

2a.

edición

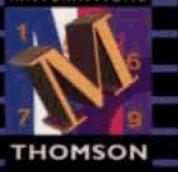
Principios
estadísticos
para
el diseño
y
análisis
de
investigaciones

Diseño de Σ xperimentos



Robert O. Kuehl

MATEMÁTICAS



THOMSON

Robert O.
Kuehl

Diseño
y desarrollo
de sistemas
informáticos

2a. edición

THOMSON
*
LEARNING

Diseño de experimentos
Principios estadísticos de diseño
y análisis de investigación
Segunda edición

Robert O. Kuehl
The University of Arizona





Diseño de experimentos
Robert O. Kuehl

Vicepresidente de Editorial y Producción:
Miguel Ángel Toledo Castellanos

Editor de desarrollo:
Pedro de la Garza Rosales

Traducción:
Marcia González Osuna

Corrección de estilo:
Sergio Antonio Durán

COPYRIGHT© 2001 por International Thomson Editores, S.A. de C.V., una división de Thomson Learning, Inc. Thomson Learning™ es una marca registrada usada bajo permiso.

Impreso en México

Printed in Mexico

1 2 3 4 03 02 01

Para mayor información contáctenos a:

Séneca 53

Col. Polanco

México, D.F. 11560

Gerente de producción:
René Garay Argueta

Editora de producción:
Patricia Pantoja Valdez

Revisión técnica:
Gabriel Morelos Borja
ITESM-Campus Querétaro

Diseño de portada:
Jesús Enríquez

Tipografía:
Juan Rico

Lecturas:
Magdalena Ruiz y
José Carlos Morales

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

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

Traducido del libro *Desing of Experiments 2nd ed.*, publicado en inglés por Duxbury, ©2000
ISBN-0-534-36834-4.

Datos para catalogación bibliográfica:
Kuehl, Robert

Diseño de experimentos 2a. Ed.
Incluye referencias bibliográficas e índice.

ISBN 970-686-048-7

1 Diseño de experimentos 2. Principios estadísticos para el análisis y diseño de investigaciones.

División Iberoamericana

México y América Central:
Thomson Learning
Séneca 53
Col. Polanco
México, D.F. 11560
Tel (525) 281 29 06
Fax (525) 281 29 06
editor@thomsonlearning.com.mx

Pacto Andino
Thomson Learning
Calle 39 No.24-09
La Soledad
Bogotá, Colombia
Tel (571) 340-9470
Fax (571) 340-9475
clithomson@andinet.com

Cono Sur:
Suipacha 774 2 C
Buenos Aires, Argentina
Tel (5411) 4325-2236
Fax (5411) 4328-1829
thomson@thomsonlearning.com.ar

El Caribe:
Thomson Learning
Home Mortgage Plaza
268 Ponce de León A
Suite 510, 5th Floor
Hato Rey, Puerto Rico
Tel (787) 758-7580
Fax (787) 758-7573
thomson@coqui.net

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

A mamá y papá,
Schleswig,
Tío Henry
y el emparedado de conejo

Contenido

1 Principios para el diseño de investigaciones 1

- 1.1 El legado de Sir Ronald A. Fisher 1
- 1.2 Planeación de la investigación 2
- 1.3 Experimentos, tratamientos y unidades experimentales 3
- 1.4 La hipótesis de investigación genera el diseño de los tratamientos 5
- 1.5 Control local de errores experimentales 8
- 1.6 Replicar para obtener experimentos válidos 16
- 1.7 ¿Cuántas réplicas? 18
- 1.8 Aleatorizar para tener inferencias válidas 20
- 1.9 Eficiencia relativa del diseño de experimentos 24
- 1.10 De los principios a la práctica: un caso de estudio 26

2 Comenzando con diseños totalmente aleatorizados 37

- 2.1 Construcción del diseño de investigación 37
- 2.2 Cómo aleatorizar 39
- 2.3 Preparación de los registros de datos para el análisis 41
- 2.4 Un modelo estadístico para la experimentación 42
- 2.5 Estimación de los parámetros del modelo con mínimos cuadrados 47
- 2.6 Suma de cuadrados para identificar fuentes de variación importantes 50
- 2.7 Modelo de efectos del tratamiento 53
- 2.8 Grados de libertad 54
- 2.9 Resumen en la tabla de análisis de varianza 55
- 2.10 Pruebas de hipótesis sobre modelos lineales 56
- 2.11 Pruebas de significancia y pruebas de hipótesis 58
- 2.12 Errores estándar e intervalos de confianza para medias de tratamientos 59
- 2.13 Diferente número de réplicas de los tratamientos 60
- 2.14 ¿Cuántas réplicas para la prueba F ? 63
- Apéndice 2A.1 Valores esperados 70
- Apéndice 2A.2 Cuadrados medios esperados 71

3 Comparación de tratamientos 73

3.1	La comparación de tratamientos responde las preguntas de la investigación	73
3.2	Planeación de comparaciones entre los tratamientos	74
3.3	Curvas de respuesta para los factores de tratamiento cuantitativos	83
3.4	Las comparaciones múltiples afectan las tasas de error	91
3.5	Inferencia estadística simultánea	94
3.6	Comparaciones múltiples con el mejor tratamiento	98
3.7	Comparación de todos los tratamientos con un control	104
3.8	Comparaciones en pares de todos los tratamientos	107
3.9	Resumen de comentarios sobre las comparaciones múltiples	115
	Apéndice 3A Funciones lineales de las variables aleatorias	121

4 Diagnóstico de la concordancia entre los datos y el modelo 123

4.1	Un análisis válido depende de suposiciones válidas	123
4.2	Efectos de la falta de cumplimiento de las suposiciones	123
4.3	Los residuales son la base de las herramientas de diagnóstico	124
4.4	Búsqueda de residuos inusitados con los residuales	131
4.5	Transformaciones estabilizadoras de la varianza para datos con distribuciones conocidas	133
4.6	Transformaciones con exponentes para estabilizar varianzas	135
4.7	Generalización del modelo lineal	140
4.8	Evaluación de modelos por medio de gráficas de residuales <i>vs</i> valores ajustados	141
	Apéndice 4A Datos para el ejemplo 4.1	147

5 Experimentos para estudiar las varianzas 148

5.1	Modelos con efectos aleatorios para las varianzas	148
5.2	Un modelo estadístico para las componentes de la varianza	151
5.3	Estimaciones puntuales de las componentes de la varianza	152
5.4	Estimaciones de intervalos para las componentes de la varianza	153
5.5	Cursos de acción con estimaciones de varianzas negativas	155
5.6	La correlación intraclasses indica similitud dentro de los grupos	155
5.7	Diferentes números de observaciones en los grupos	157
5.8	¿Cuántas observaciones para estudiar las varianzas?	158
5.9	Submuestras aleatorias para reunir datos para el experimento	159
5.10	Uso de las estimaciones de la varianza para asignar muestras	163
5.11	Números diferentes de réplicas y submuestras	164
	Apéndice 5A Cálculo de los coeficientes para los cuadrados medios esperados de la tabla 5.9	174

6 Diseños factoriales 175

6.1	Experimentos eficientes con diseños factoriales	175
6.2	Tres efectos de los factores	177
6.3	Modelo estadístico para dos factores	181
6.4	Análisis para dos factores	183
6.5	Uso de curvas de respuesta para los factores de tratamiento cuantitativos	190
6.6	Tres factores de tratamiento	199
6.7	Estimación de la varianza del error con una réplica	205
6.8	¿Cuántas réplicas se requieren para probar los efectos de un factor?	208
6.9	Réplicas desiguales en los tratamientos	208
	Apéndice 6A Mínimos cuadrados para diseños factoriales	225

7 Diseños factoriales: modelos aleatorios y mixtos 232

7.1	Efectos aleatorios para diseños factoriales	232
7.2	Modelos mixtos	237
7.3	Diseños de factores anidados: una variación del tema	243
7.4	Diseños de factores cruzados y anidados	251
7.5	¿Cuántas réplicas?	255
7.6	Reglas para los cuadrados medios esperados	255

8 Diseños de bloques completos 263

8.1	Uso de bloques para aumentar la precisión	263
8.2	Los diseños de bloques completos aleatorizados usan un criterio de bloqueo	264
8.3	Los diseños de cuadrado latino usan dos criterios de bloqueo	275
8.4	Experimentos factoriales en diseño de bloques completos	289
8.5	Datos faltantes en diseños por bloques	291
8.6	Experimentos realizados varias veces	292
	Apéndice 8A Selección de cuadrados latinos	307

9 Diseños de bloques incompletos: introducción 310

9.1	Bloques incompletos de tratamientos para reducir el tamaño de los bloques	310
9.2	Diseños de bloques incompletos balanceados (BIB)	312
9.3	Cómo aleatorizar los diseños de bloques incompletos	313
9.4	Análisis de diseños BIB	315
9.5	Diseño renglón-columna para dos criterios de bloque	320
9.6	Reducción del tamaño del experimento con diseños parcialmente balanceados (BIPB)	322
9.7	Eficiencia de los diseños de bloques incompletos	325
	Apéndice 9A.1 Algunos diseños de bloques incompletos balanceados	330
	Apéndice 9A.2 Algunos diseños de cuadrados latinos incompletos	332
	Apéndice 9A.3 Estimaciones de mínimos cuadrados para diseño BIB	336

10 Diseño de bloques incompletos: diseños resolubles y cílicos 339

- 10.1 Diseños resolubles para ayudar a manejar el experimento 339
- 10.2 Diseños resolubles renglón-columna para dos criterios de bloque 342
- 10.3 Los diseños cílicos simplifican la construcción de diseños 345
- 10.4 Elección de diseños de bloques incompletos 352
- Apéndice 10A.1 Planes para diseños cílicos 360
- Apéndice 10A.2 Arreglos generadores para diseños α 361

11 Diseños de bloques incompletos: tratamientos factoriales 362

- 11.1 Aprovechamiento de los diseños de tratamientos factoriales 362
- 11.2 Factoriales 2^n para evaluar muchos factores 363
- 11.3 Diseños de bloque incompletos para factoriales 2^n 369
- 11.4 Método general para crear bloques incompletos 378
- 11.5 Diseños de bloques incompletos para factoriales 3^n 383
- 11.6 Observaciones finales 387
- Apéndice 11A Planes para diseños de bloques incompletos para factoriales 2^n 390

12 Diseños factoriales fraccionarios 391

- 12.1 Reducción del número de corridas experimentales con diseños fraccionados 391
- 12.2 Fracción de un medio del factorial 2^n 393
- 12.3 Resolución del diseño relacionada con la aleatorización 398
- 12.4 Análisis de diseños 2^{n-1} de media réplica 399
- 12.5 Fracciones de un cuarto de factoriales 2^n 406
- 12.6 Construcción de diseños 2^{n-p} con resolución III y IV 409
- 12.7 Genichi Taguchi y la mejora de la calidad 413
- 12.8 Observaciones finales 415
- Apéndice 12A Planes de diseño factoriales fraccionarios 421

13 Diseños con superficie de respuesta 423

- 13.1 Descripción de respuestas con ecuaciones y gráficas 423
- 13.2 Identificación de los factores significativos con factoriales 2^n 426
- 13.3 Diseños para estimar superficies de respuesta de segundo orden 431
- 13.4 Estimación de la superficie de respuesta cuadrática 440
- 13.5 Exploración de superficies de respuesta 444
- 13.6 Diseños para mezclas de ingredientes 447
- 13.7 Análisis de experimentos con mezclas 453
- Apéndice 13A.1 Estimación de mínimos cuadrados de los modelos de regresión 463
- Apéndice 13A.2 Localización de coordenadas para el punto estacionario 466
- Apéndice 13A.3 Forma canónica de la ecuación cuadrática 467

14 Diseños de parcelas divididas 469

- 14.1 Parcelas de distintos tamaños en el mismo experimento 469
- 14.2 Dos errores experimentales para dos tamaños de parcelas 472
- 14.3 Análisis para diseños de parcelas divididas 473
- 14.4 Errores estándar para las medias de los factores de tratamiento 478
- 14.5 Características del diseño de parcelas divididas 480
- 14.6 Eficiencia relativa de las comparaciones de subparcelas y parcelas completas 481
- 14.7 Diseño con doble subdivisión de parcelas para tres factores de tratamiento 483
- 14.8 Diseño de bloques divididos 483
- 14.9 Información adicional sobre diseños de parcelas divididas 486

15 Diseños con mediciones repetidas 492

- 15.1 Estudios de tendencias en el tiempo 492
- 15.2 Relaciones entre las mediciones repetidas 495
- 15.3 Una prueba para la suposición Huynh-Feldt 498
- 15.4 Análisis de varianza univariado para mediciones repetidas 499
- 15.5 Análisis cuando no se cumplen las suposiciones del análisis univariado 502
- 15.6 Otros experimentos con propiedades de mediciones repetidas 509
- 15.7 Otros modelos para la correlación entre mediciones repetidas 511
- Apéndice 15A.1 Prueba de esfericidad de Mahecule 518
- Apéndice 15A.2 Ajuste de grados de libertad para el análisis de varianza de mediciones repetidas 519

16 Diseños cruzados 520

- 16.1 Suministro de todos los tratamientos a cada unidad experimental 520
- 16.2 Análisis de diseños cruzados 524
- 16.3 Diseños balanceados para estudios cruzados 530
- 16.4 Diseños cruzados para dos tratamientos 536
- Apéndice 16A.1 Codificación de archivos de datos para estudios cruzados 545
- Apéndice 16A.2 Suma de cuadrados del tratamiento para los diseños balanceados 547

17 Análisis de covarianza 550

- 17.1 Control local con una covariada medida 550
- 17.2 Análisis de covarianza para diseños totalmente aleatorizados 553
- 17.3 Análisis de covarianza para diseños de experimentos bloqueados 565
- 17.4 Consecuencias prácticas del análisis de covarianza 570

Referencias 576**Apéndice de tablas 587****Respuestas a ejercicios seleccionados 633****Índice 661**

Prefacio

Objetivos

Mi objetivo en este libro es presentar los principios de diseño y análisis estadístico para estudios científicos comparativos dirigidos a estudiantes de posgrado en ciencias experimentales y estadística aplicada. En esta segunda edición se agregó el título principal *Diseño de experimentos*, para que se identificara mejor y se conservó el título original para dar un panorama del objetivo del libro. *Principios estadísticos de diseño y análisis de investigación* describe la filosofía de que los experimentos comparativos exitosos tienen objetivos definidos con claridad que pueden estudiarse con la elección adecuada de los diseños de tratamiento. Además, el *diseño* del experimento seleccionado al tomar en cuenta el *diseño* del tratamiento debe proporcionar el *diseño* más eficiente en el contexto del experimento y los recursos de que dispone el científico. Yo llamo a este proceso *diseño de la investigación* —el esfuerzo total en un estudio que incluye el planteamiento de la hipótesis de investigación, la elección del diseño de tratamientos para estudiar la hipótesis de investigación y la selección del diseño del experimento para facilitar la recolección eficiente de datos.

Orientación y antecedentes

El libro está orientado a las aplicaciones, usa los resultados de teorías establecidas y no incluye los desarrollos teóricos. Tiene una orientación de diseño clásico que intenta introducir al estudiante a los principios y diseños fundamentales. Estas bases constituyen los ingredientes necesarios para que los investigadores en estadística y quienes la practican realicen la innovación continua, el mejoramiento y la ampliación del desarrollo de estrategias útiles y eficientes para la investigación. El libro supone una familiaridad razonable con álgebra universitaria y el prerrequisito de un curso introductorio a los métodos estadísticos que incluye las ideas básicas de muestreo aleatorio, leyes de probabilidad, intervalos de confianza, pruebas *t*, análisis de variancia, regresión y pruebas *F*.

Ejemplos y ejercicios

Los ejemplos y ejercicios se basaron en estudios de investigación reales siempre que fue posible. Cerca de la mitad de ellos con conjuntos de datos comprenden aplicaciones a varias ciencias agrícolas y biológicas, mientras que la otra mitad se refiere a aplicaciones de la investigación en ingeniería, la industria y química.

Estructura y cobertura

El primer capítulo hace hincapié en el papel de las ciencias estadísticas en las etapas de planeación del estudio de investigación. Los capítulos 2 a 7 hacen énfasis en la asociación entre las hipótesis de investigación y los diseños de tratamiento y su análisis en las condiciones más sencillas de los diseños totalmente aleatorizados. Los capítulos 8 a 17 presentan desarrollos de diseños de experimentos y su análisis con estructuras de bloqueo sencillas en los primeros capítulos hasta llegar a las más complejas y especializadas en los últimos. Se tuvo cuidado de preservar el paradigma de diseño de la investigación con ejemplos científicos diversos para ilustrar la mayoría de los temas en cada capítulo. Se proporciona al lector un recordatorio consistente de la conexión entre los objetivos de investigación, el diseño estadístico y el análisis que cumple con esos objetivos. Asimismo, se dan las interpretaciones de los análisis.

Nuevo en la segunda edición

El capítulo 1 es clave para todo lo que sigue en el libro ya que establece la relación entre las hipótesis de investigación y los diseños de tratamientos, réplicas, asignaciones aleatorias y prácticas de control local para construir el escenario de los siguientes capítulos. En esta edición se da el crédito adecuado a sir Ronald A. Fisher y su artículo fundamental publicado en 1926, que consolida sus conceptos unificados para las inferencias válidas de los experimentos designados. Se agregó el estudio de un caso que ilustra la aplicación de los principios de diseño presentados en las secciones de este capítulo.

Se incluyeron presentaciones nuevas y actualizadas, notas aclaratorias (espero) y referencias actualizadas de muchos temas, en varios lugares de esta edición. Se hicieron algunos cambios para mejorar la apariencia global que incluyen un formato genérico de la tabla de análisis de varianza en lugar de las impresiones de computadora. También se eliminaron algunas ilustraciones con datos para cálculos con calculadoras manuales de la suma de cuadrados del análisis de varianza básico.

Los programas de computadora surgen de manera constante en términos de los que hacen, la interacción con el usuario y el despliegue de los resultados. Existe abundancia de programas comerciales para ayudar al diseño de experimentos para bloques completos e incompletos, factoriales fraccionarios, mezclas, superficies de respuesta y otros, incluso los esquemas de aleatorización. Los programas para el análisis son aún más numerosos. Elegir un paquete de software específico parece ser algo personal y es algo que mejor se omite en un libro de esta naturaleza.

Se proporcionan indicadores de metodología avanzada para la elección del análisis, como métodos a menudo basados en la máxima posibilidad, como los de estimación de las componentes de varianza en el capítulo 7, los programas de modelos mixtos para parcelas divididas en el capítulo 14 y opciones para las matrices de covarianza para medidas repetidas en el capítulo 15. Se agregó un fragmento sobre el modelo lineal generalizado en el capítulo 4 para que el estudiante sepa que existen modelos y métodos que no requieren la suposición de distribución normal. Incluye referencias de varios libros sobre el análisis del modelo generalizado con diseño especial para los usuarios de este modelo.

Se actualizó la inferencia estadística simultánea en el capítulo 3 con el estudio de las distintas fuerzas de inferencia disponibles al experimentador para el análisis posterior de los métodos de varianza. El nivel de intervalo de confianza de la inferencia con intervalos simultáneos es la opción preferida de inferencia en comparaciones múltiples, seguida de las desigualdades de confianza que se usaron en la primera edición. Se proporcionan algunas inferencias en el capítulo 3 para otras pruebas que requieren software (ahora incluido en algunos paquetes) para calcular los valores críticos de la prueba, pero que proporcionan alternativas mucho mejores que las antiguas.

En el capítulo 4 se agregaron las gráficas de localización-dispersión para ayudar a la evaluación de las suposiciones de varianza homogénea, y se incluyó la gráfica de dispersión de residuales ajustados para evaluar el ajuste del modelo. El modelo mixto no restringido es ahora el modelo elegido para los diseños del capítulo 7, por la razón de que es la extensión natural de los diseños balanceados y no balanceados. Quizá es trivial la preocupación de algunos, pero es un tema con una división pareja entre los modelos “restringidos” y “no restringidos”.

Se añadieron los diseños de bloques incompletos con solución, latinizados en la sección de diseño cíclico del capítulo 10. Estos diseños ofrecen una nueva dimensión para el control del bloqueo para los diseños alfa, y preservan al mismo tiempo la ortogonalidad para ese criterio de bloqueo. Se incluyeron varias páginas sobre la filosofía del método general de Taguchi en el capítulo 12. Aunque el método de Taguchi es muy especializado para la comunidad de investigación en manufactura, tiene algunos componentes de diseño y análisis que deben reconocerse. Un ejemplo con pendientes heterogéneas entre los grupos de tratamiento para la covariada se agregó en el capítulo 17 para ampliar la presentación de las suposiciones del análisis de covarianza.

Agradecimientos

Estoy en deuda con muchas personas que contribuyeron en forma directa o indirecta al desarrollo de este libro. Doy las gracias a mis mentores en las universidades de Iowa State y North Carolina State, colegas y amigos que de alguna manera han tenido una influencia positiva en mí a través de los años. En particular George, Lowell, Russ, Arnel, Charley, John, Clark, Dave y Ted merecen mención especial. Agradezco a los colegas y amigos de la University of Arizona y de otros lados que durante años me han consultado sobre problemas estimulantes que influyeron en

la enseñanza y organización de este material. En particular, les doy las gracias (con reconocimiento en el texto) por su generosa contribución de datos para los ejemplos y ejercicios. Gracias especiales a Elnora Fairbank y Helen Ferris por su dedicación en la preparación del manuscrito en sus diferentes etapas. Estoy agradecido con Rick Axelson por su apoyo en la programación estadística durante la preparación del manuscrito original y con Harika Basaran por preparar en la computadora los datos de los ejercicios y ejemplos. Estoy en deuda con John Kimmel por su confianza y apoyo en las etapas de formación del libro, y extiendo mi especial aprecio a Alex Kugushev por su dedicación a la excelencia. Deseo expresar mi gratitud a Carolyn Crockett, quien confió en la viabilidad del libro y promovió la segunda edición. Doy las gracias a Kimberly Raburn por su valiosa guía en todas las cosas importantes durante el proceso de revisión. Por último, reconozco las críticas y sugerencias útiles de muchos revisores de ambas ediciones, entre ellos Richard Alldredge, Daniel C. Coster, Shu Geng, Robert Heckard, Hui-Kuang Hsieh, David Jowett, Larry J. Ringer, Oliver Schabenberger y G. Morris Southward.

Robert O. Kuehl

1 Principios para el diseño de investigaciones

Todas nuestras actividades asociadas con planear y realizar estudios de investigación tienen implicaciones estadísticas. Los principios que se presentan en el capítulo 1 constituyen la base para la estructura de un estudio de investigación; a su vez, esa estructura define la función del estudio. Si la estructura es razonable, el estudio funcionará de manera adecuada y se obtendrá la información para la que fue diseñado. Si la estructura tiene fallas, el estudio no funcionará bien y presentará información incompleta o errónea. Los principios estadísticos son los asociados con la recolección de aquellas observaciones que proporcionen la mayor cantidad de información para el estudio de investigación de una manera eficiente. Incluyen el diseño de tratamientos, el control local de la variabilidad, el número de réplicas, la aleatorización y la eficiencia de los experimentos.

1.1 El legado de sir Ronald A. Fisher

Nadie ha tenido tanto impacto en los principios estadísticos del diseño de experimentos en su tiempo como Ronald A. Fisher. En octubre de 1919, Fisher fue contratado en Rothamsted Experimental Station, cerca de Harpenden, Inglaterra. Le pidieron que trabajara con ellos de seis meses a un año, para aplicar un exhaustivo análisis estadístico a los datos de investigaciones agrícolas que el personal había recolectado.

