y resistentes a la corrosión. Se realizó un experimento en el que se comparó del límite elástico especímenes elásticos de aleaciones de níquel cargados en una solución de ácido sulfúrico saturada con disulfuro de carbón. Se compararon dos aleaciones; una con 75% de níquel y otra con 30% de níquel. Se probaron las aleaciones en dos tiempos de carga diferentes, de 25 y 50 días. Se realizó un factorial 2^3 con los factores siguientes:

% de ácido sulfúrico: 4%, 6% (x_1)
 tiempo de carga: 25 días, 50 días (x_2)
 composición del níquel: 30%, 75% (x_3)

Se preparó un espécimen para cada una de las 8 condiciones. Como los ingenieros no estaban seguros de la naturaleza del modelo, es decir, de si se necesitarían o no términos cuadráticos, incorporaron un tercer nivel (intermedio) y realizaron 4 corridas centrales utilizando 4 especímenes con ácido sulfúrico al 5%, 37.5 días y una composición de níquel de 52.5%. A continuación se incluyen las resistencias en kilogramos por pulgada cuadrada.

Composición del níquel	Tiempo de carga			
	25 días		50 días	
	Ácido sulfúrico 4%	Ácido sulfúrico 6%	Ácido sulfúrico 4%	Ácido sulfúrico 6%
75%	52.5	56.5	47.9	47.2
30%	50.2	50.8	47.4	41.7

Las corridas centrales produjeron las siguientes resistencias:

51.6, 51.4, 52.4, 52.9

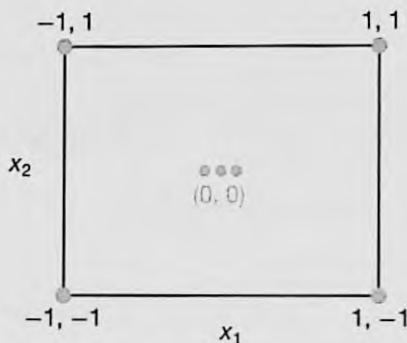


Figura 15.14: Gráfica para el ejercicio 15.17.

- Haga pruebas para determinar cuáles efectos principales e interacciones deberían incluirse en el modelo ajustado.
- Pruebe para la curvatura cuadrática.
- Si la curvatura cuadrática es significativa, ¿cuántos puntos de diseño adicionales se necesitan para determinar cuáles términos cuadráticos deberían incluirse en el modelo?

15.16 Suponga que es posible llevar a cabo una réplica del experimento del ejercicio 15.13.

- ¿Una segunda réplica del esquema de bloques del ejercicio 15.13 sería la mejor opción?
- Si la respuesta del inciso a es negativa, proporcione el diseño de una mejor opción para la segunda réplica.
- ¿Qué concepto utilizó en la elección del diseño?

15.17 Considere la figura 15.14, que representa un factorial 2^2 con 3 corridas centrales. Si la curvatura cuadrática es significativa, ¿cuáles otros puntos de diseño seleccionarían, que permitieran estimar los términos x_1^2 y x_2^2 ? Explique su respuesta.

15.6 Experimentos factoriales fraccionarios

El experimento factorial 2^k se puede volver muy demandante, en términos del número de unidades experimentales requeridas, cuando el valor de k es grande. Una de las ventajas reales de este plan experimental es que permite un grado de libertad para cada interacción. Sin embargo, en muchas situaciones experimentales se sabe que ciertas interacciones son despreciables, por lo que sería un desperdicio de esfuerzo experimental utilizar el experimento factorial completo. De hecho, el experimentador podría tener limitaciones económicas que le impidan hacer observaciones de todas las combinaciones 2^k de tratamientos. Cuando k es grande, a menudo se puede usar un **experimento factorial**

fraccionario donde quizás sea posible llevar a cabo la mitad, un cuarto o incluso un octavo del plan factorial total.

Construcción de la fracción de $\frac{1}{2}$

La construcción del diseño de media réplica es idéntica a la asignación del experimento factorial 2^k en dos bloques. Se comienza por seleccionar un contraste de definición que se sacrificará por completo. Luego se construyen los dos bloques en concordancia y se elige cualquiera de ellos como plan experimental.

A menudo la fracción de $\frac{1}{2}$ de un factorial 2^k se conoce como diseño 2^{k-1} , el cual indica el número de puntos de diseño. El primer ejemplo de un diseño 2^{k-1} será uno de $\frac{1}{2}$ o uno de 2^3 o uno de 2^{3-1} . En otras palabras, el científico o el ingeniero no puede usar el complemento completo, es decir, todo el diseño 2^3 con 8 puntos de diseño, por lo que debe apelar a un diseño con sólo cuatro puntos de diseño. La pregunta es la siguiente: de los puntos de diseño (1), a , b , ab , ac , c , bc y abc , ¿cuáles son los cuatro puntos de diseño que producirán el diseño más útil? La respuesta, junto con los conceptos importantes relacionados, aparece en la tabla de signos + y - que muestra los contrastes para el diseño 2^3 completo. Considere la tabla 15.9.

Tabla 15.9: Contrastes para los siete efectos disponibles en el caso de un experimento factorial 2^3

	Combinación de tratamientos	Efectos							
		I	A	B	C	AB	AC	BC	ABC
2^{3-1}	a	+	+	-	-	-	-	+	+
	b	+	-	+	-	-	+	-	+
	c	+	-	-	+	+	-	-	+
	abc	+	+	+	+	+	+	+	+
2^{3-1}	ab	+	+	+	-	+	-	-	-
	ac	+	+	-	+	-	+	-	-
	bc	+	-	+	+	-	-	+	-
	(1)	+	-	-	-	+	+	+	-

Observe que las dos fracciones $\frac{1}{2}$ son $\{a, b, c, abc\}$ y $\{ab, ac, bc, (1)\}$. Observe también en la tabla 15.9 que en ambos diseños ABC no tiene contraste, pero todos los demás efectos sí lo tienen. En una de las fracciones se tiene que ABC contiene todos los signos + y en la otra fracción el efecto ABC contiene todos los signos -. Como resultado, se dice que el diseño de la parte superior de la tabla es descrito por $ABC = I$, y el de la parte inferior por $ABC = -I$. La interacción ABC se denomina **generador del diseño**, y $ABC = I$ (o $ABC = -I$ para el segundo diseño) recibe el nombre de **relación definitoria**.

Alias en el 2^{3-1}

Si nos centramos en el diseño $ABC = I$ (el 2^{3-1} superior), es evidente que seis efectos contienen contrastes. Esto produce la apariencia inicial de que todos los *efectos* se pueden estudiar por separado de ABC . Sin embargo, el lector recordará que con sólo cuatro puntos de diseño, incluso si se replican, los grados de libertad disponibles (además del error experimental) son

Términos del modelo de regresión	3
Intersección	$\frac{1}{4}$

Un análisis más detallado sugiere que los siete efectos no son ortogonales y que cada contraste está representado en otro efecto. De hecho, si se emplea el símbolo \equiv para denotar **contrastes idénticos**, se tiene que

$$A \equiv BC; \quad B \equiv AC; \quad C \equiv AB.$$

Como resultado, dentro de un par no es posible estimar un efecto independiente de su "socio" alias. Los efectos

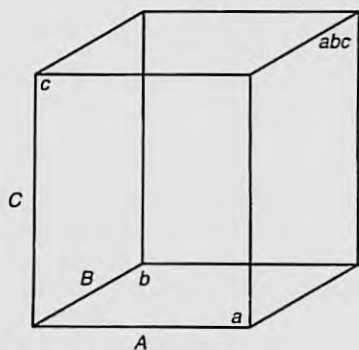
$$A = \frac{a + abc - b - c}{2} \quad \text{y} \quad BC = \frac{a + abc - b - c}{2}$$

producirán el mismo resultado numérico, de manera que contienen la misma información. De hecho, con frecuencia se dice que **comparten un grado de libertad**. En realidad, el efecto estimado verdaderamente estima la suma, es decir, $A + BC$. Se dice que A y BC son alias, al igual que B y AC , y que C y AB .

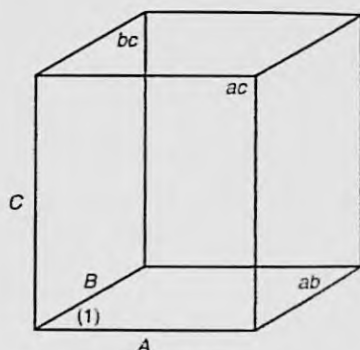
Para la fracción $ABC = -I$ se observa que los alias son los mismos que para la fracción $ABC = I$, además del signo. Así, se tiene

$$A \equiv -BC; \quad B \equiv -AC; \quad C \equiv -AB.$$

Las dos fracciones aparecen en las esquinas de los cubos de las figuras 15.15a y 15.15b.



(a) La fracción $ABC = I$



(a) La fracción $ABC = -I$

Figura 15.15: Las fracciones $\frac{1}{2}$ del factorial 2^3 .

Cómo se determinan los alias en general

En general, para un diseño 2^{k-1} , cada efecto, además de aquel definido por el generador, tendrá un *solo socio alias*. El efecto definido por el generador no tendrá alias en otro

efecto, sino que su alias será la media, ya que el estimador de mínimos cuadrados será la media. Para determinar el alias de cada efecto, sólo se comienza con la relación definitoria, digamos $ABC = I$, para el diseño 2^{3-1} . Entonces, para obtener, digamos, el alias para el efecto A , se multiplica A por ambos lados de la ecuación $ABC = I$ y se reduce cualquier exponente por el módulo 2. Por ejemplo,

$$A \cdot ABC = A, \quad \text{con lo que } BC \equiv A.$$

En forma similar,

$$B \equiv B \cdot ABC \equiv AB^2C \equiv AC,$$

y, por supuesto,

$$C \equiv C \cdot ABC \equiv ABC^2 \equiv AB.$$

Ahora, para la segunda fracción, es decir, la definida por la relación $ABC = -I$,

$$A \equiv -BC; \quad B \equiv -AC; \quad C \equiv -AB.$$

Como resultado, el valor numérico del efecto A en realidad estima $A - BC$. De manera similar, el valor de B estima $B - AC$, y el valor de C estima $C - AB$.

Construcción formal del diseño 2^{k-1}

La comprensión plena del concepto de los alias facilita el conocimiento de la construcción del diseño 2^{k-1} . Se comienza con la investigación del 2^{3-1} . Se requieren tres factores y cuatro puntos de diseño. El procedimiento comienza con un **factorial completo** en $k - 1 = 2$ factores A y B . Después se agrega un tercer factor de acuerdo con las estructuras de alias deseadas. Por ejemplo, con ABC como el generador, resulta claro que $C = \pm AB$. Así, se descubre que $C = AB$, o $C = -AB$ complementan el factorial completo en A y B . La tabla 15.10 ilustra un procedimiento que resulta muy sencillo.

Tabla 15.10: Construcción de los dos diseños 2^{3-1}

2 ² básico		2 ³⁻¹ ; ABC = I			2 ³⁻¹ ; ABC = -I		
A	B	A	B	C = AB	A	B	C = -AB
-	-	-	-	+	-	-	-
+	-	+	-	-	+	-	+
-	+	-	+	-	-	+	+
+	+	+	+	+	+	+	-

Note que ya vimos que $ABC = I$ proporciona los puntos de diseño a, b, c y abc ; en tanto que $ABC = -I$ proporciona $(1), ac, bc$ y ab . Anteriormente pudimos construir los mismos diseños usando los contrastes que se muestran en la tabla 15.9. Sin embargo, a medida que el diseño se vuelve más complicado con fracciones superiores, esas tablas de contrastes se vuelven más difíciles de trabajar.

Ahora considere un diseño 2^{4-1} , es decir, $\frac{1}{2}$ de un diseño factorial 2^4 , que incluye los factores A, B, C y D . Como en el caso del diseño 2^{3-1} , la interacción que se usa como

generador es la interacción de mayor orden, en este caso $ABCD$. Debe recordarse que $ABCD = I$, la relación definitoria sugiere que se sacrifica la información sobre $ABCD$. Aquí comenzamos con el diseño 2^3 completo en A , B y C , y se forma $D = \pm ABC$ para generar los dos diseños 2^{4-1} . La tabla 15.11 ilustra la construcción de ambos diseños.

Tabla 15.11: Construcción de los dos diseños 2^{4-1}

2^3 Básico			$2^{4-1}; ABCD = I$				$2^{4-1}; ABCD = -I$			
A	B	C	A	B	C	$D = ABC$	A	B	C	$D = -ABC$
-	-	-	-	-	-	-	-	-	-	+
+	-	-	+	-	-	+	+	-	-	-
-	+	-	-	+	-	+	-	+	-	-
+	+	-	+	+	-	-	+	+	-	+
-	-	+	-	-	+	+	-	-	+	-
+	-	+	+	-	+	-	+	-	+	+
-	+	+	-	+	+	-	-	+	+	+
+	+	+	+	+	+	+	+	+	+	-

Aquí, empleando las notaciones a, b, c , etcétera, se tienen los diseños siguientes:

$$ABCD = I, (1), ad, bd, ab, cd, ac, bc, abcd$$

$$ABCD = -I, d, a, b, abd, c, acd, bcd, abc.$$

En el caso de 2^{4-1} , los alias se obtienen como se describió antes para 2^{3-1} . Cada efecto tiene un solo socio alias que se obtiene mediante la multiplicación que se efectúa utilizando la relación definitoria. Por ejemplo, el alias de A para el diseño $ABCD = I$ es dado por

$$A = A \cdot ABCD = A^2BCD = BCD.$$

El alias para AB es dado por

$$AB = AB \cdot ABCD = A^2B^2CD = CD.$$

Como es fácil observar, los efectos principales tienen alias con interacciones de tres factores y las interacciones de dos factores tienen alias con otras interacciones de dos factores. La lista completa es dada por

$$A = BCD \quad AB = CD$$

$$B = ACD \quad AC = BD$$

$$C = ABD \quad AD = BC$$

$$D = ABC.$$

Construcción de la fracción de $\frac{1}{4}$

En el caso de la fracción de $\frac{1}{4}$, en vez de una se seleccionan dos interacciones para ser sacrificadas, y la tercera resulta al obtener la interacción generalizada de las dos

seleccionadas. Observe que esto se asemeja mucho a la construcción de cuatro bloques. La fracción que se emplea es simplemente uno de los bloques. Un ejemplo sencillo ayuda mucho a ver la conexión con la construcción de la fracción de $\frac{1}{2}$. Considere la construcción de $\frac{1}{4}$ de un factorial 2^5 , es decir, un diseño 2^{5-2} con los factores A, B, C, D y E . Un procedimiento que evita el confundir dos efectos principales es la selección de ABD y ACE como las interacciones que corresponden a los dos generadores, lo que produce $ABD = I$ y $ACE = I$ como las relaciones definitorias. La tercera interacción sacrificada sería $(ABD)(ACE) = A^2BCDE = BCDE$. Para la construcción del diseño se comienza con un factorial $2^{5-2} = 2^3$ en A, B y C . Se usan las interacciones ABD y ACE para proporcionar los generadores, de manera que el factorial 2^3 en A, B y C es proporcionado por el factor $D = \pm AB$ y $E = \pm AC$. Así, una de las fracciones es dada por

A	B	C	$D = AB$	$E = AC$	
-	-	-	+	+	<i>de</i>
+	-	-	-	-	<i>a</i>
-	+	-	-	+	<i>be</i>
+	+	-	+	-	<i>abd</i>
-	-	+	+	-	<i>cd</i>
+	-	+	-	+	<i>ace</i>
-	+	+	-	-	<i>bc</i>
+	+	+	+	+	<i>abcde</i>

Las otras tres fracciones se calculan utilizando los generadores $\{D = -AB, E = AC\}$, $\{D = AB, E = -AC\}$ y $\{D = -AB, E = -AC\}$. Considere un análisis del diseño 2^{5-2} anterior, que contiene ocho puntos de diseño para estudiar cinco factores. Los alias para los efectos principales son dados por

$A(ABD) \equiv BD$	$A(ACE) \equiv CE$	$A(BCDE) \equiv ABCDE$
$B \equiv AD$	$\equiv ABCE$	$\equiv CDE$
$C \equiv ABCD$	$\equiv AE$	$\equiv BDE$
$D \equiv AB$	$\equiv ACDE$	$\equiv BCE$
$E \equiv ABDE$	$\equiv AC$	$\equiv BCD$

Los alias para otros efectos se pueden obtener de la misma manera. El desglose de los grados de libertad es dado por (además de la réplica)

Efectos principales	5	
Falta de ajuste	$\frac{2}{7}$	$(CD = BE, BC = DE)$
Total	7	

Se listan las interacciones sólo para el grado dos en la falta de ajuste.

Ahora considere el caso de un diseño 2^{6-2} , que permite 16 puntos de diseño para estudiar seis factores. Nuevamente se eligen dos generadores de diseño. Una opción pragmática para complementar un factorial $2^{6-2} = 2^4$ completo en A, B, C y D consiste en usar $E = \pm ABC$ y $F = \pm BCD$. La construcción se muestra en la tabla 15.12.

Es evidente que con ocho puntos de diseño más que en 2^{5-2} los alias de los efectos principales no representarán un problema difícil. De hecho, observe que con las relaciones definitorias $ABCE = \pm I$, $BCDF = \pm I$, y $(ABCE)(BCDF) = ADEF = \pm I$, los

Tabla 15.12: Diseño 2^{6-2}

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E = ABC</i>	<i>F = BCD</i>	Combinación de tratamientos
-	-	-	-	-	-	(1)
+	-	-	-	+	-	<i>ae</i>
-	+	-	-	+	+	<i>bef</i>
+	+	-	-	-	+	<i>abf</i>
-	-	+	-	+	+	<i>cef</i>
+	-	+	-	-	+	<i>acf</i>
-	+	+	-	-	-	<i>bc</i>
+	+	+	-	+	-	<i>abce</i>
-	-	-	+	-	+	<i>df</i>
+	-	-	+	+	+	<i>adef</i>
-	+	-	+	+	-	<i>bde</i>
+	+	-	+	-	-	<i>abd</i>
-	-	+	+	+	-	<i>cde</i>
+	-	+	+	-	-	<i>acd</i>
-	+	+	+	-	+	<i>bcdf</i>
+	+	+	+	+	+	<i>abcdef</i>

efectos principales tendrán alias con interacciones que no son menos complejas que las de tercer orden. La estructura de los alias para los efectos principales se escribe

$$\begin{aligned}
 A &\equiv BCE \equiv ABCDF \equiv DEF, & D &\equiv ABCDE \equiv BCF \equiv AEF, \\
 B &\equiv ACE \equiv CDF \equiv ABDEF, & E &\equiv ABC \equiv BCDEF \equiv ADF, \\
 C &\equiv ABE \equiv BDF \equiv ACDEF, & F &\equiv ABCEF \equiv BCD \equiv ADE,
 \end{aligned}$$

cada uno con un solo grado de libertad. Para las interacciones de dos factores,

$$\begin{aligned}
 AB &\equiv CE \equiv ACDF \equiv BDEF, & AF &\equiv BCEF \equiv ABCD \equiv DE, \\
 AC &\equiv BE \equiv ABDF \equiv CDEF, & BD &\equiv ACDE \equiv CF \equiv ABEF, \\
 AD &\equiv BCDE \equiv ABCF \equiv EF, & BF &\equiv ACEF \equiv CD \equiv ABDE, \\
 AE &\equiv BC \equiv ABCDEF \equiv DF.
 \end{aligned}$$

Por supuesto, aquí hay algunos alias entre las interacciones de dos factores. Los dos grados de libertad restantes se explican por medio de los siguientes grupos:

$$ABD \equiv CDE \equiv ACF \equiv BEF, \quad ACD \equiv BDE \equiv ABF \equiv CEF.$$

Es evidente que antes de recomendar finalmente el plan experimental siempre debemos estar conscientes de que la estructura de alias es para un experimento fraccionario. La selección adecuada de contrastes de definición es importante, ya que es lo que determina la estructura de los alias.

15.7 Análisis de experimentos factoriales fraccionados

La dificultad para realizar pruebas formales de significancia con datos de experimentos factoriales fraccionados radica en la determinación del término del error apropiado.

A menos que se disponga de datos de experimentos anteriores, el error debe provenir de una agrupación de contrastes que representan efectos que se presume son despreciables.

Las sumas de cuadrados para los efectos individuales se calculan usando en esencia los mismos procedimientos que se emplean para obtener el factorial completo. Es posible formar un contraste en las combinaciones de tratamientos construyendo la tabla de signos positivos y negativos. Por ejemplo, para media réplica de un experimento factorial 2^3 con ABC como contraste de definición, un conjunto posible de combinaciones de tratamientos, junto con el signo algebraico apropiado para cada contraste que se usa para calcular los efectos y las sumas de cuadrados de los distintos efectos, sería como el que se presenta en la tabla 15.13.

Tabla 15.13: Signos para los contrastes en media réplica de un experimento factorial 2^3

Combinación de tratamientos	Efecto factorial						
	A	B	C	AB	AC	BC	ABC
a	+	-	-	-	-	+	+
b	-	+	-	-	+	-	+
c	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+

Observe que en la tabla 15.13 los contrastes A y BC son idénticos, lo cual ilustra los alias. Asimismo, $B \equiv AC$ y $C \equiv AB$. En esta situación se tienen tres contrastes ortogonales que representan los 3 grados de libertad disponibles. Si se obtuvieran dos observaciones para cada una de las cuatro combinaciones de tratamientos, entonces tendríamos un estimado de la varianza del error con 4 grados de libertad. Si suponemos que los efectos de interacción son despreciables, podríamos probar la significancia de todos los efectos principales.

Un ejemplo del efecto y la suma de cuadrados correspondientes es

$$A = \frac{a - b - c + abc}{2n}, \quad SCA = \frac{(a - b - c + abc)^2}{2^2 n}.$$

En general, la suma de cuadrados con un grado de libertad para cualquier efecto en una fracción 2^{-p} de un experimento factorial 2^k ($p < k$) se obtiene elevando al cuadrado los contrastes en los totales de los tratamientos seleccionados y dividiendo entre $2^{k-p}n$, donde n es el número de réplicas de estas combinaciones de tratamientos.

Ejemplo 15.6: Suponga que se desea emplear una media réplica para estudiar los efectos de cinco factores, cada uno en dos niveles, sobre alguna respuesta, y que se conoce que cualquiera que sea el efecto de cada factor, será constante para cada nivel de los demás factores. En otras palabras, no hay interacciones. Sea el contraste de definición $ABCDE$ lo que ocasiona que los efectos principales tengan alias con interacciones de cuatro factores. El agrupamiento de contrastes que incluyen interacciones proporciona $15 - 5 = 10$ grados de libertad para el error. Realice un análisis de varianza con los datos de la tabla 15.14 y pruebe todos los efectos principales a un nivel de significancia de 0.05.

Solución: Las sumas de cuadrados y los efectos para los efectos principales son

$$SCA = \frac{(11.3 - 15.6 - \dots - 14.7 + 13.2)^2}{2^{5-1}} = \frac{(-17.5)^2}{16} = 19.14.$$

Tabla 15.14: Datos para el ejemplo 15.6

Tratamiento	Respuesta	Tratamiento	Respuesta
<i>a</i>	11.3	<i>bcd</i>	14.1
<i>b</i>	15.6	<i>abe</i>	14.2
<i>c</i>	12.7	<i>ace</i>	11.7
<i>d</i>	10.4	<i>ade</i>	9.4
<i>e</i>	9.2	<i>bce</i>	16.2
<i>abc</i>	11.0	<i>bde</i>	13.9
<i>abd</i>	8.9	<i>cde</i>	14.7
<i>acd</i>	9.6	<i>abcde</i>	13.2

$$A = -\frac{17.5}{8} = -2.19,$$

$$SCB = \frac{(-11.3 + 15.6 - \dots - 14.7 + 13.2)^2}{2^{5-1}} = \frac{(18.1)^2}{16} = 20.48,$$

$$B = \frac{18.1}{8} = 2.26,$$

$$SCC = \frac{(-11.3 - 15.6 + \dots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(10.3)^2}{16} = 6.63,$$

$$C = \frac{10.3}{8} = 1.21,$$

$$SCD = \frac{(-11.3 - 15.6 - \dots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(-7.7)^2}{16} = 3.71,$$

$$D = \frac{-7.7}{8} = -0.96,$$

$$SCE = \frac{(-11.3 - 15.6 - \dots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(8.9)^2}{16} = 4.95,$$

$$E = \frac{8.9}{8} = 1.11.$$

Todos los demás cálculos y pruebas de significancia se resumen en la tabla 15.15. Las pruebas indican que el factor *A* tiene un efecto negativo significativo sobre la respuesta; mientras que el factor *B* tiene un efecto positivo significativo. Los factores *C*, *D* y *E* no son significativos al nivel de significancia de 0.05. \blacksquare

Ejercicios

15.18 Liste los alias de los diferentes efectos en un experimento factorial 2^5 cuando el contraste de definición es *ACDE*.

15.19 *a)* Obtenga una fracción de $\frac{1}{2}$ de un diseño factorial 2^4 usando *BCD* como el contraste de definición.

b) Divida la fracción de $\frac{1}{2}$ en dos bloques de cuatro unidades cada uno confundiendo *ABC*.

c) Construya la tabla de análisis de varianza (fuentes de variación y grados de libertad) para probar todos los efectos principales no confundidos, si se acepta que todas las interacciones de los efectos son despreciables.

15.20 Construya una fracción de $\frac{1}{4}$ de un diseño factorial 2^6 utilizando *ABCD* y *BDEF* como los contrastes de definición. Diga cuáles efectos tienen alias con los seis efectos principales.

Tabla 15.15: Análisis de varianza para los datos de media réplica de un experimento factorial 2^5

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Efecto principal:				
<i>A</i>	19.14	1	19.14	6.21
<i>B</i>	20.48	1	20.48	6.65
<i>C</i>	6.63	1	6.63	2.15
<i>D</i>	3.71	1	3.71	1.20
<i>E</i>	4.95	1	4.95	1.61
Error	30.83	10	3.08	
Total	85.74	15		

15.21 a) Con los contrastes de definición *ABCE* y *ABDF* obtenga una fracción de $\frac{1}{4}$ de un diseño 2^6 .

b) Muestre la tabla del análisis de varianza (fuentes de variación y grados de libertad) para todas las pruebas apropiadas, suponiendo que *E* y *F* no interactúan y que las interacciones de tres factores y mayores son despreciables.

15.22 En un experimento que implica sólo 16 ensayos se varían siete factores en dos niveles. Se utiliza un experimento factorial 2^7 con una fracción de $\frac{1}{8}$, con los contrastes de definición *ACD*, *BEF* y *CEG*. Los datos son los siguientes:

Combinación de tratamientos	Respuesta	Combinación de tratamientos	Respuesta
(1)	31.6	<i>acg</i>	31.1
<i>ad</i>	28.7	<i>cdg</i>	32.0
<i>abce</i>	33.1	<i>beg</i>	32.8
<i>cdef</i>	33.6	<i>adefg</i>	35.3
<i>acef</i>	33.7	<i>efg</i>	32.4
<i>bcde</i>	34.2	<i>abdeg</i>	35.3
<i>abdf</i>	32.5	<i>bcdfg</i>	35.6
<i>bf</i>	27.8	<i>abcfg</i>	35.1

Realice un análisis de varianza sobre los siete efectos principales, suponiendo que las interacciones son despreciables. Use un nivel de significancia de 0.05.

15.23 Se lleva a cabo un experimento para que un ingeniero adquiera conocimiento acerca de cómo influye la temperatura de sellado *A*, la temperatura de enfriamiento de una barra *B*, el porcentaje de aditivo de policetileno *C* y la presión *D* sobre la resistencia del sello (en gramos por pulgada) de un lote de envoltura para pan. Se utiliza un experimento factorial 2^4 con fracción de $\frac{1}{2}$ con un contraste de definición *ABCD*. A continuación se presentan los datos. Realice un análisis de varianza sólo sobre los efectos principales usando $\alpha = 0.05$

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	Respuesta
-1	-1	-1	-1	6.6
1	-1	-1	1	6.9
-1	1	-1	1	7.9
1	1	-1	-1	6.1
-1	-1	1	1	9.2
1	-1	1	-1	6.8
-1	1	1	-1	10.4
1	1	1	1	7.3

15.24 En un experimento realizado en el Departamento de Ingeniería Mecánica, y analizado por el Centro de Consultoría en Estadística de Virginia Tech, un sensor detecta una carga eléctrica cada vez que las aspas de una turbina completan un giro. Luego, el sensor mide la amplitud de la corriente eléctrica. Seis factores son rpm *A*, temperatura *B*, distancia entre las aspas *C*, distancia entre las aspas y la carcasa *D*, ubicación de la entrada *E*, y ubicación del detector *F*. Se utiliza un experimento factorial 2^6 con fracción de $\frac{1}{2}$, con contrastes de definición *ABCE* y *BCDF*. Los datos son los siguientes:

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	Respuesta
-1	-1	-1	-1	-1	-1	3.89
1	-1	-1	-1	1	-1	10.46
-1	1	-1	-1	1	1	25.98
1	1	-1	-1	-1	1	39.88
-1	-1	1	-1	1	1	61.88
1	-1	1	-1	-1	1	3.22
-1	1	1	-1	-1	-1	8.94
1	1	1	-1	1	-1	20.29
-1	-1	-1	1	-1	1	32.07
1	-1	-1	1	1	1	50.76
-1	1	-1	1	1	-1	2.80
1	1	-1	1	-1	-1	8.15
-1	-1	1	1	1	-1	16.80
1	-1	1	1	-1	-1	25.47
-1	1	1	1	-1	1	44.44
1	1	1	1	1	1	2.45

Realice un análisis de varianza de los efectos principales y las interacciones de dos factores, si se acepta que las interacciones de tres factores o más son despreciables. Use $\alpha = 0.05$.

15.25 En el estudio denominado *Durability of Rubber to Steel Adhesively Bonded Joints*, efectuado por el Departamento de Ciencias del Ambiente y Mecánica, y analizado por el Centro de Consultoría en Estadística de Virginia Tech, un experimentador midió el número de roturas en un sello adhesivo. Se planteó que la concentración de agua marina A , la temperatura B , el pH C , el voltaje D y la tensión E influyen en el rompimiento de un sello adhesivo. Se utilizó un experimento factorial 2^5 con fracción de $\frac{1}{2}$ y con el contraste de definición $AB CDE$. Los datos son los siguientes:

A	B	C	D	E	Respuesta
-1	-1	-1	-1	1	462
1	-1	-1	-1	-1	746
-1	1	-1	-1	-1	714
1	1	-1	-1	1	1070
-1	-1	1	-1	-1	474
1	-1	1	-1	1	832
-1	1	1	-1	1	764
1	1	1	-1	-1	1087
-1	-1	-1	1	-1	522
1	-1	-1	1	1	854
-1	1	-1	1	1	773
1	1	-1	1	-1	1068
-1	-1	1	1	1	572
1	-1	1	1	-1	831
-1	1	1	1	-1	819
1	1	1	1	1	1104

Realice un análisis de varianza de los efectos principales y de las interacciones de dos factores AD , AE , BD , BE ; suponga que las interacciones de tres o más factores son despreciables. Use $\alpha = 0.05$.

15.26 Considere un diseño 2^{5-1} con los factores A , B , C , D y E . Construya el diseño comenzando con un

diseño 2^4 y use $E = ABCD$ como generador. Indique todos los alias.

15.27 Hay seis factores y sólo se pueden usar ocho puntos de diseño. Construya un diseño 2^{6-3} , comenzando con un diseño 2^3 , y utilice $D = AB$, $E = -AC$ y $F = BC$ como generadores.

15.28 Considere el ejercicio 15.27. Construya otro 2^{6-3} que sea diferente del diseño elegido en el ejercicio 15.27.

15.29 Para el ejercicio 15.27 proporcione todos los alias para los seis efectos principales.

15.30 En Myers, Montgomery y Anderson-Cook (2009) se analiza una aplicación en la cual a un ingeniero le interesan los efectos del agrietamiento de una aleación de titanio. Los tres factores son A , temperatura; B , contenido de titanio; y C , cantidad de refinador en grano. La siguiente tabla presenta una parte del diseño y la respuesta, la longitud de las grietas inducida en la muestra de la aleación.

A	B	C	Respuesta
-1	-1	-1	0.5269
1	1	-1	2.3380
1	-1	1	4.0060
1	1	1	3.3640

- ¿Cuál es la relación de definición?
- Proporcione alias para los tres efectos principales asumiendo que las interacciones de dos factores pueden ser reales.
- Si suponemos que las interacciones son despreciables, ¿cuál será el factor principal más importante?
- ¿Qué nivel sugeriría para el factor obtenido en el inciso c en la producción final, alto o bajo?
- ¿Qué niveles sugeriría para los demás factores en la producción final?
- ¿Qué riesgos hay en las recomendaciones que hizo en los incisos d y e ? Responda de manera detallada.

15.8 Diseños de fracciones superiores y de filtrado

Algunas situaciones industriales requieren que el analista determine cuáles factores controlables, de entre un número grande de ellos, tienen un efecto sobre alguna respuesta importante. Los factores pueden ser cualitativos o variables de clase, variables de regresión o una mezcla de ambas. El procedimiento analítico puede requerir un análisis de varianza, una regresión o ambos. A menudo el modelo de regresión utilizado sólo incluye los efectos lineales principales, aunque tal vez sea posible estimar algunas interacciones. La situación exige la selección de variables y los diseños experimentales resultantes se denominan **diseños de filtrado**. Es evidente que los diseños ortogonales de dos niveles saturados o casi saturados son candidatos viables.

Resolución del diseño

A menudo los diseños ortogonales de dos niveles se clasifican según su **resolución**, la cual es determinada por la siguiente definición.

Definición 15.1: La **resolución** de un diseño ortogonal de dos niveles es la longitud de la interacción más pequeña (menos compleja) de entre el conjunto de contrastes de definición.

Si el diseño se construye como un factorial completo o fraccionado, ya sea un diseño 2^k , o bien, 2^{k-p} , $p = 1, 2, \dots, k - 1$, el concepto de resolución del diseño es un auxiliar para determinar el efecto de los alias. Por ejemplo, un diseño de resolución II sería de poca utilidad, ya que habría al menos un caso de alias de un efecto principal con otro. Un diseño de resolución III tendría todos sus efectos principales (lineales) ortogonales entre sí. No obstante, habrá algunos alias entre los efectos lineales y las interacciones de dos factores. Entonces, es evidente que si el analista está interesado en estudiar los efectos principales (lineales en el caso de la regresión) y no hay interacciones de dos factores, entonces se requiere un diseño cuya resolución sea de al menos III.

15.9 Construcción de diseños de resolución III y IV, con 8, 16 y 32 puntos de diseño

Es posible construir diseños útiles con resoluciones III y IV para 2 a 7 variables con 8 puntos de diseño. Empezamos con un factorial 2^3 que haya sido saturado simbólicamente con interacciones.

x_1	x_2	x_3	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	$x_1 x_2 x_3$
-1	-1	-1	1	1	1	-1
1	-1	-1	-1	-1	1	1
-1	1	-1	-1	1	-1	1
-1	-1	1	1	-1	-1	1
1	1	-1	1	-1	-1	-1
1	-1	1	-1	1	-1	-1
-1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1

Es evidente que, con sólo reemplazar las columnas de interacción por nuevos efectos principales para las siete variables, se puede construir un diseño de resolución III. Por ejemplo, podríamos definir

$$\begin{aligned}
 x_4 &= x_1 x_2 && \text{(contraste de definición ABD)} \\
 x_5 &= x_1 x_3 && \text{(contraste de definición ACE)} \\
 x_6 &= x_2 x_3 && \text{(contraste de definición BCF)} \\
 x_7 &= x_1 x_2 x_3 && \text{(contraste de definición ABCG)}
 \end{aligned}$$

y obtendríamos una fracción 2^{-4} de un factorial 2^7 . Las expresiones anteriores identifican los contrastes de definición elegidos. Resultan once contrastes de definición adicionales y todos contienen al menos tres letras. Así, el diseño es de resolución III. Es evidente que si se comienza con un *subconjunto* de columnas aumentadas y se concluye con un diseño

Tabla 15.16: Algunos diseños 2^{k-p} de resoluciones III, IV, V, VI y VII

Número de factores	Diseño	Número de puntos	Generadores
3	2 ³⁻¹ _{III}	4	$C = \pm AB$
4	2 ⁴⁻¹ _{IV}	8	$D = \pm ABC$
5	2 ⁵⁻² _{III}	8	$D = \pm AB ; E = \pm AC$
6	2 ⁶⁻¹ _{VI}	32	$F = \pm ABCDE$
	2 ⁶⁻² _{IV}	16	$E = \pm ABC ; F = \pm BCD$
	2 ⁶⁻³ _{III}	8	$D = \pm AB ; F = \pm BC ; E = \pm AC$
7	2 ⁷⁻¹ _{VII}	64	$G = \pm ABCDEF$
	2 ⁷⁻² _{IV}	32	$E = \pm ABC ; G = \pm ABDE$
	2 ⁷⁻³ _{IV}	16	$E = \pm ABC ; F = \pm BCD ; G = \pm ACD$
	2 ⁷⁻⁴ _{III}	8	$D = \pm AB ; E = \pm AC ; F = \pm BC ; G = \pm ABC$
8	2 ⁸⁻² _V	64	$G = \pm ABCD ; H = \pm ABEF$
	2 ⁸⁻³ _{IV}	32	$F = \pm ABC ; G = \pm ABD ; H = \pm BCDE$
	2 ⁸⁻⁴ _{IV}	16	$E = \pm BCD ; F = \pm ACD ; G = \pm ABC ; H = \pm ABD$

que incluye menos de 7 variables de diseño, el resultado es un diseño de resolución III en menos de siete variables.

Es posible construir un conjunto similar de diseños posibles para 16 puntos de diseño, comenzando con un diseño 2⁴ saturado con interacciones. Las definiciones de las variables que corresponden a estas interacciones producen diseños de resolución III por medio de 15 variables. De manera similar, se pueden construir diseños que contengan 32 corridas, comenzando con un diseño 2⁵.

La tabla 15.16 proporciona lineamientos para construir diseños de 8, 16, 32 y 64 puntos, con resolución III, IV e incluso V. La tabla proporciona el número de factores, el número de corridas y los generadores que se utilizan para producir los diseños 2^{k-p}. El generador dado se emplea para **aumentar el factorial completo** que contiene $k - p$ factores.

15.10 Otros diseños de resolución III de dos niveles; los diseños de Plackett-Burman

Una familia de diseños desarrollada por Plackett y Burman (1946, véase la bibliografía) llena el vacío del tamaño de la muestra que existe con los factoriales fraccionados. Éstos son útiles con muestras de tamaño 2^r, es decir, incluyen muestras de tamaños 4, 8, 16, 32, 64, ... Los diseños de Plackett-Burman incluyen 4r puntos de diseño, por lo que se dispone de diseños de tamaño 12, 20, 24, 28, etcétera. Estos diseños de Plackett-Burman de dos niveles son diseños de resolución III y son muy fáciles de construir. Se proporcionan "renglones básicos" para cada tamaño de muestra. Estos renglones de signos + y - son $n - 1$ en número. Para construir las columnas de la matriz de diseño se comienza con el renglón básico y se hace una permutación cíclica sobre las columnas, hasta que se forman k columnas (el número deseado de variables). Después se llena el último

renglón con signos negativos. El resultado será un diseño de resolución III con k variables ($k = 1, 2, \dots, N$). Los renglones básicos son los siguientes:

$N = 12$	+	+	-	+	+	+	-	-	-	+	-												
$N = 16$	+	+	+	+	-	+	-	+	+	-	-	+	-	-	-								
$N = 20$	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	+	+	-					
$N = 24$	+	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-

Ejemplo 15.7 Construya un diseño depurado de dos niveles con 6 variables que contengan 12 puntos de diseño.

Solución: Comience con el renglón básico en la columna inicial. La segunda columna se forma llevando la entrada inferior de la primera columna a la parte superior de la segunda, y repitiendo la primera. La tercera columna se forma del mismo modo, utilizando las entradas de la segunda columna. Cuando haya un número suficiente de columnas sencillamente se llena el último renglón con signos negativos. El diseño resultante es como sigue:

	x_1	x_2	x_3	x_4	x_5	x_6
	+	-	+	-	-	-
	+	+	-	+	-	-
	-	+	+	-	+	-
	+	-	+	+	-	+
	+	+	-	+	+	-
	+	+	+	-	+	+
	-	+	+	+	-	+
	-	-	+	+	+	-
	-	-	-	+	+	+
	+	-	-	-	+	+
	-	+	-	-	-	+
	-	-	-	-	-	-

Los diseños de Plackett-Burman son populares en la industria para situaciones de filtrado. Como se trata de diseños de resolución III, todos los efectos lineales son ortogonales. Para cualquier tamaño de muestra el usuario dispone de un diseño para $k = 2, 3, \dots, N - 1$ variables.

La estructura de alias para el diseño de Plackett-Burman es muy complicada, por lo que el usuario no puede construir el diseño con un control completo de la estructura de alias, como en el caso de los diseños 2^k o 2^{k-p} . Sin embargo, en el caso de modelos de regresión el diseño de Plackett-Burman acepta interacciones (aunque no serán ortogonales) cuando se dispone de suficientes grados de libertad. ■

15.11 Introducción a la metodología de superficie de respuesta

En el estudio de caso 15.2 se ajustó un modelo de regresión a un conjunto de datos con la meta específica de encontrar condiciones en esas variables de diseño que optimizaran (maximizaran) la eficiencia de purificación del carbón. El modelo incluía tres efectos principales lineales, tres términos de interacción de dos factores y un término de interacción de tres factores. La respuesta del modelo era la eficiencia de la purificación, y las condiciones óptimas de x_1 , x_2 y x_3 se obtuvieron utilizando los signos y la magnitud

de los coeficientes del modelo. En este ejemplo se utilizó un diseño de dos niveles para mejorar el proceso o para optimizarlo. En muchas áreas de la ciencia y de la ingeniería la aplicación se extiende para incluir modelos y diseños más complicados a los que, en conjunto, se les denomina **metodología de superficie de respuesta (MSR)**. Esta metodología abarca tanto métodos gráficos como analíticos. El término *superficie de respuesta* se deriva de la apariencia de la superficie multidimensional de la respuesta estimada constante de un modelo de segundo orden, es decir, un modelo con términos de primer y segundo orden. A continuación se presenta un ejemplo.

El modelo de superficie de respuesta de segundo orden

En muchos ejemplos industriales de optimización de procesos se utiliza un *modelo de superficie de respuesta de segundo orden*. Para el caso de, digamos $k = 2$ variables de proceso o variables de diseño, y una sola respuesta y , el modelo es dado por

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon.$$

Aquí se tienen $k = 2$ términos de primer orden, dos términos puros de segundo orden o cuadráticos y un término de interacción dado por $\beta_{12} x_1 x_2$. Los términos x_1 y x_2 se codifican en la forma conocida de ± 1 . El término ϵ denota al acostumbrado error del modelo. En general, para k variables de diseño el modelo contendrá $1 + k + k + \binom{k}{2}$ términos del modelo y , por lo tanto, el diseño experimental debe contener al menos un número similar de puntos de diseño. Además, los términos cuadráticos requieren que las variables de diseño estén fijas en el diseño con al menos tres niveles. Al diseño resultante se le denomina *diseño de segundo orden*. A continuación se presenta un ejemplo.

El siguiente **diseño central compuesto (DCC)** y el ejemplo fueron tomados de Myers, Montgomery y Anderson-Cook (2009). Quizás la clase más popular de diseños de segundo orden sea la clase de los diseños centrales compuestos. El ejemplo que se presenta en la tabla 15.17 se refiere a un proceso químico en el que la temperatura de reacción, ξ_1 , y la concentración del reactante, ξ_2 , se muestran en sus niveles naturales y también de forma codificada. Cada factor tiene cinco niveles. Además, se incluye el orden en que se realizaron las observaciones de x_1 y x_2 . La columna de la derecha proporciona los valores de la respuesta y , el porcentaje de conversión del proceso. Los primeros cuatro puntos de diseño representan los conocidos puntos factoriales en los niveles ± 1 . Los siguientes cuatro puntos se conocen como puntos axiales, los cuales van seguidos por las corridas centrales que se explicaron y ejemplificaron antes en este capítulo. De esta manera, los cinco niveles de cada uno de los dos factores son -1 , $+1$, -1.414 , $+1.414$ y 0 . En la figura 15.16 se presenta una imagen clara de la geometría del diseño central compuesto para este ejemplo de $k = 2$. En esta figura se ilustra la fuente del término **puntos axiales**. Estos cuatro puntos se localizan sobre los ejes factoriales, a una distancia axial de $\alpha = \sqrt{2} = 1.414$ a partir del centro del diseño. De hecho, para este DCC en particular, los puntos del perímetro, axiales y factoriales, se encuentran todos a la distancia $\sqrt{2}$ del centro del diseño, y como resultado tenemos ocho puntos equidistantes sobre un círculo más cuatro réplicas en el centro del diseño.

Ejemplo 15.8: **Análisis de superficie de respuesta:** Un análisis de los datos en el ejemplo de las dos variables podría implicar el ajuste de una función de respuesta de segundo orden. La superficie de respuesta resultante se puede utilizar de forma analítica o gráfica para determinar el impacto que tienen x_1 y x_2 sobre el porcentaje de conversión del proceso. Los coeficientes en la función de respuesta están determinados por medio del método de

Tabla 15.17: Diseño central compuesto para el ejemplo 15.8

Observación	Corrida	Temperatura (°C)		Concentración (%)		y
		ξ_1	ξ_2	x_1	x_2	
1	4	200	15	-1	-1	43
2	12	250	15	1	-1	78
3	11	200	25	-1	1	69
4	5	250	25	1	1	73
5	6	189.65	20	-1.414	0	48
6	7	260.35	20	1.414	0	78
7	1	225	12.93	0	-1.414	65
8	3	225	27.07	0	1.414	74
9	8	225	20	0	0	76
10	10	225	20	0	0	79
11	9	225	20	0	0	83
12	2	225	20	0	0	81

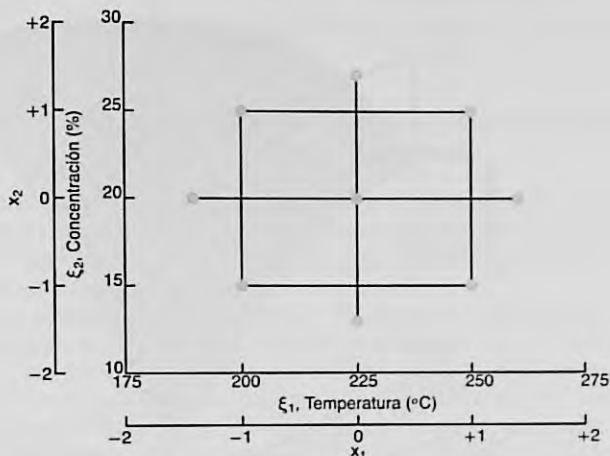


Figura 15.16: Diseño central compuesto para el ejemplo 15.8.

mínimos cuadrados que estudiamos en el capítulo 12 y que ejemplificamos a lo largo de este capítulo. El modelo resultante de respuesta de segundo orden es dado en las variables codificadas como

$$\hat{y} = 79.75 + 10.18x_1 + 4.22x_2 - 8.50x_1^2 - 5.25x_2^2 - 7.75x_1x_2,$$

mientras que en las variables naturales es dado por

$$\hat{y} = -1080.22 + 7.7671\xi_1 + 23.1932\xi_2 - 0.0136\xi_1^2 - 0.2100\xi_2^2 - 0.0620\xi_1\xi_2.$$

Como este ejemplo sólo incluye dos variables de diseño, el método más esclarecedor para determinar la naturaleza de la superficie de respuesta en la región del diseño

consiste en utilizar gráficas de dos o tres dimensiones. Sería interesante determinar cuáles niveles de temperatura x_1 y concentración x_2 producen un estimado deseable del porcentaje de conversión \hat{y} . La función de respuesta estimada anterior se graficó en tres dimensiones y la *superficie de respuesta* resultante se presenta en la figura 15.17. La altura de la superficie es \hat{y} expresada en porcentaje. En esta figura es fácil observar por qué se utiliza el término **superficie de respuesta**. En el caso en que sólo se utilizan dos variables de diseño las gráficas bidimensionales de curvas pueden ser útiles. Observe en la figura 15.18 que las curvas de la conversión constante estimada se ven como rodajas de la superficie de respuesta. Observe que cualquiera de las dos figuras indica con facilidad cuáles coordenadas de temperatura y concentración producen el mayor porcentaje de conversión estimado. En las gráficas las coordenadas se presentan tanto en unidades codificadas como en unidades naturales. Observe que la mayor conversión estimada se encuentra en aproximadamente 240°C y una concentración de 20%. La respuesta máxima estimada (o pronosticada) en esa ubicación es 82.47%.

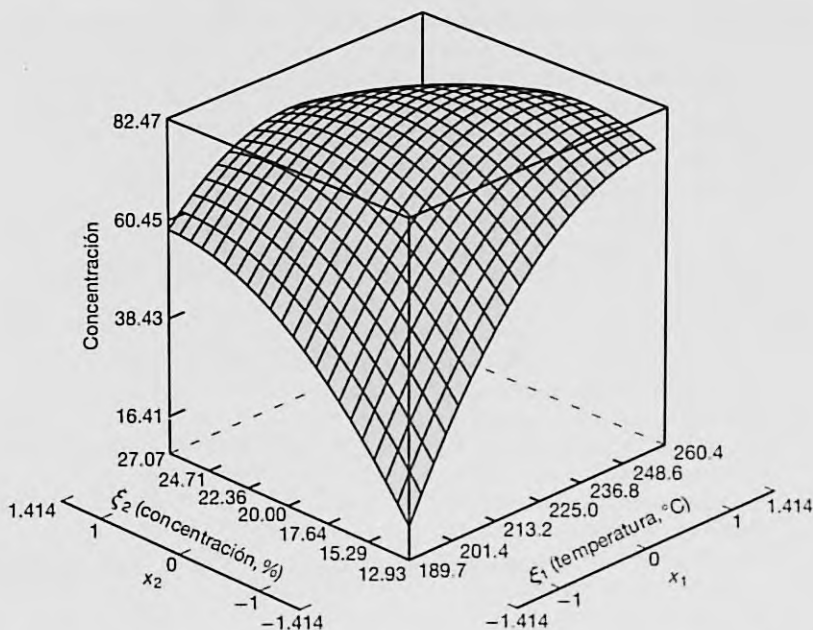


Figura 15.17: Gráfica de la superficie de respuesta de la conversión pronosticada para el ejemplo 15.8.

Otros comentarios acerca del análisis de superficie de respuesta

El libro de Myers, Montgomery y Anderson-Cook (2009) proporciona una gran cantidad de información sobre el análisis y el diseño de la metodología de superficie de respuesta. La ilustración gráfica que se utilizó aquí podría ampliarse con resultados analíticos que brindan información acerca de la naturaleza de la superficie de respuesta dentro de la región del diseño.

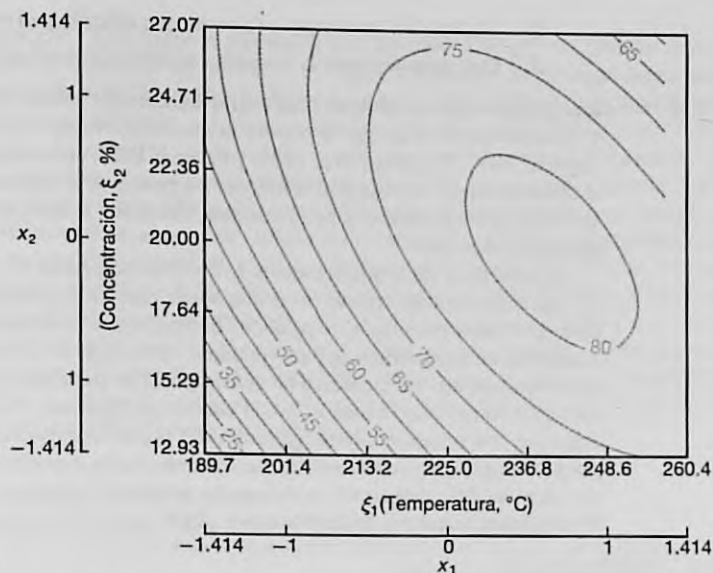


Figura 15.18: Gráfica de curvas de la conversión pronosticada para el ejemplo 15.8.

Se pueden usar otros cálculos para determinar si la ubicación de las condiciones óptimas está dentro o muy lejos de la región del diseño experimental. Existen muchos aspectos importantes a tomar en cuenta cuando se necesita determinar las condiciones apropiadas para la operación futura de un proceso.

Otras secciones del libro de Myers, Montgomery y Anderson-Cook (2009) abordan otros aspectos del diseño experimental. Por ejemplo, el diseño central compuesto, aunque es el tipo de diseño más útil, no es el único que se utiliza en la metodología de superficie de respuesta. En el libro mencionado se analizan muchos otros tipos. Además, el diseño central compuesto al que aquí nos referimos es un caso especial en el que $k = 2$. El caso más general $k > 2$ se analiza en Myers, Montgomery y Anderson-Cook (2009).

15.12 Diseño robusto de parámetros

En este capítulo se destacó el concepto del uso del diseño de experimentos (DE) para adquirir conocimientos sobre procesos de ingeniería y científicos. En el caso en que un proceso incluye un producto es posible usar el DE para mejorar el producto o la calidad. Como se expuso en el capítulo 1, se ha dado mucha importancia al empleo de métodos estadísticos para mejorar los productos. Un aspecto importante de este esfuerzo por mejorar la calidad, que surgió en la década de 1980 y continuó a lo largo de la década de 1990, consiste en incluir la calidad en los procesos y productos en la etapa de investigación o de diseño del proceso. A menudo se requiere del DE para desarrollar procesos con las siguientes propiedades:

1. Insensibles (robustos) a las condiciones ambientales

2. Insensibles (robustos) a factores que dificultan el control
3. Que proporcionen la mínima variación de desempeño

Los métodos que se utilizan para lograr las características deseables en los puntos 1, 2 y 3 forman parte de lo que se conoce como *diseño robusto de parámetros* o DRP (véase Taguchi, 1991; Taguchi y Wu, 1985; y Kackar, 1985, en la bibliografía). En este contexto el término *diseño* se refiere al diseño de los procesos o sistemas, en tanto que *parámetro* se refiere a los parámetros en el sistema. Éstos son a los que nos hemos referido como *factores* o *variables*.

Queda muy claro que las metas 1, 2 y 3 mencionadas son muy nobles. Por ejemplo, un ingeniero petrolero puede tener una buena mezcla de gasolina que se desempeñe muy bien en condiciones ideales y estables. Sin embargo, el desempeño se deteriorará debido a cambios en las condiciones ambientales, como tipo de conductor, factores climáticos, tipo de motor, etc. Un científico de una empresa de alimentos podría tener una muy buena mezcla para pasteles, a menos que el usuario no siga con exactitud las instrucciones del empaque con respecto a la temperatura del horno, tiempo de horneado, entre otros. Un producto o proceso cuyo desempeño sea consistente cuando se expone a esas condiciones ambientales cambiantes se denomina **producto robusto** o **proceso robusto**. (Véase Myers, Montgomery y Anderson-Cook, 2009, en la bibliografía).

Variables de control y ruido

Taguchi (1991) destacó la idea de utilizar dos clases de variables de diseño en un estudio que incluye un diseño de superficie de respuesta (DSR): *factores de control* y *factores de ruido*.

Definición 15.2: Los **factores de control** son variables que se pueden controlar tanto en el experimento como en el proceso. Los **factores de ruido** son variables que pueden o no controlarse en el experimento, pero que no pueden controlarse en el proceso (o que no pueden controlarse bien).

Un método importante consiste en usar variables de control y variables de ruido en el mismo experimento, como efectos fijos. Para lograr esto con frecuencia se utilizan los diseños o arreglos ortogonales.

Meta del diseño robusto de parámetros	La meta del diseño robusto de parámetros es elegir los niveles de las variables de control, es decir, el diseño del proceso, que sean más robustos (insensibles) a los cambios en las variables de ruido.
---------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Debe señalarse que los *cambios en las variables de ruido* en realidad implican cambios durante el proceso, cambios en el campo, cambios en el ambiente, cambios en el manejo o uso por parte del consumidor, etcétera.

Arreglo del producto

Un enfoque del diseño de experimentos que incluye tanto variables de control como de ruido consiste en utilizar un plan experimental que requiere un diseño ortogonal para las variables de control y de ruido, por separado. Entonces, el experimento completo es simplemente el producto o cruce de estos dos diseños ortogonales. El siguiente es un ejemplo sencillo de un arreglo de productos con dos variables de control y dos de ruido.

Ejemplo 15.9: En el artículo "The Taguchi Approach to Parameter Design" en *Quality Progress*, de diciembre de 1987, D. M. Byrne y S. Taguchi analizan un ejemplo interesante en el que se busca un método para ensamblar un conector electrométrico a un tubo de nailon que entrega el rendimiento de arranque requerido para una aplicación de motor automotriz. El objetivo es encontrar condiciones controlables que maximicen la fuerza de arranque. Entre las variables controlables están *A*, el espesor de la pared del conector, y *B*, la profundidad de inserción. Durante la operación rutinaria existen diversas variables que no se pueden controlar, aunque se controlan durante el experimento. Entre ellas están *C*, el tiempo de acondicionamiento, y *D*, la temperatura de acondicionamiento. Se toman tres niveles para cada variable de control y dos para cada variable de ruido. Como resultado, el arreglo cruzado es el siguiente. Se trata de un arreglo de control de 3×3 y el de ruido es el conocido factorial 2^2 con (1), *c*, *d* y *cd* que representan las combinaciones de los cuatro factores. El propósito del factor de ruido es crear la *clase de variabilidad de la respuesta, la fuerza de arranque, que se podría esperar en la operación diaria con el proceso*. En la tabla 15.18 se muestra el diseño. ■

Tabla 15.18: Diseño para el ejemplo 15.9

		<i>B</i> (profundidad)		
		Superficial	Media	Profunda
<i>A</i> (espesor de pared)	Delgado	(1)	(1)	(1)
		<i>c</i>	<i>c</i>	<i>c</i>
		<i>d</i>	<i>d</i>	<i>d</i>
	Medio	<i>cd</i>	<i>cd</i>	<i>cd</i>
		(1)	(1)	(1)
		<i>c</i>	<i>c</i>	<i>c</i>
	Grueso	<i>d</i>	<i>d</i>	<i>d</i>
		<i>cd</i>	<i>cd</i>	<i>cd</i>
		(1)	(1)	(1)
		<i>c</i>	<i>c</i>	<i>c</i>
		<i>d</i>	<i>d</i>	<i>d</i>
		<i>cd</i>	<i>cd</i>	<i>cd</i>

Estudio de caso 15.3: **Optimización de proceso de soldadura.** En un experimento que Schmidt y Launsby describen en *Understanding Industrial Designed Experiments* (1991; véase la bibliografía), en una planta de ensamble de circuitos integrados se lleva a cabo la optimización de un proceso de soldadura. Las partes se insertan a mano o en forma automática en una tarjeta que tiene impreso un circuito. Una vez que las partes se insertan, la tarjeta se coloca en una máquina soldadora de ola que se emplea para conectar todos los elementos del circuito. Las tarjetas se colocan en un transportador y pasan por una serie de etapas. Se lavan en una mezcla fundente para eliminar el óxido. Para minimizar la torsión se precientan antes de aplicar la soldadura, la cual se realiza conforme las tarjetas se mueven a través de la ola de soldadura. El objetivo del experimento consiste en minimizar el número de defectos de soldadura por millón de uniones. Los factores y los niveles de control se incluyen en la tabla 15.19.

Tabla 15.19: Factores de control para el estudio de caso 15.3

Factor	(-1)	(+1)
A, temperatura del crisol para soldar (°F)	480	510
B, velocidad del transportador (pies/min)	7.2	10
C, densidad del fundente	0.9°	1.0°
D, temperatura de precalentado	150	200
E, altura de la ola (pulgadas)	0.5	0.6

A nivel experimental es fácil controlar estos factores, pero en la planta o en el proceso es mucho más difícil. ▮

Factores de ruido: tolerancias sobre los factores de control

A menudo, en procesos como éste los factores naturales de ruido son las tolerancias sobre los factores de control. Por ejemplo, en el proceso real en línea la temperatura del crisol para soldar y la velocidad de la banda transportadora son difíciles de controlar. Se sabe que el control de la temperatura está dentro de $\pm 5^\circ\text{F}$, y que el control de la velocidad de la banda está dentro de ± 0.2 pies/min. Es posible que la variabilidad de la respuesta del producto (desempeño de la soldadura) se incremente debido a la incapacidad de controlar esos dos factores en ciertos niveles nominales. El tercer factor de ruido es el tipo de ensamble involucrado. En la práctica se utilizan uno de dos tipos de ensambles. Así, se tienen los factores de ruido que se presentan en la tabla 15.20.

Tabla 15.20: Factores de ruido para el estudio de caso 15.3

Factor	(-1)	(+1)
A*, tolerancia de la temperatura del crisol para soldar (°F), (desviación de la nominal)	-5	+5
B*, tolerancia de la velocidad del transportador (pies/min), (desviación del ideal)	-0.2	+0.2
C*, tipo de ensamble	1	2

Se eligieron factoriales fraccionados tanto para el arreglo de control (arreglo interior) como para el de ruido (arreglo exterior): el primero es $\frac{1}{4}$ de un diseño 2⁵, y el segundo es $\frac{1}{2}$ de un diseño 2³. El arreglo cruzado y los valores de respuesta se presentan en la tabla 15.21. Las primeras tres columnas del arreglo interior representan un diseño 2³. La cuarta y la quinta columnas están formadas por $D = -AC$ y $E = -BC$. Así, las interacciones de definición para el arreglo interior son ACD , BCE y $ABDF$. El arreglo exterior es una fracción estándar de resolución III de un diseño 2³. Observe que cada punto del arreglo interior contiene corridas del arreglo exterior. Así, se observan cuatro valores de respuesta en cada combinación del arreglo de control. La figura 15.19 muestra gráficas que revelan el efecto de la temperatura y la densidad sobre la respuesta media.

Tabla 15.21: Arreglos cruzados y valores de respuesta para el estudio de caso 15.3

Arreglo interior					Arreglo exterior					
A	B	C	D	E	(1)	$a*b^*$	$a*c^*$	$b*c^*$	\bar{y}	s_y
1	1	1	-1	-1	194	197	193	275	214.75	40.20
1	1	-1	1	1	136	136	132	136	135.00	2.00
1	-1	1	-1	1	185	261	264	264	243.50	39.03
1	-1	-1	1	-1	47	125	127	42	85.25	47.11
-1	1	1	1	-1	295	216	204	293	252.00	48.75
-1	1	-1	-1	1	234	159	231	157	195.25	43.04
-1	-1	1	1	1	328	326	247	322	305.75	39.25
-1	-1	-1	-1	-1	186	187	105	104	145.50	47.35

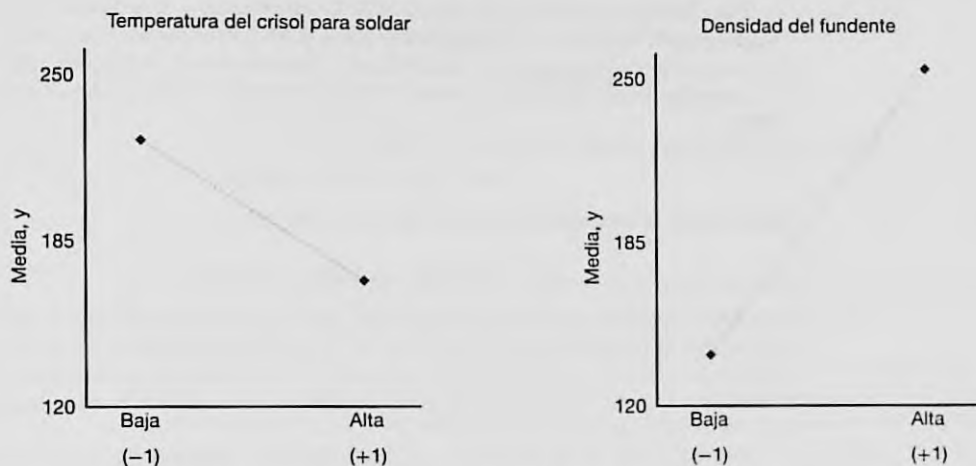


Figura 15.19: Gráfica que muestra la influencia de los factores sobre la respuesta media.

