

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

$$\begin{aligned} SCA &= bcn \sum_{i=1}^a (\bar{y}_{i...} - \bar{y}_{....})^2 & SC(AB) &= cn \sum_i \sum_j (\bar{y}_{ij..} - \bar{y}_{i...} - \bar{y}_{.j.} + \bar{y}_{....})^2 \\ SCB &= acn \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{....})^2 & SC(AC) &= bn \sum_i \sum_k (\bar{y}_{i.k.} - \bar{y}_{i...} - \bar{y}_{..k.} + \bar{y}_{....})^2 \\ SCC &= abn \sum_{k=1}^c (\bar{y}_{..k.} - \bar{y}_{....})^2 & SC(BC) &= an \sum_j \sum_k (\bar{y}_{.jk.} - \bar{y}_{.j.} - \bar{y}_{..k.} + \bar{y}_{....})^2 \\ SC(ABC) &= n \sum_i \sum_j \sum_k (\bar{y}_{ijk.} - \bar{y}_{ij..} - \bar{y}_{i.k.} - \bar{y}_{.jk.} + \bar{y}_{i...} + \bar{y}_{.j.} + \bar{y}_{..k.} - \bar{y}_{....})^2 \\ STC &= \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{....})^2 & SCE &= \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{ijk.})^2 \end{aligned}$$

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	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:				
AB	$SC(AB)$	$(a - 1)(b - 1)$	s_4^2	$f_4 = \frac{s_4^2}{s^2}$
AC	$SC(AC)$	$(a - 1)(c - 1)$	s_5^2	$f_5 = \frac{s_5^2}{s^2}$
BC	$SC(BC)$	$(b - 1)(c - 1)$	s_6^2	$f_6 = \frac{s_6^2}{s^2}$
Interacción de 3 factores:				
ABC	$SC(ABC)$	$(a - 1)(b - 1)(c - 1)$	s_7^2	$f_7 = \frac{s_7^2}{s^2}$
Error	SCE	$abc(n - 1)$	s^2	
Total	STC	$abcn - 1$		

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	<i>f</i> calculada	Valor <i>P</i>
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.

a) Elabore la tabla ANOVA completa. Agrupe todas las “interacciones” con los bloques dentro del error.

b) ¿Cuáles interacciones parecen ser significativas?

c) 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.

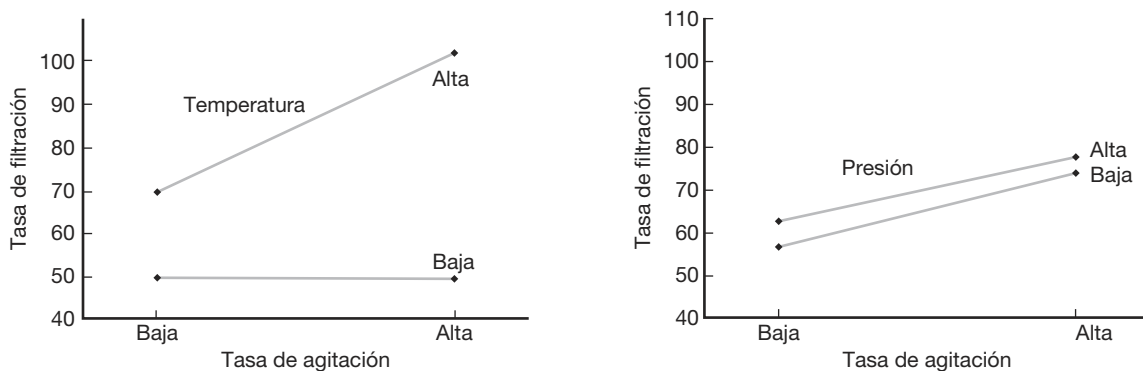


Figura 14.8: Gráficas de interacción para el ejemplo 14.5.

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 $y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijkl}$. 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

- a) Haga una prueba de significancia sobre la interacción BC al nivel $\alpha = 0.05$.
- b) 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

- a) Haga pruebas de significancia sobre todas las interacciones a un nivel $\alpha = 0.05$.
- b) Realice pruebas de significancia sobre los efectos principales a un nivel $\alpha = 0.05$.
- c) 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 Ordinance 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 amoníaco 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

- a) Realice un análisis de varianza con $\alpha = 0.01$ para probar la existencia de efectos principales y de interacción significativos.
- b) Analice la influencia de los 3 factores sobre el porcentaje de peso del perclorato de amoníaco. 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 cortante, ambos en 3 niveles. A continuación se presentan los datos de fatiga expresados en miles de ciclos antes de la falla.

- a) Realice un análisis de varianza con $\alpha = 0.01$ para probar la existencia de efectos principales y de interacción significativos.
- b) 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
		3549	1017	937
	Medio (50–60% RH)	2858	799	314
		8829	3471	244
		10,914	685	261
		4067	810	522
		2595	3409	739
	Alto (86–91% RH)	6489	1862	1344
		5248	2710	1027
		6816	2632	663
		5860	2131	1216
		5901	2470	1097
Cromado	Bajo (20–25% RH)	5395	4035	130
		2768	2022	841
		1821	914	1595
		3604	2036	1482
		4106	3524	529
	Medio (50–60% RH)	4833	1847	252
		7414	1684	105
		10,022	3042	847
		7463	4482	874
		21,906	996	755
	Alto (86–91% RH)	3287	1319	586
		5200	929	402
		5493	1263	846
		4145	2236	524
		3336	1392	751

14.20 Para un estudio de la dureza de los empastes dentales de oro se eligieron 5 dentistas al azar y se asignaron a combinaciones de 3 métodos de condensación y 2 tipos de oro. Se midió la dureza. (Véase Hoaglin, Mosteller y Tukey, 1991). Permita que los dentistas desempeñen el papel de bloques. Los datos se presentan a continuación.

- Proponga el modelo adecuado con las suposiciones.
- ¿Hay una interacción significativa entre el método de condensación y el tipo de material de empaste de oro?
- ¿Hay un método de condensación que parezca mejor? Explique su respuesta.

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.)

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 copiadoras 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.

	Estado superficial de la adhesión en el rodillo	Dureza del rodillo de la prensa					
		20		40		60	
Temp. 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
Temp. 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
Temp. 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.

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
			78	75	72
	Líquido	Caliente	75	75	73
		Tibia			
Y	En polvo	Caliente	90	92	92
		Tibia	88	86	88
			78	76	70
	Líquido	Caliente	76	77	76
		Tibia			
Z	En polvo	Caliente	85	87	88
		Tibia	76	74	78
			60	70	68
	Líquido	Caliente	55	57	54
		Tibia			

- a) ¿Existen efectos de la interacción significativos a un nivel $\alpha = 0.05$?
- b) ¿Hay diferencias significativas entre las tres marcas de detergente?
- c) ¿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		Temp. del molde	
	220	250	220	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

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.

- a) 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.
- b) 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.
- c) 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	

- a) Haga un análisis de varianza y determine cuáles factores e interacciones influyen en la proporción de gel.
- b) 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.
- c) Haga una gráfica de probabilidad normal de los residuales y comente.