Fue durante su ejercicio en Rothamsted, donde permaneció hasta 1933, que desarrolló y consolidó los principios básicos de diseño y análisis que hasta la fecha son prácticas necesarias para llegar a resultados de investigación válidos. De 1919 a 1925 estudió y analizó experimentos relativos al trigo que se habían realizado desde 1843. De sus investigaciones estadísticas de éstos y otros experimentos, Fisher desarrolló el análisis de varianza y unificó sus ideas básicas sobre los principios del diseño de experimentos.

En 1926 publicó el primer resumen completo de sus ideas en el artículo “The Arrangement of Field Experiments” (Fisher, 1926); en ese importante artículo describió tres componentes fundamentales de los experimentos en el área de pruebas agrícolas: **control local** de las condiciones de campo para reducir el error experimental, **replicación** como un medio para estimar la varianza del error experimen-

tal y **aleatorización** para obtener una estimación válida de esa varianza. Aunque la replicación y el control local se aplicaban en ese tiempo, sus justificaciones respecto a la varianza del error experimental eran conceptos relativamente nuevos. La aleatorización fue un concepto radicalmente nuevo que se encontró con el escepticismo y la resistencia de sus contemporáneos, quienes en su mayor parte no entendían sus implicaciones estadísticas. Estos conceptos se estudiarán con detalle en las siguientes secciones de este capítulo.

Dos libros de Fisher que surgieron a partir de sus experiencias en Rothamsted, se convirtieron en consulta obligada para los investigadores durante el diseño y análisis de estudios experimentales. *Statistical Principles for Research Workers (Principios estadísticos para investigadores)* se publicó por primera vez en 1925 (Fisher, 1925), con 13 ediciones subsecuentes; y *Design of Experiments (Diseño de experimentos)*, se publicó 10 años más tarde (Fisher, 1935), con 7 ediciones posteriores. Sus contribuciones al diseño de experimentos fueron sólo unas cuantas entre las que hizo a la ciencia de la estadística. Una biografía con sus experiencias, escrita por una de sus hijas (Box, 1978), le rinde homenaje como científico consumado.

1.2 Planeación de la investigación

Un **programa de investigación** es un esfuerzo organizado de un científico para adquirir conocimientos sobre un proceso natural o artificial. El programa completo puede necesitar de muchos estudios individuales, cada uno con objetivos específicos. Normalmente, los estudios individuales responden preguntas y proporcionan piezas de información afines que, en conjunto, satisfacen las metas del programa. El diseño y el análisis de los estudios de investigación individuales son el objeto de atención de este libro.

La buena planeación ayuda al científico a organizar las tareas necesarias para un estudio de investigación. El estudio individual exige que el científico tome una serie de decisiones críticas. Examinemos a un científico de nutrición que desea mejorar el método normal para evaluar la calidad nutricional de diferentes fuentes de proteínas. Aunque existía un procedimiento estándar de evaluación aceptable, con una definición bastante rígida dada por sus colegas, planteó la hipótesis de que la sustitución de ratas por ratones como animal de prueba era más eficiente en término de costo y tiempo. Un examen de su hipótesis requería un estudio para determinar si los ratones eran más eficientes. Entre las decisiones críticas que debía tomar estaban el número de ratones y ratas que usaría, la cantidad de proteína en las dietas, las distintas fuentes proteínicas necesarias para validar el nuevo protocolo, la duración del experimento y el número de réplicas del mismo.

1.3 E

Los planes documentados previenen omisiones

No es posible exagerar la importancia de desarrollar un plan por escrito. La consulta frecuente de un documento previene las omisiones graves; también será útil para insertar notas y cualquier alteración relacionada con los detalles específicos del plan original.

Al iniciar el estudio, el investigador sagaz desarrolla una lista de verificación de aspectos concretos; algunos de los que suelen incluirse son:

- objetivos específicos del experimento
- identificación de los factores que influyen y cuáles de ellos varían y cuáles permanecen constantes
- características a medir
- procedimientos particulares para realizar las pruebas o medir las características
- número de repeticiones del experimento básico a realizar
- recursos y materiales disponibles.

Bicking (1954) presenta una detallada lista de verificación para planear un estudio de investigación que se puede consultar como guía para desarrollar un plan escrito.

Preguntas sencillas para enfocar las actividades

Las preguntas sencillas, pero que impliquen un reto, auxilian en el proceso de diseño, aunque se tenga una hipótesis de investigación bien definida como móvil del estudio de investigación.

Las preguntas que centran nuestra atención a través del proceso de diseño incluyen: “¿Cuál es mi objetivo?”, “¿Qué quiero saber?” y “¿Por qué quiero saberlo?”. Las preguntas de seguimiento productivo para cada actividad en el proceso, tales como: “¿Cómo voy a realizar esta tarea?” y “¿Por qué hago esta tarea?”, dirigen la atención a definir el papel de cada actividad en el estudio de investigación.

Las componentes del estudio se analizarán por separado en las siguientes secciones, pero están interrelacionadas y en un estudio real el investigador debe integrarlas. Para empezar, a continuación se establece el vocabulario concreto que se usará para expresar las ideas.

1.3 Experimentos, tratamientos y unidades experimentales

La comunicación correcta requiere que ambas partes respondan a un vocabulario común con un mismo significado. Esta sección establece la interpretación de algunos términos y conceptos comunes al aplicarlos a los estudios de investigación científica.

Para el propósito de este libro, un **experimento** debe limitarse a investigaciones que establecen un conjunto particular de circunstancias, bajo un protocolo específico para observar y evaluar las implicaciones de las observaciones resultantes. El investigador determina y controla los protocolos de un experimento para evaluar y probar algo que en su mayor parte no se conoce hasta ese momento.

El **experimento comparativo** es el tipo de experimento que utilizan los investigadores en áreas como biología, medicina, agricultura, ingeniería, sicología y

otras ciencias experimentales. El adjetivo *comparativo* implica que se establezca más de un conjunto de circunstancias en el experimento y que se comparan entre sí las respuestas a las diferentes circunstancias.

Los **tratamientos** son el conjunto de circunstancias creados para el experimento, en respuesta a la hipótesis de investigación y son el centro de la misma. Entre los ejemplos de tratamientos se encuentran dietas de animales, producción de variedades de cultivos, temperaturas, tipos de suelo y cantidades de nutrientes. En un estudio comparativo se usan dos o más tratamientos y se comparan sus efectos en el sujeto de estudio.

La **unidad experimental** es la entidad física o el sujeto expuesto al tratamiento independientemente de otras unidades. La unidad experimental, una vez expuesta al tratamiento, constituye una sola réplica del tratamiento.

El **error experimental** describe la variación entre las unidades experimentales tratadas de manera idéntica e independiente. Los distintos orígenes del error experimental son: 1) la variación natural entre unidades experimentales; 2) la variabilidad en la medición de la respuesta; 3) la imposibilidad de reproducir las condiciones del tratamiento con exactitud de una unidad a otra; 4) la interacción de los tratamientos con las unidades experimentales, y 5) cualquier otro factor externo que influya en las características medidas.

Una prueba de alimentación a reses proporciona un ejemplo de la variación natural entre las unidades experimentales; dos reses de la misma cría y vacada reciben la misma cantidad de alimento, pero durante el lapso de un mes una engorda 2 libras al día y la otra 2.3.

La imposibilidad de reproducir el tratamiento con exactitud se da cuando las réplicas de los tubos de ensayo se preparan de modo independiente, con la misma mezcla de compuestos, y el peso de los productos químicos obtenidos en cada tubo difiere en 0.1 microgramo (μg). Los procesos en pipetas o básculas no son exactos, por lo tanto, introducen una pequeña variación durante la preparación del tratamiento.

Un importante objetivo de los cálculos estadísticos es lograr una estimación de la *varianza del error experimental*. En su forma más simple, el error experimental es la varianza en unidades de las observaciones del experimento, cuando las diferencias entre éstas se pueden atribuir sólo al error experimental. Muchos de los procedimientos estadísticos usados requieren una estimación de esta varianza, como las estimaciones del intervalo de confianza para una media y las pruebas *t* de Student de dos muestras, para la hipótesis de que no hay diferencias entre las medias de dos poblaciones en tratamiento.

Los **estudios por observación comparativos** son aquellos para los que desearíamos hacer un experimento, pero no es posible por razones prácticas o éticas. El investigador tiene en mente condiciones o tratamientos que tienen efectos causales en sujetos para los que no es posible efectuar experimentos hasta obtener respuestas. Los investigadores de las ciencias sociales, ecología, vida silvestre, vida marina y otros recursos naturales, a menudo han de llevar a cabo estudios por observación en vez de la experimentación directa. Tal vez la unidad básica de estudio en la investigación sean sujetos humanos, animales individuales, hábitats u otros microcosmos, que tienen el papel de unidad experimental en los experimentos diseñados.

Los sujetos se autoseleccionan en grupos identificables, o sencillamente existen en sus circunstancias particulares; estos grupos o circunstancias se usan como clasificaciones del tratamiento en el estudio por observación. Por el contrario, el investigador asigna los tratamientos para las unidades experimentales del experimento diseñado. Por ejemplo, para estudiar la nitrificación de suelos a partir de grupos puros y mixtos de pinos y robles, un científico escoge grupos ya existentes de las dos especies y recolecta las observaciones necesarias en los sitios seleccionados. En este caso, un experimento real es impráctico porque requiere una gran cantidad de tiempo establecer grupos de árboles maduros.

En ocasiones, las consideraciones éticas impiden el uso de experimentos en sustitución de los estudios por observación. Considerando un estudio para comparar la severidad de las lesiones por accidentes automovilísticos con y sin el cinturón de seguridad puesto; sería clara la falta de ética si se asignara al azar un tratamiento de personas “con cinturón” y “sin cinturón” y después se provocaran colisiones de los autos contra una pared de concreto, tampoco habría quién se prestara para esto. En su lugar, los investigadores se apoyan en los datos de lesiones en los accidentes y comparan los datos de “cinturón puesto” y “sin cinturón”.

La naturaleza de la inferencia científica es la diferencia principal entre un experimento diseñado y un estudio por observación. En el experimento diseñado suele ser posible asignar las relaciones causales entre las respuestas y los tratamientos; mientras que los estudios por observación se limitan a asociar las relaciones entre las respuestas y las condiciones del tratamiento.

1.4 La hipótesis de investigación genera el diseño de los tratamientos

La hipótesis de investigación establece un conjunto de circunstancias y sus consecuencias. Los tratamientos son una creación de las circunstancias para el experimento. Así, es importante identificar los tratamientos con el papel que cada uno tiene en la evaluación de la hipótesis de investigación. Si no se logra delinear con claridad esta hipótesis y el objetivo del estudio, puede haber dificultades en la selección de los tratamientos y experimentos sin éxito.

Relación entre los tratamientos y las hipótesis

Cuando se eligen los tratamientos adecuados en respuesta a una hipótesis de investigación, es posible comprender mejor los mecanismos fundamentales, ya sean físicos, químicos biológicos o sociales. En algunos casos, el objetivo puede ser “escoger al ganador” para encontrar un tratamiento que proporcione la respuesta deseada; en otros casos, el experimento se usa para aclarar los mecanismos básicos asociados con los tratamientos, en términos de su influencia sobre las variables de respuesta medidas. En este último caso, una hipótesis sólida motiva la selección de los tratamientos.

Es forzoso que el investigador se asegure de que los tratamientos elegidos concuerden con la hipótesis de investigación. Puede ser suficiente —y más senci-

llo— diseñar un estudio sólo para descubrir el mejor tratamiento. Sin embargo, con poco esfuerzo adicional puede obtenerse aún más información fundamental del experimento en respuesta a la hipótesis de investigación.

A continuación se presentan tratamientos, usados en el marco de investigaciones reales, generados por hipótesis de investigación:

- Se estudió la cinética de bebida de las abejas productoras de miel a diferentes temperaturas ambientales, para responder a la hipótesis de que la energía requerida al reunir comida para la colonia era independiente de la temperatura.
- La supervivencia de siembras de *Euphorbia* atacadas por un patógeno del suelo se determinó para distintos tipos de tratamientos de fungicida, en respuesta a la hipótesis de que no todos los fungicidas tienen la misma eficiencia para controlar tal patógeno.
- En ingeniería de tránsito, se evaluaron varios métodos para medir los retrasos del tránsito en intersecciones con diferentes tipos de configuraciones en los semáforos, en respuesta a la hipótesis de que el método para medir retrasos dependía del tipo de configuración usada en la señalización.
- El desarrollo de la adaptación social en niños pequeños se midió según su relación con: 1) educación de los padres, 2) ingresos de los padres, 3) estructura familiar y 4) edad del niño, para analizar una compleja hipótesis de investigación de qué ciertas características demográficas familiares afectan de manera favorable el desarrollo de un niño.

Observe que en algunas situaciones de investigación los tratamientos son condiciones impuestas por el investigador, como las del estudio de las abejas y la supervivencia de las semillas; mientras que en los estudios de ingeniería de tránsito y del desarrollo de los niños, los tratamientos corresponden a las condiciones presentes. Ya sea que el investigador realice un experimento diseñado o un estudio por observación, su tarea es seleccionar los tratamientos apropiados para responder a la hipótesis de investigación.

Muchas veces se requieren tratamientos adicionales para evaluar por completo las consecuencias de la hipótesis. Una componente importante de muchos diseños de tratamiento es el *tratamiento de control*, que se explica en la siguiente sección.

Los tratamientos de control son un punto de referencia

El **tratamiento de control** es un punto necesario para evaluar el efecto de los tratamientos experimentales; existen diversas circunstancias en las que el tratamiento de control es útil y necesario.

La condiciones del experimento pueden ser un obstáculo para la efectividad de tratamientos experimentales que, en general, han sido efectivos. Un *control* al que no se da *tratamiento* revelará las condiciones en las que se efectuó el experimento. Por ejemplo, los fertilizantes con nitrógeno suelen ser efectivos, pero no producirán respuestas en campos con alta fertilidad. Un *control de fertilizante sin nitrógeno* señalará las condiciones básicas de fertilidad del experimento.

En ocasiones, los tratamientos requieren la manipulación de las unidades experimentales o sujetos cuando esa sola manipulación puede producir una respuesta. En estos casos, los controles *placebo* establecen una base para la efectividad del tratamiento. La unidad o sujeto placebo se procesa de la misma manera que las unidades en tratamiento, pero sin incluir en su protocolo el tratamiento activo. Uno de los experimentos de salud más famosos, las pruebas de campo en 1954 de la vacuna Salk contra la poliomielitis, usó controles placebo o inocuos en cerca de la mitad de las áreas de prueba en Estados Unidos. El placebo se preparó con una apariencia igual a la de la vacuna, pero sin la actividad antigénica. Los sujetos placebo se inocularon de la misma forma que los sujetos que recibieron la vacuna (Tanur *et al.*, 1978).

Por último, el control puede representar una práctica *normal* con la que se puede comparar el método experimental. En algunas situaciones es necesario incluir dos tipos de control distintos. Por ejemplo, el no tratamiento y el tratamiento placebo pueden indicar el efecto de manipular la unidad experimental en ausencia de tratamiento.

El diseño de tratamientos con múltiples factores amplía las inferencias

En el artículo “The Arrangement of Field Experiments”(*Arreglo de los experimentos de campo*), Fisher (1926) observó que el proverbio más repetido respecto a los experimentos de campo era el que decía:

Debemos hacer algunas preguntas a la naturaleza, y de preferencia, una a la vez.

Fisher estaba convencido de que se trataba de un punto de vista equivocado. A este respecto escribió: “la naturaleza... responderá mejor a un cuestionario concebido lógica y cuidadosamente; en realidad, si le planteamos una sola pregunta, se rehusará a responder hasta que se analice algún otro tema”.

Él entendía que en los sistemas naturales no se sabe si la influencia de un tratamiento es independiente de otra o si su influencia se relaciona con la variación de otros tratamientos. En consecuencia, las condiciones en las que se comparan los tratamientos pueden ser aspectos importantes del diseño.

Por ejemplo, en un estudio de la cantidad de nitrógeno que la bacteria *Risobium* fija en el suelo, una comparación interesante fue la cantidad de nitrógeno producida en suelos salinos, sódicos y normales. Sin embargo, se sabía que las condiciones de temperatura y humedad en el suelo afectaban la producción de nitrógeno. De hecho, las condiciones óptimas de temperatura y humedad bien podían ser diferentes para cada tipo de suelo. En consecuencia, el experimento se estableció para examinar la producción de nitrógeno a varias temperaturas, combinadas con distintas condiciones de humedad para cada suelo. Los diseños del tratamiento en este tipo de experimentos se conocen como **tratamientos con diseños factoriales**, en el que un conjunto de tratamientos, digamos suelos, se prueba con uno o más conjuntos de tratamientos, como humedad y temperatura.

Un **factor** es un grupo específico de tratamientos: temperatura, humedad y tipos de suelo se consideran un factor cada uno. Las diversas categorías de un factor se denominan *niveles* del factor. Los niveles de temperatura son 20°C, 30°C

y 40°C; mientras que los niveles del tipo de suelo son normal, salino y sódico. Un *factor cuantitativo* tiene niveles asociados con puntos ordenados en alguna escala de medición, como temperatura. Los niveles de un *factor cualitativo* representan distintas categorías o clasificaciones, como tipo de suelo, que no se pueden acomodar conforme alguna magnitud.

El arreglo factorial consiste de todas las combinaciones posibles de los niveles de los factores de tratamiento. Por ejemplo, el arreglo factorial de tres niveles de temperatura con tres niveles de tipo de suelo consiste en $3 \times 3 = 9$ combinaciones de tratamiento factorial, que son:

(20°C, normal)	(30°C, normal)	(40°C, normal)
(20°C, salino)	(30°C, salino)	(40°C, salino)
(20°C, sódico)	(30°C, sódico)	(40°C, sódico)

Con este arreglo, el investigador pudo evaluar la producción de nitrógeno en cada tipo de suelo y determinar la influencia de las condiciones de temperatura en la producción relativa por tipo de suelo. Además, fue posible evaluar por separado la influencia de la temperatura en la producción de nitrógeno. El análisis estadístico de los arreglos factoriales se estudiará en los capítulos 6 y 7.

1.5 Control local de errores experimentales

Los principales objetivos de la mayoría de los experimentos son las comparaciones claras y exactas entre tratamientos a través de un conjunto apropiado de condiciones. Estos objetivos requieren estimaciones precisas de las medias y poderosas pruebas estadísticas; al reducir la varianza del error experimental aumenta la posibilidad de lograrlos. El **control local** describe las acciones que emplea un investigador para reducir o controlar el error experimental, incrementar la exactitud de las observaciones y establecer la base de la inferencia de un estudio.

El investigador controla: 1) técnica, 2) selección de unidades experimentales, 3) bloquización o aseguramiento de la uniformidad de información en todos los tratamientos, 4) selección del diseño experimental y 5) medición de covariados. A continuación se analizará cada aspecto con más detalle.

La técnica afecta la variación y el sesgo

Si las tareas experimentales se realizan sin una planeación cuidadosa, las observaciones mostrarán un incremento en la variación. Las *técnicas* incluyen tareas sencillas como una medición exacta, la preparación de medios, obtención de soluciones o calibración de instrumentos. En un nivel más complejo, el investigador puede contar con varios métodos o instrumentos de laboratorio para medir propiedades químicas o físicas. Los métodos pueden variar en exactitud, precisión y alcance de aplicación. El investigador debe elegir el método o instrumento que le proporcione las observaciones más precisas dentro de los recursos presupuestados.

Cuando la técnica tiene un efecto adverso en la precisión, las varianzas estimadas del error experimental se incrementan de manera innecesaria. Las observa-

ciones lejanas de la media, causadas por el registro de errores o por condiciones ambientales extremas, pueden aumentar la variación. Cualquiera que sea la causa, el investigador debe decidir si incluir o no estas observaciones en el análisis.

Por otro lado, tal vez no exista un patrón discernible en las observaciones distinto a un incremento general en su variabilidad. Este tipo de aumento en la variación puede señalar una técnica equivocada en el curso del experimento. El investigador deberá verificar la selección de las unidades experimentales para el estudio, el protocolo del tratamiento, las técnicas de medición y el personal, en busca de fuentes de error y luego intentar hacer los ajustes necesarios.

Las técnicas pobres pueden afectar la exactitud de las observaciones y sesgar los resultados. Más aún, la variación introducida en las observaciones por las técnicas pobres, no necesariamente es aleatoria y por lo tanto, no está sujeta a las mismas leyes de probabilidad que se asocian con la inferencia estadística.

La aplicación uniforme de los tratamientos durante el experimento aumenta la probabilidad de mediciones no sesgadas de sus efectos. Por ejemplo, se requiere la ingestión de cantidades uniformes de alimento para medir con exactitud las diferencias entre las dietas de los animales. Es necesario aplicar cantidades uniformes de fertilizante a las parcelas para medir con precisión las diferencias en la cosecha debidas a éste.

Selección de unidades experimentales uniformes

Las unidades experimentales heterogéneas producen valores grandes en la varianza del error experimental. La comparación precisa entre los tratamientos requiere la selección de unidades experimentales uniformes para reducir el error experimental. Sin embargo, una selección demasiado restringida puede producir condiciones de uniformidad artificiales. Un conjunto estrecho de condiciones restringe la base de inferencia del estudio. Entonces, para asegurar una confiabilidad razonable de las conclusiones del experimento, es deseable que las unidades representen una variedad suficiente de condiciones, sin que aumente innecesariamente la heterogeneidad de las unidades experimentales.

La naturaleza del experimento señala el equilibrio entre la variedad de las condiciones y la uniformidad de las unidades. Por ejemplo, la selección de plantas en un estudio de su crecimiento se puede examinar en el conjunto de condiciones en el cual se espera cultivar las variedades; así, las condiciones pueden ser muy diversas. Si las selecciones se examinan en varios sitios lejanos entre sí, la uniformidad en cada lugar cobra importancia. La uniformidad de las unidades en un experimento con vacas lecheras requiere elegir vacas de la misma cría, en la misma etapa de lactancia y con un número similar de lactancias.

La segregación en bloques (bloquización) para reducir la variación del error experimental

Fisher (1926) aseguró que no había que desistir de alguna ventaja para obtener una estimación válida del error, pero que eran necesarias dos cosas:

- a) hacer una distinción clara entre las componentes del error que debían eliminarse en el campo y las que no debían eliminarse... b) modificar el proceso estadístico de la estimación del error de manera que los componentes del error que se eliminaron en el campo sean igualmente eliminados en el laboratorio estadístico.

La *bloquización* proporciona control local del ambiente para reducir el error experimental. Las unidades experimentales se agrupan de manera que su variabilidad dentro de los grupos sea menor que entre las unidades antes de agruparlas. La práctica de bloquizar o agrupar las unidades experimentales en conjuntos homogéneos, va de la mano con la selección de las unidades experimentales para tener uniformidad. Los tratamientos se comparan entre sí dentro de los grupos de unidades en un entorno más uniforme y las diferencias entre ellos no se confunden con las grandes discrepancias entre las unidades experimentales. En el análisis estadístico es posible separar el error experimental de la variabilidad asociada con las diferencias del entorno entre los grupos de unidades.

Las unidades experimentales se bloquizan en grupos de unidades similares, con base en un factor o factores que se espera o se sabe que tienen alguna relación con la variable de respuesta o con la medición que se supone responde de manera diferente a los diversos tratamientos.

Cuatro importantes criterios para la bloquización

Los cuatro criterios que se usan con más frecuencia para bloquizar unidades experimentales son: 1) proximidad (parcelas vecinas), 2) características físicas (edad o peso), 3) tiempo, 4) administración de tareas en el experimento.

La práctica de bloquización clásica se originó en los experimentos agrícolas; en los que las parcelas contiguas pertenecían a un grupo y cada uno de los tratamientos se asignaba a una parcela de ese grupo. Después, se usaba el segundo grupo de parcelas contiguas de la misma manera, y así sucesivamente, hasta tener un diseño de *bloquización completa*.