Análisis simultáneo de la media y varianza del proceso

En la mayoría de los ejemplos que utilizan DSR el analista se interesa por encontrar condiciones para las variables de control que proporcionen valores adecuados para la respuesta media \bar{y} . Sin embargo, la variación de las variables de ruido proporciona información acerca de la varianza del proceso σ_y^2 que podría anticiparse en el mismo. Es evidente que un producto robusto es aquel para el que el proceso es consistente y, por lo tanto, tiene poca varianza. El DSR puede incluir el análisis simultáneo de \bar{y} y s_y .

Resulta que la temperatura y la densidad del fundente son los factores más importantes en el estudio de caso 15.3, y al parecer influyen en s_y y \bar{y} . Por fortuna, para ambas es preferible una *alta temperatura* y una *baja densidad del fundente*. De acuerdo con la figura 15.19 las condiciones "óptimas" son

$$\text{temperatura de soldadura} = 510^\circ\text{F}, \quad \text{densidad del fundente} = 0.9^\circ.$$

Enfoques alternativos al diseño robusto de parámetros

Un enfoque sugerido por muchos estudiosos consiste en modelar la media y la varianza muestrales por separado. Con frecuencia el modelado separado ayuda al experimentador a comprender mejor el proceso involucrado. En el siguiente ejemplo se ilustra este enfoque con el experimento del proceso de soldadura.

Estudio de caso 15.4: Considere los datos del estudio de caso 15.3. Un método alternativo consiste en ajustar modelos separados para la media \bar{y} y la desviación estándar muestral. Suponga que se usa el código habitual $+1$ y -1 para los factores de control. Con base en la importancia aparente de la temperatura del crisol para soldar x_1 y la densidad del fundente x_2 , la regresión lineal sobre la respuesta (número de errores por millón de uniones) produce

$$\hat{y} = 197.125 - 27.5x_1 + 57.875x_2.$$

Para obtener los niveles más robustos de la temperatura y la densidad del fundente es conveniente establecer un compromiso entre la respuesta media y la variabilidad, y para esto es necesario modelar la variabilidad. Una herramienta importante para hacerlo es la transformación logarítmica (véase Bartlett y Kendall, 1946, o Carroll y Ruppert, 1988):

$$\ln s^2 = \gamma_0 + \gamma_1(x_1) + \gamma_2(x_2).$$

Este proceso de modelado produce el siguiente resultado:

$$\widehat{\ln s^2} = 6.6975 - 0.7458x_1 + 0.6150x_2.$$

El modelo *logarítmico lineal* tiene un amplio uso en el modelado de la varianza muestral, ya que la transformación logarítmica de la varianza muestral se presta al uso del método de mínimos cuadrados. Esto resulta del hecho de que las suposiciones de normalidad y de varianza homogénea a menudo son muy buenas cuando se utiliza $\ln s^2$ en lugar de s^2 como respuesta del modelo.

El análisis que es importante para el científico o el ingeniero echa mano de los dos modelos al mismo tiempo. Un método gráfico puede ser muy útil. La figura 15.20 presenta al mismo tiempo gráficas sencillas de los modelos de la media y de la desviación estándar. Como se esperaba, la ubicación de la temperatura y la densidad del fundente que minimizan el número medio de errores es la misma que la que minimiza la variabilidad, es decir, temperatura alta y densidad del flujo baja. El método gráfico de la *superficie múltiple de respuesta* permite que el usuario perciba intercambios entre la media del proceso y su variabilidad. Para este ejemplo es probable que el ingeniero se sienta insatisfecho con las condiciones extremas de la temperatura de la soldadura y la densidad del fundente. La figura ofrece estimados de lo que se pierde a medida que uno se aleja de las condiciones óptimas de la media y la variabilidad hacia cualquier condición intermedia.

En el estudio de caso 15.4 para las variables de control se eligieron valores que proporcionaran condiciones deseables tanto para la media como para la varianza del proceso. Se tomaron la media y la varianza a través de la distribución de las variables de ruido en el proceso y se modelaron por separado, y se encontraron condiciones apropiadas por medio de un método doble de superficie de respuesta. Como el estudio de caso 15.4 incluye dos modelos (media y varianza) podría considerarse un análisis doble de superficie de respuesta. Por fortuna, en este ejemplo las mismas condiciones de las

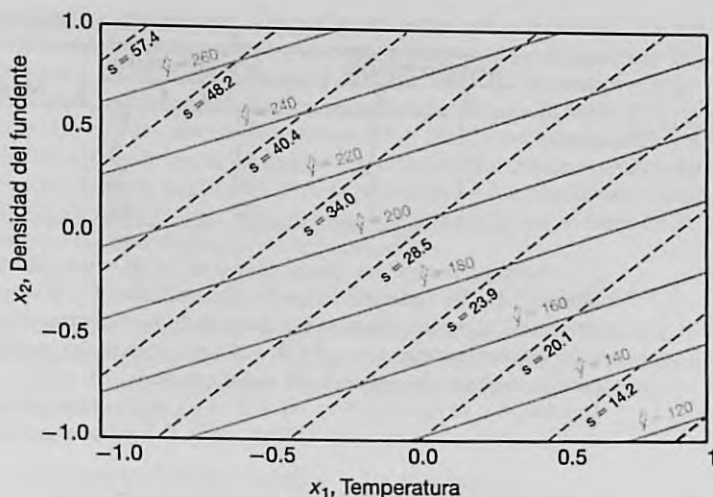


Figura 15.20: Media y desviación estándar del estudio de caso 15.4.

dos variables de control relevantes, la temperatura y la densidad del fundente, eran las óptimas para la media y la varianza del proceso. En la práctica la mayoría de las veces es necesario apelar a algún tipo de compromiso entre la media y la varianza.

El método que se ilustra en el estudio de caso 15.4 implica encontrar condiciones óptimas para el proceso cuando los datos que se utilizan provienen de un tipo de diseño experimental con arreglo de producto (o arreglo cruzado). Con frecuencia el uso de un arreglo de producto, un cruce entre dos diseños, es muy costoso. Sin embargo, el desarrollo de modelos dobles de superficie de respuesta, es decir, un modelo para la media y otro para la varianza, se puede lograr sin un arreglo de producto. El diseño que incluye tanto variables de control como de ruido se conoce como *arreglo combinado*. Este tipo de diseño y el análisis resultante se puede usar para determinar cuáles condiciones de las variables de control son las más robustas (insensibles) a la variación de las variables de ruido. Esto se puede considerar equivalente a encontrar niveles de control que minimicen la varianza del proceso producida por el movimiento de las variables de ruido.

El papel de la interacción control por ruido

La estructura de la varianza del proceso es determinada en gran medida por la naturaleza de la interacción control por ruido. La naturaleza de la falta de homogeneidad de la varianza del proceso depende de cuáles variables de control interactúan con cuáles variables de ruido. De manera específica, como se ilustrará, aquellas variables de control que interactúan con una o más variables de ruido podrían ser objeto del análisis. Por ejemplo, considere un caso citado por Myers, Montgomery y Anderson-Cook (2009), el cual incluye dos variables de control y una variable de ruido con los datos que se incluyen en la tabla 15.22. A y B son las variables de control y C es la variable de ruido.

Tabla 15.22: Datos experimentales en un arreglo cruzado

Arreglo interno		Arreglo externo		Media de respuesta
A	B	C = -1	C = +1	
-1	-1	11	15	13.0
-1	1	7	8	7.5
1	-1	10	26	18.0
1	1	10	14	12.0

Podemos ilustrar las interacciones AC y BC con gráficas, como se observa en la figura 15.21. Es necesario entender que mientras A y B se mantienen constantes en el proceso, C sigue una distribución de probabilidad durante el mismo. Dada esta información, queda claro que $A = -1$ y $B = +1$ son niveles que producen valores más pequeños para la varianza del proceso, en tanto que $A = +1$ y $B = -1$ producen valores más grandes. Así, se dice que $A = -1$ y $B = +1$ son valores robustos, es decir, insensibles a cambios inevitables en la variable de ruido C durante el proceso.

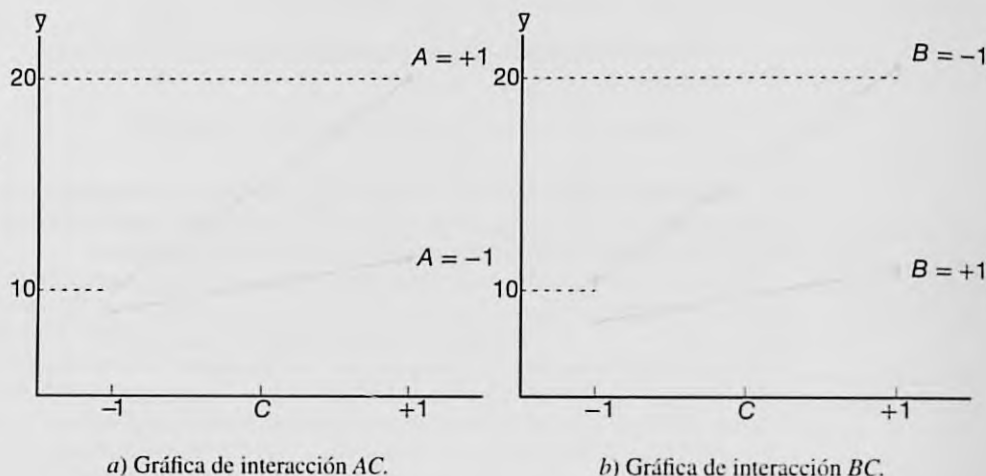


Figura 15.21: Gráficas de interacción para los datos de la tabla 15.22.

En el ejemplo anterior se dice que tanto A como B son efectos de dispersión, es decir, que ambos factores afectan la varianza del proceso. Asimismo, ambos factores son efectos de la ubicación, ya que la media de y cambia conforme los dos factores pasan de -1 a $+1$.

Análisis que incluye el modelo que contiene variables de control y de ruido

Aunque se ha hecho énfasis en que las variables de ruido no permanecen constantes durante el funcionamiento del proceso, el análisis da como resultado condiciones deseables

o incluso óptimas y proporciona de manera directa e indirecta información útil sobre el proceso. El modelo de respuesta es, en realidad, un modelo de superficie de respuesta en el vector x y en el vector z , donde x contiene variables de control y z las variables de ruido. Ciertas operaciones permiten generar modelos para la media y la varianza del proceso similares a los del estudio de caso 15.4. En Myers, Montgomery y Anderson-Cook (2009) se proporcionan los detalles; aquí se ilustrará con un ejemplo muy sencillo. Considere los datos de la tabla 15.22 de la página 650 con las variables de control A y B y la variable de ruido C . Hay ocho corridas experimentales en un factorial $2^2 \times 2$ o 2^3 . Así, podemos escribir el modelo de respuesta como

$$y(x, z) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 z + \beta_{1z} x_1 z + \beta_{2z} x_2 z + \epsilon.$$

No se incluirán las interacciones de tres factores en el modelo de regresión. A , B y C de la tabla 15.22 están representados en el modelo por medio de x_1 , x_2 y z , respectivamente. Se supone que el término del error ϵ posee las propiedades acostumbradas de independencia y varianza constante.

Las superficies de respuesta de la media y la varianza

Es más fácil comprender las superficies de respuesta de la media y la varianza del proceso si consideramos la esperanza y la varianza de z a lo largo del proceso. Se supone que la variable de ruido C [denotada por z en $y(x, z)$] es continua, con media igual a cero y varianza σ_z^2 . Los modelos de la media y la varianza del proceso se pueden considerar como

$$E_z [y(x, z)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{1z} x_1 z,$$

$$\text{Var}_z [y(x, z)] = \sigma^2 + \sigma_z^2 (\beta_3 + \beta_{1z} x_1 + \beta_{2z} x_2)^2 = \sigma^2 + \sigma_z^2 I_x^2,$$

donde I_x es la pendiente $\frac{\partial y(x, z)}{\partial z}$ en la dirección de z . Como se indicó antes, debemos observar que las interacciones de los factores A y B con la variable de ruido C son componentes fundamentales de la varianza del proceso.

Aunque ya se analizó este ejemplo por medio de las gráficas de la figura 15.21, las cuales revelan el papel de las interacciones AB y AC , es aleccionador ver el análisis considerando $E_z [y(x, z)]$ y $\text{Var}_z [y(x, z)]$. En este ejemplo el lector puede verificar fácilmente que el estimado b_{1z} para β_{1z} es $15/8$, mientras que el estimado de b_{2z} para β_{2z} es $-15/8$. El coeficiente $b_3 = 25/8$. Así, la condición $x_1 = +1$ y $x_2 = -1$ resulta en un estimado de la varianza del proceso de

$$\widehat{\text{Var}}_z [y(x, z)] = \sigma^2 + \sigma_z^2 (b_3 + b_{1z} x_1 + b_{2z} x_2)^2$$

$$= \sigma^2 + \sigma_z^2 \left[\frac{25}{8} + \left(\frac{15}{8} \right) (1) + \left(\frac{-15}{8} \right) (-1) \right]^2 = \sigma^2 + \sigma_z^2 \left(\frac{55}{8} \right)^2,$$

en tanto que para $x_1 = -1$ y $x_2 = 1$ tenemos

$$\widehat{\text{Var}}_z [y(x, z)] = \sigma^2 + \sigma_z^2 (b_3 + b_{1z} x_1 + b_{2z} x_2)^2$$

$$= \sigma^2 + \sigma_z^2 \left[\frac{25}{8} + \left(\frac{15}{8} \right) (-1) + \left(\frac{15}{8} \right) (1) \right]^2 = \sigma^2 + \sigma_z^2 \left(\frac{-5}{8} \right)^2.$$

De esta manera, para la condición más deseable (robusta) de $x_1 = -1$ y $x_2 = 1$, la varianza del proceso estimada debido a la variable de ruido C (o z) es $(25/64)\sigma_z^2$.

La condición más indeseable, la de máxima varianza del proceso, es decir, $x_1 = +1$ y $x_2 = -1$, produce una varianza del proceso estimada de $(3025/64)\sigma_z^2$. En lo que se refiere a la respuesta media, la figura 15.21 indica que si se desea una respuesta máxima, entonces $x_1 = +1$ y $x_2 = -1$ produce el mejor resultado.

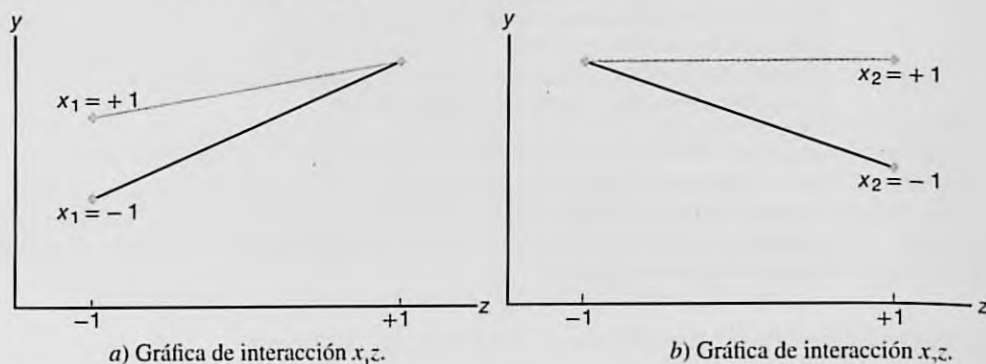


Figura 15.22: Gráficas de interacción para los datos del ejercicio 15.31.

Ejercicios

15.31 Considere un ejemplo en el que hay dos variables de control: x_1 y x_2 , y una variable de ruido z . El objetivo consiste en determinar los niveles de x_1 y x_2 , que son robustos ante los cambios de z , es decir, los niveles de x_1 y x_2 que minimizan la varianza producida en la respuesta y cuando z se mueve entre -1 a $+1$. Las variables x_1 y x_2 se encuentran a dos niveles, -1 y $+1$ en el experimento. Los datos producen las gráficas de la figura 15.22. Observe que x_1 y x_2 interactúan con la variable de ruido z . ¿Qué parámetros de x_1 y x_2 (-1 o $+1$ para cada uno) producen la varianza mínima en y ? Explique sus resultados.

15.32 Considere el siguiente factorial 2³ con variables de control x_1 y x_2 y variable de ruido z . ¿Es posible elegir x_1 y x_2 en niveles que minimicen a $\text{Var}(y)$? Explique su respuesta.

	$z = -1$		$z = +1$	
	$x_2 = -1$	$x_2 = +1$	$x_2 = -1$	$x_2 = +1$
$x_1 = -1$	4	6	8	10
$x_1 = +1$	1	3	3	5

15.33 Considere el estudio de caso 15.1 del moldeado por inyección. Suponga que es difícil controlar la temperatura de moldeado y , por lo tanto, que se puede asumir que en el proceso sigue una distribución normal con media igual a cero y varianza σ_z^2 .

El interés se centra en la varianza de la respuesta de contracción del propio proceso. Dentro del análisis de la figura 15.7 es evidente que la temperatura de moldeado, la velocidad de inyección y la interacción de ambos son los únicos factores importantes.

- ¿El parámetro de la velocidad se podría usar para crear algún tipo de control de la varianza del proceso en la contracción que surja debido a la imposibilidad de controlar la temperatura? Explique su respuesta.
- Utilice los estimados de parámetros de la figura 15.7 y proporcione un estimado de los siguientes modelos:
 - contracción media a lo largo de la distribución de la temperatura;
 - varianza de la contracción como función de σ_z^2 .
- Utilice el modelo de la varianza estimada para determinar el nivel de velocidad que minimiza la varianza de la contracción.
- Utilice el modelo de la contracción media para determinar qué valor de la velocidad minimiza la contracción media.
- ¿Los resultados anteriores son consistentes con su análisis de la gráfica de interacción de la figura 15.6? Explique su respuesta.

15.34 En el estudio de caso 15.2 acerca de los datos de la purificación de carbón se sabe que el porcentaje

de sólidos en el sistema del proceso varía de manera incontrolable durante el proceso y es considerado como un factor de ruido con media igual a 0 y varianza σ^2 . La respuesta, la eficiencia de la pureza, tiene una media y una varianza que cambian de comportamiento durante el proceso. Utilice sólo términos significativos en los siguientes incisos.

- Utilice los estimados de la figura 15.9 para desarrollar los modelos de la varianza y la eficiencia media del proceso.
- ¿Qué factor (o factores) podrían controlarse a ciertos niveles para controlar o minimizar la varianza del proceso?
- ¿Qué condiciones o factores B y C dentro de la región del diseño maximizan la media estimada?
- ¿Qué nivel de C sugeriría para minimizar la varianza del proceso cuando $B = 1$? ¿Y cuando $B = -1$?

15.35 Use los datos de purificación del carbón del ejercicio 15.2 de la página 609 para ajustar un modelo del tipo

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3,$$

Ejercicios de repaso

15.39 Se utilizó un diseño de Plackett-Burman para estudiar las propiedades reológicas de los copolímeros de alto peso molecular. En el experimento se fijaron dos niveles para cada una de seis variables. La respuesta es la viscosidad del polímero. Los datos fueron analizados en el Centro de Consultoría en Estadística de Virginia Tech, por personal del Departamento de Ingeniería Química de la universidad. Las variables son las siguientes: química del bloque duro x_1 , tasa de flujo de nitrógeno x_2 , tiempo de calentamiento x_3 , porcentaje de compresión x_4 , mediciones (alta y baja) x_5 , porcentaje de esfuerzo x_6 . A continuación se presentan los datos

Observación	x_1	x_2	x_3	x_4	x_5	x_6	y
1	1	-1	1	-1	-1	-1	194,700
2	1	1	-1	1	-1	-1	588,400
3	-1	1	1	-1	1	-1	7533
4	1	-1	1	1	-1	1	514,100
5	1	1	-1	1	1	-1	277,300
6	1	1	1	-1	1	1	493,500
7	-1	1	1	1	-1	1	8969
8	-1	-1	1	1	1	-1	18,340
9	-1	-1	-1	1	1	1	6793
10	1	-1	-1	-1	1	1	160,400
11	-1	1	-1	-1	-1	1	7008
12	-1	-1	-1	-1	-1	-1	3637

Construya una ecuación de regresión que relacione la viscosidad con los niveles de las seis variables.

donde los niveles son

- x_1 , porcentaje de sólidos: 8, 12
 x_2 , tasa de flujo: 150, 250 gal/min
 x_3 , pH: 5, 6

Centre y escale las variables a las unidades de diseño. Asimismo, realice una prueba para la falta de ajuste y haga comentarios acerca de lo adecuado del modelo de regresión lineal.

15.36 Se utiliza un plan factorial 2^5 para construir un modelo de regresión que contenga coeficientes de primer orden y términos del modelo para todas las interacciones de dos factores. Para cada factor se realizan corridas duplicadas. Construya la tabla de análisis de varianza que muestre los grados de libertad para la regresión, la falta de ajuste y el error puro.

15.37 Considere la fracción $\frac{1}{16}$ del factorial 2^7 que se estudió en la sección 15.9. Liste los 11 contrastes de definición adicionales.

15.38 Construya un diseño de Plackett-Burman para 10 variables que contengan 24 corridas experimentales.

Realice pruebas t para todos los efectos principales. Recomiende los factores a conservar para estudios futuros y los factores a eliminar. Use el cuadrado medio residual (5 grados de libertad) como medida del error experimental.

15.40 Una empresa petrolera grande del suroeste lleva a cabo experimentos de manera regular para probar los aditivos de los fluidos de perforación. La viscosidad plástica es una medición reológica que refleja el espesor del fluido. Se agregan varios polímeros al fluido para incrementar su viscosidad. A continuación se presenta un conjunto de datos en el que se usaron dos polímeros, con dos niveles cada uno, y se midió la viscosidad. La concentración de los polímeros se indica como "baja" y "alta". Haga un análisis del experimento factorial 2^2 . Pruebe los efectos e interacción de los dos polímeros.

Polímero 2	Polímero 1	
	Baja	Alta
Baja	3.0	3.5
Alta	11.7	12.0

15.41 Se analiza un experimento factorial 2^2 en el Centro de Consultoría en Estadística de Virginia Tech. El cliente es miembro del Department of Housing, Interior Design, and Resource Management y le interesa comparar hornos de arranque en frío y de

precalentamiento en términos de la energía total que se entrega al producto, y, además, comparar las condiciones de convección con el modo regular. Se hicieron cuatro corridas experimentales con cada una de las cuatro combinaciones de los factores. A continuación se presentan los datos del experimento:

	Precalentamiento		Frío	
Modo de convección	618	619.3	575	573.7
	629	611	574	572
Modo regular	581	585.7	558	562
	581	595	562	566

Haga un análisis de varianza para estudiar la interacción y los efectos principales. Saque sus conclusiones.

15.42 En el estudio "The Use of Regression Analysis for Correcting Matrix Effects in the X-Ray Fluorescence Analysis of Pyrotechnic Compositions", publicado en *Proceedings of the Tenth Conference on the Design of Experiments in Army Research Development and Testing*, ARO-D Report 65-3 (1965), se realizó un experimento donde se hicieron variar las concentraciones de cuatro componentes de una mezcla de propulsor y los pesos de partículas finas y gruesas en la suspensión. Los factores *A*, *B*, *C* y *D*, cada uno en dos niveles, representan las concentraciones de los cuatro componentes, y los factores *E* y *F*, también en dos niveles, representan los pesos de las partículas finas y gruesas que hay en la suspensión. El objetivo del análisis era determinar si las relaciones de intensidad de rayos X asociadas con el componente 1 del propulsor eran influidas en forma significativa por la variación de las concentraciones de los distintos componentes y los pesos de las partículas, según su tamaño, en la mezcla. Se utilizó una fracción de $\frac{1}{8}$ de un experimento factorial 2⁶ con los contrastes de definición *ADE*, *BCE* y *ACF*. Los datos siguientes representan el total de un par de lecturas de intensidad.

El cuadrado medio del error agrupado con 8 grados de libertad es dado por 0.02005. Analice los datos utili-

zando un nivel de significancia de 0.05 para determinar si las concentraciones de los componentes y los pesos de las partículas finas y gruesas presentes en la suspensión influyen de manera significativa en las relaciones de intensidad asociadas con el componente 1. Suponga que no existe interacción entre los seis factores.

Lote	Combinación de tratamientos	Relación total de intensidad
1	<i>abef</i>	2.2480
2	<i>cdef</i>	1.8570
3	(1)	2.2428
4	<i>ace</i>	2.3270
5	<i>bde</i>	1.8830
6	<i>abcd</i>	1.8078
7	<i>adf</i>	2.1424
8	<i>bef</i>	1.9122

15.43 Utilice la tabla 15.16 para construir un diseño de 16 corridas con 8 factores que tenga resolución IV.

15.44 En el ejercicio de repaso 15.43, compruebe que el diseño en efecto tiene resolución IV.

15.45 Construya un diseño que contenga 9 puntos de diseño, sea ortogonal, contenga un total de 12 corridas y 3 grados de libertad para el error de réplica, y también que permita hacer una prueba de falta de ajuste para la curvatura cuadrática pura.

15.46 Considere un diseño 2_{III}^{3-1} con 2 corridas centrales. Considere \bar{y} como la respuesta promedio en el parámetro de diseño y \bar{y}_0 como la respuesta promedio en el centro del diseño. Suponga que el verdadero modelo de la regresión es

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2$$

- Proporcione (y compruebe) $E(\bar{y}_i - \bar{y}_0)$.
- Explique lo que haya aprendido del resultado del inciso a

15.13 Posibles riesgos y errores conceptuales; relación con el material de otros capítulos

En el empleo de experimentos factoriales fraccionados uno de los aspectos más importantes que debe atender el analista es la *resolución del diseño*. Un diseño de resolución baja es más pequeño y, por lo tanto, menos costoso que uno de mayor resolución. Sin embargo, se paga un precio por el diseño más barato. El diseño de menor resolución tiene alias más pesados que uno de resolución mayor. Por ejemplo, si el investigador sospecha que las interacciones de dos factores son importantes, entonces no debería emplear la resolución III. Un diseño de resolución III es estrictamente un **plan de efectos principales**.

Capítulo 16

Estadística no paramétrica

16.1 Pruebas no paramétricas

La mayoría de los procedimientos de prueba de hipótesis que se presentaron en los capítulos anteriores se basan en la suposición de que las muestras aleatorias se seleccionan de poblaciones normales. Por fortuna la mayor parte de estas pruebas aún son confiables cuando existen ligeras desviaciones de la normalidad, en particular cuando el tamaño de la muestra es grande. Tradicionalmente, a tales procedimientos de prueba se les denomina **métodos paramétricos**. En este capítulo consideramos varios procedimientos de prueba alternativos, llamados **métodos no paramétricos** o **de distribución libre**, que a menudo no suponen conocimiento de ninguna clase acerca de las distribuciones de las poblaciones subyacentes, excepto, quizá, que éstas son continuas.

Los analistas de datos están usando procedimientos no paramétricos o de distribución libre cada vez con mayor frecuencia. En la ciencia y la ingeniería hay muchas aplicaciones en las que los datos no se reportan como valores de un continuo, sino, más bien, como una **escala ordinal** en la que es natural asignar rangos a los datos. De hecho, en este capítulo el lector notará muy pronto que los métodos de distribución libre aquí descritos implican un *análisis de rangos*. La mayoría de los analistas consideran que los cálculos involucrados en los métodos no paramétricos son muy atractivos e intuitivos.

Para revisar un ejemplo donde se aplica una prueba no paramétrica considere la situación en que dos jueces deben clasificar cinco marcas de cerveza de alta calidad asignando la categoría 1 a la marca que se considera que tiene la mejor calidad general, la categoría 2 a la segunda mejor, y así sucesivamente. Luego se puede utilizar una prueba no paramétrica para determinar si existe algún acuerdo entre los dos jueces.

También debemos señalar que las pruebas no paramétricas tienen asociadas varias desventajas. La primera es que no utilizan toda la información que proporciona la muestra, por lo tanto, cuando se pueden aplicar ambos métodos, estas últimas muestran ser menos eficientes que el procedimiento paramétrico correspondiente. En consecuencia, para lograr la misma potencia que la prueba paramétrica correspondiente, una prueba no paramétrica requerirá un tamaño muestral mayor que el que requeriría la primera.

Como antes indicamos, ligeras desviaciones de la normalidad dan como resultado desviaciones menores del ideal para las pruebas paramétricas estándar. Esto es particularmente cierto para la prueba t y la prueba F . En el caso de la prueba t y la prueba F , el

valor P citado podría tener un ligero error si se transgrediera moderadamente la suposición de normalidad.

En resumen, si se puede aplicar tanto una prueba paramétrica como una no paramétrica al mismo conjunto de datos, se debe aplicar la técnica paramétrica más eficiente. Sin embargo, es importante reconocer que a menudo no es posible justificar las suposiciones de normalidad, y que no siempre contamos con medidas cuantitativas. Es una ventaja que los estadísticos nos brinden diversos procedimientos no paramétricos útiles. Armado con las técnicas no paramétricas, el analista de datos tiene más herramientas para adaptar una variedad más amplia de situaciones experimentales. Se debe señalar que incluso basándose en las suposiciones de la teoría normal estándar, la eficiencia de las técnicas no paramétricas se acerca mucho más a la del procedimiento paramétrico correspondiente. Por otro lado, las grandes desviaciones de la normalidad hacen que el método no paramétrico sea mucho más eficiente que el procedimiento paramétrico.

Prueba de signo

El lector debería recordar que los procedimientos que se estudiaron en la sección 10.4 para probar la hipótesis nula de que $\mu = \mu_0$ son válidos sólo si la población es aproximadamente normal o si la muestra es grande. Sin embargo, si $n < 30$ y la población decididamente no es normal, debemos recurrir a una prueba no paramétrica.

La prueba de signo se utiliza para probar hipótesis sobre una *mediana* de la población. En el caso de muchos de los procedimientos no paramétricos, la media es reemplazada por la mediana como el **parámetro de ubicación** pertinente a probar. Recuerde que la mediana muestral se definió en la sección 1.3. El equivalente de la población, que se denota con $\tilde{\mu}$, tiene una definición análoga. Dada una variable aleatoria X , $\tilde{\mu}$ se define de modo que $P(X > \tilde{\mu}) \leq 0.5$ y $P(X < \tilde{\mu}) \leq 0.5$. En el caso continuo,

$$P(X > \tilde{\mu}) = P(X < \tilde{\mu}) = 0.5.$$

Por supuesto, si la distribución es simétrica, la media y la mediana de la población son iguales. Al probar la hipótesis nula H_0 de que $\tilde{\mu} = \tilde{\mu}_0$ en comparación con la hipótesis alternativa adecuada, con base en una muestra aleatoria de tamaño n , reemplazamos cada valor de la muestra que exceda a $\tilde{\mu}_0$ con un signo *más*, y cada valor de la muestra menor que $\tilde{\mu}_0$ con un signo *menos*. Si la hipótesis nula es verdadera y la población es simétrica, la suma de los signos más debería ser casi igual a la suma de los signos menos. Cuando un signo aparece con más frecuencia de lo que debería, con base sólo en el azar, rechazamos la hipótesis de que la mediana de la población $\tilde{\mu}$ es igual a $\tilde{\mu}_0$.

En teoría, la prueba de signo sólo se puede aplicar en situaciones en las que $\tilde{\mu}_0$ no puede ser igual al valor de cualquiera de las observaciones. Aunque la probabilidad de obtener una observación muestral exactamente igual a $\tilde{\mu}_0$ cuando la población es continua es de cero, en la práctica un valor de la muestra igual a $\tilde{\mu}_0$ ocurre con frecuencia debido a una falta de precisión en el registro de los datos. Cuando se observan valores de la muestra iguales a $\tilde{\mu}_0$, se excluyen del análisis, lo cual da como resultado que se reduzca el tamaño de la muestra.

El estadístico de prueba adecuado para la prueba de signo es la variable aleatoria binomial X , que representa el número de signos más en la muestra aleatoria. Si la hipótesis nula de que $\tilde{\mu} = \tilde{\mu}_0$ es verdadera, la probabilidad de que un valor muestral dé como resultado un signo más o uno menos es igual a $1/2$. Por lo tanto, para probar la hipótesis nula de que $\tilde{\mu} = \tilde{\mu}_0$, en realidad probamos la hipótesis nula de que el número de signos

más es un valor de una variable aleatoria que tiene una distribución binomial con el parámetro $p = 1/2$. Por lo tanto, los valores P para las alternativas unilateral y bilateral se pueden calcular usando esta distribución binomial. Por ejemplo, probando

$$H_0: \bar{\mu} = \bar{\mu}_0,$$

$$H_1: \bar{\mu} < \bar{\mu}_0,$$

se rechaza H_0 a favor de H_1 sólo si la proporción de signos más es lo suficientemente menor que $1/2$, es decir, cuando el valor x de la variable aleatoria es pequeño. Por lo tanto, si el valor P que se calcula

$$P = P(X \leq x \text{ cuando } p = 1/2)$$

es menor o igual que algún nivel de significancia α preseleccionado, se rechaza H_0 a favor de H_1 . Por ejemplo, cuando $n = 15$ y $x = 3$, en la tabla A.1 encontramos que

$$P = P(X \leq 3 \text{ cuando } p = 1/2) = \sum_{x=0}^3 b\left(x; 15, \frac{1}{2}\right) = 0.0176,$$

de manera que la hipótesis nula $\bar{\mu} = \bar{\mu}_0$ realmente se puede rechazar a un nivel de significancia de 0.05 pero no a un nivel de 0.01.

Para probar la hipótesis

$$H_0: \bar{\mu} = \bar{\mu}_0,$$

$$H_1: \bar{\mu} > \bar{\mu}_0,$$

se rechaza H_0 a favor de H_1 sólo si la proporción de signos más es suficientemente mayor que $1/2$, es decir, cuando x es grande. En consecuencia, si el valor P calculado

$$P = P(X \geq x \text{ cuando } p = 1/2)$$

es menor que α , se rechaza H_0 a favor de H_1 . Finalmente, para probar la hipótesis

$$H_0: \bar{\mu} = \bar{\mu}_0,$$

$$H_1: \bar{\mu} \neq \bar{\mu}_0,$$

se rechaza H_0 a favor de H_1 cuando la proporción de signos más es significativamente menor o mayor que $1/2$. Esto, por supuesto, es equivalente a que x sea tan pequeña o tan grande como se requiere. Por lo tanto, si $x < n/2$ y el valor P calculado

$$P = 2P(X \leq x \text{ cuando } p = 1/2)$$

es menor o igual que α , o si $x > n/2$ y el valor P calculado

$$P = 2P(X \geq x \text{ cuando } p = 1/2)$$

es menor o igual que α , se rechaza H_0 a favor de H_1 .

Siempre que $n > 10$, las probabilidades binomiales con $p = 1/2$ se pueden aproximar a partir de la curva normal, ya que $np = nq > 5$. Suponga, por ejemplo, que deseamos probar la hipótesis

$$H_0: \bar{\mu} = \bar{\mu}_0,$$

$$H_1: \bar{\mu} < \bar{\mu}_0,$$

a un nivel de significancia $\alpha = 0.05$ para una muestra aleatoria de tamaño $n = 20$ que produce $x = 6$ signos más. Si utilizamos la aproximación de la curva normal con

$$\bar{\mu} = np = (20)(0.5) = 10$$

y

$$\sigma = \sqrt{npq} = \sqrt{(20)(0.5)(0.5)} = 2.236,$$

encontramos que

$$z = \frac{6.5 - 10}{2.236} = -1.57.$$

Por lo tanto,

$$P = P(X \leq 6) \approx P(Z < -1.57) = 0.0582,$$

que conduce a no rechazar la hipótesis nula.

Ejemplo 16.1: Los siguientes datos representan el número de horas que funciona una desbrozadora antes de requerir una recarga:

1.5, 2.2, 0.9, 1.3, 2.0, 1.6, 1.8, 1.5, 2.0, 1.2, 1.7.

A un nivel de significancia de 0.05 utilice la prueba de signo para probar la hipótesis de que esta desbrozadora específica funciona con una mediana de 1.8 horas antes de requerir una recarga.

- Solución:**
1. $H_0: \bar{\mu} = 1.8$.
 2. $H_1: \bar{\mu} \neq 1.8$.
 3. $\alpha = 0.05$.
 4. Estadístico de prueba: variable binomial X con $p = \frac{1}{2}$.
 5. Cálculos: Al reemplazar cada valor con el símbolo "+" si excede 1.8, con el símbolo "-" si es menor que 1.8 y descartar las mediciones que sean iguales a 1.8, obtenemos la siguiente secuencia

- + - - + - - + - -

para la cual $n = 10$, $x = 3$ y $n/2 = 5$. Por lo tanto, el valor P que se obtiene de la tabla A.1 es

$$P = 2P\left(X \leq 3 \text{ cuando } p = \frac{1}{2}\right) = 2 \sum_{x=0}^3 b\left(x; 10, \frac{1}{2}\right) = 0.3438 > 0.05.$$

6. Decisión: No se rechaza la hipótesis nula y se concluye que la mediana del tiempo de funcionamiento no difiere significativamente de 1.8 horas. \blacksquare

También se puede utilizar la prueba de signo para probar la hipótesis nula $\bar{\mu}_1 - \bar{\mu}_2 = d_0$ para observaciones de pares. Aquí se reemplaza cada diferencia, d_i , con un signo más o un signo menos, dependiendo de si la diferencia ajustada, $d_i - d_0$, es positiva o negativa. A lo largo de esta sección hemos asumido que las poblaciones son simétricas. No obstante, aun si las poblaciones fueran asimétricas, podríamos llevar a cabo el mismo procedimiento de prueba, pero las hipótesis se referirían a las medianas de la población en vez de a las medias.

Ejemplo 16.2: Una empresa de taxis intenta decidir si utilizar neumáticos radiales en vez de neumáticos regulares con cinturón le serviría para ahorrar combustible. Se equipan 16 automóviles con neumáticos radiales y se conducen por un recorrido de prueba establecido. Después se equipan los mismos automóviles con los neumáticos regulares con cinturón y se hace que los mismos conductores vuelvan a realizar el recorrido de prueba. El consumo de gasolina, en kilómetros por litro, se presenta en la tabla 16.1.

¿Podemos concluir a un nivel de significancia de 0.05 que los automóviles equipados con neumáticos radiales ahorran más combustible que los equipados con neumáticos regulares con cinturón?

Tabla 16.1: Datos para el ejemplo 16.2

Automóvil	1	2	3	4	5	6	7	8
Neumáticos radiales	4.2	4.7	6.6	7.0	6.7	4.5	5.7	6.0
Neumáticos con cinturón	4.1	4.9	6.2	6.9	6.8	4.4	5.7	5.8
Automóvil	9	10	11	12	13	14	15	16
Neumáticos radiales	7.4	4.9	6.1	5.2	5.7	6.9	6.8	4.9
Neumáticos con cinturón	6.9	4.9	6.0	4.9	5.3	6.5	7.1	4.8

Solución: Sean $\bar{\mu}_1$ y $\bar{\mu}_2$ la mediana de los kilómetros por litro para los automóviles equipados con neumáticos radiales y con cinturón, respectivamente.

1. $H_0: \bar{\mu}_1 - \bar{\mu}_2 = 0$.
2. $H_1: \bar{\mu}_1 - \bar{\mu}_2 > 0$.
3. $\alpha = 0.05$.
4. Estadístico de prueba: variable binomial X con $p = 1/2$.
5. Cálculos: después de reemplazar cada diferencia positiva con un símbolo "+" y cada diferencia negativa con un símbolo "-", y después de descartar las dos diferencias de cero, obtenemos la secuencia

+ - + + - + + + + + + + - +

para la que $n = 14$ y $x = 11$. Si usamos la aproximación de la curva normal, encontramos que

$$z = \frac{10.5 - 7}{\sqrt{(14)(0.5)(0.5)}} = 1.87,$$

y entonces

$$P = P(X \geq 11) \approx P(Z > 1.87) = 0.0307.$$

6. Decisión: Se rechaza H_0 y se concluye que, en promedio, los neumáticos radiales ahorran más combustible.

La prueba de signo no sólo es uno de los procedimientos no paramétricos más fáciles de aplicar, sino que tiene la ventaja adicional de poder aplicarse a datos dicotómicos que no se pueden registrar en una escala numérica, pero que se pueden representar mediante respuestas positivas y negativas. Por ejemplo, la prueba de signo se aplica en experimentos donde se registra una respuesta cualitativa como “éxito” o “fracaso”; y en experimentos de tipo sensorial donde se registra un signo más o un signo menos, dependiendo de si el catador del sabor identifica de manera correcta o incorrecta el ingrediente que se desea.

Intentaremos hacer comparaciones entre varios de los procedimientos no paramétricos y las pruebas paramétricas correspondientes. En el caso de la prueba de signo la competencia es, desde luego, la prueba t . Si se toman muestras de una distribución normal, al utilizar la prueba t se obtendrá como resultado la potencia más grande de la prueba. Si la distribución sólo es simétrica, aunque no sea normal, en términos de potencia se prefiere la prueba t , a menos que la distribución tenga “colas muy pesadas” en comparación con la distribución normal.

16.2 Prueba de rango con signo

El lector debe notar que la prueba de signo sólo utiliza los signos más y menos de las diferencias entre las observaciones y $\tilde{\mu}_0$ en el caso de una muestra, o los signos más y menos de las diferencias entre los pares de observaciones en el caso de muestras en pares; no se toma en cuenta la magnitud de esas diferencias. Una prueba que utiliza dirección y magnitud, propuesta en 1945 por Frank Wilcoxon, ahora se conoce comúnmente como **prueba de rango con signo de Wilcoxon**.

El analista puede extraer más información de los datos de manera no paramétrica si es razonable aplicar una restricción adicional a la distribución de la que se toman los datos. La prueba de rango con signo de Wilcoxon se aplica en el caso de una **distribución continua simétrica**. En esta condición se prueba la hipótesis nula $\tilde{\mu} = \tilde{\mu}_0$. Primero restamos $\tilde{\mu}_0$ de cada valor muestral y descartamos todas las diferencias iguales a cero. Las diferencias restantes se ordenan sin importar el signo. Se asigna una categoría de 1 a la diferencia absoluta más pequeña, es decir, sin signo, una categoría de 2 a la siguiente más pequeña, y así sucesivamente. Cuando el valor absoluto de dos o más diferencias es el mismo, se asigna a cada uno el promedio de los rangos que se asignarían si las diferencias fueran distinguibles. Por ejemplo, si la quinta y la sexta diferencias más pequeñas tienen el mismo valor absoluto, a cada una se le asignaría una categoría de 5.5. Si la hipótesis $\tilde{\mu} = \tilde{\mu}_0$ es verdadera, el total de los rangos que corresponden a las diferencias positivas debería ser casi igual al total de los rangos que corresponden a las diferencias negativas. Representemos estos totales con w_+ y w_- , respectivamente. Designamos el más pequeño de w_+ y w_- con w .

Al seleccionar muestras repetidas esperaríamos que w_+ y w_- y, por lo tanto, w varíara. De esta manera, consideramos a w_+ , w_- y w como valores de las correspondientes variables aleatorias W_+ , W_- y W . La hipótesis nula $\tilde{\mu} = \tilde{\mu}_0$ se puede rechazar a favor de la hipótesis alternativa $\tilde{\mu} < \tilde{\mu}_0$ sólo si w_+ es pequeña y w_- es grande. De igual manera, la hipótesis alternativa $\tilde{\mu} > \tilde{\mu}_0$ se puede aceptar sólo si w_+ es grande y w_- es pequeña. Para una alternativa bilateral se puede rechazar H_0 a favor de H_1 si w_+ o w_- y, en consecuencia, w son suficientemente pequeñas. Por lo tanto, no importa cuál sea la hipótesis alternativa,

cuando el valor del estadístico adecuado W_+ , W_- o W es suficientemente pequeño, se rechaza la hipótesis nula.

Dos muestras con observaciones en pares

Con el fin de probar la hipótesis nula de que se toman muestras de dos poblaciones simétricas continuas con $\tilde{\mu}_1 = \tilde{\mu}_2$ para el caso de muestras en pares, se ordenan las diferencias de las observaciones en pares sin importar el signo y se procede como en el caso de una sola muestra. Los diversos procedimientos de prueba para los casos de una sola muestra y de muestras en pares se resumen en la tabla 16.2.

Tabla 16.2: Prueba de rango con signo

| H_0 | H_1 | Calcular |
|---------------------------------|------------------------------------|----------|
| $\tilde{\mu} = \tilde{\mu}_0$ | $\tilde{\mu} < \tilde{\mu}_0$ | w_+ |
| | $\tilde{\mu} > \tilde{\mu}_0$ | w_- |
| | $\tilde{\mu} \neq \tilde{\mu}_0$ | w |
| $\tilde{\mu}_1 = \tilde{\mu}_2$ | $\tilde{\mu}_1 < \tilde{\mu}_2$ | w_+ |
| | $\tilde{\mu}_1 > \tilde{\mu}_2$ | w_- |
| | $\tilde{\mu}_1 \neq \tilde{\mu}_2$ | w |

No es difícil mostrar que siempre que $n < 5$ y que el nivel de significancia no exceda a 0.05 para una prueba de una cola, o a 0.10 para una prueba de dos colas, todos los valores posibles de w_+ , w_- o w conducirán a la aceptación de la hipótesis nula. Sin embargo, cuando $5 \leq n \leq 30$, la tabla A.16 muestra valores críticos aproximados de W_+ y W_- a niveles de significancia iguales a 0.01, 0.025 y 0.05 para una prueba de una cola y valores críticos de W a niveles de significancia iguales a 0.02, 0.05 y 0.10 para una prueba de dos colas. Se rechaza la hipótesis nula si el valor calculado w_+ , w_- o w es **menor o igual que** el valor tabulado apropiado. Por ejemplo, cuando $n = 12$, la tabla A.16 indica que se requiere un valor de $w_+ \leq 17$ para que la hipótesis alternativa unilateral $\tilde{\mu} < \tilde{\mu}_0$ sea significativa al nivel 0.05.

Ejemplo 16.3: Repita el ejemplo 16.1 usando la prueba de rango con signo.

- Solución:**
- $H_0: \tilde{\mu} = 1.8$.
 - $H_1: \tilde{\mu} \neq 1.8$.
 - $\alpha = 0.05$.
 - Región crítica: Como $n = 10$, después de descartar la medida que es igual a 1.8, la tabla A.16 indica que la región crítica es $w \leq 8$.
 - Cálculos: Al restar 1.8 a cada medida y después ordenar las diferencias sin hacer caso del signo, tenemos

| | | | | | | | | | | |
|--------|------|-----|------|------|-----|------|------|-----|------|------|
| d_i | -0.3 | 0.4 | -0.9 | -0.5 | 0.2 | -0.2 | -0.3 | 0.2 | -0.6 | -0.1 |
| Rangos | 5.5 | 7 | 10 | 8 | 3 | 3 | 5.5 | 3 | 9 | 1 |

Ahora bien, $w_+ = 13$ y $w_- = 42$, de manera que $w = 13$, el menor de w_+ y w_- .

6. Decisión: Como antes, no se rechaza H_0 y se concluye que la mediana del tiempo de operación no difiere significativamente de 1.8 horas. \square

La prueba de rango con signo también se puede utilizar para probar la hipótesis nula de que $\tilde{\mu}_1 - \tilde{\mu}_2 = d_0$. En este caso las poblaciones no necesitan ser simétricas. Como ocurre con la prueba de signo, restamos d_0 de cada diferencia, ordenamos las diferencias ajustadas sin importar el signo y aplicamos el mismo procedimiento anterior.

Ejemplo 16.4: Se afirma que, si se le proporcionan ejemplos de problemas con antelación, un estudiante universitario de último año puede aumentar en al menos 50 puntos su calificación en el área de especialidad del examen para ingresar a posgrado. Para probar esta afirmación se divide a un grupo de 20 estudiantes del último año en 10 pares, de manera que cada par tenga casi la misma calificación promedio durante sus 3 primeros años en la universidad. Los ejemplos de problemas y las respuestas se proporcionan al azar a un miembro de cada par una semana antes del examen. Las calificaciones del examen se presentan en la tabla 16.3.

Tabla 16.3: Datos para el ejemplo 16.4

| | Par | | | | | | | | | |
|---------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Con ejemplos de problemas | 531 | 621 | 663 | 579 | 451 | 660 | 591 | 719 | 543 | 575 |
| Sin ejemplos de problemas | 509 | 540 | 688 | 502 | 424 | 683 | 568 | 748 | 530 | 524 |

A un nivel de significancia de 0.05 pruebe la hipótesis nula de que los ejemplos de problemas aumentan las calificaciones en 50 puntos, en comparación con la hipótesis alternativa de que aumentan menos de 50 puntos.

Solución: Representemos con $\tilde{\mu}_1$ y $\tilde{\mu}_2$ la mediana de las calificaciones de todos los estudiantes que resuelven el examen en cuestión con y sin ejemplos de problemas, respectivamente.

1. $H_0: \tilde{\mu}_1 - \tilde{\mu}_2 = 50$.
2. $H_1: \tilde{\mu}_1 - \tilde{\mu}_2 < 50$.
3. $\alpha = 0.05$.
4. Región crítica: Como $n = 10$, la tabla A.16 indica que la región crítica es $w_+ \leq 11$.
5. Cálculos:

| | Par | | | | | | | | | |
|-------------|-----|----|-----|-----|-----|-----|-----|-----|-----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| d_i | 22 | 81 | -25 | 77 | 27 | -23 | 23 | -29 | 13 | 51 |
| $d_i - d_0$ | -28 | 31 | -75 | 27 | -23 | -73 | -27 | -79 | -37 | 1 |
| Rangos | 5 | 6 | 9 | 3.5 | 2 | 8 | 3.5 | 10 | 7 | 1 |

Se obtiene que $w_+ = 6 + 3.5 + 1 = 10.5$.

6. Decisión: Rechazar H_0 y concluir que los ejemplos de problemas, en promedio, no aumentan las calificaciones del examen para ingresar a posgrado hasta en 50 puntos. \square

Aproximación normal para muestras grandes

Cuando $n \geq 15$ la distribución muestral de W_+ (o W_-) se aproxima a la distribución normal, con media y varianza dadas por

$$\mu_{W_+} = \frac{n(n+1)}{4} \text{ y } \sigma_{W_+}^2 = \frac{n(n+1)(2n+1)}{24}$$

Por lo tanto, cuando n excede al valor más grande en la tabla A.16 se utiliza el estadístico

$$Z = \frac{W_+ - \mu_{W_+}}{\sigma_{W_+}}$$

para determinar la región crítica para la prueba.

Ejercicios

16.1 Los siguientes datos representan el tiempo, en minutos, que un paciente tiene que esperar durante 12 visitas al consultorio de un médico antes de ser atendido:

17 15 20 20 32 28
12 26 25 25 35 24

Utilice la prueba de signo a un nivel de significancia de 0.05 para probar la afirmación del médico de que la mediana del tiempo de espera de sus pacientes no es mayor de 20 minutos.

16.2 Los siguientes datos representan el número de horas de vuelo de entrenamiento que 18 estudiantes de piloto reciben de cierto instructor antes de su primer vuelo solos:

9 12 18 14 12 14 12 10 16
11 9 11 13 11 13 15 13 14

Con las probabilidades binomiales de la tabla A.1 realice una prueba de signo a un nivel de significancia de 0.02 para probar la afirmación del instructor de que la mediana del tiempo de vuelo de entrenamiento que sus estudiantes requieren antes de volar solos es de 12 horas.

16.3 Un inspector de alimentos examina 16 latas de cierta marca de jamón para determinar el porcentaje de impurezas externas. Se registraron los siguientes datos:

2.4 2.3 3.1 2.2 2.3 1.2 1.0 2.4
1.7 1.1 4.2 1.9 1.7 3.6 1.6 2.3

Utilice una aproximación normal a la distribución binomial y realice una prueba de signo a un nivel de significancia de 0.05 para probar la hipótesis nula de que la mediana del porcentaje de impurezas en esta marca de jamón es de 2.5%, en comparación con la hipótesis alternativa de que la mediana del porcentaje de impurezas no es de 2.5%.

16.4 Un proveedor de pintura acrílica afirma que un nuevo aditivo reducirá el tiempo de secado de su pintura. Para probar esta afirmación se pintaron 12 paneles de madera; la mitad de cada panel se pintó con la pintura que contiene el aditivo regular y la otra mitad con

la pintura que contiene el nuevo aditivo. Los tiempos de secado, en horas, son los siguientes:

| Panel | Tiempo de secado (horas) | |
|-------|--------------------------|-----------------|
| | Aditivo nuevo | Aditivo regular |
| 1 | 6.4 | 6.6 |
| 2 | 5.8 | 5.8 |
| 3 | 7.4 | 7.8 |
| 4 | 5.5 | 5.7 |
| 5 | 6.3 | 6.0 |
| 6 | 7.8 | 8.4 |
| 7 | 8.6 | 8.8 |
| 8 | 8.2 | 8.4 |
| 9 | 7.0 | 7.3 |
| 10 | 4.9 | 5.8 |
| 11 | 5.9 | 5.8 |
| 12 | 6.5 | 6.5 |

Utilice la prueba de signo a un nivel de 0.05 para probar la hipótesis nula de que el nuevo aditivo no disminuye el tiempo que tarda en secar la pintura con el aditivo regular.

16.5 Se afirma que una nueva dieta reducirá el peso de una persona en 4.5 kilogramos, en promedio, en un periodo de dos semanas. Se registran los pesos de 10 mujeres que siguen esta dieta, antes y después de un periodo de dos semanas, y se obtienen los siguientes datos:

| Mujer | Peso antes | Peso después |
|-------|------------|--------------|
| 1 | 58.5 | 60.0 |
| 2 | 60.3 | 54.9 |
| 3 | 61.7 | 58.1 |
| 4 | 69.0 | 62.1 |
| 5 | 64.0 | 58.5 |
| 6 | 62.6 | 59.9 |
| 7 | 56.7 | 54.4 |
| 8 | 63.6 | 60.2 |
| 9 | 68.2 | 62.3 |
| 10 | 59.4 | 58.7 |

Utilice la prueba de signo a un nivel de significancia de 0.05 para probar la hipótesis de que la dieta reduce la mediana del peso en 4.5 kilogramos, en comparación con la hipótesis alternativa de que la mediana de la pérdida de peso es menor que 4.5 kilogramos.

16.6 En un experimento de contaminación atmosférica se comparan dos tipos de instrumentos para medir la cantidad de monóxido de azufre en la atmósfera. Se registraron las siguientes lecturas diarias durante dos semanas:

| Día | Monóxido de azufre | |
|-----|--------------------|---------------|
| | Instrumento A | Instrumento B |
| 1 | 0.96 | 0.87 |
| 2 | 0.82 | 0.74 |
| 3 | 0.75 | 0.63 |
| 4 | 0.61 | 0.55 |
| 5 | 0.89 | 0.76 |
| 6 | 0.64 | 0.70 |
| 7 | 0.81 | 0.69 |
| 8 | 0.68 | 0.57 |
| 9 | 0.65 | 0.53 |
| 10 | 0.84 | 0.88 |
| 11 | 0.59 | 0.51 |
| 12 | 0.94 | 0.79 |
| 13 | 0.91 | 0.84 |
| 14 | 0.77 | 0.63 |

Utilice la aproximación normal a la distribución binomial y realice una prueba de signo para determinar si los diferentes instrumentos conducen a diferentes resultados. Utilice un nivel de significancia de 0.05.

16.7 Las siguientes cifras indican la presión sanguínea sistólica de 16 corredores antes y después de una carrera de ocho kilómetros:

| Corredor | Antes | Después |
|----------|-------|---------|
| 1 | 158 | 164 |
| 2 | 149 | 158 |
| 3 | 160 | 163 |
| 4 | 155 | 160 |
| 5 | 164 | 172 |
| 6 | 138 | 147 |
| 7 | 163 | 167 |
| 8 | 159 | 169 |
| 9 | 165 | 173 |
| 10 | 145 | 147 |
| 11 | 150 | 156 |
| 12 | 161 | 164 |
| 13 | 132 | 133 |
| 14 | 155 | 161 |
| 15 | 146 | 154 |
| 16 | 159 | 170 |

Utilice una prueba de signo a un nivel de significancia de 0.05 para probar la hipótesis nula de que correr ocho kilómetros aumenta la mediana de la presión sanguínea sistólica en ocho puntos, en comparación con la hipóte-

sis alternativa de que el aumento en la mediana es menor que ocho puntos.

16.8 Analice los datos del ejercicio 16.1 usando la prueba de rango con signo.

16.9 Analice los datos del ejercicio 16.2 usando la prueba de rango con signo.

16.10 Los pesos de 5 personas, en kilogramos, antes de dejar de fumar y cinco semanas después de dejar de fumar, son los siguientes:

| | Individuo | | | | |
|---------|-----------|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 |
| Antes | 66 | 80 | 69 | 52 | 75 |
| Después | 71 | 82 | 68 | 56 | 73 |

Utilice la prueba de rango con signo para observaciones en pares y pruebe la hipótesis, a un nivel de significancia de 0.05, de que dejar de fumar no influye en el peso de una persona, en comparación con la hipótesis alternativa de que al dejar de fumar se aumenta de peso.

16.11 Repita el ejercicio 16.5 usando la prueba de rango con signo.

16.12 Los siguientes son los números de recetas surtidas por dos farmacias en un periodo de 20 días:

| Día | Farmacia A | Farmacia B |
|-----|------------|------------|
| 1 | 19 | 17 |
| 2 | 21 | 15 |
| 3 | 15 | 12 |
| 4 | 17 | 12 |
| 5 | 24 | 16 |
| 6 | 12 | 15 |
| 7 | 19 | 11 |
| 8 | 14 | 13 |
| 9 | 20 | 14 |
| 10 | 18 | 21 |
| 11 | 23 | 19 |
| 12 | 21 | 15 |
| 13 | 17 | 11 |
| 14 | 12 | 10 |
| 15 | 16 | 20 |
| 16 | 15 | 12 |
| 17 | 20 | 13 |
| 18 | 18 | 17 |
| 19 | 14 | 16 |
| 20 | 22 | 18 |

A un nivel de significancia de 0.01 utilice la prueba de rango con signo para determinar si las dos farmacias surten el mismo número de recetas, "en promedio", en comparación con la hipótesis alternativa de que la farmacia A surte más recetas que la farmacia B.

16.13 Repita el ejercicio 16.7 usando la prueba de rango con signo.

16.14 Repita el ejercicio 16.6 con la prueba de rango con signo.

16.3 Prueba de la suma de rangos de Wilcoxon

Como antes indicamos, el procedimiento no paramétrico por lo general es una alternativa adecuada para la prueba de la teoría normal cuando la suposición de normalidad no es válida. Cuando nos interesa probar la igualdad de las medias de dos distribuciones continuas que evidentemente no son normales, y las muestras son independientes, es decir, que no hay emparejamiento de observaciones, la **prueba de la suma de rangos de Wilcoxon** o la **prueba de dos muestras de Wilcoxon** es una alternativa apropiada a la prueba t de dos muestras que se describe en el capítulo 10.

Probaremos la hipótesis nula H_0 de que $\bar{\mu}_1 = \bar{\mu}_2$ en comparación con alguna hipótesis alternativa adecuada. Primero seleccionamos una muestra aleatoria de cada una de las poblaciones. Sea n_1 el número de observaciones en la muestra más pequeña y n_2 el número de observaciones en la muestra más grande. Cuando las muestras son de igual tamaño n_1 y n_2 se pueden asignar de manera aleatoria. Se ordenan las $n_1 + n_2$ observaciones de las muestras combinadas en orden ascendente y se sustituye un rango de 1, 2, ..., $n_1 + n_2$ para cada observación. En el caso de empates (observaciones idénticas), se reemplazan las observaciones por la media de los rangos que tendrían las observaciones si fueran distinguibles. Por ejemplo, si la séptima y octava observaciones fueran idénticas, se asignaría un rango de 7.5 a cada una de las dos observaciones.

La suma de los rangos que corresponden a las n_1 observaciones en la muestra más pequeña se denota con w_1 . De manera similar, el valor w_2 representa la suma de los n_2 rangos que corresponden a la muestra más grande. El total $w_1 + w_2$ depende sólo del número de observaciones en las dos muestras y de ninguna manera resulta afectado por los resultados del experimento. Por lo tanto, si $n_1 = 3$ y $n_2 = 4$, entonces $w_1 + w_2 = 1 + 2 + \dots + 7 = 28$, sin importar los valores numéricos de las observaciones. En general,

$$w_1 + w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2},$$

la suma aritmética de los enteros 1, 2, ..., $n_1 + n_2$. Una vez que se determina w_1 , es más fácil calcular w_2 mediante la fórmula

$$w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - w_1.$$

Al elegir muestras repetidas de tamaños n_1 y n_2 esperaríamos que w_1 y, por lo tanto, w_2 , varíen. Así, podríamos considerar a w_1 y w_2 como valores de las variables aleatorias W_1 y W_2 , respectivamente. La hipótesis nula $\bar{\mu}_1 = \bar{\mu}_2$ se rechazará a favor de la hipótesis alternativa $\bar{\mu}_1 < \bar{\mu}_2$ sólo si w_1 es pequeña y w_2 es grande. De igual manera, la hipótesis alternativa $\bar{\mu}_1 > \bar{\mu}_2$ se puede aceptar sólo si w_1 es grande y w_2 es pequeña. Para una prueba de dos colas podemos rechazar H_0 a favor de H_1 si w_1 es pequeña y w_2 es grande, o si w_1 es grande y w_2 es pequeña. En otras palabras, se acepta la hipótesis alternativa $\bar{\mu}_1 < \bar{\mu}_2$ si w_1 es suficientemente pequeña; la hipótesis alternativa $\bar{\mu}_1 > \bar{\mu}_2$ se acepta si w_2 es suficientemente pequeña; y la hipótesis alternativa $\bar{\mu}_1 \neq \bar{\mu}_2$ se acepta si el mínimo de w_1 y w_2 es tan pequeño como se requiere. En la práctica real por lo general basamos nuestra decisión en el valor

$$u_1 = w_1 - \frac{n_1(n_1 + 1)}{2} \quad \text{o} \quad u_2 = w_2 - \frac{n_2(n_2 + 1)}{2}$$

del estadístico relacionado U_1 o U_2 , o en el valor u del estadístico U , el mínimo de U_1 y U_2 . Dichos estadísticos simplifican la construcción de tablas de valores críticos, dado

que U_1 y U_2 tienen distribuciones muestrales simétricas y toman valores en el intervalo de 0 a $n_1 n_2$, tales que $u_1 + u_2 = n_1 n_2$.

De las fórmulas para u_1 y u_2 vemos que u_1 será pequeña cuando w_1 es pequeña, y u_2 será pequeña cuando w_2 sea pequeña. En consecuencia, la hipótesis nula se rechazará siempre que los estadísticos apropiados U_1 , U_2 o U tomen un valor menor o igual que el valor crítico deseado dado en la tabla A.17. Los diversos procedimientos de prueba se resumen en la tabla 16.4.

Tabla 16.4: Prueba de la suma de rangos

| H_0 | H_1 | Calcular |
|---------------------------------|------------------------------------|----------|
| $\tilde{\mu}_1 = \tilde{\mu}_2$ | $\tilde{\mu}_1 < \tilde{\mu}_2$ | u_1 |
| | $\tilde{\mu}_1 > \tilde{\mu}_2$ | u_2 |
| | $\tilde{\mu}_1 \neq \tilde{\mu}_2$ | u |

La tabla A.17 proporciona valores críticos de U_1 y U_2 para niveles de significancia iguales a 0.001, 0.01, 0.025 y 0.05 para una prueba de una cola, y valores críticos de U para niveles de significancia iguales a 0.002, 0.02, 0.05 y 0.10 para una prueba de dos colas. Si el valor observado de u_1 , u_2 o u es **menor o igual que** el valor crítico tabulado, se rechaza la hipótesis nula al nivel de significancia que se indica en la tabla. Suponga, por ejemplo, que deseamos probar la hipótesis nula de que $\tilde{\mu}_1 = \tilde{\mu}_2$ en comparación con la hipótesis alternativa unilateral de que $\tilde{\mu}_1 < \tilde{\mu}_2$, a un nivel de significancia de 0.05 para muestras aleatorias de tamaños $n_1 = 3$ y $n_2 = 5$, que producen el valor $w_1 = 8$. Se sigue que

$$u_1 = 8 - \frac{(3)(4)}{2} = 2.$$

Nuestra prueba de una sola cola se basa en el estadístico U_1 . Si se usa la tabla A.17, se rechaza la hipótesis nula de medias iguales cuando $u_1 \leq 1$. Como $u_1 = 2$ no cae en la región de rechazo, no se puede rechazar la hipótesis nula.