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 $A_i, B_j, (AB)_{ij}$ y ϵ_{ijk} son variables aleatorias independientes con medias igual a 0 y varianzas $\sigma^2_\alpha, \sigma^2_\beta, \sigma^2_{\alpha\beta}$ y σ^2 , respectivamente. Las sumas de cuadrados para experimentos de efectos aleatorios se calculan

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

$$\begin{aligned} H_0: \sigma_{\alpha}^2 &= 0, \\ H_1: \sigma_{\alpha}^2 &\neq 0, \end{aligned}$$

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_1^2/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 (total) = 84.5564,
 SCA (operadores) = 18.2106,
 $SC(AB)$ (interacción) = 5.5161.

SCE (error) = 10.6733,
 SCB (lotes) = 50.1564,

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

$$\begin{aligned} H'_0: \sigma_\alpha^2 &= 0, & H''_0: B_1 = B_2 = \dots = B_b &= 0, & H'''_0: \sigma_{\alpha\beta}^2 &= 0, \\ H'_1: \sigma_\alpha^2 &\neq 0, & H''_1: \text{Al menos una de las } B_j &\text{no es igual a 0}, & H'''_1: \sigma_{\alpha\beta}^2 &\neq 0. \end{aligned}$$

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
<i>A</i>	$\sigma^2 + bc n \sigma_{\alpha}^2$	$\sigma^2 + cn \sigma_{\alpha\beta}^2 + bc n \sigma_{\alpha}^2$
<i>B</i>	$\sigma^2 + cn \sigma_{\alpha\beta}^2 + acn \sum_{j=1}^b \frac{B_j^2}{b-1}$	$\sigma^2 + cn \sigma_{\alpha\beta}^2 + acn \sigma_{\beta}^2$
<i>C</i>	$\sigma^2 + bn \sigma_{\alpha\gamma}^2 + abn \sum_{k=1}^c \frac{C_k^2}{c-1}$	$\sigma^2 + n \sigma_{\alpha\beta\gamma}^2 + an \sigma_{\beta\gamma}^2 + bn \sigma_{\alpha\gamma}^2 + abn \sum_{k=1}^c \frac{C_k^2}{c-1}$
<i>AB</i>	$\sigma^2 + cn \sigma_{\alpha\beta}^2$	$\sigma^2 + cn \sigma_{\alpha\beta}^2$
<i>AC</i>	$\sigma^2 + bn \sigma_{\alpha\gamma}^2$	$\sigma^2 + n \sigma_{\alpha\beta\gamma}^2 + bn \sigma_{\alpha\gamma}^2$
<i>BC</i>	$\sigma^2 + n \sigma_{\alpha\beta\gamma}^2 + an \sum_j \sum_k \frac{(BC)_{jk}^2}{(b-1)(c-1)}$	$\sigma^2 + n \sigma_{\alpha\beta\gamma}^2 + an \sigma_{\beta\gamma}^2$
<i>AB</i>	$\sigma^2 + n \sigma_{\alpha\beta\gamma}^2$	$\sigma^2 + n \sigma_{\alpha\beta\gamma}^2$
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.

- Escriba un modelo adecuado, con suposiciones.
- Utilice análisis de varianza para probar las hipótesis apropiadas para los inspectores, el nivel de inspección y la interacción.

Inspector	Nivel de inspección					
	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,

- utilice un estimado agrupado del error cuando esto sea apropiado;
- 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.

- Escriba el modelo apropiado.
- Evalúe los componentes de la varianza para el operador y la hora.
- 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

- ¿Cómo se le llama a este tipo de modelo?
- Analice los datos usando el modelo apropiado.
- ¿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

- Proporcione el modelo adecuado y liste las suposiciones que se hacen.
- ¿Hay un efecto significativo debido al ajuste de potencia?
- ¿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	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 aná-

lisis 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 mejorías 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 2^2

		A	Media
B	$\left\{ \begin{array}{l} b \\ (1) \end{array} \right.$	ab	$\frac{b+ab}{2n}$
		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 2^2 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 **contrastes**. 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

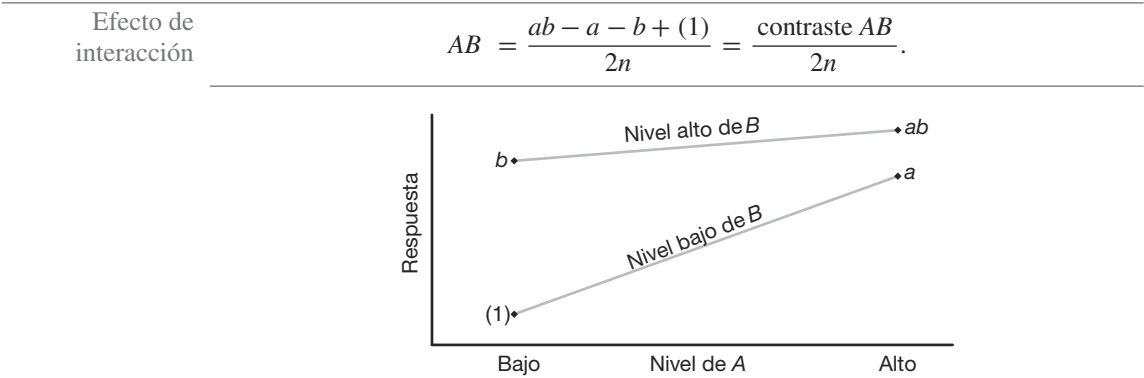


Figura 15.1: Respuesta que sugiere una interacción aparente.

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.

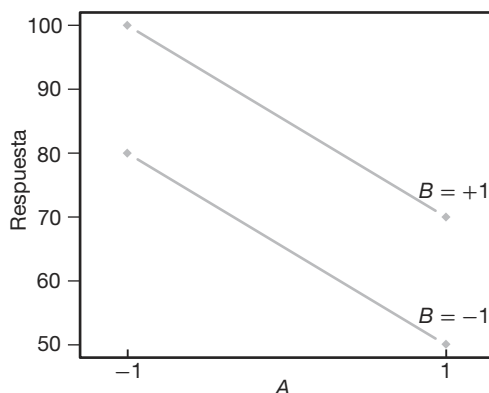


Figura 15.2: Gráfica de interacción para los datos de la tabla 15.2.

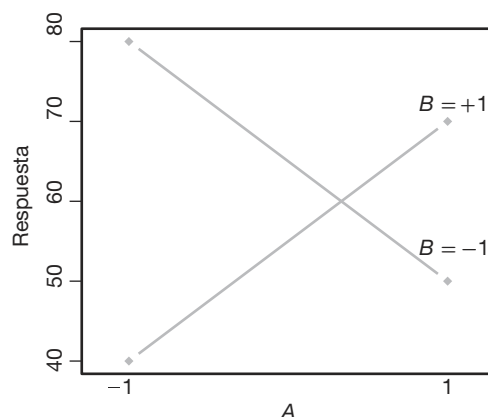


Figura 15.3: Gráfica de interacción para los datos de la tabla 15.3.