El motivo principal para este tipo de bloquización es que las parcelas cercanas en campos de cultivo se parecen más que las separadas por mayores distancias. Algunos patrones de variabilidad pueden requerir arreglos diferentes para reducir el error experimental.

Otra unidad de bloquización natural está determinada por la camada de los animales. El tamaño de la camada en algunas especies permite asignarle varios tratamientos. Se aprovecha el peso como un factor de bloquización en experimentos con animales si su variación se amplifica en la variable de respuesta.

Los experimentos industriales requieren lotes homogéneos de materia prima. La replicación de un experimento puede necesitar más materia prima que la proporcionada por un lote y la variación de un lote a otro puede aumentar el error experimental. Un lote lo suficientemente grande para una réplica de todos los tratamientos puede servir como unidad de bloquización.

La bloquización se usa para dividir el experimento en unidades de tamaño razonable y administrar de manera uniforme el tiempo o las tareas. Los días son

unidades de bloquización convenientes sólo si se puede cultivar en el campo o procesar en el laboratorio una réplica de los tratamientos durante un solo día.

Los técnicos pueden servir como unidades de bloquización individual para evitar confundir la variabilidad del observador o técnico con la de los tratamientos. Puede asignarse a cada persona una réplica de cada tratamiento, cuando se dispone de varias para el registro de datos o el análisis del laboratorio.

Una demostración de la reducción de la varianza mediante bloquización

Una **prueba de uniformidad** muestra la posible efectividad que puede tener la bloquización en la reducción de la varianza dentro de un estudio de investigación. En esencia, la prueba de uniformidad es un experimento en el que las unidades experimentales se miden sin someterlas a un tratamiento. Por ejemplo, una prueba de uniformidad clásica en agricultura es un campo de trigo de la misma variedad, dividido en parcelas con las mismas dimensiones. Se mide la cosecha de trigo en cada parcela. Como la variación en los campos agrícolas normalmente ocurre en gradientes, se determina qué grupos de parcelas adyacentes tienen la menor varianza. En los experimentos de los años siguientes es posible asignar los tratamientos en grupos de parcelas similares, según los resultados de la prueba de uniformidad.

De manera similar, en la medida de interés o alguna variable conocida que se sabe que tiene una relación estrecha con esa medida, la observación de la línea base previa a la aplicación de tratamientos es equivalente a una prueba de uniformidad para propósitos de bloquización. Por ejemplo, las medidas de peso, edad, composición química o niveles de colesterol previas a la administración del tratamiento, pueden ser adecuadas como criterio de bloquización si tienen una relación estrecha con la medida de interés o son en sí una medida de interés.

Supongamos que en una prueba de uniformidad se tienen observaciones de diez unidades o de mediciones previas al tratamiento: 43, 72, 46, 66, 49, 68, 50, 76, 42 y 69. La media y la varianza de las observaciones en estas diez unidades son $\bar{x} = 58$ y $s^2 = 175$. Al agrupar las unidades en dos bloques según el tamaño de las que se midieron, se tiene

$$\begin{array}{lll} \text{Bloque 1: } & 43, 46, 49, 50, 42 & \bar{x} = 46 \quad s^2 = 12.5 \\ \text{Bloque 2: } & 72, 66, 68, 79, 69 & \bar{x} = 70 \quad s^2 = 15.2 \end{array}$$

Mientras que la varianza total entre las diez unidades era 175, en cada bloque se redujo a 12.5 y 15.2, respectivamente. El componente de error eliminado por la bloquización se reflejará en la varianza entre las medias de los dos bloques, 46 y 70. Se supone ahora que la variabilidad dentro de los bloques representa la variación natural que existe entre las unidades experimentales relativamente uniformes, sin la restricción de las diferencias ambientales controlables. Del mismo modo, la comparación entre los tratamientos dentro de esos bloques no tendrá influencia de esas diferencias ambientales controlables.

Estrategias de asignación por compatibilidad

El agrupamiento de las unidades a menudo usa estrategias de compatibilidad para equiparar unidades similares. Los sujetos o unidades se eligen para cada tratamiento por su similitud en todos los factores que pueden influir. Cada variable que pueda influir en el valor de las características a observar en los sujetos es un candidato como variable a controlar. Los sujetos asignados tienen valores comunes para las variables a controlar, con la excepción del tratamiento que se administra.

Durante el diseño del estudio, las estrategias para asignar intentan lograr que los sujetos sean similares en todos los factores que pueden sesgar en exceso la comparación de los tratamientos. *La asignación por compatibilidad uno a uno* y *compatibilidad por clase* son dos estrategias generales empleadas para la elección de unidades o sujetos.

Con la estrategia por compatibilidad uno a uno (por parejas), para cada sujeto de cada tratamiento se identifican parejas con sujetos que tienen los mismos valores de las variables a controlar y se asignan a cada uno de los otros tratamientos. Los valores de las variables a controlar pueden elegirse como: 1) *valores exactos* o 2) *valores calibrados*. Es posible encontrar valores exactos con sujetos humanos en variables como género, profesión, uso del cinturón de seguridad y nivel de educación. Los valores exactos en un estudio del desempeño del tránsito en tramos de carretera son factibles en variables como número de carriles, ancho de los mismos, presencia o ausencia de carril central y límites de velocidad.

Los valores calibrados permiten cierta tolerancia en los valores de las variables comparadas. Es posible que un estudio sobre la manipulación de un ecosistema boscoso requiera que los sitios sean compatibles en composición de especies, pendiente y aspecto. Es difícil lograr la compatibilidad exacta entre estos factores. Sin embargo, quizás se consigan sitios con valores similares para algunas o todas las variables de control comparadas, al grado de que la variación no tenga efectos serios en la comparación de los tratamientos.

Una compatibilidad por clase se puede lograr a través de una estrategia de: 1) *frecuencias* o de 2) *medias*. El método de frecuencia estratifica las unidades en grupos con base en las variables de control. Supongamos que la edad es una variable de influencia potencial en un estudio con seres humanos. Estos sujetos se pueden estratificar a partir de una distribución de frecuencias para sus edades, de manera que haya un número suficiente de sujetos en cada estrato de edad para asignarlos a todos los tratamientos.

La estrategia de compatibilidad por clase basada en las medias agrupa a los sujetos o unidades de manera que tengan los mismos valores promedio de las variables de control en cada grupo de tratamiento. Los animales experimentales se pueden agrupar cuando su peso promedio es el mismo para todos los tratamientos.

La naturaleza de las investigaciones señala la estrategia de compatibilidad más efectiva y si ésta es el protocolo deseable. Los detalles de los métodos de compatibilidad, sus ventajas y desventajas para estudios por observación comparativos, se pueden encontrar en Cochran (1983) y Fleiss (1981).

El diseño experimental asigna los tratamientos

El **diseño experimental** es el arreglo de las unidades experimentales utilizado para controlar el error experimental, a la vez que acomoda los tratamientos. Existe en la literatura una amplia variedad de arreglos diseñados para controlar el error experimental y se observa una tendencia natural a diseñar los experimentos de acuerdo con diseños ya existentes. Pero desarrollar un diseño de experimento que satisfaga la demanda del experimento que se está realizando es una actitud más adecuada.

El logro de la máxima información, precisión y exactitud en los resultados, junto con el uso más eficiente de los recursos existentes, es un principio a seguir en la elección del diseño adecuado del experimento.

La asociación entre la asignación de tratamientos y el diseño experimental se ilustra en dos contextos diferentes. La primera ilustración muestra cómo se asocian tres tratamientos con seis unidades experimentales, donde cada tratamiento se asigna a dos unidades. La segunda muestra la asignación de tres tratamientos a las seis unidades experimentales, después de bloquizarlas en dos conjuntos de tres unidades homogéneas.

Diseño de experimentos sin bloquizar

Consideremos un experimento para comparar los efectos de tres aditivos de gasolina, según la emisión de monóxido de carbono. Los motores de automóvil se usan como unidades experimentales y cada aditivo de gasolina se debe usar en dos motores. La figura 1.1 muestra una representación esquemática del diseño. Los cuadros representan motores como unidades experimentales. Se administra uno de los aditivos a un motor independientemente de los demás motores en el experimento. Los tres tratamientos con aditivo se asignan al azar a los seis motores, dos unidades para cada tratamiento.

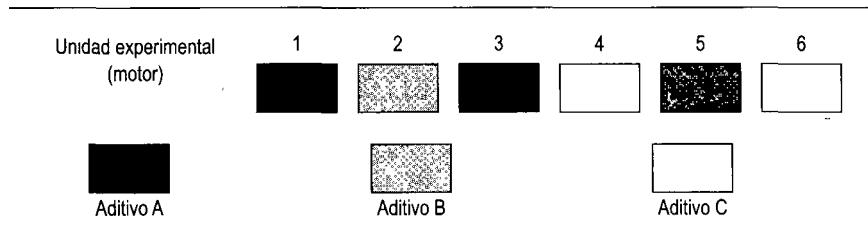


Figura 1.1 Ilustración de un diseño totalmente aleatorio, con tres tratamientos, con dos unidades experimentales cada uno

Este diseño, conocido como **diseño totalmente aleatorizado**, es el más sencillo. Los tratamientos se asignan a las unidades experimentales al azar. Cada unidad experimental tiene la misma posibilidad de recibir cualquier tratamiento. La función de aleatorización en el diseño de experimentos se estudia en la sección 1.8.

Experimentos con un solo criterio de bloquización

El diseño totalmente aleatorio proporciona poco control sobre la variación del entorno; varias clases de diseños genéricos emplean bloquización o el agrupamiento de unidades experimentales para controlarla. El diseño más sencillo es la **bloquización totalmente aleatoria** con un solo criterio de bloquización. Este diseño usa una restricción única sobre la asignación aleatoria de los tratamientos a las unidades experimentales; todos los tratamientos deben ocurrir igual número de veces en cada bloque.

Consideremos el arreglo de las unidades en un diseño de bloqueo totalmente aleatorio en un experimento para estudiar los efectos de tres dietas en el crecimiento de ratones de laboratorio. Se eligieron tres ratones del mismo género entre dos camadas. Éstas se usaron como los bloques mostrados en la figura 1.2. Las tres dietas se asignaron a los ratones de cada cría al azar.

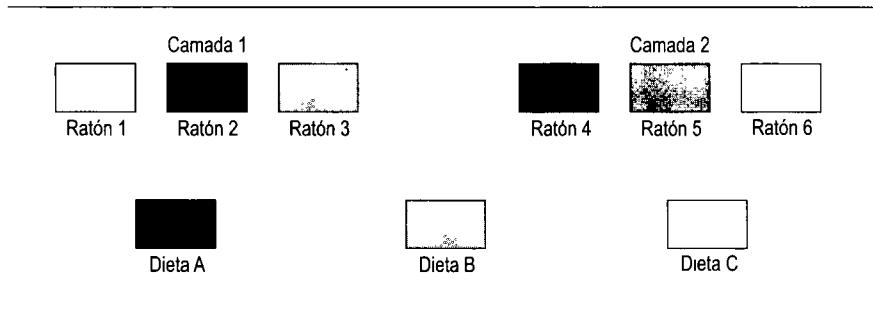


Figura 1.2 Ilustración de un diseño de bloques totalmente aleatorio, con tres dietas experimentales en dos crías de ratones

El uso de una camada como bloque reduce la variación del error experimental porque aísla la variación entre camadas de la variación entre las dietas. Éstas se pueden comparar en condiciones uniformes cuando cada una de ellas se prueba en la misma camada. Los detalles del diseño de bloques totalmente aleatorio y su análisis se presentan en el capítulo 8.

Covariadas para el control estadístico de la variación

Las **covariadas** son variables relacionadas con la variable de respuesta que nos interesa. La información de las covariadas se usa para establecer un control estadístico sobre la varianza del error experimental, mediante un procedimiento conocido como *análisis de covarianza*.

Anteriormente se sugirió que el peso de un cuerpo puede usarse como un criterio de bloqueo. Una vez agrupados los animales según su peso no se usan más esos valores. Pero el peso real se puede usar eficazmente para reducir las estimaciones del error en el modelo estadístico, en lugar o además de emplearlo para la

bloquización. El peso del cuerpo sería una covariada del aumento de peso en el experimento. Otros ejemplos de covariadas pueden ser las calificaciones prestablecidas, la fertilidad de las parcelas, la cosecha del año anterior en siembras perennes o la pureza de la materia prima en un proceso químico; todas pueden variar de una unidad a otra. Cualquier atributo medible y que pueda tener una relación estadística con la variable de interés primario, es un candidato para ajuste covariado.

El primer requisito es que el tratamiento específico usado en el experimento no afecte a la covariada. En la práctica, las covariadas se miden antes de aplicar los tratamientos, antes de que éstos tengan tiempo de provocar una respuesta en la unidad, o cuando se supone que el tratamiento nunca tiene efectos sobre la covariada.

Las observaciones en el experimento consisten en pares de observaciones (x, y) en cada unidad experimental, donde y es la variable de interés en el experimento y x es la covariada. Supongamos que existen seis pares de observaciones para cada uno de dos grupos de tratamiento y que los datos son los de la figura 1.3.

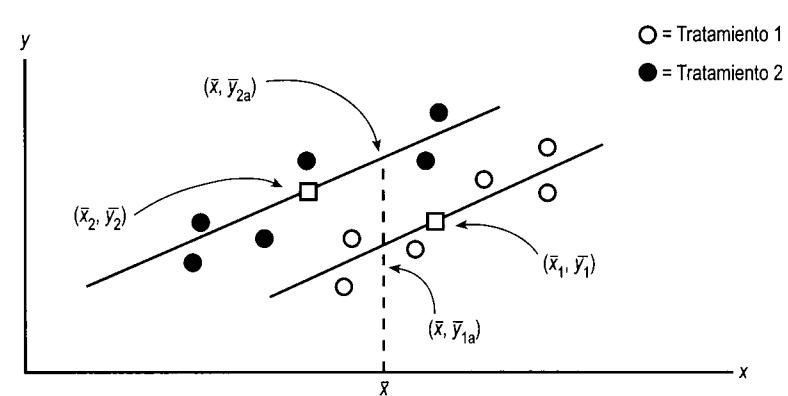


Figura 1.3 Ilustración de un ajuste covariado para dos grupos en tratamiento

Una parte de la variación en y está asociada con x y con los efectos del tratamiento si existe una relación estadística entre la variable de interés, y , y la covariada x . El análisis de covarianza estima la relación de regresión entre y y la covariada x , para reducir estadísticamente la varianza del error experimental. La respuesta promedio a cada tratamiento, \bar{y}_1 y \bar{y}_2 , se ajusta al mismo valor de la covariada, por lo general la gran media \bar{x} como se muestra en la figura 1.3. Una comparación entre las medias ajustadas de los tratamientos, \bar{y}_{1a} y \bar{y}_{2a} , elimina la influencia de la covariada x en la comparación. Por ejemplo, si el aumento en el peso tiene como covariada el peso inicial, la respuesta al tratamiento del aumento promedio se ajusta para eliminar la variación asociada con el peso inicial. Las medias ajustadas del aumento de peso representan el aumento que se obtendría si todos los animales tuvieran el mismo peso inicial. El análisis de covarianza se estudia en el capítulo 17.

1.6 Replicar para obtener experimentos válidos

La comunidad científica considera la réplica de un experimento como el primer requisito para obtener resultados experimentales válidos. La **réplica** implica una repetición independiente del experimento básico. Dicho de manera más específica, cada tratamiento se aplica de manera independiente a dos o más unidades experimentales.

Existen diversas razones para hacer réplicas de un experimento, las más notables son:

- Demuestra que se pueden reproducir los resultados, al menos bajo las condiciones experimentales actuales.
- Proporciona cierto grado de seguridad contra resultados anormales en el experimento, debidos a accidentes no previstos.
- Proporciona las medias para estimar la varianza del error experimental. Aun cuando la experimentación previa proporcione estimaciones de la varianza, la estimación a partir del experimento en curso puede ser más exacta porque refleja el comportamiento actual de las observaciones.
- Proporciona la posibilidad de aumentar la precisión en la estimación de las medias de los tratamientos. Al incrementar las réplicas, r disminuye $s_{\bar{y}} = s^2/r$, lo que aumenta la precisión de \bar{y} .

Las unidades de observación y las experimentales pueden ser claramente distintas

La *unidad de observación* puede no ser equivalente a la unidad experimental. La primera puede ser una muestra de la última, como muestras individuales de plantas de una parcela o muestras del plasma de un sujeto.

La varianza en las observaciones de las unidades experimentales es la varianza del error experimental; es una medida válida de la variación entre las unidades experimentales que recibieron tratamiento de manera *independiente*.

Muchas veces la varianza entre las observaciones múltiples de la misma unidad experimental se usa equivocadamente como medida del error experimental para comparar los tratamientos. Los siguientes ejemplos quizás aclaren la diferencia entre las réplicas de unidades experimentales y las observaciones múltiples de la misma unidad.

Ejemplo 1.1 Un estudio sencillo de raciones de dieta animal tiene un corral con seis animales asignados a la ración A y otro con seis más asignados a la ración B. Se reúnen datos sobre el aumento de peso o algún otro aspecto adecuado para examinar la eficacia de las raciones. Al final del estudio, se toman de cada animal las medidas necesarias. El esquema en la figura 1.4 ilustra el diseño.

Por lo general, se puede utilizar la prueba t de Student, que usa la diferencia entre las medias de los dos corrales, para probar la hipótesis de que no hay diferencia entre las raciones. Sin embargo, las diferencias entre los dos corrales tal vez se

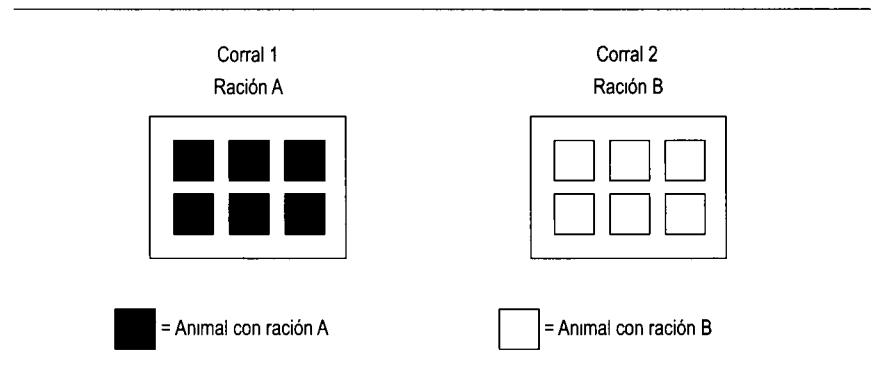


Figura 1.4 Ilustración de un experimento sin réplica

deban al efecto de otros factores además de los tratamientos. La respuesta a una ración dada puede variar animales de diferentes corrales. Es posible que esta variación se deba al efecto de uno o más factores que contribuyen al error experimental.

La naturaleza de la variación en la respuesta de un corral a otro, o *efecto del corral*, contribuye al error experimental. Además, la variación en la preparación y presentación del tratamiento en cada corral puede causar una interacción tratamiento-corral. Por lo tanto, cualesquiera diferencias entre las dos raciones del ejemplo 1.1 no se pueden atribuir únicamente a las raciones. Las diferencias se pueden atribuir a combinaciones de los efectos de los tratamientos, de los corrales y de las interacciones entre ambos.

El experimento no resolvería sin duda la pregunta de si difieren las dos raciones en su efecto sobre el aumento de peso. El experimento sólo tiene una réplica verdadera. El *corral* es la *unidad experimental* porque ésa es la unidad a la que se administró el tratamiento independiente. Los *animales* en el corral son las *unidades de observación*. La varianza calculada entre las observaciones de los animales dentro de los corrales es sólo una estimación del error en las observaciones y no una estimación de la varianza entre las unidades experimentales.

Ejemplo 1.2 Supongamos que el experimento del ejemplo 1.1 se reestructura de manera que los animales se dividen de manera aleatoria en cuatro corrales de tres animales cada uno. Más aún, las dos raciones se asignan al azar a los dos corrales como se muestra en la figura 1.5. Las raciones se administran en forma independiente a cada corral.

En el ejemplo 1.2 cada ración tiene una réplica adecuada. Las *unidades experimentales* son los *corrales* en los que se administran raciones independientes (dos corrales por ración) y los *animales* dentro de los corrales siguen siendo las *unidades de observación*. Entonces, la respuesta de la unidad experimental es el promedio de la respuesta del corral.

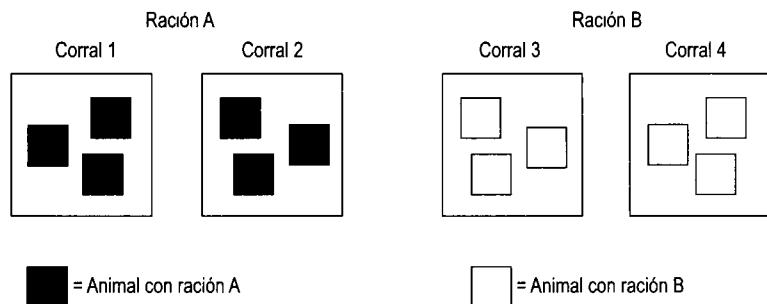


Figura 1.5 Ilustración de la reproducción de un experimento

La estimación de la varianza del error experimental, s^2 , calculada como la varianza entre las medias de los corrales dentro de cada ración, es la varianza adecuada para la prueba t de Student. La varianza entre los animales de cada corral, digamos s_w^2 , es una medida de la variabilidad en las unidades en observación dentro de las réplicas de corrales.

Así, en este tipo de estudio están presentes dos niveles de variación: 1) la variación entre las unidades de observación *dentro* de cada unidad experimental (s_w^2) y 2) la variación *entre* las unidades experimentales (s^2).

Es importante distinguir qué unidades del estudio constituyen la unidad experimental y en consecuencia, qué unidades constituyen una verdadera réplica del tratamiento.

Addelman (1970) presenta más detalles de los modelos y el análisis estadísticos para la seudorréplica y la réplica verdadera mostrados en los ejemplos 1.1 y 1.2. En Addelman Nelson y Rawlings (1983) se encuentran ejemplos de seudorréplicas en estudios de agronomía. Hulbert (1984) proporciona numerosos ejemplos de seudorréplicas en experimentos en campos ecológicos.

1.7 ¿Cuántas réplicas?

El número de réplicas en un estudio de investigación afecta la precisión de las estimaciones de las medias de los tratamientos y la potencia de las pruebas estadísticas para detectar las diferencias entre las medias de los grupos en tratamiento. Sin embargo, el costo de conducir estudios de investigación restringe las réplicas a un número razonable. Entonces, el número de réplicas está determinado por las restricciones prácticas que se pueden asignar al problema.

Número de réplicas para pruebas de hipótesis

El método para determinar el número de réplicas con frecuencia se basa en un examen de la hipótesis sobre las diferencias entre las medias de los grupos en

tratamiento. Aquí se utiliza un método elemental para los experimentos con dos muestras independientes, para ilustrar algunos atributos del problema del número de réplicas.

Este método se basa en una prueba de hipótesis acerca de las diferencias entre las medias de dos grupos de tratamiento $d = m_i - m_j$, con una varianza de error experimental conocida s^2 , utilizando la prueba estadística de distribución normal. Este método determina el número de réplicas necesario para probar la diferencia entre dos medias muestrales con errores especificados tipo I y tipo II.

El número de réplicas necesario está influido primordialmente por cuatro factores que se requieren para los cálculos:

- la varianza (σ^2)
- el tamaño de la diferencia (que tiene un significado físico) entre las dos medias (δ)
- el nivel de significancia de la prueba (α), o la probabilidad del error tipo I
- la potencia de la prueba $1 - \beta$, o la probabilidad de detectar δ , donde β es la probabilidad del error tipo II.