Ejemplo 16.5: Se encontró que el contenido de nicotina de dos marcas de cigarrillos, medido en miligramos, es el siguiente:

| | | | | | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Marca A | 2.1 | 4.0 | 6.3 | 5.4 | 4.8 | 3.7 | 6.1 | 3.3 | | |
| Marca B | 4.1 | 0.6 | 3.1 | 2.5 | 4.0 | 6.2 | 1.6 | 2.2 | 1.9 | 5.4 |

A un nivel de significancia de 0.05 pruebe la hipótesis de que las medianas del contenido de nicotina de las dos marcas son iguales, en comparación con la hipótesis alternativa de que son diferentes.

- Solución:**
1. $H_0: \tilde{\mu}_1 = \tilde{\mu}_2$.
 2. $H_1: \tilde{\mu}_1 \neq \tilde{\mu}_2$.
 3. $\alpha = 0.05$.
 4. Región crítica: $u \leq 17$ (de la tabla A.17).
 5. Cálculos: Las observaciones se acomodan en orden ascendente y se les asignan rangos del 1 al 18.

| Datos originales | Rangos | Datos originales | Rangos |
|------------------|--------|------------------|--------|
| 0.6 | 1 | 4.0 | 10.5* |
| 1.6 | 2 | 4.0 | 10.5 |
| 1.9 | 3 | 4.1 | 12 |
| 2.1 | 4* | 4.8 | 13* |
| 2.2 | 5 | 5.4 | 14.5* |
| 2.5 | 6 | 5.4 | 14.5 |
| 3.1 | 7 | 6.1 | 16* |
| 3.3 | 8* | 6.2 | 17 |
| 3.7 | 9* | 6.3 | 18* |

*Los rangos marcados con asterisco pertenecen a la muestra A.

Ahora

$$w_1 = 4 + 8 + 9 + 10.5 + 13 + 14.5 + 16 + 18 = 93$$

y

$$w_2 = \frac{(18)(19)}{2} - 93 = 78.$$

Por lo tanto,

$$u_1 = 93 - \frac{(8)(9)}{2} = 57, \quad u_2 = 78 - \frac{(10)(11)}{2} = 23.$$

6. Decisión: no se rechaza la hipótesis nula H_0 y se concluye que no hay diferencia significativa en las medianas del contenido de nicotina en las dos marcas de cigarrillos. ▮

Teoría normal de aproximación para dos muestras

Cuando n_1 y n_2 exceden a 8, la distribución muestral de U_1 (o U_2) se aproxima a la distribución normal con media y varianza dadas por

$$\mu_{U_1} = \frac{n_1 n_2}{2} \quad \text{y} \quad \sigma_{U_1}^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}.$$

En consecuencia, cuando n_2 es mayor que 20, el valor máximo en la tabla A.17, y n_1 es al menos 9, se puede utilizar el estadístico

$$Z = \frac{U_1 - \mu_{U_1}}{\sigma_{U_1}}$$

para la prueba, con la región crítica que cae ya sea en alguna o en ambas colas de la distribución normal estándar, dependiendo de la forma de H_1 .

El uso de la prueba de suma de rangos de Wilcoxon no se restringe a poblaciones no normales. Se puede utilizar en vez de la prueba t de dos muestras cuando las poblaciones son normales, aunque la potencia será menor. La prueba de suma de rangos de Wilcoxon siempre es superior a la prueba t para poblaciones definitivamente no normales.

16.4 Prueba de Kruskal-Wallis

En los capítulos 13, 14 y 15 la técnica del análisis de varianza resalta como técnica analítica para probar la igualdad de $k \geq 2$ medias de la población. Sin embargo, el lector debería recordar que para que la prueba F sea teóricamente correcta se debe suponer normalidad. En esta sección investigamos una alternativa no paramétrica al análisis de varianza.

La **prueba de Kruskal-Wallis**, también llamada **prueba H de Kruskal-Wallis**, es una generalización de la prueba de la suma de rangos para el caso de $k > 2$ muestras. Se utiliza para probar la hipótesis nula H_0 de que k muestras independientes provienen de poblaciones idénticas. Presentada en 1952 por W. H. Kruskal y W. A. Wallis, la prueba constituye un procedimiento no paramétrico para probar la igualdad de las medias, en el análisis de varianza de un factor, cuando el experimentador desea evitar la suposición de que las muestras se seleccionaron de poblaciones normales.

Sea n_i ($i = 1, 2, \dots, k$) el número de observaciones en la i -ésima muestra. Primero combinamos todas las k muestras y acomodamos las $n = n_1 + n_2 + \dots + n_k$ observaciones en orden ascendente, y sustituimos el rango apropiado de $1, 2, \dots, n$ para cada observación. En el caso de empates (observaciones idénticas), seguimos el procedimiento acostumbrado de reemplazar las observaciones por la media de los rangos que tendrían las observaciones si fueran distinguibles. La suma de los rangos que corresponde a las n_i observaciones en la i -ésima muestra se denota mediante la variable aleatoria R_i . Consideremos ahora el estadístico

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1),$$

que se aproxima muy bien mediante una distribución chi cuadrada con $k - 1$ grados de libertad, cuando H_0 es verdadera, siempre y cuando cada muestra conste de al menos 5 observaciones. El hecho de que h , el supuesto valor de H , sea grande cuando las muestras independientes provienen de poblaciones que no son idénticas nos permite establecer el siguiente criterio de decisión para probar H_0 :

| | |
|--------------------------|----------------------------------------------------------------------------------------------------------------------|
| Prueba de Kruskal-Wallis | Para probar la hipótesis nula H_0 de que k muestras independientes provienen de poblaciones idénticas se calcula |
|--------------------------|----------------------------------------------------------------------------------------------------------------------|

$$h = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{r_i^2}{n_i} - 3(n+1),$$

donde r_i es el valor supuesto de R_i para $i = 1, 2, \dots, k$. Si h cae en la región crítica $H > \chi_{\alpha}^2$ con $v = k - 1$ grados de libertad, se rechaza H_0 al nivel de significancia α ; de otra manera no se rechaza H_0 .

Ejemplo 16.6: En un experimento para determinar cuál de tres diferentes sistemas de misiles es preferible, se mide la tasa de combustión del propulsor. Los datos, después de codificarlos, se presentan en la tabla 16.5. Utilice la prueba de Kruskal-Wallis y un nivel de significancia de $\alpha = 0.05$ para probar la hipótesis de que las tasas de combustión del propulsor son iguales para los tres sistemas de misiles.

Tabla 16.5: Tasas de combustión del propulsor

| Sistema de misiles | | | | | | | | |
|--------------------|------|------|------|------|------|------|------|------|
| 1 | | | 2 | | | 3 | | |
| 24.0 | 16.7 | 22.8 | 23.2 | 19.8 | 18.1 | 18.4 | 19.1 | 17.3 |
| 19.8 | 18.9 | | 17.6 | 20.2 | 17.8 | 17.3 | 19.7 | 18.9 |
| | | | | | | 18.8 | 19.3 | |

- Solución:**
1. $H_0: \mu_1 = \mu_2 = \mu_3$.
 2. H_1 : las tres medias son diferentes.
 3. $\alpha = 0.05$.
 4. Región crítica: $h > \chi_{0.05}^2 = 5.991$, para $v = 2$ grados de libertad.
 5. Cálculos: En la tabla 16.6 convertimos las 19 observaciones a rangos y sumamos los rangos para cada sistema de misiles.

Tabla 16.6: Rangos para las tasas de combustión del propulsor

| Sistema de misiles | | |
|--------------------|--------------|--------------|
| 1 | 2 | 3 |
| 19 | 18 | 7 |
| 1 | 14.5 | 11 |
| 17 | 6 | 2.5 |
| 14.5 | 4 | 2.5 |
| 9.5 | 16 | 13 |
| $r_1 = 61.0$ | 5 | 9.5 |
| | $r_2 = 63.5$ | 8 |
| | | 12 |
| | | $r_3 = 65.5$ |

Ahora, al sustituir $n_1 = 5$, $n_2 = 6$, $n_3 = 8$ y $r_1 = 61.0$, $r_2 = 63.5$, $r_3 = 65.5$, el estadístico de prueba H toma el valor

$$h = \frac{12}{(19)(20)} \left(\frac{61.0^2}{5} + \frac{63.5^2}{6} + \frac{65.5^2}{8} \right) - (3)(20) = 1.66.$$

6. Decisión: Como $h = 1.66$ no cae en la región crítica $h > 5.991$, no hay evidencia suficiente para rechazar la hipótesis de que las tasas de combustión del propulsor son iguales para los tres sistemas de misiles. ▀

Ejercicios

16.15 Un fabricante de cigarrillos afirma que el contenido de alquitrán de la marca de cigarrillos *B* es menor que la de la marca *A*. Para probar esta afirmación se registraron las siguientes medidas del contenido de alquitrán, en miligramos:

| | | | | | | |
|----------------|---|----|---|----|----|----|
| Marca <i>A</i> | 1 | 12 | 9 | 13 | 11 | 14 |
| Marca <i>B</i> | 8 | 10 | 7 | | | |

Utilice la prueba de suma de rangos con $\alpha = 0.05$ para probar si la afirmación es válida.

16.16 Para averiguar si un nuevo suero detendrá la leucemia se seleccionan nueve pacientes que se encuentran en una etapa avanzada de la enfermedad. Cinco pacientes reciben el tratamiento y cuatro no. Los tiempos de supervivencia, en años, a partir del momento en que comienza el experimento son

| | | | | | |
|-----------------|-----|-----|-----|-----|-----|
| Con tratamiento | 2.1 | 5.3 | 1.4 | 4.6 | 0.9 |
| Sin tratamiento | 1.9 | 0.5 | 2.8 | 3.1 | |

Utilice la prueba de suma de rangos a un nivel de significancia de 0.05 para determinar si el suero es eficaz.

16.17 Los siguientes datos representan el número de horas que operan dos diferentes tipos de calculadoras científicas de bolsillo antes de que necesiten recargarse.

| | | | | | | | | | |
|----------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Calculadora <i>A</i> | 5.5 | 5.6 | 6.3 | 4.6 | 5.3 | 5.0 | 6.2 | 5.8 | 5.1 |
| Calculadora <i>B</i> | 3.8 | 4.8 | 4.3 | 4.2 | 4.0 | 4.9 | 4.5 | 5.2 | 4.5 |

Utilice la prueba de la suma de rangos con $\alpha = 0.01$ para determinar si la calculadora *A* opera más tiempo que la calculadora *B* con una carga completa de la batería.

16.18 Se fabrica un hilo para pesca usando dos procesos. Para determinar si hay una diferencia en la resistencia media a la rotura de los hilos, se seleccionan 10 piezas de cada proceso y después se prueba la resistencia a la rotura de cada una. Los resultados son los siguientes:

| | | | | | |
|-----------|------|------|------|------|------|
| Proceso 1 | 10.4 | 9.8 | 11.5 | 10.0 | 9.9 |
| | 9.6 | 10.9 | 11.8 | 9.3 | 10.7 |
| Proceso 2 | 8.7 | 11.2 | 9.8 | 10.1 | 10.8 |
| | 9.5 | 11.0 | 9.8 | 10.5 | 9.9 |

Utilice la prueba de suma de rangos con $\alpha = 0.1$ para determinar si hay diferencia entre las resistencias medias a la rotura de los hilos fabricados mediante los dos procesos.

16.19 De una clase de matemáticas de 12 estudiantes que tienen las mismas capacidades y utilizan material programado se seleccionan cinco al azar para propor-

cionarles enseñanza adicional. Los resultados del examen final son los siguientes:

| Conseñanza adicional | Calificación | | | | | | |
|-------------------------|--------------|----|----|----|----|----|----|
| | 87 | 69 | 78 | 91 | 80 | | |
| Sin enseñanza adicional | 75 | 88 | 64 | 82 | 93 | 79 | 67 |

Utilice la prueba de la suma de rangos con $\alpha = 0.05$ para determinar si la enseñanza adicional influye en la calificación promedio.

16.20 Los siguientes datos representan los pesos, en kilogramos, del equipaje personal que llevan, en diferentes vuelos, un jugador de un equipo de beisbol y un jugador de un equipo de basquetbol.

| Peso del equipaje (kilogramos) | | | | | |
|--------------------------------|------|------|-----------------------|------|--|
| Jugador de beisbol | | | Jugador de basquetbol | | |
| 16.3 | 20.0 | 18.6 | 15.4 | 16.3 | |
| 18.1 | 15.0 | 15.4 | 17.7 | 18.1 | |
| 15.9 | 18.6 | 15.6 | 18.6 | 16.8 | |
| 14.1 | 14.5 | 18.3 | 12.7 | 14.1 | |
| 17.7 | 19.1 | 17.4 | 15.0 | 13.6 | |
| 16.3 | 13.6 | 14.8 | 15.9 | 16.3 | |
| 13.2 | 17.2 | 16.5 | | | |

Utilice la prueba de la suma de rangos con $\alpha = 0.05$ para probar la hipótesis nula de que los dos atletas llevan la misma cantidad de equipaje en promedio, en comparación con la hipótesis alternativa de que el peso promedio del equipaje de los dos atletas es diferente.

16.21 Los siguientes datos representan los tiempos de funcionamiento, en horas, para tres tipos de calculadoras científicas de bolsillo, antes de que requieran recarga:

| Calculadora | | | | | | | | |
|-------------|-----|-----|----------|-----|-----|----------|-----|-----|
| <i>A</i> | | | <i>B</i> | | | <i>C</i> | | |
| 4.9 | 6.1 | 4.3 | 5.5 | 5.4 | 6.2 | 6.4 | 6.8 | 5.6 |
| 4.6 | 5.2 | | 5.8 | 5.5 | 5.2 | 6.5 | 6.3 | 6.6 |
| | | | 4.8 | | | | | |

Utilice la prueba de Kruskal-Wallis a un nivel de significancia de 0.01, para probar la hipótesis de que los tiempos de funcionamiento de las tres calculadoras son iguales.

16.22 En el ejercicio 13.6 de la página 519 utilice la prueba de Kruskal-Wallis, a un nivel de significancia de 0.05, para determinar si los solventes químicos orgánicos difieren de manera significativa en su tasa de absorción.

16.5 Pruebas de rachas

Al aplicar los diversos conceptos estadísticos que se presentan a lo largo de este libro siempre asumimos que los datos muestrales se reunieron mediante algún procedimiento aleatorio. Las **pruebas de rachas**, que se basan en el orden en el que se obtienen las observaciones muestrales, constituyen una técnica útil para probar la hipótesis nula H_0 de que las observaciones en realidad se obtuvieron al azar.

Para ilustrar las pruebas de rachas suponga que se encuesta a 12 personas para saber si utilizan cierto producto. Se cuestionaría seriamente la supuesta aleatoriedad de la muestra si las 12 personas fueran del mismo sexo. Designaremos a un hombre y a una mujer con los símbolos H y M , respectivamente, y registraremos los resultados de acuerdo con su género en el orden en que ocurren. Una secuencia común para el experimento sería

$$\underbrace{M M}_{\text{racha}} \underbrace{F F F}_{\text{racha}} \underbrace{M}_{\text{racha}} \underbrace{F F}_{\text{racha}} \underbrace{M M M M}_{\text{racha}},$$

donde agrupamos las subsecuencias de símbolos idénticos. Tales agrupamientos se llaman **rachas**.

Definición 16.1: Una **racha** es una subsecuencia de uno o más símbolos idénticos que representan una propiedad común de los datos.

Sin importar si las mediciones de la muestra representan datos cualitativos o cuantitativos, la prueba de rachas divide los datos en dos categorías mutuamente excluyentes: hombre o mujer, defectuoso o no defectuoso, cara o cruz, arriba o abajo de la mediana, etcétera. En consecuencia, una secuencia siempre estará limitada a dos símbolos distintos. Sea n_1 el número de símbolos asociados con la categoría de menor ocurrencia, y n_2 el número de símbolos que pertenecen a la otra categoría. Entonces, el tamaño de la muestra $n = n_1 + n_2$.

Para los $n = 12$ símbolos en nuestra encuesta tenemos cinco rachas, donde la primera incluye dos H , la segunda tres M , y así sucesivamente. Si el número de rachas es mayor o menor que el que esperaríamos por el azar, se debe rechazar la hipótesis de que la muestra se extrajo al azar. Ciertamente, una muestra que tiene como resultado sólo dos corridas,

$$H H H H H H H M M M M M$$

o la inversa, es muy improbable que provenga de un proceso de selección aleatorio. Este resultado indicaría que las primeras siete personas entrevistadas son hombres, seguidos de cinco mujeres. Asimismo, si la muestra tiene como resultado el número máximo de 12 rachas, como en la secuencia alternada

$$H M H M H M H M H M H M,$$

de nuevo sospecharíamos del orden en que se seleccionaron los individuos para la encuesta.

La prueba de rachas para la aleatoriedad se basa en la variable aleatoria V , el número total de rachas que suceden en la secuencia completa del experimento. En la tabla A.18 se dan valores de $P(V \leq v^*$ cuando H_0 es verdadera) para $v^* = 2, 3, \dots, 20$ rachas y valores

de n_1 y n_2 menores o iguales que 10. Los valores P tanto para pruebas de una cola como de dos colas se pueden obtener usando estos valores tabulados.

En la encuesta anterior presentamos un total de 5 M y 7 H . De aquí, con $n_1 = 5$, $n_2 = 7$ y $v = 5$, en la tabla A.18 observamos que el valor P para una prueba de dos colas es

$$P = 2P(V \leq 5 \text{ cuando } H_0 \text{ es verdadera}) = 0.394 > 0.05.$$

Es decir, el valor $v = 5$ es razonable a un nivel de significancia de 0.05 cuando H_0 es verdadera y, por lo tanto, no tenemos suficiente evidencia para rechazar la hipótesis de aleatoriedad de nuestra muestra.

Cuando el número de rachas es grande, por ejemplo, cuando $v = 11$ y $n_1 = 5$ y $n_2 = 7$, entonces el valor P en una prueba de dos colas es

$$\begin{aligned} P &= 2P(V \geq 11 \text{ cuando } H_0 \text{ es verdadera}) = 2[1 - P(V \leq 10 \text{ cuando } H_0 \text{ es verdadera})] \\ &= 2(1 - 0.992) = 0.016 < 0.05, \end{aligned}$$

que nos lleva a rechazar la hipótesis de que los valores de la muestra ocurren al azar.

La prueba de rachas también sirve para detectar desviaciones en la aleatoriedad de una secuencia de mediciones cuantitativas a lo largo del tiempo, ocasionadas por tendencias o periodos. Al reemplazar cada medición en el orden en que se obtiene, con un símbolo *más* si caen por arriba de la mediana, o con un símbolo *menos* si caen por debajo de la mediana, y omitiendo todas las mediciones que son exactamente iguales a la mediana, se genera una secuencia de signos de más y menos que se somete a prueba para verificar su aleatoriedad, como se ilustra en el siguiente ejemplo.

Ejemplo 16.7: Una máquina vierte adelgazador de pintura acrílica en un contenedor. ¿Si se mide el contenido de los siguientes 15 contenedores y los resultados son 3.6, 3.9, 4.1, 3.6, 3.8, 3.7, 3.4, 4.0, 3.8, 4.1, 3.9, 4.0, 3.8, 4.2 y 4.1 litros, diría que la cantidad de adelgazador de pintura que despacha la máquina varía de forma aleatoria? Utilice un nivel de significancia de 0.1.

- Solución:**
1. H_0 : La secuencia es aleatoria.
 2. H_1 : La secuencia no es aleatoria.
 3. $\alpha = 0.1$.
 4. Estadístico de prueba: V , número total de rachas.
 5. Cálculos: Para la muestra dada encontramos $\bar{x} = 3.9$. Al reemplazar cada medición por el símbolo "+", si cae por arriba de 3.9, por el signo "-" si cae por debajo de 3.9, y si se omiten las dos mediciones que son iguales a 3.9, obtenemos la secuencia

- + - - - + - + + - + +

para la que $n_1 = 6$, $n_2 = 7$ y $v = 8$. Por lo tanto, de la tabla A.18, el valor P calculado es

$$\begin{aligned} P &= 2P(V \geq 8 \text{ cuando } H_0 \text{ es verdadera}) \\ &= 2[1 - P(V \leq 8 \text{ cuando } H_0 \text{ es verdadera})] = 2(0.5) = 1. \end{aligned}$$

6. Decisión: No se rechaza la hipótesis de que la secuencia de mediciones varía de forma aleatoria. J

La prueba de rachas, aunque menos poderosa, también se utiliza como una alternativa a la prueba de dos muestras de Wilcoxon para probar la afirmación de que dos muestras aleatorias provienen de poblaciones que tienen la misma distribución y, por lo tanto, medias iguales. Si las poblaciones son simétricas, el rechazo de la afirmación de distribuciones iguales es equivalente a aceptar la hipótesis alternativa de que las medias no son iguales. Para hacer la prueba primero se combinan las observaciones de ambas muestras y se acomodan en orden ascendente. Ahora se asigna la letra A a cada observación tomada de una de las poblaciones, y la letra B a cada observación de la otra población, generando así una secuencia que consta de los símbolos A y B . Si las observaciones de una población se vinculan con las observaciones de la otra población, la secuencia de símbolos A y B que se genera no será única y, en consecuencia, es poco probable que el número de rachas sea único. Los procedimientos para romper los empates por lo general dan como resultado tediosos cálculos adicionales, por lo que siempre que ocurran dichas situaciones sería preferible aplicar la prueba de la suma de rangos de Wilcoxon.

Con el fin de ilustrar el uso de las rachas al probar la igualdad de medias, considere los tiempos de supervivencia de los pacientes de leucemia del ejercicio 16.16 de la página 670, para los que tenemos

| | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0.5 | 0.9 | 1.4 | 1.9 | 2.1 | 2.8 | 3.1 | 4.6 | 5.3 |
| B | A | A | B | A | B | B | A | A |

que resultan en $v = 6$ rachas. Si las dos poblaciones simétricas tienen medias iguales, las observaciones de las dos muestras estarán entremezcladas, lo cual dará como resultado muchas rachas. Sin embargo, si las medias de la población son significativamente diferentes, esperaríamos que la mayoría de las observaciones de una de las dos muestras fueran más pequeñas que las de la otra muestra. En el caso extremo de que las poblaciones no se traslapan, obtendríamos una secuencia de la forma

$$A A A A A B B B B \text{ o } B B B B A A A A A$$

y en cualquier caso sólo habría dos rachas. En consecuencia, la hipótesis de medias de la población iguales se rechazará a un nivel de significancia α sólo cuando v sea suficientemente pequeña, de modo que

$$P = P(V \leq v \text{ cuando } H_0 \text{ es verdadera}) \leq \alpha,$$

lo que implica una prueba de una cola.

Si regresamos a los datos del ejercicio 16.16 de la página 670, para los que $n_1 = 4$, $n_2 = 5$ y $v = 6$, en la tabla A.18 encontramos que

$$P = P(V \leq 6 \text{ cuando } H_0 \text{ es verdadera}) = 0.786 > 0.05$$

y, por lo tanto, no se rechaza la hipótesis nula de medias iguales. De aquí concluimos que el nuevo suero no prolonga la vida, ya que no detiene la leucemia.

Cuando n_1 y n_2 aumentan en tamaño, la distribución de muestreo de V se aproxima a la distribución normal con media y varianza dadas por

$$\mu_V = \frac{2n_1n_2}{n_1 + n_2} + 1 \text{ y } \sigma_V^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}.$$

En consecuencia, cuando n_1 y n_2 son ambos mayores que 10, se puede utilizar el estadístico

$$Z = \frac{V - \mu_V}{\sigma_V}$$

con el fin de establecer la región crítica para la prueba de rachas.

16.6 Límites de tolerancia

En el capítulo 9 se analizaron los límites de tolerancia para una distribución normal de mediciones. En esta sección consideramos un método para construir intervalos de tolerancia que sean independientes de la forma de la distribución subyacente. Como se podría sospechar, para un grado de confianza razonable serán considerablemente más grandes que los que se construyen cuando se supone normalidad, y el tamaño de la muestra que se requiere es por lo general muy grande. Los límites de tolerancia no paramétricos se establecen en términos de las observaciones más grande y más pequeña en nuestra muestra.

Límites de tolerancia bilaterales Para cualquier distribución de mediciones los límites de tolerancia bilaterales son indicados por las observaciones más grande y más pequeña en una muestra de tamaño n , donde n se determina de manera que se asegure, con $100(1 - \gamma)\%$ de confianza, que **al menos** la proporción $1 - \alpha$ de la distribución está incluida entre los extremos de la muestra.

La tabla A.19 proporciona los tamaños de la muestra requeridos para los valores seleccionados de γ y $1 - \alpha$. Por ejemplo, cuando $\gamma = 0.01$ y $1 - \alpha = 0.95$, debemos seleccionar una muestra aleatoria de tamaño $n = 130$ para tener 99% de confianza en que al menos 95% de la distribución de mediciones está incluido entre los extremos de la muestra.

En vez de determinar un tamaño muestral n tal que una proporción específica de mediciones esté contenida entre los extremos de la muestra, en muchos procesos industriales es deseable determinar un tamaño de la muestra tal que una proporción fija de la población caiga por debajo de la observación más grande (o por arriba de la más pequeña) de la muestra. Tales límites se denominan límites de tolerancia unilaterales.

Límites de tolerancia unilaterales Para cualquier distribución de mediciones un límite de tolerancia unilateral se determina mediante la observación más pequeña (o más grande) en una muestra de tamaño n , donde n se determina de manera que se pueda asegurar con $100(1 - \gamma)\%$ de confianza que **al menos** la proporción $1 - \alpha$ de la distribución excederá a la observación más pequeña (menor que la mayor) de la muestra.

La tabla A.20 muestra los tamaños de la muestra requeridos, correspondientes a valores seleccionados de γ y $1 - \alpha$. De aquí, cuando $\gamma = 0.05$ y $1 - \alpha = 0.70$, debemos elegir una muestra de tamaño $n = 9$ para tener 95% de confianza en que 70% de nuestra distribución de mediciones excederá la observación más pequeña de la muestra.

16.7 Coeficiente de correlación de rango

En el capítulo 11 utilizamos el coeficiente de correlación muestral r para medir el coeficiente de correlación poblacional ρ , la relación lineal entre dos variables continuas X y Y . Si los rangos $1, 2, \dots, n$ se asignan a las observaciones x en orden de magnitud y de manera similar a las observaciones y , y si estos rangos se sustituyen después con los valores numéricos reales en la fórmula para el coeficiente de correlación del capítulo 11, obtenemos el equivalente no paramétrico del coeficiente de correlación convencional. Un coeficiente de correlación calculado de esta forma se conoce como **coeficiente de correlación de rangos de Spearman** y se denota con r_s . Cuando no hay empates entre ambos conjuntos de mediciones la fórmula para r_s se reduce a una expresión mucho más simple que incluye las diferencias d_i entre los rangos asignados a los n pares de x y y que establecemos ahora.

Coeficiente de correlación de rangos Una medida no paramétrica de la asociación entre dos variables X y Y es dada por el **coeficiente de correlación de rango**

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n d_i^2,$$

donde d_i es la diferencia entre los rangos asignados a x_i y y_i , y n es el número de pares de datos.

En la práctica, la fórmula anterior también se usa cuando hay empates entre las observaciones x o y . Los rangos para observaciones empatadas se asignan de la misma manera que en la prueba de rango con signo al promediar los rangos que se habrían asignado si las observaciones fueran distinguibles.

El valor de r_s por lo general se acercará al valor que se obtiene al calcular r con base en mediciones numéricas y se interpreta de forma muy similar. Como antes, el valor de r_s irá de -1 a $+1$. Un valor de $+1$ o -1 indica una asociación perfecta entre X y Y ; el signo más ocurre para rangos idénticos y el signo menos para rangos inversos. Cuando r_s se acerca a cero, se concluye que las variables no están correlacionadas.

Ejemplo 16.8: Las cifras que se listan en la tabla 16.7, publicadas por la Comisión Federal de Comercio, muestran los miligramos de alquitrán y nicotina que se encontraron en 10 marcas de cigarrillos. Calcule el coeficiente de correlación de rangos para medir el grado de relación entre el contenido de alquitrán y de nicotina en cigarrillos.

Tabla 16.7: Contenidos de alquitrán y nicotina

| Marca de cigarrillos | Contenido de alquitrán | Contenido de nicotina |
|----------------------|------------------------|-----------------------|
| Viceroy | 14 | 0.9 |
| Marlboro | 17 | 1.1 |
| Chesterfield | 28 | 1.6 |
| Kool | 17 | 1.3 |
| Kent | 16 | 1.0 |
| Raleigh | 13 | 0.8 |
| Old Gold | 24 | 1.5 |
| Philip Morris | 25 | 1.4 |
| Oasis | 18 | 1.2 |
| Players | 31 | 2.0 |

Solución: Sean X y Y los contenidos de alquitrán y nicotina, respectivamente. Primero asignamos rangos a cada conjunto de medidas, con el rango de 1 asignado al número más bajo en cada conjunto, el rango de 2 al segundo número más bajo en cada conjunto, y así sucesivamente, hasta que se asigna el rango 10 al número más grande. La tabla 16.8 muestra los rangos individuales de las mediciones y las diferencias en rangos para los 10 pares de observaciones.

Tabla 16.8: Rangos para los contenidos de alquitrán y nicotina

| Marca de cigarrillos | x_i | y_i | d_i |
|----------------------|-------|-------|-------|
| Viceroy | 2.0 | 2.0 | 0.0 |
| Marlboro | 4.5 | 4.0 | 0.5 |
| Chesterfield | 9.0 | 9.0 | 0.0 |
| Kool | 4.5 | 6.0 | -1.5 |
| Kent | 3.0 | 3.0 | 0.0 |
| Raleigh | 1.0 | 1.0 | 0.0 |
| Old Gold | 7.0 | 8.0 | -1.0 |
| Philip Morris | 8.0 | 7.0 | 1.0 |
| Oasis | 6.0 | 5.0 | 1.0 |
| Players | 10.0 | 10.0 | 0.0 |

Al sustituir en la fórmula para r_s , encontramos que

$$r_s = 1 - \frac{(6)(5.50)}{(10)(100 - 1)} = 0.967,$$

lo que indica una correlación positiva alta entre las cantidades de alquitrán y de nicotina que se encuentra en los cigarrillos. ■

Hay algunas ventajas al usar r_s en vez de r . Por ejemplo, ya no suponemos que la relación fundamental entre X y Y es lineal, por lo tanto, cuando los datos poseen una relación curvilínea distinta, el coeficiente de correlación de rangos probablemente será más confiable que la medida convencional. Una segunda ventaja del uso del coeficiente de correlación de rangos es el hecho de que no se hacen suposiciones de normalidad respecto a las distribuciones de X y Y . Quizá la mayor ventaja ocurre cuando no somos capaces de hacer mediciones numéricas significativas y, sin embargo, se pueden establecer rangos. Tal es el caso, por ejemplo, cuando diferentes jueces clasifican a un grupo de individuos de acuerdo con algún atributo. El coeficiente de correlación de rangos se puede utilizar en esta situación como una medida de la consistencia de los dos jueces.

Para probar la hipótesis de que $\rho = 0$ utilizando un coeficiente de correlación de rangos, se necesita considerar la distribución muestral de los valores r_s , con base en la suposición de que no hay correlación. En la tabla A.21 aparecen valores críticos calculados para $\alpha = 0.05, 0.025, 0.01$ y 0.005 . La elaboración de esta tabla es similar a la tabla de valores críticos para la distribución t , excepto por la columna izquierda, que ahora proporciona el número de pares de observaciones en vez de los grados de libertad. Como la distribución de los valores r_s es simétrica alrededor de cero cuando $\rho = 0$, el valor r_s que deja una área de α a la izquierda es igual al negativo del valor r_s que deja una área de α a la derecha. Para una hipótesis alternativa bilateral la región crítica de tamaño α cae igualmente en las dos colas de la distribución. Para una prueba en la que la hipótesis alternativa es negativa, la región crítica está completamente en la cola izquierda de la distribución y, cuando la hipótesis alternativa es positiva, la región crítica se coloca por completo en la cola derecha.

Ejemplo 16.9: Remítase al ejemplo 16.8 y pruebe la hipótesis de que la correlación entre la cantidad de alquitrán y nicotina encontrada en los cigarrillos es cero en comparación con la hipótesis alternativa de que es mayor que cero. Utilice un nivel de significancia de 0.01.

- Solución:**
1. $H_0: \rho = 0$.
 2. $H_1: \rho > 0$.
 3. $\alpha = 0.01$.
 4. Región crítica: $r_s > 0.745$, de la tabla A.21.
 5. Cálculos: Del ejemplo 16.8, $r_s = 0.967$.
 6. Decisión: Se rechaza H_0 y se concluye que hay una correlación significativa entre la cantidad de alquitrán y nicotina que se encuentra en los cigarrillos. ■

Con base en la suposición de que no hay correlación, se puede demostrar que la distribución de los valores r_s se aproxima a una distribución normal, con una media igual a cero y una desviación estándar de $1/\sqrt{n-1}$ conforme aumenta n . En consecuencia, cuando n excede a los valores dados en la tabla A.21 se puede probar si existe una correlación significativa calculando

$$z = \frac{r_s - 0}{1/\sqrt{n-1}} = r_s \sqrt{n-1}$$

y comparando con los valores críticos de la distribución normal estándar que se presentan en la tabla A.3.

Ejercicios

16.23 Con el fin de estimar la proporción de votantes que favorecen a cierto candidato para alcalde, se selecciona una muestra aleatoria de 15 adultos que viven en una pequeña ciudad. También se le pregunta a cada individuo si se graduó de la universidad. Al denotar con S y N las respuestas "sí" y "no", respectivamente, a la pregunta sobre la escolaridad, se obtuvo la siguiente secuencia:

$N N N N Y Y N Y Y N Y N N N N$

Utilice la prueba de rachas a un nivel de significancia de 0.1 para determinar si la secuencia apoya la afirmación de que la muestra se seleccionó al azar.

16.24 Se utiliza un proceso de platingo para cubrir cierto tipo de charola de servicio. Cuando el proceso está bajo control el espesor de la plata sobre la charola variará de forma aleatoria siguiendo una distribución normal con una media de 0.02 milímetros y una desviación estándar de 0.005 milímetros. Suponga que las siguientes 12 charolas examinadas muestran los siguientes espesores de plata: 0.019, 0.021, 0.020, 0.019, 0.020, 0.018, 0.023, 0.021, 0.024, 0.022, 0.023, 0.022. Utilice la prueba de rachas para determinar si las fluc-

tuaciones en el espesor de una charola a otra son aleatorias. Utilice $\alpha = 0.05$.

16.25 Use la prueba de rachas a un nivel de 0.01 para probar si hay una diferencia en el tiempo promedio de funcionamiento para las dos calculadoras del ejercicio 16.17 de la página 670.

16.26 En una línea de producción industrial los artículos se inspeccionan de forma periódica en busca de defectos. La siguiente es una secuencia de artículos defectuosos, D , y no defectuosos, N , producidos por esta línea:

$D D N N N D N N D D N N N N$

$N D D D N N D N N N N D N D$

Utilice la teoría de muestras grandes para la prueba de rachas a un nivel de significancia de 0.05 para determinar si los artículos defectuosos ocurren al azar.

16.27 Suponga que las mediciones del ejercicio 1.14 de la página 30 se registraron en renglones sucesivos de izquierda a derecha conforme se reunieron. Utilice la prueba de rachas con $\alpha = 0.05$ para probar la hipótesis de que los datos representan una secuencia aleatoria.

16.28 ¿Qué tan grande debe ser una muestra para tener 95% de confianza en que al menos 85% de la distribución de medidas se incluye entre los extremos de la muestra?

16.29 ¿Cuál es la probabilidad de que el rango de una muestra aleatoria de tamaño 24 incluya al menos a 90% de la población?

16.30 ¿Qué tan grande debe ser una muestra para tener 99% de confianza en que al menos 80% de la población será menor que la observación más grande de la muestra?

16.31 ¿Cuál es la probabilidad de que al menos 95% de una población exceda al valor más pequeño en una muestra aleatoria de tamaño $n = 135$?

16.32 En la siguiente tabla se presentan las calificaciones registradas de 10 estudiantes en un examen de medio curso y las del examen final en un curso de cálculo:

| Estudiante | Examen de medio curso | Examen final |
|------------|-----------------------|--------------|
| L.S.A. | 84 | 73 |
| W.P.B. | 98 | 63 |
| R.W.K. | 91 | 87 |
| J.R.L. | 72 | 66 |
| J.K.L. | 86 | 78 |
| D.L.P. | 93 | 78 |
| B.L.P. | 80 | 91 |
| D.W.M. | 0 | 0 |
| M.N.M. | 92 | 88 |
| R.H.S. | 87 | 77 |

- Calcule el coeficiente de correlación de rangos.
- Pruebe la hipótesis nula de que $\rho = 0$ en comparación con la hipótesis alternativa de que $\rho > 0$. Utilice $\alpha = 0.025$.

16.33 Refiérase a los datos del ejercicio 11.1 de la página 398 y

- calcule el coeficiente de correlación de rangos;
- a un nivel de significancia de 0.05 pruebe la hipótesis nula de que $\rho = 0$, en comparación con la hipótesis alternativa de que $\rho \neq 0$. Compare sus resultados con los obtenidos en el ejercicio 11.44 de la página 435.

16.34 Calcule el coeficiente de correlación de rangos para la precipitación pluvial diaria y la cantidad de partículas eliminadas en el ejercicio 11.13 de la página 400.

16.35 Refiérase a los datos del ejercicio 11.47 de la página 436 respecto al peso y tamaño de tórax de los bebés, y

- calcule el coeficiente de correlación de rangos;
- a un nivel de significancia de 0.025, pruebe la hipótesis de que $\rho = 0$ en comparación con la hipótesis alternativa de que $\rho > 0$.

16.36 Un grupo de consumidores prueba la calidad general de nueve marcas de hornos de microondas. Los rangos asignados por el grupo y los precios de venta al menudeo sugeridos son los siguientes:

| Fabricante | Clasificación del grupo | Precio sugerido (\$) |
|------------|-------------------------|----------------------|
| A | 6 | 480 |
| B | 9 | 395 |
| C | 2 | 575 |
| D | 8 | 550 |
| E | 5 | 510 |
| F | 1 | 545 |
| G | 7 | 400 |
| H | 4 | 465 |
| I | 3 | 420 |

¿Existe una relación significativa entre la calidad y el precio de un horno de microondas? Utilice un nivel de significancia de 0.05.

16.37 En un desfile de regreso a clases dos jueces califican ocho carros alegóricos en el siguiente orden:

| | Carro alegórico | | | | | | | |
|--------|-----------------|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Juez A | 5 | 8 | 4 | 3 | 6 | 2 | 7 | 1 |
| Juez B | 7 | 5 | 4 | 2 | 8 | 1 | 6 | 3 |

- Calcule el coeficiente de correlación de rangos.
- Pruebe la hipótesis nula de que $\rho = 0$ en comparación con la hipótesis alternativa de que $\rho > 0$. Use $\alpha = 0.05$.

16.38 En el artículo titulado "Risky Assumptions" de Paul Slovic, Baruch Fischhoff y Sarah Lichtenstein, publicado en *Psychology Today* (junio de 1980), miembros de la Liga de Mujeres Votantes y expertos profesionalmente implicados en la evaluación de riesgos clasificaron el riesgo de muerte, en Estados Unidos, de realizar 30 actividades y utilizar tecnologías. Las puntuaciones se presentan en la tabla 16.9.

- Calcule el coeficiente de correlación de rangos.
- Pruebe la hipótesis nula de cero correlación entre las clasificaciones de la Liga de Mujeres Votantes y de los expertos en comparación con la hipótesis alternativa de que la correlación no es igual a cero. Utilice un nivel de significancia de 0.05.

Tabla 16.9: Rango de datos para el ejercicio 16.38

| Riesgo de la actividad o tecnología | Riesgo de la actividad o tecnología | | Riesgo de la actividad o tecnología | Riesgo de la actividad o tecnología | |
|-------------------------------------|-------------------------------------|----------|-------------------------------------|-------------------------------------|----------|
| | Votantes | Expertos | | Votantes | Expertos |
| Energía nuclear | 1 | 20 | Vehículos de motor | 2 | 1 |
| Armas de fuego | 3 | 4 | Tabaquismo | 4 | 2 |
| Motocicletas | 5 | 6 | Bebidas alcohólicas | 6 | 3 |
| Aviación privada | 7 | 12 | Trabajo policiaco | 8 | 17 |
| Pesticidas | 9 | 8 | Cirugía | 10 | 5 |
| Bombero | 11 | 18 | Construcción grande | 12 | 13 |
| Cacería | 13 | 23 | Latas de aerosol | 14 | 26 |
| Montañismo | 15 | 29 | Bicicletas | 16 | 15 |
| Aviación comercial | 17 | 16 | Energía eléctrica | 18 | 9 |
| Natación | 19 | 10 | Anticonceptivos | 20 | 11 |
| Esquí | 21 | 30 | Rayos X | 22 | 7 |
| Futbol americano | 23 | 27 | Ferrocarriles | 24 | 19 |
| Conservadores de alimentos | 25 | 14 | Colorantes de alimentos | 26 | 21 |
| Podadoras | 27 | 28 | Antibióticos | 28 | 24 |
| Electrodomésticos | 29 | 22 | Vacunas | 30 | 25 |

Ejercicios de repaso

16.39 Un estudio de una empresa química compara las propiedades de desecación de dos diferentes polímeros. Se utilizaron 10 lodos diferentes y se permitió que ambos polímeros secan cada lodo. El secado libre se midió en mL/min.

| Tipo de lodo | Polímero A | Polímero B |
|--------------|------------|------------|
| 1 | 12.7 | 12.0 |
| 2 | 14.6 | 15.0 |
| 3 | 18.6 | 19.2 |
| 4 | 17.5 | 17.3 |
| 5 | 11.8 | 12.2 |
| 6 | 16.9 | 16.6 |
| 7 | 19.9 | 20.1 |
| 8 | 17.6 | 17.6 |
| 9 | 15.6 | 16.0 |
| 10 | 16.0 | 16.1 |

- a) Utilice la prueba de signos a un nivel de 0.05 para probar la hipótesis nula de que el polímero A tiene la misma mediana de secado que el polímero B.
- b) Utilice la prueba de rangos con signo para probar la hipótesis del inciso a.

16.40 En el ejercicio de repaso 13.45 de la página 555 use la prueba de Kruskal-Wallis, a un nivel de significancia de 0.05, para determinar si los análisis químicos realizados por los cuatro laboratorios producen, en promedio, los mismos resultados.

16.41 Use los datos del ejercicio 13.14 de la página 530 para ver si la cantidad mediana de pérdida de nitrógeno en la transpiración difiere para los tres niveles de proteína dietética.

Capítulo 17

Control estadístico de la calidad

17.1 Introducción

La idea de usar técnicas de muestreo y análisis estadístico en un entorno de producción tuvo sus comienzos en la década de 1920. El objetivo de este concepto tan exitoso es reducir de manera sistemática la variabilidad y el aislamiento asociados con las fuentes de dificultades *durante la producción*. En 1924 Walter A. Shewhart, de la empresa Bell Telephone Laboratories, desarrolló el concepto de gráfica de control. Sin embargo, fue hasta la Segunda Guerra Mundial cuando se generalizó el uso de este tipo de gráficas debido a la importancia que durante ese periodo tuvo el mantenimiento de la calidad en los procesos de producción. En las décadas de 1950 y 1960 el desarrollo del control de calidad y el área general de seguridad de la calidad crecieron con rapidez, en particular con el surgimiento del programa espacial en Estados Unidos. En Japón hubo un amplio y exitoso uso del control de calidad gracias a los esfuerzos de W. Edwards Deming, quien trabajó como consultor en Japón después de la Segunda Guerra Mundial. El control de calidad ha sido, y es, un elemento importante en el desarrollo de la industria y la economía de Japón.

El control de calidad está recibiendo cada vez más atención como una herramienta de administración en la cual se observan y evalúan las características importantes de un producto en comparación con algún tipo de estándar. Los diversos procedimientos en el control de calidad implican un uso considerable de los procedimientos de muestreo y los principios estadísticos expuestos en capítulos anteriores. Los principales usuarios del control de calidad son, por supuesto, las corporaciones industriales. Es evidente que un programa eficaz de control de calidad mejora la calidad del artículo que se produce y aumenta las utilidades. Esto es particularmente cierto en la actualidad, pues los productos se fabrican en volúmenes altos. Antes de que surgiera el movimiento hacia los métodos de control de calidad, a menudo ésta se veía afectada debido a la falta de eficiencia, lo cual, por supuesto, incrementaba los costos.

La gráfica de control

El objetivo de una gráfica de control es determinar si el desempeño de un proceso se mantiene en un nivel aceptable de calidad. Se espera, desde luego, que cualquier proceso experimente una variabilidad natural, es decir, una variabilidad debida esencialmente a fuentes de variación poco importantes e incontrolables. Por otro lado, un proceso puede experimentar formas más severas de variabilidad en mediciones de desempeño fundamentales.

Estas fuentes de variabilidad pueden surgir de uno de varios tipos de “causas asignables” no aleatorias, como errores del operador o indicadores mal ajustados en una máquina. Un proceso que opera en dicho estado se denomina **fuera de control**. Se dice que un proceso que sólo experimenta variaciones aleatorias está en **control estadístico**. Desde luego, un proceso de producción exitoso puede operar en un estado de control durante un periodo largo. Se supone que durante este periodo el proceso elabora un producto aceptable. Sin embargo, podría ocurrir un “cambio” gradual o repentino que requiera detección.

El propósito de una gráfica de control es que funcione como un dispositivo para detectar el estado no aleatorio o fuera de control de un proceso. La gráfica de control suele adoptar la forma que se indica en la figura 17.1. Cuando ocurre un cambio en el proceso es importante detectarlo con rapidez, de manera que se pueda corregir el problema. Evidentemente, si el cambio no se detecta de inmediato, se producirán muchos artículos defectuosos o que no cumplen las especificaciones, lo cual dará como resultado un desperdicio significativo y un incremento en los costos.

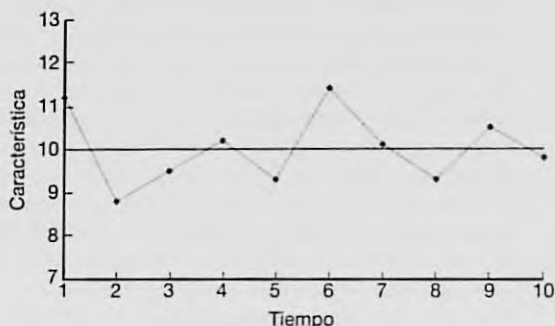


Figura 17.1: Gráfica de control típica.

Se deben considerar ciertos tipos de características de la calidad y se deben tomar muestras de las unidades del proceso a medida que pasa el tiempo. Digamos, por ejemplo, que la característica de un cojinete de motor es la circunferencia. La línea central representa el valor promedio de la característica cuando el proceso está bajo control. Los puntos que se indican en la figura representarían los resultados de, digamos, los promedios muestrales de tal característica, con muestras tomadas en diferentes momentos. Los límites de control superior e inferior se eligen de tal manera que se esperaría que, si el proceso está bajo control, todos los puntos muestrales queden cubiertos por estos límites. Como resultado, la forma general de los puntos graficados a lo largo del tiempo determina si se concluye que el proceso está bajo control. La evidencia de que está “dentro de control” se obtiene de un patrón aleatorio de puntos con todos los valores graficados dentro de los límites de control. Cuando un punto cae fuera de los límites de control, se considera como evidencia de que un proceso está fuera de control, en cuyo caso se sugiere una búsqueda para determinar la causa. Además, un patrón no aleatorio de puntos se debe considerar sospechoso y, evidentemente, un indicador de que es necesario investigar para encontrar la medida correctiva adecuada.

17.2 Naturaleza de los límites de control

Las ideas fundamentales en las que se basan las gráficas de control son similares en estructura a la prueba de hipótesis. Los límites de control se establecen para controlar la probabilidad de cometer el error de concluir que el proceso está fuera de control, cuando de hecho no lo está. Esto corresponde a la probabilidad de cometer un error tipo I si probáramos la hipótesis nula de que el proceso está bajo control. Por otro lado, debemos estar atentos al error del segundo tipo, es decir, el de no encontrar el proceso fuera de control cuando de hecho sí lo está (error tipo II). De esta manera, la elección de los límites de control es similar a la elección de una región crítica.

Como en el caso de la prueba de hipótesis, el tamaño de la muestra en cada punto es importante. La elección del tamaño de la muestra depende en gran medida de la sensibilidad o potencia de detección del estado fuera de control. En esta aplicación, el concepto de *potencia* es muy similar al de la situación de la prueba de hipótesis. Queda claro que cuanto más grande sea la muestra en cada periodo, más rápida será la detección de un proceso fuera de control. En cierto sentido los límites de control en realidad definen lo que el usuario considera como estar *bajo control*. En otras palabras, la amplitud dada por los límites de control debe depender en cierto sentido de la variabilidad del proceso. Como resultado, el cálculo de los límites de control dependerá de manera natural de los datos que se tomen de los resultados del proceso. De esta forma, cualquier aplicación del control de calidad debe comenzar con el cálculo de una muestra o conjunto de muestras preliminar, que establecerá tanto la línea central como los límites del control de calidad.

17.3 Objetivos de la gráfica de control

Un propósito evidente de la gráfica de control es la vigilancia del proceso, o sea determinar si es o no necesario realizar cambios. Además, la constante y sistemática obtención de datos a menudo permite a la administración evaluar la capacidad del proceso. Es evidente que, si una sola característica de desempeño es importante, el muestreo y la estimación continuos de la media y la desviación estándar de esa característica de desempeño ofrecen la actualización de lo que el proceso puede hacer en términos de desempeño promedio y variación aleatoria. Esto es valioso incluso cuando el proceso permanece bajo control durante periodos largos. La estructura sistemática y formal de la gráfica de control a menudo puede prevenir una reacción desmesurada ante cambios que representen sólo fluctuaciones aleatorias. Obviamente, en muchas situaciones los cambios realizados por una reacción desmesurada pueden crear graves problemas que son difíciles de resolver.

Las características de calidad de las gráficas de control por lo general caen en *dos* categorías: **variables** y **atributos**. Como resultado, los tipos de gráficas de control con frecuencia tienen las mismas clasificaciones. En el caso de la gráfica de los tipos de variables, la característica suele ser una medida sobre un continuo, como el diámetro o el peso. En el caso de la gráfica de atributos, lo que refleja la característica es si el producto individual se *ajusta a las especificaciones* (si está o no defectuoso). Las aplicaciones para estas dos situaciones distintas son evidentes.

En el caso de la gráfica de variables se debe ejercer control sobre la tendencia central y la variabilidad. Lo que a un analista de control de calidad le debe preocupar es si existe o no, *en promedio*, un cambio en los valores de la característica de desempeño. Además, siempre habrá interés por saber si algún cambio en las condiciones del proceso

provoca que disminuya la precisión, es decir, que aumente la variabilidad. Para manejar estos dos conceptos es esencial utilizar gráficas de control separadas. La tendencia central es controlada por la *gráfica \bar{X}* , donde las medias de muestras relativamente pequeñas se dibujan en la gráfica de control. La variabilidad alrededor de la media se controla mediante el *rango* en la muestra, o la *desviación estándar de la muestra*. En el caso de muestreo de atributos a menudo la cantidad que se grafica es la *proporción de artículos defectuosos* de una muestra. En la siguiente sección analizamos el desarrollo de gráficas de control para los tipos de variables de las características del desempeño.

17.4 Gráficas de control para variables

Un ejemplo es una forma relativamente sencilla de explicar los rudimentos de la gráfica \bar{X} para variables. Suponga que en un proceso de fabricación de cierta parte de un motor se deben utilizar las gráficas de control de calidad. Suponga también que la media del proceso es $\mu = 50$ mm y que la desviación estándar es $\sigma = 0.01$ mm. Imagine que se toman muestras en grupos de 5 cada hora y que los valores de la *media muestral \bar{X}* se registran y grafican como en la figura 17.2. Los límites para las gráficas \bar{X} se basan en la desviación estándar de la variable aleatoria \bar{X} . Sabemos, a partir de lo expuesto en el capítulo 8, que para el promedio de observaciones independientes en una muestra de tamaño n ,

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}},$$

donde σ es la desviación estándar de una observación individual. Los límites de control están diseñados para dar como resultado una pequeña probabilidad de que un valor dado de \bar{X} esté fuera de los límites dado que, en realidad, el proceso está bajo control, es decir, $\mu = 50$. Si recurrimos al teorema del límite central, tendremos que, en las condiciones en las que el proceso está controlado,

$$\bar{X} \sim N\left(50, \frac{0.01}{\sqrt{5}}\right).$$

Como resultado, $100(1 - \alpha)\%$ de los valores \bar{X} cae dentro de los límites cuando el proceso está bajo control si utilizamos los límites

$$\text{LCI} = \mu - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 50 - z_{\alpha/2} (0.0045), \quad \text{LCS} = \mu + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 50 + z_{\alpha/2} (0.0045).$$

Aquí LCI y LCS representan el límite de control inferior y el límite de control superior, respectivamente. Con frecuencia las gráficas \bar{X} se basan en límites denominados “tres-sigma”, refiriéndonos, por supuesto, a $z_{\alpha/2} = 3$ y a límites que se convierten en

$$\mu \pm 3 \frac{\sigma}{\sqrt{n}}.$$

En nuestro ejemplo, los límites superior e inferior son

$$\text{LCI} = 50 - 3(0.0045) = 49.9865, \quad \text{LCS} = 50 + 3(0.0045) = 50.0135.$$

Por consiguiente, si vemos la estructura de los límites 3σ desde el punto de vista de la prueba de hipótesis para un punto muestral dado, encontraremos que hay una probabilidad

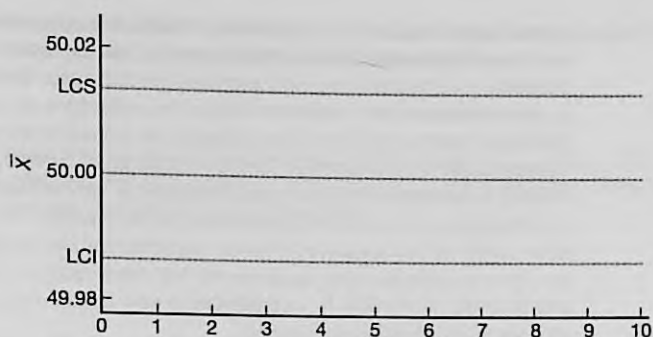


Figura 17.2: Los límites de control 3σ para el ejemplo de la parte del motor.

de 0.0026 de que el valor \bar{X} caiga fuera de los límites de control, dado que el proceso está bajo control. Ésta es la probabilidad de que el analista determine *de manera errónea* que el proceso está fuera de control (véase la tabla A.3).

El ejemplo anterior no sólo ilustra la gráfica \bar{X} para las variables, también proporciona al lector una idea general de la naturaleza de las gráficas de control. La línea central por lo general refleja el valor ideal de un parámetro importante. Los límites de control se establecen a partir del conocimiento de las propiedades de muestreo del estadístico que estima el parámetro en cuestión. Con mucha frecuencia implican un múltiplo de la desviación estándar del estadístico. Se ha generalizado el uso de límites 3σ . En el caso de la gráfica \bar{X} que se presenta aquí, el teorema del límite central brinda al usuario una buena aproximación de la probabilidad de determinar de forma errónea que el proceso está fuera de control. En general, sin embargo, es probable que el usuario no confíe en la normalidad del estadístico sobre la línea central. Lo anterior podría dar como resultado que no se conozca la probabilidad exacta de cometer un "error tipo I". A pesar de esto se ha vuelto muy común utilizar los límites $k\sigma$. Aunque los límites 3σ se utilizan ampliamente, en ocasiones el usuario utilizará otro método. Cuando es importante detectar de forma rápida una situación fuera de control podría ser apropiado utilizar un múltiplo menor de σ . Si se toman en cuenta los costos de producción, cabe señalar que permitir que un proceso continúe funcionando fuera de control, incluso por periodos cortos, puede resultar más costoso que invertir en la investigación y corrección de las causas de la pérdida del control en el proceso. En este caso es evidente que los límites apropiados son los límites de control que son más estrictos que los límites 3σ .

Subgrupos racionales

Los valores de la muestra que se utilizan para el control de calidad se dividen en subgrupos, en los que una *muestra* representa un subgrupo. Como antes indicamos, el orden en el tiempo de producción es en realidad una base natural para la selección de los subgrupos. Podríamos considerar el esfuerzo de control de calidad de manera muy simple como 1) muestreo, 2) detección de un estado fuera de control y 3) búsqueda de las causas atribuibles que puedan ocurrir con el tiempo. Tal vez parezca que la selección de la base para estos grupos muestrales es muy sencilla, pero la elección de estos subgrupos de información muestral podría tener un efecto importante en el éxito del programa de control de calidad. Estos subgrupos con frecuencia se denominan **subgrupos racionales**. En

general, si el analista está interesado en detectar un *cambio de ubicación*, se considera que los subgrupos se deben elegir de manera que la variabilidad dentro del subgrupo sea pequeña, y de manera que haya mayores posibilidades de detectar las causas atribuibles, si se presentaran. Así, deseamos elegir los subgrupos de forma que se maximice la variabilidad entre subgrupos. Por ejemplo, un método razonable es elegir unidades en un subgrupo que se producen de forma cercana en el tiempo. Por otro lado, las gráficas de control a menudo se utilizan para controlar la variabilidad, en cuyo caso el estadístico de desempeño es la *variabilidad dentro de la muestra*. Por consiguiente, es más importante elegir los subgrupos racionales para maximizar la variabilidad dentro de la muestra. En este caso las observaciones en los subgrupos se deberían comportar más como una muestra aleatoria y la variabilidad dentro de las muestras necesita ser una descripción de la variabilidad del proceso.

Es importante señalar que las gráficas de control sobre la variabilidad se deben establecer antes de construir gráficas sobre el centro de ubicación (digamos, gráficas \bar{X}). Cualquier gráfica de control sobre el centro de ubicación en realidad dependerá de la variabilidad. Por ejemplo, vimos un ejemplo de la gráfica de tendencia central y ésta depende de σ . En las secciones que siguen se analizará un estimado de σ a partir de los datos.

Gráfica \bar{X} con parámetros estimados

Con anterioridad ilustramos las nociones de la gráfica \bar{X} que usa el teorema del límite central y emplea valores *conocidos* de la media y desviación estándar del proceso. Como al principio se indicó, se utilizan los límites de control

$$LCI = \mu - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad LCS = \mu + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

y un valor \bar{X} que cae fuera de estos límites se considera evidencia de un cambio en la media μ , y, por lo tanto, de la posibilidad de que el proceso esté fuera de control.

En muchas situaciones prácticas no es razonable suponer que conocemos μ y σ . Como resultado, se deben proporcionar estimados de los datos que se obtienen cuando el proceso está bajo control. Por lo general los estimados se determinan durante un periodo en el que se reúne *información antecedente o de inicio*. Se elige una base para subgrupos racionales y se reúnen los datos con muestras de tamaño n en cada subgrupo. Los tamaños de la muestra por lo general son pequeños, digamos, 4, 5 o 6, y se toman k muestras, con k al menos igual a 20. Durante este periodo, en el que se supone que el proceso está bajo control, el usuario establece los estimados de μ y σ en los que se basa la gráfica de control. La información importante reunida durante este periodo incluye las medias muestrales en el subgrupo, la media general y el rango de la muestra en cada subgrupo. En los siguientes párrafos señalaremos cómo se utiliza esta información para producir la gráfica de control.

Una parte de la información muestral de estas k muestras toma la forma $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k$, donde la variable aleatoria \bar{X}_i es el promedio de los valores en la i -ésima muestra. Evidentemente, el promedio global es la variable aleatoria

$$\bar{\bar{X}} = \frac{1}{k} \sum_{i=1}^k \bar{X}_i.$$

Éste es el estimador adecuado de la media del proceso y, por consiguiente, es la línea central en la gráfica de control \bar{X} . En aplicaciones de control de calidad a menudo es

conveniente estimar σ a partir de la información relacionada con los rangos en las muestras, en vez de las desviaciones estándar muestrales. Definamos

$$R_i = X_{\text{máx}, i} - X_{\text{mín}, i}$$

como el rango para los datos en la i -ésima muestra. Aquí $X_{\text{máx}, i}$ y $X_{\text{mín}, i}$ son, respectivamente, la observación más grande y la más pequeña en la muestra. El estimado apropiado de σ es una función del rango promedio

$$\bar{R} = \frac{1}{k} \sum_{i=1}^k R_i.$$

Un estimado de σ , digamos $\hat{\sigma}$, se obtiene mediante

$$\hat{\sigma} = \frac{\bar{R}}{d_2},$$

donde d_2 es una constante que depende del tamaño de la muestra. Los valores de d_2 se muestran en la tabla A.22.

El uso del rango para producir un estimado de σ tiene sus raíces en aplicaciones similares a la del control de calidad, en particular debido a que, en la época en que aún era muy difícil lograr cálculos precisos, el rango era muy fácil de calcular en comparación con otros estimados de variabilidad. La suposición de normalidad de las observaciones individuales está implícita en la gráfica \bar{X} . Por supuesto, la existencia del teorema del límite central es ciertamente útil a este respecto. Bajo la suposición de normalidad, usamos una variable aleatoria llamada rango relativo dada por

$$W = \frac{R}{\sigma}.$$

De la cual resulta que los momentos de W son funciones simples del tamaño de la muestra n (véase la referencia de Montgomery, 2000b, en la bibliografía). El valor esperado de W a menudo se denomina d_2 . Así, al tomar el valor esperado de la W anterior,

$$\frac{E(R)}{\sigma} = d_2,$$

la cual facilita la comprensión del fundamento para el estimado $\hat{\sigma} = \bar{R}/d_2$. Se sabe bien que el método del rango produce un estimador eficiente de σ en muestras hasta cierto punto pequeñas. Esto hace que el estimador sea particularmente atractivo en aplicaciones de control de calidad, ya que los tamaños de la muestra en los subgrupos por lo general son pequeños. El uso del método del rango para la estimación de σ tiene como resultado gráficas de control con los siguientes parámetros:

$$\text{LCS} = \bar{\bar{X}} + \frac{3\bar{R}}{d_2\sqrt{n}}, \quad \text{línea central} = \bar{\bar{X}}, \quad \text{LCI} = \bar{\bar{X}} - \frac{3\bar{R}}{d_2\sqrt{n}}.$$

Al definir la cantidad

$$A_2 = \frac{3}{d_2\sqrt{n}},$$

tenemos que

$$LCS = \bar{\bar{X}} + A_2 \bar{R}, \quad LCI = \bar{\bar{X}} - A_2 \bar{R}.$$

Para simplificar la estructura el usuario de las gráficas \bar{X} a menudo encuentra valores tabulados de A_2 . En la tabla A.22 se incluyen valores de A_2 para varios tamaños de la muestra.

Gráficas R para control de variación

Hasta aquí todos los ejemplos y detalles tuvieron que ver con el intento del analista de control de calidad de detectar condiciones fuera de control producidas por un *cambio en la media*. Los límites de control se basan en la distribución de la variable aleatoria \bar{X} y dependen de la suposición de normalidad de las observaciones individuales. Es importante que el control se aplique tanto a la variabilidad como al centro de ubicación. De hecho, muchos expertos consideran que el control de la variabilidad de la característica del desempeño es más importante y que es necesario establecerlo antes de considerar el centro de ubicación. La variabilidad del proceso se puede controlar usando *gráficas del rango muestral*. Una gráfica de los rangos muestrales a lo largo del tiempo se denomina **gráfica R** . Se puede utilizar la misma estructura general, como en el caso de la gráfica \bar{X} , donde \bar{R} es la *línea central* y los límites de control dependen de que se estime la desviación estándar de la variable aleatoria R . Por lo tanto, como en el caso de la gráfica \bar{X} , se establecen límites 3σ donde “ 3σ ” implica $3\sigma_R$. La cantidad σ_R se debe estimar a partir de los datos, tal como se estima $\sigma_{\bar{X}}$.