Cálculo de las sumas de cuadrados

Se aprovecha el hecho de que en el factorial 2^2 , 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 2^2

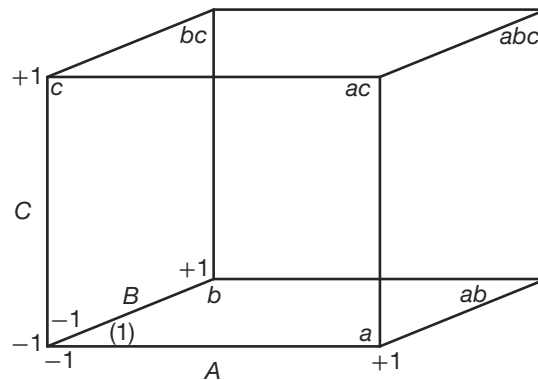
Combinación de tratamientos	Efecto factorial		
	A	B	AB
(1)	–	–	+
a	+	–	–
b	–	+	–
ab	+	+	+

El factorial 2^3

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 2^3 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 2^3

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 2^3 .

Es útil analizar e ilustrar la geometría del factorial 2^3 del mismo modo que se hizo para el factorial 2^2 en la figura 15.1. Para el 2^3 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 2^2 . 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 2^3 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 2^4 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 2^4 sin réplica.

Tabla 15.6: Datos para el estudio de caso 15.1

Combinación de factores	Respuesta (cm $\times 10^4$)	Combinación de factores	Respuesta (cm $\times 10^4$)
(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.

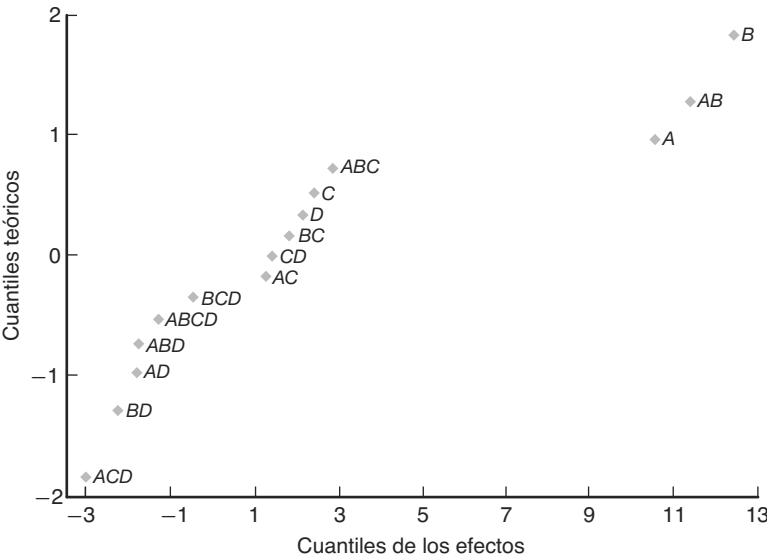


Figura 15.5: Gráfica cuantil-cuantil normal de los efectos para el estudio de caso del ejemplo 15.1.

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 esperaba, 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.

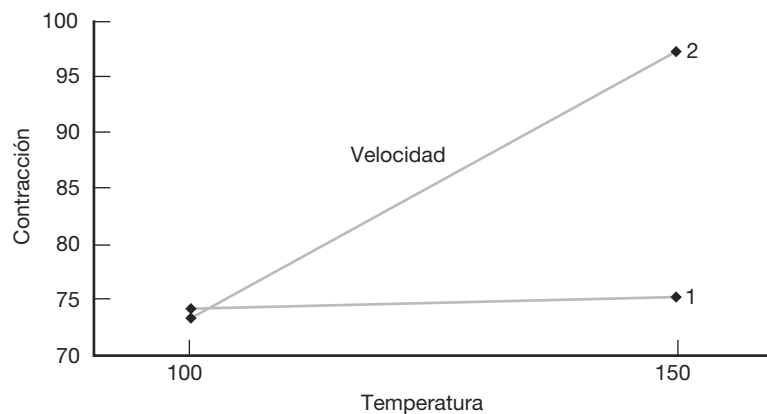


Figura 15.6: Gráfica de la interacción para el estudio de caso 15.1.

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						
		Sum of				
Source	DF	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	
Standard						
Parameter	Estimate	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	Réplica 2	Réplica 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 de fósforo		
	Réplica 1	Réplica 2	Total
(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	<i>f</i> calculada	Valor <i>P</i>
Efecto principal:						
<i>A</i>	−1.2000	11.52	1	11.52	4.68	0.0459
<i>B</i>	−1.2250	12.01	1	12.01	4.88	0.0421
<i>C</i>	−2.2250	39.61	1	39.61	16.10	0.0010
<i>D</i>	1.4875	17.70	1	17.70	7.20	0.0163
Interacción de dos factores:						
<i>AB</i>	0.9875	7.80	1	7.80	3.17	0.0939
<i>AC</i>	0.6125	3.00	1	3.00	1.22	0.2857
<i>AD</i>	−1.3250	14.05	1	14.05	5.71	0.0295
<i>BC</i>	1.1875	11.28	1	11.28	4.59	0.0480
<i>BD</i>	0.6250	3.13	1	3.13	1.27	0.2763
<i>CD</i>	0.7000	3.92	1	3.92	1.59	0.2249
Interacción de tres factores:						
<i>ABC</i>	−0.5500	2.42	1	2.42	0.98	0.3360
<i>ABD</i>	1.7375	24.15	1	24.15	9.82	0.0064
<i>ACD</i>	1.4875	17.70	1	17.70	7.20	0.0163
<i>BCD</i>	−0.8625	5.95	1	5.95	2.42	0.1394
Interacción de cuatro factores:						
<i>ABCD</i>	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
(1)	7.9	9.6
<i>a</i>	9.1	10.2
<i>b</i>	8.6	5.8
<i>c</i>	10.4	12.0
<i>d</i>	7.1	8.3
<i>ab</i>	11.1	12.3
<i>ac</i>	16.4	15.5
<i>ad</i>	7.1	8.7
<i>bc</i>	12.6	15.2
<i>bd</i>	4.7	5.8
<i>cd</i>	7.4	10.9
<i>abc</i>	21.9	21.9
<i>abd</i>	9.8	7.8
<i>acd</i>	13.8	11.2
<i>bcd</i>	10.2	11.1
<i>abcd</i>	12.8	14.3

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.

B	H	132	149
		137	152
	L	145	154
		147	150
		L	H
		A	

- Suponga un modelo que contiene dos efectos principales y una interacción y calcule los tres efectos. ¿Tiene usted alguna interpretación en este momento?
- Realice un análisis de varianza y haga pruebas de interacción. Saque conclusiones.
- 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 interprételas.

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).

- Calcule los coeficientes de regresión.
- ¿Cómo se relacionan los coeficientes b_1 , b_2 y b_{12} con los efectos que encontró en el ejercicio 15.6a)?
- 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.