El número de réplicas necesario para cada grupo en tratamiento, r , para alternativas bilaterales se estima mediante:

$$r \geq 2[z_{\alpha/2} + z_{\beta}]^2 \left(\frac{\sigma}{\delta} \right)^2 \quad (1.1)$$

donde $z_{\alpha/2}$ es una variable normal estándar excedida con probabilidad $\alpha/2$ y z_{β} está excedida con probabilidad β . Las probabilidades para la variable normal estándar se encuentran en la tabla 1 del apéndice.

Es posible estimar el número de réplicas si se conoce el coeficiente porcentual de variación, %CV. El %CV se sustituye por σ en la ecuación (1.1), donde %CV = $100(\sigma/\mu)$. La diferencia (δ) debe expresarse como porcentaje de la media global esperada del experimento, % δ = $100(\delta/\mu)$, en la ecuación (1.1).

La influencia del coeficiente de variación (%CV), la diferencia porcentual (% δ), la potencia ($1 - \beta$) y el nivel de significancia (α) sobre el número de réplicas necesario se muestra en la tabla 1.1. Aunque los valores en ésta sólo se aplican a dos muestras independientes, los efectos son similares en experimentos más complejos.

El número de réplicas necesario generalmente aumenta si:

- la varianza, %CV, o σ^2 , aumenta
- el tamaño de la diferencia entre dos medias, % δ o δ , disminuye
- el nivel de significancia de la prueba, α , disminuye
- la potencia de la prueba, $1 - \beta$, aumenta

Los valores calculados para el número de réplicas necesario son estimaciones y aproximaciones. Con frecuencia se determinan con base en las estimaciones de la varianza asociadas a estudios previos y no a las del estudio real que se usarán

Tabla 1.1 Número de réplicas necesario para un coeficiente de variación dado (% CV) y probabilidad ($1 - \beta$) de obtener una diferencia significativa de % δ entre dos medias de tratamiento, con una prueba bilateral a un nivel de significancia α

		$\alpha = .05$		$\alpha = .01$	
		% δ		% δ	
% CV	$1 - \beta$	10	20	10	20
5	.80	4	1	6	2
	.95	7	2	9	3
10	.80	16	4	24	6
	.95	26	7	36	9

para calcular los intervalos de confianza y pruebas de hipótesis. La determinación del número de réplicas para las aplicaciones del análisis de varianza se considerarán en los capítulos 2, 5, 6 y 7. Se dispone de programas de cómputo comerciales para determinar el número de réplicas para muchos tipos de diferentes estudios experimentales.

1.8 Aleatorizar para tener inferencias válidas

Al reconciliar de esta manera las dos necesidades de reducción del error y de una estimación válida del error,... ningún principio se encuentra comprometido en el menor grado. Un experimento admite una estimación válida del error, o no lo hace; lo haga o no, no depende del arreglo real de las parcelas, sino sólo del modo en que se llegó a ese arreglo (Fisher, 1926).

La réplica de un experimento proporciona los datos para estimar la varianza del error experimental. La bloquización proporciona un medio para reducir el error experimental. Sin embargo, las réplicas y la bloquización por sí solos no garantizan estimaciones *válidas* de la varianza del error experimental o de las comparaciones de tratamientos.

Fisher (1926) señaló que la sola **aleatorización** proporciona estimaciones válidas de la varianza del error para los métodos de inferencia estadística justificados para la estimación y pruebas de hipótesis en el experimento. La aleatorización es la asignación aleatoria de tratamientos a las unidades experimentales.

Razonamiento para la aleatorización

El análisis de datos a partir de un experimento supone que las observaciones constituyen una muestra aleatoria de una población con distribución normal. Esta suposición es aceptable para los estudios comparativos por observación que usan muestras aleatorias de las unidades de observación disponibles de las distintas poblaciones en tratamiento. Sin embargo, cuando se hizo una selección cuidadosa, controlada y supervisada durante el experimento, es cuestionable que las unidades experimentales se puedan considerar una muestra aleatoria.

Unidad:	1	2	3	4	5	6	7
Tratamiento:	B	B	A	A	B	A	A
Respuesta:	14	16	19	17	15	13	17

Si no hay diferencia en los efectos de los tratamientos A y B, entonces éstos son sólo etiquetas en las unidades experimentales y no afectan los resultados. Por ejemplo, si la hipótesis nula es cierta, la respuesta de la unidad 1 será 14 sin importar qué tratamiento se aplicó. Lo mismo se cumple para las otras unidades bajo la hipótesis nula.

Las etiquetas A y B (cuatro con A y tres con B) se pueden asignar a las siete unidades experimentales en $7!/4!3! = 35$ arreglos posibles. Éstos son los 35 experimentos posibles al asignar de manera aleatoria los tratamientos a las unidades.

Los 35 arreglos posibles se muestran en la tabla 1.2, junto con la diferencia entre las medias de los grupos ($\bar{y}_A - \bar{y}_B$) con base en las etiquetas asignadas a las unidades en cada arreglo. Esas diferencias ($\bar{y}_A - \bar{y}_B$) son 35 diferencias posibles para las 35 aleatorizaciones posibles si se cumple la hipótesis nula. Constituyen la distribución aleatoria bajo la hipótesis nula.

Tabla 1.2 Treinta y cinco arreglos posibles de cuatro tratamientos A y tres B, en siete unidades experimentales con una diferencia de medias $\bar{y}_A - \bar{y}_B$

Arreglo	Unidad: Respuesta:	1 14	2 16	3 19	4 17	5 15	6 13	7 17	$\bar{y}_A - \bar{y}_B$
1		A	A	B	B	A	A	B	-3.17
2		A	B	B	A	A	A	B	-2.58
3		A	B	B	B	A	A	A	-2.58
4		A	A	B	A	B	A	B	-2.00
5		A	A	B	B	B	A	A	-2.00
6		A	B	A	B	A	A	B	-1.42
7		A	B	B	A	B	A	A	-1.42
8		B	A	B	A	A	A	B	-1.42
9		B	A	B	B	A	A	A	-1.42
10		A	A	A	B	B	A	B	-0.83
11		A	A	B	A	A	B	B	-0.83
12		A	A	B	B	A	B	A	-0.83
13		B	B	B	A	A	A	A	-0.83
14		A	B	A	A	B	A	B	-0.25
15		A	B	A	B	B	A	A	-0.25
16		A	B	B	A	A	B	A	-0.25
17		B	A	A	B	A	A	B	-0.25
18		B	A	B	A	B	A	A	-0.25

(continúa en la siguiente página)

Tabla 1.2 (Continuación)

<i>Arreglo</i>	<i>Unidad:</i> <i>Respuesta:</i>	1 14	2 16	3 19	4 17	5 15	6 13	7 17	$\bar{y}_A - \bar{y}_B$
19		A	A	A	B	A	B	B	0.33
20		A	A	B	A	B	B	A	0.33
21		B	B	A	A	A	A	B	0.33
22		B	B	A	B	A	A	A	0.33
23		A	B	A	A	A	B	B	0.92
24		A	B	A	B	A	B	A	0.92
25		B	A	A	A	B	A	B	0.92
26		B	A	A	B	B	A	A	0.92
27		B	A	B	A	A	B	A	0.92
28		A	A	A	A	B	B	B	1.50
29		A	A	A	B	B	B	A	1.50
30		B	B	A	A	B	A	A	1.50
31		A	B	A	A	B	B	A	2.08
32		B	A	A	A	A	B	B	2.08
33		B	A	A	B	A	B	A	2.08
34		B	B	A	A	A	B	A	2.67
35		B	A	A	A	B	B	A	3.25

Consideremos ahora una hipótesis alternativa $H_a: \mu_A - \mu_B \neq 0$ a la hipótesis nula $H_0: \mu_A - \mu_B = 0$. El arreglo del experimento real es el 30, con diferencia de medias $(\bar{y}_A - \bar{y}_B) = 1.50$. Puede darse una diferencia absoluta de 1.50 o mayor en 13 arreglos. Con la hipótesis nula se da una diferencia de medias absoluta de 1.50 o mayor con una frecuencia de $\frac{13}{35}$, que conduce a un nivel de significancia de 0.37. Con base en los resultados observados del experimento, el arreglo 30, no hay razón para rechazar la hipótesis nula.

Con la suposición de una hipótesis nula cierta, $H_0: \mu_A - \mu_B = 0$, la prueba de aleatorización permite evaluar el estadístico de la prueba, a partir del experimento real, $(\bar{y}_A - \bar{y}_B) = 1.50$, contra los valores de $\bar{y}_A - \bar{y}_B$ del resto de los miembros de la población de 35 experimentos posibles.

Las pruebas de la teoría normal aproximan a las pruebas de aleatorización

Fisher (1935) demostró primero que las pruebas de teoría normal son una buena aproximación a las pruebas de aleatorización, siempre que se haya hecho una asignación aleatoria de los tratamientos a las unidades experimentales y los tamaños de las muestras sean razonablemente grandes. Las aproximaciones a las pruebas de aleatorización mediante las de teoría normal mejoran al incrementar el tamaño de la muestra. Se puede encontrar una guía práctica de pruebas de aleatorización para diversas situaciones experimentales en Edgington (1987) y Manly (1991).

Kempthorne (1952), Scheffé (1959), Mead (1988) y Hinkelmann y Kempthorne (1994) proporcionan tratamientos rigurosos de los modelos de aleatorización y las pruebas de significancia en el diseño de experimentos. Un análisis más detallado de aleatorización relacionada con la inferencia estadística se puede encontrar en Kempthorne (1966, 1975).

Se ha convertido en una práctica común describir los modelos estadísticos para los estudios experimentales, en términos de los modelos de teoría normal. Las formalidades en estos modelos son más directas que en los de aleatorización. Sin embargo, el uso de los modelos de teoría normal en experimentos se justifica sólo bajo la protección de la aleatorización.

Aleatorización restringida en circunstancias difíciles

El resultado de la aleatorización puede ser un arreglo de tratamientos no satisfactorio desde el punto de vista de la validez científica. Es posible obtener arreglos secuenciales, como AAA BBB CCC o ABC ABC ABC, en una asignación aleatoria. Sin embargo, la disposición secuencial de los tratamientos puede conducir a problemas de sesgo.

Hurlbert (1984) cuestiona el uso ciego de la aleatorización en estudios ecológicos de pequeña escala. En los estudios pequeños existe una alta posibilidad de obtener un esquema aleatorio con una marcada segregación de los tratamientos en tiempo y espacio. La segregación puede conducir a resultados ficticios que confundan los efectos del espacio con los del tratamiento. Hurlbert (1984) afirma que debe haber cierta separación de los tratamientos para evitar la segregación sistemática en los experimentos pequeños.

Yates (1948) y Youden (1956), en forma independiente, introdujeron la aleatorización restringida como solución al problema de malos patrones que puede generar la aleatorización completa. La aleatorización restringida omite ciertos arreglos de los tratamientos que, en opinión del experimentador, sean inaceptables para el estudio específico.

Los experimentos industriales pueden requerir un desmantelamiento elaborado del aparato experimental entre ciertos tipos de tratamiento. Youden (1972) ofrece ejemplos de experimentos industriales y de laboratorio en los que el costo de cambiar de un tratamiento a otro es mayor que las ventajas de una aleatorización completa.

Bailey (1986, 1987) analiza investigaciones recientes e históricas sobre temas de aleatorización restringida. Se han desarrollado esquemas para la asignación restringida de tratamientos que admiten el uso del análisis de varianza normal (Bailey, 1986; Ypuden, 1972).

1.9 Eficiencia relativa del diseño de experimentos

La **eficiencia relativa** mide la efectividad de la bloquización para reducir la varianza del error experimental en el diseño de experimentos. En la práctica, la eficiencia

relativa se mide para determinar la eficiencia del diseño usado *en realidad* respecto a otro diseño más sencillo que *pudo usarse* pero no se usó. Por ejemplo, la eficiencia de un diseño de bloques totalmente aleatorizado, se determina en relación con el diseño aleatorizado por completo.

La varianza de la media de un tratamiento $\sigma_{\bar{y}}^2 = \sigma^2/r$ es una medida de la precisión de las medias del tratamiento estimadas. Esta precisión se controla mediante la magnitud de σ^2 y el número de réplicas, r , que en cierto grado están bajo el control del investigador. Éste puede aumentar el número de réplicas para disminuir $\sigma_{\bar{y}}^2$ y aumentar la precisión en la estimación de la media. El investigador también puede intentar reducir σ^2 a través de varias actividades de control local (como bloquización), e incrementar con esto la precisión del experimento. En esta sección se analiza un método para medir la efectividad de la bloquización.

El uso de $\sigma_{\bar{y}}^2$ como medida de precisión proporciona un medio para comparar la precisión relativa de dos diseños experimentales. Suponiendo que un diseño tiene una varianza verdadera del error experimental de $\sigma_1^2 = 1$, y que un segundo diseño tiene un error de varianza de $\sigma_2^2 = 2$. El valor de $\sigma_{\bar{y}}^2 = \sigma^2/r$ será el mismo para ambos diseños, si el segundo tiene el doble de réplicas que el primero. Es decir, la varianza de la media del tratamiento en cada diseño es:

$$\text{Diseño 1: } \sigma_{\bar{y}_1}^2 = \frac{\sigma_1^2}{r_1} = \frac{1}{r_1}$$

y

$$\text{Diseño 2: } \sigma_{\bar{y}_2}^2 = \frac{\sigma_2^2}{r_2} = \frac{2}{r_2}$$

Las varianzas $\sigma_{\bar{y}_1}^2$ y $\sigma_{\bar{y}_2}^2$ serán iguales sólo si $r_2 = 2r_1$, o sea si el diseño 2 tiene el doble de réplicas que el diseño 1. Por lo tanto, el diseño 1 es más eficiente que el diseño 2 en cuanto al número de réplicas necesarias y con la misma precisión, para una estimación de la media del tratamiento.

En la práctica, no se conoce σ^2 para cada diseño y debe estimarse a partir de los datos. Además, los grados de libertad para la estimación de la varianza cambian con los diseños. En estas circunstancias, la precisión relativa de los dos diseños se determina según el concepto de *información* (Fisher, 1960). Fisher calculó la cantidad de información que proporciona la diferencia estimada entre dos medias respecto a la diferencia real entre las medias de las poblaciones. La información calculada con este concepto es:

$$I = \frac{(f+1)}{(f+3)} \frac{1}{s^2} \tag{1.2}$$

donde s^2 es la varianza estimada del error experimental con f grados de libertad. Si σ^2 se conoce, entonces $I = 1/\sigma^2$ y el coeficiente $(f+1)/(f+3)$ se sustituye por la unidad. Para cualquier reducción en la variabilidad, s^2 , existe un incremento acor-

de en la información que se tiene de la diferencia en las medias poblacionales. Tanto la precisión como la información aumentan al disminuir la variabilidad.

La *eficiencia relativa* de dos diseños de experimentos se define como la razón de la información en los dos diseños. Supongamos que:

$$I_1 = \frac{(f_1 + 1)}{(f_1 + 3)} \cdot \frac{1}{s_1^2} \quad \text{e} \quad I_2 = \frac{(f_2 + 1)}{(f_2 + 3)} \cdot \frac{1}{s_2^2}$$

son las medidas de información estimadas de los diseños 1 y 2, respectivamente. Su eficiencia relativa se estima como:

$$RE = \frac{I_1}{I_2} = \frac{(f_1 + 1)(f_2 + 3)}{(f_1 + 3)(f_2 + 1)} \cdot \frac{s_2^2}{s_1^2} \quad (1.3)$$

Si $RE = 1$, entonces la información de los dos diseños es la misma y cada diseño requiere el mismo número de réplicas para tener la misma varianza en las medias de los tratamientos σ_y^2 . Si $RE > 1$, entonces el diseño 1 es más eficiente que el 2. Por ejemplo, si $RE = 1.5$, entonces el diseño 2 requiere 1.5 veces el número de réplicas que el diseño 1 para tener la misma varianza en la media del tratamiento.

1.10 De los principios a la práctica: un caso de estudio

El diseño de un estudio médico experimental publicado por Moon *et al.* (1955), ilustra el proceso de llevar los principios a la práctica. Las componentes de diseño del estudio proporcionan ejemplos de cómo los principios de diseño de la investigación, cubiertos en este capítulo, ayudaron a preparar al investigador para establecer su hipótesis. A continuación se expone una breve descripción de los principales elementos para el diseño de la investigación de ese estudio. Se pueden encontrar otros detalles en la publicación misma.

El problema

El cáncer de piel sin melanoma, que incluye carcinomas de células basales (BCC squamous cell carcinoma) y escamosas (SCC basal cell carcinoma), es el tipo de cáncer más común. Los médicos residentes del estado de Arizona han detectado una incidencia de estos cánceres entre 3 y 7 veces superior que en el resto de la población general de Estados Unidos. Aunque el cáncer sin melanoma no amenaza la vida, tiene un elevado costo en tratamientos y horas hombre.

Una historia de queratosa actínica (AK), un tipo de lesión en la piel, se ha aceptado como indicio para identificar a los individuos con riesgo de contraer cáncer en la piel. En muchos casos, las lesiones AK se convierten en cáncer de la piel sin melanoma y, por lo tanto, se clasifican como lesiones premalignas. Aque-

llos individuos que tienen una historia de queratosa actínica, pero poco o ningún cáncer sin melanoma en la piel, se consideran con riesgo moderado. Además, las personas mayores con piel clara y una larga historia de exposición al sol, tienen mayor riesgo de desarrollar cáncer en la piel.

Los resultados de recientes estudios clínicos y de laboratorio sugieren que la vitamina A y otros retinoides tienen un efecto preventivo sobre el tejido epitelial, como el de la piel. Sin embargo, estos estudios incluyeron un número pequeño de individuos y no han producido estimaciones confiables del efecto de la vitamina A en la prevención primaria del cáncer en la piel humana.

Durante cinco años se llevará a cabo una prueba clínica en el sur de Arizona, para evaluar la efectividad de la vitamina A o algún complemento de retinol como restrictor del riesgo de contraer cáncer sin melanoma en la piel, en los individuos que ya tienen un riesgo moderado.

Hipótesis de investigación

Un complemento de retinol (vitamina A) reduce la incidencia del cáncer en la piel en individuos con riesgo moderado y una historia de por lo menos diez lesiones de queratosa actínica.

Diseño del tratamiento

Las consideraciones importantes para establecer la dosis de retinol (vitamina A) incluyeron la necesidad de elevar por encima de lo normal la ingestión diaria de retinol en la mayoría de los adultos y evitar una dosis que indujera los posibles efectos colaterales adversos relacionados con la ingestión excesiva de retinol. Era necesario un tratamiento inocuo como grupo de comparación con el grupo tratado; éstos eran individuos con las mismas características que los sujetos tratados y con el mismo protocolo de seguimiento en el curso del estudio. Los sujetos no sabrían a qué grupo de tratamiento estaban asignados para mantener a ambos grupos bajo el mismo régimen.

Tratamiento: Complemento dietético diario de 25,000 UI (Unidades Internacionales) de retinol en cápsulas, autoadministrado.

Placebo: Cápsula inerte, diaria, autoadministrada.

Medidas de interés

La hipótesis se refería a la relación del riesgo de contraer cáncer en la piel con los niveles de retinol. Por lo tanto, la medida de interés era si un sujeto desarrollaba cáncer en la piel durante el curso del estudio. El análisis podía considerar varios enfoques para probar la hipótesis. Éstos incluían si se desarrollaba un cáncer, resultado binario; cuántos cánceres se desarrollaban, variable de conteo; el tiempo transcurrido hasta el desarrollo del cáncer, que es una medida de tiempo para el evento usada en análisis de supervivencia. Cada enfoque podía extraerse mediante

un registro del tiempo que tomaba el desarrollo de un cáncer, si es que ocurría. Así, las variables medidas eran:

- el tiempo transcurrido hasta la aparición del primer carcinoma de célula basal (**BCC**, basal cell carcinoma) y cada uno de los subsecuentes, si ocurrían.
- el tiempo transcurrido hasta la aparición del primer carcinoma de célula escamosa (**SCC**, squamous cell carcinoma) y cada uno de los subsecuentes, si ocurrían.

Selección de sujetos con características comunes

Se requería que los sujetos de este estudio fueran representativos de la población de adultos saludables con un riesgo moderado de cáncer de la piel sin melanoma, deseosos de participar en el estudio y que en ese momento no ingirieran una cantidad excesiva de vitamina A en su dieta. Se investigó a más de 11,000 individuos para el estudio; cerca del 25% eran elegibles. A continuación se exponen algunos criterios de selección usados por los investigadores.

Los sujetos se reclutaron a través de consultas a dermatólogos y anuncios en los medios de comunicación. Podían ser hombres o mujeres con una historia de al menos diez lesiones de queratosa actínica con diagnóstico clínico, la más reciente durante el año previo. No debían tener más de dos ocurrencias anteriores de SCC o BCC y cero diagnósticos de cáncer diferentes a los SCC o BCC en el año anterior.

Los individuos elegibles debían tener entre 21 y 84 años, ser autosuficientes, capaces de moverse por sí mismos, sin diagnósticos de enfermedades que amenazaran su vida, con intención de residir en Arizona los siguientes cinco años y dispuestos a asistir semestralmente a la clínica para la visita de seguimiento. También debían estar dispuestos a limitar el complemento de vitamina A que no correspondiera al estudio a no más de 10,000 UI diarias. Además, debían estar dentro del 95% de los niveles normales de colesterol total, funcionamiento del hígado, conteo de leucocitos, hemoglobina y plaquetas.

Algunas técnicas para reducir el error no aleatorio y el sesgo

Para asegurarse de reducir al mínimo el error no aleatorio y el sesgo en las respuestas del estudio debían tomarse varias precauciones. Éstas incluyeron la seguridad de que los sujetos tomaran la medicina con regularidad, continuara su deseo de permanecer en el estudio, efectuaran sus visitas de evaluación y no conocieran a qué grupo de tratamiento estaban asignados.

Era de esperarse que si los sujetos sabían que pertenecían al grupo placebo, tal vez se autoadministraran vitamina A con la esperanza de reducir su riesgo personal e involuntariamente sesgaran la comparación con el tratamiento. Los investigadores se aseguraron de que esto no ocurriera al ocultar al sujeto y al clínico que suministraba las cápsulas, de qué tipo eran. Esto se conoce como prueba *doble ciega*.

Se estableció un periodo de tres meses para evaluar la capacidad y disposición de los individuos para adherirse al protocolo del estudio. Recibieron una botella con cien cápsulas inocuas, que deberían tomar una por día. Se asignaron a los

grupos de tratamiento o régimen inocuo los sujetos que habían tomado al menos el 75% de las cápsulas en ese periodo y deseaban continuar con el estudio, recibieron las cápsulas apropiadas para seis meses y programaron su visita de seguimiento.

Los sujetos regresaron a la clínica cada seis meses para ser examinados en cuanto a cualquier síntoma de BCC o SCC. Como medida de seguridad, también fueron examinados en busca de los posibles efectos colaterales causados por la ingestión elevada de retinol. El medicamento restante de los sujetos se pesaba para evaluar su adhesión a la dosis. Durante la entrevista, se respondía a la preguntas del sujeto y se le motivaba a que continuara con el programa. Luego recibían el suministro de cápsulas para los siguientes seis meses y se programaba su visita subsecuente. Cuando se aproximaba la cita, recibían un recordatorio por correo o teléfono.

Una vez al año se recolectaba una prueba de sangre de una muestra aleatoria de los sujetos, para analizar los niveles de palmitato de retinol a fin de obtener información adicional sobre la adhesión del grupo complemento de retinol. Si se seguía el régimen de ingestión de las cápsulas, los niveles de palmitato de retinol debían ser mayores en el grupo en tratamiento que en el grupo placebo.

Réplicas

La prueba se llevó a cabo en dos clínicas, una en Tucson y otra en Phoenix, Arizona. El número de sujetos necesarios para el estudio se basó en suposiciones sobre la incidencia promedio anual del cáncer en la piel en los grupos placebo y en tratamiento, y la ocurrencia esperada de sujetos que se salían de los protocolos prescritos por el estudio. El tamaño necesario de la muestra se determinó en 1,118 sujetos para cada grupo de tratamiento y se basó en una potencia de .80 y la tasa de error tipo I bilateral de .80 de .05, usando una técnica específica para estudios de medición de tiempo para evento.