El estimado de σ_R , la desviación estándar, también se basa en la distribución del rango relativo

$$W = \frac{R}{\sigma}.$$

La desviación estándar de W es una función conocida del tamaño de la muestra y por lo general se denota por d_3 . Esto da como resultado,

$$\sigma_R = \sigma d_3.$$

Ahora podemos reemplazar σ por $\hat{\sigma} = \bar{R}/d_2$, y de esta forma el estimador de σ_R es

$$\hat{\sigma}_R = \frac{\bar{R}d_3}{d_2}.$$

Por consiguiente, las cantidades que definen la gráfica R son

$$LCS = \bar{R}D_4, \quad \text{línea central} = \bar{R}, \quad LCI = \bar{R}D_3,$$

donde las constantes D_4 y D_3 (que dependen sólo de n) son

$$D_4 = 1 + 3 \frac{d_3}{d_2}, \quad D_3 = 1 - 3 \frac{d_3}{d_2}.$$

Las constantes D_4 y D_3 se encuentran tabuladas en la tabla A.22.

Gráficas \bar{X} y R para variables

Se controla un proceso de fabricación de partes componentes para misiles, donde la característica de desempeño es la resistencia a la tensión, en libras por pulgada cuadrada. Se toman muestras de tamaño 5 cada hora y se reportan 25 muestras. Los datos se muestran en la tabla 17.1.

Tabla 17.1: Información muestral de los datos de resistencia a la tensión

| Número de muestra | Observaciones | | | | | \bar{X}_i | R_i |
|-------------------|---------------|------|------|------|------|-------------|-------|
| 1 | 1515 | 1518 | 1512 | 1498 | 1511 | 1510.8 | 20 |
| 2 | 1504 | 1511 | 1507 | 1499 | 1502 | 1504.6 | 12 |
| 3 | 1517 | 1513 | 1504 | 1521 | 1520 | 1515.0 | 17 |
| 4 | 1497 | 1503 | 1510 | 1508 | 1502 | 1504.0 | 13 |
| 5 | 1507 | 1502 | 1497 | 1509 | 1512 | 1505.4 | 15 |
| 6 | 1519 | 1522 | 1523 | 1517 | 1511 | 1518.4 | 12 |
| 7 | 1498 | 1497 | 1507 | 1511 | 1508 | 1504.2 | 14 |
| 8 | 1511 | 1518 | 1507 | 1503 | 1509 | 1509.6 | 15 |
| 9 | 1506 | 1503 | 1498 | 1508 | 1506 | 1504.2 | 10 |
| 10 | 1503 | 1506 | 1511 | 1501 | 1500 | 1504.2 | 11 |
| 11 | 1499 | 1503 | 1507 | 1503 | 1501 | 1502.6 | 8 |
| 12 | 1507 | 1503 | 1502 | 1500 | 1501 | 1502.6 | 7 |
| 13 | 1500 | 1506 | 1501 | 1498 | 1507 | 1502.4 | 9 |
| 14 | 1501 | 1509 | 1503 | 1508 | 1503 | 1504.8 | 8 |
| 15 | 1507 | 1508 | 1502 | 1509 | 1501 | 1505.4 | 8 |
| 16 | 1511 | 1509 | 1503 | 1510 | 1507 | 1508.0 | 8 |
| 17 | 1508 | 1511 | 1513 | 1509 | 1506 | 1509.4 | 7 |
| 18 | 1508 | 1509 | 1512 | 1515 | 1519 | 1512.6 | 11 |
| 19 | 1520 | 1517 | 1519 | 1522 | 1516 | 1518.8 | 6 |
| 20 | 1506 | 1511 | 1517 | 1516 | 1508 | 1511.6 | 11 |
| 21 | 1500 | 1498 | 1503 | 1504 | 1508 | 1502.6 | 10 |
| 22 | 1511 | 1514 | 1509 | 1508 | 1506 | 1509.6 | 8 |
| 23 | 1505 | 1508 | 1500 | 1509 | 1503 | 1505.0 | 9 |
| 24 | 1501 | 1498 | 1505 | 1502 | 1505 | 1502.2 | 7 |
| 25 | 1509 | 1511 | 1507 | 1500 | 1499 | 1505.2 | 12 |

Como antes indicamos, es importante comenzar por establecer las condiciones de variabilidad "bajo control". La línea central calculada para la gráfica R es

$$\bar{R} = \frac{1}{25} \sum_{i=1}^{25} R_i = 10.72.$$

En la tabla A.22 encontramos que para $n = 5$, $D_3 = 0$ y $D_4 = 2.114$. Como resultado, los límites de control para la gráfica R son

$$LCI = \bar{R}D_3 = (10.72)(0) = 0,$$

$$LCS = \bar{R}D_4 = (10.72)(2.114) = 22.6621.$$

En la figura 17.3 se muestra la gráfica R . Ninguno de los rangos graficados cae fuera de los límites de control. Como resultado, no hay nada que indique la existencia de una situación fuera de control.

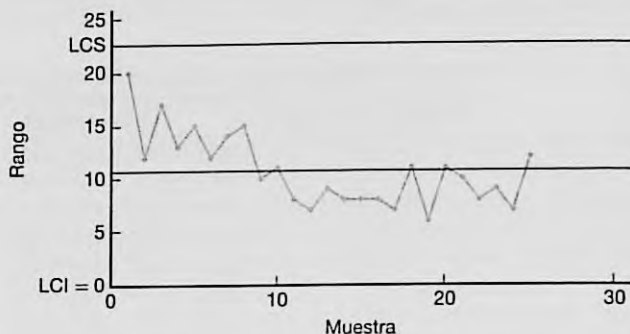


Figura 17.3: Gráfica R para el ejemplo de resistencia a la tensión.

Ahora se puede construir la gráfica \bar{X} para las lecturas de la resistencia a la tensión. La línea central es

$$\bar{\bar{X}} = \frac{1}{25} \sum_{i=1}^{25} \bar{X}_i = 1507.328.$$

En la tabla A.22 encontramos que, para muestras de tamaño 5, $A_2 = 0.577$. De esta forma, los límites de control son

$$\text{LCS} = \bar{\bar{X}} + A_2 \bar{R} = 1507.328 + (0.577)(10.72) = 1513.5134,$$

$$\text{LCI} = \bar{\bar{X}} - A_2 \bar{R} = 1507.328 - (0.577)(10.72) = 1501.1426.$$

En la figura 17.4 se muestra la gráfica \bar{X} . Como el lector puede observar, tres valores caen fuera de los límites de control, lo cual es una señal de que no se deberían usar los límites de control de \bar{X} para el control de calidad de la línea.

Más comentarios acerca de las gráficas de control para variables

Un proceso podría parecer estar bajo control y, de hecho, permanecer así durante un periodo largo. ¿Esto significaría necesariamente que el proceso está funcionando de manera exitosa? Un proceso que opera *bajo control* es simplemente aquel en el que la media y la variabilidad del proceso permanecen estables, indicando, aparentemente, que no han ocurrido cambios graves. “Bajo control” implica que el proceso permanece consistente con variabilidad natural. Las gráficas de control de calidad pueden verse como un método en el que la variabilidad natural inherente rige la amplitud de los límites de control. Sin embargo, no determinan hasta qué punto un proceso bajo control satisface las *especificaciones* predeterminadas que requiere el proceso. Las especificaciones son límites que establece el consumidor. Si la variabilidad natural del proceso actual es mayor

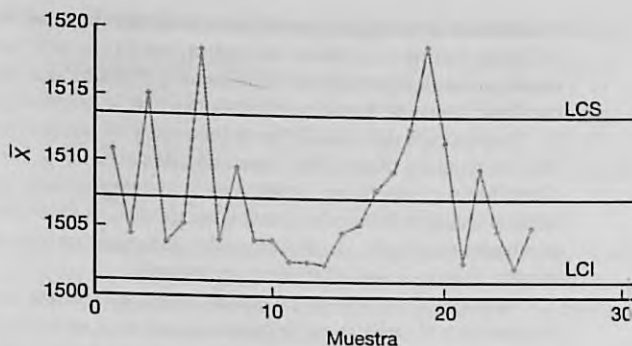


Figura 17.4: Gráfica \bar{X} para el ejemplo de resistencia a la tensión.

que la que determinan las especificaciones, aunque el proceso permanezca estable y esté bajo control, con demasiada frecuencia producirá artículos que no cumplirán las especificaciones.

Aludimos a la suposición de normalidad para las observaciones individuales en una gráfica de control de variables. Para la gráfica \bar{X} , si las observaciones individuales son normales, el estadístico \bar{X} es normal. Como resultado, el analista de control de calidad en este caso tiene control sobre la probabilidad de un error tipo I. Si las X individuales no son normales, \bar{X} es aproximadamente normal, por lo tanto, existe un control aproximado sobre la probabilidad de un error tipo I para el caso en el que se conoce σ . Sin embargo, utilizar o no el método del rango para estimar la desviación estándar también depende de la suposición de normalidad. Estudios respecto a la robustez de la gráfica \bar{X} para desviaciones de la normalidad indican que, para las muestras de tamaño $k \geq 4$, la gráfica \bar{X} da como resultado un riesgo α cercano al anunciado (véase el trabajo de Montgomery, 2000b, y Schilling y Nelson, 1976, en la bibliografía). Indicamos antes que la aproximación $\pm k\sigma_R$ a la gráfica R es una cuestión de conveniencia y tradición. Incluso si la distribución de observaciones individuales es normal, la distribución de R no es normal. De hecho, la distribución de R no es ni siquiera simétrica. Los límites de control simétricos de $\pm k\sigma_R$ sólo proporcionan una aproximación al riesgo α y, en algunos casos, la aproximación no es particularmente buena.

Elección del tamaño de la muestra (función característica de operación) en el caso de la gráfica \bar{X}

Los científicos e ingenieros que manejan el control de calidad a menudo se refieren a los factores que influyen en el *diseño de la gráfica de control*. Los componentes que determinan el diseño de la gráfica incluyen el tamaño de la muestra que se toma en cada subgrupo, la amplitud de los límites de control y la frecuencia del muestreo. Todos estos factores dependen en gran medida de consideraciones económicas y prácticas. La frecuencia de muestreo evidentemente depende del costo del muestreo y del costo en el que se incurre si el proceso continúa fuera de control durante un periodo largo. Estos mismos factores afectan la amplitud de la región "bajo control". El costo asociado con la investigación y la búsqueda de las causas atribuibles de la pérdida de control repercute en

la amplitud de la región y en la frecuencia de muestreo. Se ha puesto mucha atención en el diseño óptimo de gráficas de control, por lo que aquí no se darán mayores detalles. Se remite al lector al trabajo de Montgomery (2000b), que se cita en la bibliografía, para un excelente recuento histórico de gran parte de esta investigación.

La elección del tamaño de la muestra y la frecuencia de muestreo implican equilibrar los recursos disponibles para estos dos esfuerzos. En muchos casos es probable que el analista necesite hacer cambios en la estrategia hasta lograr el equilibrio adecuado. El analista siempre debe estar consciente de que, si el costo de producción de artículos no adecuados es grande, la estrategia adecuada sería una alta frecuencia de muestreo con un tamaño de la muestra relativamente pequeño.

Al elegir el tamaño de una muestra hay que tomar en cuenta muchos factores. En la ilustración y el análisis enfatizamos el uso de $n = 4, 5$ o 6 . Estos valores se consideran relativamente pequeños para problemas generales en inferencia estadística, pero serían tamaños de muestra apropiados para el control de calidad. Una justificación, por supuesto, es que el control de calidad es un proceso continuo y los resultados producidos por una muestra o un conjunto de unidades serán seguidos por resultados de muchas más. Así, el tamaño de la muestra "eficaz" de todo el esfuerzo de control de calidad es muchas veces mayor que el tamaño que se utiliza en un subgrupo. Por lo general se considera más efectivo *tomar muestras frecuentemente* con un tamaño muestral pequeño.

El analista puede utilizar el concepto de potencia de una prueba para obtener información de la eficacia del tamaño de la muestra elegido. Esto es especialmente importante, ya que por lo general se utilizan muestras de tamaño pequeño en cada subgrupo. Remítase a los capítulos 10 y 13 para un análisis de la potencia de pruebas formales sobre las medias y el análisis de varianza. Aunque en el control de calidad en realidad no se realizan pruebas formales de hipótesis, se puede tratar la información como si la estrategia en cada subgrupo fuera la de probar una hipótesis, ya sea sobre la media de la población μ o sobre la desviación estándar σ . Es de interés la *probabilidad de detectar* una condición fuera de control para una muestra dada y, quizá más importante, el número esperado de corridas requeridas para detectarla. La probabilidad de detectar una condición fuera de control específica corresponde a la potencia de una prueba. No es nuestra intención demostrar el desarrollo de la potencia para todos los tipos de gráficas de control que aquí se presentan, más bien, lo que deseamos es mostrar el desarrollo de la gráfica \bar{X} y presentar los resultados de potencia para la gráfica R .

Considere la gráfica \bar{X} cuando se conoce el valor de σ . Suponga que el estado bajo control tiene $\mu = \mu_0$. Un estudio del papel que desempeña el tamaño de la muestra del subgrupo equivale a investigar el riesgo β , es decir, la probabilidad de que un valor \bar{X} permanezca dentro de los límites de control cuando realmente ha ocurrido un cambio en la media. Suponga que la forma que toma el cambio es

$$\mu = \mu_0 + r\sigma.$$

De nuevo, al utilizar la normalidad de \bar{X} tenemos

$$\beta = P(\text{LCI} \leq \bar{X} \leq \text{LCS} \mid \mu = \mu_0 + r\sigma).$$

Para el caso de límites $k\sigma$,

$$\text{LCI} = \mu_0 - \frac{k\sigma}{\sqrt{n}} \quad \text{y} \quad \text{LCS} = \mu_0 + \frac{k\sigma}{\sqrt{n}}.$$

Como resultado, si denotamos con Z la variable aleatoria normal estándar

$$\begin{aligned}\beta &= P \left[Z < \left(\frac{\mu_0 + k\sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}} \right) \right] - P \left[Z < \left(\frac{\mu_0 - k\sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}} \right) \right] \\ &= P \left\{ Z < \left[\frac{\mu_0 + k\sigma/\sqrt{n} - (\mu + r\sigma)}{\sigma/\sqrt{n}} \right] \right\} - P \left\{ Z < \left[\frac{\mu_0 - k\sigma/\sqrt{n} - (\mu + r\sigma)}{\sigma/\sqrt{n}} \right] \right\} \\ &= P(Z < k - r\sqrt{n}) - P(Z < -k - r\sqrt{n}).\end{aligned}$$

Observe el papel que desempeñan n , r y k en la expresión para el riesgo β . La probabilidad de no detectar un cambio específico, como se esperaba, aumenta claramente con un incremento en k . β disminuye con un aumento en r , la magnitud del cambio, y disminuye con un incremento en el tamaño de la muestra, n .

Se debería enfatizar que la expresión anterior da como resultado el riesgo β (probabilidad de un error tipo II) para el caso de una *sola muestra*. Por ejemplo, suponga que, en el caso de una muestra de tamaño 4, ocurre un cambio de σ en la media. La probabilidad de detectar el cambio (potencia) en la primera muestra después del cambio es, suponiendo límites 3σ :

$$1 - \beta = 1 - [P(Z < 1) - P(Z < -5)] = 0.1587.$$

Por otro lado, la probabilidad de detectar un cambio de 2σ es

$$1 - \beta = 1 - [P(Z < -1) - P(Z < -7)] = 0.8413.$$

Los resultados anteriores ilustran una muy modesta probabilidad de detectar un cambio de magnitud σ y una alta probabilidad de detectar un cambio de magnitud 2σ . En la figura 17.5 se observa la imagen completa de cómo se desempeñan los límites de control 3σ para la gráfica \bar{X} que aquí se describe. En lugar de graficar las funciones de potencia se presenta una gráfica de β contra r , donde el cambio en la media tiene una magnitud $r\sigma$. Por supuesto, los tamaños de la muestra de $n = 4, 5, 6$ dan como resultado una pequeña probabilidad de detectar un cambio de 1.0σ o incluso 1.5σ en la primera muestra después del cambio.

Pero si el muestreo se realiza con frecuencia, la probabilidad podría no ser tan importante como el número promedio o esperado de corridas que se requiere antes de detectar un cambio. Una detección rápida es importante y ciertamente posible, aunque no hay muchas probabilidades de lograrlo en la primera muestra. Resulta que las gráficas \bar{X} con estas muestras pequeñas conducirán a una detección relativamente rápida. Si β es la probabilidad de no detectar un cambio en la primera muestra después del cambio, entonces la probabilidad de detectarlo en la muestra s -ésima después de que ocurre es, suponiendo que las muestras son independientes:

$$P_s = (1 - \beta)\beta^{s-1}.$$

El lector debe reconocer que ésta es una aplicación de la distribución geométrica. El valor promedio o esperado del número de muestras que se requieren para la detección es

$$\sum_{s=1}^{\infty} s\beta^{s-1}(1 - \beta) = \frac{1}{1 - \beta}.$$

Por consiguiente, el número esperado de muestras que se requieren para detectar el cambio en la media es el *recíproco de la potencia*, es decir, la probabilidad de detección en la primera muestra después del cambio.

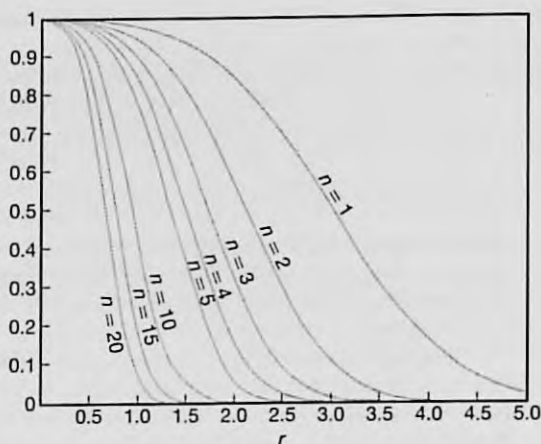


Figura 17.5: Curvas características de operación para la gráfica \bar{X} con límites 3σ . Aquí, β es el error de probabilidad tipo II en la primera muestra después de que ocurre un cambio en la media de $r\sigma$.

Ejemplo 17.1: En cierto esfuerzo por controlar la calidad es importante que el analista detecte con rapidez los cambios en la media de $\pm\sigma$ utilizando una gráfica de control 3σ con una muestra de tamaño $n = 4$. El número esperado de muestras que se requieren después del cambio para detectar el estado fuera de control podría ser útil en la evaluación del procedimiento de control de calidad.

En la figura 17.5, para $n = 4$ y $r = 1$, se puede ver que $\beta \approx 0.84$. Si utilizamos s para denotar el número de muestras que se requieren para detectar el cambio, la media de s es

$$E(s) = \frac{1}{1 - \beta} = \frac{1}{0.16} = 6.25.$$

De esta manera, se requieren siete subgrupos, en promedio, antes de detectar un cambio de $\pm\sigma$. ▮

Elección del tamaño de la muestra para la gráfica R

La curva CO de la gráfica R se muestra en la figura 17.6. Como la gráfica R se utiliza para controlar la desviación estándar del proceso, y la desviación estándar después de que el proceso se sale de control, el riesgo β se grafica como una función de la desviación estándar bajo control, σ_0 . La última desviación estándar se denotará con σ_1 . Sea

$$\lambda = \frac{\sigma_1}{\sigma_0}.$$

Para varios tamaños muestrales se grafica β contra λ .

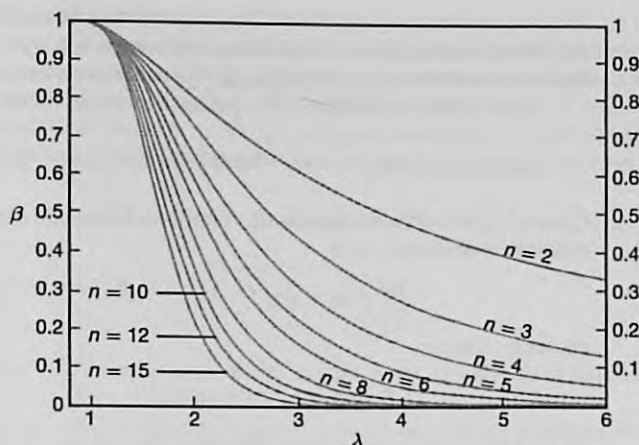


Figura 17.6: Curva característica operativa para las gráficas R con límites de 3σ .

Gráficas \bar{X} y S para variables

Para el estudiante de estadística es natural anticipar el uso de la varianza muestral en la gráfica \bar{X} y en una gráfica para el control de la variabilidad. El rango es un estimador eficiente de σ , pero esta eficiencia disminuye a medida que aumenta el tamaño de la muestra. Para una n tan grande como 10 se debe utilizar el tan conocido estadístico

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

en la gráfica de control, tanto para la media como para la variabilidad. El lector debe recordar que en el capítulo 9 se expuso que S^2 es un estimador no sesgado de σ^2 , pero que S no es no sesgado para σ . Para evitar sesgos se acostumbra corregir S en las aplicaciones de la gráfica de control. Sabemos, en general, que

$$E(S) \neq \sigma.$$

En el caso en que las X_i sean independientes y estén distribuidas de forma normal con media μ y varianza σ^2 ,

$$E(S) = c_4 \sigma, \quad \text{donde} \quad c_4 = \left(\frac{2}{n-1}\right)^{1/2} \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]}$$

y $\Gamma(\cdot)$ se refiere a la función gamma (véase el capítulo 6). Por ejemplo, para $n = 5$, $c_4 = (3/8)\sqrt{2\pi}$. Además, la varianza del estimador S es

$$\text{Var}(S) = \sigma^2(1 - c_4^2).$$

Establezcimos las propiedades de S que nos permitirán escribir límites de control para \bar{X} y S . Para construir una estructura adecuada comenzamos por suponer que conocemos σ . Después presentamos la estimación de σ a partir de un conjunto de muestras preliminar.

Al graficar el estadístico S , los parámetros evidentes de la gráfica de control son

$$LCS = c_4\sigma + 3\sigma\sqrt{1 - c_3^2}, \quad \text{línea central} = c_4\sigma, \quad LCI = c_4\sigma - 3\sigma\sqrt{1 - c_3^2}.$$

Como de costumbre, los límites de control se definen de manera más sucinta utilizando constantes tabuladas. Sean

$$B_5 = c_4 - 3\sqrt{1 - c_3^2}, \quad B_6 = c_4 + 3\sqrt{1 - c_3^2},$$

entonces, tenemos

$$LCS = B_6\sigma, \quad \text{línea central} = c_4\sigma, \quad LCI = B_5\sigma.$$

En la tabla A.22 se encuentran tabulados los valores de B_5 y B_6 para varios tamaños muestrales.

Ahora, por supuesto, los límites de control anteriores sirven como base para el desarrollo de los parámetros de control de calidad en la situación que con más frecuencia se observa en la práctica, a saber, en la que se desconoce σ . Debemos suponer una vez más que para producir un estimado de σ durante lo que se supone es un periodo "bajo control" se toma un conjunto de *muestras base* o muestras preliminares. Las desviaciones estándar de las muestras S_1, S_2, \dots, S_m se obtienen a partir de muestras que son, cada una, de tamaño n . A menudo se utiliza un estimador no sesgado del tipo

$$\frac{\bar{S}}{c_4} = \left(\frac{1}{m} \sum_{i=1}^m S_i \right) / c_4$$

para σ . Aquí, desde luego, \bar{S} , el valor promedio de la desviación estándar muestral en la muestra preliminar, es la línea central lógica en la gráfica de control para el control de la variabilidad. Los límites de control superior e inferior son estimadores no sesgados de los límites de control adecuados para el caso en el que se conoce σ . Como

$$E\left(\frac{\bar{S}}{c_4}\right) = \sigma,$$

el estadístico \bar{S} es una línea central apropiada (como un estimador no sesgado de $c_4\sigma$) y las cantidades

$$\bar{S} - 3\frac{\bar{S}}{c_4}\sqrt{1 - c_3^2} \quad \text{y} \quad \bar{S} + 3\frac{\bar{S}}{c_4}\sqrt{1 - c_3^2}$$

son los límites de control 3σ inferior y superior apropiados, respectivamente. Como resultado, la línea central y los límites para la gráfica S de control de variabilidad son

$$LCI = B_3\bar{S}, \quad \text{línea central} = \bar{S}, \quad LCS = B_4\bar{S},$$

donde

$$B_3 = 1 - \frac{3}{c_4}\sqrt{1 - c_3^2}, \quad B_4 = 1 + \frac{3}{c_4}\sqrt{1 - c_3^2}.$$

las constantes B_3 y B_4 aparecen en la tabla A.22.

Ahora podemos escribir los parámetros de la gráfica \bar{X} correspondiente que implican el uso de la desviación estándar muestral. Supongamos que podemos disponer de S y \bar{X} de la muestra base preliminar. La línea central continúa siendo $\bar{\bar{X}}$ y los límites 3σ son simplemente de la forma $\bar{\bar{X}} \pm 3\hat{\sigma}/\sqrt{n}$, donde $\hat{\sigma}$ es un estimador insesgado. Simplemente proporcionamos \bar{S}/c_4 como un estimador de σ , y de esta forma tenemos

$$\text{LCI} = \bar{\bar{X}} - A_3 \bar{S}, \quad \text{línea central} = \bar{\bar{X}}, \quad \text{LCS} = \bar{\bar{X}} + A_3 \bar{S},$$

donde

$$A_3 = \frac{3}{c_4 \sqrt{n}}.$$

En la tabla A.22 aparece la constante A_3 para varios tamaños de la muestra.

Ejemplo 17.2: Se producen contenedores mediante un proceso en el que el volumen de éstos es sometido a un control de calidad. Se utilizaron 25 muestras de tamaño 5 para establecer los parámetros de control de calidad. En la tabla 17.2 se documenta la información de estas muestras.

En la tabla A.22 se observa que $B_3 = 0$, $B_4 = 2.089$ y $A_3 = 1.427$. Como resultado, los límites de control para \bar{X} son dados por

$$\text{LCS} = \bar{\bar{X}} + A_3 \bar{S} = 62.3771, \quad \text{LCI} = \bar{\bar{X}} - A_3 \bar{S} = 62.2741,$$

y los límites de control para la gráfica S son

$$\text{LCI} = B_3 \bar{S} = 0, \quad \text{LCS} = B_4 \bar{S} = 0.0754.$$

Las figuras 17.7 y 17.8 muestran las gráficas de control para este ejemplo, \bar{X} y S , respectivamente. En las gráficas se representa la información de las 25 muestras en el conjunto de datos preliminar. Al parecer, el control se establece después de las primeras muestras. ▮

17.5 Gráficas de control para atributos

Como indicamos al principio de este capítulo, muchas aplicaciones industriales de control de calidad requieren que la característica de calidad indique sólo que el artículo "se ajusta". En otras palabras, no hay una medición continua que sea crucial para el desempeño del artículo. Una ilustración evidente de este tipo de muestreo, denominado **muestreo por atributos**, es el desempeño de una bombilla que funciona o no de manera satisfactoria. El artículo **está o no defectuoso**. Las piezas metálicas fabricadas pueden tener deformaciones; los contenedores de una línea de producción pueden tener fugas. En ambos casos un artículo defectuoso impide su uso por parte del consumidor. La gráfica de control estándar para esta situación es la gráfica p , o *gráfica para la fracción de defectuosos*. Como se podría esperar, la distribución de probabilidad que interviene es la distribución binomial. Se remite al lector al capítulo 5 para información básica de la distribución binomial.

Tabla 17.2: Volumen de contenedores para 25 muestras en una muestra preliminar (en centímetros cúbicos)

| Muestra | Observaciones | | | | \bar{X}_i | S_i | |
|---------|---------------|--------|--------|--------|-------------|---------------------------|--------|
| 1 | 62.255 | 62.301 | 62.289 | 62.189 | 62.311 | 62.269 | 0.0495 |
| 2 | 62.187 | 62.225 | 62.337 | 62.297 | 62.307 | 62.271 | 0.0622 |
| 3 | 62.421 | 62.377 | 62.257 | 62.295 | 62.222 | 62.314 | 0.0829 |
| 4 | 62.301 | 62.315 | 62.293 | 62.317 | 62.409 | 62.327 | 0.0469 |
| 5 | 62.400 | 62.375 | 62.295 | 62.272 | 62.372 | 62.343 | 0.0558 |
| 6 | 62.372 | 62.275 | 62.315 | 62.372 | 62.302 | 62.327 | 0.0434 |
| 7 | 62.297 | 62.303 | 62.337 | 62.392 | 62.344 | 62.335 | 0.0381 |
| 8 | 62.325 | 62.362 | 62.351 | 62.371 | 62.397 | 62.361 | 0.0264 |
| 9 | 62.327 | 62.297 | 62.318 | 62.342 | 62.318 | 62.320 | 0.0163 |
| 10 | 62.297 | 62.325 | 62.303 | 62.307 | 62.333 | 62.313 | 0.0153 |
| 11 | 62.315 | 62.366 | 62.308 | 62.318 | 62.319 | 62.325 | 0.0232 |
| 12 | 62.297 | 62.322 | 62.344 | 62.342 | 62.313 | 62.324 | 0.0198 |
| 13 | 62.375 | 62.287 | 62.362 | 62.319 | 62.382 | 62.345 | 0.0406 |
| 14 | 62.317 | 62.321 | 62.297 | 62.372 | 62.319 | 62.325 | 0.0279 |
| 15 | 62.299 | 62.307 | 62.383 | 62.341 | 62.394 | 62.345 | 0.0431 |
| 16 | 62.308 | 62.319 | 62.344 | 62.319 | 62.378 | 62.334 | 0.0281 |
| 17 | 62.319 | 62.357 | 62.277 | 62.315 | 62.295 | 62.313 | 0.0300 |
| 18 | 62.333 | 62.362 | 62.292 | 62.327 | 62.314 | 62.326 | 0.0257 |
| 19 | 62.313 | 62.387 | 62.315 | 62.318 | 62.341 | 62.335 | 0.0313 |
| 20 | 62.375 | 62.321 | 62.354 | 62.342 | 62.375 | 62.353 | 0.0230 |
| 21 | 62.399 | 62.308 | 62.292 | 62.372 | 62.299 | 62.334 | 0.0483 |
| 22 | 62.309 | 62.403 | 62.318 | 62.295 | 62.317 | 62.328 | 0.0427 |
| 23 | 62.293 | 62.293 | 62.342 | 62.315 | 62.349 | 62.318 | 0.0264 |
| 24 | 62.388 | 62.308 | 62.315 | 62.392 | 62.303 | 62.341 | 0.0448 |
| 25 | 62.324 | 62.318 | 62.315 | 62.295 | 62.319 | 62.314 | 0.0111 |
| | | | | | | $\bar{\bar{X}} = 62.3256$ | |
| | | | | | | $\bar{S} = 0.0361$ | |

Gráfica p para la fracción de artículos defectuosos

Cualquier artículo fabricado puede tener varias características que son importantes y deben ser examinadas por un inspector. Sin embargo, todo el procedimiento se enfoca aquí en una sola característica. Suponga que para todos los artículos la probabilidad de encontrar uno defectuoso es p , y que todos los artículos se producen de forma independiente. Entonces, en una muestra aleatoria de n artículos producidos, con X como el número de artículos defectuosos, tenemos

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n.$$

Como se podría suponer, la media y varianza de la variable aleatoria binomial desempeñarán un papel importante en el desarrollo de la gráfica de control. El lector debería recordar que

$$E(X) = np \quad \text{y} \quad \text{Var}(X) = np(1-p).$$

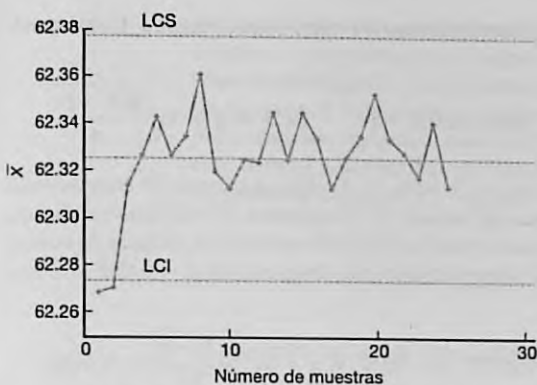


Figura 17.7: Gráfica \bar{X} con límites de control establecidos con los datos del ejemplo 17.2.

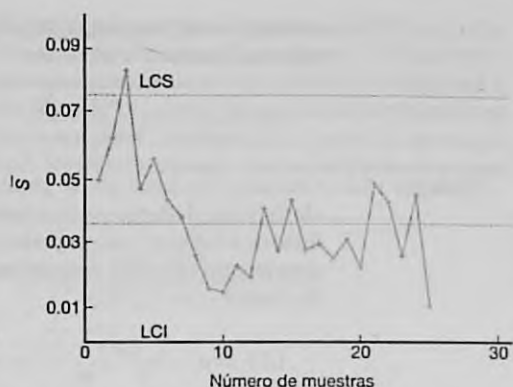


Figura 17.8: Gráfica \bar{S} con límites de control establecidos con los datos del ejemplo 17.2.

Un estimador no sesgado de p es la **fracción de defectuosos** o la **proporción de defectuosos**, \hat{p} , donde

$$\hat{p} = \frac{\text{número de defectuosos en la muestra de tamaño } n}{n}$$

Como en el caso de las gráficas de control de variables, las propiedades de distribución de p son importantes para la creación de la gráfica de control. Sabemos que

$$E(\hat{p}) = p, \quad \text{Var}(\hat{p}) = \frac{p(1-p)}{n}$$

Aquí aplicamos los mismos principios 3σ que utilizamos para las gráficas de variables. Supongamos inicialmente que conocemos p . Entonces, la estructura de las gráficas de control implica utilizar límites 3σ con

$$\hat{\sigma} = \sqrt{\frac{p(1-p)}{n}}$$

De esta manera, los límites son

$$\text{LCI} = p - 3\sqrt{\frac{p(1-p)}{n}}, \quad \text{LCS} = p + 3\sqrt{\frac{p(1-p)}{n}}$$

con el proceso considerado bajo control cuando los valores \hat{p} de la muestra caen dentro de los límites de control.

En general, por supuesto, no se conoce el valor de p y se debe estimar a partir de un conjunto base de muestras de forma muy similar al caso de μ y σ en las gráficas de variables. Suponga que hay m muestras preliminares de tamaño n . Para una muestra dada, cada una de las n observaciones se reporta como “defectuosa” o “no defectuosa”. El estimador no sesgado evidente para p que se utiliza en la gráfica de control es

$$\bar{p} = \frac{1}{m} \sum_{i=1}^m \hat{p}_i$$

donde \hat{p}_i es la proporción de artículos defectuosos en la i -ésima muestra. Como resultado, los límites de control son

$$LCI = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}, \quad \text{línea central} = \bar{p}, \quad LCS = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}.$$

Ejemplo 17.3: Considere los datos que se presentan en la tabla 17.3 sobre el número de componentes electrónicos defectuosos en muestras de tamaño 50. Se tomaron 20 muestras con la finalidad de establecer valores preliminares para la gráfica de control. Las gráficas de control determinadas por este periodo preliminar tendrán una línea central $\hat{p} = 0.088$ y límites de control

$$LCI = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{50}} = -0.0322 \quad \text{y} \quad LCS = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{50}} = 0.2082.$$

Tabla 17.3: Datos para el ejemplo 17.3 que permiten establecer límites de control en gráficas p , con muestras de tamaño 50

| Muestra | Número de componentes defectuosos | Fracción de defectuosos \hat{p}_i |
|---------|-----------------------------------|-------------------------------------|
| 1 | 8 | 0.16 |
| 2 | 6 | 0.12 |
| 3 | 5 | 0.10 |
| 4 | 7 | 0.14 |
| 5 | 2 | 0.04 |
| 6 | 5 | 0.10 |
| 7 | 3 | 0.06 |
| 8 | 8 | 0.16 |
| 9 | 4 | 0.08 |
| 10 | 4 | 0.08 |
| 11 | 3 | 0.06 |
| 12 | 1 | 0.02 |
| 13 | 5 | 0.10 |
| 14 | 4 | 0.08 |
| 15 | 4 | 0.08 |
| 16 | 2 | 0.04 |
| 17 | 3 | 0.06 |
| 18 | 5 | 0.10 |
| 19 | 6 | 0.12 |
| 20 | 3 | 0.06 |
| | | $\bar{p} = 0.088$ |

Evidentemente, con un valor calculado negativo, el LCI se ajusta a cero. A partir de los valores de los límites de control al parecer el proceso está bajo control durante este periodo preliminar. ▮

Selección del tamaño de la muestra para la gráfica p

La elección del tamaño de la muestra para la gráfica p de atributos incluye los mismos tipos generales de consideraciones que los de la gráfica para variables. Se requiere un tamaño de la muestra tan grande como para tener una alta probabilidad de detectar una

condición fuera de control cuando, de hecho, ha ocurrido un cambio específico en p . No existe un *mejor método* para elegir el tamaño de la muestra. Sin embargo, Duncan (1986; véase la bibliografía) sugirió una aproximación razonable que consiste en elegir una n tal que haya 0.5 de probabilidades de detectar un cambio de una cantidad particular en p . La solución resultante para n es bastante simple. Suponga que se aplica la aproximación normal a la distribución binomial. Deseamos, siempre que la condición de p haya cambiado a, digamos, $p_1 > p_0$, que

$$P(\hat{p} \geq \text{LCS}) = P\left[Z \geq \frac{\text{LCS} - p_1}{\sqrt{p_1(1-p_1)/n}}\right] = 0.5.$$

Como $P(Z > 0) = 0.5$, se establece

$$\frac{\text{LCS} - p_1}{\sqrt{p_1(1-p_1)/n}} = 0.$$

Al sustituir,

$$p + 3\sqrt{\frac{p(1-p)}{n}} = \text{LCS},$$

tenemos

$$(p - p_1) + 3\sqrt{\frac{p(1-p)}{n}} = 0.$$

Ahora podemos calcular n , el tamaño de cada muestra:

$$n = \frac{9}{\Delta^2} p(1-p),$$

donde, desde luego, Δ es el "cambio" en el valor de p , y p es la probabilidad de un artículo defectuoso sobre la que se basan los límites de control. Sin embargo, si las gráficas de control se basan en límites $k\sigma$, entonces

$$n = \frac{k^2}{\Delta^2} p(1-p).$$

Ejemplo 17.4: Suponga que se diseña una gráfica de control de calidad de atributos con un valor de $p = 0.01$ para la probabilidad de tener bajo control un artículo defectuoso. ¿Cuál es el tamaño de la muestra por subgrupo que produce una probabilidad de 0.5 de que se detecte un cambio en el proceso para $p = p_1 = 0.05$? La gráfica p resultante incluirá límites 3σ .

Solución: Aquí tenemos $\Delta = 0.04$. El tamaño adecuado de la muestra es

$$n = \frac{9}{(0.04)^2} (0.01)(0.99) = 55.69 \approx 56.$$

Gráficas de control para artículos defectuosos (uso del modelo de Poisson)

En el procedimiento anterior supusimos que el artículo bajo consideración es uno que está defectuoso, es decir, que no funciona, o uno que no tiene defecto, en cuyo caso el artículo funciona y, por lo tanto, es aceptable para el consumidor. En muchas situaciones este método del artículo "defectuoso o no" es demasiado simplista. Las unidades pueden contener defectos o no cumplir con las especificaciones, y aun así funcionar bastante bien para el consumidor. En realidad, en este caso sería importante ejercer control sobre el número de defectos o número de artículos que no cumplen las especificaciones. Este tipo de control de calidad tiene aplicación cuando las unidades no son simplistas ni grandes. Por ejemplo, el número de defectos puede ser muy útil como objeto de control cuando el artículo o unidad es, digamos, una computadora personal. Otro ejemplo es una unidad definida por 50 pies de tubería fabricada, donde el número de soldaduras defectuosas es el objeto del control de calidad; el número de defectos en 50 pies de alfombra fabricada o el número de "burbujas" en una hoja grande de vidrio fabricado.

A partir de lo aquí descrito queda claro que en este caso no es apropiada la distribución binomial. El número total de artículos que no cumplen las especificaciones en una unidad o el número promedio por unidad se podría usar como la medida para la gráfica de control. A menudo se supone que el número de artículos que no cumplen las especificaciones en una muestra tiene una distribución de Poisson. A este tipo de gráfica con frecuencia se le llama **gráfica C**.

Suponga que el número de defectos X en una unidad de producto tiene una distribución de Poisson con parámetro λ . (Aquí $t = 1$ para el modelo de Poisson). Recuerde que para la distribución de Poisson,

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

Aquí, la variable aleatoria X es el número de artículos que no cumplen las especificaciones. En el capítulo 5 vimos que tanto la media como la varianza de la variable aleatoria de Poisson son λ . Por consiguiente, si la gráfica de control de calidad se estructurara de acuerdo con los límites 3σ acostumbrados, si conociéramos λ tendríamos,

$$LCS = \lambda + 3\sqrt{\lambda}, \quad \text{línea central} = \lambda, \quad LCI = \lambda - 3\sqrt{\lambda}.$$

Como de costumbre, λ a menudo debe provenir de un estimador de los datos. Un estimado no sesgado de λ es el número *promedio* de artículos que no cumplen las especificaciones por muestra. Este estimado se denota mediante $\hat{\lambda}$. Así, la gráfica de control tiene los límites

$$LCS = \hat{\lambda} + 3\sqrt{\hat{\lambda}}, \quad \text{línea central} = \hat{\lambda}, \quad LCI = \hat{\lambda} - 3\sqrt{\hat{\lambda}}.$$

Ejemplo 17.5: La tabla 17.4 representa el número de defectos en 20 muestras sucesivas de rollos de hoja metálica, cada uno con 100 pies de largo. Para controlar el número de defectos en tales muestras se debe desarrollar una gráfica de control a partir de estos datos preliminares. El estimado del parámetro de Poisson λ es dado por $\hat{\lambda} = 5.95$. Como resultado, los límites de control sugeridos por estos datos preliminares son

$$LCS = \hat{\lambda} + 3\sqrt{\hat{\lambda}} = 13.2678 \quad \text{y} \quad LCI = \hat{\lambda} - 3\sqrt{\hat{\lambda}} = -1.3678,$$

donde LCI se iguala a cero.

Tabla 17.4: Datos para el ejemplo 17.5; el control implica el número de defectos en rollos de hojas metálicas

| Número de muestra | Número de defectos | Número de muestra | Número de defectos |
|-------------------|--------------------|-------------------|--------------------|
| 1 | 8 | 11 | 3 |
| 2 | 7 | 12 | 7 |
| 3 | 5 | 13 | 5 |
| 4 | 4 | 14 | 9 |
| 5 | 4 | 15 | 7 |
| 6 | 7 | 16 | 7 |
| 7 | 6 | 17 | 8 |
| 8 | 4 | 18 | 6 |
| 9 | 5 | 19 | 7 |
| 10 | 6 | 20 | 4 |
| | | Prom. 5.95 | |

La figura 17.9 presenta una gráfica de los datos preliminares con los límites de control.

La tabla 17.5 incluye datos adicionales tomados del proceso de producción. Para cada muestra se inspeccionó la unidad en la que se basó la gráfica, a saber, 100 pies del metal. Se incluye la información de 20 muestras. La figura 17.10 muestra una gráfica de los datos adicionales de producción. Es evidente que el proceso está bajo control, o al menos lo estaba en el periodo en el que se tomaron los datos. ■

Tabla 17.5: Datos adicionales del proceso de producción del ejemplo 17.5

| Número de muestra | Número de defectos | Número de muestra | Número de defectos |
|-------------------|--------------------|-------------------|--------------------|
| 1 | 3 | 11 | 7 |
| 2 | 5 | 12 | 5 |
| 3 | 8 | 13 | 9 |
| 4 | 5 | 14 | 4 |
| 5 | 8 | 15 | 6 |
| 6 | 4 | 16 | 5 |
| 7 | 3 | 17 | 3 |
| 8 | 6 | 18 | 2 |
| 9 | 5 | 19 | 1 |
| 10 | 2 | 20 | 6 |

En el ejemplo 17.5 dejamos muy claro que la unidad de muestreo o de inspección son 100 pies de metal. En muchos casos en los que el artículo es específico, como en el caso de una computadora personal o el de un tipo específico de dispositivo electrónico, la unidad de inspección podría ser un *conjunto de artículos*. Por ejemplo, el analista decide utilizar 10 computadoras en cada subgrupo y de esta forma observar un conteo del número total de defectos encontrados. Por consiguiente, la muestra preliminar para construir la gráfica de control implica utilizar varias muestras, cada una de 10 computadoras. La elección del tamaño de la muestra puede depender de muchos factores. A menudo deseamos un tamaño de la muestra que asegure un LCI positivo.

El analista podría utilizar el número promedio de defectos por unidad de muestreo como la medida básica de la gráfica de control. Por ejemplo, para el caso de la compu-

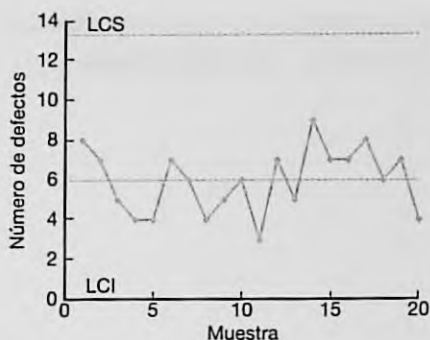


Figura 17.9: Datos preliminares representados en la gráfica de control para el ejemplo 17.5.

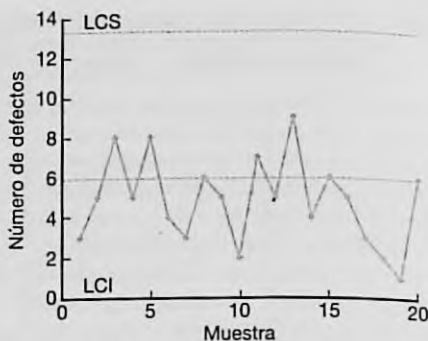


Figura 17.10: Datos adicionales de producción para el ejemplo 17.5.

tadora personal, sea la variable aleatoria el número total de defectos

$$U = \frac{\text{número total de defectos}}{n}$$

que se mide para cada muestra de, digamos, $n = 10$. Si suponemos que el número de defectos por unidad de muestreo es de Poisson con parámetro λ , podemos utilizar el método de las funciones generadoras de momento para demostrar que U es una variable aleatoria de Poisson (véase el ejercicio de repaso 17.1). De esta manera, la gráfica de control para esta situación se caracteriza por lo siguiente:

$$\text{LCS} = \bar{U} + 3\sqrt{\frac{\bar{U}}{n}}, \quad \text{línea central} = \bar{U}, \quad \text{LCI} = \bar{U} - 3\sqrt{\frac{\bar{U}}{n}}.$$

Aquí, desde luego, \bar{U} es el promedio de los valores U en el conjunto de datos preliminares o base. El término \bar{U}/n se deriva del resultado que

$$E(U) = \lambda, \quad \text{Var}(U) = \frac{\lambda}{n},$$

y por ello \bar{U} es un estimado no sesgado de $E(U) = \lambda$ y \bar{U}/n es un estimado no sesgado de $\text{Var}(U) = \lambda/n$. Este tipo de gráfica de control a menudo se denomina **gráfica U** .

En esta sección basamos toda la explicación de las gráficas de control en el modelo de probabilidad de Poisson. Este modelo se ha utilizado en combinación con el concepto 3σ . Como explicamos antes en este capítulo, el concepto de límites 3σ tiene sus raíces en la aproximación normal, aunque muchos usuarios consideran que el concepto funciona bien como herramienta pragmática incluso si la normalidad no es siquiera aproximadamente correcta. La dificultad, desde luego, radica en el hecho de que, en ausencia de normalidad, no es posible controlar la probabilidad de una especificación incorrecta de un estado fuera de control. En el caso del modelo de Poisson, cuando λ es pequeña la distribución es bastante asimétrica, una condición que puede producir resultados indeseables si se utiliza el método 3σ .

17.6 Gráficas de control de cusum

La desventaja de las gráficas de control similares a las de Shewhart, que se explicaron y ejemplificaron en las secciones anteriores, radica en su incapacidad para detectar pequeños cambios en la media. Un mecanismo de control de calidad que ha recibido mucha atención en la literatura estadística y que se ha utilizado extensamente en la industria es la **gráfica de suma acumulada (cusum)**. El método de la gráfica de suma acumulada es sencillo y, por lo tanto, atractivo. Para el lector debe ser evidente por qué es más sensible a pequeños cambios en la media. Considere una gráfica de control para la media con un nivel de referencia establecido en el valor W . Considere las observaciones particulares X_1, X_2, \dots, X_r . Las primeras cusum r son

$$\begin{aligned} S_1 &= X_1 - W \\ S_2 &= S_1 + (X_2 - W) \\ S_3 &= S_2 + (X_3 - W) \\ &\vdots \\ S_r &= S_{r-1} + (X_r - W). \end{aligned}$$

Es evidente que la cusum es simplemente la acumulación de las diferencias del nivel de referencia. Es decir,

$$S_k = \sum_{i=1}^k (X_i - W), \quad k = 1, 2, \dots$$

La gráfica cusum es, entonces, una gráfica de S_k contra el tiempo.

Suponga que consideramos que el nivel de referencia W es un valor aceptable de la media μ . Salta a la vista que, si no hay cambio en μ , la gráfica cusum debería ser aproximadamente horizontal, con algunas fluctuaciones menores balanceadas alrededor de cero. Ahora, si sólo hay un cambio moderado en la media, debe resultar un cambio más o menos grande en la *pendiente* de la gráfica cusum, dado que cada nueva observación tiene la probabilidad de contribuir a un cambio y la medida que se grafica acumula esos cambios. Desde luego, la señal de que la media ha cambiado reside en la naturaleza de la pendiente de la gráfica cusum. El objetivo de la gráfica es detectar cambios que se alejan del nivel de referencia. Una pendiente diferente de cero (en cualquier dirección) representa un cambio a partir del nivel de referencia. Una pendiente positiva indica un aumento en la media por arriba del nivel de referencia, en tanto que una pendiente negativa señala una disminución.

Las gráficas cusum a menudo se diseñan con un *nivel de calidad aceptable* definido (NCA) y un *nivel de calidad rechazable* (NCR) preestablecido por el usuario. Ambos representan valores de la media. Se podría considerar que éstos desempeñan papeles similares a los de las medias nula y alternativa en la prueba de hipótesis. Considere una situación en la que el analista desea detectar un aumento en el valor de la media del proceso. Usaremos la notación μ_0 para NCA y μ_1 para NCR, y $\mu_1 > \mu_0$. El nivel de referencia se fija ahora en

$$W = \frac{\mu_0 + \mu_1}{2}.$$

Los valores de S_r ($r = 1, 2, \dots$) tendrán una pendiente negativa si la media del proceso está en μ_0 y una pendiente positiva si la media del proceso está en μ_1 .

Regla de decisión para las gráficas cusum

Como antes se expuso, la pendiente de la gráfica cusum proporciona la señal de acción para el analista de control de calidad. La regla de decisión exige tomar medidas si, en el r -ésimo periodo de muestreo,

$$d_r > h,$$

donde h es un valor preestablecido que se denomina **longitud del intervalo de decisión** y

$$d_r = S_r - \min_{1 \leq i \leq r-1} S_i.$$

En otras palabras, se toman medidas si los datos revelan que el valor de la cusum real excede en una cantidad específica al valor previo de la cusum más pequeño.

Una modificación en la mecánica que se describió antes facilita el uso del método. Describimos un procedimiento que grafica las cusum y calcula las diferencias. Una modificación simple implica graficar las diferencias de manera directa y permitir la verificación contra el intervalo de decisión. La expresión general para d_r es muy sencilla. Para el procedimiento de cusum, con el que se detectan aumentos en la media,

$$d_r = \max[0, d_{r-1} + (X_r - W)].$$

La elección del valor de h es, por supuesto, muy importante. En este libro no se proporcionan los detalles que aparecen en la literatura que trata de esta elección. Para una exposición más completa se remite al lector a Ewan y Kemp, 1960, y a Montgomery, 2000b (véase la bibliografía). Una consideración importante es la **longitud esperada de la corrida**. De manera ideal, la longitud esperada de la corrida es bastante grande bajo $\mu = \mu_0$ y muy pequeña cuando $\mu = \mu_1$.

Ejercicios de repaso

17.1 Considere X_1, X_2, \dots, X_n , como variables aleatorias de Poisson independientes con parámetros $\mu_1, \mu_2, \dots, \mu_n$. Utilice las propiedades de las funciones generadoras de momento para demostrar que la variable aleatoria $\sum_{i=1}^n X_i$ es una variable aleatoria de Poisson con media $\sum_{i=1}^n \mu_i$ y varianza $\sum_{i=1}^n \mu_i$.

17.2 Considere los siguientes datos tomados en subgrupos de tamaño 5. Los datos contienen 20 promedios y rangos del diámetro (en milímetros) de una parte importante de un motor. Elabore gráficas \bar{X} y R . ¿Parecería que el proceso está bajo control?

| Muestra | \bar{X} | R |
|---------|-----------|--------|
| 1 | 2.3972 | 0.0052 |
| 2 | 2.4191 | 0.0117 |
| 3 | 2.4215 | 0.0062 |
| 4 | 2.3917 | 0.0089 |
| 5 | 2.4151 | 0.0095 |
| 6 | 2.4027 | 0.0101 |
| 7 | 2.3921 | 0.0091 |
| 8 | 2.4171 | 0.0059 |

| Muestra | \bar{X} | R |
|---------|-----------|--------|
| 9 | 2.3951 | 0.0068 |
| 10 | 2.4215 | 0.0048 |
| 11 | 2.3887 | 0.0082 |
| 12 | 2.4107 | 0.0032 |
| 13 | 2.4009 | 0.0077 |
| 14 | 2.3992 | 0.0107 |
| 15 | 2.3889 | 0.0025 |
| 16 | 2.4107 | 0.0138 |
| 17 | 2.4109 | 0.0037 |
| 18 | 2.3944 | 0.0052 |
| 19 | 2.3951 | 0.0038 |
| 20 | 2.4015 | 0.0017 |

17.3 En el ejercicio de repaso 17.2 suponga que el comprador fija especificaciones para la parte. Las especificaciones exigen que el diámetro caiga en el rango cubierto por 2.40000 ± 0.0100 mm. ¿Qué proporción de unidades producidas por este proceso no cumplirán con las especificaciones?

17.4 Para la situación del ejercicio de repaso 17.2 proporcione estimados numéricos de la media y de la desviación estándar del diámetro para la parte que se fabrica en el proceso.

17.5 Considere los datos de la tabla 17.1. Suponga que se toman muestras adicionales de tamaño 5 y que se registra la resistencia a la tensión. El muestreo produce los siguientes resultados (en libras por pulgada cuadrada).

| Muestra | \bar{X} | R |
|---------|-----------|-----|
| 1 | 1511 | 22 |
| 2 | 1508 | 14 |
| 3 | 1522 | 11 |
| 4 | 1488 | 18 |
| 5 | 1519 | 6 |
| 6 | 1524 | 11 |
| 7 | 1519 | 8 |
| 8 | 1504 | 7 |
| 9 | 1500 | 8 |
| 10 | 1519 | 14 |

- a) Grafique los datos, utilice las gráficas \bar{X} y R para los datos preliminares de la tabla 17.1.
 b) ¿Parecería que el proceso está bajo control? Si no es así, explique por qué.

17.6 Considere un proceso bajo control con media $\mu = 25$ y $\sigma = 1.0$. Suponga que se usan subgrupos de tamaño 5 con límites de control $\mu \pm 3\sigma/\sqrt{n}$ y línea central en μ . Suponga que ocurre un cambio en la media, y que la nueva media es $\mu = 26.5$.

- a) ¿Cuál es el número promedio de muestras requeridas (después del cambio) para detectar la situación fuera de control?
 b) ¿Cuál es la desviación estándar del número de corridas requeridas?

17.7 Considere la situación del ejemplo 17.2. Se toman los siguientes datos de muestras adicionales de tamaño 5. Grafique los valores \bar{X} y S sobre las gráficas \bar{X} y S que se dibujaron con los datos en la muestra preliminar. ¿Parecería que el proceso está bajo control? Explique su respuesta.

| Muestra | \bar{X} | S_i |
|---------|-----------|-------|
| 1 | 62.280 | 0.062 |
| 2 | 62.319 | 0.049 |
| 3 | 62.297 | 0.077 |
| 4 | 62.318 | 0.042 |
| 5 | 62.315 | 0.038 |
| 6 | 62.389 | 0.052 |
| 7 | 62.401 | 0.059 |
| 8 | 62.315 | 0.042 |
| 9 | 62.298 | 0.036 |
| 10 | 62.337 | 0.068 |

17.8 Cada hora se toman muestras de tamaño 50 de un proceso que produce cierto tipo de artículo que se considera que está defectuoso o que no tiene defecto. Se toman 20 muestras.

- a) Construya una gráfica de control para controlar la proporción de artículos defectuosos.
 b) ¿Parecería que el proceso está bajo control? Explique su respuesta.

| Número de artículos defectuosos | | Número de artículos defectuosos | |
|---------------------------------|---|---------------------------------|---|
| Muestra | | Muestra | |
| 1 | 4 | 11 | 2 |
| 2 | 3 | 12 | 4 |
| 3 | 5 | 13 | 1 |
| 4 | 3 | 14 | 2 |
| 5 | 2 | 15 | 3 |
| 6 | 2 | 16 | 1 |
| 7 | 2 | 17 | 1 |
| 8 | 1 | 18 | 2 |
| 9 | 4 | 19 | 3 |
| 10 | 3 | 20 | 1 |

17.9 Para la situación del ejercicio de repaso 17.8 suponga que se reúnen los siguientes datos adicionales:

| Muestra | Número de artículos defectuosos |
|---------|---------------------------------|
| 1 | 3 |
| 2 | 4 |
| 3 | 2 |
| 4 | 2 |
| 5 | 3 |
| 6 | 1 |
| 7 | 3 |
| 8 | 5 |
| 9 | 7 |
| 10 | 7 |

¿Parecería que el proceso está bajo control? Explique su respuesta.

17.10 Se aplica un programa de control de calidad para un proceso, donde se fabrican grandes placas de acero, con un interés especial por los defectos superficiales. El objetivo es establecer una gráfica de control de calidad para el número de defectos por placa. Los datos se presentan a continuación. Elabore la gráfica de control apropiada utilizando esta información. ¿Parecería que el proceso está bajo control?

| Número de defectos | | Número de defectos | |
|--------------------|---|--------------------|---|
| Muestra | | Muestra | |
| 1 | 4 | 11 | 1 |
| 2 | 2 | 12 | 2 |
| 3 | 1 | 13 | 2 |
| 4 | 3 | 14 | 3 |
| 5 | 0 | 15 | 1 |
| 6 | 4 | 16 | 4 |
| 7 | 5 | 17 | 3 |
| 8 | 3 | 18 | 2 |
| 9 | 2 | 19 | 1 |
| 10 | 2 | 20 | 3 |

Capítulo 18

Estadística bayesiana

18.1 Conceptos bayesianos

Los métodos clásicos de estimación que hemos estudiado hasta ahora se basan sólo en la información que brinda la muestra aleatoria. Estos métodos en esencia interpretan probabilidades como frecuencias relativas. Por ejemplo, para obtener un intervalo de confianza de 95% para μ , interpretamos la aseveración

$$P(-1.96 < Z < 1.96) = 0.95$$

para afirmar que, en experimentos repetidos, Z caerá 95% de las veces entre -1.96 y 1.96 . Dado que

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

para una muestra normal con varianza conocida, el enunciado de probabilidad aquí significa que 95% de los intervalos aleatorios $(\bar{X} - 1.96\sigma/\sqrt{n}, \bar{X} + 1.96\sigma/\sqrt{n})$ contienen la media μ verdadera. Otro enfoque de los métodos estadísticos de estimación se denomina **metodología bayesiana**. La idea principal del método proviene de la regla de Bayes, que examinamos en la sección 2.7. La diferencia fundamental entre el enfoque bayesiano y el clásico o frecuente es que en los conceptos bayesianos los parámetros se consideran variables aleatorias.

Probabilidad subjetiva

La probabilidad subjetiva es el fundamento de los conceptos bayesianos. En el capítulo 2 analizamos dos acercamientos posibles a la probabilidad, es decir, el método de la frecuencia relativa y el método de la indiferencia. El primero determina una probabilidad como una consecuencia de experimentos repetidos. Por ejemplo, para decidir el porcentaje de tiros libres de un jugador de basquetbol, podemos registrar el número de tiros que hace y el número total de intentos que tal jugador ha hecho hasta el momento. La probabilidad de que este jugador acierte un tiro libre se puede calcular como el cociente de estos dos números. Por otro lado, si no sabemos acerca de cualquier sesgo en un dado, la probabilidad de que aparezca un 3 en el siguiente lanzamiento será de $1/6$. Dicho enfoque en la interpretación de la probabilidad se basa en la regla de la indiferencia.

Sin embargo, en muchas situaciones no es posible aplicar las interpretaciones de probabilidad anteriores. Por ejemplo, considere las siguientes preguntas: “¿Qué probabilidad hay de que llueva mañana?” “¿Qué tan probable es que el precio de estas acciones aumente a fin de mes?” y “¿Cuál es la probabilidad de que dos empresas se fusionen?”. Estas preguntas difícilmente se podrían interpretar mediante los enfoques anteriores, y las respuestas podrían ser diferentes para distintas personas. No obstante, este tipo de preguntas se plantean constantemente en la vida diaria y el enfoque utilizado para explicar esas probabilidades se llama *probabilidad subjetiva*, ya que refleja opiniones subjetivas.

Perspectiva condicional

Recuerde que en los capítulos 9 a 17 todas las inferencias estadísticas se basaban en el hecho de que los parámetros se desconocen pero son cantidades fijas, excepto los revisados en la sección 9.14, en donde los parámetros se trataron como variables y los estimados de máxima verosimilitud (EMV) se calcularon con base en la muestra de datos observados. En la estadística bayesiana los parámetros no sólo se manejan como variables, como en los cálculos de EMV, sino que también se manejan como aleatorios.

Puesto que los datos observados son los únicos resultados experimentales para el profesionista, la inferencia estadística se basa en los datos reales observados a partir de un experimento dado. A esta visión se le llama *perspectiva condicional*. Más aún, en los conceptos bayesianos, dado que los parámetros se manejan como aleatorios, es factible especificar una distribución de probabilidad, por lo general utilizando la *probabilidad subjetiva* para el parámetro. Este tipo de distribución se denomina *distribución previa* y comúnmente refleja la creencia previa del experimentador acerca del parámetro. En la perspectiva bayesiana, una vez que se realiza un experimento y se observan los datos, todo el conocimiento acerca de un parámetro está contenido en los datos reales observados, así como en la información previa.

Aplicaciones bayesianas

Aunque la regla de Bayes se atribuye a Thomas Bayes, las aplicaciones bayesianas fueron utilizadas por primera vez por el científico francés Pierre Simon Laplace, quien publicó un artículo sobre el uso de la inferencia bayesiana en las proporciones binomiales desconocidas (para revisar la distribución binomial véase la sección 5.2).

A partir de la introducción del paquete para el cálculo de la cadena Markov de Monte Carlo (MCMC) para el análisis bayesiano a principios de la década de 1990, los métodos bayesianos se han vuelto cada vez más populares para los modelos estadísticos y el análisis de datos. Al mismo tiempo, la metodología que utiliza conceptos bayesianos ha avanzado mucho y se aplica en campos como la bioinformática, la biología, los negocios, la ingeniería, las ciencias ambientales y la ecología, así como en la ciencia de la vida y la salud, entre otros.