- ¿Parece válida la suposición de que las interacciones de tres y cuatro factores son despreciables?
- ¿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

- Calcule los siete efectos. Con base en su magnitud, ¿cuál parece ser importante?
- Haga un análisis de varianza y observe los valores *P*.
- ¿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
<i>a</i>	43
<i>b</i>	35
<i>c</i>	44
<i>abc</i>	39

Con sólo estas corridas los signos para los contrastes son dados por

	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
<i>a</i>	+	−	−	−	−	+	+
<i>b</i>	−	+	−	−	+	−	+
<i>c</i>	−	−	+	+	−	−	+
<i>abc</i>	+	+	+	+	+	+	+

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

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}.$$

La matriz **X** se denomina **matriz del modelo**. Suponga, por ejemplo, que se utiliza un experimento factorial 2^3 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 \mathbf{X} se vuelve

$$\mathbf{X} = \begin{array}{cccc} & x_1 & x_2 & x_3 & \text{Identificación del diseño} \\ \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} & & & \begin{array}{l} (1) \\ a \\ b \\ c \\ ab \\ ac \\ bc \\ abc \end{array} \end{array}$$

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 \mathbf{X} en el ejemplo 2^3 son *ortogonales*. Como resultado, el cálculo de los coeficientes de regresión que se describió en la sección 12.3 se convierte en

$$\begin{aligned} b &= \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} = \left(\frac{1}{8}\mathbf{I}\right) \mathbf{X}'\mathbf{y} \\ &= \frac{1}{8} \begin{bmatrix} a + ab + ac + abc + (1) + b + c + bc \\ a + ab + ac + abc - (1) - b - c - bc \\ b + ab + bc + abc - (1) - a - c - ac \\ c + ac + bc + abc - (1) - a - b - ab \end{bmatrix}, \end{aligned}$$

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:

$$\begin{aligned} \text{Efecto} &= \frac{\text{contraste}}{2^{k-1}(n)} \\ \text{Coeficiente de regresión} &= \frac{\text{contraste}}{2^k(n)} = \frac{\text{efecto}}{2}. \end{aligned}$$

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

$$\begin{bmatrix} 1 & x_1 & x_2 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

con los coeficientes de regresión

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} \frac{(1) + a + b + ab}{4} \\ \frac{a + ab - (1) - b}{4} \\ \frac{b + ab - (1) - a}{4} \end{bmatrix} = \begin{bmatrix} 36.25 \\ 6.25 \\ 2.75 \end{bmatrix}.$$

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 2^2 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									
		Parameter	Standard						
Variable	DF	Estimate	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 $x_1x_2x_3$ (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 \mathbf{X} . Considere la matriz \mathbf{X} para el factorial 2^2 del ejemplo 15.2. Observe que las tres columnas son mutuamente ortogonales. La matriz \mathbf{X} del factorial 2^3 también contiene columnas ortogonales. El factorial 2^3 con interacciones produciría una matriz \mathbf{X} del tipo

$$\mathbf{X} = \begin{bmatrix} & x_1 & x_2 & x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 & x_1 x_2 x_3 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

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

$$\sigma^2 A^{-1} = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}.$$

En el caso del experimento factorial 2^k las columnas de \mathbf{X} son mutuamente ortogonales,

lo que impone una estructura muy especial. En general, para 2^k se puede escribir

$$\mathbf{X} = [\mathbf{1} \quad \pm \mathbf{1} \quad \pm \mathbf{1} \quad \cdots \quad \pm \mathbf{1} \quad \pm \mathbf{1} \quad \cdots],$$

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:

$$\mathbf{X} = \begin{matrix} & x_1 & x_2 & x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 \\ \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

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^k 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

$$t^2 = \frac{(\text{contraste})^2}{s^2 2^k n} = f.$$

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 requerido 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.

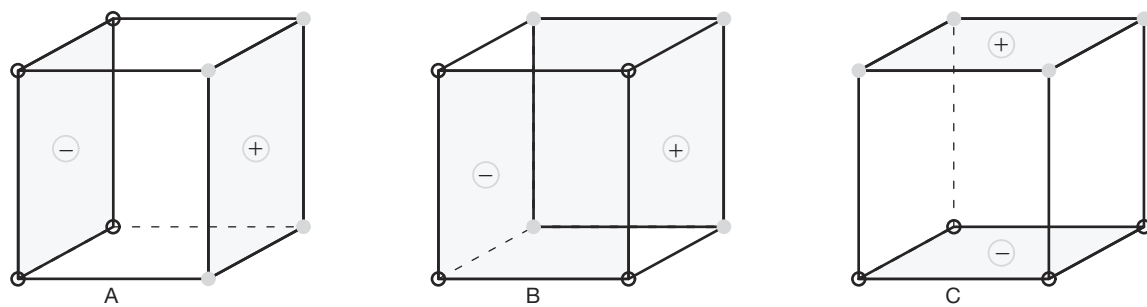
Condiciones para varianzas mínimas de los coeficientes	Si el modelo de regresión contiene términos no mayores de primer orden, y si los rangos de las variables son dados por $x_j \in [-1, +1]$ para $j = 1, 2, \dots, k$, entonces $\text{Var}(b_j)/\sigma^2$, para $j = 1, 2, \dots, k$, se minimiza si el 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.

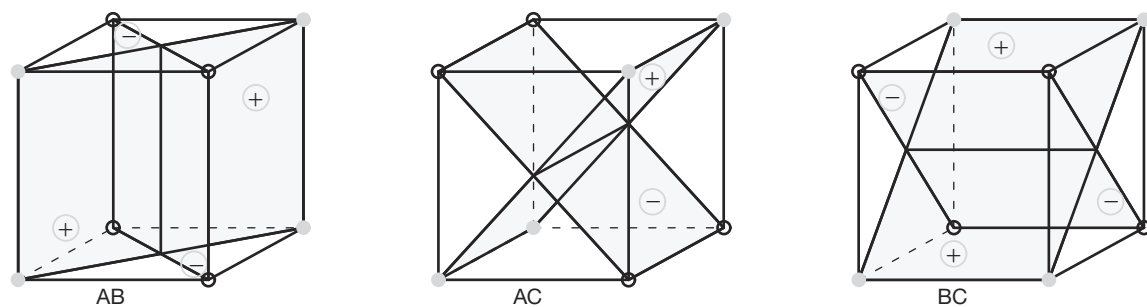
Otro método para lograr una mejor comprensión del “balance” proporcionado por el factorial 2^3 consiste en observar la situación mediante una gráfica. En la figura 15.10 se aprecia cada uno de los contrastes ortogonales y, por lo tanto, mutuamente independientes. En las gráficas se comparan los planos 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

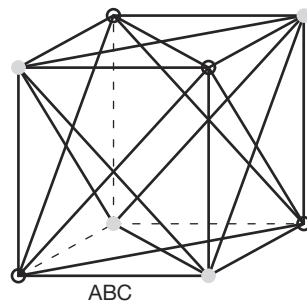
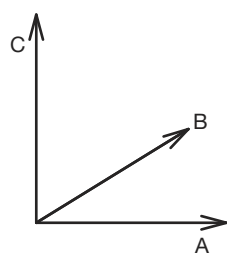
En la situación en que se aplica el diseño 2^k con variables **continuas** de diseño y se busca ajustar un modelo de regresión lineal, el uso de réplicas de corridas en el **diseño central** puede ser sumamente útil. De hecho, además de las ventajas que se analizarán a continuación, la mayoría de los científicos e ingenieros considerarían que las corridas



(a) Efectos principales



(b) Interacción de dos factores



(c) Interacción de tres factores

● = + corridas
○ = - 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.