Formación de bloques para reducir el error experimental

Se previó que el riesgo de un sujeto al cáncer de piel sin melanoma está relacionado con el tiempo que pasa bajo el sol y si tiene la piel clara. Cualquiera que pase más tiempo bajo el sol puede tener un riesgo mayor de contraer cáncer de piel. En el primer contacto, se recopiló información de cada sujeto en cuanto al tiempo de exposición semanal al sol y la reacción esperada de quemadura en la piel después de 30 minutos.

Éstos eran los factores de interferencia más probables con las comparaciones entre los grupos placebo y en tratamiento, por lo que se usaron como factores de bloqueo previos a la asignación de sujetos a los grupos. Los sujetos se clasificaron según dos niveles de exposición al sol: < 10 horas contra \geq 10 horas por semana. También se agruparon según los niveles de reacción de la piel después de 30 minutos: quemaduras siempre o casi siempre contra quemaduras en forma moderada, rara vez o nunca.

Así, los sujetos se clasificaron en uno de cuatro bloques construidos con dos factores en cada combinación de niveles. Por ejemplo, los sujetos expuestos al sol

< 10 horas semanales y cuya piel se quema en forma moderada, rara vez o nunca después de 30 minutos de exposición al sol, se colocaban en el mismo bloque. Se asignó un número igual de sujetos al tratamiento de retinol y placebo en cada bloque; entonces, cualesquiera posibles diferencias de riesgo de los sujetos, debidas a la exposición al sol y sensibilidad de la piel, no interferirían con las diferencias entre los grupos de tratamiento y placebo.

Aleatorización

Los individuos se involucran con el tiempo en pruebas clínicas conforme sus médicos los identifican como candidatos para el estudio y ellos responden al llamado como voluntarios. Por lo tanto, en ocasiones transcurre un año o más para la asignación de los sujetos a los tratamientos, hasta que se logra el número necesario de sujetos (réplicas). Para este estudio, el periodo de reclutamiento fue de más de cuatro años.

Los sujetos se asignaron al tipo de bloque adecuado (descrito antes) cuando se unieron al estudio. Se asignaron al azar al tratamiento placebo o con retinol en grupos de cuatro, con el mismo criterio de bloques. Por ejemplo, si los primeros dos sujetos de un bloque se asignaban al azar al grupo placebo, los siguientes dos recibían automáticamente el tratamiento de retinol. La aleatorización se iniciaba de nuevo para los siguientes cuatro sujetos de cualquiera de los bloques. Este método de asignación en bloques de cuatro sujetos aseguraba el mismo número de sujetos en cada tratamiento durante el periodo de reclutamiento de la prueba y eliminaba la posibilidad de tener un número mayor en un tratamiento y por más tiempo que en el otro.

La aleatorización se hizo por separado en las clínicas de Tucson y Arizona. De esta manera, las clínicas se convirtieron, de hecho, en un factor de bloque para el estudio.

Medidas covariadas para el control estadístico del error experimental

Se podía concebir que numerosos factores tuvieran alguna relación con el riesgo de desarrollar cáncer en la piel. Aquellos factores que se pensó que podían tener un mayor potencial para afectar la comparación de riesgo entre los grupos con retinol y placebo se usaron como factores de bloque: exposición al sol y sensibilidad de la piel a esa exposición. Incluir más factores en el esquema de bloques pudo haber causado tareas tediosas innecesarias, en particular si no se tenía evidencia que sugiriera que otros factores tuvieran un impacto importante en el riesgo de contraer cáncer en la piel.

Hubo otros factores que los investigadores consideraron que podían influir, por lo que se registraron datos como edad, género, cánceres de piel anteriores (0, 1 o 2), número de lunares y pecas, uso de vitaminas, vitamina A en la dieta al inicio del estudio (determinada en la entrevista), y palmitato de retinol en el plasma al inicio del estudio.

Estos factores pudieron usarse como covariadas en el análisis de datos para reducir el error experimental en la comparación de los grupos. Se realizó una verificación de la aleatorización después del periodo de cuatro años de reclutamiento

para determinar si los sujetos en los grupos tenían una distribución igual respecto a estos factores. La tabulación cruzada que se muestra en la tabla 1.3 indica una distribución muy parecida de los sujetos en los grupos de tratamiento y placebo para las covariadas más importantes consideradas en el estudio. Observe que hay 26 sujetos con más de 2 cánceres de piel previos incluidos en el estudio, aunque no cumplían con los criterios de selección de no más de dos. Por supuesto, puede deberse a errores de registro o a descuido de los clínicos que permitieron el reclutamiento. La razón de este tipo de errores sigue siendo un misterio.

EJERCICIOS PARA EL CAPÍTULO 1

1. a. Busque las definiciones de *investigación* e *hipótesis* en un diccionario y formule una definición del término *hipótesis de investigación*.
b. ¿En qué difiere la hipótesis de investigación de la *hipótesis estadística* formulada para las pruebas estadísticas como $H_0: \mu = 0$ contra $H_a: \mu \neq 0$?
2. Elija un artículo de una publicación en su campo de estudio que proporcione los resultados de un experimento comparativo o un estudio por observación. Identifique y describa brevemente (una o dos oraciones) lo siguiente:
 - a. Hipótesis de investigación
 - b. Tratamientos
 - c. Unidades experimentales (de observación)
 - d. Tipo de diseño de experimentos
 - e. Criterios para agrupamiento, bloquización o compatibilidad en el estudio
 - f. Proporcione la cita del artículo
3. a. Elija una situación práctica en su campo de estudio y describa un problema cuya solución deba determinarse en forma experimental
b. Indique lo siguiente para el problema descrito en el inciso a):
 - (i) Una hipótesis de investigación
 - (ii) Los tratamientos necesarios para evaluar la hipótesis
 - (iii) Qué constituye una unidad experimental
4. Elija un artículo de una publicación de su campo de interés y revíselo a fin de evaluar la aplicación de un buen diseño de investigación. Muchos aspectos del mismo se han presentado por separado en este capítulo en cuanto a su efecto en la inferencia estadística y científica. El caso de estudio en la sección 1.10 ilustra los elementos de un estudio de investigación con reportes, que deben identificarse para validar las conclusiones de un proyecto de investigación.

Seleccione un artículo sobre un experimento realizado para comparar dos o más tratamientos, o bien, un estudio comparativo por observación llevado a cabo para comparar dos o más condiciones de “tratamiento” ya existentes.

Algunas preguntas que debe contestar para hacer una crítica son: “¿Incluye este estudio todos los elementos importantes de un buen diseño de investigación?, si es así, ¿Los pusieron en práctica de manera adecuada?, ¿Se describen los elementos del trabajo de manera que yo pueda entender o duplicar lo que hicieron?”

En su revisión dedique especial atención a los siguientes aspectos:

- a. Revisión de la literatura.
- b. Establecimiento del problema.
- c. Hipótesis de investigación y objetivos del estudio: analice si se presentaron y eran razonables.
- d. Diseño de tratamiento: describa su relación con la hipótesis y los objetivos.
- e. Diseño de experimentos o del estudio por observación.
- f. La aleatorización (experimentos) o el muestreo aleatorio (estudios por observación) y las réplicas.
- g. Hipótesis estadística y procedimientos de análisis estadístico.
- h. Conclusiones y confiabilidad estadística de las mismas.
- i. Autoevaluación del estudio por el autor(es) y potencial para investigaciones futuras.
- j. Proporcione la cita del artículo.

Su crítica debe describir y evaluar el enfoque del autor, en la investigación misma y en el artículo, respecto a cada elemento importante de la investigación con base en la lista anterior. Incluya la referencia de la página del artículo a la que se refieren sus comentarios.

5. Se planea un estudio de la fisiología de los ejercicios con voluntarios humanos. Los dos tratamientos del estudio son dos métodos de entrenamiento para ejercicios aeróbicos (llamados métodos A y B). Al final de diez semanas de ejercicio, cada individuo normal será sometido a una prueba de las medidas respiratorias y cardiovasculares normales en una caminadora.

<i>Individuo</i>	<i>Sexo</i>	<i>Edad</i>	<i>Individuo</i>	<i>Sexo</i>	<i>Edad</i>
1	M	54	10	M	18
2	M	38	11	M	31
3	F	41	12	F	18
4	F	18	13	M	58
5	F	19	14	M	74
6	F	39	15	F	58
7	M	51	16	F	21
8	F	44	17	M	35
9	M	62	18	M	34
			19	F	38

La tabla presenta 19 voluntarios con su sexo y edad. Todos los voluntarios gozan de buena salud y están en el rango normal de peso para su edad, sexo y estatura. Se probarán ocho individuos en cada método, por lo que sólo se usarán 16 de los 19 voluntarios; un sujeto participará sólo en un método.

- a. Explique cómo agruparía a los individuos antes de asignarles el tratamiento, para mantener mínimo el valor de la varianza del error experimental.
 - b. Explique por qué agrupó de esa manera.
 - c. Muestre su asignación final de individuos en los grupos de tratamiento.
6. En una prueba de tratamiento de textiles con planchado permanente se planea un experimento para comparar tres tratamientos aplicados a telas, para producir camisas sin arrugas. En el pasado se usaba formaldehído para este tipo de telas, pero se consideró un tratamiento químico no deseable. Este estudio debe considerar tres químicos alternativos: a) PCA (ácido 1-2-3 propano fénico), b) BTCA (ácido tetracarboxílico de butano) y c) Ácido cítrico.

Se usarán cuatro camisas en cada tratamiento. Primero, se aplican los tratamientos a las camisas, que luego se someten a una simulación de uso y lavado en una máquina. Los tratamientos químicos no contaminarán a los otros si se colocan todos en la misma lavadora durante la prueba. La máquina puede manejar de una a cuatro camisas en una corrida de simulación. Al final de la corrida se mide la resistencia al rompimiento de la tela de cada camisa y si no tiene arrugas después de someterla a la simulación. La comparación entre los tratamientos puede verse afectada por: a) la variación natural de una camisa a otra, b) errores de medición, c) variación en la aplicación del tratamiento de planchado permanente, y d) variación en la corrida de simulación en la máquina de uso y lavado. A continuación se presenta una breve descripción de los tres métodos propuestos para realizar este sencillo experimento.

Método I. Las camisas se dividen al azar en tres grupos de cuatro camisas. Cada grupo recibe un tratamiento de planchado permanente como un lote y después cada uno se procesa en una corrida de simulación. Cada corrida tiene cuatro camisas que recibieron el mismo tratamiento. Se hacen tres corridas de simulación.

Método II. Las camisas se dividen al azar en tres grupos de tratamiento de cuatro camisas y se aplica el proceso de planchado permanente a cada camisa en forma independiente. Las camisas se agrupan en cuatro conjuntos de tres, una con cada tipo de tratamiento y cada conjunto se somete a una corrida en la máquina de simulación. Se realizan cuatro corridas de simulación.

Método III. Las camisas se dividen al azar en tres grupos de cuatro camisas. Se aplica el proceso de planchado permanente a cada camisa en forma independiente. La simulación de uso y lavado se realiza como en el método I.

- a. ¿Qué método favorecería usted?
 - b. ¿Por qué favorece ese método?
 - c. Explique en forma breve cuál es la desventaja de los otros dos métodos.
7. Explique qué significa el término *réplica* en el contexto de: a) un experimento en el que se examina la efectividad de varios antibióticos en animales de laboratorio y b) un estudio por observación para determinar las diferencias en especies de pasto presentes en cosechas de mesquite puras y en cosechas roble puras en el sur de Arizona.

8. Se planea un experimento para comparar tres métodos de enseñanza. Cada uno se aplica en un salón con 25 estudiantes. Se usará un instructor diferente en cada salón y en consecuencia para cada método.
 - a. Escriba una crítica corta del experimento propuesto.
 - b. ¿Cómo puede mejorarse el experimento?

9. Se planea un experimento para comparar la resistencia de tres mezclas de asfalto para carretera. Se fabricará un solo lote de cada mezcla. Se harán varios especímenes de asfalto de cada mezcla y se probará su fuerza de tensión.
 - a. Escriba una crítica corta del experimento propuesto.
 - b. ¿Cómo puede mejorarse el experimento?

10. Suponga que desea aleatorizar la asignación de dos tratamientos en 16 unidades experimentales. ¿Cuántas aleatorizaciones son posibles si deben asignarse 8 unidades a cada tratamiento?, ¿Cuántas si se quiere asignar 6 unidades a un tratamiento y 10 al otro?

11. Un experimento con cuatro tratamientos y cinco réplicas de cada uno requiere 20 unidades experimentales. ¿Cuántas aleatorizaciones son posibles para este experimento?

12. Un experimento con dos tratamientos y tres réplicas de cada uno tiene la siguiente aleatorización de tratamientos a las unidades experimentales y se muestra junto con la respuesta medida en cada unidad:

Unidad:	1	2	3	4	5	6
Tratamiento:	A	B	B	A	A	B
Respuesta:	7	10	9	5	10	12

Conduzca una prueba de aleatorización de la hipótesis nula, H_0 : no hay diferencia en el efecto de los tratamientos A y B, contra la alternativa, H_a : el efecto del tratamiento B es mayor que el del A. Utilice el estadístico de la prueba $\bar{y}_B - \bar{y}_A$. (*Sugerencia*: sólo es necesaria la identificación directa de la mitad de las aleatorizaciones. Cada una tiene una aleatorización “espejo” en la que las letras A y B se intercambian. Por ejemplo, el “espejo” de la aleatorización en la tabla anterior es B, A, A, B, B, A.)

13. En la tabla se muestran los coeficientes de variación y las eficiencias relativas (bloque totalmente aleatorizado contra aleatorización completa) del mismo experimento realizado en los cuatro sitios dados. Cada prueba usó un diseño de bloque totalmente aleatorizado.

Lugar	Coeficiente de variación (%)	Eficiencia relativa (%)
Tucson	10	100
Phoenix	10	150
Los Ángeles	20	200
San Francisco	20	125

36 CAPÍTULO 1 PRINCIPIOS DE DISEÑO EN INVESTIGACIÓN

- a. ¿Cuántas réplicas más de un diseño totalmente aleatorizado serán necesarias en Los Ángeles para obtener la misma precisión que en el diseño de bloque totalmente aleatorizado, al estimar las medias de los tratamientos? Explique su respuesta.
- b. Si se hiciera la misma pregunta del inciso a) respecto a San Francisco, ¿serían necesarias más o menos réplicas que en Los Ángeles? Explique su respuesta.
- c. Suponga que en Tucson se requieren cuatro réplicas en el diseño de bloque totalmente aleatorizado para detectar diferencias de $\delta = 20\%$ con una prueba con un nivel de significancia de 0.5 y una probabilidad (potencia) de .90. ¿Cuántas réplicas se requieren en Phoenix con el mismo criterio para un diseño de bloque totalmente aleatorizado? Explique su respuesta.
- d. ¿Se requerirán más o menos réplicas en los Ángeles que en Tucson con el mismo criterio para el diseño de bloque aleatorizado por completo? Explique su respuesta.

2 Comenzando con diseños totalmente aleatorizados

En el capítulo 1 se presentaron los principios del diseño experimental relacionados con las metas establecidas por la hipótesis de investigación (la exactitud y precisión de las observaciones y la validez de los resultados del análisis). En este capítulo se describen algunos de esos principios mediante un experimento con un diseño totalmente aleatorizado. Se desarrolla un modelo estadístico con parámetros que describen el experimento de acuerdo con la hipótesis de investigación y luego se estiman los parámetros con el método de mínimos cuadrados. Se calcula la varianza del error experimental y se utiliza para estimar los errores estándar y los intervalos de confianza para los parámetros y para probar las hipótesis estadísticas sobre ellos. Se deriva la partición fundamental para la suma de los cuadrados de las observaciones y se resume en la típica tabla para análisis de varianza.

2.1 Construcción del diseño de investigación

La *hipótesis de investigación, el diseño del tratamiento y el diseño del estudio experimental o por observación*, constituyen el **diseño de investigación** para el estudio. Los tratamientos se diseñan para resolver preguntas e hipótesis específicas que surgen en los programas de investigación. Por ejemplo, si un microbiólogo plantea la hipótesis de que la actividad de los microbios del suelo depende de las condiciones de humedad, se establecen tratamientos con distintos niveles de humedad para medir la actividad de los microbios y evaluar la hipótesis. Si un ingeniero de tránsito plantea la hipótesis de que la velocidad del tránsito se relaciona con el ancho de los carriles en las calles, para evaluar la hipótesis se seleccionan carriles con diferente anchura y se mide la velocidad de los automóviles en cada uno.

El diseño del tratamiento debe encontrarse dentro del diseño del experimento. El investigador debe decidir qué constituye una unidad experimental, cuántas réplicas de unidades experimentales exige cada tratamiento y qué tratamiento asignar a cada una de ellas. El investigador también debe determinar si agrupará por

bloques las unidades experimentales en grupos homogéneos para controlar el error experimental.

De manera similar, el estudio comparativo por observación que asocia la velocidad del tránsito con el ancho de los carriles, requiere que el investigador determine cuántas calles independientes con cada ancho exige el estudio y cómo agruparlas para controlar las variables.

Los pormenores de cálculo para el análisis estadístico de los distintos diseños pueden variar de un diseño a otro, aunque muchos procedimientos estadísticos usados en el análisis son comunes a la mayoría de los diseños existentes. Esto se debe a que los procedimientos mismos generalmente se relacionan con los diseños del tratamiento específico, cada uno de los cuales puede aparecer en varias configuraciones del diseño experimental.

La intención de este capítulo y el siguiente es introducir procedimientos estadísticos útiles para diversos estudios comparativos. En los capítulos subsecuentes, los procedimientos se extienden al diseño de tratamientos más complejos y se muestran sus aplicaciones en otros diseños que se van presentando. En este capítulo se ilustran primero los procedimientos con una clasificación unilateral de los tratamientos, en un diseño totalmente aleatorizado con igual número de réplicas.

Ejemplo 2.1 Supresión del crecimiento bacterial en carnes almacenadas

La vida de anaquel de las carnes almacenadas es el tiempo que un corte previamente empacado es sano, nutritivo y vendible. Un paquete normal expuesto al aire ambiental tiene una vida aproximada de 48 horas, después de las cuales la carne comienza a deteriorarse por contaminación de microbios, degradación del color y encogimiento. El empaque al vacío es efectivo para suprimir el desarrollo de microbios; sin embargo, continúan siendo un problema los otros aspectos.

Algunos estudios recientes sugieren las atmósferas controladas de gas, como alternativa a los empaques actuales. Dos atmósferas que prometen combinar la capacidad de suprimir el desarrollo de microbios con la conservación de las cualidades de la carne son: 1) dióxido de carbono puro (CO_2), y 2) mezclas de monóxido de carbono (CO), oxígeno (O_2) y nitrógeno (N).

Hipótesis de investigación: Con base en esta nueva información, el investigador plantea la hipótesis de que alguna forma de atmósfera controlada proporcionará un entorno más efectivo de empaque para el almacenamiento de carne.

Diseño del tratamiento: El diseño del tratamiento desarrollado por el investigador para evaluar la hipótesis incluyó empaques con: 1) aire del ambiente con un empaque comercial de plástico; 2) al vacío; 3) una mezcla de gases con 1% CO, 40% O_2 , y 59% N y 4) 100% CO_2 . Los empaques con el aire del ambiente y al vacío sirven como tratamientos de control, ya que ambos son estándares con cuya efectividad se puede comparar la de los nuevos empaques.

Diseño del experimento: Se usó un diseño totalmente aleatorizado para el experimento. A cada conjunto de condiciones de empaque se le asignaron al azar 3

cortes del mismo tamaño (75 g). (El método de aleatorización se presenta en la sección 2.2.) Cada corte se empacó por separado en las condiciones asignadas. En este ejemplo, se evalúa la efectividad de cada tratamiento para suprimir el desarrollo bacterial. Después de 9 días de almacenamiento a 4°C en una instalación normal, se midió el número de bacterias siccotrópicas en la carne. Las bacterias siccotrópicas se encuentran en la superficie de la carne y se asocian con la carne deteriorada.

Los resultados se muestran en la tabla 2.1. El crecimiento bacterial se expresa como el logaritmo del número de bacterias por centímetro cuadrado.

Tabla 2.1 Bacterias siccotrópicas [$\log(N/cm^2)$] en muestras de carne almacenadas en cuatro condiciones de empaque, durante nueve días

<i>Condiciones de empaque</i>	<i>Bacteria siccotrópica</i>		
	<i>Log(N/cm²)</i>	<i>Total</i>	<i>Media</i>
Empaque comercial de plástico	7.66, 6.98, 7.80	22.44	7.48
Empaque al vacío	5.26, 5.44, 5.80	16.50	5.50
1% CO, 40% O ₂ , 59% N	7.41, 7.33, 7.04	21.78	7.26
100% CO ₂	3.51, 2.91, 3.66	10.08	3.36

Fuente: B. Nichols (1980), *Comparison of Grain-Fed and Grass-Fed Beef for Quality Changes When Packaged in Various Gas Atmospheres and Vacuum*, Tesis de Maestría, Departamento de Ciencias Animales, Universidad de Arizona.

2.2 Cómo aleatorizar

Aleatorización de tratamientos en el diseño de experimentos

Los cortes de carne usados para el experimento eran unidades relativamente homogéneas y se usó un diseño totalmente aleatorizado para evitar la asignación subjetiva de tratamientos a los cortes. El procedimiento adecuado de aleatorización para este tipo de diseño se ilustra con el estudio de almacenamiento de carne.

Paso 1. Asignar la secuencia de números del 1 al 12 a las unidades experimentales, los cortes de carne.

Paso 2. Obtener una permutación aleatoria de los números 1 al 12 y anotarlos en el orden de la permutación. Una permutación aleatoria se obtiene tomando una sucesión de números con dos o tres dígitos de una tabla de números aleatorios (tabla XII del apéndice) y ordenándolos de menor (1) a mayor (N). Los números ordenados constituyen una permutación aleatoria. Cada número del proceso de permutación se iguala con el número de corte. Supongamos que la permutación ordenada es:

1 6 7 12 5 3 10 9 2 8 4 11

Paso 3. Asignar los primeros tres cortes en la lista (1, 6 y 7) al tratamiento A. Los siguientes tres cortes (12, 5 y 3) se asignan al tratamiento B y así sucesivamente. La asignación final de cortes a tratamientos es:

Corte:	1	6	7	12	5	3	10	9	2	8	4	11
Tratamiento:	A	A	A	B	B	B	C	C	C	D	D	D

La permutación aleatoria de los números 1 al 12 asegura que cada una de las asignaciones de tratamiento posible tiene la misma probabilidad de ocurrencia. Muchos programas comerciales de cómputo para estadística incluyen rutinas de permutación y aleatorización.

A falta de un programa de computadora o de una tabla de números aleatorios se puede usar un método equivalente de aleatorización. Después de asignar números a las unidades experimentales como en el paso 1, se elaboran tarjetas de papel con los mismos números y se sacan al azar de un recipiente. Los primeros r números son las unidades experimentales asignadas al primer tratamiento. Los segundos r números corresponden a las que se asignan al segundo tratamiento, y así sucesivamente. Hader (1973) presenta un análisis de los métodos adecuados e inadecuados de asignación aleatoria de tratamientos a las unidades.

Selección de las unidades experimentales para estudios comparativos por observación

El planteamiento del ingeniero de tránsito sobre el ancho de las calles es un estudio comparativo por observación. En éste el investigador no puede asignar al azar una unidad a un grupo de tratamiento. Según el tipo de programa de investigación, las unidades básicas se autoseleccionan, o bien, existen en sus grupos característicos.