18.2 Inferencias bayesianas

Considere el problema de calcular un estimado puntual del parámetro θ para la población con distribución $f(x|\theta)$, dado θ . Denote con $\pi(\theta)$ la distribución previa de θ . Suponga que se observa una muestra aleatoria de tamaño n denotada con $\mathbf{x} = (x_1, x_2, \dots, x_n)$.

Definición 18.1: La distribución de θ , dado x , que se denomina distribución posterior, es dada por

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{g(x)},$$

donde $g(x)$ es la distribución marginal de x .

La distribución marginal de x en la definición anterior se puede calcular usando la siguiente fórmula:

$$g(x) = \begin{cases} \sum_{\theta} f(x|\theta)\pi(\theta), & \theta \text{ es discreta,} \\ \int_{-\infty}^{\infty} f(x|\theta)\pi(\theta) d\theta, & \theta \text{ es continua.} \end{cases}$$

Ejemplo 18.1: Suponga que la distribución previa para la proporción de artículos defectuosos que produce una máquina es

| | | |
|----------|-----|-----|
| p | 0.1 | 0.2 |
| $\pi(p)$ | 0.6 | 0.4 |

Denote con x el número de artículos defectuosos en una muestra aleatoria de tamaño 2. Calcule la distribución de probabilidad posterior de p , dado que se observa x .

Solución: La variable aleatoria X sigue una distribución binomial

$$f(x|p) = b(x; 2, p) = \binom{2}{x} p^x q^{2-x}, \quad x = 0, 1, 2.$$

La distribución marginal de x se puede calcular como

$$\begin{aligned} g(x) &= f(x|0.1)\pi(0.1) + f(x|0.2)\pi(0.2) \\ &= \binom{2}{x} [(0.1)^x (0.9)^{2-x} (0.6) + (0.2)^x (0.8)^{2-x} (0.4)]. \end{aligned}$$

Por lo tanto, para $x = 0, 1, 2$ obtenemos las siguientes probabilidades marginales

| | | | |
|--------|-------|-------|-------|
| x | 0 | 1 | 2 |
| $g(x)$ | 0.742 | 0.236 | 0.022 |

La probabilidad posterior de $p = 0.1$, dado x , es

$$\pi(0.1|x) = \frac{f(x|0.1)\pi(0.1)}{g(x)} = \frac{(0.1)^x (0.9)^{2-x} (0.6)}{(0.1)^x (0.9)^{2-x} (0.6) + (0.2)^x (0.8)^{2-x} (0.4)},$$

y $\pi(0.2|x) = 1 - \pi(0.1|x)$.

Suponga que se observa $x = 0$.

$$\pi(0.1|0) = \frac{f(0|0.1)\pi(0.1)}{g(0)} = \frac{(0.1)^0 (0.9)^{2-0} (0.6)}{0.742} = 0.6550,$$

y $\pi(0.2|0) = 0.3450$. Si se observa $x = 1$, $\pi(0.1|1) = 0.4576$ y $\pi(0.2|1) = 0.5424$. Por último, $\pi(0.1|2) = 0.2727$ y $\pi(0.2|2) = 0.7273$. ▮

La distribución previa del ejemplo 18.1 es discreta, aunque el rango natural de p va de 0 a 1. Considere el siguiente ejemplo, en el cual tenemos una distribución previa que abarca el espacio completo de p .

Ejemplo 18.2: Suponga que la distribución previa de p es uniforme, es decir, $\pi(p) = 1$, para $0 < p < 1$. Use la misma variable aleatoria X que en el ejemplo 18.1 para calcular la distribución posterior de p .

Solución: Como en el ejemplo 18.1, tenemos

$$f(x|p) = b(x; 2, p) = \binom{2}{x} p^x q^{2-x}, \quad x = 0, 1, 2.$$

La distribución marginal de x se puede calcular como

$$g(x) = \int_0^1 f(x|p)\pi(p) dp = \binom{2}{x} \int_0^1 p^x (1-p)^{2-x} dp.$$

La integral anterior se puede evaluar en cada x directamente como $g(0) = 1/3$, $g(1) = 1/3$ y $g(2) = 1/3$. Por lo tanto, la distribución posterior de p , dado x , es

$$\pi(p|x) = \frac{\binom{2}{x} p^x (1-p)^{2-x}}{1/3} = 3 \binom{2}{x} p^x (1-p)^{2-x}, \quad 0 < p < 1.$$

La distribución posterior anterior es en realidad una distribución beta (véase la sección 6.8) con parámetros $\alpha = x + 1$ y $\beta = 3 - x$. Por lo tanto, si se observa $x = 0$, la distribución posterior de p es una distribución beta con parámetros $(1, 3)$. La media posterior es $\mu = \frac{1}{1+3} = \frac{1}{4}$ y la varianza posterior es $\sigma^2 = \frac{(1)(3)}{(1+3)^2(1+3+1)} = \frac{3}{80}$. \blacksquare

Si utilizamos la distribución posterior, podemos estimar directamente el (los) parámetro(s) en una población. Al calcular las distribuciones posteriores es muy útil estar familiarizado con las distribuciones que se estudiaron en los capítulos 5 y 6. Observe que en la definición 18.1 la *variable* en la distribución posterior es θ , en tanto se proporcione x . Por consiguiente, podemos tratar a $g(x)$ como una constante cuando calculamos la distribución posterior de θ . Entonces, la distribución posterior se puede expresar como

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta),$$

donde el símbolo " \propto " significa *proporcional a*. En el cálculo que se hizo de la distribución posterior podríamos dejar los factores que no dependen de θ fuera de la constante de normalización, esto es, la densidad marginal $g(x)$.

Ejemplo 18.3 Suponga que las variables aleatorias X_1, \dots, X_n son independientes y provienen de una distribución de Poisson con media λ . Suponga que la distribución previa de λ es exponencial con media 1. Calcule la distribución posterior de λ cuando $\bar{x} = 3$ con $n = 10$.

Solución: La función de densidad de $\mathbf{X} = (X_1, \dots, X_n)$ es

$$f(\mathbf{x}|\lambda) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!},$$

y la distribución previa es

$$\pi(\theta) = e^{-\lambda}, \text{ para } \lambda > 0.$$

En consecuencia, utilizando la definición 18.1 se obtiene la siguiente distribución posterior de λ

$$\pi(\lambda|x) \propto f(x|\lambda)\pi(\lambda) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-\lambda} \propto e^{-(n+1)\lambda} \lambda^{\sum_{i=1}^n x_i}.$$

Si nos remitimos a la distribución gamma en la sección 6.6, concluimos que la distribución posterior de λ sigue una distribución gamma con parámetros $1 + \sum_{i=1}^n x_i$ y $\frac{1}{n+1}$. Por lo tanto, tenemos la media y la varianza posterior de λ como $\frac{\sum_{i=1}^n x_i + 1}{n+1}$ y $\frac{\sum_{i=1}^n x_i + 1}{(n+1)^2}$. Así, cuando $\bar{x} = 3$ con $n = 10$, tenemos $\sum_{i=1}^{10} x_i = 30$. Por lo tanto, la distribución posterior de λ es una distribución gamma con parámetros 31 y $1/11$. ■

A partir del ejemplo 18.3 observamos que en ocasiones es muy conveniente usar la técnica "proporcional a" para calcular la distribución posterior, especialmente cuando el resultado se puede formar para una distribución de uso común como las que se describen en los capítulos 5 y 6.

Estimación puntual mediante la distribución posterior

Una vez que hemos derivado la distribución posterior, fácilmente podemos usar el resumen de la distribución posterior para hacer inferencias sobre los parámetros de la población. Por ejemplo, la media, la mediana y la moda posteriores son útiles para estimar el parámetro.

Ejemplo 18.4: Suponga que en el ejemplo 18.2 se observa $x = 1$. Determine la media y la moda posteriores.

Solución: Cuando $x = 1$, la distribución posterior de p se puede expresar como

$$\pi(p|1) = 6p(1-p), \quad \text{para } 0 < p < 1.$$

Para calcular la media de esta distribución necesitamos encontrar

$$\int_0^1 6p^2(1-p) dp = 6 \left(\frac{1}{3} - \frac{1}{4} \right) = \frac{1}{2}.$$

Para determinar la moda posterior se requiere obtener un valor p tal que se maximice la distribución posterior. Si tomamos la derivada de $\pi(p)$ respecto a p , obtenemos $6 - 12p$. Al despejar p en $6 - 12p = 0$, obtenemos $p = \frac{1}{2}$. La segunda derivada es -12 , la cual implica que la moda posterior se logra en $p = \frac{1}{2}$. ■

Los métodos bayesianos de estimación respecto a la media μ de una población normal se basan en el siguiente ejemplo.

Ejemplo 18.5: Si \bar{x} es la media de una muestra aleatoria de tamaño n tomada de una población normal con varianza conocida σ^2 , y la distribución previa de la media poblacional es una distribución normal con media conocida μ_0 y varianza conocida σ_0^2 , demuestre que la distribución

posterior de la media poblacional es también una distribución normal con media μ^* y desviación estándar σ^* , donde

$$\mu^* = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n} \bar{x} + \frac{\sigma^2/n}{\sigma_0^2 + \sigma^2/n} \mu_0 \quad \text{y} \quad \sigma^* = \sqrt{\frac{\sigma_0^2 \sigma^2}{n\sigma_0^2 + \sigma^2}}.$$

Solución: La función de densidad de la muestra es

$$f(x_1, x_2, \dots, x_n | \mu) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2 \right],$$

para $-\infty < x_i < \infty$ e $i = 1, 2, \dots, n$, y la previa es

$$\pi(\mu) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp \left[-\frac{1}{2} \left(\frac{\mu - \mu_0}{\sigma_0} \right)^2 \right], \quad -\infty < \mu < \infty.$$

Entonces la distribución posterior de μ es

$$\begin{aligned} \pi(\mu|x) &\propto \exp \left\{ -\frac{1}{2} \left[\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2 + \left(\frac{\mu - \mu_0}{\sigma_0} \right)^2 \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[\frac{n(\bar{x} - \mu)^2}{\sigma^2} + \frac{(\mu - \mu_0)^2}{\sigma_0^2} \right] \right\}, \end{aligned}$$

debido a

$$\sum_{i=1}^n (x_i - \mu)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2$$

de la sección 8.5. Al completar los cuadrados para μ se obtiene la distribución posterior

$$\pi(\mu|x) \propto \exp \left[-\frac{1}{2} \left(\frac{\mu - \mu^*}{\sigma^*} \right)^2 \right],$$

donde

$$\mu^* = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2}, \quad \sigma^* = \sqrt{\frac{\sigma_0^2\sigma^2}{n\sigma_0^2 + \sigma^2}}.$$

Esta es una distribución normal con media μ^* y desviación estándar σ^* . ▮

El teorema del límite central nos permite utilizar el ejemplo 18.5 también cuando seleccionamos muestras aleatorias suficientemente grandes ($n \geq 30$ para muchos casos de experimentación en ingeniería), a partir de poblaciones no normales (la distribución no dista mucho de ser simétrica), y cuando la distribución previa de la media es aproximadamente normal.

Resulta pertinente hacer algunos comentarios acerca del ejemplo 18.5. La media posterior μ^* también se puede escribir como

$$\mu^* = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n} \bar{x} + \frac{\sigma^2/n}{\sigma_0^2 + \sigma^2/n} \mu_0,$$

que es el promedio ponderado de la media muestral \bar{x} y la media previa μ_0 . Como ambos coeficientes están entre 0 y 1 y se suman a 1, la media posterior μ^* siempre se encuentra

entre \bar{x} y μ_0 . Esto significa que tanto \bar{x} como μ_0 influyen en la estimación posterior de μ . Además, la ponderación de \bar{x} depende de la varianza previa, así como de la varianza de la media muestral. Para un problema con una muestra grande ($n \rightarrow \infty$), la media posterior $\mu^* \rightarrow \bar{x}$. Esto significa que la media previa no desempeña ningún papel en la estimación de la media poblacional μ utilizando la distribución posterior. Esto es muy razonable, puesto que indica que cuando una cantidad de datos es sustancial, la información a partir de los datos dominará la información de μ proporcionada por la previa. Por otro lado, cuando la varianza previa es grande ($\sigma_0^2 \rightarrow \infty$), la media posterior μ^* también va hacia \bar{x} . Observe que para una distribución normal, cuanto mayor es la varianza, más plana será la función de densidad. El carácter plano de la distribución normal en este caso significa que casi no hay información previa subjetiva disponible del parámetro μ antes de reunir los datos. Por lo tanto, es razonable que la estimación posterior μ^* sólo dependa del valor de los datos de \bar{x} .

Ahora considere la desviación estándar posterior σ^* . Este valor también se escribe como

$$\sigma^* = \sqrt{\frac{\sigma_0^2 \sigma^2 / n}{\sigma_0^2 + \sigma^2 / n}}.$$

Es evidente que el valor σ^* es menor que σ_0 y que σ/\sqrt{n} , la desviación estándar previa y la desviación estándar de \bar{x} , respectivamente. Esto sugiere que la estimación posterior es más precisa que la previa y que los datos muestrales. En consecuencia, la incorporación tanto de los datos como de la información previa produce una mejor información posterior que si se utiliza cualquiera de los datos o la información previa por sí solos. Esto es un fenómeno común en la inferencia bayesiana. Además, para calcular μ^* y σ^* mediante las fórmulas del ejemplo 18.5 suponemos que se conoce σ^2 . Como por lo general éste no es el caso, deberemos reemplazar σ^2 por la varianza de la muestra s^2 siempre que $n \geq 30$.

Estimación del intervalo bayesiano

De manera similar al intervalo de confianza clásico, en el análisis bayesiano podemos calcular un intervalo bayesiano del $100(1 - \alpha)\%$ empleando la distribución posterior.

Definición 18.2: El intervalo $a < \theta < b$ se denomina **intervalo de Bayes** del $100(1 - \alpha)\%$ para θ si

$$\int_{-\infty}^a \pi(\theta|x) d\theta = \int_b^{\infty} \pi(\theta|x) d\theta = \frac{\alpha}{2}.$$

Recuerde que, de acuerdo con el enfoque frecuentista, la probabilidad de un intervalo de confianza, digamos de 95%, se interpreta como una probabilidad de cobertura, esto significa que, si un experimento se repite una y otra vez (con considerables datos no observados), la probabilidad de que los intervalos calculados, de acuerdo con la regla, cubran el parámetro verdadero es de 95%. Sin embargo, en la interpretación del intervalo bayesiano, digamos para un intervalo de 95%, podemos decir que la probabilidad de que el parámetro desconocido caiga dentro del intervalo calculado (que sólo depende de los datos observados) es de 95%.

Ejemplo 18.6: Suponga que $X \sim b(x; n, p)$ con $n = 2$ conocida, y la distribución previa de p es uniforme $\pi(p) = 1$ para $0 < p < 1$. Calcule el intervalo de Bayes de 95% para p .

Solución: Como en el ejemplo 18.2, cuando $x = 0$ la distribución posterior es una distribución beta con parámetros 1 y 3, es decir, $\pi(p|0) = 3(1-p)^2$, para $0 < p < 1$. Por consiguiente, necesitamos despegar a y b utilizando la definición 18.2, lo que produce lo siguiente:

$$0.025 = \int_0^a 3(1-p)^2 dp = 1 - (1-a)^3$$

y

$$0.025 = \int_b^1 3(1-p)^2 dp = (1-b)^3.$$

Las soluciones a las ecuaciones anteriores dan como resultado $a = 0.0084$ y $b = 0.7076$. Por lo tanto, la probabilidad de que p caiga dentro de $(0.0084, 0.7076)$ es de 95%. \blacksquare

Para la población normal y el caso previo normal descrito en el ejemplo 18.5, la media posterior μ^* es el estimado de Bayes de la media poblacional μ , y se puede construir un **intervalo bayesiano** para μ de $100(1 - \alpha)\%$ calculando el intervalo

$$\mu^* - z_{\alpha/2}\sigma^* < \mu < \mu^* + z_{\alpha/2}\sigma^*,$$

que se centra en la media posterior y contiene $100(1 - \alpha)\%$ de la probabilidad posterior.

Ejemplo 18.7: Una empresa de equipo eléctrico fabrica bombillas con una duración distribuida de forma aproximadamente normal y una desviación estándar de 100 horas. La experiencia previa nos conduce a creer que μ es un valor de una variable aleatoria normal con una media $\mu_0 = 800$ horas y una desviación estándar $\sigma_0 = 10$ horas. Si una muestra aleatoria de 25 bombillas tiene una duración promedio de 780 horas, calcule un intervalo bayesiano de 95% para μ .

Solución: De acuerdo con el ejemplo 18.5, la distribución posterior de la media también es una distribución normal con media

$$\mu^* = \frac{(25)(780)(10)^2 + (800)(100)^2}{(25)(10)^2 + (100)^2} = 796$$

y desviación estándar

$$\sigma^* = \sqrt{\frac{(10)^2(100)^2}{(25)(10)^2 + (100)^2}} = \sqrt{80}.$$

El intervalo bayesiano de 95% para μ es dado entonces por

$$796 - 1.96\sqrt{80} < \mu < 796 + 1.96\sqrt{80},$$

o

$$778.5 < \mu < 813.5.$$

En consecuencia, estamos 95% seguros de que μ estará entre 778.5 y 813.5.

Por otro lado, si desconocemos la información previa acerca de μ , procedemos como en la sección 9.4 para construir el intervalo de confianza clásico de 95%.

$$780 - (1.96)\left(\frac{100}{\sqrt{25}}\right) < \mu < 780 + (1.96)\left(\frac{100}{\sqrt{25}}\right),$$

o $740.8 < \mu < 819.2$, el cual se ve que es más amplio que el intervalo bayesiano correspondiente. \blacksquare

18.3 Estimados bayesianos mediante la teoría de decisión

Con la metodología bayesiana se puede obtener la distribución posterior del parámetro. Los estimados bayesianos también se pueden derivar usando la distribución posterior y una función de pérdida cuando se incurre en una pérdida. Una función de pérdida es aquella que describe el costo de una decisión asociada con un suceso de interés. Aquí sólo se citan unas cuantas funciones de pérdida de uso común y sus estimados de Bayes asociados.

Pérdida del cuadrado del error

Definición 18.3: La función de pérdida del cuadrado del error es

$$L(\theta, a) = (\theta - a)^2,$$

donde θ es el parámetro (o estado natural) y a una acción (o estimado).

Un estimado de Bayes minimiza la pérdida posterior esperada dada en los datos muestrales observados.

Teorema 18.1: La media de la distribución posterior $\pi(\theta|x)$, denotada con θ^* , es el estimado de Bayes de θ bajo la función de pérdida del cuadrado del error.

Ejemplo 18.8: Calcule el estimado de Bayes de p para todos los valores de x en el ejemplo 18.1 cuando se utiliza la función de pérdida del cuadrado del error.

Solución: Cuando $x = 0$, $p^* = (0.1)(0.6550) + (0.2)(0.3450) = 0.1345$.

Cuando $x = 1$, $p^* = (0.1)(0.4576) + (0.2)(0.5424) = 0.1542$.

Cuando $x = 2$, $p^* = (0.1)(0.2727) + (0.2)(0.7273) = 0.1727$.

Observe que el estimado clásico de p es $\hat{p} = x/n = 0, 1/2$ y 1 , respectivamente, para los valores de x en $0, 1$ y 2 . Estos estimados clásicos son muy diferentes de los estimados de Bayes correspondientes. ■

Ejemplo 18.9: Repita el ejemplo 18.8 en la situación del ejemplo 18.2.

Solución: Puesto que la distribución posterior de p es una distribución $B(x + 1, 3 - x)$ (véase la sección 6.8 en la página 201), el estimado de Bayes de p es

$$p^* = E^{\pi(p|x)}(p) = 3 \binom{2}{x} \int_0^1 p^{x+1} (1-p)^{2-x} dp,$$

que produce $p^* = 1/4$ para $x = 0$, $p^* = 1/2$ para $x = 1$, y $p^* = 3/4$ para $x = 2$, respectivamente. Advierta que cuando se observa $x = 1$, el estimado de Bayes y el estimado clásico de \hat{p} son equivalentes. ■

Para la situación normal que se describe en el ejemplo 18.5 el estimado de Bayes de μ bajo la pérdida del cuadrado del error será la media posterior μ^* .

Ejemplo 18.10: Suponga que la distribución muestral de una variable aleatoria X es de Poisson con parámetro λ . Suponga que la distribución previa de λ sigue una distribución gamma con

parámetros (α, β) . Calcule el estimado de Bayes de λ bajo la función de pérdida del cuadrado del error.

Solución: Si utilizamos el ejemplo 18.3, concluimos que la distribución posterior de λ sigue una distribución gamma con parámetros $(x + \alpha, (1 + 1/\beta)^{-1})$. Por medio del teorema 6.4 obtenemos la media posterior

$$\hat{\lambda} = \frac{x + \alpha}{1 + 1/\beta}.$$

Como la media posterior es el estimado de Bayes bajo la pérdida del cuadrado del error, $\hat{\lambda}$ es nuestro estimado de Bayes. \blacksquare

Pérdida del error absoluto

La pérdida del cuadrado del error descrita antes es similar al concepto de los mínimos cuadrados que se analizó en relación con la regresión en los capítulos 11 y 12. En esta sección presentamos otra función de pérdida como sigue.

Definición 18.4: La función de pérdida del error absoluto se define como

$$L(\theta, a) = |\theta - a|,$$

donde θ es el parámetro y a una acción.

Teorema 18.2: La mediana de la distribución posterior $\pi(\theta|x)$, denotada con θ^* , es el estimado de Bayes de θ bajo la función de pérdida del error absoluto.

Ejemplo 18.11: Bajo la pérdida del error absoluto calcule el estimador de Bayes para el ejemplo 18.9 cuando se observa $x = 1$.

Solución: Nuevamente, la distribución posterior de p es $B(x + 1, 3 - x)$. Cuando $x = 1$ se trata de una distribución beta con densidad $\pi(p|x=1) = 6x(1-x)$ para $0 < x < 1$ y 0 en otro caso. La mediana de esta distribución es un valor de p^* tal que

$$\frac{1}{2} = \int_0^{p^*} 6p(1-p) dp = 3p^{*2} - 2p^{*3},$$

que produce la respuesta $p^* = \frac{1}{2}$. Por lo tanto, el estimado de Bayes en este caso es 0.5.

Ejercicios

18.1 Estime la proporción de artículos defectuosos que produce la máquina del ejemplo 18.1 si la muestra aleatoria de tamaño 2 produce dos artículos defectuosos.

18.2 Supongamos que la distribución previa para la proporción p de bebidas de una máquina despachadora que se derraman al servirse es

| | | | |
|----------|------|------|------|
| p | 0.05 | 0.10 | 0.15 |
| $\pi(p)$ | 0.3 | 0.5 | 0.2 |

Si dos de las siguientes 9 bebidas de esta máquina se derraman, calcule

- la distribución posterior para la proporción p ;
- el estimado de Bayes de p .

18.3 Repita el ejercicio 18.2 cuando una de las siguientes 4 bebidas se derrama y la distribución uniforme previa es

$$\pi(p) = 10, \quad 0.05 < p < 0.15.$$

18.4 Las llamadas de servicio llegan a un centro de mantenimiento de acuerdo con un proceso de Poisson con λ llamadas por minuto. Un conjunto de datos de 20 periodos de un minuto producen un promedio de 1.8 llamadas. Si la distribución previa de λ sigue una distribución exponencial con media 2, determine la distribución posterior de λ .

18.5 Un estudio previo indica que el porcentaje de fumadores empedernidos, p , que tienen cáncer de pulmón sigue una distribución beta (véase la sección 6.8) con media de 70% y desviación estándar de 10%. Suponga que un nuevo conjunto de datos recolectado indica que 81 de 120 fumadores empedernidos tiene cáncer de pulmón.

- Determine la distribución posterior del porcentaje de fumadores empedernidos que tienen cáncer de pulmón combinando los nuevos datos y la información previa.
- ¿Cuál es la probabilidad posterior de que p sea mayor que 50%?

18.6 El constructor de un nuevo complejo de condominios afirma que 3 de 5 compradores preferirá un departamento de dos recámaras, mientras que su banquero afirma que sería más correcto decir que 7 de 10 compradores preferirán uno de dos recámaras. En las predicciones previas de este tipo el banquero ha sido dos veces más confiable que el constructor. Si 12 de los siguientes 15 condominios que se venden en este complejo son de dos recámaras, calcule

- las probabilidades posteriores que se asocian con las afirmaciones del constructor y del banquero;
- un estimado puntual de la proporción de compradores que prefieren un condominio de dos recámaras.

18.7 El tiempo en que se consume la primera etapa de un cohete es una variable aleatoria normal con una desviación estándar de 0.8 minutos. Suponga una distribución previa normal para μ con una media de ocho minutos y una desviación estándar de 0.2 minutos. Si se lanzan 10 de estos cohetes y la primera etapa tiene un tiempo de consumo promedio de 9 minutos, calcule un intervalo bayesiano de 95% para μ .

18.8 La utilidad diaria de una máquina despachadora de jugos, ubicada en un edificio de oficinas, es un valor de una variable aleatoria normal, con media μ y varianza σ^2 desconocidas. Desde luego, la media variará un poco de un edificio a otro, y el distribuidor considera que estas utilidades promedio diarias se pueden describir mejor usando una distribución normal con media

$\mu_0 = \$30.00$ y desviación estándar $\sigma_0 = \$1.75$. Si una de estas máquinas despachadoras de jugo, ubicada en cierto edificio, muestra una utilidad promedio diaria de $\bar{x} = \$24.90$, durante los primeros 30 días con una desviación estándar de $s = \$2.10$, calcule

- un estimado de Bayes de la utilidad promedio diaria verdadera para este edificio;
- un intervalo bayesiano de 95% de μ para este edificio;
- la probabilidad de que la utilidad promedio diaria de la máquina en este edificio sea de entre \$24.00 y \$26.00.

18.9 El departamento de matemáticas de una universidad grande diseña un examen de colocación para aplicarlo a los grupos de nuevo ingreso a primer año. Los miembros del departamento consideran que la calificación promedio para este examen variará de un grupo de primer año a otro. Esta variación de la calificación promedio del grupo se expresa de manera subjetiva mediante una distribución normal, con una media $\mu_0 = 72$ y una varianza $\sigma_0^2 = 5.76$.

- ¿Qué probabilidad previa existe de que la calificación promedio real, que asigna el departamento para los alumnos de nuevo ingreso del siguiente año, caiga entre 71.8 y 73.4?
- Construya un intervalo bayesiano de 95% para μ en el caso de que el examen se aplicara a una muestra aleatoria de 100 estudiantes de primer grado del siguiente grupo de nuevo ingreso y tuviera como resultado una calificación promedio de 70 con una varianza de 64.
- ¿Qué probabilidad posterior debería asignar el departamento al evento del inciso a)?

18.10 Suponga que en el ejemplo 18.7 la empresa de equipo eléctrico no tiene suficiente información previa respecto a la duración media poblacional que le permita suponer una distribución normal para μ . La empresa cree, sin embargo, que μ seguramente estará entre 770 y 830 horas, y considera que una aproximación bayesiana más realista sería suponer una distribución previa

$$\pi(\mu) = \frac{1}{60}, \quad 770 < \mu < 830.$$

Si una muestra aleatoria de 25 bombillas tiene una vida promedio de 780 horas, siga los pasos de la demostración del ejemplo 18.5 para encontrar la distribución posterior.

$$\pi(\mu \mid x_1, x_2, \dots, x_{25}).$$

18.11 Suponga que el tiempo T antes de que falle cierta bisagra es una variable aleatoria exponencial con densidad de probabilidad

$$f(t) = \theta e^{-\theta t}, \quad t > 0.$$

Por experiencia, nos inclinamos a pensar que θ es un valor de una variable aleatoria exponencial con densidad de probabilidad

$$\pi(\theta) = 2e^{-2\theta}, \quad \theta > 0.$$

Si tenemos una muestra de n observaciones de T , demuestre que la distribución posterior de Θ es una distribución gamma

$$\alpha = n + 1 \quad \text{y} \quad \beta = \left(\sum_{i=1}^n t_i + 2 \right)^{-1}.$$

18.12 Suponga que una muestra consta de 5, 6, 6, 7, 5, 6, 4, 9 y 3, y 6 proviene de una población de Poisson con media λ . Suponga que el parámetro λ sigue una distribución gamma con parámetros (3, 2). Bajo la función de pérdida del cuadrado del error, calcule el estimado de Bayes de λ .

18.13 Una variable aleatoria X sigue una distribución binomial negativa con parámetros $k = 5$ y p , es decir, $b^*(x; 5, p)$. Además, se sabe que p sigue una distribución uniforme en el intervalo (0, 1). Calcule el es-

timado de Bayes de p bajo la función de pérdida del cuadrado del error.

18.14 Una variable aleatoria X sigue una distribución exponencial con media $1/\beta$. Suponga que la distribución previa de β es otra distribución exponencial con media 2.5. Determine el estimado de Bayes de β bajo la función de pérdida del error absoluto.

18.15 Una muestra aleatoria X_1, \dots, X_n proviene de una población con distribución uniforme (véase la sección 6.1) con θ desconocida. Los datos se presentan a continuación:

0.13, 1.06, 1.65, 1.73, 0.95, 0.56, 2.14, 0.33, 1.22,
0.20, 1.55, 1.18, 0.71, 0.01, 0.42, 1.03, 0.43, 1.02,
0.83, 0.88

Suponga que la distribución previa de θ tiene la densidad

$$\pi(\theta) = \begin{cases} \frac{1}{\theta^2}, & \theta > 1, \\ 0, & \theta \leq 1. \end{cases}$$

Determine el estimador de Bayes bajo la función de pérdida del error absoluto.

Bibliografía

- 1 Bartlett, M. S. y Kendall, D. G. (1946). "The Statistical Analysis of Variance Heterogeneity and Logarithmic Transformation", *Journal of the Royal Statistical Society*, Ser. B. 8, 128-138.
- 2 Bowker, A. H. y Lieberman, G. J. (1972). *Engineering Statistics*, 2.a, ed. Upper Saddle River, N.J.: Prentice Hall.
- 3 Box, G. E. P. (1988). "Signal to Noise Ratios, Performance Criteria and Transformations (with discussion)", *Technometrics*, 30, 1-17.
- 4 Box, G. E. P. y Fung, C. A. (1986). "Studies in Quality Improvement: Minimizing Transmitted Variation by Parameter Design", Informe 8. University of Wisconsin-Madison, Center for Quality and Productivity Improvement.
- 5 Box, G. E. P., Hunter, W. G. y Hunter, J. S. (1978). *Statistics for Experimenters*. Nueva York: John Wiley & Sons.
- 6 Brownlee, K. A. (1984). *Statistical Theory and Methodology: In Science and Engineering*, 2.a, ed., Nueva York: John Wiley & Sons.
- 7 Carroll, R. J. y Ruppert, D. (1988). *Transformation and Weighting in Regression*. Nueva York: Chapman y Hall.
- 8 Chatterjee, S., Hadi, A. S. y Price, B. (1999). *Regression Analysis by Example*, 3.a, ed., Nueva York: John Wiley & Sons.
- 9 Cook, R. D. y Weisberg, S. (1982). *Residuals and Influence in Regression*. Nueva York: Chapman y Hall.
- 10 Daniel, C. y Wood, F. S. (1999). *Fitting Equations to Data: Computer Analysis of Multifactor Data*, 2.a, ed., Nueva York: John Wiley & Sons.
- 11 Daniel, W. W. (1989). *Applied Nonparametric Statistics*, 2.a, ed. Belmont, Calif.: Wadsworth Publishing Company.
- 12 Devore, J. L. (2003). *Probability and Statistics for Engineering and the Sciences*, 6.a, ed., Belmont, Calif: Duxbury Press.
- 13 Dixon, W. J. (1983). *Introduction to Statistical Analysis*, 4.a, ed., Nueva York: McGraw-Hill.
- 14 Draper, N. R. y Smith, H. (1998). *Applied Regression Analysis*, 3.a, ed., Nueva York: John Wiley & Sons.

- 15 Duncan, A. (1986). *Quality Control and Industrial Statistics*, 5.a, ed., Homewood, Ill.: Irwin.
- 16 Dyer, D. D., y Keating, J. P. (1980). "On the Determination of Critical Values for Bartlett's Test", *Journal of the American Statistical Association*, **75**, 313-319.
- 17 Ewan, W. D. y Kemp, K. W. (1960). "Sampling Inspection of Continuous Processes with No Autocorrelation between Successive Results", *Biometrika*, **47**, 363-380.
- 18 Geary, R. C. (1947). "Testing for Normality", *Biometrika*, **34**, 209-242.
- 19 Gunst, R. F. y Mason, R. L. (1980). *Regression Analysis and Its Application: A Data-Oriented Approach*. Nueva York: Marcel Dekker.
- 20 Guttman, I., Wilks, S. S. y Hunter, J. S. (1971). *Introductory Engineering Statistics*. Nueva York: John Wiley & Sons.
- 21 Harville, D. A. (1977). "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems", *Journal of the American Statistical Association*, **72**, 320-338.
- 22 Hicks, C. R. y Turner, K. V. (1999). *Fundamental Concepts in the Design of Experiments*, 5.a, ed., Oxford: Oxford University Press.
- 23 Hoaglin, D. C., Mosteller, F. y Tukey, J. W. (1991). *Fundamentals of Exploratory Analysis of Variance*. Nueva York: John Wiley & Sons.
- 24 Hocking, R. R. (1976). "The Analysis and Selection of Variables in Linear Regression", *Biometrics*, **32**, 1-49.
- 25 Hodges, J. L. y Lehmann, E. L. (2005). *Basic Concepts of Probability and Statistics*, 2.a, ed. Filadelfia: Society for Industrial and Applied Mathematics.
- 26 Hoerl, A. E. y Wennard, R. W. (1970). "Ridge Regression: Applications to Nonorthogonal Problems", *Technometrics*, **12**, 55-67.
- 27 Hogg, R. V. y Ledolter, J. (1992). *Applied Statistics for Engineers and Physical Scientists*, 2.a, ed., Upper Saddle River, N.J.: Prentice Hall.
- 28 Hogg, R. V., McKean, J. W. y Craig, A. (2005). *Introduction to Mathematical Statistics*, 6.a, ed., Upper Saddle River, N.J.: Prentice Hall.
- 29 Hollander, M. y Wolfe, D. (1999). *Nonparametric Statistical Methods*. Nueva York: John Wiley & Sons.
- 30 Johnson, N. L. y Leone, F. C. (1977). *Statistics and Experimental Design: In Engineering and the Physical Sciences*, 2.a, ed., Vols. I y II, Nueva York: John Wiley & Sons.
- 31 Kackar, R. (1985). "Off-Line Quality Control, Parameter Design, and the Taguchi Methods", *Journal of Quality Technology*, **17**, 176-188.
- 32 Koopmans, L. H. (1987). *An Introduction to Contemporary Statistics*, 2.a, ed., Boston: Duxbury Press.
- 33 Kutner, M. H., Nachtsheim, C. J., Neter, J. y Li, W. (2004). *Applied Linear Regression Models*, 5.a, ed., Nueva York: McGraw-Hill/Irwin.

- 34 Larsen, R. J. y Morris, M. L. (2000). *An Introduction to Mathematical Statistics and Its Applications*, 3.a, ed., Upper Saddle River, N.J.: Prentice Hall.
- 35 Lehmann, E. L. y D'Abrera, H. J. M. (1998). *Nonparametrics: Statistical Methods Based on Ranks*, ed. rev., Upper Saddle River, N.J.: Prentice Hall.
- 36 Lentner, M. y Bishop, T. (1986). *Design and Analysis of Experiments*, 2.a, ed., Blacksburg, Va.: Valley Book Co.
- 37 Mallows, C. L. (1973). "Some Comments on C_p ", *Technometrics*, **15**, 661-675.
- 38 McClave, J. T., Dietrich, F. H. y Sincich, T. (1997). *Statistics*, 7.a, ed. Upper Saddle River, N.J.: Prentice Hall.
- 39 Montgomery, D. C. (2008a). *Design and Analysis of Experiments*, 7.a, ed., Nueva York: John Wiley & Sons.
- 40 Montgomery, D. C. (2008b). *Introduction to Statistical Quality Control*, 6.a, ed., Nueva York: John Wiley & Sons.
- 41 Mosteller, F. y Tukey, J. (1977). *Data Analysis and Regression*. Reading, Mass.: Addison-Wesley Publishing Co.
- 42 Myers, R. H. (1990). *Classical and Modern Regression with Applications*, 2.a, ed., Boston: Duxbury Press.
- 43 Myers, R. H., Khuri, A. I. y Vining, G. G. (1992). "Response Surface Alternatives to the Taguchi Robust Parameter Design Approach", *The American Statistician*, **46**, 131-139.
- 44 Myers, R. H., Montgomery, D. C. y Anderson-Cook, C. M. (2009). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 3.a, ed., Nueva York: John Wiley & Sons.
- 45 Myers, R. H., Montgomery, D. C., Vining, G. G. y Robinson, T. J. (2008). *Generalized Linear Models with Applications in Engineering and the Sciences*, 2.a, ed., Nueva York: John Wiley & Sons.
- 46 Noether, G. E. (1976). *Introduction to Statistics: A Nonparametric Approach*, 2.a, ed., Boston: Houghton Mifflin Company.
- 47 Olkin, I., Gleser, L. J. y Derman, C. (1994). *Probability Models and Applications*, 2.a, ed., Nueva York: Prentice Hall.
- 48 Ott, R. L. y Longnecker, M. T. (2000). *An Introduction to Statistical Methods and Data Analysis*, 5.a, ed., Boston: Duxbury Press.
- 49 Pacansky, J., England, C. D. y Wattman, R. (1986). "Infrared Spectroscopic Studies of Poly (perfluoropropyleneoxide) on Gold Substrate: A Classical Dispersion Analysis for the Refractive Index". *Applied Spectroscopy*, **40**, 8-16.
- 50 Plackett, R. L. y Burman, J. P. (1946). "The Design of Multifactor Experiments", *Biometrika*, **33**, 305-325.
- 51 Ross, S. M. (2002). *Introduction to Probability Models*, 9.a, ed., Nueva York: Academic Press, Inc.

- 52 Satterthwaite, F. E. (1946). "An Approximate Distribution of Estimates of Variance Components", *Biometrics*, **2**, 110-114.
- 53 Schilling, E. G. y Nelson, P. R. (1976). "The Effect of Nonnormality on the Control Limits of \bar{X} Charts", *Journal of Quality Technology*, **8**, 347-373.
- 54 Schmidt, S. R. y Launsby, R. G. (1991). *Understanding Industrial Designed Experiments*. Colorado Springs, Col. Air Academy Press.
- 55 Shoemaker, A. C., Tsui, K.-L. y Wu, C. F. J. (1991). "Economical Experimentation Methods for Robust Parameter Design", *Technometrics*, **33**, 415-428.
- 56 Snedecor, G. W. y Cochran, W. G. (1989). *Statistical Methods*, 8a ed., Allies, Iowa: The Iowa State University Press.
- 57 Steel, R. G. D., Torrie, J. H. y Dickey, D. A. (1996). *Principles and Procedures of Statistics: A Biometrical Approach*, 3.a, ed., Nueva York: McGraw-Hill.
- 58 Taguchi, G. (1991). *Introduction to Quality Engineering*. White Plains, N.Y.: Unipub/Kraus International.
- 59 Taguchi, G. y Wu, Y. (1985). *Introduction to Off-Line Quality Control*. Nagoya, Japan: Central Japan Quality Control Association.
- 60 Thompson, W. O. y Cady, F. B. (1973). *Proceedings of the University of Kentucky Conference on Regression with a Large Number of Predictor Variables*. Lexington, Ken.: University of Kentucky Press.
- 61 Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley Publishing Co.
- 62 Vining, G. G. y Myers, R. H. (1990). "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach", *Journal of Quality Technology*, **22**, 38-45.
- 63 Welch, W. J., Yu, T. K., Kang, S. M. y Sacks, J. (1990). "Computer Experiments for Quality Control by Parameter Design", *Journal of Quality Technology*, **22**, 15-22.
- 64 Winer, B. J. (1991). *Statistical Principles in Experimental Design*, 3.a, ed., Nueva York: McGraw-Hill.

Apéndice A

Tablas y demostraciones estadísticas

Tabla A.1 (continuación) Sumas de probabilidad binomial $\sum_{x=0}^r b(x; n, p)$

| n | r | p | | | | | | | | | |
|----|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
| 8 | 0 | 0.4305 | 0.1678 | 0.1001 | 0.0576 | 0.0168 | 0.0039 | 0.0007 | 0.0001 | 0.0000 | |
| | 1 | 0.8131 | 0.5033 | 0.3671 | 0.2553 | 0.1064 | 0.0352 | 0.0085 | 0.0013 | 0.0001 | |
| | 2 | 0.9619 | 0.7969 | 0.6785 | 0.5518 | 0.3154 | 0.1445 | 0.0498 | 0.0113 | 0.0012 | 0.0000 |
| | 3 | 0.9950 | 0.9437 | 0.8862 | 0.8059 | 0.5941 | 0.3633 | 0.1737 | 0.0580 | 0.0104 | 0.0004 |
| | 4 | 0.9996 | 0.9896 | 0.9727 | 0.9420 | 0.8263 | 0.6367 | 0.4059 | 0.1941 | 0.0563 | 0.0050 |
| | 5 | 1.0000 | 0.9988 | 0.9958 | 0.9887 | 0.9502 | 0.8555 | 0.6846 | 0.4482 | 0.2031 | 0.0381 |
| | 6 | | 0.9999 | 0.9996 | 0.9987 | 0.9915 | 0.9648 | 0.8936 | 0.7447 | 0.4967 | 0.1869 |
| | 7 | | 1.0000 | 1.0000 | 0.9999 | 0.9993 | 0.9961 | 0.9832 | 0.9424 | 0.8322 | 0.5695 |
| 8 | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |
| 9 | 0 | 0.3874 | 0.1342 | 0.0751 | 0.0404 | 0.0101 | 0.0020 | 0.0003 | 0.0000 | | |
| | 1 | 0.7748 | 0.4362 | 0.3003 | 0.1960 | 0.0705 | 0.0195 | 0.0038 | 0.0004 | 0.0000 | |
| | 2 | 0.9470 | 0.7382 | 0.6007 | 0.4628 | 0.2318 | 0.0898 | 0.0250 | 0.0043 | 0.0003 | 0.0000 |
| | 3 | 0.9917 | 0.9144 | 0.8343 | 0.7297 | 0.4826 | 0.2539 | 0.0994 | 0.0253 | 0.0031 | 0.0001 |
| | 4 | 0.9991 | 0.9804 | 0.9511 | 0.9012 | 0.7334 | 0.5000 | 0.2666 | 0.0988 | 0.0196 | 0.0009 |
| | 5 | 0.9999 | 0.9969 | 0.9900 | 0.9747 | 0.9006 | 0.7461 | 0.5174 | 0.2703 | 0.0856 | 0.0083 |
| | 6 | 1.0000 | 0.9997 | 0.9987 | 0.9957 | 0.9750 | 0.9102 | 0.7682 | 0.5372 | 0.2618 | 0.0530 |
| | 7 | | 1.0000 | 0.9999 | 0.9996 | 0.9962 | 0.9805 | 0.9295 | 0.8040 | 0.5638 | 0.2252 |
| | 8 | | | 1.0000 | 1.0000 | 0.9997 | 0.9980 | 0.9899 | 0.9596 | 0.8658 | 0.6126 |
| 9 | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |
| 10 | 0 | 0.3487 | 0.1074 | 0.0563 | 0.0282 | 0.0060 | 0.0010 | 0.0001 | 0.0000 | | |
| | 1 | 0.7361 | 0.3758 | 0.2440 | 0.1493 | 0.0464 | 0.0107 | 0.0017 | 0.0001 | 0.0000 | |
| | 2 | 0.9298 | 0.6778 | 0.5256 | 0.3828 | 0.1673 | 0.0547 | 0.0123 | 0.0016 | 0.0001 | |
| | 3 | 0.9872 | 0.8791 | 0.7759 | 0.6496 | 0.3823 | 0.1719 | 0.0548 | 0.0106 | 0.0009 | 0.0000 |
| | 4 | 0.9984 | 0.9672 | 0.9219 | 0.8497 | 0.6331 | 0.3770 | 0.1662 | 0.0473 | 0.0064 | 0.0001 |
| | 5 | 0.9999 | 0.9936 | 0.9803 | 0.9527 | 0.8338 | 0.6230 | 0.3669 | 0.1503 | 0.0328 | 0.0016 |
| | 6 | 1.0000 | 0.9991 | 0.9965 | 0.9894 | 0.9452 | 0.8281 | 0.6177 | 0.3504 | 0.1209 | 0.0128 |
| | 7 | | 0.9999 | 0.9996 | 0.9984 | 0.9877 | 0.9453 | 0.8327 | 0.6172 | 0.3222 | 0.0702 |
| | 8 | | 1.0000 | 1.0000 | 0.9999 | 0.9983 | 0.9893 | 0.9536 | 0.8507 | 0.6242 | 0.2639 |
| | 9 | | | | 1.0000 | 0.9999 | 0.9990 | 0.9940 | 0.9718 | 0.8926 | 0.6513 |
| 10 | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |
| 11 | 0 | 0.3138 | 0.0859 | 0.0422 | 0.0198 | 0.0036 | 0.0005 | 0.0000 | | | |
| | 1 | 0.6974 | 0.3221 | 0.1971 | 0.1130 | 0.0302 | 0.0059 | 0.0007 | 0.0000 | | |
| | 2 | 0.9104 | 0.6174 | 0.4552 | 0.3127 | 0.1189 | 0.0327 | 0.0059 | 0.0006 | 0.0000 | |
| | 3 | 0.9815 | 0.8389 | 0.7133 | 0.5696 | 0.2963 | 0.1133 | 0.0293 | 0.0043 | 0.0002 | |
| | 4 | 0.9972 | 0.9496 | 0.8854 | 0.7897 | 0.5328 | 0.2744 | 0.0994 | 0.0216 | 0.0020 | 0.0000 |
| | 5 | 0.9997 | 0.9883 | 0.9657 | 0.9218 | 0.7535 | 0.5000 | 0.2465 | 0.0782 | 0.0117 | 0.0003 |
| | 6 | 1.0000 | 0.9980 | 0.9924 | 0.9784 | 0.9006 | 0.7256 | 0.4672 | 0.2103 | 0.0504 | 0.0028 |
| | 7 | | 0.9998 | 0.9988 | 0.9957 | 0.9707 | 0.8867 | 0.7037 | 0.4304 | 0.1611 | 0.0185 |
| | 8 | | 1.0000 | 0.9999 | 0.9994 | 0.9941 | 0.9673 | 0.8811 | 0.6873 | 0.3826 | 0.0896 |
| | 9 | | | 1.0000 | 1.0000 | 0.9993 | 0.9941 | 0.9698 | 0.8870 | 0.6779 | 0.3026 |
| | 10 | | | | | 1.0000 | 0.9995 | 0.9964 | 0.9802 | 0.9141 | 0.6862 |
| 11 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |

Tabla A.1 (continuación) Sumas de probabilidad binomial $\sum_{x=0}^r b(x; n, p)$

| n | r | p | | | | | | | | | | | | | |
|----|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--|--|--|--|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | | | | |
| 15 | 0 | 0.2059 | 0.0352 | 0.0134 | 0.0047 | 0.0005 | 0.0000 | | | | | | | | |
| | 1 | 0.5490 | 0.1671 | 0.0802 | 0.0353 | 0.0052 | 0.0005 | 0.0000 | | | | | | | |
| | 2 | 0.8159 | 0.3980 | 0.2361 | 0.1268 | 0.0271 | 0.0037 | 0.0003 | 0.0000 | | | | | | |
| | 3 | 0.9444 | 0.6482 | 0.4613 | 0.2969 | 0.0905 | 0.0176 | 0.0019 | 0.0001 | 0.0000 | | | | | |
| | 4 | 0.9873 | 0.8358 | 0.6865 | 0.5155 | 0.2173 | 0.0592 | 0.0093 | 0.0007 | 0.0000 | 0.0000 | | | | |
| | 5 | 0.9978 | 0.9389 | 0.8516 | 0.7216 | 0.4032 | 0.1509 | 0.0338 | 0.0037 | 0.0001 | | | | | |
| | 6 | 0.9997 | 0.9819 | 0.9434 | 0.8689 | 0.6098 | 0.3036 | 0.0950 | 0.0152 | 0.0008 | | | | | |
| | 7 | 1.0000 | 0.9958 | 0.9827 | 0.9500 | 0.7869 | 0.5000 | 0.2131 | 0.0500 | 0.0042 | 0.0000 | | | | |
| | 8 | | 0.9992 | 0.9958 | 0.9848 | 0.9050 | 0.6964 | 0.3902 | 0.1311 | 0.0181 | 0.0003 | | | | |
| | 9 | | 0.9999 | 0.9992 | 0.9963 | 0.9662 | 0.8491 | 0.5968 | 0.2784 | 0.0611 | 0.0022 | | | | |
| | 10 | | 1.0000 | 0.9999 | 0.9993 | 0.9907 | 0.9408 | 0.7827 | 0.4845 | 0.1642 | 0.0127 | | | | |
| | 11 | | | 1.0000 | 0.9999 | 0.9981 | 0.9824 | 0.9095 | 0.7031 | 0.3518 | 0.0556 | | | | |
| | 12 | | | | 1.0000 | 0.9997 | 0.9963 | 0.9729 | 0.8732 | 0.6020 | 0.1841 | | | | |
| | 13 | | | | | 1.0000 | 0.9995 | 0.9948 | 0.9647 | 0.8329 | 0.4510 | | | | |
| | 14 | | | | | | 1.0000 | 0.9995 | 0.9953 | 0.9648 | 0.7941 | | | | |
| 15 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | | | | |
| 16 | 0 | 0.1853 | 0.0281 | 0.0100 | 0.0033 | 0.0003 | 0.0000 | | | | | | | | |
| | 1 | 0.5147 | 0.1407 | 0.0635 | 0.0261 | 0.0033 | 0.0003 | 0.0000 | | | | | | | |
| | 2 | 0.7892 | 0.3518 | 0.1971 | 0.0994 | 0.0183 | 0.0021 | 0.0001 | | | | | | | |
| | 3 | 0.9316 | 0.5981 | 0.4050 | 0.2459 | 0.0651 | 0.0106 | 0.0009 | 0.0000 | | | | | | |
| | 4 | 0.9830 | 0.7982 | 0.6302 | 0.4499 | 0.1666 | 0.0384 | 0.0049 | 0.0003 | | | | | | |
| | 5 | 0.9967 | 0.9183 | 0.8103 | 0.6598 | 0.3288 | 0.1051 | 0.0191 | 0.0016 | 0.0000 | | | | | |
| | 6 | 0.9995 | 0.9733 | 0.9204 | 0.8247 | 0.5272 | 0.2272 | 0.0583 | 0.0071 | 0.0002 | | | | | |
| | 7 | 0.9999 | 0.9930 | 0.9729 | 0.9256 | 0.7161 | 0.4018 | 0.1423 | 0.0257 | 0.0015 | 0.0000 | | | | |
| | 8 | 1.0000 | 0.9985 | 0.9925 | 0.9743 | 0.8577 | 0.5982 | 0.2839 | 0.0744 | 0.0070 | 0.0001 | | | | |
| | 9 | | 0.9998 | 0.9984 | 0.9929 | 0.9417 | 0.7728 | 0.4728 | 0.1753 | 0.0267 | 0.0005 | | | | |
| | 10 | | 1.0000 | 0.9997 | 0.9984 | 0.9809 | 0.8949 | 0.6712 | 0.3402 | 0.0817 | 0.0033 | | | | |
| | 11 | | | 1.0000 | 0.9997 | 0.9951 | 0.9616 | 0.8334 | 0.5501 | 0.2018 | 0.0170 | | | | |
| | 12 | | | | 1.0000 | 0.9991 | 0.9894 | 0.9349 | 0.7541 | 0.4019 | 0.0684 | | | | |
| | 13 | | | | | 0.9999 | 0.9979 | 0.9817 | 0.9006 | 0.6482 | 0.2108 | | | | |
| | 14 | | | | | 1.0000 | 0.9997 | 0.9967 | 0.9739 | 0.8593 | 0.4853 | | | | |
| | 15 | | | | | | 1.0000 | 0.9997 | 0.9967 | 0.9719 | 0.8147 | | | | |
| 16 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | | | | |

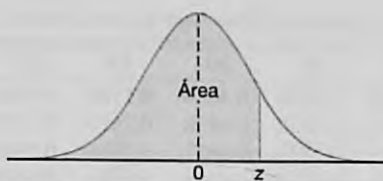
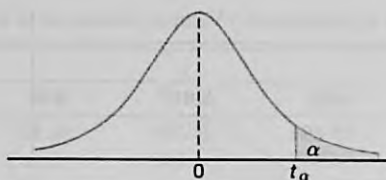


Tabla A.3 Áreas bajo la curva normal

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .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 | 0.0018 | 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 | 0.0048 |
| -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 | 0.0351 | 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 | 0.0721 | 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 | 0.3483 |
| -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 |

Tabla A.4 Valores críticos de la distribución t

| ν | α | | | | | | |
|----------|----------|-------|-------|-------|-------|-------|--------|
| | 0.40 | 0.30 | 0.20 | 0.15 | 0.10 | 0.05 | 0.025 |
| 1 | 0.325 | 0.727 | 1.376 | 1.963 | 3.078 | 6.314 | 12.706 |
| 2 | 0.289 | 0.617 | 1.061 | 1.386 | 1.886 | 2.920 | 4.303 |
| 3 | 0.277 | 0.584 | 0.978 | 1.250 | 1.638 | 2.353 | 3.182 |
| 4 | 0.271 | 0.569 | 0.941 | 1.190 | 1.533 | 2.132 | 2.776 |
| 5 | 0.267 | 0.559 | 0.920 | 1.156 | 1.476 | 2.015 | 2.571 |
| 6 | 0.265 | 0.553 | 0.906 | 1.134 | 1.440 | 1.943 | 2.447 |
| 7 | 0.263 | 0.549 | 0.896 | 1.119 | 1.415 | 1.895 | 2.365 |
| 8 | 0.262 | 0.546 | 0.889 | 1.108 | 1.397 | 1.860 | 2.306 |
| 9 | 0.261 | 0.543 | 0.883 | 1.100 | 1.383 | 1.833 | 2.262 |
| 10 | 0.260 | 0.542 | 0.879 | 1.093 | 1.372 | 1.812 | 2.228 |
| 11 | 0.260 | 0.540 | 0.876 | 1.088 | 1.363 | 1.796 | 2.201 |
| 12 | 0.259 | 0.539 | 0.873 | 1.083 | 1.356 | 1.782 | 2.179 |
| 13 | 0.259 | 0.538 | 0.870 | 1.079 | 1.350 | 1.771 | 2.160 |
| 14 | 0.258 | 0.537 | 0.868 | 1.076 | 1.345 | 1.761 | 2.145 |
| 15 | 0.258 | 0.536 | 0.866 | 1.074 | 1.341 | 1.753 | 2.131 |
| 16 | 0.258 | 0.535 | 0.865 | 1.071 | 1.337 | 1.746 | 2.120 |
| 17 | 0.257 | 0.534 | 0.863 | 1.069 | 1.333 | 1.740 | 2.110 |
| 18 | 0.257 | 0.534 | 0.862 | 1.067 | 1.330 | 1.734 | 2.101 |
| 19 | 0.257 | 0.533 | 0.861 | 1.066 | 1.328 | 1.729 | 2.093 |
| 20 | 0.257 | 0.533 | 0.860 | 1.064 | 1.325 | 1.725 | 2.086 |
| 21 | 0.257 | 0.532 | 0.859 | 1.063 | 1.323 | 1.721 | 2.080 |
| 22 | 0.256 | 0.532 | 0.858 | 1.061 | 1.321 | 1.717 | 2.074 |
| 23 | 0.256 | 0.532 | 0.858 | 1.060 | 1.319 | 1.714 | 2.069 |
| 24 | 0.256 | 0.531 | 0.857 | 1.059 | 1.318 | 1.711 | 2.064 |
| 25 | 0.256 | 0.531 | 0.856 | 1.058 | 1.316 | 1.708 | 2.060 |
| 26 | 0.256 | 0.531 | 0.856 | 1.058 | 1.315 | 1.706 | 2.056 |
| 27 | 0.256 | 0.531 | 0.855 | 1.057 | 1.314 | 1.703 | 2.052 |
| 28 | 0.256 | 0.530 | 0.855 | 1.056 | 1.313 | 1.701 | 2.048 |
| 29 | 0.256 | 0.530 | 0.854 | 1.055 | 1.311 | 1.699 | 2.045 |
| 30 | 0.256 | 0.530 | 0.854 | 1.055 | 1.310 | 1.697 | 2.042 |
| 40 | 0.255 | 0.529 | 0.851 | 1.050 | 1.303 | 1.684 | 2.021 |
| 60 | 0.254 | 0.527 | 0.848 | 1.045 | 1.296 | 1.671 | 2.000 |
| 120 | 0.254 | 0.526 | 0.845 | 1.041 | 1.289 | 1.658 | 1.980 |
| ∞ | 0.253 | 0.524 | 0.842 | 1.036 | 1.282 | 1.645 | 1.960 |

Tabla A.4 (continuación) Valores críticos de la distribución t

| ν | α | | | | | | |
|----------|----------|--------|--------|--------|--------|---------|---------|
| | 0.02 | 0.015 | 0.01 | 0.0075 | 0.005 | 0.0025 | 0.0005 |
| 1 | 15.894 | 21.205 | 31.821 | 42.433 | 63.656 | 127.321 | 636.578 |
| 2 | 4.849 | 5.643 | 6.965 | 8.073 | 9.925 | 14.089 | 31.600 |
| 3 | 3.482 | 3.896 | 4.541 | 5.047 | 5.841 | 7.453 | 12.924 |
| 4 | 2.999 | 3.298 | 3.747 | 4.088 | 4.604 | 5.598 | 8.610 |
| 5 | 2.757 | 3.003 | 3.365 | 3.634 | 4.032 | 4.773 | 6.869 |
| 6 | 2.612 | 2.829 | 3.143 | 3.372 | 3.707 | 4.317 | 5.959 |
| 7 | 2.517 | 2.715 | 2.998 | 3.203 | 3.499 | 4.029 | 5.408 |
| 8 | 2.449 | 2.634 | 2.896 | 3.085 | 3.355 | 3.833 | 5.041 |
| 9 | 2.398 | 2.574 | 2.821 | 2.998 | 3.250 | 3.690 | 4.781 |
| 10 | 2.359 | 2.527 | 2.764 | 2.932 | 3.169 | 3.581 | 4.587 |
| 11 | 2.328 | 2.491 | 2.718 | 2.879 | 3.106 | 3.497 | 4.437 |
| 12 | 2.303 | 2.461 | 2.681 | 2.836 | 3.055 | 3.428 | 4.318 |
| 13 | 2.282 | 2.436 | 2.650 | 2.801 | 3.012 | 3.372 | 4.221 |
| 14 | 2.264 | 2.415 | 2.624 | 2.771 | 2.977 | 3.326 | 4.140 |
| 15 | 2.249 | 2.397 | 2.602 | 2.746 | 2.947 | 3.286 | 4.073 |
| 16 | 2.235 | 2.382 | 2.583 | 2.724 | 2.921 | 3.252 | 4.015 |
| 17 | 2.224 | 2.368 | 2.567 | 2.706 | 2.898 | 3.222 | 3.965 |
| 18 | 2.214 | 2.356 | 2.552 | 2.689 | 2.878 | 3.197 | 3.922 |
| 19 | 2.205 | 2.346 | 2.539 | 2.674 | 2.861 | 3.174 | 3.883 |
| 20 | 2.197 | 2.336 | 2.528 | 2.661 | 2.845 | 3.153 | 3.850 |
| 21 | 2.189 | 2.328 | 2.518 | 2.649 | 2.831 | 3.135 | 3.819 |
| 22 | 2.183 | 2.320 | 2.508 | 2.639 | 2.819 | 3.119 | 3.792 |
| 23 | 2.177 | 2.313 | 2.500 | 2.629 | 2.807 | 3.104 | 3.768 |
| 24 | 2.172 | 2.307 | 2.492 | 2.620 | 2.797 | 3.091 | 3.745 |
| 25 | 2.167 | 2.301 | 2.485 | 2.612 | 2.787 | 3.078 | 3.725 |
| 26 | 2.162 | 2.296 | 2.479 | 2.605 | 2.779 | 3.067 | 3.707 |
| 27 | 2.158 | 2.291 | 2.473 | 2.598 | 2.771 | 3.057 | 3.689 |
| 28 | 2.154 | 2.286 | 2.467 | 2.592 | 2.763 | 3.047 | 3.674 |
| 29 | 2.150 | 2.282 | 2.462 | 2.586 | 2.756 | 3.038 | 3.660 |
| 30 | 2.147 | 2.278 | 2.457 | 2.581 | 2.750 | 3.030 | 3.646 |
| 40 | 2.123 | 2.250 | 2.423 | 2.542 | 2.704 | 2.971 | 3.551 |
| 60 | 2.099 | 2.223 | 2.390 | 2.504 | 2.660 | 2.915 | 3.460 |
| 120 | 2.076 | 2.196 | 2.358 | 2.468 | 2.617 | 2.860 | 3.373 |
| ∞ | 2.054 | 2.170 | 2.326 | 2.432 | 2.576 | 2.807 | 3.290 |