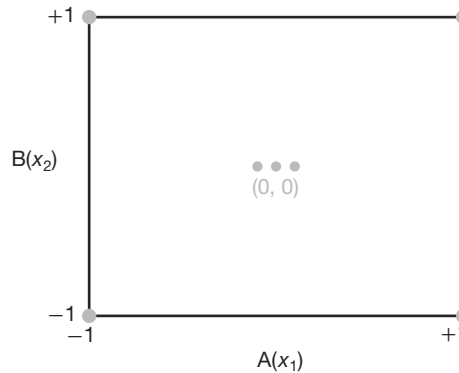


Figura 15.11: Un diseño 2^2 con corridas centrales.

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

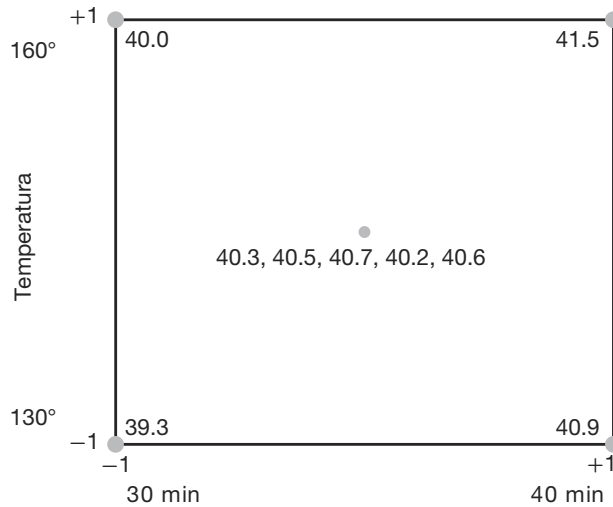
$$t_{n_c-1} = \frac{\bar{y}_f - \bar{y}_0}{s_{\bar{y}_f - \bar{y}_0}} = \frac{\bar{y}_f - \bar{y}_0}{\sqrt{CME(1/n_f + 1/n_c)}},$$

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.

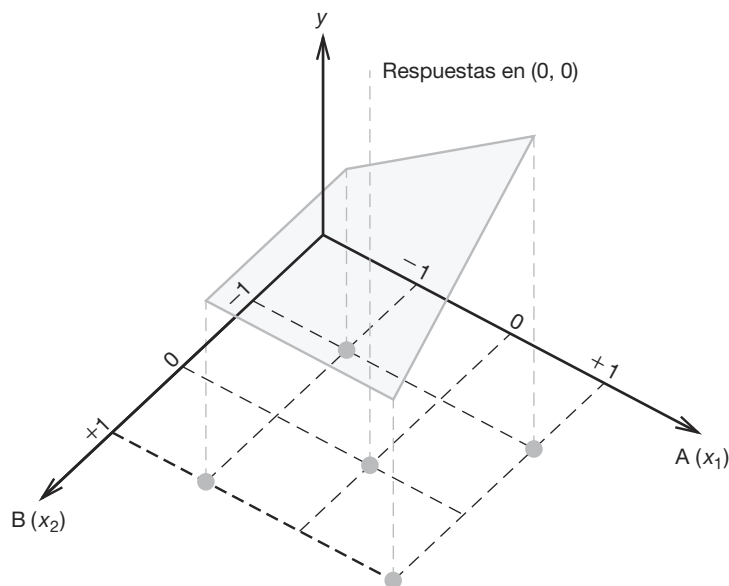


Figura 15.13: Factorial 2^2 con corridas en $(0, 0)$.

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

- a) Haga pruebas para determinar cuáles efectos principales e interacciones deberían incluirse en el modelo ajustado.
- b) Pruebe para la curvatura cuadrática.
- c) 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.

- a) ¿Una segunda réplica del esquema de bloques del ejercicio 15.13 sería la mejor opción?
- b) Si la respuesta del inciso a es negativa, proporcione el diseño de una mejor opción para la segunda réplica.
- c) ¿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 seleccionaría, que permitieran estimar los términos x_1^2 y x_2^2 ? Explique su respuesta.

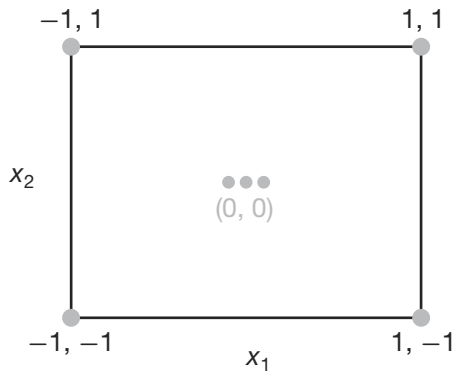


Figura 15.14: Gráfica para el ejercicio 15.17.

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>I</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
2^{3-1}	<i>a</i>	+	+	−	−	−	−	+	+
	<i>b</i>	+	−	+	−	−	+	−	+
	<i>c</i>	+	−	−	+	+	−	−	+
	<i>abc</i>	+	+	+	+	+	+	+	+
2^{3-1}	<i>ab</i>	+	+	+	−	+	−	−	−
	<i>ac</i>	+	+	−	+	−	+	−	−
	<i>bc</i>	+	−	+	+	−	−	+	−
	(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

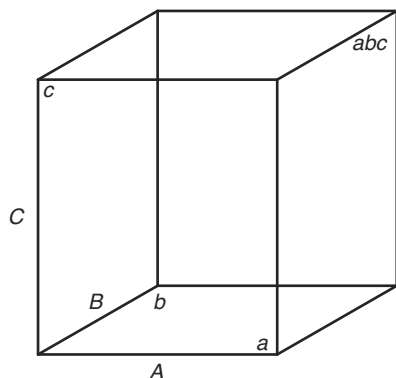
$$A = \frac{a + abc - b - c}{2} \quad \text{y} \quad BC = \frac{a + abc - b - c}{2}$$

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 .