Debe seleccionarse una muestra probabilística de unidades entre los miembros disponibles de cada población en tratamiento. Las unidades se eligen de cada población de manera que cada unidad tenga la misma oportunidad de entrar a la muestra. Observe que cada población representa una clasificación de tratamiento diferente y que el muestreo aleatorio se lleva a cabo sólo *dentro* de la población.

El primer paso requiere una identificación de las poblaciones que representan las condiciones o tratamientos de interés para el estudio por observación. Se construye una lista de todas las unidades disponibles en cada población. Por ejemplo, en un estudio ecológico quizás los dos tratamientos sean comunidades agrícolas de pastizales y de roble. Las poblaciones incluyen todos los sitios con pastizales y todos los sitios con robles dentro del área que se estudia. Se asigna un código de identificación único a cada sitio en las poblaciones. Por ejemplo, se asignan los números 0 al 99 a los 100 sitios disponibles para cada clasificación.

Supongamos que el investigador se propone establecer parcelas en diez pastizales y en diez sembradíos de roble. El método para muestrear diez sitios de cada área de estudio comienza con la selección de diez números con dos dígitos para cada sitio, de una tabla de números aleatorios (tabla XII del apéndice). Imaginemos que el primer conjunto de números con dos dígitos es 12, 63, 34, 05, 97, 72,

42, 44, 82, 51. Los sitios en el área de estudio que corresponden a esos números se usan para establecer las parcelas de medición. Se sigue el mismo procedimiento para cada grupo de tratamiento. Si de la tabla se obtiene el mismo número aleatorio más de una vez, se puede sacar un número adicional para completar la muestra. Muchos programas comerciales de cómputo para estadística incluyen rutinas para obtener un muestreo aleatorio a partir de una lista.

2.3 Preparación de los registros de datos para el análisis

En la mayor parte de los estudios se reúne una gran cantidad de datos que deben organizarse antes de que los programas de computadora los usen. La ejecución de los programas de estadística requiere un *archivo de datos*. El archivo puede capturarse en una terminal cuando el programa empieza o puede ser un archivo previamente almacenado en la computadora mediante un programa de captura de datos.

Se identifica con claridad cada observación en el archivo con una unidad experimental específica y un tratamiento del estudio. El archivo de datos tiene un formato conveniente para su revisión y escrutinio en busca de irregularidades ocurridas ya sea en las mediciones o en la captura de las observaciones. El archivo para el ejemplo 2.1 puede contener una secuencia de números que identifica cada corte de carne, el grupo de tratamiento al que se asignó y la cuenta de bacterias observada, como se puede ver en el cuadro 2.1.

Cuadro 2.1 Datos para el experimento de almacenamiento de carne

<i>Corte</i>	<i>Tratamiento</i>	<i>Log (N/cm²)</i>
1	Comercial	7.66
6	Comercial	6.98
7	Comercial	7.80
12	Al vacío	5.26
5	Al vacío	5.44
3	Al vacío	5.80
10	Mezcla de gases	7.41
9	Mezcla de gases	7.33
2	Mezcla de gases	7.04
8	CO ₂	3.51
4	CO ₂	2.91
11	CO ₂	3.66

El archivo de datos consiste de 12 líneas (no se incluyen los títulos), cada una con la información requerida por el programa para mantener los registros y hacer los cálculos.

Para los programas, cada línea del archivo es un *caso* u *observación*; los datos de almacenamiento de carne tienen 12 casos u observaciones. Cada columna es una *variable*, las variables en este archivo son *corte*, *tratamiento* y *log (N/cm²)*. Los valores reales de cada línea del archivo reciben el nombre de *valores de los datos*.

Las únicas variables que requiere el programa de estadística para realizar los cálculos relativos a este ejemplo son: *tratamiento* y *log (N/cm²)*.

En principio, la variable *corte* es una variable de referencia usada en el archivo de datos para identificar el caso con una unidad experimental en el estudio real. Las variables de referencia, que en los archivos de datos pueden incluir el código de la unidad experimental, el nombre del tratamiento, fechas y número de experimento, son muy útiles para manejar archivos grandes o cuando éstos se utilizan mucho para refrescar la memoria tiempo después de su creación.

2.4 Un modelo estadístico para la experimentación

El análisis estadístico se basa en un modelo estadístico formal subyacente. La interpretación adecuada del análisis requiere la comprensión del modelo. En los estudios comparativos, la característica de las unidades o sujetos medida en la observación es la *variable de respuesta*, identificada como la variable *y*. En el experimento de almacenamiento de carne, la cuenta de bacterias es la variable de respuesta.

El modelo estadístico para los estudios comparativos supone que existe una *población de referencia* de los sujetos o unidades experimentales. En la mayoría de los casos la población es conceptual, aunque es posible imaginar una población de motores de automóvil, tiendas, parcelas, corrales o carne empacada. Cada unidad individual en la población tiene un valor para la variable de respuesta *y*, y esta variable tiene una media μ y una varianza σ^2 .

Se supone una población de referencia para cada condición de tratamiento en el estudio, y también se supone que las unidades son sus representantes seleccionados al azar como resultado de una aleatorización. En los estudios por observación se infiere que las unidades se seleccionan al azar de las poblaciones de tratamiento.

El modelo estadístico se ilustra en la figura 2.1 con cuatro poblaciones de tratamiento hipotéticas. Cada población tiene una distribución normal en la variable de respuesta y cada una tiene una media diferente. Tal situación existe si los cuatro métodos de empaque tienen distinta capacidad para la inhibición del desarrollo de bacterias.

Se supone que la varianza σ^2 es la misma para cada población y no la afecta el tratamiento, como se muestra en la figura 2.1. Es decir, se supone que la varianza de las poblaciones de tratamiento es homogénea.

Uso del modelo de medias de celdas para describir las observaciones

Las observaciones se expresan como la suma de las medias poblacionales de tratamiento y de los errores experimentales con el modelo de *medias de celdas*

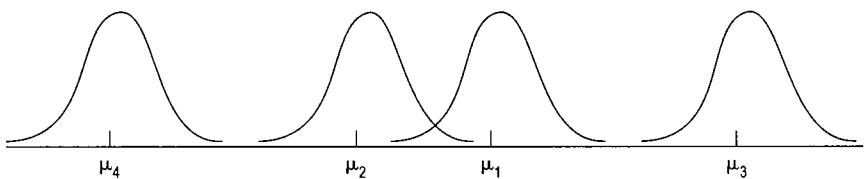


Figura 2.1 Ilustración de las poblaciones en tratamiento

$$y_{ij} = \mu_i + e_{ij} \quad (2.1)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde y_{ij} denota la j -ésima observación del i -ésimo grupo en tratamiento; μ_i es la media de la i -ésima población en tratamiento, y e_{ij} es el error experimental. Éste es un **modelo estadístico lineal** para la clasificación de tratamientos de un factor, en un diseño de experimentos totalmente aleatorizado.

El modelo tiene cierta tolerancia para la variación entre las observaciones de un grupo de tratamiento dado. Debido al error experimental, cada observación se desvía de la media de su población μ_i en una cantidad e_{ij} . La varianza del error experimental σ^2 es la varianza de e_{ij} , y se supone que es la misma para toda la población en tratamiento.

Las observaciones del experimento de almacenamiento de carne se muestran en la tabla 2.2, con $t = 4$ tratamientos y $r = 3$ réplicas con su identificación y_{ij} y su representación del modelo estadístico.

Uso de modelos alternos para describir otras hipótesis estadísticas

El modelo estadístico para el experimento refleja las creencias respecto a la relación entre los tratamientos y las observaciones. El modelo de medias de celdas es un *modelo completo*, $y_{ij} = \mu_i + e_{ij}$, que incluye una media distinta para cada población en tratamiento. Si no hay diferencias entre las medias de los tratamientos, se usa un modelo con un conjunto reducido de parámetros. El *modelo reducido*, $y_{ij} = \mu + e_{ij}$, establece que todas las observaciones pertenecen a la misma población con media μ .

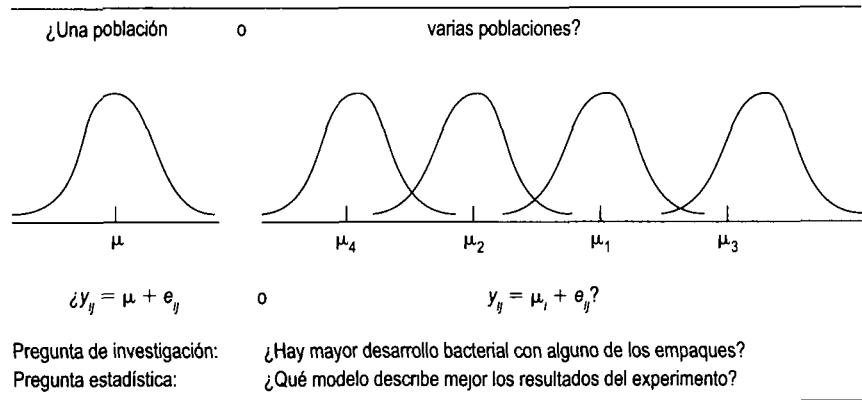
Los dos modelos representan la hipótesis estadística alternativa adecuada para el experimento. El modelo reducido representa la condición de la hipótesis nula sin diferencias entre las medias de los tratamientos, se establece como $H_0: \mu_1 = \mu_2 = \dots = \mu_t$. El modelo completo representa la condición de la hipótesis alterna cuando existen sólo algunas diferencias entre las medias de los tratamientos; se establece como $H_a: \mu_i \neq \mu_k$, donde $i \neq k$.

El investigador del experimento con carne debe determinar si el desarrollo de bacterias difiere con los distintos métodos de empaque o si ninguno es mejor en

Tabla 2.2 Identificación de los valores observados en el experimento de almacenamiento de carnes y su representación con el modelo estadístico lineal

Corte	Tratamiento	Observación	$\log (N/cm^2)$	y_{ij}	Modelo
1	1	1	7.66	y_{11}	$\mu_1 + e_{11}$
6	1	2	6.98	y_{12}	$\mu_1 + e_{12}$
7	1	3	7.80	y_{13}	$\mu_1 + e_{13}$
12	2	1	5.26	y_{21}	$\mu_2 + e_{21}$
5	2	2	5.44	y_{22}	$\mu_2 + e_{22}$
3	2	3	5.80	y_{23}	$\mu_2 + e_{23}$
10	3	1	7.41	y_{31}	$\mu_3 + e_{31}$
9	3	2	7.33	y_{32}	$\mu_3 + e_{32}$
2	3	3	7.04	y_{33}	$\mu_3 + e_{33}$
8	4	1	3.51	y_{41}	$\mu_4 + e_{41}$
4	4	2	2.91	y_{42}	$\mu_4 + e_{42}$
11	4	3	3.66	y_{43}	$\mu_4 + e_{43}$

cuanto suprimir el desarrollo de bacterias. Desde el punto de vista del modelo estadístico, el investigador debe determinar cuál de los dos modelos, el completo o el reducido, caracteriza mejor los datos del experimento. Las preguntas de investigación se traducen en preguntas sobre las poblaciones estadísticas trazadas en la figura 2.2.

**Figura 2.2** Preguntas de investigación, preguntas estadísticas y modelos alternos para el experimento de almacenamiento de carne

Para tomar una decisión con respecto a los tratamientos, el investigador necesita de un método estadístico para estimar los parámetros de los dos modelos y, con base en el criterio de observación, determinar qué hipótesis o modelos estadísticos se ajustan mejor a los datos del experimento.

Modelo lineal estadístico general

El modelo de medias de celdas es un caso especial del **modelo lineal general**. El modelo más general describe las relaciones entre dos tipos de variables como una función lineal en un conjunto de parámetros. Un tipo de variable es la de respuesta y , considerada dependiente del segundo tipo de variable: las variables de diseño x_1, x_2, \dots, x_k . Estas x_i pueden fijarse mediante el diseño del tratamiento, como la temperatura del tratamiento, o pueden ser covariadas medidas, como la edad de los sujetos. Las x_i también pueden representar categorías del tratamiento, como las del experimento de carnes.

El modelo estadístico relaciona y con las x_i a través de un conjunto de parámetros $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, con una relación lineal, es decir:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + e \quad (2.2)$$

Algunos ejemplos sencillos aclararán cómo puede el investigador desarrollar un modelo único para un estudio específico.

Un experimento mide la tasa de una reacción química, y , como respuesta a la temperatura, T , en la cámara de reacción. El investigador plantea la hipótesis de que el incremento en la tasa es lineal respecto a la temperatura. Si $x_1 = T$, la sencilla ecuación de una línea recta describe la relación lineal como:

$$y = \beta_0 + \beta_1 x_1 + e$$

que es el modelo de regresión lineal. Si el investigador piensa que la relación entre tasa y temperatura es cuadrática, entonces $x_1 = T$ y $x_2 = T^2$ y el modelo se convierte en

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$$

¿Cómo funciona el modelo si los tratamientos son categorías que no se pueden representar con valores métricos para las x_i ? En este caso, el investigador puede hacer de las x_i sus **variables indicadoras**. La variable indicadora hace justo lo que su nombre implica, indica el grupo de tratamiento al que pertenece la observación. En un esquema de variables indicadoras, $x = 1$ señala que la observación pertenece a cierto grupo y $x = 0$ que no pertenece a ese grupo.

El experimento de almacenamiento de carne tiene cuatro tratamientos y el modelo, a su vez, tiene cuatro variables indicadoras, x_1, x_2, x_3 y x_4 , que toman los valores 0 y 1 de la siguiente manera

$$x_1 = \begin{cases} 1 & \text{con empaque comercial} \\ 0 & \text{otro} \end{cases}$$

$$x_2 = \begin{cases} 1 & \text{con empaque al vacío} \\ 0 & \text{otro} \end{cases}$$

$$x_3 = \begin{cases} 1 & \text{con mezcla de gases} \\ 0 & \text{otro} \end{cases}$$

$$x_4 = \begin{cases} 1 & \text{con CO}_2 \text{ puro} \\ 0 & \text{otro} \end{cases}$$

El modelo puede escribirse como:

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + e$$

Si la observación se hizo en el tratamiento con empaque comercial, con $x_1 = 1$ y $x_2 = x_3 = x_4 = 0$:

$$y = \beta_1 + e$$

Si la observación se hizo en el tratamiento al vacío, con $x_2 = 1$ y $x_1 = x_3 = x_4 = 0$

$$y = \beta_2 + e$$

y así sucesivamente. Observe que $\beta_0 = 0$, ya que en este caso no es necesario para una descripción de las observaciones.

Entonces, el modelo $y = \beta_1 + e$ describe a la observación que tiene un valor de $\beta_1 + e$, si proviene del tratamiento con empaque comercial. Así $\beta_1 = \mu_1$, la media de la población con empaque comercial, que se ilustra en la figura 2.2.

El modelo de medias de la ecuación (2.1) se deriva directamente de esta representación si se hace $\beta_1 = \mu_1$, $\beta_2 = \mu_2$, $\beta_3 = \mu_3$ y $\beta_4 = \mu_4$. Dado que x_i solamente toma los valores 0 o 1; y si $\beta_1 = \mu_1$, entonces una observación del tratamiento de empaque comercial se modela como:

$$y = \mu_1 + e$$

Si se añaden subíndices para identificar las observaciones específicas, tendremos:

$$y_{1j} = \mu_1 + e_{1j} \quad j = 1, 2, \dots, r$$

que es el modelo de las observaciones del tratamiento con empaque comercial. Si se generaliza el modelo para incluir todos los grupos de tratamiento se obtiene el modelo de la ecuación (2.1), es decir:

$$y_{ij} = \mu_i + e_{ij} \\ i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

Esta expresión del modelo experimental con grupos de tratamiento es más tradicional y proporciona una descripción más específica del experimento.

La flexibilidad del modelo lineal general de la ecuación (2.2), permite al investigador incluir las covariadas medidas junto con las variables del grupo de tratamiento. Por ejemplo, en el experimento de almacenamiento de carnes, al investigador le preocupaba que las fluctuaciones en el contenido de humedad entre

las muestras de carne pudieran alterar el desarrollo de bacterias. Antes del empaque, se midió el contenido de humedad en la carne de cada unidad experimental.

Si x = contenido de humedad (%), entonces se puede agregar al modelo un término que representa una relación lineal entre la cuenta de bacterias y la humedad, como:

$$y_{ij} = \mu_i + \beta x_{ij} + e_{ij}$$

donde x_{ij} es el contenido de humedad en el j -ésimo corte del i -ésimo tratamiento y β es el coeficiente de regresión lineal. Éste es el modelo tradicional del **análisis de covarianza** que supone que el coeficiente de regresión es el mismo para todos los grupos en tratamiento. El análisis de covarianza se estudia en el capítulo 17.

2.5 Estimación de los parámetros del modelo con mínimos cuadrados

El **método de mínimos cuadrados** es quizás el que se usa con más frecuencia para estimar los parámetros del modelo lineal. Las estimaciones de los mínimos cuadrados son las estimaciones de μ_i , resultado de la suma más pequeña del cuadrado de los errores experimentales. Si éstos son independientes, tienen una media de cero y varianzas homogéneas, las estimaciones con mínimos cuadrados son sin sesgo con varianza mínima. Como se estudió en el capítulo 1, el muestreo aleatorio para los estudios por observación y la aleatorización en experimentos, garantizan la suposición de independencia. Los métodos para evaluar el supuesto de una varianza homogénea se presentan en el capítulo 4.

Estimadores para el modelo completo

Para el modelo de medias de celdas, los errores experimentales son la diferencia entre las medias de las observaciones de la población $e_{ij} = y_{ij} - \mu_i$, donde las observaciones y_{ij} son las únicas cantidades conocidas. Si se denotan los estimadores de mínimos cuadrados de μ_i como $\hat{\mu}_i$ para el modelo completo, entonces los estimadores de los errores experimentales son $\hat{e}_{ij} = y_{ij} - \hat{\mu}_i$. La suma mínima de los cuadrados es:

$$SCE = \sum_{i=1}^t \sum_{j=1}^r \hat{e}_{ij}^2 = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \hat{\mu}_i)^2 \quad (2.3)$$

La SCE es la suma de cuadrados estimada para el error experimental y es una medida de qué tan bien se ajusta el modelo a los datos.

Se usa un método de cálculo diferencial para determinar los estimadores $\hat{\mu}_i$, que minimizan la suma de cuadrados:

$$Q = \sum e_{ij}^2 = \sum (y_{ij} - \mu_i)^2$$

El método produce un conjunto de ecuaciones que deben resolverse para los estimadores. El nombre convencional de estas ecuaciones es *ecuaciones normales*; sin embargo, la designación no tiene relación con la distribución de probabilidad normal.

Para el tratamiento de t grupos con r réplicas por tratamiento, existen t ecuaciones normales, una por cada media de tratamiento. Las ecuaciones normales se encuentran obteniendo primero la diferencial de:

$$\sum_{i=1}^t \sum_{j=1}^r e_y^2 = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \mu_i)^2$$

con respecto a cada una de las μ_i , e igualando a cero el resultado. La derivada parcial para una ecuación típica es:

$$\frac{\partial}{\partial \mu_i} \sum_{j=1}^r (y_{ij} - \mu_i)^2 = -2 \sum_{j=1}^r (y_{ij} - \mu_i) = 0$$

Simplificando la ecuación y sustituyendo $\hat{\mu}_i$, en lugar de μ_i , se obtiene:

$$\sum_{j=1}^r \hat{\mu}_i = \sum_{j=1}^r y_{ij}$$

o

$$r\hat{\mu}_i = y_i$$

donde y_i es la suma total de las observaciones en el i -ésimo tratamiento.¹

La solución para el estimador de mínimos cuadrados de una media de tratamiento μ_i , es:

$$\hat{\mu}_i = \frac{y_i}{r} = \bar{y}_i \quad i = 1, 2, \dots, t \quad (2.4)$$

Esto causa que los estimadores de las medias poblacionales de los tratamientos basadas en el criterio de mínimos cuadrados sean las medias de los grupos de tratamientos observados.

Las ecuaciones normales para el estudio de almacenamiento de carne son:

$$\begin{aligned} 3\hat{\mu}_1 &= 22.44 \\ 3\hat{\mu}_2 &= 16.50 \\ 3\hat{\mu}_3 &= 21.78 \\ 3\hat{\mu}_4 &= 10.08 \end{aligned}$$

¹ Se usa la notación con puntos para simplificar la presentación de las sumas. El total de observaciones para el i -ésimo tratamiento se denota y_i donde el punto indica que se sumaron todas las observaciones del i -ésimo grupo de tratamiento para obtener este total, es decir:

$$y_i = \sum_{j=1}^r y_{ij} = y_{i1} + y_{i2} + \dots + y_{ir}$$

Además, el total de todas las observaciones se denota $y..$ con los dos puntos indicando que se ha concluido la suma sobre los dos subíndices, es decir:

$$y.. = \sum_{i=1}^t \sum_{j=1}^r y_{ij}$$

Las estimaciones con mínimos cuadrados para el experimento de carnes son:

$$\hat{\mu}_1 \frac{22.44}{3} = 7.48$$

$$\hat{\mu}_2 \frac{16.50}{3} = 5.50$$

$$\hat{\mu}_3 \frac{21.78}{3} = 7.26$$

$$\hat{\mu}_4 \frac{10.08}{3} = 3.36$$

La suma de cuadrados para el error experimental con el modelo $y_{ij} = \mu_i + e_{ij}$ es:

$$SCE = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \hat{\mu}_i)^2 = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 \quad (2.5)$$

Observe que la SCE es la suma completa de los cuadrados dentro de cada grupo de tratamiento. La varianza de la muestra para el i -ésimo grupo es:

$$s_i^2 = \sum_{j=1}^r \frac{(y_{ij} - \bar{y}_i)^2}{(r-1)}$$

y es una estimación de σ^2 , a partir de los datos del i -ésimo grupo de tratamiento.

Si es posible suponer que σ^2 es homogénea, es decir, que es la misma para todos los grupos en tratamiento, entonces:

$$s^2 = \frac{\sum_{i=1}^t \left[\sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 \right]}{t(r-1)} = \frac{SCE}{t(r-1)} \quad (2.6)$$

es una suma completa de σ^2 para todos los datos del experimento. Las sumas de cuadrados y la varianza calculadas para el experimento de carnes se muestran en la tabla 2.3.

Estimadores para el modelo reducido

Cuando no existen diferencias entre las medias de las poblaciones en tratamiento, se usa un modelo $y_{ij} = \mu + e_{ij}$ más sencillo o reducido para describir los datos. El estimador de mínimos cuadrados de μ es la gran media de todas las observaciones en el experimento:

$$\hat{\mu} = \bar{y} = \frac{y}{N} \quad (2.7)$$

donde $N = rt$.

En el modelo reducido bajo la hipótesis nula, la suma de cuadrados mínima para el error experimental es:

Tabla 2.3 Observaciones, medias y sumas de cuadrados dentro de los grupos para el experimento de almacenamiento de carne

	<i>Comercial</i>	<i>Al vacío</i>	<i>CO, O₂, N</i>	<i>CO₂</i>
	7.66	5.26	7.41	3.51
	6.98	5.44	7.33	2.91
	7.80	5.80	7.04	3.66
$\hat{\mu}_i = \bar{y}_i$	7.48	5.50	7.26	3.36
$\sum_{j=1}^r (y_{ij} - \bar{y}_i)^2$	0.3848	0.1512	0.0758	0.3150

$$SCE = 0.3848 + 0.1512 + 0.0758 + 0.3150 = 0.9268$$

$$s^2 = \frac{SCE}{t(r-1)} = \frac{0.9268}{4(2)} = 0.11585$$

$$SCE_r = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \hat{\mu}_i)^2 = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 \quad (2.8)$$

La SCE_r es la suma total del cuadrado de todas las observaciones, expresadas como desviaciones estándar de la gran media.