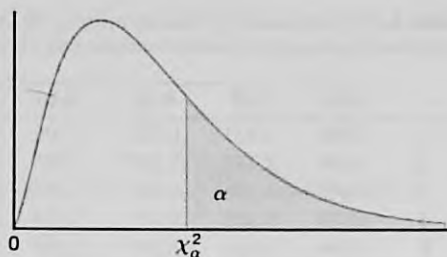


Tabla A.5 Valores críticos de la distribución chi cuadrada

| ν | α | | | | | | | | | |
|-------|----------|--------|--------|--------|---------|--------|--------|--------|--------|--------|
| | 0.995 | 0.99 | 0.98 | 0.975 | 0.95 | 0.90 | 0.80 | 0.75 | 0.70 | 0.50 |
| 1 | 0.00393 | 0.0157 | 0.0328 | 0.0982 | 0.00393 | 0.0158 | 0.0642 | 0.102 | 0.148 | 0.455 |
| 2 | 0.0100 | 0.0201 | 0.0404 | 0.0506 | 0.103 | 0.211 | 0.446 | 0.575 | 0.713 | 1.386 |
| 3 | 0.0717 | 0.115 | 0.185 | 0.216 | 0.352 | 0.584 | 1.005 | 1.213 | 1.424 | 2.366 |
| 4 | 0.207 | 0.297 | 0.429 | 0.484 | 0.711 | 1.064 | 1.649 | 1.923 | 2.195 | 3.357 |
| 5 | 0.412 | 0.554 | 0.752 | 0.831 | 1.145 | 1.610 | 2.343 | 2.675 | 3.000 | 4.351 |
| 6 | 0.676 | 0.872 | 1.134 | 1.237 | 1.635 | 2.204 | 3.070 | 3.455 | 3.828 | 5.348 |
| 7 | 0.989 | 1.239 | 1.564 | 1.690 | 2.167 | 2.833 | 3.822 | 4.255 | 4.671 | 6.346 |
| 8 | 1.344 | 1.647 | 2.032 | 2.180 | 2.733 | 3.490 | 4.594 | 5.071 | 5.527 | 7.344 |
| 9 | 1.735 | 2.088 | 2.532 | 2.700 | 3.325 | 4.168 | 5.380 | 5.899 | 6.393 | 8.343 |
| 10 | 2.156 | 2.558 | 3.059 | 3.247 | 3.940 | 4.865 | 6.179 | 6.737 | 7.267 | 9.342 |
| 11 | 2.603 | 3.053 | 3.609 | 3.816 | 4.575 | 5.578 | 6.989 | 7.584 | 8.148 | 10.341 |
| 12 | 3.074 | 3.571 | 4.178 | 4.404 | 5.226 | 6.304 | 7.807 | 8.438 | 9.034 | 11.340 |
| 13 | 3.565 | 4.107 | 4.765 | 5.009 | 5.892 | 7.041 | 8.634 | 9.299 | 9.926 | 12.340 |
| 14 | 4.075 | 4.660 | 5.368 | 5.629 | 6.571 | 7.790 | 9.467 | 10.165 | 10.821 | 13.339 |
| 15 | 4.601 | 5.229 | 5.985 | 6.262 | 7.261 | 8.547 | 10.307 | 11.037 | 11.721 | 14.339 |
| 16 | 5.142 | 5.812 | 6.614 | 6.908 | 7.962 | 9.312 | 11.152 | 11.912 | 12.624 | 15.338 |
| 17 | 5.697 | 6.408 | 7.255 | 7.564 | 8.672 | 10.085 | 12.002 | 12.792 | 13.531 | 16.338 |
| 18 | 6.265 | 7.015 | 7.906 | 8.231 | 9.390 | 10.865 | 12.857 | 13.675 | 14.440 | 17.338 |
| 19 | 6.844 | 7.633 | 8.567 | 8.907 | 10.117 | 11.651 | 13.716 | 14.562 | 15.352 | 18.338 |
| 20 | 7.434 | 8.260 | 9.237 | 9.591 | 10.851 | 12.443 | 14.578 | 15.452 | 16.266 | 19.337 |
| 21 | 8.034 | 8.897 | 9.915 | 10.283 | 11.591 | 13.240 | 15.445 | 16.344 | 17.182 | 20.337 |
| 22 | 8.643 | 9.542 | 10.600 | 10.982 | 12.338 | 14.041 | 16.314 | 17.240 | 18.101 | 21.337 |
| 23 | 9.260 | 10.196 | 11.293 | 11.689 | 13.091 | 14.848 | 17.187 | 18.137 | 19.021 | 22.337 |
| 24 | 9.886 | 10.856 | 11.992 | 12.401 | 13.848 | 15.659 | 18.062 | 19.037 | 19.943 | 23.337 |
| 25 | 10.520 | 11.524 | 12.697 | 13.120 | 14.611 | 16.473 | 18.940 | 19.939 | 20.867 | 24.337 |
| 26 | 11.160 | 12.198 | 13.409 | 13.844 | 15.379 | 17.292 | 19.820 | 20.843 | 21.792 | 25.336 |
| 27 | 11.808 | 12.878 | 14.125 | 14.573 | 16.151 | 18.114 | 20.703 | 21.749 | 22.719 | 26.336 |
| 28 | 12.461 | 13.565 | 14.847 | 15.308 | 16.928 | 18.939 | 21.588 | 22.657 | 23.647 | 27.336 |
| 29 | 13.121 | 14.256 | 15.574 | 16.047 | 17.708 | 19.768 | 22.475 | 23.567 | 24.577 | 28.336 |
| 30 | 13.787 | 14.953 | 16.306 | 16.791 | 18.493 | 20.599 | 23.364 | 24.478 | 25.508 | 29.336 |
| 40 | 20.707 | 22.164 | 23.838 | 24.433 | 26.509 | 29.051 | 32.345 | 33.66 | 34.872 | 39.335 |
| 50 | 27.991 | 29.707 | 31.664 | 32.357 | 34.764 | 37.689 | 41.449 | 42.942 | 44.313 | 49.335 |
| 60 | 35.534 | 37.485 | 39.699 | 40.482 | 43.188 | 46.459 | 50.641 | 52.294 | 53.809 | 59.335 |

Tabla A.5 (continuación) Valores críticos de la distribución chi cuadrada

| ν | α | | | | | | | | | |
|-------|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 0.30 | 0.25 | 0.20 | 0.10 | 0.05 | 0.025 | 0.02 | 0.01 | 0.005 | 0.001 |
| 1 | 1.074 | 1.323 | 1.642 | 2.706 | 3.841 | 5.024 | 5.412 | 6.635 | 7.879 | 10.827 |
| 2 | 2.408 | 2.773 | 3.219 | 4.605 | 5.991 | 7.378 | 7.824 | 9.210 | 10.597 | 13.815 |
| 3 | 3.665 | 4.108 | 4.642 | 6.251 | 7.815 | 9.348 | 9.837 | 11.345 | 12.838 | 16.266 |
| 4 | 4.878 | 5.385 | 5.989 | 7.779 | 9.488 | 11.143 | 11.668 | 13.277 | 14.860 | 18.466 |
| 5 | 6.064 | 6.626 | 7.289 | 9.236 | 11.070 | 12.832 | 13.388 | 15.086 | 16.750 | 20.515 |
| 6 | 7.231 | 7.841 | 8.558 | 10.645 | 12.592 | 14.449 | 15.033 | 16.812 | 18.548 | 22.457 |
| 7 | 8.383 | 9.037 | 9.803 | 12.017 | 14.067 | 16.013 | 16.622 | 18.475 | 20.278 | 24.321 |
| 8 | 9.524 | 10.219 | 11.030 | 13.362 | 15.507 | 17.535 | 18.168 | 20.090 | 21.955 | 26.124 |
| 9 | 10.656 | 11.389 | 12.242 | 14.684 | 16.919 | 19.023 | 19.679 | 21.666 | 23.589 | 27.877 |
| 10 | 11.781 | 12.549 | 13.442 | 15.987 | 18.307 | 20.483 | 21.161 | 23.209 | 25.188 | 29.588 |
| 11 | 12.899 | 13.701 | 14.631 | 17.275 | 19.675 | 21.920 | 22.618 | 24.725 | 26.757 | 31.264 |
| 12 | 14.011 | 14.845 | 15.812 | 18.549 | 21.026 | 23.337 | 24.054 | 26.217 | 28.300 | 32.909 |
| 13 | 15.119 | 15.984 | 16.985 | 19.812 | 22.362 | 24.736 | 25.471 | 27.688 | 29.819 | 34.527 |
| 14 | 16.222 | 17.117 | 18.151 | 21.064 | 23.685 | 26.119 | 26.873 | 29.141 | 31.319 | 36.124 |
| 15 | 17.322 | 18.245 | 19.311 | 22.307 | 24.996 | 27.488 | 28.259 | 30.578 | 32.801 | 37.698 |
| 16 | 18.418 | 19.369 | 20.465 | 23.542 | 26.296 | 28.845 | 29.633 | 32.000 | 34.267 | 39.252 |
| 17 | 19.511 | 20.489 | 21.615 | 24.769 | 27.587 | 30.191 | 30.995 | 33.409 | 35.718 | 40.791 |
| 18 | 20.601 | 21.605 | 22.760 | 25.989 | 28.869 | 31.526 | 32.346 | 34.805 | 37.156 | 42.312 |
| 19 | 21.689 | 22.718 | 23.900 | 27.204 | 30.144 | 32.852 | 33.687 | 36.191 | 38.582 | 43.819 |
| 20 | 22.775 | 23.828 | 25.038 | 28.412 | 31.410 | 34.170 | 35.020 | 37.566 | 39.997 | 45.314 |
| 21 | 23.858 | 24.935 | 26.171 | 29.615 | 32.671 | 35.479 | 36.343 | 38.932 | 41.401 | 46.796 |
| 22 | 24.939 | 26.039 | 27.301 | 30.813 | 33.924 | 36.781 | 37.659 | 40.289 | 42.796 | 48.268 |
| 23 | 26.018 | 27.141 | 28.429 | 32.007 | 35.172 | 38.076 | 38.968 | 41.638 | 44.181 | 49.728 |
| 24 | 27.096 | 28.241 | 29.553 | 33.196 | 36.415 | 39.364 | 40.270 | 42.980 | 45.558 | 51.179 |
| 25 | 28.172 | 29.339 | 30.675 | 34.382 | 37.652 | 40.646 | 41.566 | 44.314 | 46.928 | 52.619 |
| 26 | 29.246 | 30.435 | 31.795 | 35.563 | 38.885 | 41.923 | 42.856 | 45.642 | 48.290 | 54.051 |
| 27 | 30.319 | 31.528 | 32.912 | 36.741 | 40.113 | 43.195 | 44.140 | 46.963 | 49.645 | 55.475 |
| 28 | 31.391 | 32.620 | 34.027 | 37.916 | 41.337 | 44.461 | 45.419 | 48.278 | 50.994 | 56.892 |
| 29 | 32.461 | 33.711 | 35.139 | 39.087 | 42.557 | 45.722 | 46.693 | 49.588 | 52.335 | 58.301 |
| 30 | 33.530 | 34.800 | 36.250 | 40.256 | 43.773 | 46.979 | 47.962 | 50.892 | 53.672 | 59.702 |
| 40 | 44.165 | 45.616 | 47.269 | 51.805 | 55.758 | 59.342 | 60.436 | 63.691 | 66.766 | 73.403 |
| 50 | 54.723 | 56.334 | 58.164 | 63.167 | 67.505 | 71.420 | 72.613 | 76.154 | 79.490 | 86.660 |
| 60 | 65.226 | 66.981 | 68.972 | 74.397 | 79.082 | 83.298 | 84.58 | 88.379 | 91.952 | 99.608 |

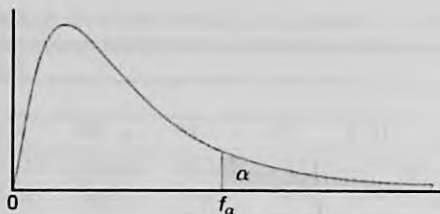


Tabla A.6 Valores críticos de la distribución F

| | | $f_{\alpha,05}(v_1, v_2)$ | | | | | | | | |
|----------|--------|---------------------------|--------|--------|--------|--------|--------|--------|--------|--|
| | | v_1 | | | | | | | | |
| v_2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| 1 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 236.77 | 238.88 | 240.54 | |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 | |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 | |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 | |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 | |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 | |
| 12 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 | |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 | |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 | |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 | |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 | |
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 | |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 | |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 | |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 | |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 | |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 | |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 | |
| 25 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 | |
| 26 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.27 | |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.25 | |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.45 | 2.36 | 2.29 | 2.24 | |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.55 | 2.43 | 2.35 | 2.28 | 2.22 | |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 | |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.12 | |
| 60 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 | |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.18 | 2.09 | 2.02 | 1.96 | |
| ∞ | 3.84 | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.88 | |

Tabla A.6 (continuación) Valores críticos de la distribución F

| v_2 | $f_{\alpha,05}(v_1, v_2)$ | | | | | | | | | |
|----------|---------------------------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| | v_1 | | | | | | | | | |
| | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
| 1 | 241.88 | 243.91 | 245.95 | 248.01 | 249.05 | 250.10 | 251.14 | 252.20 | 253.25 | 254.31 |
| 2 | 19.40 | 19.41 | 19.43 | 19.45 | 19.45 | 19.46 | 19.47 | 19.48 | 19.49 | 19.50 |
| 3 | 8.79 | 8.74 | 8.70 | 8.66 | 8.64 | 8.62 | 8.59 | 8.57 | 8.55 | 8.53 |
| 4 | 5.96 | 5.91 | 5.86 | 5.80 | 5.77 | 5.75 | 5.72 | 5.69 | 5.66 | 5.63 |
| 5 | 4.74 | 4.68 | 4.62 | 4.56 | 4.53 | 4.50 | 4.46 | 4.43 | 4.40 | 4.36 |
| 6 | 4.06 | 4.00 | 3.94 | 3.87 | 3.84 | 3.81 | 3.77 | 3.74 | 3.70 | 3.67 |
| 7 | 3.64 | 3.57 | 3.51 | 3.44 | 3.41 | 3.38 | 3.34 | 3.30 | 3.27 | 3.23 |
| 8 | 3.35 | 3.28 | 3.22 | 3.15 | 3.12 | 3.08 | 3.04 | 3.01 | 2.97 | 2.93 |
| 9 | 3.14 | 3.07 | 3.01 | 2.94 | 2.90 | 2.86 | 2.83 | 2.79 | 2.75 | 2.71 |
| 10 | 2.98 | 2.91 | 2.85 | 2.77 | 2.74 | 2.70 | 2.66 | 2.62 | 2.58 | 2.54 |
| 11 | 2.85 | 2.79 | 2.72 | 2.65 | 2.61 | 2.57 | 2.53 | 2.49 | 2.45 | 2.40 |
| 12 | 2.75 | 2.69 | 2.62 | 2.54 | 2.51 | 2.47 | 2.43 | 2.38 | 2.34 | 2.30 |
| 13 | 2.67 | 2.60 | 2.53 | 2.46 | 2.42 | 2.38 | 2.34 | 2.30 | 2.25 | 2.21 |
| 14 | 2.60 | 2.53 | 2.46 | 2.39 | 2.35 | 2.31 | 2.27 | 2.22 | 2.18 | 2.13 |
| 15 | 2.54 | 2.48 | 2.40 | 2.33 | 2.29 | 2.25 | 2.20 | 2.16 | 2.11 | 2.07 |
| 16 | 2.49 | 2.42 | 2.35 | 2.28 | 2.24 | 2.19 | 2.15 | 2.11 | 2.06 | 2.01 |
| 17 | 2.45 | 2.38 | 2.31 | 2.23 | 2.19 | 2.15 | 2.10 | 2.06 | 2.01 | 1.96 |
| 18 | 2.41 | 2.34 | 2.27 | 2.19 | 2.15 | 2.11 | 2.06 | 2.02 | 1.97 | 1.92 |
| 19 | 2.38 | 2.31 | 2.23 | 2.16 | 2.11 | 2.07 | 2.03 | 1.98 | 1.93 | 1.88 |
| 20 | 2.35 | 2.28 | 2.20 | 2.12 | 2.08 | 2.04 | 1.99 | 1.95 | 1.90 | 1.84 |
| 21 | 2.32 | 2.25 | 2.18 | 2.10 | 2.05 | 2.01 | 1.96 | 1.92 | 1.87 | 1.81 |
| 22 | 2.30 | 2.23 | 2.15 | 2.07 | 2.03 | 1.98 | 1.94 | 1.89 | 1.84 | 1.78 |
| 23 | 2.27 | 2.20 | 2.13 | 2.05 | 2.01 | 1.96 | 1.91 | 1.86 | 1.81 | 1.76 |
| 24 | 2.25 | 2.18 | 2.11 | 2.03 | 1.98 | 1.94 | 1.89 | 1.84 | 1.79 | 1.73 |
| 25 | 2.24 | 2.16 | 2.09 | 2.01 | 1.96 | 1.92 | 1.87 | 1.82 | 1.77 | 1.71 |
| 26 | 2.22 | 2.15 | 2.07 | 1.99 | 1.95 | 1.90 | 1.85 | 1.80 | 1.75 | 1.69 |
| 27 | 2.20 | 2.13 | 2.06 | 1.97 | 1.93 | 1.88 | 1.84 | 1.79 | 1.73 | 1.67 |
| 28 | 2.19 | 2.12 | 2.04 | 1.96 | 1.91 | 1.87 | 1.82 | 1.77 | 1.71 | 1.65 |
| 29 | 2.18 | 2.10 | 2.03 | 1.94 | 1.90 | 1.85 | 1.81 | 1.75 | 1.70 | 1.64 |
| 30 | 2.16 | 2.09 | 2.01 | 1.93 | 1.89 | 1.84 | 1.79 | 1.74 | 1.68 | 1.62 |
| 40 | 2.08 | 2.00 | 1.92 | 1.84 | 1.79 | 1.74 | 1.69 | 1.64 | 1.58 | 1.51 |
| 60 | 1.99 | 1.92 | 1.84 | 1.75 | 1.70 | 1.65 | 1.59 | 1.53 | 1.47 | 1.39 |
| 120 | 1.91 | 1.83 | 1.75 | 1.66 | 1.61 | 1.55 | 1.50 | 1.43 | 1.35 | 1.25 |
| ∞ | 1.83 | 1.75 | 1.67 | 1.57 | 1.52 | 1.46 | 1.39 | 1.32 | 1.22 | 1.00 |

Tabla A.6 (continuación) Valores críticos de la distribución F

| | | $f_{0.01}(v_1, v_2)$ | | | | | | | | |
|----------|---------|----------------------|---------|---------|---------|---------|---------|---------|---------|--|
| | | v_1 | | | | | | | | |
| v_2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| 1 | 4052.18 | 4999.50 | 5403.35 | 5624.58 | 5763.65 | 5858.99 | 5928.36 | 5981.07 | 6022.47 | |
| 2 | 98.50 | 99.00 | 99.17 | 99.25 | 99.30 | 99.33 | 99.36 | 99.37 | 99.39 | |
| 3 | 34.12 | 30.82 | 29.46 | 28.71 | 28.24 | 27.91 | 27.67 | 27.49 | 27.35 | |
| 4 | 21.20 | 18.00 | 16.69 | 15.98 | 15.52 | 15.21 | 14.98 | 14.80 | 14.66 | |
| 5 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.46 | 10.29 | 10.16 | |
| 6 | 13.75 | 10.92 | 9.78 | 9.15 | 8.75 | 8.47 | 8.26 | 8.10 | 7.98 | |
| 7 | 12.25 | 9.55 | 8.45 | 7.85 | 7.46 | 7.19 | 6.99 | 6.84 | 6.72 | |
| 8 | 11.26 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.18 | 6.03 | 5.91 | |
| 9 | 10.56 | 8.02 | 6.99 | 6.42 | 6.06 | 5.80 | 5.61 | 5.47 | 5.35 | |
| 10 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 | |
| 11 | 9.65 | 7.21 | 6.22 | 5.67 | 5.32 | 5.07 | 4.89 | 4.74 | 4.63 | |
| 12 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.39 | |
| 13 | 9.07 | 6.70 | 5.74 | 5.21 | 4.86 | 4.62 | 4.44 | 4.30 | 4.19 | |
| 14 | 8.86 | 6.51 | 5.56 | 5.04 | 4.69 | 4.46 | 4.28 | 4.14 | 4.03 | |
| 15 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.89 | |
| 16 | 8.53 | 6.23 | 5.29 | 4.77 | 4.44 | 4.20 | 4.03 | 3.89 | 3.78 | |
| 17 | 8.40 | 6.11 | 5.18 | 4.67 | 4.34 | 4.10 | 3.93 | 3.79 | 3.68 | |
| 18 | 8.29 | 6.01 | 5.09 | 4.58 | 4.25 | 4.01 | 3.84 | 3.71 | 3.60 | |
| 19 | 8.18 | 5.93 | 5.01 | 4.50 | 4.17 | 3.94 | 3.77 | 3.63 | 3.52 | |
| 20 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.46 | |
| 21 | 8.02 | 5.78 | 4.87 | 4.37 | 4.04 | 3.81 | 3.64 | 3.51 | 3.40 | |
| 22 | 7.95 | 5.72 | 4.82 | 4.31 | 3.99 | 3.76 | 3.59 | 3.45 | 3.35 | |
| 23 | 7.88 | 5.66 | 4.76 | 4.26 | 3.94 | 3.71 | 3.54 | 3.41 | 3.30 | |
| 24 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.50 | 3.36 | 3.26 | |
| 25 | 7.77 | 5.57 | 4.68 | 4.18 | 3.85 | 3.63 | 3.46 | 3.32 | 3.22 | |
| 26 | 7.72 | 5.53 | 4.64 | 4.14 | 3.82 | 3.59 | 3.42 | 3.29 | 3.18 | |
| 27 | 7.68 | 5.49 | 4.60 | 4.11 | 3.78 | 3.56 | 3.39 | 3.26 | 3.15 | |
| 28 | 7.64 | 5.45 | 4.57 | 4.07 | 3.75 | 3.53 | 3.36 | 3.23 | 3.12 | |
| 29 | 7.60 | 5.42 | 4.54 | 4.04 | 3.73 | 3.50 | 3.33 | 3.20 | 3.09 | |
| 30 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.30 | 3.17 | 3.07 | |
| 40 | 7.31 | 5.18 | 4.31 | 3.83 | 3.51 | 3.29 | 3.12 | 2.99 | 2.89 | |
| 60 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.95 | 2.82 | 2.72 | |
| 120 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.79 | 2.66 | 2.56 | |
| ∞ | 6.63 | 4.61 | 3.78 | 3.32 | 3.02 | 2.80 | 2.64 | 2.51 | 2.41 | |

Tabla A.6 (continuación) Valores críticos de la distribución F

| v_2 | $f_{0.01}(v_1, v_2)$ | | | | | | | | | |
|----------|----------------------|---------|---------|---------|---------|---------|---------|---------|---------|----------|
| | v_1 | | | | | | | | | |
| | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
| 1 | 6055.85 | 6106.32 | 6157.28 | 6208.73 | 6234.63 | 6260.65 | 6286.78 | 6313.03 | 6339.39 | 6365.86 |
| 2 | 99.40 | 99.42 | 99.43 | 99.45 | 99.46 | 99.47 | 99.47 | 99.48 | 99.49 | 99.50 |
| 3 | 27.23 | 27.05 | 26.87 | 26.69 | 26.60 | 26.50 | 26.41 | 26.32 | 26.22 | 26.13 |
| 4 | 14.55 | 14.37 | 14.20 | 14.02 | 13.93 | 13.84 | 13.75 | 13.65 | 13.56 | 13.46 |
| 5 | 10.05 | 9.89 | 9.72 | 9.55 | 9.47 | 9.38 | 9.29 | 9.20 | 9.11 | 9.02 |
| 6 | 7.87 | 7.72 | 7.56 | 7.40 | 7.31 | 7.23 | 7.14 | 7.06 | 6.97 | 6.88 |
| 7 | 6.62 | 6.47 | 6.31 | 6.16 | 6.07 | 5.99 | 5.91 | 5.82 | 5.74 | 5.65 |
| 8 | 5.81 | 5.67 | 5.52 | 5.36 | 5.28 | 5.20 | 5.12 | 5.03 | 4.95 | 4.86 |
| 9 | 5.26 | 5.11 | 4.96 | 4.81 | 4.73 | 4.65 | 4.57 | 4.48 | 4.40 | 4.31 |
| 10 | 4.85 | 4.71 | 4.56 | 4.41 | 4.33 | 4.25 | 4.17 | 4.08 | 4.00 | 3.91 |
| 11 | 4.54 | 4.40 | 4.25 | 4.10 | 4.02 | 3.94 | 3.86 | 3.78 | 3.69 | 3.60 |
| 12 | 4.30 | 4.16 | 4.01 | 3.86 | 3.78 | 3.70 | 3.62 | 3.54 | 3.45 | 3.36 |
| 13 | 4.10 | 3.96 | 3.82 | 3.66 | 3.59 | 3.51 | 3.43 | 3.34 | 3.25 | 3.17 |
| 14 | 3.94 | 3.80 | 3.66 | 3.51 | 3.43 | 3.35 | 3.27 | 3.18 | 3.09 | 3.00 |
| 15 | 3.80 | 3.67 | 3.52 | 3.37 | 3.29 | 3.21 | 3.13 | 3.05 | 2.96 | 2.87 |
| 16 | 3.69 | 3.55 | 3.41 | 3.26 | 3.18 | 3.10 | 3.02 | 2.93 | 2.84 | 2.75 |
| 17 | 3.59 | 3.46 | 3.31 | 3.16 | 3.08 | 3.00 | 2.92 | 2.83 | 2.75 | 2.65 |
| 18 | 3.51 | 3.37 | 3.23 | 3.08 | 3.00 | 2.92 | 2.84 | 2.75 | 2.66 | 2.57 |
| 19 | 3.43 | 3.30 | 3.15 | 3.00 | 2.92 | 2.84 | 2.76 | 2.67 | 2.58 | 2.49 |
| 20 | 3.37 | 3.23 | 3.09 | 2.94 | 2.86 | 2.78 | 2.69 | 2.61 | 2.52 | 2.42 |
| 21 | 3.31 | 3.17 | 3.03 | 2.88 | 2.80 | 2.72 | 2.64 | 2.55 | 2.46 | 2.36 |
| 22 | 3.26 | 3.12 | 2.98 | 2.83 | 2.75 | 2.67 | 2.58 | 2.50 | 2.40 | 2.31 |
| 23 | 3.21 | 3.07 | 2.93 | 2.78 | 2.70 | 2.62 | 2.54 | 2.45 | 2.35 | 2.26 |
| 24 | 3.17 | 3.03 | 2.89 | 2.74 | 2.66 | 2.58 | 2.49 | 2.40 | 2.31 | 2.21 |
| 25 | 3.13 | 2.99 | 2.85 | 2.70 | 2.62 | 2.54 | 2.45 | 2.36 | 2.27 | 2.17 |
| 26 | 3.09 | 2.96 | 2.81 | 2.66 | 2.58 | 2.50 | 2.42 | 2.33 | 2.23 | 2.13 |
| 27 | 3.06 | 2.93 | 2.78 | 2.63 | 2.55 | 2.47 | 2.38 | 2.29 | 2.20 | 2.10 |
| 28 | 3.03 | 2.90 | 2.75 | 2.60 | 2.52 | 2.44 | 2.35 | 2.26 | 2.17 | 2.06 |
| 29 | 3.00 | 2.87 | 2.73 | 2.57 | 2.49 | 2.41 | 2.33 | 2.23 | 2.14 | 2.03 |
| 30 | 2.98 | 2.84 | 2.70 | 2.55 | 2.47 | 2.39 | 2.30 | 2.21 | 2.11 | 2.01 |
| 40 | 2.80 | 2.66 | 2.52 | 2.37 | 2.29 | 2.20 | 2.11 | 2.02 | 1.92 | 1.80 |
| 60 | 2.63 | 2.50 | 2.35 | 2.20 | 2.12 | 2.03 | 1.94 | 1.84 | 1.73 | 1.60 |
| 120 | 2.47 | 2.34 | 2.19 | 2.03 | 1.95 | 1.86 | 1.76 | 1.66 | 1.53 | 1.38 |
| ∞ | 2.32 | 2.18 | 2.04 | 1.88 | 1.79 | 1.70 | 1.59 | 1.47 | 1.32 | 1.00 |

Tabla A.7 Factores de tolerancia para distribuciones normales

| n | Intervalos bilaterales | | | | | | Intervalos unilaterales | | | | | |
|----------|------------------------|--------|--------|-----------------|---------|---------|-------------------------|--------|--------|-----------------|---------|---------|
| | $\gamma = 0.05$ | | | $\gamma = 0.01$ | | | $\gamma = 0.05$ | | | $\gamma = 0.01$ | | |
| | 1 - α | | | 1 - α | | | 1 - α | | | 1 - α | | |
| | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 | 0.90 | 0.95 | 0.99 |
| 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 | 17.170 | 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 | 3.001 |
| 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 |

Adaptada de C. Eisenhart, M. W. Hastay y W. A. Wallis, *Techniques of Statistical Analysis*, capítulo 2, McGraw-Hill Book Company, Nueva York, 1947. Se utiliza con autorización de McGraw-Hill Book Company.

Tabla A.8 Tamaño muestral para la prueba t de la media

| | | Nivel de la prueba t | | | | | | | | | | | | | | | | | | | | | |
|-------------------|------|------------------------|-----|-----|-----|-----------------|-----|-----|-----|------------------|----|-----|-----|-----------------|-----|----|-----|-----|-----|-----|-----|----|----|
| | | $\alpha = 0.005$ | | | | $\alpha = 0.01$ | | | | $\alpha = 0.025$ | | | | $\alpha = 0.05$ | | | | | | | | | |
| Prueba unilateral | | $\alpha = 0.01$ | | | | $\alpha = 0.02$ | | | | $\alpha = 0.05$ | | | | $\alpha = 0.1$ | | | | | | | | | |
| Prueba bilateral | | $\alpha = 0.01$ | | | | $\alpha = 0.02$ | | | | $\alpha = 0.05$ | | | | $\alpha = 0.1$ | | | | | | | | | |
| $\beta = 0.1$ | | .01 | .05 | .1 | .2 | .5 | .01 | .05 | .1 | .2 | .5 | .01 | .05 | .1 | .2 | .5 | .01 | .05 | .1 | .2 | .5 | | |
| | 0.05 | | | | | | | | | | | | | | | | | | | | | | |
| | 0.10 | | | | | | | | | | | | | | | | | | | | | | |
| | 0.15 | | | | | | | | | | | | | | | | | | | | 122 | | |
| | 0.20 | | | | | | | | | 139 | | | | | | 99 | | | | | 70 | | |
| | 0.25 | | | | | 110 | | | | 90 | | | | | 128 | 64 | | | 139 | 101 | 45 | | |
| | 0.30 | | | | 134 | 78 | | | | 115 | 63 | | | | 119 | 90 | 45 | | 122 | 97 | 71 | 32 | |
| | 0.35 | | | 125 | 99 | 58 | | | | 109 | 85 | 47 | | | 109 | 88 | 67 | 34 | | 90 | 72 | 52 | 24 |
| | 0.40 | | 115 | 97 | 77 | 45 | | | 101 | 85 | 66 | 37 | 117 | | 84 | 68 | 51 | 26 | 101 | 70 | 55 | 40 | 19 |
| | 0.45 | | 92 | 77 | 62 | 37 | 110 | | 81 | 68 | 53 | 30 | 93 | 67 | 54 | 41 | 21 | 80 | 55 | 44 | 33 | 15 | |
| | 0.50 | 100 | 75 | 63 | 51 | 30 | 90 | 66 | 55 | 43 | 25 | 76 | 54 | 44 | 34 | 18 | 65 | 45 | 36 | 27 | 13 | | |
| | 0.55 | 83 | 63 | 53 | 42 | 26 | 75 | 55 | 46 | 36 | 21 | 63 | 45 | 37 | 28 | 15 | 54 | 38 | 30 | 22 | 11 | | |
| | 0.60 | 71 | 53 | 45 | 36 | 22 | 63 | 47 | 39 | 31 | 18 | 53 | 38 | 32 | 24 | 13 | 46 | 32 | 26 | 19 | 9 | | |
| | 0.65 | 61 | 46 | 39 | 31 | 20 | 55 | 41 | 34 | 27 | 16 | 46 | 33 | 27 | 21 | 12 | 39 | 28 | 22 | 17 | 8 | | |
| | 0.70 | 53 | 40 | 34 | 28 | 17 | 47 | 35 | 30 | 24 | 14 | 40 | 29 | 24 | 19 | 10 | 34 | 24 | 19 | 15 | 8 | | |
| | 0.75 | 47 | 36 | 30 | 25 | 16 | 42 | 31 | 27 | 21 | 1 | 35 | 26 | 21 | 16 | 9 | 30 | 21 | 17 | 13 | 7 | | |
| | 0.80 | 41 | 32 | 27 | 22 | 14 | 37 | 28 | 24 | 19 | 12 | 31 | 22 | 19 | 15 | 9 | 27 | 19 | 15 | 12 | 6 | | |
| | 0.85 | 37 | 29 | 24 | 20 | 13 | 33 | 25 | 21 | 17 | 11 | 28 | 21 | 17 | 13 | 8 | 24 | 17 | 14 | 11 | 6 | | |
| | 0.90 | 34 | 26 | 22 | 18 | 12 | 29 | 23 | 19 | 16 | 10 | 25 | 19 | 16 | 12 | 7 | 21 | 15 | 13 | 10 | 5 | | |
| | 0.95 | 31 | 24 | 20 | 17 | 11 | 27 | 21 | 18 | 14 | 9 | 23 | 17 | 14 | 11 | 7 | 19 | 14 | 11 | 9 | 5 | | |
| | 1.00 | 28 | 22 | 19 | 16 | 10 | 25 | 19 | 16 | 13 | 9 | 21 | 16 | 13 | 10 | 6 | 18 | 13 | 11 | 8 | 5 | | |
| | 1.1 | 24 | 19 | 16 | 14 | 9 | 21 | 16 | 14 | 12 | 8 | 18 | 13 | 11 | 9 | 6 | 15 | 11 | 9 | 7 | | | |
| | 1.2 | 21 | 16 | 14 | 12 | 8 | 18 | 14 | 12 | 10 | 7 | 15 | 12 | 10 | 8 | 5 | 13 | 10 | 8 | 6 | | | |
| | 1.3 | 18 | 15 | 13 | 11 | 8 | 16 | 13 | 11 | 9 | 6 | 13 | 10 | 9 | 7 | | 11 | 8 | 7 | 6 | | | |
| | 1.4 | 16 | 13 | 12 | 10 | 7 | 14 | 11 | 10 | 9 | 6 | 12 | 9 | 8 | 7 | | 10 | 8 | 7 | 5 | | | |
| | 1.5 | 15 | 12 | 11 | 9 | 7 | 13 | 10 | 9 | 8 | 6 | 11 | 8 | 7 | 6 | | 9 | 7 | 6 | | | | |
| | 1.6 | 13 | 11 | 10 | 8 | 6 | 12 | 10 | 9 | 7 | 5 | 10 | 8 | 7 | 6 | | 8 | 6 | 6 | | | | |
| | 1.7 | 12 | 10 | 9 | 8 | 6 | 11 | 9 | 8 | 7 | | 9 | 7 | 6 | 5 | | 8 | 6 | 5 | | | | |
| | 1.8 | 12 | 10 | 9 | 8 | 6 | 10 | 8 | 7 | 7 | | 8 | 7 | 6 | | | 7 | 6 | | | | | |
| | 1.9 | 11 | 9 | 8 | 7 | 6 | 10 | 8 | 7 | 6 | | 8 | 6 | 6 | | | 7 | 5 | | | | | |
| | 2.0 | 10 | 8 | 8 | 7 | 5 | 9 | 7 | 7 | 6 | | 7 | 6 | 5 | | | 6 | | | | | | |
| | 2.1 | 10 | 8 | 7 | 7 | | 8 | 7 | 6 | 6 | | 7 | 6 | | | | 6 | | | | | | |
| | 2.2 | 9 | 8 | 7 | 6 | | 8 | 7 | 6 | 5 | | 7 | 6 | | | | 6 | | | | | | |
| | 2.3 | 9 | 7 | 7 | 6 | | 8 | 6 | 6 | | | 6 | 5 | | | | 5 | | | | | | |
| | 2.4 | 8 | 7 | 7 | 6 | | 7 | 6 | 6 | | | 6 | | | | | | | | | | | |
| | 2.5 | 8 | 7 | 6 | 6 | | 7 | 6 | 6 | | | 6 | | | | | | | | | | | |
| | 3.0 | 7 | 6 | 6 | 5 | | 6 | 5 | 5 | | | 5 | | | | | | | | | | | |
| | 3.5 | 6 | 5 | 5 | | | 5 | | | | | | | | | | | | | | | | |
| | 4.0 | 6 | | | | | | | | | | | | | | | | | | | | | |

Reproducida con autorización de O. L. Davies, ed., *Design and Analysis of Industrial Experiments*, Oliver & Boyd, Edimburgo, 1956.

Tabla A.10 Valores críticos para la prueba de Bartlett

| $b_{\lambda}(0.01; n)$ | | | | | | | | | |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Número de poblaciones, k | | | | | | | | | |
| n | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 3 | 0.1411 | 0.1672 | | | | | | | |
| 4 | 0.2843 | 0.3165 | 0.3475 | 0.3729 | 0.3937 | 0.4110 | | | |
| 5 | 0.3984 | 0.4304 | 0.4607 | 0.4850 | 0.5046 | 0.5207 | 0.5343 | 0.5458 | 0.5558 |
| 6 | 0.4850 | 0.5149 | 0.5430 | 0.5653 | 0.5832 | 0.5978 | 0.6100 | 0.6204 | 0.6293 |
| 7 | 0.5512 | 0.5787 | 0.6045 | 0.6248 | 0.6410 | 0.6542 | 0.6652 | 0.6744 | 0.6824 |
| 8 | 0.6031 | 0.6282 | 0.6518 | 0.6704 | 0.6851 | 0.6970 | 0.7069 | 0.7153 | 0.7225 |
| 9 | 0.6445 | 0.6676 | 0.6892 | 0.7062 | 0.7197 | 0.7305 | 0.7395 | 0.7471 | 0.7536 |
| 10 | 0.6783 | 0.6996 | 0.7195 | 0.7352 | 0.7475 | 0.7575 | 0.7657 | 0.7726 | 0.7786 |
| 11 | 0.7063 | 0.7260 | 0.7445 | 0.7590 | 0.7703 | 0.7795 | 0.7871 | 0.7935 | 0.7990 |
| 12 | 0.7299 | 0.7483 | 0.7654 | 0.7789 | 0.7894 | 0.7980 | 0.8050 | 0.8109 | 0.8160 |
| 13 | 0.7501 | 0.7672 | 0.7832 | 0.7958 | 0.8056 | 0.8135 | 0.8201 | 0.8256 | 0.8303 |
| 14 | 0.7674 | 0.7835 | 0.7985 | 0.8103 | 0.8195 | 0.8269 | 0.8330 | 0.8382 | 0.8426 |
| 15 | 0.7825 | 0.7977 | 0.8118 | 0.8229 | 0.8315 | 0.8385 | 0.8443 | 0.8491 | 0.8532 |
| 16 | 0.7958 | 0.8101 | 0.8235 | 0.8339 | 0.8421 | 0.8486 | 0.8541 | 0.8586 | 0.8625 |
| 17 | 0.8076 | 0.8211 | 0.8338 | 0.8436 | 0.8514 | 0.8576 | 0.8627 | 0.8670 | 0.8707 |
| 18 | 0.8181 | 0.8309 | 0.8429 | 0.8523 | 0.8596 | 0.8655 | 0.8704 | 0.8745 | 0.8780 |
| 19 | 0.8275 | 0.8397 | 0.8512 | 0.8601 | 0.8670 | 0.8727 | 0.8773 | 0.8811 | 0.8845 |
| 20 | 0.8360 | 0.8476 | 0.8586 | 0.8671 | 0.8737 | 0.8791 | 0.8835 | 0.8871 | 0.8903 |
| 21 | 0.8437 | 0.8548 | 0.8653 | 0.8734 | 0.8797 | 0.8848 | 0.8890 | 0.8926 | 0.8956 |
| 22 | 0.8507 | 0.8614 | 0.8714 | 0.8791 | 0.8852 | 0.8901 | 0.8941 | 0.8975 | 0.9004 |
| 23 | 0.8571 | 0.8673 | 0.8769 | 0.8844 | 0.8902 | 0.8949 | 0.8988 | 0.9020 | 0.9047 |
| 24 | 0.8630 | 0.8728 | 0.8820 | 0.8892 | 0.8948 | 0.8993 | 0.9030 | 0.9061 | 0.9087 |
| 25 | 0.8684 | 0.8779 | 0.8867 | 0.8936 | 0.8990 | 0.9034 | 0.9069 | 0.9099 | 0.9124 |
| 26 | 0.8734 | 0.8825 | 0.8911 | 0.8977 | 0.9029 | 0.9071 | 0.9105 | 0.9134 | 0.9158 |
| 27 | 0.8781 | 0.8869 | 0.8951 | 0.9015 | 0.9065 | 0.9105 | 0.9138 | 0.9166 | 0.9190 |
| 28 | 0.8824 | 0.8909 | 0.8988 | 0.9050 | 0.9099 | 0.9138 | 0.9169 | 0.9196 | 0.9219 |
| 29 | 0.8864 | 0.8946 | 0.9023 | 0.9083 | 0.9130 | 0.9167 | 0.9198 | 0.9224 | 0.9246 |
| 30 | 0.8902 | 0.8981 | 0.9056 | 0.9114 | 0.9159 | 0.9195 | 0.9225 | 0.9250 | 0.9271 |
| 40 | 0.9175 | 0.9235 | 0.9291 | 0.9335 | 0.9370 | 0.9397 | 0.9420 | 0.9439 | 0.9455 |
| 50 | 0.9339 | 0.9387 | 0.9433 | 0.9468 | 0.9496 | 0.9518 | 0.9536 | 0.9551 | 0.9564 |
| 60 | 0.9449 | 0.9489 | 0.9527 | 0.9557 | 0.9580 | 0.9599 | 0.9614 | 0.9626 | 0.9637 |
| 80 | 0.9586 | 0.9617 | 0.9646 | 0.9668 | 0.9685 | 0.9699 | 0.9711 | 0.9720 | 0.9728 |
| 100 | 0.9669 | 0.9693 | 0.9716 | 0.9734 | 0.9748 | 0.9759 | 0.9769 | 0.9776 | 0.9783 |

Reproducida de D. D. Dyer y J. P. Keating, "On the Determination of Critical Values for Bartlett's Test", *J. Am. Stat. Assoc.*, 75, 1980, con autorización del consejo de directores.

Tabla A.10 (continuación) Valores críticos para la prueba de Bartlett

| $b_k(0.05; n)$ | | | | | | | | | |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Número de poblaciones, k | | | | | | | | | |
| n | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 3 | 0.3123 | 0.3058 | 0.3173 | 0.3299 | | | | | |
| 4 | 0.4780 | 0.4699 | 0.4803 | 0.4921 | 0.5028 | 0.5122 | 0.5204 | 0.5277 | 0.5341 |
| 5 | 0.5845 | 0.5762 | 0.5850 | 0.5952 | 0.6045 | 0.6126 | 0.6197 | 0.6260 | 0.6315 |
| 6 | 0.6563 | 0.6483 | 0.6559 | 0.6646 | 0.6727 | 0.6798 | 0.6860 | 0.6914 | 0.6961 |
| 7 | 0.7075 | 0.7000 | 0.7065 | 0.7142 | 0.7213 | 0.7275 | 0.7329 | 0.7376 | 0.7418 |
| 8 | 0.7456 | 0.7387 | 0.7444 | 0.7512 | 0.7574 | 0.7629 | 0.7677 | 0.7719 | 0.7757 |
| 9 | 0.7751 | 0.7686 | 0.7737 | 0.7798 | 0.7854 | 0.7903 | 0.7946 | 0.7984 | 0.8017 |
| 10 | 0.7984 | 0.7924 | 0.7970 | 0.8025 | 0.8076 | 0.8121 | 0.8160 | 0.8194 | 0.8224 |
| 11 | 0.8175 | 0.8118 | 0.8160 | 0.8210 | 0.8257 | 0.8298 | 0.8333 | 0.8365 | 0.8392 |
| 12 | 0.8332 | 0.8280 | 0.8317 | 0.8364 | 0.8407 | 0.8444 | 0.8477 | 0.8506 | 0.8531 |
| 13 | 0.8465 | 0.8415 | 0.8450 | 0.8493 | 0.8533 | 0.8568 | 0.8598 | 0.8625 | 0.8648 |
| 14 | 0.8578 | 0.8532 | 0.8564 | 0.8604 | 0.8641 | 0.8673 | 0.8701 | 0.8726 | 0.8748 |
| 15 | 0.8676 | 0.8632 | 0.8662 | 0.8699 | 0.8734 | 0.8764 | 0.8790 | 0.8814 | 0.8834 |
| 16 | 0.8761 | 0.8719 | 0.8747 | 0.8782 | 0.8815 | 0.8843 | 0.8868 | 0.8890 | 0.8909 |
| 17 | 0.8836 | 0.8796 | 0.8823 | 0.8856 | 0.8886 | 0.8913 | 0.8936 | 0.8957 | 0.8975 |
| 18 | 0.8902 | 0.8865 | 0.8890 | 0.8921 | 0.8949 | 0.8975 | 0.8997 | 0.9016 | 0.9033 |
| 19 | 0.8961 | 0.8926 | 0.8949 | 0.8979 | 0.9006 | 0.9030 | 0.9051 | 0.9069 | 0.9086 |
| 20 | 0.9015 | 0.8980 | 0.9003 | 0.9031 | 0.9057 | 0.9080 | 0.9100 | 0.9117 | 0.9132 |
| 21 | 0.9063 | 0.9030 | 0.9051 | 0.9078 | 0.9103 | 0.9124 | 0.9143 | 0.9160 | 0.9175 |
| 22 | 0.9106 | 0.9075 | 0.9095 | 0.9120 | 0.9144 | 0.9165 | 0.9183 | 0.9199 | 0.9213 |
| 23 | 0.9146 | 0.9116 | 0.9135 | 0.9159 | 0.9182 | 0.9202 | 0.9219 | 0.9235 | 0.9248 |
| 24 | 0.9182 | 0.9153 | 0.9172 | 0.9195 | 0.9217 | 0.9236 | 0.9253 | 0.9267 | 0.9280 |
| 25 | 0.9216 | 0.9187 | 0.9205 | 0.9228 | 0.9249 | 0.9267 | 0.9283 | 0.9297 | 0.9309 |
| 26 | 0.9246 | 0.9219 | 0.9236 | 0.9258 | 0.9278 | 0.9296 | 0.9311 | 0.9325 | 0.9336 |
| 27 | 0.9275 | 0.9249 | 0.9265 | 0.9286 | 0.9305 | 0.9322 | 0.9337 | 0.9350 | 0.9361 |
| 28 | 0.9301 | 0.9276 | 0.9292 | 0.9312 | 0.9330 | 0.9347 | 0.9361 | 0.9374 | 0.9385 |
| 29 | 0.9326 | 0.9301 | 0.9316 | 0.9336 | 0.9354 | 0.9370 | 0.9383 | 0.9396 | 0.9406 |
| 30 | 0.9348 | 0.9325 | 0.9340 | 0.9358 | 0.9376 | 0.9391 | 0.9404 | 0.9416 | 0.9426 |
| 40 | 0.9513 | 0.9495 | 0.9506 | 0.9520 | 0.9533 | 0.9545 | 0.9555 | 0.9564 | 0.9572 |
| 50 | 0.9612 | 0.9597 | 0.9606 | 0.9617 | 0.9628 | 0.9637 | 0.9645 | 0.9652 | 0.9658 |
| 60 | 0.9677 | 0.9665 | 0.9672 | 0.9681 | 0.9690 | 0.9698 | 0.9705 | 0.9710 | 0.9716 |
| 80 | 0.9758 | 0.9749 | 0.9754 | 0.9761 | 0.9768 | 0.9774 | 0.9779 | 0.9783 | 0.9787 |
| 100 | 0.9807 | 0.9799 | 0.9804 | 0.9809 | 0.9815 | 0.9819 | 0.9823 | 0.9827 | 0.9830 |

Tabla A.11 Valores críticos para la prueba de Cochran

| | | $\alpha = 0.01$ | | | | | | | | | | | | | |
|----------|--|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| | | n | | | | | | | | | | | | | |
| k | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 17 | 37 | 145 | ∞ |
| 2 | | 0.9999 | 0.9950 | 0.9794 | 0.9586 | 0.9373 | 0.9172 | 0.8988 | 0.8823 | 0.8674 | 0.8539 | 0.7949 | 0.7067 | 0.6062 | 0.5000 |
| 3 | | 0.9933 | 0.9423 | 0.8831 | 0.8335 | 0.7933 | 0.7606 | 0.7335 | 0.7107 | 0.6912 | 0.6743 | 0.6059 | 0.5153 | 0.4230 | 0.3333 |
| 4 | | 0.9676 | 0.8643 | 0.7814 | 0.7212 | 0.6761 | 0.6410 | 0.6129 | 0.5897 | 0.5702 | 0.5536 | 0.4884 | 0.4057 | 0.3251 | 0.2500 |
| 5 | | 0.9279 | 0.7885 | 0.6957 | 0.6329 | 0.5875 | 0.5531 | 0.5259 | 0.5037 | 0.4854 | 0.4697 | 0.4094 | 0.3351 | 0.2644 | 0.2000 |
| 6 | | 0.8828 | 0.7218 | 0.6258 | 0.5635 | 0.5195 | 0.4866 | 0.4608 | 0.4401 | 0.4229 | 0.4084 | 0.3529 | 0.2858 | 0.2229 | 0.1667 |
| 7 | | 0.8376 | 0.6644 | 0.5685 | 0.5080 | 0.4659 | 0.4347 | 0.4105 | 0.3911 | 0.3751 | 0.3616 | 0.3105 | 0.2494 | 0.1929 | 0.1429 |
| 8 | | 0.7945 | 0.6152 | 0.5209 | 0.4627 | 0.4226 | 0.3932 | 0.3704 | 0.3522 | 0.3373 | 0.3248 | 0.2779 | 0.2214 | 0.1700 | 0.1250 |
| 9 | | 0.7544 | 0.5727 | 0.4810 | 0.4251 | 0.3870 | 0.3592 | 0.3378 | 0.3207 | 0.3067 | 0.2950 | 0.2514 | 0.1992 | 0.1521 | 0.1111 |
| 10 | | 0.7175 | 0.5358 | 0.4469 | 0.3934 | 0.3572 | 0.3308 | 0.3106 | 0.2945 | 0.2813 | 0.2704 | 0.2297 | 0.1811 | 0.1376 | 0.1000 |
| 12 | | 0.6528 | 0.4751 | 0.3919 | 0.3428 | 0.3099 | 0.2861 | 0.2680 | 0.2535 | 0.2419 | 0.2320 | 0.1961 | 0.1535 | 0.1157 | 0.0833 |
| 15 | | 0.5747 | 0.4069 | 0.3317 | 0.2882 | 0.2593 | 0.2386 | 0.2228 | 0.2104 | 0.2002 | 0.1918 | 0.1612 | 0.1251 | 0.0934 | 0.0667 |
| 20 | | 0.4799 | 0.3297 | 0.2654 | 0.2288 | 0.2048 | 0.1877 | 0.1748 | 0.1646 | 0.1567 | 0.1501 | 0.1248 | 0.0960 | 0.0709 | 0.0500 |
| 24 | | 0.4247 | 0.2871 | 0.2295 | 0.1970 | 0.1759 | 0.1608 | 0.1495 | 0.1406 | 0.1338 | 0.1283 | 0.1060 | 0.0810 | 0.0595 | 0.0417 |
| 30 | | 0.3632 | 0.2412 | 0.1913 | 0.1635 | 0.1454 | 0.1327 | 0.1232 | 0.1157 | 0.1100 | 0.1054 | 0.0867 | 0.0658 | 0.0480 | 0.0333 |
| 40 | | 0.2940 | 0.1915 | 0.1508 | 0.1281 | 0.1135 | 0.1033 | 0.0957 | 0.0898 | 0.0853 | 0.0816 | 0.0668 | 0.0503 | 0.0363 | 0.0250 |
| 60 | | 0.2151 | 0.1371 | 0.1069 | 0.0902 | 0.0796 | 0.0722 | 0.0668 | 0.0625 | 0.0594 | 0.0567 | 0.0461 | 0.0344 | 0.0245 | 0.0167 |
| 120 | | 0.1225 | 0.0759 | 0.0585 | 0.0489 | 0.0429 | 0.0387 | 0.0357 | 0.0334 | 0.0316 | 0.0302 | 0.0242 | 0.0178 | 0.0125 | 0.0083 |
| ∞ | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Reproducida de C. Eisenhart, M. W. Hastay y W. A. Wallis, *Techniques of Statistical Analysis*, capítulo 15, McGraw-Hill Book Company, Nueva York, 1947. Utilizada con autorización de McGraw-Hill Book Company.

Tabla A.11 (continuación) Valores críticos para la prueba de Cochran

| $\alpha = 0.05$ | | | | | | | | | | | | | | |
|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| k | n | | | | | | | | | | | | | |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 17 | 37 | 145 | ∞ |
| 2 | 0.9985 | 0.9750 | 0.9392 | 0.9057 | 0.8772 | 0.8534 | 0.8332 | 0.8159 | 0.8010 | 0.7880 | 0.7341 | 0.6602 | 0.5813 | 0.5000 |
| 3 | 0.9669 | 0.8709 | 0.7977 | 0.7457 | 0.7071 | 0.6771 | 0.6530 | 0.6333 | 0.6167 | 0.6025 | 0.5466 | 0.4748 | 0.4031 | 0.3333 |
| 4 | 0.9065 | 0.7679 | 0.6841 | 0.6287 | 0.5895 | 0.5598 | 0.5365 | 0.5175 | 0.5017 | 0.4884 | 0.4366 | 0.3720 | 0.3093 | 0.2500 |
| 5 | 0.8412 | 0.6838 | 0.5981 | 0.5441 | 0.5065 | 0.4783 | 0.4564 | 0.4387 | 0.4241 | 0.4118 | 0.3645 | 0.3066 | 0.2513 | 0.2000 |
| 6 | 0.7808 | 0.6161 | 0.5321 | 0.4803 | 0.4447 | 0.4184 | 0.3980 | 0.3817 | 0.3682 | 0.3568 | 0.3135 | 0.2612 | 0.2119 | 0.1667 |
| 7 | 0.7271 | 0.5612 | 0.4800 | 0.4307 | 0.3974 | 0.3726 | 0.3535 | 0.3384 | 0.3259 | 0.3154 | 0.2756 | 0.2278 | 0.1833 | 0.1429 |
| 8 | 0.6798 | 0.5157 | 0.4377 | 0.3910 | 0.3595 | 0.3362 | 0.3185 | 0.3043 | 0.2926 | 0.2829 | 0.2462 | 0.2022 | 0.1616 | 0.1250 |
| 9 | 0.6385 | 0.4775 | 0.4027 | 0.3584 | 0.3286 | 0.3067 | 0.2901 | 0.2768 | 0.2659 | 0.2568 | 0.2226 | 0.1820 | 0.1446 | 0.1111 |
| 10 | 0.6020 | 0.4450 | 0.3733 | 0.3311 | 0.3029 | 0.2823 | 0.2666 | 0.2541 | 0.2439 | 0.2353 | 0.2032 | 0.1655 | 0.1308 | 0.1000 |
| 12 | 0.5410 | 0.3924 | 0.3264 | 0.2880 | 0.2624 | 0.2439 | 0.2299 | 0.2187 | 0.2098 | 0.2020 | 0.1737 | 0.1403 | 0.1100 | 0.0833 |
| 15 | 0.4709 | 0.3346 | 0.2758 | 0.2419 | 0.2195 | 0.2034 | 0.1911 | 0.1815 | 0.1736 | 0.1671 | 0.1429 | 0.1144 | 0.0889 | 0.0667 |
| 20 | 0.3894 | 0.2705 | 0.2205 | 0.1921 | 0.1735 | 0.1602 | 0.1501 | 0.1422 | 0.1357 | 0.1303 | 0.1108 | 0.0879 | 0.0675 | 0.0500 |
| 24 | 0.3434 | 0.2354 | 0.1907 | 0.1656 | 0.1493 | 0.1374 | 0.1286 | 0.1216 | 0.1160 | 0.1113 | 0.0942 | 0.0743 | 0.0567 | 0.0417 |
| 30 | 0.2929 | 0.1980 | 0.1593 | 0.1377 | 0.1237 | 0.1137 | 0.1061 | 0.1002 | 0.0958 | 0.0921 | 0.0771 | 0.0604 | 0.0457 | 0.0333 |
| 40 | 0.2370 | 0.1576 | 0.1259 | 0.1082 | 0.0968 | 0.0887 | 0.0827 | 0.0780 | 0.0745 | 0.0713 | 0.0595 | 0.0462 | 0.0347 | 0.0250 |
| 60 | 0.1737 | 0.1131 | 0.0895 | 0.0765 | 0.0682 | 0.0623 | 0.0583 | 0.0552 | 0.0520 | 0.0497 | 0.0411 | 0.0316 | 0.0234 | 0.0167 |
| 120 | 0.0998 | 0.0632 | 0.0495 | 0.0419 | 0.0371 | 0.0337 | 0.0312 | 0.0292 | 0.0279 | 0.0266 | 0.0218 | 0.0165 | 0.0120 | 0.0083 |
| ∞ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Tabla A.12 Puntos porcentuales superiores de la distribución de rangos estudentizados: valores de $q(0.05; k, \nu)$

| Grados de libertad, ν | Número de tratamientos, k | | | | | | | | | |
|---------------------------|-----------------------------|------|------|------|------|------|------|------|------|--|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| 1 | 18.0 | 27.0 | 32.8 | 37.2 | 40.5 | 43.1 | 45.1 | 47.1 | 49.1 | |
| 2 | 6.09 | 5.33 | 4.80 | 4.37 | 4.03 | 3.77 | 3.56 | 3.39 | 3.25 | |
| 3 | 4.50 | 3.91 | 3.57 | 3.29 | 3.01 | 2.79 | 2.61 | 2.47 | 2.35 | |
| 4 | 3.93 | 3.40 | 3.12 | 2.88 | 2.63 | 2.44 | 2.29 | 2.18 | 2.08 | |
| 5 | 3.64 | 3.16 | 2.92 | 2.70 | 2.49 | 2.32 | 2.19 | 2.10 | 2.01 | |
| 6 | 3.46 | 2.99 | 2.76 | 2.56 | 2.37 | 2.21 | 2.09 | 2.01 | 1.93 | |
| 7 | 3.34 | 2.88 | 2.66 | 2.47 | 2.29 | 2.14 | 2.03 | 1.95 | 1.87 | |
| 8 | 3.26 | 2.81 | 2.59 | 2.41 | 2.24 | 2.09 | 1.98 | 1.91 | 1.83 | |
| 9 | 3.20 | 2.75 | 2.54 | 2.36 | 2.19 | 2.04 | 1.93 | 1.86 | 1.78 | |
| 10 | 3.15 | 2.70 | 2.49 | 2.31 | 2.14 | 2.00 | 1.89 | 1.82 | 1.74 | |
| 11 | 3.11 | 2.66 | 2.45 | 2.27 | 2.10 | 1.95 | 1.84 | 1.77 | 1.69 | |
| 12 | 3.08 | 2.63 | 2.42 | 2.24 | 2.07 | 1.92 | 1.81 | 1.74 | 1.66 | |
| 13 | 3.06 | 2.61 | 2.40 | 2.22 | 2.05 | 1.90 | 1.79 | 1.72 | 1.64 | |
| 14 | 3.03 | 2.59 | 2.38 | 2.20 | 2.03 | 1.88 | 1.77 | 1.70 | 1.62 | |
| 15 | 3.01 | 2.57 | 2.36 | 2.18 | 2.01 | 1.86 | 1.75 | 1.68 | 1.60 | |
| 16 | 3.00 | 2.56 | 2.35 | 2.17 | 1.99 | 1.84 | 1.73 | 1.66 | 1.58 | |
| 17 | 2.98 | 2.55 | 2.34 | 2.16 | 1.98 | 1.83 | 1.72 | 1.65 | 1.57 | |
| 18 | 2.97 | 2.54 | 2.33 | 2.15 | 1.97 | 1.82 | 1.71 | 1.64 | 1.56 | |
| 19 | 2.96 | 2.53 | 2.32 | 2.14 | 1.96 | 1.81 | 1.70 | 1.63 | 1.55 | |
| 20 | 2.95 | 2.52 | 2.31 | 2.13 | 1.95 | 1.80 | 1.69 | 1.62 | 1.54 | |
| 24 | 2.92 | 2.50 | 2.29 | 2.11 | 1.93 | 1.78 | 1.67 | 1.60 | 1.52 | |
| 30 | 2.89 | 2.48 | 2.27 | 2.09 | 1.91 | 1.76 | 1.65 | 1.58 | 1.50 | |
| 40 | 2.86 | 2.45 | 2.24 | 2.06 | 1.88 | 1.73 | 1.62 | 1.55 | 1.47 | |
| 60 | 2.83 | 2.43 | 2.22 | 2.04 | 1.86 | 1.71 | 1.60 | 1.53 | 1.45 | |
| 120 | 2.80 | 2.40 | 2.19 | 2.01 | 1.83 | 1.68 | 1.57 | 1.50 | 1.42 | |
| ∞ | 2.77 | 2.32 | 2.11 | 1.93 | 1.75 | 1.60 | 1.49 | 1.42 | 1.34 | |

Tabla A.13 Rangos estudentizados significativos mínimos $r_p(0.05; p, v)$

| $\alpha = 0.05$ | | | | | | | | | |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| p | | | | | | | | | |
| v | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 |
| 2 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 |
| 3 | 4.501 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 |
| 4 | 3.927 | 4.013 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 |
| 5 | 3.635 | 3.749 | 3.797 | 3.814 | 3.814 | 3.814 | 3.814 | 3.814 | 3.814 |
| 6 | 3.461 | 3.587 | 3.649 | 3.68 | 3.694 | 3.697 | 3.697 | 3.697 | 3.697 |
| 7 | 3.344 | 3.477 | 3.548 | 3.588 | 3.611 | 3.622 | 3.626 | 3.626 | 3.626 |
| 8 | 3.261 | 3.399 | 3.475 | 3.521 | 3.549 | 3.566 | 3.575 | 3.579 | 3.579 |
| 9 | 3.199 | 3.339 | 3.420 | 3.470 | 3.502 | 3.523 | 3.536 | 3.544 | 3.547 |
| 10 | 3.151 | 3.293 | 3.376 | 3.430 | 3.465 | 3.489 | 3.505 | 3.516 | 3.522 |
| 11 | 3.113 | 3.256 | 3.342 | 3.397 | 3.435 | 3.462 | 3.48 | 3.493 | 3.501 |
| 12 | 3.082 | 3.225 | 3.313 | 3.370 | 3.410 | 3.439 | 3.459 | 3.474 | 3.484 |
| 13 | 3.055 | 3.200 | 3.289 | 3.348 | 3.389 | 3.419 | 3.442 | 3.458 | 3.470 |
| 14 | 3.033 | 3.178 | 3.268 | 3.329 | 3.372 | 3.403 | 3.426 | 3.444 | 3.457 |
| 15 | 3.014 | 3.160 | 3.25 | 3.312 | 3.356 | 3.389 | 3.413 | 3.432 | 3.446 |
| 16 | 2.998 | 3.144 | 3.235 | 3.298 | 3.343 | 3.376 | 3.402 | 3.422 | 3.437 |
| 17 | 2.984 | 3.130 | 3.222 | 3.285 | 3.331 | 3.366 | 3.392 | 3.412 | 3.429 |
| 18 | 2.971 | 3.118 | 3.210 | 3.274 | 3.321 | 3.356 | 3.383 | 3.405 | 3.421 |
| 19 | 2.960 | 3.107 | 3.199 | 3.264 | 3.311 | 3.347 | 3.375 | 3.397 | 3.415 |
| 20 | 2.950 | 3.097 | 3.190 | 3.255 | 3.303 | 3.339 | 3.368 | 3.391 | 3.409 |
| 24 | 2.919 | 3.066 | 3.160 | 3.226 | 3.276 | 3.315 | 3.345 | 3.370 | 3.390 |
| 30 | 2.888 | 3.035 | 3.131 | 3.199 | 3.250 | 3.290 | 3.322 | 3.349 | 3.371 |
| 40 | 2.858 | 3.006 | 3.102 | 3.171 | 3.224 | 3.266 | 3.300 | 3.328 | 3.352 |
| 60 | 2.829 | 2.976 | 3.073 | 3.143 | 3.198 | 3.241 | 3.277 | 3.307 | 3.333 |
| 120 | 2.800 | 2.947 | 3.045 | 3.116 | 3.172 | 3.217 | 3.254 | 3.287 | 3.314 |
| ∞ | 2.772 | 2.918 | 3.017 | 3.089 | 3.146 | 3.193 | 3.232 | 3.265 | 3.294 |

Condensada de H. L. Harter, "Critical Values for Duncan's New Multiple Range Test", *Biometrics*, 16, núm. 4, 1960, con autorización del autor y del editor.