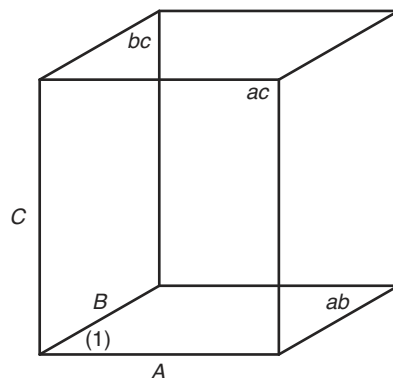
Para la fracción $ABC = -I$ se observa que los alias son los mismos que para la fracción $ABC = I$, además del signo. Así, se tiene

$$A \equiv -BC; \quad B \equiv -AC; \quad C \equiv -AB.$$

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,

$$A \cdot ABC = A, \quad \text{con lo que} \quad BC \equiv A.$$

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>bcd</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						
	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
<i>a</i>	+	−	−	−	−	+	+
<i>b</i>	−	+	−	−	+	−	+
<i>c</i>	−	−	+	+	−	−	+
<i>abc</i>	+	+	+	+	+	+	+

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

$$SCA = \frac{(11.3 - 15.6 - \cdots - 14.7 + 13.2)^2}{2^{5-1}} = \frac{(-17.5)^2}{16} = 19.14,$$

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

$$A = -\frac{17.5}{8} = -2.19,$$

$$SCB = \frac{(-11.3 + 15.6 - \cdots - 14.7 + 13.2)^2}{2^{5-1}} = \frac{(18.1)^2}{16} = 20.48,$$

$$B = \frac{18.1}{8} = 2.26,$$

$$SCC = \frac{(-11.3 - 15.6 + \cdots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(10.3)^2}{16} = 6.63,$$

$$C = \frac{10.3}{8} = 1.21,$$

$$SCD = \frac{(-11.3 - 15.6 - \cdots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(-7.7)^2}{16} = 3.71,$$

$$D = \frac{-7.7}{8} = -0.96,$$

$$SCE = \frac{(-11.3 - 15.6 - \cdots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(8.9)^2}{16} = 4.95,$$

$$E = \frac{8.9}{8} = 1.11.$$

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. ■

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>bcdg</i>	35.6
<i>bf</i>	27.8	<i>abcf</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 adquiriera conocimiento acerca de cómo influye la temperatura de sellado *A*, la temperatura de enfriamiento de una barra *B*, el porcentaje de aditivo de polietileno *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 $\frac{1}{4}$, 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 $ABCDE$. 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_{III}^{3-1}	4	$C = \pm AB$
4	2_{IV}^{4-1}	8	$D = \pm ABC$
5	2_{III}^{5-2}	8	$D = \pm AB ; E = \pm AC$
6	2_{VII}^{6-1}	32	$F = \pm ABCDE$
	2_{IV}^{6-2}	16	$E = \pm ABC ; F = \pm BCD$
	2_{III}^{6-3}	8	$D = \pm AB ; F = \pm BC ; E = \pm AC$
7	2_{VII}^{7-1}	64	$G = \pm ABCDEF$
	2_{IV}^{7-2}	32	$E = \pm ABC ; G = \pm ABDE$
	2_{IV}^{7-3}	16	$E = \pm ABC ; F = \pm BCD ; G = \pm ACD$
	2_{III}^{7-4}	8	$D = \pm AB ; E = \pm AC ; F = \pm BC ; G = \pm ABC$
8	2_V^{8-2}	64	$G = \pm ABCD ; H = \pm AB EF$
	2_{IV}^{8-3}	32	$F = \pm ABC ; G = \pm ABD ; H = \pm BCDE$
	2_{IV}^{8-4}	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 2^4 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 2^5 .

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:

$N = 12$	+	+	-	+	+	+	-	-	-	+	-																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							</
----------	---	---	---	---	---	---	---	---	---	---	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

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 (%)	x_1	x_2	y
		ξ_1	ξ_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

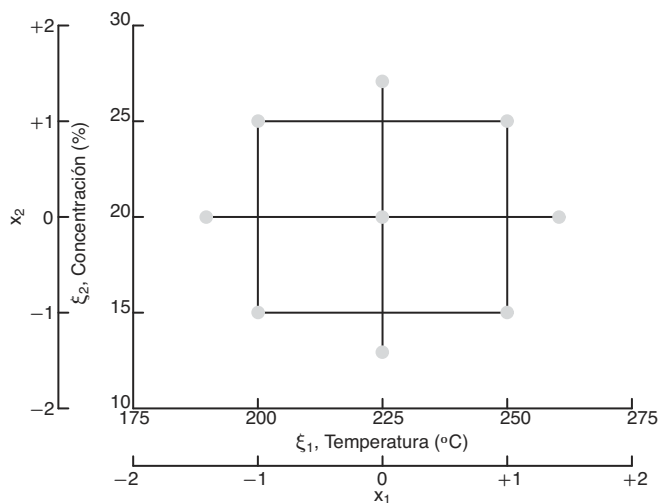


Figura 15.16: Diseño central compuesto para el ejemplo 15.8.

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%.

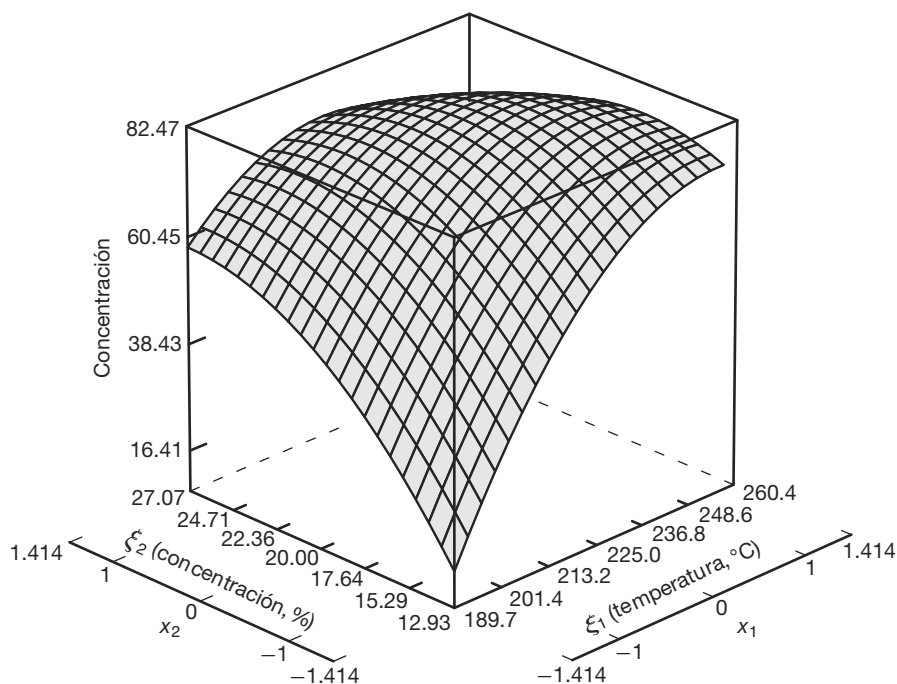


Figura 15.17: Gráfica de la superficie de respuesta de la conversión pronosticada para el ejemplo 15.8.

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.