La estimación de la gran media para el experimento de carnes es:

$$\hat{\mu} = \bar{y} = \frac{70.80}{12} = 5.90$$

y

$$SCE_r = (7.66 - 5.90)^2 + \dots + (3.66 - 5.90)^2 = 33.7996$$

2.6 Suma de cuadrados para identificar fuentes de variación importantes

Se pueden usar las diferencias de las sumas de los cuadrados del error experimental para que los dos modelos sean una partición de la variación total en el experimento. Estas particiones aclararán y explicarán los resultados del experimento. Las sumas de los cuadrados del error experimental calculadas para los dos modelos de empaque de carnes eran muy diferentes: $SCE_f = 0.9268$ para el modelo completo con cuatro medias poblacionales de tratamiento ($y_{ij} = \mu_i + e_{ij}$), y $SCE_r = 33.7996$ para el modelo reducido con una sola media poblacional ($y_{ij} = \mu + e_{ij}$).

La menor suma de cuadrados para el modelo completo indica que los errores experimentales estimados en el modelo completo ($\hat{e}_{ij} = y_{ij} - \hat{\mu}_i$), serán generalmente valores más pequeños que sus contrapartes en el modelo reducido. La diferencia entre las observaciones y sus medias de grupo separadas $\hat{\mu}_i$, mostradas en la tabla 2.4, son:

Tabla 2.4 Valores observados, estimaciones y sus desviaciones, con los modelos reducido y completo

Tratamiento	Observado y	Modelo reducido $y_y = \mu + e_y$		Modelo completo $y_y = \mu_i + e_y$	
		Estimado $\hat{\mu}$	Diferencia $(y_y - \hat{\mu})$	Estimado $\hat{\mu}_i$	Diferencia $(y_y - \hat{\mu}_i)$
Comercial	7.66	5.90	1.76	7.48	0.18
	6.98	5.90	1.08	7.48	-0.50
	7.80	5.90	1.90	7.48	0.32
Al vacío	5.26	5.90	-0.64	5.50	-0.24
	5.44	5.90	-0.46	5.50	-0.06
	5.80	5.90	-0.10	5.50	0.30
CO, O ₂ , N	7.41	5.90	1.51	7.26	0.15
	7.33	5.90	1.43	7.26	0.07
	7.04	5.90	1.14	7.26	-0.22
100% CO ₂	3.51	5.90	-2.39	3.36	0.15
	2.91	5.90	-2.99	3.36	-0.45
	3.66	5.90	-2.24	3.36	0.30
			$SCE_r = 33.7996$		
				$SCE_c = 0.9268$	

menores que las diferencias entre las observaciones y la gran media, $\hat{\mu} = 5.90$, estimadas con el modelo reducido, con excepción del tratamiento al vacío. A continuación se calculan las dos sumas de cuadrados de los errores experimentales y su diferencia.

$$\text{Modelo reducido: } SCE_r = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_.)^2 = 33.7996$$

$$\text{Modelo completo: } SCE_c = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 = 0.9268$$

$$\begin{aligned} \text{Diferencia: } SCE_r - SCE_c &= \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_.)^2 - \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 \\ &= \sum_{i=1}^t \sum_{j=1}^r (\bar{y}_i - \bar{y}_.)^2 = r \sum_{i=1}^t (\bar{y}_i - \bar{y}_.)^2 \\ &= 32.8728 \end{aligned}$$

La suma de cuadrados de la **diferencia** es la suma de los cuadrados de las diferencias entre las medias de los grupos en tratamiento \bar{y}_i y la gran media \bar{y} . La suma de cuadrados de la diferencia, conocida como **suma de los cuadrados de tratamientos**,

representa una reducción en la *SCE* una vez incluidos los tratamientos en el modelo; así, con frecuencia se le llama *reducción de la suma de cuadrados debida a los tratamientos*. La **suma de cuadrados total** del experimento es SCE_t , ya que se trata de la suma de los cuadrados de la diferencia entre todas las observaciones y la gran media \bar{y} .

Partición fundamental: con el modelo reducido, la suma total de cuadrados SCE_t , es la sumatoria de la suma del cuadrado de los tratamientos; con el modelo completo, es la suma de cuadrados del error experimental SCE_e . Por lo tanto, se tiene la relación:

$$\sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y})^2 = \sum_{i=1}^t \sum_{j=1}^r (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 \quad (2.9)$$

o

$$SC \text{ total} = SC \text{ tratamiento} + SC \text{ error}$$

Se ha hecho una partición de la suma de cuadrados total en dos partes:

- *SC tratamiento* es la suma del cuadrado de las diferencias entre las medias del grupo en tratamiento y la gran media.
- *SC error* es la suma del cuadrado de las diferencias entre las observaciones en el grupo y la media del grupo.

Las fórmulas para la suma de cuadrados se derivan de una identidad para la desviación de la gran media que tenga cualquier observación. La ecuación:

$$(y_{ij} - \bar{y}) = (\bar{y}_i - \bar{y}) + (y_{ij} - \bar{y}_i) \quad (2.10)$$

hace una partición en dos de la desviación que tenga cualquier observación de la gran media. Es una suma de: 1) la desviación de la media del grupo de la gran media ($\bar{y}_i - \bar{y}$), y 2) la desviación de la observación de la media del grupo ($y_{ij} - \bar{y}_i$), donde esta última es la medida del error experimental asociado con la observación. Al elevar al cuadrado y sumar ambos lados de la ecuación (2.10) se obtiene:

$$\begin{aligned} \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y})^2 &= \sum_{i=1}^t \sum_{j=1}^r (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 \\ &\quad + 2 \sum_{i=1}^t \sum_{j=1}^r (\bar{y}_i - \bar{y})(y_{ij} - \bar{y}_i) \end{aligned}$$

Pero el término del producto cruzado suma cero, por lo que la expresión que resulta:

$$\sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y})^2 = \sum_{i=1}^t \sum_{j=1}^r (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2$$

es idéntica a la partición de la suma de cuadrados que muestra la ecuación (2.9).

Un resumen de las fórmulas para las sumas de cuadrados equivalentes a la fórmula de definición de la ecuación (2.9) es:

$$SC \text{ total} = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_.)^2$$

$$SC \text{ tratamiento} = r \sum_{i=1}^t (\bar{y}_{i\cdot} - \bar{y}_.)^2$$

$$SC \text{ error} = SC \text{ total} - SC \text{ tratamiento}$$

2.7 Modelo de efectos del tratamiento

El **efecto de un tratamiento** indica cuánto cambia una medición en una unidad experimental al someterse a tratamiento. Es posible expresar el modelo de medias de celda de manera que refleje los efectos del tratamiento sobre las unidades experimentales. Si se reordena la ecuación (2.10) de manera que exprese la observación en términos de una función de la gran media y las dos desviaciones, se obtiene:

$$y_{ij} = \bar{y}_. + (\bar{y}_{i\cdot} - \bar{y}_.) + (y_{ij} - \bar{y}_{i\cdot}) \quad (2.11)$$

Un modelo de población equivalente sería:

$$y_{ij} = \bar{\mu} + (\mu_{i\cdot} - \bar{\mu}) + (y_{ij} - \mu_{i\cdot}) \quad (2.12)$$

donde $\bar{\mu} = \sum_i \mu_{i\cdot} / t$ es el promedio de las medias poblacionales para el modelo de medias de celdas $y_{ij} = \mu_{i\cdot} + e_{ij}$.

La desviación que muestran las medias del grupo con respecto a la gran media ($\mu_{i\cdot} - \bar{\mu}$) se conoce como *efecto del tratamiento*, y el modelo en la ecuación (2.12) se escribe con frecuencia como:

$$y_{ij} = \mu + \tau_i + e_{ij} \quad (2.13)$$

donde $\mu = \bar{\mu}$, $\tau_i = (\mu_{i\cdot} - \bar{\mu})$, y $e_{ij} = (y_{ij} - \mu_{i\cdot})$. Las diversas expresiones para los efectos del tratamiento se muestran en el cuadro 2.2. La figura 2.3 contiene una representación gráfica de los efectos del tratamiento para el experimento de carnes.

Cuadro 2.2 Efectos del tratamiento

	<i>Tratamiento 1</i>	<i>Tratamiento 2</i>	...	<i>Tratamiento t</i>
Muestra	$(\bar{y}_1 - \bar{y}_.)$	$(\bar{y}_2 - \bar{y}_.)$...	$(\bar{y}_t - \bar{y}_.)$
Población	$(\mu_1 - \bar{\mu})$	$(\mu_2 - \bar{\mu})$...	$(\mu_t - \bar{\mu})$
Efecto	τ_1	τ_2	...	τ_t

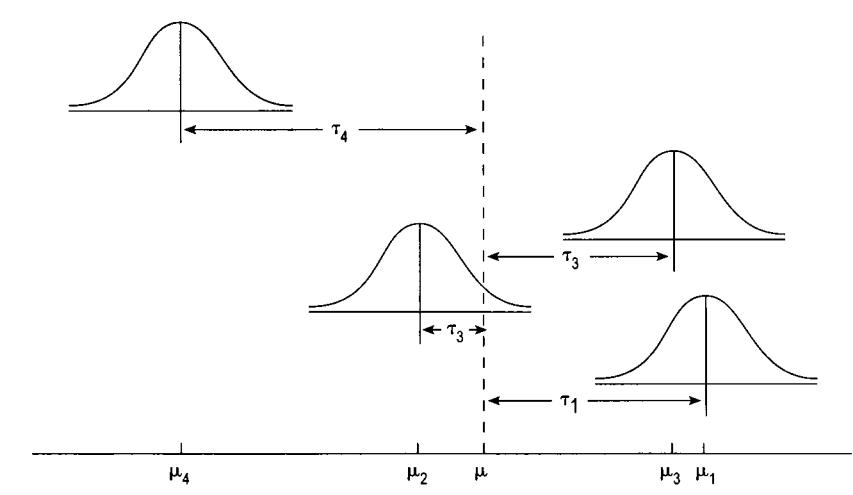


Figura 2.3 Representación gráfica de los efectos del tratamiento

El modelo en la ecuación (2.13) tiene $(t + 1)$ parámetros de población, que son $\mu, \tau_1, \tau_2, \dots, \tau_t$. Su suma es igual a cero como consecuencia de las definiciones de los efectos del tratamiento.

$$\sum_{i=1}^t \tau_i = \sum_{i=1}^t (\mu_i - \bar{\mu}) = 0 \quad (2.14)$$

2.8 Grados de libertad

Se puede pensar en los **grados de libertad** como el número de elementos estadísticamente independientes en las sumas de cuadrados. El valor de los grados de libertad representa el número de piezas de información independientes en las sumas de cuadrados. La suma de cuadrados de todas las observaciones $\sum y_{ij}^2$ tiene N elementos estadísticamente independientes y, por lo tanto, N grados de libertad.

Una vez estimado el parámetro μ mediante $\bar{y}_{..}$ a partir de los datos, la suma de cuadrados del error para el modelo reducido es $SCE_r = SC \text{ total} = \sum (y_{ij} - \bar{y}_{..})^2$. Las $(y_{ij} - \bar{y}_{..})$ en SCE_r no son N elementos estadísticamente independientes porque suman cero y alguna de ellas es el negativo de la suma de los otros $(N - 1)$ valores. Esta restricción lineal sobre las observaciones es consecuencia de la estimación de un parámetro, μ , en el modelo reducido. En general, los grados de libertad para la SCE después de ajustar el modelo es el número de observaciones menos el número de parámetros estimados a partir de los datos.

Existen t parámetros $(\mu_1, \mu_2, \dots, \mu_t)$ estimados para el modelo completo. En consecuencia, el número de elementos estadísticamente independientes en la suma de cuadrados del error para el modelo completo es $(N - t)$, de manera que SCE_c tiene $(N - 1)$ grados de libertad.

La suma de cuadrados del tratamiento se determina a partir de la diferencia entre las SCE para dos modelos:

$$SC \text{ tratamiento} = SCE_r - SCE_c$$

Los grados de libertad para estas diferencias se pueden determinar como la diferencia entre los grados de libertad de SCE_r y SCE_c

$$(N - 1) - (N - t) = (t - 1)$$

Así, se asocian $(t - 1)$ grados de libertad con la reducción de la suma de cuadrados debida a los tratamientos.

2.9 Resumen en la tabla de análisis de varianza

La **tabla de análisis de varianza** resume el conocimiento acerca de la variabilidad en las observaciones del experimento. Se ha hecho una partición en dos de la suma de cuadrados total, una representa la variación entre las medias de tratamientos, la otra al error experimental.

La varianza del error experimental σ^2 se estima mediante $s^2 = SCE/(N - t)$, donde s^2 se conoce como el cuadrado medio del error (CME). La otra media al cuadrado de importancia es el cuadrado medio de tratamientos (CMT), que se calcula con:

$$CMT = \frac{SC \text{ tratamiento}}{(t - 1)}$$

Las particiones de la suma de cuadrados, los grados de libertad y los cuadrados de las medias se resumen en la tabla de análisis de varianza como la que se muestra en la tabla 2.5.

Tabla 2.5 Tabla de análisis de varianza para tratamientos con un solo factor, en un diseño totalmente aleatorizado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Total	$N - 1$	$SC \text{ total}$	
Tratamientos	$t - 1$	$SC \text{ tratamiento}$	$CMT = \frac{SCT}{t - 1}$
Error	$N - t$	$SC \text{ error}$	$CME = \frac{SCE}{N - t}$

El *CME* es un estimador sin sesgo de la varianza del error experimental σ^2 ; es decir, el valor esperado del *CME* es igual a σ^2 , o bien:

$$E(CME) = \sigma^2 \quad (2.15)$$

El valor esperado del *CMT* es:

$$E(CMT) = \sigma^2 + r\theta_t^2 \quad (2.16)$$

donde:

$$\theta_t^2 = \frac{\sum_{i=1}^t (\mu_i - \bar{\mu})^2}{(t - 1)}$$

es la varianza de las medias de los tratamientos. En consecuencia, *CMT* estima una combinación de la varianza del error experimental y la varianza entre las medias de los tratamientos, en el modelo lineal de la hipótesis $y_{ij} = \mu_i + e_{ij}$. Los desarrollos algebraicos de $E(CME)$ y $E(CMT)$ se presentan en el apéndice 2A.2 de este capítulo.

2.10 Pruebas de hipótesis sobre modelos lineales

En la tabla 2.6 se muestra toda la tabla de análisis de varianza correspondiente al experimento de carnes. El análisis de varianza resume la magnitud de las fuentes de variación en el experimento. Si una variación debida a los tratamientos es significativamente mayor que el error experimental aleatorio, entonces se requiere una prueba de hipótesis.

Tabla 2.6 Análisis de varianza para log (Nº/cm²) de la bacteria siccotrópica en el experimento de almacenamiento de carne

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	11	33.800			
Empaque	3	32.873	10.958	94.58	.000
Error	8	0.927	0.116		

Se puede usar una prueba de aleatorización basada en el análisis estadístico de la información de la varianza. Sin embargo, como se explica en la sección 1.8, las pruebas estadísticas basadas en la teoría de la distribución normal tienen la misma validez siempre que se cumplan las suposiciones de distribución normal. En consecuencia, se supone que las observaciones y_{ij} son independientes y tienen una distribución normal con media μ_i y varianza σ^2 como se muestra en la figura 2.1. Los métodos para evaluar la suposición de observaciones con distribución normal se presentan en el capítulo 4.

Diferencia de suma de cuadrados para comparar dos modelos

La SCE_c es una medida de cuánto se ajusta el modelo completo a los datos observados, y la SCE_r es la medida equivalente para el modelo reducido. Entonces, la diferencia ($SCE_r - SCE_c$) es una medida de la superioridad del modelo completo sobre el modelo reducido. En consecuencia, la razón proporciona un medio para evaluar la ventaja relativa y forma parte esencial del conocido criterio de la prueba F :

$$\frac{(SCE_r - SCE_c)}{SCE_c}$$

Estadístico F para probar una hipótesis del modelo

De la teoría estadística, se sabe que las sumas de los cuadrados de variables aleatorias con distribución normal se asocian con la distribución ji-cuadrada. Se puede demostrar que SCE_c/σ^2 se distribuye como la variable ji-cuadrada con $(N - t)$ grados de libertad. La diferencia $SCT = SCE_r - SCE_c$, con $(t - 1)$ grados de libertad, representa la reducción en la suma de cuadrados total debida a las diferencias entre las medias de los tratamientos. Cuando las medias de los tratamientos son iguales, $\mu_1 = \mu_2 = \dots = \mu_t$, se puede demostrar que este criterio, $(SCE_r - SCE_c)/\sigma^2 = SCT/\sigma^2$, también tiene una distribución ji-cuadrada con $(t - 1)$ grados de libertad y que es independiente de la distribución de SCE_c/σ^2 .

El cociente:

$$F = \frac{(SCE_r - SCE_c)(t - 1)}{SCE_c/(N - t)} \quad (2.17)$$

es el cociente de dos distribuciones ji-cuadrado, cada una dividida entre sus respectivos grados de libertad. Bajo la hipótesis nula de que no hay diferencia entre los tratamientos, el cociente tiene una distribución F con $(t - 1)$ y $(N - t)$ grados de libertad, respectivamente, en el numerador y el denominador.

El estadístico de prueba calculado a partir de la tabla de análisis de varianza para probar la hipótesis nula $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ es:

$$F_0 = \frac{CMT}{CME} \quad (2.18)$$

que tiene distribución F cuando la hipótesis nula es cierta.

La suma de cuadrados esperada en las ecuaciones (2.15) y (2.16) muestra que CME es un estimador sin sesgo de σ^2 , bajo cualquier hipótesis. Pero CMT es un estimador sin sesgo de σ^2 , sólo bajo el modelo de la hipótesis nula o del modelo reducido; es decir, si $\mu_1 = \mu_2 = \dots = \mu_t$, entonces $\theta_t^2 = 0$ y $E(CMT) = \sigma^2$. Con la hipótesis alterna, el valor esperado de CMT es mayor que σ^2 , como se ve en la ecuación (2.16), en consecuencia, el valor esperado del numerador del estadístico F_0 será mayor que el del denominador. Los valores grandes de F_0 sugieren el rechazo de la hipótesis nula.

Para la prueba de hipótesis se utiliza una región crítica de un lado en la cola de la derecha. La hipótesis nula H_0 se rechaza para una probabilidad de un error tipo I de α si:

$$F_0 > F_{\alpha,(t-1),(N-t)}$$

donde $F_0 = CMT/CME$ y $F_{\alpha,(t-1),(N-t)}$ es el valor de la distribución F que es excedido con probabilidad α . Los valores críticos de la distribución F se encuentran en la tabla IV del apéndice.

Una prueba de hipótesis de que no hay diferencia entre los cuatro tratamientos de empaque de carne en cuanto al crecimiento de bacterias se ilustra en el cuadro 2.3. Del análisis de varianza (tabla 2.6) los cuadrados medios requeridos son $CMT = 10.958$ con 3 grados de libertad y $CME = 0.116$ con 8 grados de libertad.

Cuadro 2.3 Prueba de hipótesis para el experimento de carnes

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_\alpha : \mu_i \neq \mu_k \text{ para al menos una } i \neq k$$

$$\alpha = .05 \quad \text{Región crítica: } F_0 > F_{05,3,8} = 4.07$$

$$F_0 = \frac{CMT}{CME} = \frac{10.958}{0.116} = 94.58$$

Como $F_0 = 94.58$ cae en la región crítica $F_0 > 4.07$, se rechaza la hipótesis nula y se concluye que los tratamientos difieren respecto al número de bacterias sicotrópicas observadas en la carne almacenada bajo condiciones diferentes.

2.11 Pruebas de significancia y pruebas de hipótesis

Una práctica común en las pruebas de hipótesis es determinar qué tan significativa es una prueba estadística. Este nivel de significancia es la probabilidad de exceder el valor del estadístico de prueba bajo las condiciones de la hipótesis nula. Observe la columna con título $Pr > F$ de la tabla 2.6. El valor .000 es la probabilidad de que el estadístico F con 3 y 8 grados de libertad sea mayor que $F_0 = 94.58$ y suele denominarse “valor P ”. Como el valor reportado es .000, se sabe que la probabilidad de exceder $F_0 = 94.58$ es menor que .0001, o $P < .0001$. Muchos investigadores usan la magnitud del valor P para decidir la significancia estadística de la prueba F en el análisis de varianza. Es común que se proporcione este valor en el análisis de los resultados. Por ejemplo, la prueba F para el caso actual puede darse como “significativa en el nivel de significancia $P < .0001$ ”. Si el valor de la probabilidad es menor que los niveles de significancia tradicionales de .01 y .05, la hipótesis nula será rechazada porque el estadístico observado está en la región crítica.

La mayoría de los programas de cómputo para análisis de varianza incluyen este valor de probabilidad para el estadístico F en los resultados impresos. En

Hung *et al.* (1977) y las referencias ahí citadas, se encuentra una discusión técnica del valor P y sus relaciones con el número de réplicas y con la hipótesis alterna.

2.12 Errores estándar e intervalos de confianza para medias de tratamientos

En la sección 2.5 está determinado que los estimadores de mínimos cuadrados de las medias poblacionales eran las medias observadas de los grupos en tratamiento \bar{y}_i . Las medias observadas son los promedios de r observaciones independientes, de manera que la varianza de la media de un grupo de tratamiento es $\sigma_{\bar{y}_i}^2 = \sigma^2/r$. El estimador de la varianza es:

$$s_{\bar{y}_i}^2 = \frac{s^2}{r} \quad (2.19)$$

donde, a partir del análisis de varianza, $s^2 = ECM$. El estimador del error estándar de la media es:

$$s_{\bar{y}_i} = \sqrt{\frac{s^2}{r}} \quad (2.20)$$

Se construye un estimado de $100(1 - \alpha)\%$ para el intervalo de confianza (IC), con límites superior e inferior respectivamente, para cada media de grupo en tratamiento; donde $t_{\alpha/2,(N-t)}$ es un estadístico t de Student excedido con una probabilidad $\alpha/2$ y $(N - t)$ son los grados de libertad para los CME:

$$\bar{y}_i + t_{\alpha/2,(N-t)}(s_{\bar{y}_i}) \quad y \quad \bar{y}_i - t_{\alpha/2,(N-t)}(s_{\bar{y}_i}) \quad (2.21)$$

El error estándar de la media para cualquier grupo en tratamiento del estudio de almacenamiento de carne es:

$$s_{\bar{y}_i} = \sqrt{\frac{0.116}{3}} = 0.197$$

y $t_{0.025,8} = 2.306$. Las medias de los grupos en tratamiento se muestran en el cuadro 2.4 con una estimación del IC de 95%.

Cuadro 2.4 Medias, errores estándar e intervalos de confianza del 95% para el experimento de empaques de carne

Tratamiento	Media	Error estándar	IC del 95%	(superior, inferior)
Comercial	7.48	0.197	7.48 ± 0.454	(7.03, 7.93)
Al vacío	5.50	0.197	5.50 ± 0.454	(5.05, 5.95)
CO, O ₂ , N	7.26	0.197	7.26 ± 0.454	(6.81, 3.81)
100% CO ₂	3.36	0.197	3.36 ± 0.454	(2.91, 3.81)

Para cualquier media de grupo, el IC del 95% es

$$\bar{y}_i \pm (2.306)(0.197) \quad \text{o} \quad \bar{y}_i \pm 0.454$$

2.13 Diferente número de réplicas de los tratamientos

Pueden ocurrir diferente número de réplicas debido a la pérdida de algunas unidades experimentales durante el estudio, a que había un número insuficiente de sujetos disponibles para todos los grupos de estudio o por extravío, destrucción o nulidad de los datos recolectados. Los efectos más notables visibles de las réplicas desiguales aparecen en los cálculos necesarios para el análisis. Una consideración de la mayor importancia es asegurarse de que se da seguimiento a la información desigual en todos los grupos de tratamiento. La pérdida de observaciones de un grupo de tratamiento tiene como consecuencia una pérdida proporcional en la precisión de las estimaciones de la media de ese grupo con respecto a las de los grupos con todos los datos.