Tabla A.13 (continuación) Rangos estudentizados significativos mínimos $r_p(0.01; p, v)$

| $\alpha = 0.01$ | | | | | | | | | |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| v | p | | | | | | | | |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 |
| 2 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 |
| 3 | 8.261 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 |
| 4 | 6.512 | 6.677 | 6.740 | 6.756 | 6.756 | 6.756 | 6.756 | 6.756 | 6.756 |
| 5 | 5.702 | 5.893 | 5.989 | 6.040 | 6.065 | 6.074 | 6.074 | 6.074 | 6.074 |
| 6 | 5.243 | 5.439 | 5.549 | 5.614 | 5.655 | 5.680 | 5.694 | 5.701 | 5.703 |
| 7 | 4.949 | 5.145 | 5.260 | 5.334 | 5.383 | 5.416 | 5.439 | 5.454 | 5.464 |
| 8 | 4.746 | 4.939 | 5.057 | 5.135 | 5.189 | 5.227 | 5.256 | 5.276 | 5.291 |
| 9 | 4.596 | 4.787 | 4.906 | 4.986 | 5.043 | 5.086 | 5.118 | 5.142 | 5.160 |
| 10 | 4.482 | 4.671 | 4.790 | 4.871 | 4.931 | 4.975 | 5.010 | 5.037 | 5.058 |
| 11 | 4.392 | 4.579 | 4.697 | 4.780 | 4.841 | 4.887 | 4.924 | 4.952 | 4.975 |
| 12 | 4.320 | 4.504 | 4.622 | 4.706 | 4.767 | 4.815 | 4.852 | 4.883 | 4.907 |
| 13 | 4.260 | 4.442 | 4.560 | 4.644 | 4.706 | 4.755 | 4.793 | 4.824 | 4.850 |
| 14 | 4.210 | 4.391 | 4.508 | 4.591 | 4.654 | 4.704 | 4.743 | 4.775 | 4.802 |
| 15 | 4.168 | 4.347 | 4.463 | 4.547 | 4.610 | 4.660 | 4.700 | 4.733 | 4.760 |
| 16 | 4.131 | 4.309 | 4.425 | 4.509 | 4.572 | 4.622 | 4.663 | 4.696 | 4.724 |
| 17 | 4.099 | 4.275 | 4.391 | 4.475 | 4.539 | 4.589 | 4.630 | 4.664 | 4.693 |
| 18 | 4.071 | 4.246 | 4.362 | 4.445 | 4.509 | 4.560 | 4.601 | 4.635 | 4.664 |
| 19 | 4.046 | 4.220 | 4.335 | 4.419 | 4.483 | 4.534 | 4.575 | 4.610 | 4.639 |
| 20 | 4.024 | 4.197 | 4.312 | 4.395 | 4.459 | 4.510 | 4.552 | 4.587 | 4.617 |
| 24 | 3.956 | 4.126 | 4.239 | 4.322 | 4.386 | 4.437 | 4.480 | 4.516 | 4.546 |
| 30 | 3.889 | 4.056 | 4.168 | 4.250 | 4.314 | 4.366 | 4.409 | 4.445 | 4.477 |
| 40 | 3.825 | 3.988 | 4.098 | 4.180 | 4.244 | 4.296 | 4.339 | 4.376 | 4.408 |
| 60 | 3.762 | 3.922 | 4.031 | 4.111 | 4.174 | 4.226 | 4.270 | 4.307 | 4.340 |
| 120 | 3.702 | 3.858 | 3.965 | 4.044 | 4.107 | 4.158 | 4.202 | 4.239 | 4.272 |
| ∞ | 3.643 | 3.796 | 3.900 | 3.978 | 4.040 | 4.091 | 4.135 | 4.172 | 4.205 |

Tabla A.14 Valores de $d_{\alpha/2}(k, \nu)$ para comparaciones bilaterales entre k tratamientos y un control

| $\alpha = 0.05$ | | | | | | | | | |
|----------------------------------------------------------------------|------|------|------|------|------|------|------|------|------|
| $k = \text{número de medias de tratamiento (no incluye el control)}$ | | | | | | | | | |
| ν | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 2.57 | 3.03 | 3.29 | 3.48 | 3.62 | 3.73 | 3.82 | 3.90 | 3.97 |
| 6 | 2.45 | 2.86 | 3.10 | 3.26 | 3.39 | 3.49 | 3.57 | 3.64 | 3.71 |
| 7 | 2.36 | 2.75 | 2.97 | 3.12 | 3.24 | 3.33 | 3.41 | 3.47 | 3.53 |
| 8 | 2.31 | 2.67 | 2.88 | 3.02 | 3.13 | 3.22 | 3.29 | 3.35 | 3.41 |
| 9 | 2.26 | 2.61 | 2.81 | 2.95 | 3.05 | 3.14 | 3.20 | 3.26 | 3.32 |
| 10 | 2.23 | 2.57 | 2.76 | 2.89 | 2.99 | 3.07 | 3.14 | 3.19 | 3.24 |
| 11 | 2.20 | 2.53 | 2.72 | 2.84 | 2.94 | 3.02 | 3.08 | 3.14 | 3.19 |
| 12 | 2.18 | 2.50 | 2.68 | 2.81 | 2.90 | 2.98 | 3.04 | 3.09 | 3.14 |
| 13 | 2.16 | 2.48 | 2.65 | 2.78 | 2.87 | 2.94 | 3.00 | 3.06 | 3.10 |
| 14 | 2.14 | 2.46 | 2.63 | 2.75 | 2.84 | 2.91 | 2.97 | 3.02 | 3.07 |
| 15 | 2.13 | 2.44 | 2.61 | 2.73 | 2.82 | 2.89 | 2.95 | 3.00 | 3.04 |
| 16 | 2.12 | 2.42 | 2.59 | 2.71 | 2.80 | 2.87 | 2.92 | 2.97 | 3.02 |
| 17 | 2.11 | 2.41 | 2.58 | 2.69 | 2.78 | 2.85 | 2.90 | 2.95 | 3.00 |
| 18 | 2.10 | 2.40 | 2.56 | 2.68 | 2.76 | 2.83 | 2.89 | 2.94 | 2.98 |
| 19 | 2.09 | 2.39 | 2.55 | 2.66 | 2.75 | 2.81 | 2.87 | 2.92 | 2.96 |
| 20 | 2.09 | 2.38 | 2.54 | 2.65 | 2.73 | 2.80 | 2.86 | 2.90 | 2.95 |
| 24 | 2.06 | 2.35 | 2.51 | 2.61 | 2.70 | 2.76 | 2.81 | 2.86 | 2.90 |
| 30 | 2.04 | 2.32 | 2.47 | 2.58 | 2.66 | 2.72 | 2.77 | 2.82 | 2.86 |
| 40 | 2.02 | 2.29 | 2.44 | 2.54 | 2.62 | 2.68 | 2.73 | 2.77 | 2.81 |
| 60 | 2.00 | 2.27 | 2.41 | 2.51 | 2.58 | 2.64 | 2.69 | 2.73 | 2.77 |
| 120 | 1.98 | 2.24 | 2.38 | 2.47 | 2.55 | 2.60 | 2.65 | 2.69 | 2.73 |
| ∞ | 1.96 | 2.21 | 2.35 | 2.44 | 2.51 | 2.57 | 2.61 | 2.65 | 2.69 |

Reproducida de Charles W. Dunnett, "New Tables for Multiple Comparison with a Control", *Biometrics*, 20, núm. 3, 1964, con autorización del autor y del editor.

Tabla A.14 (continuación) Valores de $d_{\alpha/2}(k, \nu)$ para comparaciones bilaterales entre k tratamientos y un control

| $\alpha = 0.01$ | | | | | | | | | |
|----------------------------------------------------------------------|------|------|------|------|------|------|------|------|------|
| $k = \text{número de medias de tratamiento (no incluye el control)}$ | | | | | | | | | |
| ν | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 4.03 | 4.63 | 4.98 | 5.22 | 5.41 | 5.56 | 5.69 | 5.80 | 5.89 |
| 6 | 3.71 | 4.21 | 4.51 | 4.71 | 4.87 | 5.00 | 5.10 | 5.20 | 5.28 |
| 7 | 3.50 | 3.95 | 4.21 | 4.39 | 4.53 | 4.64 | 4.74 | 4.82 | 4.89 |
| 8 | 3.36 | 3.77 | 4.00 | 4.17 | 4.29 | 4.40 | 4.48 | 4.56 | 4.62 |
| 9 | 3.25 | 3.63 | 3.85 | 4.01 | 4.12 | 4.22 | 4.30 | 4.37 | 4.43 |
| 10 | 3.17 | 3.53 | 3.74 | 3.88 | 3.99 | 4.08 | 4.16 | 4.22 | 4.28 |
| 11 | 3.11 | 3.45 | 3.65 | 3.79 | 3.89 | 3.98 | 4.05 | 4.11 | 4.16 |
| 12 | 3.05 | 3.39 | 3.58 | 3.71 | 3.81 | 3.89 | 3.96 | 4.02 | 4.07 |
| 13 | 3.01 | 3.33 | 3.52 | 3.65 | 3.74 | 3.82 | 3.89 | 3.94 | 3.99 |
| 14 | 2.98 | 3.29 | 3.47 | 3.59 | 3.69 | 3.76 | 3.83 | 3.88 | 3.93 |
| 15 | 2.95 | 3.25 | 3.43 | 3.55 | 3.64 | 3.71 | 3.78 | 3.83 | 3.88 |
| 16 | 2.92 | 3.22 | 3.39 | 3.51 | 3.60 | 3.67 | 3.73 | 3.78 | 3.83 |
| 17 | 2.90 | 3.19 | 3.36 | 3.47 | 3.56 | 3.63 | 3.69 | 3.74 | 3.79 |
| 18 | 2.88 | 3.17 | 3.33 | 3.44 | 3.53 | 3.60 | 3.66 | 3.71 | 3.75 |
| 19 | 2.86 | 3.15 | 3.31 | 3.42 | 3.50 | 3.57 | 3.63 | 3.68 | 3.72 |
| 20 | 2.85 | 3.13 | 3.29 | 3.40 | 3.48 | 3.55 | 3.60 | 3.65 | 3.69 |
| 24 | 2.80 | 3.07 | 3.22 | 3.32 | 3.40 | 3.47 | 3.52 | 3.57 | 3.61 |
| 30 | 2.75 | 3.01 | 3.15 | 3.25 | 3.33 | 3.39 | 3.44 | 3.49 | 3.52 |
| 40 | 2.70 | 2.95 | 3.09 | 3.19 | 3.26 | 3.32 | 3.37 | 3.41 | 3.44 |
| 60 | 2.66 | 2.90 | 3.03 | 3.12 | 3.19 | 3.25 | 3.29 | 3.33 | 3.37 |
| 120 | 2.62 | 2.85 | 2.97 | 3.06 | 3.12 | 3.18 | 3.22 | 3.26 | 3.29 |
| ∞ | 2.58 | 2.79 | 2.92 | 3.00 | 3.06 | 3.11 | 3.15 | 3.19 | 3.22 |

Tabla A.15 Valores de $d_{\alpha}(k, \nu)$ para comparaciones unilaterales entre k tratamientos y un control

| $\alpha = 0.05$ | | | | | | | | | |
|----------------------------------------------------------------------|------|------|------|------|------|------|------|------|------|
| $k = \text{número de medias de tratamiento (no incluye el control)}$ | | | | | | | | | |
| ν | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 2.02 | 2.44 | 2.68 | 2.85 | 2.98 | 3.08 | 3.16 | 3.24 | 3.30 |
| 6 | 1.94 | 2.34 | 2.56 | 2.71 | 2.83 | 2.92 | 3.00 | 3.07 | 3.12 |
| 7 | 1.89 | 2.27 | 2.48 | 2.62 | 2.73 | 2.82 | 2.89 | 2.95 | 3.01 |
| 8 | 1.86 | 2.22 | 2.42 | 2.55 | 2.66 | 2.74 | 2.81 | 2.87 | 2.92 |
| 9 | 1.83 | 2.18 | 2.37 | 2.50 | 2.60 | 2.68 | 2.75 | 2.81 | 2.86 |
| 10 | 1.81 | 2.15 | 2.34 | 2.47 | 2.56 | 2.64 | 2.70 | 2.76 | 2.81 |
| 11 | 1.80 | 2.13 | 2.31 | 2.44 | 2.53 | 2.60 | 2.67 | 2.72 | 2.77 |
| 12 | 1.78 | 2.11 | 2.29 | 2.41 | 2.50 | 2.58 | 2.64 | 2.69 | 2.74 |
| 13 | 1.77 | 2.09 | 2.27 | 2.39 | 2.48 | 2.55 | 2.61 | 2.66 | 2.71 |
| 14 | 1.76 | 2.08 | 2.25 | 2.37 | 2.46 | 2.53 | 2.59 | 2.64 | 2.69 |
| 15 | 1.75 | 2.07 | 2.24 | 2.36 | 2.44 | 2.51 | 2.57 | 2.62 | 2.67 |
| 16 | 1.75 | 2.06 | 2.23 | 2.34 | 2.43 | 2.50 | 2.56 | 2.61 | 2.65 |
| 17 | 1.74 | 2.05 | 2.22 | 2.33 | 2.42 | 2.49 | 2.54 | 2.59 | 2.64 |
| 18 | 1.73 | 2.04 | 2.21 | 2.32 | 2.41 | 2.48 | 2.53 | 2.58 | 2.62 |
| 19 | 1.73 | 2.03 | 2.20 | 2.31 | 2.40 | 2.47 | 2.52 | 2.57 | 2.61 |
| 20 | 1.72 | 2.03 | 2.19 | 2.30 | 2.39 | 2.46 | 2.51 | 2.56 | 2.60 |
| 24 | 1.71 | 2.01 | 2.17 | 2.28 | 2.36 | 2.43 | 2.48 | 2.53 | 2.57 |
| 30 | 1.70 | 1.99 | 2.15 | 2.25 | 2.33 | 2.40 | 2.45 | 2.50 | 2.54 |
| 40 | 1.68 | 1.97 | 2.13 | 2.23 | 2.31 | 2.37 | 2.42 | 2.47 | 2.51 |
| 60 | 1.67 | 1.95 | 2.10 | 2.21 | 2.28 | 2.35 | 2.39 | 2.44 | 2.48 |
| 120 | 1.66 | 1.93 | 2.08 | 2.18 | 2.26 | 2.32 | 2.37 | 2.41 | 2.45 |
| ∞ | 1.64 | 1.92 | 2.06 | 2.16 | 2.23 | 2.29 | 2.34 | 2.38 | 2.42 |

Reproducida de Charles W. Dunnett, "A Multiple Comparison Procedure for Comparing Several Treatments with a Control", *J. Am. Stat. Assoc.*, 50, 1955, 1096-1121, con autorización del autor y del editor.

Tabla A.15 (continuación) Valores de $d_{\alpha}(k, \nu)$ para comparaciones unilaterales entre k tratamientos y un control

| $\alpha = 0.01$ | | | | | | | | | |
|----------------------------------------------------------------------|------|------|------|------|------|------|------|------|------|
| $k = \text{número de medias de tratamiento (no incluye el control)}$ | | | | | | | | | |
| ν | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 3.37 | 3.90 | 4.21 | 4.43 | 4.60 | 4.73 | 4.85 | 4.94 | 5.03 |
| 6 | 3.14 | 3.61 | 3.88 | 4.07 | 4.21 | 4.33 | 4.43 | 4.51 | 4.59 |
| 7 | 3.00 | 3.42 | 3.66 | 3.83 | 3.96 | 4.07 | 4.15 | 4.23 | 4.30 |
| 8 | 2.90 | 3.29 | 3.51 | 3.67 | 3.79 | 3.88 | 3.96 | 4.03 | 4.09 |
| 9 | 2.82 | 3.19 | 3.40 | 3.55 | 3.66 | 3.75 | 3.82 | 3.89 | 3.94 |
| 10 | 2.76 | 3.11 | 3.31 | 3.45 | 3.56 | 3.64 | 3.71 | 3.78 | 3.83 |
| 11 | 2.72 | 3.06 | 3.25 | 3.38 | 3.48 | 3.56 | 3.63 | 3.69 | 3.74 |
| 12 | 2.68 | 3.01 | 3.19 | 3.32 | 3.42 | 3.50 | 3.56 | 3.62 | 3.67 |
| 13 | 2.65 | 2.97 | 3.15 | 3.27 | 3.37 | 3.44 | 3.51 | 3.56 | 3.61 |
| 14 | 2.62 | 2.94 | 3.11 | 3.23 | 3.32 | 3.40 | 3.46 | 3.51 | 3.56 |
| 15 | 2.60 | 2.91 | 3.08 | 3.20 | 3.29 | 3.36 | 3.42 | 3.47 | 3.52 |
| 16 | 2.58 | 2.88 | 3.05 | 3.17 | 3.26 | 3.33 | 3.39 | 3.44 | 3.48 |
| 17 | 2.57 | 2.86 | 3.03 | 3.14 | 3.23 | 3.30 | 3.36 | 3.41 | 3.45 |
| 18 | 2.55 | 2.84 | 3.01 | 3.12 | 3.21 | 3.27 | 3.33 | 3.38 | 3.42 |
| 19 | 2.54 | 2.83 | 2.99 | 3.10 | 3.18 | 3.25 | 3.31 | 3.36 | 3.40 |
| 20 | 2.53 | 2.81 | 2.97 | 3.08 | 3.17 | 3.23 | 3.29 | 3.34 | 3.38 |
| 24 | 2.49 | 2.77 | 2.92 | 3.03 | 3.11 | 3.17 | 3.22 | 3.27 | 3.31 |
| 30 | 2.46 | 2.72 | 2.87 | 2.97 | 3.05 | 3.11 | 3.16 | 3.21 | 3.24 |
| 40 | 2.42 | 2.68 | 2.82 | 2.92 | 2.99 | 3.05 | 3.10 | 3.14 | 3.18 |
| 60 | 2.39 | 2.64 | 2.78 | 2.87 | 2.94 | 3.00 | 3.04 | 3.08 | 3.12 |
| 120 | 2.36 | 2.60 | 2.73 | 2.82 | 2.89 | 2.94 | 2.99 | 3.03 | 3.06 |
| ∞ | 2.33 | 2.56 | 2.68 | 2.77 | 2.84 | 2.89 | 2.93 | 2.97 | 3.00 |

Tabla A.16 Valores críticos para la prueba de rangos con signo

| n | Unilateral $\alpha = 0.01$
Bilateral $\alpha = 0.02$ | Unilateral $\alpha = 0.025$
Bilateral $\alpha = 0.05$ | Unilateral $\alpha = 0.05$
Bilateral $\alpha = 0.1$ |
|-----|---------------------------------------------------------|----------------------------------------------------------|--------------------------------------------------------|
| 5 | | | 1 |
| 6 | | 1 | 2 |
| 7 | 0 | 2 | 4 |
| 8 | 2 | 4 | 6 |
| 9 | 3 | 6 | 8 |
| 10 | 5 | 8 | 11 |
| 11 | 7 | 11 | 14 |
| 12 | 10 | 14 | 17 |
| 13 | 13 | 17 | 21 |
| 14 | 16 | 21 | 26 |
| 15 | 20 | 25 | 30 |
| 16 | 24 | 30 | 36 |
| 17 | 28 | 35 | 41 |
| 18 | 33 | 40 | 47 |
| 19 | 38 | 46 | 54 |
| 20 | 43 | 52 | 60 |
| 21 | 49 | 59 | 68 |
| 22 | 56 | 66 | 75 |
| 23 | 62 | 73 | 83 |
| 24 | 69 | 81 | 92 |
| 25 | 77 | 90 | 101 |
| 26 | 85 | 98 | 110 |
| 27 | 93 | 107 | 120 |
| 28 | 102 | 117 | 130 |
| 29 | 111 | 127 | 141 |
| 30 | 120 | 137 | 152 |

Reproducida de F. Wilcoxon y R. A. Wilcox, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Company, Pearl River, N. Y., 1964, con autorización de la American Cyanamid Company.

Tabla A.17 Valores críticos para la prueba de suma de rangos de Wilcoxon

| Prueba de una cola con $\alpha = 0.001$ o prueba de dos colas con $\alpha = 0.002$ | | | | | | | | | | | | | | | |
|------------------------------------------------------------------------------------|-------|---|---|---|----|----|----|----|----|----|----|----|----|----|----|
| n_1 | n_2 | | | | | | | | | | | | | | |
| | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1 | | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | 0 | 0 | 0 | 0 |
| 4 | | | | | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 3 | 3 | 3 |
| 5 | | 0 | 0 | 1 | 1 | 2 | 2 | 3 | 3 | 4 | 5 | 5 | 6 | 7 | 7 |
| 6 | 0 | 1 | 2 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 7 | | 2 | 3 | 3 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 13 | 14 | 15 | 16 |
| 8 | | | 5 | 5 | 6 | 8 | 9 | 11 | 12 | 14 | 15 | 17 | 18 | 20 | 21 |
| 9 | | | | 7 | 8 | 10 | 12 | 14 | 15 | 17 | 19 | 21 | 23 | 25 | 26 |
| 10 | | | | | 10 | 12 | 14 | 17 | 19 | 21 | 23 | 25 | 27 | 29 | 32 |
| 11 | | | | | | 15 | 17 | 20 | 22 | 24 | 27 | 29 | 32 | 34 | 37 |
| 12 | | | | | | | 20 | 23 | 25 | 28 | 31 | 34 | 37 | 40 | 42 |
| 13 | | | | | | | | 26 | 29 | 32 | 35 | 38 | 42 | 45 | 48 |
| 14 | | | | | | | | | 32 | 36 | 39 | 43 | 46 | 50 | 54 |
| 15 | | | | | | | | | | 40 | 43 | 47 | 51 | 55 | 59 |
| 16 | | | | | | | | | | | 48 | 52 | 56 | 60 | 65 |
| 17 | | | | | | | | | | | | 57 | 61 | 66 | 70 |
| 18 | | | | | | | | | | | | | 66 | 71 | 76 |
| 19 | | | | | | | | | | | | | | 77 | 82 |
| 20 | | | | | | | | | | | | | | | 88 |

Prueba de una cola con $\alpha = 0.01$ o prueba de dos colas con $\alpha = 0.02$

| n_1 | n_2 | | | | | | | | | | | | | | | |
|-------|-------|---|---|----|----|----|----|----|----|----|----|----|----|----|-----|-----|
| | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1 | | | | | | | | | | | | | | | | |
| 2 | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 3 | | | | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 3 | 3 | 4 | 4 | 4 | 5 |
| 4 | 0 | 1 | 1 | 2 | 3 | 3 | 4 | 5 | 5 | 6 | 7 | 7 | 8 | 9 | 9 | 10 |
| 5 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 6 | | 3 | 4 | 6 | 7 | 8 | 9 | 11 | 12 | 13 | 15 | 16 | 18 | 19 | 20 | 22 |
| 7 | | | 6 | 8 | 9 | 11 | 12 | 14 | 16 | 17 | 19 | 21 | 23 | 24 | 26 | 28 |
| 8 | | | | 10 | 11 | 13 | 15 | 17 | 20 | 22 | 24 | 26 | 28 | 30 | 32 | 34 |
| 9 | | | | | 14 | 16 | 18 | 21 | 23 | 26 | 28 | 31 | 33 | 36 | 38 | 40 |
| 10 | | | | | | 19 | 22 | 24 | 27 | 30 | 33 | 36 | 38 | 41 | 44 | 47 |
| 11 | | | | | | | 25 | 28 | 31 | 34 | 37 | 41 | 44 | 47 | 50 | 53 |
| 12 | | | | | | | | 31 | 35 | 38 | 42 | 46 | 49 | 53 | 56 | 60 |
| 13 | | | | | | | | | 39 | 43 | 47 | 51 | 55 | 59 | 63 | 67 |
| 14 | | | | | | | | | | 47 | 51 | 56 | 60 | 65 | 69 | 73 |
| 15 | | | | | | | | | | | 56 | 61 | 66 | 70 | 75 | 80 |
| 16 | | | | | | | | | | | | 66 | 71 | 76 | 82 | 87 |
| 17 | | | | | | | | | | | | | 77 | 82 | 88 | 93 |
| 18 | | | | | | | | | | | | | | 88 | 94 | 100 |
| 19 | | | | | | | | | | | | | | | 101 | 107 |
| 20 | | | | | | | | | | | | | | | | 114 |

Basada en parte en las tablas 1, 3, 5 y 7 de D. Auble, "Extended Tables for the Mann-Whitney Statistic", *Bulletin of the Institute of Educational Research at Indiana University*, 1, núm. 2, 1953, con autorización del director.

Tabla A.18 $P(V \leq v^*$ cuando H_0 es verdadera) en la prueba de rachas

| (n_1, n_2) | v^* | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| (2, 3) | 0.200 | 0.500 | 0.900 | 1.000 | | | | | |
| (2, 4) | 0.133 | 0.400 | 0.800 | 1.000 | | | | | |
| (2, 5) | 0.095 | 0.333 | 0.714 | 1.000 | | | | | |
| (2, 6) | 0.071 | 0.286 | 0.643 | 1.000 | | | | | |
| (2, 7) | 0.056 | 0.250 | 0.583 | 1.000 | | | | | |
| (2, 8) | 0.044 | 0.222 | 0.533 | 1.000 | | | | | |
| (2, 9) | 0.036 | 0.200 | 0.491 | 1.000 | | | | | |
| (2, 10) | 0.030 | 0.182 | 0.455 | 1.000 | | | | | |
| (3, 3) | 0.100 | 0.300 | 0.700 | 0.900 | 1.000 | | | | |
| (3, 4) | 0.057 | 0.200 | 0.543 | 0.800 | 0.971 | 1.000 | | | |
| (3, 5) | 0.036 | 0.143 | 0.429 | 0.714 | 0.929 | 1.000 | | | |
| (3, 6) | 0.024 | 0.107 | 0.345 | 0.643 | 0.881 | 1.000 | | | |
| (3, 7) | 0.017 | 0.083 | 0.283 | 0.583 | 0.833 | 1.000 | | | |
| (3, 8) | 0.012 | 0.067 | 0.236 | 0.533 | 0.788 | 1.000 | | | |
| (3, 9) | 0.009 | 0.055 | 0.200 | 0.491 | 0.745 | 1.000 | | | |
| (3, 10) | 0.007 | 0.045 | 0.171 | 0.455 | 0.706 | 1.000 | | | |
| (4, 4) | 0.029 | 0.114 | 0.371 | 0.629 | 0.886 | 0.971 | 1.000 | | |
| (4, 5) | 0.016 | 0.071 | 0.262 | 0.500 | 0.786 | 0.929 | 0.992 | 1.000 | |
| (4, 6) | 0.010 | 0.048 | 0.190 | 0.405 | 0.690 | 0.881 | 0.976 | 1.000 | |
| (4, 7) | 0.006 | 0.033 | 0.142 | 0.333 | 0.606 | 0.833 | 0.954 | 1.000 | |
| (4, 8) | 0.004 | 0.024 | 0.109 | 0.279 | 0.533 | 0.788 | 0.929 | 1.000 | |
| (4, 9) | 0.003 | 0.018 | 0.085 | 0.236 | 0.471 | 0.745 | 0.902 | 1.000 | |
| (4, 10) | 0.002 | 0.014 | 0.068 | 0.203 | 0.419 | 0.706 | 0.874 | 1.000 | |
| (5, 5) | 0.008 | 0.040 | 0.167 | 0.357 | 0.643 | 0.833 | 0.960 | 0.992 | 1.000 |
| (5, 6) | 0.004 | 0.024 | 0.110 | 0.262 | 0.522 | 0.738 | 0.911 | 0.976 | 0.998 |
| (5, 7) | 0.003 | 0.015 | 0.076 | 0.197 | 0.424 | 0.652 | 0.854 | 0.955 | 0.992 |
| (5, 8) | 0.002 | 0.010 | 0.054 | 0.152 | 0.347 | 0.576 | 0.793 | 0.929 | 0.984 |
| (5, 9) | 0.001 | 0.007 | 0.039 | 0.119 | 0.287 | 0.510 | 0.734 | 0.902 | 0.972 |
| (5, 10) | 0.001 | 0.005 | 0.029 | 0.095 | 0.239 | 0.455 | 0.678 | 0.874 | 0.958 |
| (6, 6) | 0.002 | 0.013 | 0.067 | 0.175 | 0.392 | 0.608 | 0.825 | 0.933 | 0.987 |
| (6, 7) | 0.001 | 0.008 | 0.043 | 0.121 | 0.296 | 0.500 | 0.733 | 0.879 | 0.966 |
| (6, 8) | 0.001 | 0.005 | 0.028 | 0.086 | 0.226 | 0.413 | 0.646 | 0.821 | 0.937 |
| (6, 9) | 0.000 | 0.003 | 0.019 | 0.063 | 0.175 | 0.343 | 0.566 | 0.762 | 0.902 |
| (6, 10) | 0.000 | 0.002 | 0.013 | 0.047 | 0.137 | 0.288 | 0.497 | 0.706 | 0.864 |
| (7, 7) | 0.001 | 0.004 | 0.025 | 0.078 | 0.209 | 0.383 | 0.617 | 0.791 | 0.922 |
| (7, 8) | 0.000 | 0.002 | 0.015 | 0.051 | 0.149 | 0.296 | 0.514 | 0.704 | 0.867 |
| (7, 9) | 0.000 | 0.001 | 0.010 | 0.035 | 0.108 | 0.231 | 0.427 | 0.622 | 0.806 |
| (7, 10) | 0.000 | 0.001 | 0.006 | 0.024 | 0.080 | 0.182 | 0.355 | 0.549 | 0.743 |
| (8, 8) | 0.000 | 0.001 | 0.009 | 0.032 | 0.100 | 0.214 | 0.405 | 0.595 | 0.786 |
| (8, 9) | 0.000 | 0.001 | 0.005 | 0.020 | 0.069 | 0.157 | 0.319 | 0.500 | 0.702 |
| (8, 10) | 0.000 | 0.000 | 0.003 | 0.013 | 0.048 | 0.117 | 0.251 | 0.419 | 0.621 |
| (9, 9) | 0.000 | 0.000 | 0.003 | 0.012 | 0.044 | 0.109 | 0.238 | 0.399 | 0.601 |
| (9, 10) | 0.000 | 0.000 | 0.002 | 0.008 | 0.029 | 0.077 | 0.179 | 0.319 | 0.510 |
| (10, 10) | 0.000 | 0.000 | 0.001 | 0.004 | 0.019 | 0.051 | 0.128 | 0.242 | 0.414 |

Reproducida de C. Eisenhart y R. Swed, "Tables for Testing Randomness of Grouping in a Sequence of Alternatives", *Ann. Math. Stat.*, 14, 1943, con autorización del editor.

Tabla A.18 (continuación) $P(V \leq v^*$ cuando H_0 es verdadera) en la prueba de rachas

| (n_1, n_2) | v^* | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| (2, 3) | | | | | | | | | | |
| (2, 4) | | | | | | | | | | |
| (2, 5) | | | | | | | | | | |
| (2, 6) | | | | | | | | | | |
| (2, 7) | | | | | | | | | | |
| (2, 8) | | | | | | | | | | |
| (2, 9) | | | | | | | | | | |
| (2, 10) | | | | | | | | | | |
| (3, 3) | | | | | | | | | | |
| (3, 4) | | | | | | | | | | |
| (3, 5) | | | | | | | | | | |
| (3, 6) | | | | | | | | | | |
| (3, 7) | | | | | | | | | | |
| (3, 8) | | | | | | | | | | |
| (3, 9) | | | | | | | | | | |
| (3, 10) | | | | | | | | | | |
| (4, 4) | | | | | | | | | | |
| (4, 5) | | | | | | | | | | |
| (4, 6) | | | | | | | | | | |
| (4, 7) | | | | | | | | | | |
| (4, 8) | | | | | | | | | | |
| (4, 9) | | | | | | | | | | |
| (4, 10) | | | | | | | | | | |
| (5, 5) | | | | | | | | | | |
| (5, 6) | 1.000 | | | | | | | | | |
| (5, 7) | 1.000 | | | | | | | | | |
| (5, 8) | 1.000 | | | | | | | | | |
| (5, 9) | 1.000 | | | | | | | | | |
| (5, 10) | 1.000 | | | | | | | | | |
| (6, 6) | 0.998 | 1.000 | | | | | | | | |
| (6, 7) | 0.992 | 0.999 | 1.000 | | | | | | | |
| (6, 8) | 0.984 | 0.998 | 1.000 | | | | | | | |
| (6, 9) | 0.972 | 0.994 | 1.000 | | | | | | | |
| (6, 10) | 0.958 | 0.990 | 1.000 | | | | | | | |
| (7, 7) | 0.975 | 0.996 | 0.999 | 1.000 | | | | | | |
| (7, 8) | 0.949 | 0.988 | 0.998 | 1.000 | 1.000 | | | | | |
| (7, 9) | 0.916 | 0.975 | 0.994 | 0.999 | 1.000 | | | | | |
| (7, 10) | 0.879 | 0.957 | 0.990 | 0.998 | 1.000 | | | | | |
| (8, 8) | 0.900 | 0.968 | 0.991 | 0.999 | 1.000 | 1.000 | | | | |
| (8, 9) | 0.843 | 0.939 | 0.980 | 0.996 | 0.999 | 1.000 | 1.000 | | | |
| (8, 10) | 0.782 | 0.903 | 0.964 | 0.990 | 0.998 | 1.000 | 1.000 | | | |
| (9, 9) | 0.762 | 0.891 | 0.956 | 0.988 | 0.997 | 1.000 | 1.000 | 1.000 | | |
| (9, 10) | 0.681 | 0.834 | 0.923 | 0.974 | 0.992 | 0.999 | 1.000 | 1.000 | 1.000 | |
| (10, 10) | 0.586 | 0.758 | 0.872 | 0.949 | 0.981 | 0.996 | 0.999 | 1.000 | 1.000 | 1.000 |

Tabla A.19 Tamaño muestral para límites de tolerancia no paramétricos bilaterales

| $1 - \alpha$ | $1 - \gamma$ | | | | | |
|--------------|--------------|------|------|------|------|-------|
| | 0.50 | 0.70 | 0.90 | 0.95 | 0.99 | 0.995 |
| 0.995 | 336 | 488 | 777 | 947 | 1325 | 1483 |
| 0.99 | 168 | 244 | 388 | 473 | 662 | 740 |
| 0.95 | 34 | 49 | 77 | 93 | 130 | 146 |
| 0.90 | 17 | 24 | 38 | 46 | 64 | 72 |
| 0.85 | 11 | 16 | 25 | 30 | 42 | 47 |
| 0.80 | 9 | 12 | 18 | 22 | 31 | 34 |
| 0.75 | 7 | 10 | 15 | 18 | 24 | 27 |
| 0.70 | 6 | 8 | 12 | 14 | 20 | 22 |
| 0.60 | 4 | 6 | 9 | 10 | 14 | 16 |
| 0.50 | 3 | 5 | 7 | 8 | 11 | 12 |

Tabla A-25d de Wilfrid J. Dixon y Frank J. Massey, Jr., *Introduction to Statistical Analysis*, 3a. ed., McGraw-Hill, 1969. Reproducida con autorización de The McGraw-Hill Companies, Inc.

Tabla A.20 Tamaño muestral para límites de tolerancia no paramétricos unilaterales

| $1 - \alpha$ | $1 - \gamma$ | | | | |
|--------------|--------------|------|------|------|-------|
| | 0.50 | 0.70 | 0.95 | 0.99 | 0.995 |
| 0.995 | 139 | 241 | 598 | 919 | 1379 |
| 0.99 | 69 | 120 | 299 | 459 | 688 |
| 0.95 | 14 | 24 | 59 | 90 | 135 |
| 0.90 | 7 | 12 | 29 | 44 | 66 |
| 0.85 | 5 | 8 | 19 | 29 | 43 |
| 0.80 | 4 | 6 | 14 | 21 | 31 |
| 0.75 | 3 | 5 | 11 | 7 | 25 |
| 0.70 | 2 | 4 | 9 | 13 | 20 |
| 0.60 | 2 | 3 | 6 | 10 | 14 |
| 0.50 | 1 | 2 | 5 | 7 | 10 |

Tabla A-25e de Wilfrid J. Dixon y Frank J. Massey, Jr., *Introduction to Statistical Analysis*, 3a. ed., McGraw-Hill, 1969. Reproducida con autorización de The McGraw-Hill Companies, Inc.

Tabla A.21 Valores críticos del coeficiente de correlación de rangos de Spearman

| n | $\alpha = 0.05$ | $\alpha = 0.025$ | $\alpha = 0.01$ | $\alpha = 0.005$ |
|-----|-----------------|------------------|-----------------|------------------|
| 5 | 0.900 | | | |
| 6 | 0.829 | 0.886 | 0.943 | |
| 7 | 0.714 | 0.786 | 0.893 | |
| 8 | 0.643 | 0.738 | 0.833 | 0.881 |
| 9 | 0.600 | 0.683 | 0.783 | 0.833 |
| 10 | 0.564 | 0.648 | 0.745 | 0.794 |
| 11 | 0.523 | 0.623 | 0.736 | 0.818 |
| 12 | 0.497 | 0.591 | 0.703 | 0.780 |
| 13 | 0.475 | 0.566 | 0.673 | 0.745 |
| 14 | 0.457 | 0.545 | 0.646 | 0.716 |
| 15 | 0.441 | 0.525 | 0.623 | 0.689 |
| 16 | 0.425 | 0.507 | 0.601 | 0.666 |
| 17 | 0.412 | 0.490 | 0.582 | 0.645 |
| 18 | 0.399 | 0.476 | 0.564 | 0.625 |
| 19 | 0.388 | 0.462 | 0.549 | 0.608 |
| 20 | 0.377 | 0.450 | 0.534 | 0.591 |
| 21 | 0.368 | 0.438 | 0.521 | 0.576 |
| 22 | 0.359 | 0.428 | 0.508 | 0.562 |
| 23 | 0.351 | 0.418 | 0.496 | 0.549 |
| 24 | 0.343 | 0.409 | 0.485 | 0.537 |
| 25 | 0.336 | 0.400 | 0.475 | 0.526 |
| 26 | 0.329 | 0.392 | 0.465 | 0.515 |
| 27 | 0.323 | 0.385 | 0.456 | 0.505 |
| 28 | 0.317 | 0.377 | 0.448 | 0.496 |
| 29 | 0.311 | 0.370 | 0.440 | 0.487 |
| 30 | 0.305 | 0.364 | 0.432 | 0.478 |

Reproducida de E. G. Olds, "Distribution of Sums of Squares of Rank Differences for Small Samples", *Ann. Math. Stat.*, 9, 1938, con autorización del editor.

Tabla A.22 Factores para la elaboración de gráficas de control

| Observaciones
en la muestra | Gráfica para
promedios | | Gráfica para desviaciones estándar | | | | | | Gráfica para rangos | | | | |
|--------------------------------|-----------------------------------------|-------|------------------------------------|--------|--------------------------------------|-------|-------|-------|-----------------------------------|--------|-----------------------------------------|-------|-------|
| | Factores para los
límites de control | | Factores para la
línea central | | Factores para los límites de control | | | | Factores para la
línea central | | Factores para los límites
de control | | |
| | n | A_2 | A_3 | c_4 | $1/c_4$ | B_3 | B_4 | B_5 | B_6 | d_2 | $1/d_2$ | d_3 | D_3 |
| 2 | 1.880 | 2.659 | 0.7979 | 1.2533 | 0 | 3.267 | 0 | 2.606 | 1.128 | 0.8865 | 0.853 | 0 | 3.267 |
| 3 | 1.023 | 1.954 | 0.8862 | 1.1284 | 0 | 2.568 | 0 | 2.276 | 1.693 | 0.5907 | 0.888 | 0 | 2.574 |
| 4 | 0.729 | 1.628 | 0.9213 | 1.0854 | 0 | 2.266 | 0 | 2.088 | 2.059 | 0.4857 | 0.880 | 0 | 2.282 |
| 5 | 0.577 | 1.427 | 0.9400 | 1.0638 | 0 | 2.089 | 0 | 1.964 | 2.326 | 0.4299 | 0.864 | 0 | 2.114 |
| 6 | 0.483 | 1.287 | 0.9515 | 1.0510 | 0.030 | 1.970 | 0.029 | 1.874 | 2.534 | 0.3946 | 0.848 | 0 | 2.004 |
| 7 | 0.419 | 1.182 | 0.9594 | 1.0423 | 0.118 | 1.882 | 0.113 | 1.806 | 2.704 | 0.3698 | 0.833 | 0.076 | 1.924 |
| 8 | 0.373 | 1.099 | 0.9650 | 1.0363 | 0.185 | 1.815 | 0.179 | 1.751 | 2.847 | 0.3512 | 0.820 | 0.136 | 1.864 |
| 9 | 0.337 | 1.032 | 0.9693 | 1.0317 | 0.239 | 1.761 | 0.232 | 1.707 | 2.970 | 0.3367 | 0.808 | 0.184 | 1.816 |
| 10 | 0.308 | 0.975 | 0.9727 | 1.0281 | 0.284 | 1.716 | 0.276 | 1.669 | 3.078 | 0.3249 | 0.797 | 0.223 | 1.777 |
| 11 | 0.285 | 0.927 | 0.9754 | 1.0252 | 0.321 | 1.679 | 0.313 | 1.637 | 3.173 | 0.3152 | 0.787 | 0.256 | 1.744 |
| 12 | 0.266 | 0.886 | 0.9776 | 1.0229 | 0.354 | 1.646 | 0.346 | 1.610 | 3.258 | 0.3069 | 0.778 | 0.283 | 1.717 |
| 13 | 0.249 | 0.850 | 0.9794 | 1.0210 | 0.382 | 1.618 | 0.374 | 1.585 | 3.336 | 0.2998 | 0.770 | 0.307 | 1.693 |
| 14 | 0.235 | 0.817 | 0.9810 | 1.0194 | 0.406 | 1.594 | 0.399 | 1.563 | 3.407 | 0.2935 | 0.763 | 0.328 | 1.672 |
| 15 | 0.223 | 0.789 | 0.9823 | 1.0180 | 0.428 | 1.572 | 0.421 | 1.544 | 3.472 | 0.2880 | 0.756 | 0.347 | 1.653 |
| 16 | 0.212 | 0.763 | 0.9835 | 1.0168 | 0.448 | 1.552 | 0.440 | 1.526 | 3.532 | 0.2831 | 0.750 | 0.363 | 1.637 |
| 17 | 0.203 | 0.739 | 0.9845 | 1.0157 | 0.466 | 1.534 | 0.458 | 1.511 | 3.588 | 0.2787 | 0.744 | 0.378 | 1.622 |
| 18 | 0.194 | 0.718 | 0.9854 | 1.0148 | 0.482 | 1.518 | 0.475 | 1.496 | 3.640 | 0.2747 | 0.739 | 0.391 | 1.608 |
| 19 | 0.187 | 0.698 | 0.9862 | 1.0140 | 0.497 | 1.503 | 0.490 | 1.483 | 3.689 | 0.2711 | 0.734 | 0.403 | 1.597 |
| 20 | 0.180 | 0.680 | 0.9869 | 1.0133 | 0.510 | 1.490 | 0.504 | 1.470 | 3.735 | 0.2677 | 0.729 | 0.415 | 1.585 |
| 21 | 0.173 | 0.663 | 0.9876 | 1.0126 | 0.523 | 1.477 | 0.516 | 1.459 | 3.778 | 0.2647 | 0.724 | 0.425 | 1.575 |
| 22 | 0.167 | 0.647 | 0.9882 | 1.0119 | 0.534 | 1.466 | 0.528 | 1.448 | 3.819 | 0.2618 | 0.720 | 0.434 | 1.566 |
| 23 | 0.162 | 0.633 | 0.9887 | 1.0114 | 0.545 | 1.455 | 0.539 | 1.438 | 3.858 | 0.2592 | 0.716 | 0.443 | 1.557 |
| 24 | 0.157 | 0.619 | 0.9892 | 1.0109 | 0.555 | 1.445 | 0.549 | 1.429 | 3.895 | 0.2567 | 0.712 | 0.451 | 1.548 |
| 25 | 0.153 | 0.606 | 0.9896 | 1.0105 | 0.565 | 1.435 | 0.559 | 1.420 | 3.931 | 0.2544 | 0.708 | 0.459 | 1.541 |

Tabla A.23 La 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 | 0.6320 | 0.2640 | 0.0800 | 0.0190 | 0.0040 | 0.0010 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 2 | 0.8650 | 0.5940 | 0.3230 | 0.1430 | 0.0530 | 0.0170 | 0.0050 | 0.0010 | 0.0000 | 0.0000 |
| 3 | 0.9500 | 0.8010 | 0.5770 | 0.3530 | 0.1850 | 0.0840 | 0.0340 | 0.0120 | 0.0040 | 0.0010 |
| 4 | 0.9820 | 0.9080 | 0.7620 | 0.5670 | 0.3710 | 0.2150 | 0.1110 | 0.0510 | 0.0210 | 0.0080 |
| 5 | 0.9930 | 0.9600 | 0.8750 | 0.7350 | 0.5600 | 0.3840 | 0.2380 | 0.1330 | 0.0680 | 0.0320 |
| 6 | 0.9980 | 0.9830 | 0.9380 | 0.8490 | 0.7150 | 0.5540 | 0.3940 | 0.2560 | 0.1530 | 0.0840 |
| 7 | 0.9990 | 0.9930 | 0.9700 | 0.9180 | 0.8270 | 0.6990 | 0.5500 | 0.4010 | 0.2710 | 0.1700 |
| 8 | 1.0000 | 0.9970 | 0.9860 | 0.9580 | 0.9000 | 0.8090 | 0.6870 | 0.5470 | 0.4070 | 0.2830 |
| 9 | | 0.9990 | 0.9940 | 0.9790 | 0.9450 | 0.8840 | 0.7930 | 0.6760 | 0.5440 | 0.4130 |
| 10 | | 1.0000 | 0.9970 | 0.9900 | 0.9710 | 0.9330 | 0.8700 | 0.7800 | 0.6670 | 0.5420 |
| 11 | | | 0.9990 | 0.9950 | 0.9850 | 0.9620 | 0.9210 | 0.8570 | 0.7680 | 0.6590 |
| 12 | | | 1.0000 | 0.9980 | 0.9920 | 0.9800 | 0.9540 | 0.9110 | 0.8450 | 0.7580 |
| 13 | | | | 0.9990 | 0.9960 | 0.9890 | 0.9740 | 0.9460 | 0.9000 | 0.8340 |
| 14 | | | | 1.0000 | 0.9980 | 0.9940 | 0.9860 | 0.9680 | 0.9380 | 0.8910 |
| 15 | | | | | 0.9990 | 0.9970 | 0.9920 | 0.9820 | 0.9630 | 0.9300 |

A.24 Demostración de la media de la distribución hipergeométrica

Para calcular la media de la distribución hipergeométrica escribimos

$$\begin{aligned}
 E(X) &= \sum_{x=0}^n x \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} = k \sum_{x=1}^n \frac{(k-1)!}{(x-1)!(k-x)!} \cdot \frac{\binom{N-k}{n-x}}{\binom{N}{n}} \\
 &= k \sum_{x=1}^n \frac{\binom{k-1}{x-1} \binom{N-k}{n-x}}{\binom{N}{n}}.
 \end{aligned}$$

Puesto que

$$\binom{N-k}{n-1-y} = \binom{(N-1)-(k-1)}{n-1-y} \quad \text{y} \quad \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N}{n} \binom{N-1}{n-1},$$

y con $y = x - 1$, obtenemos

$$\begin{aligned}
 E(X) &= k \sum_{y=0}^{n-1} \frac{\binom{k-1}{y} \binom{N-k}{n-1-y}}{\binom{N}{n}} \\
 &= \frac{nk}{N} \sum_{y=0}^{n-1} \frac{\binom{k-1}{y} \binom{(N-1)-(k-1)}{n-1-y}}{\binom{N-1}{n-1}} = \frac{nk}{N},
 \end{aligned}$$

ya que la sumatoria representa el total de todas las probabilidades en un experimento hipergeométrico cuando

$N - 1$ artículos se seleccionan al azar de $N - 1$, de los cuales $k - 1$ se etiqueta como éxitos.

A.25 Demostración de la media y la varianza de la distribución de Poisson

Sea $\mu = \lambda t$.

$$E(X) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu} \mu^x}{x!} = \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu} \mu^x}{x!} = \mu \sum_{x=1}^{\infty} \frac{e^{-\mu} \mu^{x-1}}{(x-1)!}.$$

Puesto que la sumatoria en el último término de la expresión anterior es la probabilidad total de una variable aleatoria de Poisson con media μ , la cual puede verse con facilidad con $y = x - 1$, es igual a 1. Por lo tanto, $E(X) = \mu$. Para calcular la varianza de X observe que

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\mu} \mu^x}{x!} = \mu^2 \sum_{x=2}^{\infty} \frac{e^{-\mu} \mu^{x-2}}{(x-2)!}.$$

Nuevamente, sea $y = x - 2$, la sumatoria en el último término de la expresión anterior es la probabilidad total de una variable aleatoria de Poisson con media μ . En consecuencia, obtenemos

$$\sigma^2 = E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2 = \mu^2 + \mu - \mu^2 = \mu = \lambda t.$$

A.26 Prueba de la media y la varianza de la distribución gamma

Para calcular la media y la varianza de la distribución gamma comenzamos por calcular

$$E(X^k) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{\infty} x^{\alpha+k-1} e^{-x/\beta} dx = \frac{\beta^{\alpha+k} \Gamma(\alpha+k)}{\beta^\alpha \Gamma(\alpha)} \int_0^{\infty} \frac{x^{\alpha+k-1} e^{-x/\beta}}{\beta^{\alpha+k} \Gamma(\alpha+k)} dx,$$

para $k = 0, 1, 2, \dots$. Puesto que el integrando en el último término de la expresión anterior es una función de densidad gamma, con parámetros $\alpha + k$ y β , es igual a 1. Por lo tanto,

$$E(X^k) = \beta^k \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}.$$

Si utilizamos la fórmula de recursividad de la función gamma de la página 194, obtenemos

$$\mu = \beta \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha\beta \quad \text{y} \quad \sigma^2 = E(X^2) - \mu^2 = \beta^2 \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} - \mu^2 = \beta^2 \alpha(\alpha+1) - (\alpha\beta)^2 = \alpha\beta^2.$$

Apéndice B

Respuestas a los ejercicios impares (no de repaso)

Capítulo 1

- 1.1 a) Tamaño de la muestra = 15
 b) Media de la muestra = 3.787
 c) Mediana de la muestra = 3.6
 e) $\bar{x}_{tr(20)} = 3.678$
 f) Son casi iguales.
- 1.3 b) Sí, el proceso de envejecimiento redujo la resistencia a la tensión.
 c) $\bar{x}_{\text{Con envejecimiento}} = 209.90$,
 $\bar{x}_{\text{Sin envejecimiento}} = 222.10$
 d) $\bar{x}_{\text{Con envejecimiento}} = 210.00$,
 $\bar{x}_{\text{Sin envejecimiento}} = 221.50$
 Las medias y las medianas son similares en cada grupo.
- 1.5 b) Control $\bar{x} = 5.60$, $\bar{x} = 5.00$, $\bar{x}_{tr(10)} = 5.13$.
 Tratamiento: $\bar{x} = 7.60$, $\bar{x} = 4.50$, $\bar{x}_{tr(10)} = 5.63$.
 c) El valor extremo de 37 en el grupo de tratamiento desempeña un papel significativo en el cálculo de la media.
- 1.7 Varianza de la muestra = 0.943
 Desviación estándar de la muestra = 0.971
- 1.9 a) Sin envejecimiento: varianza de la muestra = 23.66,
 desviación estándar de la muestra = 4.86.
 Con envejecimiento: varianza de la muestra = 42.10,
 desviación estándar de la muestra = 6.49.
 b) Con base en las cifras del inciso a), la variación en la situación "con envejecimiento" es menor que en la situación "sin envejecimiento", aunque la diferencia no sea evidente en la gráfica.

- 1.11 Control: varianza de la muestra = 69.38,
 desviación estándar de la muestra = 8.33.
 Tratamiento: varianza de la muestra = 128.04,
 desviación estándar de la muestra = 11.32.
- 1.13 a) Media = 124.3, mediana = 120
 b) 175 es una observación extrema.
- 1.15 Sí, el valor $P = 0.03125$; probabilidad de obtener $HHHHH$ con una moneda legal.
- 1.17 a) Las medias muestrales de no fumadores y fumadores son 30.32 y 43.70, respectivamente.
 b) Las desviaciones estándar de la muestra de no fumadores y fumadores son 7.13 y 16.93, respectivamente.
 d) Parece que a los fumadores les toma más tiempo quedarse dormidos. El tiempo que tardan los fumadores en quedarse dormidos es más variable.

1.19 a)

| Tallo | Hojas | Frecuencia |
|-------|----------|------------|
| 0 | 22233457 | 8 |
| 1 | 023558 | 6 |
| 2 | 035 | 3 |
| 3 | 03 | 2 |
| 4 | 057 | 3 |
| 5 | 0569 | 4 |
| 6 | 0005 | 4 |

b)

| Intervalo de clase | Punto medio de la clase | Frecuencia | Frecuencia relativa |
|--------------------|-------------------------|------------|---------------------|
| 0.0-0.9 | 0.45 | 8 | 0.267 |
| 1.0-1.9 | 1.45 | 6 | 0.200 |
| 2.0-2.9 | 2.45 | 3 | 0.100 |
| 3.0-3.9 | 3.45 | 2 | 0.067 |
| 4.0-4.9 | 4.45 | 3 | 0.100 |
| 5.0-5.9 | 5.45 | 4 | 0.133 |
| 6.0-6.9 | 6.45 | 4 | 0.133 |

- c) Media muestral = 2.7967
 Rango muestral = 6.3
 Desviación estándar de la muestra = 2.2273

1.21 a) $\bar{x} = 74.02$ y $\bar{y} = 78$
 b) $s = 39.26$

- 1.23 b) $\bar{x}_{1980} = 395.10$, $\bar{x}_{1990} = 160.15$
 c) Las emisiones medias cayeron entre 1980 y 1990; la variabilidad también disminuyó porque no hubo emisiones mucho más grandes.

- 1.25 a) Media muestral = 33.31
 b) Mediana muestral = 26.35
 d) $\bar{x}_{tr(10)} = 30.97$

Capítulo 2

- 2.1 a) $S = \{8, 16, 24, 32, 40, 48\}$
 b) $S = \{-5, 1\}$
 c) $S = \{T, HT, HHT, HHH\}$
 d) $S = \{\text{África, Antártida, Asia, Australia, Europa, Norteamérica, Sudamérica}\}$
 e) $S = \phi$

2.3 $A = C$

- 2.5 Si utilizamos un diagrama de árbol, obtenemos
 $S = \{1HH, 1HT, 1TH, 1TT, 2H, 2T, 3HH, 3HT, 3TH, 3TT, 4H, 4T, 5HH, 5HT, 5TH, 5TT, 6H, 6T\}$

- 2.7 $S_1 = \{HHHH, HHHM, HHMH, HMHH, MHHH, HHMM, HMHM, HMMH, MHMH, MMHH, MHHM, HMMM, MHMM, MMHM, MMMH, MMMM\}$;
 $S_2 = \{0, 1, 2, 3, 4\}$

- 2.9 a) $A = \{1HH, 1HT, 1TH, 1TT, 2H, 2T\}$
 b) $B = \{1TT, 3TT, 5TT\}$
 c) $A' = \{3HH, 3HT, 3TH, 3TT, 4H, 4T, 5HH, 5HT, 5TH, 5TT, 6H, 6T\}$
 d) $A' \cap B = \{3TT, 5TT\}$
 e) $A \cup B = \{1HH, 1HT, 1TH, 1TT, 2H, 2T, 3TT, 5TT\}$

- 2.11 a) $S = \{H_1H_2, H_1M_1, H_1M_2, H_2H_1, H_2M_1, H_2M_2, M_1H_1, M_1H_2, M_1M_2, M_2H_1, M_2H_2, M_2M_1\}$

- b) $A = \{H_1H_2, H_1M_1, H_1M_2, H_2H_1, H_2M_1, H_2M_2\}$
 c) $B = \{H_1M_1, H_1M_2, H_2M_1, H_2M_2, M_1H_1, M_1H_2, M_2H_1, M_2H_2\}$
 d) $C = \{M_1M_2, M_2M_1\}$
 e) $A \cap B = \{H_1M_1, H_1M_2, H_2M_1, H_2M_2\}$
 f) $A \cup B = \{H_1H_2, H_1M_1, H_1M_2, H_2H_1, H_2M_1, H_2M_2, M_1H_1, M_1H_2, M_2H_1, M_2H_2, M_2M_1\}$

- 2.15 a) {nitrógeno, potasio, uranio, oxígeno}
 b) {cobre, sodio, zinc, oxígeno}
 c) {cobre, sodio, nitrógeno, potasio, uranio, zinc}
 d) {cobre uranio, zinc}
 e) ϕ
 f) {oxígeno}

- 2.19 a) La familia experimentará fallas mecánicas, pero no recibirá una infracción por cometer una falta de tránsito, y no llegará a un lugar para acampar que esté lleno.
 b) La familia recibirá una infracción por cometer una falta de tránsito y llegará a un lugar para acampar que esté lleno, pero no experimentará fallas mecánicas.
 c) La familia experimentará fallas mecánicas y llegará a un lugar para acampar que esté lleno.
 d) La familia recibirá una infracción por cometer una falta de tránsito, pero no llegará a un lugar para acampar que esté lleno.
 e) La familia no experimentará fallas mecánicas.

2.21 18

2.23 156

2.25 20

2.27 48

2.29 210

2.31 72

2.33 a) 1024; b) 243

2.35 362,880

2.37 2880

2.39 a) 40,320; b) 336

2.41 360

- 2.43 24
- 2.45 3360
- 2.47 56
- 2.49 a) La suma de las probabilidades excede a 1.
 b) La suma de las probabilidades es menor que 1.
 c) Una probabilidad negativa.
 d) La probabilidad de un corazón y de una carta negra es cero.
- 2.51 $S = \{\$10, \$25, \$100\}$; $P(10) = \frac{11}{20}$, $P(25) = \frac{3}{10}$,
 $P(100) = \frac{15}{100}$; $\frac{17}{20}$
- 2.53 a) 0.3; b) 0.2
- 2.55 10/117
- 2.57 a) 5/26; b) 9/26; c) 19/26
- 2.59 a) 94/54,145; b) 143/39,984
- 2.61 a) 22/25; b) 3/25; c) 17/50
- 2.63 a) 0.32; b) 0.68; vc) oficina o estudio
- 2.65 a) 0.8; b) 0.45; c) 0.55
- 2.67 a) 0.31; b) 0.93; c) 0.31
- 2.69 a) 0.009; b) 0.999; c) 0.01
- 2.71 a) 0.048; b) \$50,000; c) \$12,500
- 2.73 a) La probabilidad de que un convicto que vende drogas también cometa un robo a mano armada.
 b) La probabilidad de que un convicto que comete un robo a mano armada no venda drogas.
 c) La probabilidad de que un convicto que no vende drogas tampoco cometa un robo a mano armada.
- 2.75 a) 14/39; b) 95/112
- 2.77 a) 5/34; b) 3/8
- 2.79 a) 0.018; b) 0.614; c) 0.166; d) 0.479
- 2.81 a) 0.35; b) 0.875; c) 0.55
- 2.83 a) 9/28; b) 3/4; c) 0.91
- 2.85 0.27
- 2.87 5/8
- 2.89 a) 0.0016; b) 0.9984
- 2.91 a) 91/323; b) 91/323

- 2.93 a) 0.75112; b) 0.2045
- 2.95 0.0960
- 2.97 0.40625
- 2.99 0.1124
- 2.101 0.857

Capítulo 3

- 3.1 Discreta; continua; continua; discreta; discreta; continua.

3.3 Espacio muestral w

| | |
|-----|----|
| HHH | 3 |
| HHT | 1 |
| HTH | 1 |
| THH | 1 |
| HTT | -1 |
| THT | -1 |
| TTH | -1 |
| TTT | -3 |

- 3.5 a) 1/30; b) 1/10

- 3.7 a) 0.68; b) 0.375

- 3.9 b) 19/80

| x | 0 | 1 | 2 |
|--------|---------------|---------------|---------------|
| $f(x)$ | $\frac{2}{7}$ | $\frac{4}{7}$ | $\frac{1}{7}$ |

$$3.13 \quad F(x) = \begin{cases} 0, & \text{para } x < 0, \\ 0.41, & \text{para } 0 \leq x < 1, \\ 0.78, & \text{para } 1 \leq x < 2, \\ 0.94, & \text{para } 2 \leq x < 3, \\ 0.99, & \text{para } 3 \leq x < 4, \\ 1, & \text{para } x \geq 4 \end{cases}$$

$$3.15 \quad F(x) = \begin{cases} 0, & \text{para } x < 0, \\ \frac{2}{7}, & \text{para } 0 \leq x < 1, \\ \frac{6}{7}, & \text{para } 1 \leq x < 2, \\ 1, & \text{para } x \geq 2 \end{cases}$$

- a) 4/7; b) 5/7

- 3.17 b) 1/4; c) 0.3

$$3.19 \quad F(x) = \begin{cases} 0, & x < 1 \\ \frac{x-1}{2}, & 1 \leq x < 3; 1/4 \\ 1, & x \geq 3 \end{cases}$$

$$3.21 \text{ a) } 3/2; \text{ b) } F(x) = \begin{cases} 0, & x < 0 \\ x^{3/2}, & 0 \leq x < 1; 0.3004 \\ 1, & x \geq 1 \end{cases}$$

$$3.23 \quad F(w) = \begin{cases} 0, & \text{para } w < -3, \\ \frac{1}{27}, & \text{para } -3 \leq w < -1, \\ \frac{7}{27}, & \text{para } -1 \leq w < 1, \\ \frac{19}{27}, & \text{para } 1 \leq w < 3, \\ 1, & \text{para } w \geq 3 \end{cases}$$

a) 20/27; b) 2/3

| | | | | |
|------|------------|---------------|---------------|---------------|
| 3.25 | t | 20 | 25 | 30 |
| | $P(T = t)$ | $\frac{1}{5}$ | $\frac{3}{5}$ | $\frac{1}{5}$ |

$$3.27 \text{ a) } F(x) = \begin{cases} 0, & x < 0, \\ 1 - \exp(-x/2000), & x \geq 0 \end{cases}$$

b) 0.6065; c) 0.6321

$$3.29 \text{ b) } F(x) = \begin{cases} 0, & x < 1, \\ 1 - x^{-3}, & x \geq 1 \end{cases}$$

c) 0.0156

3.31 a) 0.2231; b) 0.2212

3.33 a) $k = 280$; b) 0.3633; c) 0.0563

3.35 a) 0.1528; b) 0.0446

3.37 a) 1/36; b) 1/15

| | | | | | | |
|------|----|-----------|----------------|-----------------|-----------------|----------------|
| 3.39 | a) | | x | | | |
| | | $f(x, y)$ | 0 | 1 | 2 | 3 |
| | | 0 | 0 | $\frac{3}{70}$ | $\frac{9}{70}$ | $\frac{3}{70}$ |
| | | 1 | $\frac{2}{70}$ | $\frac{18}{70}$ | $\frac{18}{70}$ | $\frac{2}{70}$ |
| | | 2 | $\frac{3}{70}$ | $\frac{9}{70}$ | $\frac{3}{70}$ | 0 |

b) 1/2

3.41 a) 1/16; b) $g(x) = 12x(1-x)^2$, para $0 \leq x \leq 1$;
c) 1/4

3.43 a) 3/64; b) 1/2

3.45 0.6534

3.47 a) Dependiente; b) 1/3

| | | | | | |
|------|----|--------|------|------|------|
| 3.49 | a) | x | 1 | 2 | 3 |
| | | $g(x)$ | 0.10 | 0.35 | 0.55 |

| | | | | | |
|--|----|--------|------|------|------|
| | b) | y | 1 | 2 | 3 |
| | | $h(y)$ | 0.20 | 0.50 | 0.30 |

c) 0.2857

| | | | | | | |
|------|----|-----------|----------------|-----------------|----------------|----------------|
| 3.51 | a) | | x | | | |
| | | $f(x, y)$ | 0 | 1 | 2 | 3 |
| | | 0 | $\frac{1}{55}$ | $\frac{6}{55}$ | $\frac{6}{55}$ | $\frac{1}{55}$ |
| | | 1 | $\frac{6}{55}$ | $\frac{16}{55}$ | $\frac{6}{55}$ | 0 |
| | | 2 | $\frac{6}{55}$ | $\frac{6}{55}$ | 0 | 0 |
| | | 3 | $\frac{1}{55}$ | 0 | 0 | 0 |

b) 42/55

3.53 5/8

3.55 Independiente

3.57 a) 3; b) 21/512

3.59 Dependiente

Capítulo 4

4.1 0.88

4.3 25 centavos

4.5 \$1.23

4.7 \$500

4.9 \$6900

4.11 $(\ln 4)/\pi$

4.13 100 horas

4.15 0

4.17 209

4.19 \$1855

4.21 \$833.33

4.23 a) 35.2; b) $\mu_x = 3.20$; $\mu_y = 3.00$

4.25 2

4.27 2000 horas

4.29 b) 3/2

4.31 a) 1/6; b) $(5/6)^5$

4.33 \$5,250,000

4.35 0.74

4.37 1/18; en términos de utilidad real la varianza es $\frac{1}{18} (5000)^2$

4.39 1/6

4.41 118.9

4.43 $\mu_y = 10$; $\sigma_y^2 = 144$

4.45 0.01

4.47 -0.0062

4.49 $\sigma_x^2 = 0.8456, \sigma_x = 0.9196$

4.51 $-1/\sqrt{5}$

4.53 $\mu_{g(x)} = 10.33, \sigma_{g(x)} = 6.66$

4.55 \$0.80

4.57 209

4.59 $\mu = 7/2, \sigma^2 = 15/4$

4.61 3/14

4.63 52

4.65 a) 7; b) 0; c) 12.25

4.67 46/63

4.69 a) $E(X) = E(Y) = 1/3$ y $\text{Var}(X) = \text{Var}(Y) = 4/9$; b) $E(Z) = 2/3$ y $\text{Var}(Z) = 8/9$

4.71 a) 4; b) 32; 16

4.73 Mediante cálculo directo, $E(e^Y) = 1884.32$. Si usamos la aproximación de ajuste de segundo orden, $E(e^Y) \approx 1883.38$, que se acerca mucho al valor real.