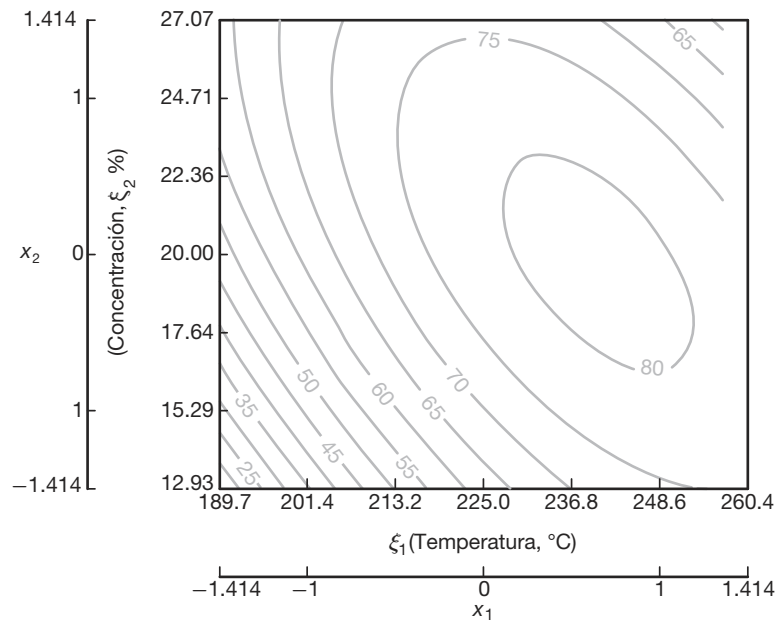


Figura 15.18: Gráfica de curvas de la conversión pronosticada para el ejemplo 15.8.

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 <i>cambios en las variables de ruido</i> 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

		B (profundidad)		
		Superficial	Media	Profunda
A (espesor de pared)	Delgado	(1) <i>c</i> <i>d</i> <i>cd</i>	(1) <i>c</i> <i>d</i> <i>cd</i>	(1) <i>c</i> <i>d</i> <i>cd</i>
	Medio	(1) <i>c</i> <i>d</i> <i>cd</i>	(1) <i>c</i> <i>d</i> <i>cd</i>	(1) <i>c</i> <i>d</i> <i>cd</i>
	Grueso	(1) <i>c</i> <i>d</i> <i>cd</i>	(1) <i>c</i> <i>d</i> <i>cd</i>	(1) <i>c</i> <i>d</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 precalientan 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 ($^{\circ}\text{F}$)	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 ($^{\circ}\text{F}$), (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 2^5 , y el segundo es $\frac{1}{2}$ de un diseño 2^3 . 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 2^3 . 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 $ABDE$. El arreglo exterior es una fracción estándar de resolución III de un diseño 2^3 . 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					
<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	(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

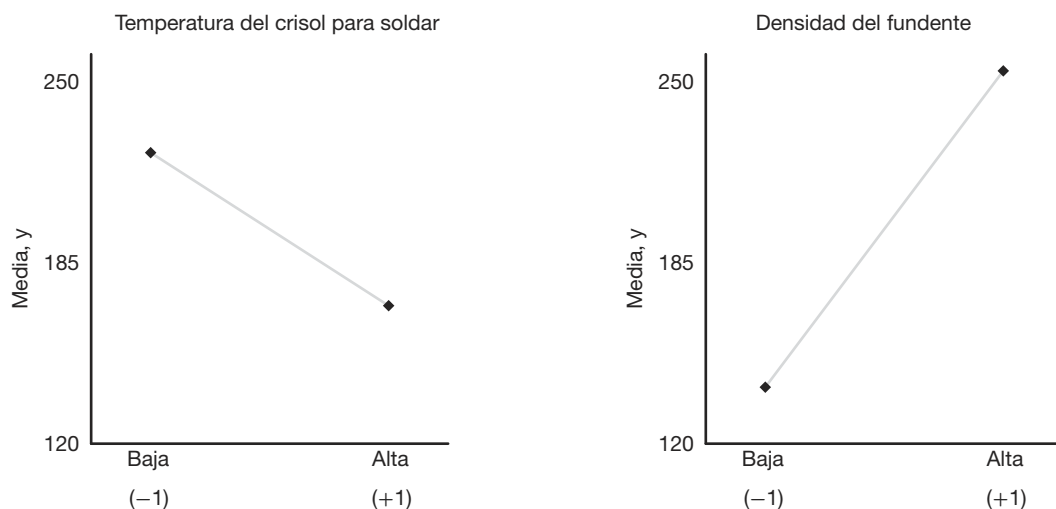


Figura 15.19: Gráfica que muestra la influencia de los factores sobre la respuesta media.

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

temperatura de soldadura = 510°F, densidad del fundente = 0.9°.

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 proporcionarían 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

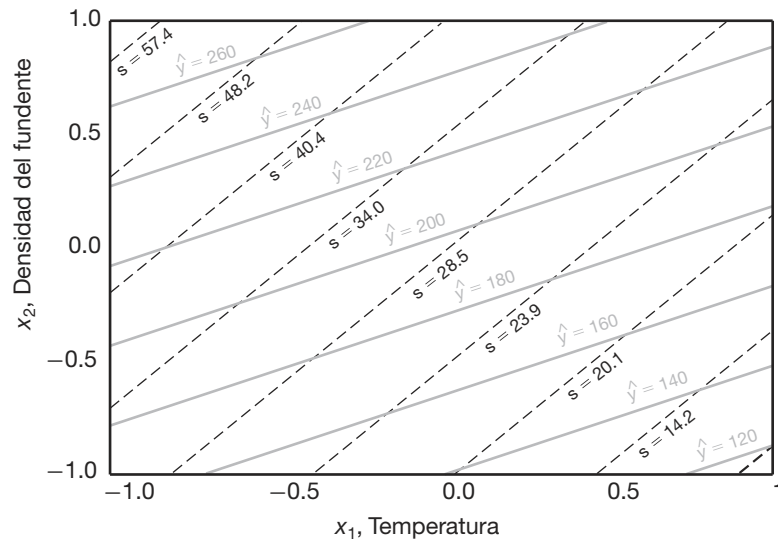


Figura 15.20: Media y desviación estándar del estudio de caso 15.4.

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.

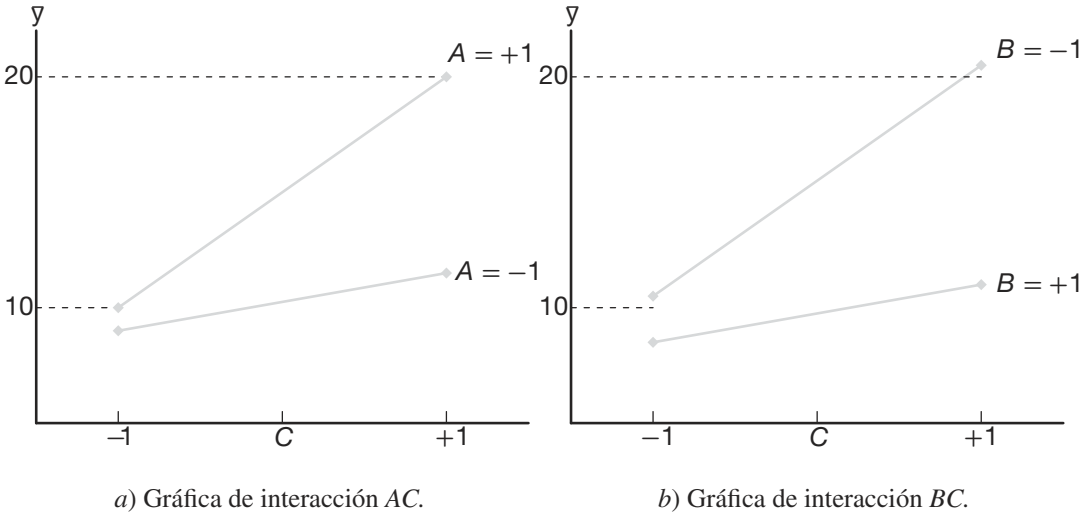


Figura 15.21: Gráficas de interacción para los datos de la tabla 15.22.

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 \mathbf{x} y en el vector \mathbf{z} , donde \mathbf{x} contiene variables de control y \mathbf{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

$$E_z[y(x, z)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{1z} x_1 z, \\ \text{Var}_z[y(x, z)] = \sigma^2 + \sigma_z^2(\beta_3 + \beta_{1z} x_1 + \beta_{2z} x_2)^2 = \sigma^2 + \sigma_z^2 l_x^2,$$

donde l_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.

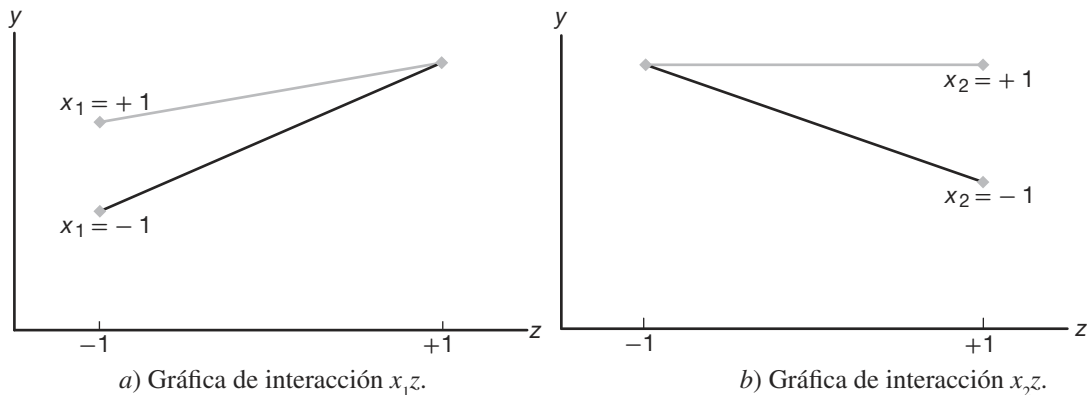


Figura 15.22: Gráficas de interacción para los datos del ejercicio 15.31.

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 2^3 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_z^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	11.3	12.0
Alta	11.7	12.0	21.7	22.4

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 2^6 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	$abef$	2.2480
2	$cdef$	1.8570
3	(1)	2.2428
4	ace	2.3270
5	bde	1.8830
6	$abcd$	1.8078
7	adf	2.1424
8	bcf	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}_j 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

$$E(Y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3, \\ + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{33}x_3^2,$$

- a) Proporcione (y compruebe) $E(\bar{y}_j - \bar{y}_0)$.
- b) 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

$$\begin{aligned} H_0: \tilde{\mu} &= \tilde{\mu}_0, \\ H_1: \tilde{\mu} &< \tilde{\mu}_0, \end{aligned}$$

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 $\tilde{\mu} = \tilde{\mu}_0$ realmente se puede rechazar a un nivel de significancia de 0.05 pero no a un nivel de 0.01.

Para probar la hipótesis

$$\begin{aligned} H_0: \tilde{\mu} &= \tilde{\mu}_0, \\ H_1: \tilde{\mu} &> \tilde{\mu}_0, \end{aligned}$$

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

$$\begin{aligned} H_0: \tilde{\mu} &= \tilde{\mu}_0, \\ H_1: \tilde{\mu} &\neq \tilde{\mu}_0, \end{aligned}$$

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

$$H_0: \tilde{\mu} = \tilde{\mu}_0,$$

$$H_1: \tilde{\mu} < \tilde{\mu}_0,$$

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

$$\tilde{\mu} = np = (20)(0.5) = 10$$

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. $H_0: \tilde{\mu} = 1.8$.
 2. $H_1: \tilde{\mu} \neq 1.8$.
 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. ─

También se puede utilizar la prueba de signo para probar la hipótesis nula $\tilde{\mu}_1 - \tilde{\mu}_2 = d_0$ 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 $\tilde{\mu}_1$ y $\tilde{\mu}_2$ la mediana de los kilómetros por litro para los automóviles equipados con neumáticos radiales y con cinturón, respectivamente.

1. H_0 : $\tilde{\mu}_1 - \tilde{\mu}_2 = 0$.
2. H_1 : $\tilde{\mu}_1 - \tilde{\mu}_2 > 0$.
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 variaría. 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: 1. H_0 : $\tilde{\mu} = 1.8$.

2. H_1 : $\tilde{\mu} \neq 1.8$.

3. $\alpha = 0.05$.

4. 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 $w \leq 8$.

5. 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. ─

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. ─

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 registran 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 $\tilde{\mu}_1 = \tilde{\mu}_2$ 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 $\tilde{\mu}_1 = \tilde{\mu}_2$ se rechazará a favor de la hipótesis alternativa $\tilde{\mu}_1 < \tilde{\mu}_2$ sólo si w_1 es pequeña y w_2 es grande. De igual manera, la hipótesis alternativa $\tilde{\mu}_1 > \tilde{\mu}_2$ 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 $\tilde{\mu}_1 < \tilde{\mu}_2$ si w_1 es suficientemente pequeña; la hipótesis alternativa $\tilde{\mu}_1 > \tilde{\mu}_2$ se acepta si w_2 es suficientemente pequeña; y la hipótesis alternativa $\tilde{\mu}_1 \neq \tilde{\mu}_2$ 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

$$\mu_{U_1} = \frac{n_1 n_2}{2} \text{ y } \sigma_{U_1}^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}.$$

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 $\nu = 2$ 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:

| | Calificación | | | | | | |
|-------------------------|--------------|----|----|----|----|----|----|
| Con enseñanza adicional | 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 béisbol | | | 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,

$$H H H H H H M M M M M$$

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 $P(V \leq v^* \text{ cuando } H_0 \text{ es verdadera})$ 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. ■

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 traslapen, obtendríamos una secuencia de la forma

$$A A A A A B B B B \quad \text{o} \quad B B B B B A A A A A$$

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

$$\mu_V = \frac{2n_1n_2}{n_1 + n_2} + 1 \quad \text{y} \quad \sigma_V^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}.$$

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,..., 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 plateado 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 |

- a) Calcule el coeficiente de correlación de rangos.
- b) 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

- a) calcule el coeficiente de correlación de rangos;
- b) 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

- a) calcule el coeficiente de correlación de rangos;
- b) 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 |

- a) Calcule el coeficiente de correlación de rangos.
- b) 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.

- a) Calcule el coeficiente de correlación de rangos.
- b) 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.