4.75 0.03125

4.77 a) A lo sumo 4/9; b) al menos 5/9; c) al menos 21/25; d) 10

5.17 $\mu = 3.5, \sigma^2 = 1.05$

5.19 $f(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3} 0.35^{x_1} 0.05^{x_2} 0.60^{x_3}$

5.21 0.0095

5.23 0.0077

5.25 0.8670

5.27 a) 0.2852; b) 0.9887; c) 0.6083

5.29 5/14

5.31 $h(x; 6, 3, 4) = \frac{\binom{4}{x} \binom{2}{3-x}}{\binom{6}{3}}$, para $x = 1, 2, 3$,
 $P(2 \leq X \leq 3) = 4/5$

5.33 a) 0.3246; b) 0.4496

5.35 0.9517

5.37 a) 0.6815; b) 0.1153

5.39 0.9453

5.41 0.6077

5.43 a) 4/33; b) 8/165

5.45 0.2315

5.47 a) 0.3991; b) 0.1316

5.49 0.0515

5.51 63/64

5.53 a) 0.3840; b) 0.0067

5.55 a) 0.0630; b) 0.9730

5.57 a) 0.1429; b) 0.1353

5.59 a) 0.1638; b) 0.032

5.61 0.2657

5.63 $\mu = 6, \sigma^2 = 6$

5.65 a) 0.2650; b) 0.9596

5.67 a) 0.8243; b) 14

5.69 4

5.71 $5.53 \times 10^{-4}; \mu = 7.5$

5.73 a) 0.0137; b) 0.0830

5.75 0.4686

Capítulo 5

5.1 $\mu = \frac{1}{k} \sum_{i=1}^k x_i, \sigma^2 = \frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2$

5.3 $f(x) = \frac{1}{10}$ para $x = 1, 2, \dots, 10$ y $f(x) = 0$ en otro caso; 3/10

5.5 a) 0.0480; b) 0.2375; c) $P(X = 5 | p = 0.3) = 0.1789, P = 0.3$ es razonable.

5.7 a) 0.0474; b) 0.0171

5.9 a) 0.7073; b) 0.4613; c) 0.1484

5.11 0.1240

5.13 0.8369

5.15 a) 0.0778; b) 0.3370; c) 0.0870

Capítulo 6

- 6.3 a) 0.6; b) 0.7; c) 0.5
- 6.5 a) 0.0823; b) 0.0250; c) 0.2424; d) 0.9236; e) 0.8133; f) 0.6435
- 6.7 a) 0.54; b) -1.72; c) 1.28
- 6.9 a) 0.1151; b) 16.1; c) 20.275; d) 0.5403
- 6.11 a) 0.0548; b) 0.4514; c) 23 tazas d) 189.95 mililitros
- 6.13 a) 0.8980; b) 0.0287; c) 0.6080
- 6.15 a) 0.0571; b) 99.11%; c) 0.3974; d) 27.952 minutos; e) 0.0092
- 6.17 6.24 años
- 6.19 a) 51%; b) \$18.37
- 6.21 a) 0.0401; b) 0.0244
- 6.23 26 estudiantes
- 6.25 a) 0.3085; b) 0.0197
- 6.27 a) 0.9514; b) 0.0668
- 6.29 a) 0.1171; b) 0.2049
- 6.31 0.1357
- 6.33 a) 0.0778; b) 0.0571; c) 0.6811
- 6.35 a) 0.8749; b) 0.0059
- 6.37 a) 0.0228; b) 0.3974
- 6.41 $2.8e^{-1.8} - 3.4e^{-2.4} = 0.1545$
- 6.43 a) $\mu = 6$; $\sigma^2 = 18$;
b) entre 0 y 14.485 millones de litros
- 6.45 $\sum_{x=4}^6 \binom{6}{x} (1 - e^{-3/4})^x (e^{-3/4})^{6-x} = 0.3968$
- 6.47 a) $\sqrt{\pi/2} = 1.2533$ años; b) e^{-2}
- 6.49 a) Media = 0.25, mediana = 0.206; b) varianza = 0.0375; c) 0.2963
- 6.51 $e^{-4} = 0.0183$
- 6.53 a) $\mu = \alpha\beta = 50$; b) $\sigma^2 = \alpha\beta^2 = 500$; $\sigma = \sqrt{500}$
c) 0.815
- 6.55 a) 0.1889; b) 0.0357

6.57 Media = e^6 , varianza = $e^{12}(e^4 - 1)$

6.59 a) e^{-5} ; b) $\beta = 0.2$

Capítulo 7

- 7.1 $g(y) = 1/3$, para $y = 1, 3, 5$
- 7.3 $g(y_1, y_2) = \left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}, 2 - y_1 \right) \times \left(\frac{1}{4} \right)^{(y_1 + y_2)/2} \left(\frac{1}{3} \right)^{(y_1 - y_2)/2} \left(\frac{5}{12} \right)^{2 - y_1}$;
para $y_1 = 0, 1, 2$; $y_2 = -2, -1, 0, 1, 2$;
 $y_2 \leq y_1$; $y_1 + y_2 = 0, 2, 4$
- 7.7 Distribución gamma con $\alpha = 3/2$ y $\beta = m/2b$
- 7.9 a) $g(y) = 32/y^3$, para $y > 4$; b) $1/4$
- 7.11 $h(z) = 2(1 - z)$, para $0 < z < 1$
- 7.13 $h(w) = 6 + 6w - 12w^{1/2}$, para $0 < w < 1$
- 7.15 $g(y) = \begin{cases} \frac{2}{9\sqrt{y}}, & 0 < y < 1, \\ \frac{\sqrt{y} + 1}{9\sqrt{y}}, & 1 < y < 4 \end{cases}$
- 7.19 Ambas son iguales a μ .
- 7.23 a) Gamma(2, 1); b) Uniforme(0, 1)

Capítulo 8

- 8.1 a) Las respuestas de todas las personas en Richmond que tienen teléfono;
b) Resultados para un número grande o infinito de lanzamientos de una moneda;
c) Periodo de vida de tal calzado deportivo cuando es utilizado en el torneo profesional;
d) Todos los posibles intervalos de tiempo para esta abogada que maneja desde su casa hasta su oficina.
- 8.3 a) $\bar{x} = 3.2$ segundos; b) $\bar{x} = 3.1$ segundos
- 8.5 a) $\bar{x} = 2.4$; b) $\bar{x} = 2$; c) $m = 3$
- 8.7 a) 53.75; b) 75 y 100
- 8.9 a) El rango es 10; b) $s = 3.307$
- 8.11 a) 2.971; b) 2.971

- 8.13 $s = 0.585$
- 8.15 a) 45.9; b) 5.1
- 8.17 0.3159
- 8.19 a) La varianza se reduce de 0.49 a 0.16;
b) La varianza se incrementa de 0.04 a 0.64.
- 8.21 Sí.
- 8.23 a) $\mu = 5.3$; $\sigma^2 = 0.81$;
b) $\mu_{\bar{x}} = 5.3$; $\sigma_{\bar{x}}^2 = 0.0225$;
c) 0.9082
- 8.25 a) 0.6898; b) 7.35
- 8.29 0.5596
- 8.31 a) La probabilidad de que el tiempo promedio de secado sea mayor que 1.0 es 0.0013; b) 13
- 8.33 a) 1/2; b) 0.3085
- 8.35 $P(\bar{X} \leq 775 | \mu = 760) = 0.9332$
- 8.37 a) 27.488; b) 18.475; c) 36.415
- 8.39 a) 0.297; b) 32.852; c) 46.928
- 8.41 a) 0.05; b) 0.94
- 8.45 a) 0.975; b) 0.10; c) 0.875; d) 0.99
- 8.47 a) 2.500; b) 1.319; c) 1.714
- 8.49 No; $\mu > 20$
- 8.51 a) 2.71; b) 3.51; c) 2.92; d) 0.47; e) 0.34
- 8.53 La razón F es 1.44. Las varianzas no son significativamente diferentes.
- 9.15 (13,075, 33,925)
- 9.17 (6.05, 16.55)
- 9.19 323.946 a 326.154
- 9.21 Límite superior de predicción: 9.42; límite superior de tolerancia: 11.72
- 9.25 Sí, el valor de 6.9 está fuera del intervalo de predicción.
- 9.27 a) (0.9876, 1.0174);
b) (0.9411, 1.0639);
c) (0.9334, 1.0716)
- 9.35 $2.9 < \mu_1 - \mu_2 < 7.1$
- 9.37 $2.80 < \mu_1 - \mu_2 < 3.40$
- 9.39 $1.5 < \mu_1 - \mu_2 < 12.5$
- 9.41 $0.70 < \mu_1 - \mu_2 < 3.30$
- 9.43 $-6536 < \mu_1 - \mu_2 < 2936$
- 9.45 (-0.74, 6.30)
- 9.47 (-6.92, 36.70)
- 9.49 $0.54652 < \mu_B - \mu_A < 1.69348$
- 9.51 Método 1: $0.194 < p < 0.262$; método 2: $0.1957 < p < 0.2639$
- 9.53 a) $0.498 < p < 0.642$; b) error ≤ 0.072
- 9.55 a) $0.739 < p < 0.961$; b) no
- 9.57 a) $0.644 < p < 0.690$; b) error ≤ 0.023
- 9.59 2576
- 9.61 160
- 9.63 9604
- 9.65 $-0.0136 < p_T - p_M < 0.0636$
- 9.67 $0.0011 < p_1 - p_2 < 0.0869$
- 9.69 (-0.0849, 0.0013); no es significativamente diferente.
- 9.71 $0.293 < \sigma^2 < 6.736$; la afirmación es válida
- 9.73 $3.472 < \sigma^2 < 12.804$
- 9.75 $9.27 < \sigma < 34.16$
- 9.77 $0.549 < \sigma_1 / \sigma_2 < 2.690$
- 9.79 $0.016 < \sigma_1^2 / \sigma_2^2 < 0.454$; no

Capítulo 9

- 9.1 56
- 9.3 $0.3097 < \mu < 0.3103$
- 9.5 a) $22,496 < \mu < 24,504$; b) error ≤ 1004
- 9.7 35
- 9.9 $10.15 < \mu < 12.45$
- 9.11 $0.978 < \mu < 1.033$
- 9.13 $47.722 < \mu < 49.278$

$$9.81 \frac{1}{n} \sum_{i=1}^n x_i$$

$$9.83 \hat{\beta} = \bar{x}/5$$

$$9.85 \hat{\theta} = \text{máx}\{x_1, \dots, x_n\}$$

9.87 $x \ln p + (1-x) \ln(1-p)$. Sea la derivada con respecto a $p = 0$; $\hat{p} = x = 1.0$

10.33 $t = 1.50$; no hay evidencia suficiente para concluir que el incremento en la concentración de sustrato causaría un incremento en la velocidad media de más de 0.5 micromoles por 30 minutos.

10.35 $t = 0.70$; no hay suficiente evidencia que apoye la conclusión de que el suero es efectivo.

10.37 $t = 2.55$; rechace H_0 ; $\mu_1 - \mu_2 > 4$ kilómetros.

10.39 $t' = 0.22$; no rechace H_0 .

10.41 $t' = 2.76$; rechace H_0 .

10.43 $t = -2.53$; rechace H_0 ; la afirmación es válida.

10.45 $t = 2.48$; valor $P < 0.02$; rechace H_0 .

10.47 $n = 6$

10.49 $78.28 \approx 79$

10.51 5

10.53 a) $H_0: M_{\text{caliente}} - M_{\text{frío}} = 0$,

$H_1: M_{\text{caliente}} - M_{\text{frío}} \neq 0$;

b) t apareada, $t = 0.99$; valor $P > 0.30$; no rechace H_0 .

10.55 valor $P = 0.4044 =$ (con una prueba de una cola); no se refuta la afirmación.

10.57 $z = 1.44$; no rechace H_0 .

10.59 $z = -5.06$ con valor $P \approx 0$; concluya que menos de una quinta parte de los hogares se calienta con petróleo.

10.61 $z = 0.93$ con valor $P = P(Z > 0.93) = 0.1762$; no hay evidencia suficiente para concluir que la nueva medicina es eficaz.

10.63 $z = 2.36$ con valor $P = 0.0182$; sí, la diferencia es significativa.

10.65 $z = 1.10$ con valor $P = 0.1357$; no tenemos evidencia suficiente para concluir que el cáncer de mama es más frecuente en las comunidades urbanas.

10.67 $\chi^2 = 18.13$ con valor $P = 0.0676$ (de los resultados por computadora); no rechace H_0 ; $\sigma^2 = 0.03$.

10.69 $\chi^2 = 63.75$ con valor $P = 0.8998$ (de los resultados por computadora); no rechace H_0 .

10.71 $\chi^2 = 42.37$ con valor $P = 0.0117$ (de los resultados por computadora); la máquina está fuera de control.

10.73 $f = 1.33$ con valor $P = 0.3095$ (de los resultados por computadora); no rechace H_0 ; $\sigma_1 = \sigma_2$.

Capítulo 10

10.1 a) Concluya que menos de 30% del público es alérgico a ciertos productos de queso cuando, de hecho, 30% o más es alérgico.

b) Concluya que al menos 30% del público es alérgico a ciertos productos de queso cuando, de hecho, menos de 30% es alérgico.

10.3 a) La empresa no es culpable;

b) la empresa es culpable.

10.5 a) 0.0559;

b) $\beta = 0.0017$; $\beta = 0.00968$; $\beta = 0.5557$

10.7 a) 0.1286;

b) $\beta = 0.0901$; $\beta = 0.0708$.

c) La probabilidad de un error tipo I es algo grande.

10.9 a) $\alpha = 0.0850$; b) $\beta = 0.3410$

10.11 a) $\alpha = 0.1357$; b) $\beta = 0.2578$

10.13 $\alpha = 0.0094$; $\beta = 0.0122$

10.15 a) $\alpha = 0.0718$; b) $\beta = 0.1151$

10.17 a) $\alpha = 0.0384$; b) $\beta = 0.5$; $\beta = 0.2776$

10.19 $z = -2.76$; sí, $\mu < 40$ meses; valor $P = 0.0029$

10.21 $z = -1.64$; valor $P = 0.10$

10.23 $t = 0.77$; no rechace H_0 .

10.25 $z = 8.97$; sí, $\mu > 20,000$ kilómetros; valor $P < 0.001$

10.27 $t = 12.72$; valor $P < 0.0005$; rechace H_0 .

10.29 $t = -1.98$; valor $P = 0.0312$; rechace H_0 .

10.31 $z = -2.60$; concluya que $\mu_A - \mu_B \leq 12$ kilogramos.

10.75 $f = 0.086$ con valor $P = 0.0328$ (de los resultados por computadora); rechace $H_0: \sigma_1 = \sigma_2$ a un nivel mayor que 0.0328.

10.77 $f = 19.67$ con valor $P = 0.0008$ (de los resultados por computadora); rechace $H_0: \sigma_1 = \sigma_2$.

10.79 $\chi^2 = 10.14$; rechace H_0 ; la razón no es 5:2:2:1.

10.81 $\chi^2 = 4.47$; no hay evidencia suficiente para afirmar que el dado esté cargado.

10.83 $\chi^2 = 3.125$; no rechace H_0 ; distribución geométrica.

10.85 $\chi^2 = 5.19$; no rechace H_0 ; distribución normal.

10.87 $\chi^2 = 5.47$; no rechace H_0 .

10.89 $\chi^2 = 124.59$; sí, la ocurrencia de estos tipos de delitos depende del distrito de la ciudad.

10.91 $\chi^2 = 5.92$ con valor $P = 0.4332$; no rechace H_0 .

10.93 $\chi^2 = 31.17$ con valor $P < 0.0001$; las actitudes no son homogéneas.

10.95 $\chi^2 = 1.84$; no rechace H_0 .

b) $4.324 < \beta_0 < 8.503$;

c) $0.446 < \beta_1 < 3.172$

11.19 a) $s^2 = 6.626$;

b) $2.684 < \beta_0 < 8.968$;

c) $0.498 < \beta_1 < 0.637$

11.21 $t = -2.24$; rechace H_0 y concluya $\beta < 6$

11.23 a) $24.438 < \mu_{Y|24.5} < 27.106$;

b) $21.88 < y_0 < 29.66$

11.25 $7.81 < \mu_{Y|1.6} < 10.81$

11.27 a) 17.1812 mpg;

b) no, el intervalo de confianza de 95% sobre la media mpg es (27.95, 29.60);

c) las millas por galón probablemente excederán a 18.

11.29 b) $\hat{y} = 3.4156x$

11.31 El valor f para probar la falta de ajuste es 1.58 y se concluye que no se rechaza H_0 . Por lo tanto, la prueba de falta de ajuste es insignificante.

11.33 a) $\hat{y} = 2.003x$;

b) $t = 1.40$, no rechace H_0 .

11.35 $f = 1.71$ y valor $P = 0.2517$; la regresión es lineal.

11.37 a) $b_0 = 10.812$, $b_1 = -0.3437$;

b) $f = 0.43$; la regresión es lineal.

11.39 a) $\hat{P} = -11.3251 - 0.0449T$;

b) sí;

c) $R^2 = 0.9355$;

d) sí

11.41 b) $\hat{N} = -175.9025 + 0.0902Y$; $R^2 = 0.3322$

11.43 $r = 0.240$

11.45 a) $r = -0.979$;

b) Valor $P = 0.0530$; no rechace H_0 a un nivel de 0.025;

c) 95.8%

11.47 a) $r = 0.784$;

b) rechace H_0 y concluya que $\rho > 0$;

c) 61.5%.

Capítulo 11

11.1 a) $b_0 = 64.529$, $b_1 = 0.561$;

b) $\hat{y} = 81.4$

11.3 a) $\hat{y} = 5.8254 + 0.5676x$;

c) $\hat{y} = 34.205$ a 50°C

11.5 a) $\hat{y} = 6.4136 + 1.8091x$;

b) $\hat{y} = 9.580$ a temperatura 1.75

11.7 b) $\hat{y} = 31.709 + 0.353x$

11.9 b) $\hat{y} = 343.706 + 3.221x$;

c) $\hat{y} = \$456$ con costos de publicidad = \$35

11.11 b) $\hat{y} = -1847.633 + 3.653x$

11.13 a) $\hat{y} = 153.175 - 6.324x$;

b) $\hat{y} = 123$ para $x = 4.8$ unidades

11.15 a) $s^2 = 176.4$;

b) $t = 2.04$; no rechace H_0 ; $\beta_1 = 0$

11.17 a) $s^2 = 0.40$;

Capítulo 12

- 12.1 $\hat{y} = 0.5800 + 2.7122x_1 + 2.0497x_2$
- 12.3 a) $\hat{y} = 27.547 + 0.922x_1 + 0.284x_2$;
b) $\hat{y} = 84$ para $x_1 = 64$ y $x_2 = 4$
- 12.5 a) $\hat{y} = -102.7132 + 0.6054x_1 + 8.9236x_2 + 1.4374x_3 + 0.0136x_4$;
b) $\hat{y} = 287.6$
- 12.7 $\hat{y} = 141.6118 - 0.2819x + 0.0003x^2$
- 12.9 a) $\hat{y} = 56.4633 + 0.1525x - 0.00008x^2$;
b) $\hat{y} = 86.7\%$ cuando la temperatura es de 225°C
- 12.11 $\hat{y} = -6.5122 + 1.9994x_1 - 3.6751x_2 + 2.5245x_3 + 5.1581x_4 + 14.4012x_5$
- 12.13 a) $\hat{y} = 350.9943 - 1.2720x_1 - 0.1539x_2$;
b) $\hat{y} = 140.9$
- 12.15 $\hat{y} = 3.3205 + 0.4210x_1 - 0.2958x_2 + 0.0164x_3 + 0.1247x_4$
- 12.17 0.1651
- 12.19 242.72
- 12.21 a) $\hat{\sigma}_{B_2}^2 = 28.0955$; b) $\hat{\sigma}_{B_1, B_2} = -0.0096$
- 12.23 $t = 5.91$ con valor $P = 0.0002$. Rechace H_0 y asegure que $\beta_1 \neq 0$.
- 12.25 $0.4516 < \mu_Y |_{x_1=900, x_2=1} < 1.2083$
y $-0.1640 < y_0 < 1.8239$
- 12.27 $263.7879 < \mu_Y |_{x_1=75, x_2=24, x_3=90, x_4=98} < 311.3357$ y $243.7175 < y_0 < 331.4062$
- 12.29 a) $t = -1.09$ con valor $P = 0.3562$;
b) $t = -1.72$ con valor $P = 0.1841$;
c) sí; no hay suficiente evidencia que demuestre que los valores de x_1 y x_2 son significativos.
- 12.31 $R^2 = 0.9997$
- 12.33 $f = 5.106$ con valor $P = 0.0303$; la regresión no es significativa en el nivel 0.01.
- 12.35 $f = 34.90$ con valor $P = 0.0002$; rechace H_0 y concluya que $\beta_1 > 0$.
- 12.37 $f = 10.18$ con valor $P < 0.01$; x_1 y x_2 son significativos en la presencia de x_3 y x_4 .
- 12.39 El modelo de dos variables es mejor.
- 12.41 Primer modelo: $R_{\text{aj}}^2 = 92.7\%$, C.V. = 9.0385.
Segundo modelo: $R_{\text{aj}}^2 = 98.1\%$, C.V. = 4.6287.
La prueba F parcial revela un valor $P = 0.0002$; el modelo 2 es mejor.
- 12.43 No hay mucha diferencia entre utilizar x_2 solo y usar x_1 y x_2 juntos, ya que R_{aj}^2 constituye 0.7696 en comparación con 0.7591, respectivamente.
- 12.45 a) $\widehat{\text{mpg}} = 5.9593 - 0.00003773 \text{ odómetro} + 0.3374 \text{ octanaje} - 12.6266z_1 - 12.9846z_2$;
b) sedán
c) no son significativamente diferentes.
- 12.47 b) $\hat{y} = 4.690$ segundos;
c) $4.450 < \mu_{Y|(180, 260)} < 4.930$
- 12.49 $\hat{y} = 2.1833 + 0.9576x_2 + 3.3253x_3$
- 12.51 a) $\hat{y} = -587.211 + 428.433x$;
b) $\hat{y} = 1180 - 191.691x + 35.20945x^2$;
c) modelo cuadrático
- 12.53 $\hat{\sigma}_{B_1}^2 = 20,588$; $\hat{\sigma}_{B_{11}}^2 = 62.6502$;
 $\hat{\sigma}_{B_1, B_{11}} = -1103.5$
- 12.55 a) Es mejor el modelo de intersección.
- 12.57 a) $\hat{y} = 3.1368 + 0.6444x_1 - 0.0104x_2 + 0.5046x_3 - 0.1197x_4 - 2.4618x_5 + 1.5044x_6$;
b) $\hat{y} = 4.6563 + 0.5133x_3 - 0.1242x_4$;
c) Criterio C_p : variables x_1 y x_2 con $s^2 = 0.7317$ y $R^2 = 0.6476$; criterio s^2 : variables x_1, x_3 y x_4 con $s^2 = 0.7251$ y $R^2 = 0.6726$;
d) $\hat{y} = 4.6563 + 0.5133x_3 - 0.1242x_4$; éste no pierde mucho en s^2 y R^2 .
e) dos observaciones tienen valores grandes de R de Student y deben verificarse.
- 12.59 a) $\hat{y} = 125.8655 + 7.7586x_1 + 0.0943x_2 - 0.0092x_3$;
b) el modelo que sólo contiene x_2 es el mejor.
- 12.61 a) $\hat{p} = (1 + e^{2.9949 - 0.0308x})^{-1}$;
b) 1.8515

Capítulo 13

- 13.1 $f = 0.31$; no hay evidencia suficiente para apoyar la hipótesis de que existen diferencias entre las 6 máquinas.
- 13.3 $f = 14.52$; sí, la diferencia es significativa.

- 13.5 $f = 8.38$; las actividades específicas promedio difieren de manera significativa.
- 13.7 $f = 2.25$; no hay evidencia suficiente para apoyar la hipótesis de que las diferentes concentraciones de $MgNH_4PO_4$ influyen significativamente en la altura que alcanzan los crisantemos.
- 13.9 $b = 0.79 > b_4(0.01, 4, 4, 9) = 0.4939$. No rechace H_0 . No hay suficiente evidencia para afirmar que las varianzas son diferentes.
- 13.11 $b = 0.7822 < b_4(0.05, 9, 8, 15) = 0.8055$. Las varianzas son significativamente diferentes.
- 13.13 a) Valor $P < 0.0001$, significativa,
b) para el contraste 1 contra 2, valor $P < 0.0001Z$, significativamente diferentes; para el contraste 3 contra 4, valor $P = 0.0648$, no es significativamente diferente.
- 13.15 A continuación se presentan los resultados para la prueba de Tukey

| | | | | |
|-------------|-------------|-------------|-------------|-------------|
| \bar{y}_4 | \bar{y}_3 | \bar{y}_1 | \bar{y}_5 | \bar{y}_2 |
| 2.98 | 4.30 | 5.44 | 6.96 | 7.90 |

- 13.17 a) valor $P = 0.0121$; sí, hay una diferencia significativa;

| | | | |
|-------------|------------|--------------|---------|
| | | Remoción | |
| | De Hess | del sustrato | |
| Disminución | modificado | de Kicknet | Kicknet |

- 13.19 $f = 70.27$ con valor $P < 0.0001$; rechace H_0 .

| | | | | |
|-------------|----------------|-----------------|----------------|----------------|
| \bar{x}_0 | \bar{x}_{25} | \bar{x}_{100} | \bar{x}_{75} | \bar{x}_{50} |
| 55.167 | 60.167 | 64.167 | 70.500 | 72.833 |

La temperatura es importante; tanto 75° como $50^\circ(C)$ producen baterías con vida activa significativamente más larga.

- 13.21 La absorción media para el agregado 4 es significativamente menor que para el otro agregado.
- 13.23 Al comparar el control con 1 y 2, significativo; al comparar el control con 3 y 4: insignificante
- 13.25 $f(\text{fertilizante}) = 6.11$; existe una diferencia significativa entre los fertilizantes
- 13.27 $f = 5.99$; el porcentaje de aditivos extranjeros no es el mismo para las tres marcas de mermelada; marca A.
- 13.29 Valor $P < 0.0001$; significativo

- 13.31 Valor $P = 0.0023$; significativo
- 13.33 Valor $P = 0.1250$; no significativo
- 13.35 Valor $P < 0.0001$;
 $f = 122.37$; la cantidad de tinta sí influye en el color de la tela.
- 13.37 a) $y_{ij} = \mu + A_i + \epsilon_{ij}$, $A_i \sim n(x; 0, \sigma_a)$,
 $\epsilon_{ij} \sim n(x; 0, \sigma)$;
b) $\hat{\sigma}_a^2 = 0$ (el componente de la varianza estimada es -0.00027 ; $\hat{\sigma}^2 = 0.0206$).
- 13.39 a) $f = 14.9$; los operadores difieren significativamente;
b) $\hat{\sigma}_a^2 = 28.91$; $s^2 = 8.32$.
- 13.41 a) $y_{ij} = \mu + A_i + \epsilon_{ij}$, $A_i \sim n(x; 0, \sigma_a)$;
b) sí, $f = 5.63$ con un valor $P = 0.0121$;
c) hay un componente significativo de varianza del telar.

Capítulo 14

- 14.1 a) $f = 8.13$; significativo;
b) $f = 5.18$; significativo;
c) $f = 1.63$; no significativo
- 14.3 a) $f = 14.81$; significativo;
b) $f = 9.04$; significativo;
c) $f = 0.61$; no significativo;
- 14.5 a) $f = 34.40$; significativo;
b) $f = 26.95$; significativo;
c) $f = 20.30$; significativo;
- 14.7 Prueba del efecto de la temperatura: $f_1 = 10.85$ con valor $P = 0.0002$;
Prueba del efecto de la cantidad de catalizador: $f_2 = 46.63$ con valor $P < 0.0001$;
Prueba del efecto de interacción: $f = 2.06$ con valor $P = 0.074$.
- 14.9 a)

| Fuente de variación | gl | Suma de cuadrados | Cuadrados medios | f | P |
|-----------------------------|----|-------------------|------------------|-------|------------|
| Velocidad de corte | 1 | 12.000 | 12.000 | 1.32 | 0.2836 |
| Geometría de la herramienta | 1 | 675.000 | 675.000 | 74.31 | < 0.0001 |
| Interacción | 1 | 192.000 | 192.000 | 21.14 | 0.0018 |
| Error | 8 | 72.667 | 9.083 | | |
| Total | 11 | 951.667 | | | |

- b) El efecto de la interacción oculta el efecto de la velocidad de corte;
- c) $f_{\text{geometría de la herramienta}=1} = 16.51$ y valor $P = 0.0036$;
- d) $f_{\text{geometría de la herramienta}=2} = 5.94$ y valor $P = 0.0407$.

14.11 a)

| Fuente de variación | gl | Suma de cuadrados | Cuadrados medios | f | P |
|---------------------|----|-------------------|------------------|-------|---------|
| Método | 1 | 0.000104 | 0.000104 | 6.57 | 0.0226 |
| Laboratorio | 6 | 0.008058 | 0.001343 | 84.70 | <0.0001 |
| Interacción | 6 | 0.000198 | 0.000033 | 2.08 | 0.1215 |
| Error | 14 | 0.000222 | 0.000016 | | |
| Total | 27 | 0.008582 | | | |

- b) La interacción no es significativa;
 c) Ambos efectos principales son significativos;
 e) $f_{\text{laboratorio}=1} = 0.01576$ y valor $P = 0.9019$; no hay diferencia significativa entre los métodos en el laboratorio 1;

$f_{\text{geometría de la herramienta}=2} = 9.081$ y valor $P = 0.0093$.

14.13 b)

| Fuente de variación | gl | Suma de cuadrados | Cuadrados medios | f | P |
|---------------------|----|-------------------|------------------|--------|---------|
| Tiempo | 1 | 0.060208 | 0.060208 | 157.07 | <0.0001 |
| Tratamiento | 1 | 0.060208 | 0.060208 | 157.07 | <0.0001 |
| Interacción | 1 | 0.000008 | 0.000008 | .02 | 0.8864 |
| Error | 8 | 0.003067 | 0.000383 | | |
| Total | 11 | 0.123492 | | | |

- c) Tanto el tiempo como el tratamiento influyen significativamente en la absorción del magnesio, aunque no existe interacción significativa entre ambos.
 d) $Y = \mu + \beta_T \text{Tiempo} + \beta_Z Z + \beta_{TZ} \text{Tiempo} Z + \epsilon$, donde $Z = 1$ cuando el tratamiento = 1 y $Z = 0$ cuando el tratamiento = 2;
 e) $f = 0.02$ con valor $P = 0.8864$; la interacción en el modelo no es significativa.

14.15 a) La interacción es significativa al nivel de 0.05, con un valor P de 0.0166.

b) Ambos efectos principales son significativos.

14.17 a) $AB: f = 3.83$; significativo;

$AC: f = 3.79$; significativo;

$BC: f = 1.31$; no es significativo;

$ABC: f = 1.63$; no es significativo;

b) $A: f = 0.54$; no es significativo;

$B: f = 6.85$; significativo;

$C: f = 2.15$; no es significativo;

c) La presencia de la interacción AC enmascara el efecto principal C .

14.19 a) Esfuerzo cortante: $f = 45.96$ con valor $P < 0.0001$;

Recubrimiento: $f = 0.05$ con valor $P = 0.8299$;

Humedad: $f = 2.13$ con valor $P = 0.1257$;

recubrimiento \times humedad: $f = 3.41$ con valor $P = 0.0385$;

recubrimiento \times esfuerzo cortante: $f = 0.08$ con valor $P = 0.9277$;

humedad \times esfuerzo cortante: $f = 3.15$ con valor $P = 0.0192$;

recubrimiento \times humedad \times esfuerzo cortante: $f = 1.93$ con valor $P = 0.1138$.

b) La mejor combinación parece ser sin recubrimiento, humedad media y nivel de esfuerzo cortante de 20.

| Efecto | f | P |
|---------------------------|-------|----------|
| Temperatura | 14.22 | < 0.0001 |
| Superficie | 6.70 | 0.0020 |
| HRC | 1.67 | 0.1954 |
| T \times S | 5.50 | 0.0006 |
| T \times HRC | 2.69 | 0.0369 |
| S \times HRC | 5.41 | 0.0007 |
| T \times S \times HRC | 3.02 | 0.0051 |

14.23 a) Sí; marca \times tipo; marca \times temperatura;

b) sí;

c) marca Y , detergente en polvo, alta temperatura.

14.25 a)

| Efecto | f | P |
|-----------------------------------------------|--------|----------|
| Tiempo | 543.53 | < 0.0001 |
| Temperatura | 209.79 | < 0.0001 |
| Solvente | 4.97 | 0.0457 |
| Tiempo \times temperatura | 2.66 | 0.1103 |
| Tiempo \times solvente | 2.04 | 0.1723 |
| Temperatura \times solvente | 0.03 | 0.8558 |
| Tiempo \times temperatura \times solvente | 6.22 | 0.0140 |

Aunque las tres interacciones bilaterales son insignificantes podrían estar enmascaradas por la interacción trilateral significativa.

14.27 a) $f = 1.49$; no hay interacción significativa;

b) $f(\text{operadores}) = 12.45$; significativo;

$f(\text{filtros}) = 8.39$; significativo;

c) $\hat{\sigma}_f^2 = 0.1777$ (filtros);

$\hat{\sigma}_f^2 = 0.3516$ (operadores);

$s^2 = 0.185$

14.29 a) $\hat{\sigma}_{\beta_1}^2, \hat{\sigma}_{\beta_2}^2, \hat{\sigma}_{\alpha_7}^2$ son significativos;

b) $\hat{\sigma}_{\beta_2}^2$ y $\hat{\sigma}_{\alpha_7}^2$ son significativos

14.31 a) Modelo combinado;

- b) Material: $f = 47.42$ con valor $P < 0.0001$;
 marca: $f = 1.73$ con valor $P = 0.2875$;
 material \times marca: $f = 16.06$ con valor $P = 0.0004$;
 c) no

Capítulo 15

15.1 B y C son significativos al nivel 0.05

15.3 Los factores A , B y C tienen efectos negativos sobre el compuesto de fósforo y el factor D tiene un efecto positivo. Sin embargo, la interpretación del efecto de los factores individuales debería implicar el uso de las gráficas de interacción.

15.5 Efectos significativos:

$$A: f = 9.98; BC: f = 19.03.$$

Efectos insignificantes:

$$B: f = 0.20; C: f = 6.54; D: f = 0.02; AB: f = 1.83;$$

$$AC: f = 0.20; AD: f = 0.57; BD: f = 1.83;$$

$CD: f = 0.02$. Como la interacción BC es significativa, se investigaría más sobre B y sobre C .

15.9 a) $b_A = 5.5$, $b_B = -3.25$ y $b_{AB} = 2.5$;

b) Los valores de los coeficientes son de la mitad de los efectos;

$$c) t_A = 5.99 \text{ con valor } P = 0.0039;$$

$$t_B = -3.54 \text{ con valor } P = 0.0241;$$

$$t_{AB} = 2.72 \text{ con valor } P = 0.0529;$$

$$t_2 = F.$$

15.11 a) $A = -0.8750$, $B = 5.8750$, $C = 9.6250$,
 $AB = -3.3750$, $AC = -9.6250$, $BC = 0.1250$
 y $ABC = -1.1250$;

B , C , AB y AC parecen importantes con base en sus magnitudes.

| Efectos | Valor P |
|---------|-----------|
| A | 0.7528 |
| B | 0.0600 |
| C | 0.0071 |
| AB | 0.2440 |
| AC | 0.0071 |
| BC | 0.9640 |
| ABC | 0.6861 |

c) Sí;

d) A un nivel alto de A , C esencialmente no tiene efecto. A un nivel bajo de A , C tiene un efecto positivo.

15.13 a)

| Máquina | | | |
|---------|---------|-------|--------|
| 1 | 2 | 3 | 4 |
| (1) | c | a | ac |
| ab | d | b | ad |
| cd | e | acd | ae |
| ce | abc | ace | bc |
| de | abd | ade | bd |
| $abcd$ | abe | bcd | be |
| $abce$ | cde | bce | $acde$ |
| $abde$ | $abcde$ | bde | $bcde$ |

b) ABD , CDE , $ABCDE$ (un posible diseño)

15.15 a) x_2, x_3, x_1x_2 y x_1x_3 ;

b) Curvatura: valor $P = 0.0038$;

c) Un punto de diseño adicional diferente de los originales.

15.17 $(0, -1)$, $(0, 1)$, $(-1, 0)$, $(1, 0)$ podría utilizarse.

15.19 a) Con BCD como el contraste de definición, el bloque principal contiene (1) , a , bc , abc , bd , abd , cd , acd ;

b) Bloque 1 Bloque 2

| | |
|-------|-------|
| (1) | a |
| bc | abc |
| abd | bd |
| acd | cd |

confundido por ABC ;

c) El contraste de definición BCD produce los siguientes alias: $A \equiv ABCD$, $B \equiv CD$, $C \equiv BD$, $D \equiv BC$, $AB \equiv ACD$, $AC \equiv ABD$ y $AD \equiv ABC$. Puesto que AD y ABC están confundidos con los bloques sólo hay dos grados de libertad para el error en las interacciones no confundidas.

| Fuente de variación | Grado de libertad |
|---------------------|-------------------|
| A | 1 |
| B | 1 |
| C | 1 |
| D | 1 |
| Bloques | 1 |
| Error | 2 |
| Total | 7 |

15.21 a) Con el contraste de definición $ABCE$ y $ABDF$ el bloque principal contiene (1) , ab , acd , bcd , ce , $abce$, ade , bde , acf , bef , df , $abdf$, acf , bef , $cdef$, $abcdef$.

- b) $A \equiv BCE \equiv BDF \equiv ACDEF$,
 $AD \equiv BCDE \equiv BF \equiv ACEF$,
 $B \equiv ACE \equiv ADF \equiv BCDEF$,
 $AE \equiv BC \equiv BDEF \equiv ACDF$,
 $C \equiv ABE \equiv ABCDF \equiv DEF$,
 $AF \equiv BCEF \equiv BD \equiv ACDE$,
 $D \equiv ABCDE \equiv ABF \equiv CEF$,
 $CE \equiv AB \equiv ABCDEF \equiv DF$,
 $E \equiv ABC \equiv ABDEF \equiv CDF$,
 $DE \equiv ABCD \equiv ABEF \equiv CF$,
 $F \equiv ABCEF \equiv ABD \equiv CDE$,
 $BCD \equiv ADE \equiv ACF \equiv BEF$,
 $AB \equiv CE \equiv DF \equiv ABCDEF$,
 $BCF \equiv AEF \equiv ACD \equiv BDE$,
 $AC \equiv BE \equiv BCDF \equiv ADEF$;

| Fuente de variación | Grados de libertad |
|---------------------|--------------------|
| A | 1 |
| B | 1 |
| C | 1 |
| D | 1 |
| E | 1 |
| F | 1 |
| AB | 1 |
| AC | 1 |
| AD | 1 |
| BC | 1 |
| BD | 1 |
| CD | 1 |
| CF | 1 |
| Error | 2 |
| Total | 15 |

15.23

| Fuente | gl | SC | CM | f | P |
|--------|----|---------|--------|------|--------|
| A | 1 | 6.1250 | 6.1250 | 5.81 | 0.0949 |
| B | 1 | 0.6050 | 0.6050 | 0.57 | 0.5036 |
| C | 1 | 4.8050 | 4.8050 | 4.56 | 0.1223 |
| D | 1 | 0.2450 | 0.2450 | 0.23 | 0.6626 |
| Error | 3 | 3.1600 | 1.0533 | | |
| Total | 7 | 14.9400 | | | |

15.25

| Fuente | gl | SC | CM | f | P |
|--------|----|------------|------------|---------|--------|
| A | 1 | 388,129.00 | 388,129.00 | 3585.49 | 0.0001 |
| B | 1 | 277,202.25 | 277,202.25 | 2560.76 | 0.0001 |
| C | 1 | 4692.25 | 4692.25 | 43.35 | 0.0006 |
| D | 1 | 9702.25 | 9702.25 | 89.63 | 0.0001 |
| E | 1 | 1806.25 | 1806.25 | 16.69 | 0.0065 |
| AD | 1 | 1406.25 | 1406.25 | 12.99 | 0.0113 |
| AE | 1 | 462.25 | 462.25 | 4.27 | 0.0843 |
| BD | 1 | 1156.00 | 1156.00 | 10.68 | 0.0171 |
| BE | 1 | 961.00 | 961.00 | 8.88 | 0.0247 |
| Error | 6 | 649.50 | 108.25 | | |
| Total | 15 | 686,167.00 | | | |

Todos los efectos principales son significativos al nivel 0.05; AD , BD y BE son también significativos al nivel 0.05.

- 15.27 El bloque principal contiene af , be , cd , abd , ace , bcf , def , $abcdef$.

- 15.29 $A \equiv BD \equiv CE \equiv CDF \equiv BEF \equiv ABCF \equiv ADEF \equiv ABCDE$;
 $B \equiv AD \equiv CF \equiv CDE \equiv AEF \equiv ABCE \equiv BDEF \equiv ABCDF$;
 $C \equiv AE \equiv BF \equiv BDE \equiv ADF \equiv CDEF \equiv ABCD \equiv ABCEF$;
 $D \equiv AB \equiv EF \equiv BCE \equiv ACF \equiv BCDF \equiv ACDE \equiv ABDEF$;
 $E \equiv AC \equiv DF \equiv ABF \equiv BCD \equiv ABDE \equiv BCEF \equiv ACDEF$;
 $F \equiv BC \equiv DE \equiv ACD \equiv ABE \equiv ACEF \equiv ABDF \equiv BCDEF$.

- 15.31 $x_1 = 1$ y $x_2 = 1$

- 15.33 a) Sí;

b) i) $E(\hat{y}) = 79.00 + 5.281A$;

ii) $\text{Var}(\hat{y}) = 6.22^2 \sigma_z^2 + 5.70^2 A^2 \sigma_z^2 + 2(6.22)(5.70)A \sigma_z^2$;

c) velocidad a bajo nivel;

d) velocidad a bajo nivel;

e) sí

- 15.35 $\hat{y} = 12.7519 + 4.7194x_1 + 0.8656x_2 - 1.4156x_3$; las unidades están centradas y a escala; prueba de falta de ajuste, $F = 81.58$, con valor $P < 0.0001$.

- 15.37 AFG , BEG , CDG , DEF , $CEFG$, $BDFG$, $BCDE$, $ADEG$, $ACDF$, $ABEF$ y $ABCDEFG$.

Capítulo 16

- 16.1 $x = 7$ con valor $P = 0.1719$; no rechace H_0 .

- 16.3 $x = 3$ con valor $P = 0.0244$; rechace H_0 .

- 16.5 $x = 4$ con valor $P = 0.3770$; no rechace H_0 .

- 16.7 $x = 4$ con valor $P = 0.1335$; no rechace H_0 .

- 16.9 $w = 43$; no rechace H_0 .

- 16.11 $w_+ = 17.5$; no rechace H_0 .

- 16.13 $w_+ = 15$ con $n = 13$; rechace H_0 a favor de $\bar{\mu}_1 - \bar{\mu}_2 < 8$.

- 16.15 $u_1 = 4$; la afirmación no es válida
- 16.17 $u_2 = 5$; A opera durante más tiempo.
- 16.19 $u = 15$; no rechace H_0 .
- 16.21 $h = 10.58$; los tiempos de operación son diferentes.
- 16.23 $v = 7$ con valor $P = 0.910$; muestra aleatoria.
- 16.25 $v = 6$ con valor $P = 0.044$; no rechace H_0 .
- 16.27 $v = 4$; muestra aleatoria.
- 16.29 0.70
- 16.31 0.995
- 16.33 a) $r_s = 0.39$; b) no rechace H_0 .
- 16.35 a) $r_s = 0.72$; b) rechace H_0 , de manera que $\rho > 0$.

- 16.37 a) $r_s = 0.71$; b) rechace H_0 , de manera que $\rho > 0$.

Capítulo 18

- 18.1 $p^* = 0.173$
- 18.3 a) $\pi(p|x=1) = 40p(1-p)^3/0.2844$; $0.05 < p < 0.15$;
b) $p^* = 0.106$
- 18.5 a) $beta(95, 45)$; b) 1
- 18.7 $8.077 < \mu < 8.692$
- 18.9 a) 0.2509; b) $68.71 < \mu < 71.69$; c) 0.0174
- 18.13 $p^* = \frac{6}{x+2}$
- 18.15 2.21

Índice analítico

A

- Análisis de varianza (ANOVA), 254, 507
 - de dos factores, 565
 - de tres factores, 579
 - de un factor, 509
 - comparación de, 520
 - contraste de, 520
 - de un solo grado de libertad, 520
 - efecto del tratamiento, 510
 - media grande, 510
 - suma de cuadrados de los contrastes, 521
 - tratamiento, 509
 - tabla de, 415
- Aplicaciones bayesianas, 710
- Aproximación
 - de binomial a hipergeométrica, 155
 - de grados de libertad de Satterthwaite, 289
 - de normal a binomial, 187, 188
 - de Poisson a binomial, 163

B

- Bernoulli
 - ensayo de, 144
 - proceso de, 144
 - variable aleatoria, 83
- Bloques, 509

C

- Cadena Markov de Monte Carlo, 710
- Coficiente
 - de confianza, 269
 - de correlación, 125, 431
 - de la población, 432
 - de rangos, 675
 - muestral, 432
 - producto-momento de Pearson, 432

- de determinación, 407, 433, 462
 - ajustado, 464
 - de variación, 471

Combinación, 50

Complemento de un evento, 39

Confianza

- coeficiente de, 269
- grado de, 269
- límites, 269, 271

Contrastes ortogonales, 522

Control de calidad, 681

- dentro de control, 682
- fuera de control, 682
- gráfica, 681, 682
- límites del, 683

Corrección de continuidad, 190

Covarianza, 119, 123

Cuadrado(s)

- medio(s), 415
 - del error, 284
 - esperados, 548

Cuantiles, 255

Curva característica de operación, 335

D

Datos históricos, 30

Desviación, 120

- estándar, 120, 122, 135
- muestral, 15, 16

Diagrama(s)

- de árbol, 36
- de dispersión, 3
- de Venn, 40

Diseño

- central compuesto, 640
- completamente aleatorizado, 8, 509
- de bloques completos aleatorizado, 533

- de experimento
 - central compuesto, 640
 - completamente aleatorizado, 532
 - contraste en el, 599
 - de bloques, 532
 - de bloques aleatorizados, 533
 - factor de ruido, 644
 - factores de control, 644
 - factorial fraccionario, 598, 612, 626, 627
 - ortogonal, 617
 - relación de definición, 627
 - resolución, 637
- Distribución, 23
 - beta, 201
 - binomial, 104, 143-145, 153, 155, 175, 188
 - negativa, 143, 158-60
 - media de la, 147
 - varianza de la, 147
 - chi cuadrada, 200
 - condicional, 99
 - conjunta, 103
 - continua
 - beta, 201
 - chi cuadrada, 200
 - de Weibull, 203, 204
 - exponencial, 195
 - gamma, 195
 - logarítmica normal, 201
 - normal, 172
 - uniforme, 171
 - de Erlang, 206-207
 - de muestreo, 232
 - de la media, 233
 - de Poisson, 143, 161, 162
 - media de la, 162
 - varianza de la, 162
 - de probabilidad, 84
 - condicional, 99
 - conjunta, 94, 95, 102
 - continua, 87
 - discreta, 84
 - marginal, 97
 - media de la, 111
 - varianza de la, 119
 - de razón de varianza, 253
 - de Weibull, 203
 - función de distribución acumulativa
 - para, 204
 - media de la, 203
 - tasa de fallas de, 204, 205
 - varianza de la, 203
 - discreta
 - binomial, 143, 144, 158, 159
 - de Poisson, 161, 162
 - geométrica, 158, 160
 - hipergeométrica, 152, 153
 - multinomial, 143, 149
 - empírica, 254
 - exponencial, 104, 194, 195
 - media de la, 196
 - negativa, 196
 - propiedad de falta de memoria de, 197
 - relación con el proceso de Poisson, 196
 - varianza de la, 196
 - F, 251-254
 - gamma, 194-195
 - media de la, 196
 - relación con el proceso de Poisson, 196
 - varianza de la, 196
 - gaussiana, 19, 172
 - geométrica, 143, 158, 160
 - media de la, 160
 - varianza de la, 160
 - hipergeométrica, 152-154, 175
 - media de la, 154
 - multivariada, 156
 - varianza de la, 154
 - hipergeométrica multivariada, 156
 - logarítmica normal, 201
 - media de la, 202
 - varianza de la, 202
 - marginal, 97, 101, 102
 - conjunta, 103
 - multinomial, 143, 149
 - normal, 19, 172, 173, 188
 - bivariada, 431
 - curva normal, 172-175
 - desviación estándar de la, 175
 - estándar, 177
 - media de la, 175
 - varianza de la, 175
 - posterior, 711
 - previa, 710

- rectangular, 171
 - sesgada, 23
 - simétrica, 23
 - t , 246-250
 - uniforme, 171
 - continua, 171
- E**
- Ecuaciones normales para la regresión lineal, 444
 - Efecto de enmascaramiento, 563
 - Eliminación hacia atrás, 479
 - Error(es)
 - en la estimación de la media, 272
 - estándar de la media, 277
 - experimental, 509
 - suma de cuadrados del, 402
 - tipo I, 322
 - tipo II, 323
 - Espacio muestral, 35
 - continuo, 83
 - discreto, 83
 - partición del, 57
 - Esperanza matemática, 111, 112, 115
 - Estadística
 - descriptiva, 3, 9
 - inferencial, 1
 - Estadístico, 228
 - C_p , 491
 - de prueba, 322
 - Estimación, 12, 142, 266
 - de dos proporciones, 300
 - de la diferencia de dos medias muestrales, 285
 - de la probabilidad máxima, 307, 308, 312
 - de la proporción de varianzas, 305
 - de observaciones en pares, 291
 - de proporciones, 296
 - de una sola varianza, 303
 - Estimación de máxima verosimilitud, 307, 308, 710
 - residual, 550
 - restringida, 550
 - Estimado(s), 12
 - agrupado de la varianza, 287
 - bayesianos, 717
 - bajo la pérdida de error absoluto, 718
 - bajo la pérdida del cuadrado del error, 717
 - de una sola media, 269
 - del intervalo, 268
 - bayesiano, 715
 - puntual, 266, 268
 - error estándar, 276
 - Estimador, 266
 - de probabilidad máxima, 308-310
 - eficiente, 267
 - insesgado, 266, 267
 - método de momentos, 314, 315
 - puntual, 266, 268
 - Estudio
 - observacional, 3, 29
 - retrospectivo, 30
 - Evento(s), 38
 - mutuamente excluyentes, 40
 - Experimento
 - binomial negativo, 158
 - de efectos aleatorios
 - componentes de la varianza, 549
 - de efectos fijos, 547
 - de Poisson, 161
 - factorial, 561
 - ANOVA de los factores, 565
 - ANOVA de tres factores, 579
 - cuadrados medios agrupados, 583
 - efectos aleatorios, 589
 - efectos de enmascaramiento, 563
 - efectos principales, 562
 - en bloques, 583
 - factor, 507
 - interacción, 562
 - modelo mixto, 591
 - nivel, 507
 - tratamiento, 507
 - factorial 2^k , 597
 - ajuste de regresión, 612
 - alias, 628
 - corridas centrales, 620
 - diseño ortogonal, 617
 - diseños de Plackett-Burman, 638
 - factorial fraccionario, 626
 - filtrado de factores, 598
 - generación del diseño, 627
 - gráficas de diagnóstico, 604

relación definitoria, 627
 resolución, 637

F

Factor, 28, 507
 Factorial, 47
 Falta de ajuste, 418
 Frecuencia relativa, 22, 31, 111
 Función(es)
 de densidad de probabilidad, 88, 89
 conjunta, 96
 de distribución acumulativa, 85, 90
 de masa de probabilidad, 84
 conjunta, 95
 de pérdida
 del cuadrado del error, 717
 del error absoluto, 718
 de probabilidad, 84, 308
 gamma, 194
 incompleta, 199
 generadoras de momentos, 218

G

Grados de libertad, 15, 16, 200, 244, 246
 aproximación de Satterthwaite de, 289
 Gráfica(s)
 de caja, 3, 24, 25
 de control
 de atributos, 697
 de variables, 684
 gráfica cusum, 705
 p, 697
 R, 688
 S, 695
 U, 704
 \bar{X} , 686
 de cuantiles, 254, 255
 cuantiles normales, 256, 257
 de probabilidad, 254
 normal, 254
 de puntos, 3, 8, 32
 de tallo y hojas, 3, 21, 22, 31
 p, 697
 R, 688
 S, 695
 U, 704

\bar{X} , 686
 función característica de operación, 691

H

Hipótesis, 320
 alternativa, 320
 estadística, 319
 nula, 320
 prueba de, 320, 321
 Histograma(s), 22
 de probabilidad, 86

I

Independencia, 62, 65, 67, 68
 estadística, 101-103
 Inferencia
 bayesiana, 710
 estadística, 3, 225, 265
 Interacción, 28, 562
 Intersección de eventos, 39
 Intervalo
 bayesiano, 715
 posterior, 317
 de confianza, 269, 270, 281, 317
 de una muestra grande, 276
 interpretación de, 289
 para el cociente de las desviaciones
 estándar, 306
 para el cociente de las varianzas, 306
 para la desviación estándar, 304
 para la diferencia de dos medias,
 285-288, 290
 para la diferencia de dos proporciones,
 300, 301
 para observaciones en pares, 293
 para una sola media, 269-272, 275
 unilateral, 273
 para una sola proporción, 297
 para una sola varianza, 304
 de predicción, 277, 278, 281
 para una observación futura, 278, 279
 unilateral, 279
 de tolerancia, 280, 281
 posterior bayesiano, 317

J

Jacobiano, de la transformación, 213

L**Límite(s)**

- de confianza unilateral, 273
- de tolerancia, 280
 - del método no paramétrico, 674
 - unilaterales, 281

M

Media, 19, 111, 112, 114, 115

- muestral, 111
- poblacional, 12, 16
- recortada, 12

Método(s)

- de distribución libre, 655
- de la regla, 37
- de mínimos cuadrados, 394, 396
- no paramétricos, 655
 - límites de tolerancia, 674
 - prueba de Kruskal-Wallis, 668
 - prueba de la suma de rangos de Wilcoxon, 665
 - prueba de rachas, 671
 - prueba de rango con signo, 660
 - prueba del signo, 656

Metodología bayesiana, 265, 709

Metodología de respuesta superficial, 447, 639, 640

factor(es)

- de control, 644
- de ruido, 644

modelo de segundo orden, 640

Moda, 713

distribución normal, 174

Modelo

- de efectos aleatorios, 547, 548
- lineal, 133

Momentos, 218

Muestra, 1, 2, 225, 226

- aleatoria, 227
 - simple, 7

desviación estándar de la, 3, 15, 16, 30, 31, 229, 230

media de la, 3, 11, 12, 19, 30-32, 225, 228

mediana de la, 3, 11, 12, 30, 31, 228

moda de la, 228

rango de la, 15, 30, 31, 229

sesgada, 7

tamaño de la, 7

varianza de la, 15, 16, 30, 225, 229

Muestreo

aleatorio, 225

de aceptación, 153

Multicolinealidad, 476

N**Nivel**

de calidad aceptable, 705

de calidad rechazable, 705

de significancia, 323, 332

O

Observaciones en pares, 291

P

Parámetro(s), 12, 142

de distribución, 104

de la población, 16, 104

Permutación, 47

circular, 49

Perspectiva

bayesiana, 710

condicional, 710

Población, 2, 4, 225, 226

media de la, 226

parámetro de la, 16, 104

tamaño de la, 226

varianza de la, 226

Potencia de una prueba, 329

Predictor lineal, 498

Probabilidad, 35, 52, 53

condicional, 62-66, 68, 75, 76

de cobertura, 715

de un evento, 52

frecuencia relativa, 55, 709

función de masa, 84

indiferencia, 55, 709

método subjetivo, 709, 710

regla aditiva, 56

- subjetiva, 709, 710
 - total, 72, 73
- Proceso de Poisson, 161, 196
 - relación con la distribución gamma, 196
- Promedio, 111
- Prueba(s)
 - de Bartlett, 516
 - de bondad de ajuste, 210, 255, 317, 370, 371
 - de Cochran, 518
 - de comparación múltiple, 523
 - de Duncan, 527
 - de Dunnett, 528
 - de Tukey, 526
 - tasa de error por experimento-familia, 525
 - de Duncan de rango múltiple, 527
 - de Dunnett, 528
 - de hipótesis, 19, 266, 319
 - bondad de ajuste, 210, 255, 370, 371
 - de dos colas, 330
 - de una cola, 330
 - dos medias con varianza desconocidas pero iguales, 343
 - dos medias con varianzas desconocidas y desiguales, 345
 - dos varianzas, 366
 - estadísticos de prueba, 326
 - F* parcial, 466
 - observaciones en pares, 345
 - propiedades importantes, 329
 - prueba de homogeneidad, 376
 - prueba de independencia, 373
 - prueba de varias proporciones, 377
 - región crítica, 322
 - selección del tamaño de la muestra, 349, 352
 - sobre dos medias, 342
 - tamaño de prueba, 323
 - una sola muestra, 336
 - una sola muestra, varianza conocida, 336
 - una sola muestra, varianza desconocida, 340
 - una sola proporción, 360
 - una sola varianza, 366
 - valor crítico, 322
 - valor *P*, 331, 333
 - de Kruskal-Wallis, 668
 - de la suma de rangos de Wilcoxon, 665
 - de rachas, 671
 - de rango con signo, 660
 - de Tukey, 526
 - del signo, 656
 - F* parcial, 466
 - para la igualdad de varianzas, 516
 - de Bartlett, 516
 - de Cochran, 518
 - sobre una sola proporción, 360
- Puntos de inflexión, distribución normal, 174
- R**
- R^2 , 407, 462
 - ajustada, 464
- Rango intercuartilar, 24, 25
- Regla
 - aditiva, 56
 - de Bayes, 72, 75
 - de eliminación, 73-75
 - de la multiplicación, 44
 - del producto, 65
 - multiplicativa, 65
- Regresión, 20
- Regresión lineal
 - a través del origen, 413
 - ANOVA, 414
 - coeficiente de determinación, 407
 - coeficiente de regresión, 392
 - correlación, 430
 - cuadrados medios, 394
 - ecuación normal, 396
 - error aleatorio, 391
 - error experimental puro, 419
 - falta de ajuste de la, 418
 - intervalo de predicción, 410, 411
 - línea ajustada, 392
 - modelo
 - empírico, 391
 - estadístico, 391
 - múltiple, 390, 443
 - ANOVA, 455
 - ecuaciones normales, 444
 - inferencia, 455
 - matriz de varianza-covarianza, 453
 - matriz TESTADA, 483
 - multilinealidad, 476

- polinomial, 446
 - R^2 ajustada, 464
 - residuales estudentizados, 483
 - residuales R de Student, 483
 - selección de variables, 456
 - suma de cuadrados de regresión, 460
 - suma de cuadrados del error, 460
 - valor extremo, 484
 - variables ortogonales, 467
 - predicción, 408
 - prueba de linealidad, 416
 - regresor, 389
 - residual, 395
 - respuesta media, 394, 409
 - selección del modelo, 476, 487
 - simple, 389, 390
 - sobreajuste, 408
 - suma
 - de cuadrados de la regresión, 461
 - de cuadrados del error, 415
 - de cuadrados total, 414
 - transformación de datos, 424
 - valor ajustado, 416
 - variable
 - categoría, 472
 - dependiente, 389
 - independiente, 389
 - Regresión logística, 497
 - dosis eficaz, 500
 - razón de probabilidad, 500
 - Regresión no lineal, 496
 - datos de conteo, 497
 - logística, 497
 - respuesta binaria, 497
 - Regresión polinomial, 443, 446
 - Regresión por etapas, 479
 - Residual, 395, 427
- S**
- Selección
- del modelo, 476
 - eliminación hacia atrás, 480
 - estadístico C_p , 491
 - métodos secuenciales, 476
 - PRESS, 487, 488
 - regresión por etapas, 480
 - selección hacia adelante, 479
 - hacia adelante, 479
 - Sesgo, 227
 - Suma de cuadrados
 - de predicción, 487, 488
 - del error, 402, 415
 - falta de ajuste, 419
 - identidad, 510, 536, 567
 - regresión, 415
 - total, 407
 - tratamiento, 511, 522, 536
 - Superficie de respuesta, 642, 648
 - diseño de parámetro robusto, 644
- T**
- Tabla de contingencia, 373
 - frecuencia marginal, 374
 - Tamaño de la muestra, 7
 - en la estimación
 - de una media, 272
 - de una proporción, 298
 - en la prueba de hipótesis, 351
 - Tasa
 - de error por experimento-familia, 525
 - de fallas, 204, 205
 - Teorema
 - de Chebyshev, 135-137, 148, 155, 180, 186
 - del límite central, 233, 234, 238
 - Transformación de variables
 - continuas, 213, 214
 - discretas, 212
 - Tratamiento
 - efecto negativo del, 563
 - efecto positivo del, 563
- U**
- Unidad experimental, 9, 286, 292, 562
 - Unión de eventos, 40
- V**
- Validación cruzada, 487
 - Valor(es)
 - esperado, 112-115
 - extremo, 24, 279, 484

- P*, 4, 109, 331-333
- Variabilidad, 8, 9, 14-16, 119, 135, 228, 251, 253
 - entre/dentro de muestras, 253, 254
- Variable
 - aleatoria, 81
 - binomial, 144, 147, 158
 - continua, 84
 - chi cuadrada, 244
 - de Bernoulli, 83, 147
 - de Poisson, 161, 162
 - discreta, 83, 84
 - función no lineal de la, 133
 - hipergeométrica, 143, 153
 - media de la, 111, 114
 - multinomial, 149
 - normal, 173
 - transformación, 211
 - uniforme continua, 171
 - uniforme discreta, 150
 - varianza de la, 119, 122
 - categorica, 472
 - ficticia, 472
 - indicadora, 472
 - ortogonales, 467
- Varianza, 119, 120, 122
 - muestral, 16
 - agrupada, 287
 - poblacional, 16



ZAW



BIBLIOTECA UTN



035372

Al elaborar la novena edición de esta obra, el interés principal de los autores no fue tan sólo incluir material nuevo, sino brindar claridad y una mejor exposición, así como conservar el equilibrio entre la teoría y las aplicaciones.

Con la finalidad de motivar al estudiante, muchos ejercicios se refieren a aplicaciones científicas y de ingeniería en la vida real. En varios capítulos se agregaron proyectos para la clase y más estudios de caso, con el fin de ayudar a los usuarios a entender los métodos estadísticos que se presentan en el contexto de una situación cotidiana. Para lograr que los estudiantes adquieran experiencia en la lectura e interpretación de listas de resultados y gráficas por computadora, los estudios de caso muestran impresiones de listas de resultados por computadora y material gráfico generado con los programas SAS y MINITAB. En algunas situaciones, los ejemplos y los estudios de caso se complementan con diversos tipos de gráficas residuales, cuantilares, de probabilidad normal y de otros tipos.

Todos los conjuntos de datos asociados con los ejercicios
están disponibles para descargar del sitio Web
<http://www.pearsonenespañol.com/walpole>

BIBLIOTECA UTN



035372

ISBN 978 607-32-1417-9



90000

9 786073 214179

Visítenos en:
www.pearsonenespañol.com