Ejemplo 2.2 Detección de flebitis durante la terapia de amiodarone

La flebitis es una inflamación de las venas que puede presentarse al administrar medicamentos por vía intravenosa. Se supuso que la droga activa es el factor principal que contribuye a la inflamación, aunque es posible que la solución excipiente usada para la administración intravenosa también contribuya.

Hipótesis de investigación: El problema de detectar la aparición de flebitis es de particular importancia para los investigadores. Este estudio se diseñó con la finalidad primordial de buscar mecanismos para la detección temprana de flebitis durante una terapia de amiodarone. Se planteó la hipótesis de que los cambios en la temperatura de los tejidos cercanos al lugar de la administración intravenosa sería señal de una inflamación inminente.

Diseño del tratamiento: Se administraron tres tratamientos intravenosos en animales de laboratorio. Éstos fueron:

- amiodarone en una solución excipiente
- sólo una solución excipiente
- una solución salina

La solución salina sirvió como tratamiento placebo de control para determinar si la sola administración afectaba la inflamación. La solución excipiente sirvió como control para separar los efectos del vehículo de los del medicamento.

Diseño del experimento: Los conejos utilizados como animales de prueba se asignaron al azar a los tres grupos de tratamiento, en un diseño totalmente aleatorizado y se les insertó una aguja en una vena de la oreja.

Un incremento en la temperatura de la oreja tratada se consideró como posible indicador temprano de flebitis. La diferencia en las temperaturas de las orejas (tratada menos no tratada) se usó como variable de respuesta. Las complicaciones con el protocolo experimental dieron como resultado un número diferente de conejos en cada grupo de tratamiento. En la tabla 2.7 se muestran las diferencias de temperatura observadas luego de 4.5 horas, para cada conejo que siguió en estudio.

Tabla 2.7 Diferencias en la temperatura de las orejas ($^{\circ}\text{C}$), tratada menos no tratada, de los conejos, 4.5 horas después del tratamiento

	<i>Amiodarone</i>	<i>Excipiente</i>	<i>Salina</i>	
	2.2	0.3	0.1	
	1.6	0.0	0.1	
	0.8	0.6	0.2	
	1.8	0.0	-0.4	
	1.4	-0.3	0.3	
	0.4	0.2	0.1	
	0.6		0.1	
	1.5		-0.5	
	0.5			
r_i	9	6	8	$N = \sum_{i=1}^t r_i = 23$
Total (y_i)	10.80	0.80	0.00	$y = 11.60$
Media (\bar{y}_i)	1.20	0.13	0.00	$\bar{y} = 0.50$

Fuente. G. Ward, Departamento de Ciencias Farmacéuticas, Universidad de Arizona

Modelo lineal para diferente número de réplicas

El modelo de medias de celda para tratamientos con un solo factor con un diseño totalmente aleatorizado con replicación desigual es:

$$y_{ij} = \mu_i + e_{ij} \quad (2.22)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r_i$$

donde r_i es el número de réplicas para el i -ésimo grupo en tratamiento. El número total de observaciones es $N = \sum_i r_i$. La interpretación y las suposiciones son las mismas que para el modelo de réplicas iguales.

El estimador de mínimos cuadrados para las medias de tratamiento determinado por los métodos descritos en la sección 2.5 es:

$$\hat{\mu}_i = \frac{y_i}{r_i} \quad i = 1, 2, \dots, t \quad (2.23)$$

o la media del grupo en tratamiento observada, como en el caso de réplicas iguales.

Análisis de varianza para réplicas desiguales

En la tabla 2.8 se muestran las particiones de las sumas de cuadrados para el análisis de varianza. Cada elemento de la suma de cuadrados para los tratamientos es un cuadrado ponderado de la desviación de la media de un tratamiento de la gran media, $r_i(\bar{y}_i - \bar{y}_.)^2$. El valor relativo r_i , es el número de réplicas para el grupo en tratamiento. Los valores relativos reflejan la cantidad de información disponible para estimar las medias de los tratamientos. La suma de cuadrados para el error experimental es la suma agrupada de los cuadrados dentro de los grupos. El cuadrado de la media esperada para los tratamientos incluye una suma ponderada del cuadrado de los efectos de cada tratamiento $\tau_i = (\mu_i - \bar{\mu})$, donde $\bar{\mu} = \sum_i r_i \mu_i / N$. Con réplicas desiguales, nuestra definición de los efectos de los tratamientos significa que su suma ponderada es igual a cero, $\sum_i r_i (\mu_i - \bar{\mu}) = 0$.

Tabla 2.8 Análisis de varianza para un diseño totalmente aleatorizado con réplicas desiguales de los tratamientos

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Valor esperado de los cuadrados medios
Total	$N - 1$	$\sum_{i=1}^t \sum_{j=1}^{r_i} (y_{ij} - \bar{y}_{..})^2$		
Tratamientos	$t - 1$	$\sum_{i=1}^t r_i (\bar{y}_i - \bar{y}_{..})^2$	CMT	$\sigma^2 + \theta_t^2$
Error	$N - t$	$\sum_{i=1}^t \sum_{j=1}^{r_i} (y_{ij} - \bar{y}_i)^2$ $\theta_t^2 = \frac{1}{(t-1)} \sum_{i=1}^t r_i (\mu_i - \bar{\mu})^2$	CME	σ^2

El análisis de varianza para el estudio de amiodarone se muestra en la tabla 2.9. Se examina la hipótesis nula de que no hay diferencia entre las medias de los tratamientos con el estadístico $F_0 = CMT/CME = 3.6081/0.2177 = 16.58$. El estadístico F_0 se encuentra en la columna de “F” en la tabla 2.9. La hipótesis nula se rechaza a un nivel de significancia de .05 porque $F_0 > F_{0.05, 2, 20} = 3.49$, de otra manera, observe que $Pr > F = .000$.

Tabla 2.9 Análisis de varianza para las diferencias de temperatura en las orejas de conejos en el estudio de amiodarone

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	22	11.5696			
Tratamiento	2	7.2162	3.6081	16.58	.000
Error	20	4.3533	0.2177		

Errores estándar para medias de tratamiento y precisión desigual

El error estándar de la media de un grupo en tratamiento se estima con:

$$s_{\bar{y}_i} = \sqrt{\frac{CME}{r_i}} \quad (2.24)$$

Los errores estándar estimados de las tres medias de tratamiento intravenoso son:

$$s_{\bar{y}_1} = \sqrt{\frac{0.2177}{9}} = 0.156$$

$$s_{\bar{y}_2} = \sqrt{\frac{0.2177}{6}} = 0.190$$

y

$$s_{\bar{y}_3} = \sqrt{\frac{0.2177}{8}} = 0.165$$

Las diferencias entre los errores estándar de las medias de los grupos tratados ilustran la precisión inferior que se obtiene cuando se pierden datos de un experimento. Por ejemplo, con sólo seis observaciones, el grupo de tratamiento con excipiente tiene un error estándar 22% mayor que el tratado con amiodarone en nueve observaciones.

2.14 ¿Cuántas réplicas para la prueba F ?

En el capítulo 1, el número de réplicas necesario para detectar alguna diferencia predeterminada entre dos medias de tratamiento se estableció con el estadístico z normal, junto con el conocimiento de la varianza, el nivel de significancia y la potencia de la prueba. En esta sección se determina el número requerido de réplicas utilizando un método basado en el estadístico F .

La potencia de la prueba de hipótesis es la probabilidad de rechazar una falsa hipótesis nula. El estadístico $F_0 = CMT/CME$ se usa para probar la hipótesis nula $H_0: \tau_i = 0$. La potencia de la prueba es $1 - \beta = P(F > F_{\alpha, \nu_1, \nu_2} | H_0 \text{ falsa})$, donde ν_1 y ν_2 son los grados de libertad respectivos del numerador y el denominador. Cuando H_0 es falsa, F_0 tiene una distribución F desplazada con ν_1 y ν_2 grados de libertad y parámetro de desplazamiento $\lambda = r \sum_i \tau_i^2 / \sigma^2$. Si la hipótesis nula es cierta, entonces el parámetro de desplazamiento tiene un valor $\lambda = 0$, puesto que todo $\tau_i = 0$ y F_0 tiene una distribución F centrada.

Se han calculado valores tabulados de la potencia de la prueba F para valores dados del nivel de significancia α ; la potencia $1 - \beta$; los grados de libertad ν_1 y ν_2 , y Φ , una función del parámetro de desplazamiento, que es:

$$\Phi = \sqrt{\frac{\lambda}{t}} = \sqrt{\frac{r \sum_{i=1}^t \tau_i^2}{t \sigma^2}} \quad (2.25)$$

En la tabla IX del apéndice están graficadas las curvas de la potencia para algunos valores del parámetro de la distribución F desplazada.

Las gráficas se usan para estimar el número de réplicas necesario para valores dados de α , $1 - \beta$, σ^2 , ν_1 , ν_2 y Φ . Un valor de Φ requiere un valor de σ^2 y valores específicos de las medias de tratamiento que conducen al rechazo de la hipótesis nula. Dados $\mu_1, \mu_2, \dots, \mu_t$, las $\tau_i = (\mu_i - \bar{\mu})$ se evalúan para Φ en la ecuación (2.25).

Ejemplo 2.3 Número de réplicas para el estudio de amiodarone

Para experimentos futuros supongamos que el investigador del estudio con amiodarone del ejemplo 2.2 se interesó en rechazar la hipótesis nula con una potencia de por lo menos .95, a un nivel de significación de .05 si la temperatura de la oreja para el grupo de tratamiento con medicamento era 0.8°C , mientras que las medias con el excipiente y la solución salina eran 0.1°C y 0°C , respectivamente. El promedio de las medias de tratamiento $\bar{\mu}$ es 0.3°C , y los efectos del tratamiento son:

$$\begin{aligned}\tau_1 &= \mu_1 - \bar{\mu} = 0.8 - 0.3 = 0.5 \\ \tau_2 &= \mu_2 - \bar{\mu} = 0.1 - 0.3 = -0.2 \\ \tau_3 &= \mu_3 - \bar{\mu} = 0.0 - 0.3 = -0.3\end{aligned}$$

Por lo tanto, $\sum_i^t \tau_i^2 = 0.38$, y se puede usar $ECM = 0.22$ del análisis de varianza como estimación de σ^2 . De la ecuación (2.25) se evalúa:

$$\Phi^2 = \frac{r \sum_{i=1}^t \tau_i^2}{t \sigma^2} = \frac{r(0.38)}{3(0.22)} = r(0.58)$$

Los valores necesarios para la curva de potencia son $\nu_1 = (t - 1) = 2$, $\nu_2 = t(r - 1) = 3(r - 1)$ y $\alpha = .05$. Si se usa $r = 5$ como un valor de la primera prueba se obtiene $\Phi = \sqrt{2.9} = 1.7$ y $\nu_2 = 12$. De la tabla IX del apéndice, la potencia aproximada de la prueba es .65, que es inferior a la requerida de .90. Si se incrementa el número de réplicas a $r = 9$, se obtiene $\Phi = \sqrt{5.22} = 2.3$ y $\nu_2 = 24$, con una potencia resultante superior a .90. En apariencia se necesitarán nueve conejos en cada grupo de tratamiento para conseguir una potencia de al menos .90.

Es difícil especificar los efectos deseados para un conjunto de tratamientos completo. Tal vez sea más sencillo especificar la diferencia entre cualesquiera dos medias de tratamiento que fuera biológica o físicamente significativa. Supongamos que se desea detectar el nivel de significancia con una diferencia de $D = \mu_i - \mu_j$. En este caso el valor mínimo de Φ^2 es:

$$\Phi^2 = \frac{rD^2}{2t\sigma^2} \quad (2.26)$$

Comentarios: Existen programas de cómputo para estimar el número de réplicas con gran facilidad, aunque el investigador debe contar con estimaciones de σ^2 , la potencia deseada, los niveles de significación y la magnitud de los efectos para declararlos significativos.

Por lo común, es difícil estimar σ^2 . Si se dispone de resultados de experimentos anteriores similares al que se considera, entonces pueden emplearse las estimaciones de la varianza para lograr valores razonables.

Cuando no es posible realizar el experimento completo en un solo periodo, o en un mismo lugar, con el número de réplicas necesario, se puede repetir varias veces o en distintos lugares. Por ejemplo, si se requieren ocho réplicas y no se pueden realizar al mismo tiempo, entonces se ejecuta dos veces con cuatro réplicas.

EJERCICIOS PARA EL CAPÍTULO 2

- Se realizó un estudio de ingeniería de tránsito sobre los retrasos en las intersecciones con semáforos en las calles de una ciudad. Se usaron tres tipos de semáforo: 1) programado, 2) semiactivado y 3) activado. Se usaron cinco intersecciones para cada tipo de semáforo. La medida de retraso utilizada fue el promedio de tiempo que cada vehículo permanece detenido en cada intersección (segundos/vehículo). Los datos son los siguientes:

Programado	Semiactivado	Activado
36.6	17.5	15.0
39.2	20.6	10.4
30.4	18.7	18.9
37.1	25.7	10.5
34.1	22.0	15.2

Fuente. W. Reilly, C. Gardner y J. Kell (1976). "A technique for measurement of delay at intersections". *Technical Report FHWA-RD-76-135*, Federal Highway Administration, Office of R & D, Washington, D. C.

- Escriba el modelo lineal estadístico para este estudio y explique sus componentes.
- Establezca las suposiciones necesarias para un análisis de varianza de los datos.
- Calcule el análisis de varianza.
- Calcule las medias de mínimos cuadrados del retraso en el tránsito y sus errores estándar para cada tipo de semáforo.

- e. Calcule el intervalo de confianza del 95% estimado para las medias de los tipos de semáforo.
 - f. Pruebe la hipótesis de que no hay diferencia entre las medias de retraso para los tipos de semáforo; a un nivel de significación de .05, con la prueba F .
 - g. Escriba las ecuaciones normales para los datos.
2. Se llevó a cabo un experimento para probar los efectos de un fertilizante nitrogenado en la producción de lechuga. Se aplicaron cinco dosis diferentes de nitrato de amonio a cuatro parcelas (réplicas) en un diseño totalmente aleatorizado. Los datos son el número de lechugas cosechadas de la parcela.

<i>Tratamiento (lb N/acre)</i>	<i>Lechugas/parcela</i>
0	104, 114, 90, 140
50	134, 130, 144, 174
100	146, 142, 152, 156
150	147, 160, 160, 163
200	131, 148, 154, 163

Fuente: Dr. B. Gardner, Department of Soil and Water Science, University of Arizona.

- 4.
- a. Escriba el modelo lineal estadístico para este estudio y explique sus componentes.
 - b. Establezca las suposiciones necesarias para un análisis de varianza de los datos.
 - c. Calcule el análisis de varianza.
 - d. Calcule las medias de mínimos cuadrados y sus errores estándar para cada medida de nitrógeno.
 - e. Calcule el intervalo de confianza del 95% estimado para las medias de los niveles de nitrógeno.
 - f. Pruebe la hipótesis de que no hay diferencia entre las medias de los niveles de nitrógeno con una prueba F a un nivel de significancia de .05.
 - g. Escriba las ecuaciones normales para los datos.
 - h. Este experimento se llevó a cabo con un diseño totalmente aleatorizado de las parcelas en un arreglo rectangular. Muestre una aleatorización de los cinco tratamientos con nitrógeno de las 20 parcelas, usando una permutación aleatoria de los números 1 a 20.
3. Un fisiólogo de animales estudió la función pituitaria de las gallinas, bajo el régimen estándar de muda de pluma forzada que usan los productores de huevo para mantenerlas en producción. Se usaron 25 gallinas en el estudio. Cinco se utilizaron para la medición, una previa al régimen de muda forzada y una al final de cada una de las cuatro etapas del régimen. Las cinco etapas del régimen fueron: 1) premuda (control), 2) ayuno de 8 días, 3) 60 gramos de salvado al día durante 10 días, 4) 80 gramos de salvado al día por 10 días y 5) mezcla de malta durante 42 días. El objetivo era dar seguimiento a las respuestas fisiológicas asociadas con la función pituitaria de las gallinas durante el régimen para explicar por qué vuelven a producir después de una muda forzada. Uno de los compuestos medidos fue la concentración de suero T3. Los datos de la tabla son las medidas de suero T3 en las cinco gallinas sacrificadas al final de cada etapa del régimen.

<i>Tratamiento</i>	<i>Suero T3, (ng/dl) × 10⁻¹</i>				
Premuda	94.09,	90.45,	99.38,	73.56,	74.39
Ayuno	98.81,	103.55,	115.23,	129.06,	117.61
60 g de salvado	197.18,	207.31,	177.50,	226.05,	222.74
80 g de salvado	102.93,	117.51,	119.92,	112.01,	101.10
Mezcla de malta	83.14,	89.59,	87.76,	96.43,	82.94

Fuente: Dr. R. Chiasson y K. Krown, Department of Veterinary Science, University of Arizona.

- Escriba el modelo lineal estadístico para este estudio y explique las componentes del modelo.
 - Establezca las suposiciones necesarias para un análisis de varianza de los datos.
 - Calcule el análisis de varianza.
 - Calcule las medias de mínimos cuadrados y sus errores estándar para cada tratamiento.
 - Calcule un intervalo de confianza de 95% estimado para las medias de los tratamientos.
 - Pruebe la hipótesis de que no hay diferencia entre las medias de los cinco tratamientos con la prueba *F* a un nivel de significancia de .05.
 - Escriba las ecuaciones normales de los datos.
 - Este experimento se llevó a cabo en un diseño totalmente aleatorizado, con una gallina en cada una de las 25 jaulas. Proporcione una asignación aleatoria de los cinco tratamientos a las 25 jaulas, con una permutación aleatoria de los números 1 a 25.
4. Se recolectaron datos de estudiantes de pedagogía en cuanto a su uso de ciertas estrategias de enseñanza estudiadas antes de sus prácticas. Había 28 estudiantes que habían aprendido las estrategias (9 en 1979, 9 en 1980 y 10 en 1981). En 1978 había 6 profesores que no habían aprendido el uso de estas estrategias y se usaron como grupo de control. El investigador registró el número promedio de estrategias por semana que cada estudiante usaba durante sus prácticas. El investigador quería saber si el número de estrategias usadas variaba con el tiempo.

<i>Número promedio de estrategias diferentes usadas</i>			
<i>Control 1978</i>	<i>1979</i>	<i>1980</i>	<i>1981</i>
6.88	7.25	10.85	7.29
5.40	10.50	7.43	14.38
16.00	8.43	6.71	6.00
9.80	8.63	7.60	5.00
7.63	8.63	7.60	5.38
5.00	7.00	5.57	14.14
	11.13	8.71	9.25
	7.25	5.86	5.71
	10.38	7.20	7.25
		10.75	

Fuente: Dr A Knorr, Family and Consumer Resources, University of Arizona.

- a. Escriba el modelo lineal estadístico para este estudio y explique las componentes del modelo.
 - b. Establezca las suposiciones necesarias para un análisis de varianza de los datos.
 - c. Calcule el análisis de varianza.
 - d. Calcule las medias de mínimos cuadrados y sus errores estándar para cada tratamiento.
 - e. Calcule un intervalo de confianza del 95% estimado para las medias de los tratamientos.
 - f. Pruebe la hipótesis de que no hay diferencia entre las medias de los cuatro tratamientos, con la prueba F a un nivel de significancia de .05.
 - g. Escriba las ecuaciones normales de los datos.
5. En cierto estudio de calibración de espectroscopia de absorción atómica, las medidas de respuesta fueron las unidades de absorción de un instrumento según la cantidad de cobre diluido en una solución ácida. Se usaron cinco niveles de cobre con cuatro réplicas del nivel cero y dos réplicas de los otros cuatro niveles. En la siguiente tabla se dan los datos de espectroscopia para cada nivel de cobre como microgramos de cobre/mililitro de solución.

<i>Cobre (mg/ml)</i>				
0.00	0.05	0.10	0.20	0.50
0.045	0.084	0.115	0.183	0.395
0.047	0.087	0.116	0.191	0.399
0.051				
0.054				

Fuente: R. J. Carroll, C. H. Spiegelman y J. Sacks (1988), "A quick and easy multiple use calibration-curve procedure", *Technometrics* 30, 137-141.

- a. Escriba el modelo lineal estadístico para este estudio y explique las componentes del modelo.
 - b. Establezca las suposiciones necesarias para un análisis de varianza de los datos.
 - c. Calcule el análisis de varianza.
 - d. Calcule las medias de mínimos cuadrados y sus errores estándar para cada tratamiento.
 - e. Calcule un intervalo de confianza del 95% estimado para las medias de los tratamientos.
 - f. Pruebe la hipótesis de que no hay diferencia entre las medias de los cinco tratamientos, con la prueba F a un nivel de significancia de .05.
 - g. Escriba las ecuaciones normales de los datos.
 - h. Un técnico preparó por separado cada solución ácida. Para prevenir errores sistemáticos de preparación de las soluciones 1 a 12, las preparó en orden aleatorio. Muestre una orden de preparación aleatoria de las 12 soluciones usando una permutación aleatoria de los números 1 al 12.
6. Considere el experimento del ejercicio 3. Suponga que se perdieron algunas gallinas durante el transcurso del mismo, lo que dio como resultado el siguiente conjunto de observaciones.

<i>Tratamiento</i>	<i>Suero T3, (ng/dl) × 10⁻¹</i>				
Premuda	94.09,	90.45,	99.38,	73.56,	
Ayuno	98.81,	103.55,	115.23,	129.06,	117.61
60 g de salvado	197.18,	207.31,	177.50,		
80 g de salvado	102.93,	117.51,	119.92,	112.01,	101.10
Mezcla de malta	82.94,	83.14,	89.59,	87.76,	

- a. Escriba el modelo lineal estadístico para este estudio y explique las componentes del modelo.
 - b. Establezca las suposiciones necesarias para un análisis de varianza de los datos.
 - c. Calcule el análisis de varianza.
 - d. Calcule las medias de mínimos cuadrados y sus errores estándar para cada tratamiento. ¿Cómo afectó la pérdida de gallinas a las estimaciones de las medias?
 - e. Calcule un intervalo de confianza de 95% estimado para las medias de los tratamientos.
 - f. Pruebe la hipótesis de que no hay diferencia entre las medias de los cinco tratamientos; con la prueba F a un nivel de significancia de .05.
 - g. Escriba las ecuaciones normales de los datos.
7. Utilice los datos del ejercicio 3 para determinar cuántas galinas necesitaría el biólogo en cada tratamiento para rechazar la hipótesis nula a un nivel de significancia de .05, con una potencia de .90, si la diferencia entre el tratamiento de control y cualquier tratamiento nuevo es de 30 unidades de T3.
8. Use los datos del ejercicio 1 para determinar cuántas intersecciones necesita el ingeniero de tránsito con cada tipo de semáforo para rechazar la hipótesis nula a un nivel de significancia de .01, con una potencia de .90 si los retrasos medios respectivos en los tres tipos de señal fueron 20, 18 y 16 segundos.
9. Éste es un pequeño ejercicio para ayudarle a comprender cómo funciona el principio de mínimos cuadrados para proporcionar una suma de cuadrados mínima del error experimental.
- a. Con la ecuación (2.3) y los datos del ejercicio 2.1 calcule SCE en los siguientes casos:
 - (i) Use la observación más pequeña en cada grupo de tratamiento como $\hat{\mu}$, en la ecuación (2.3) para calcular SCE .
 - (ii) Utilice la observación mayor en cada grupo de tratamiento como $\hat{\mu}$, en la ecuación (2.3) para calcular SCE .
 - (iii) Use la observación media en cada grupo de tratamiento como $\hat{\mu}$, en la ecuación (2.3) para calcular CME .
 - (iv) Sustituya otro valor entre el medio y el mayor como $\hat{\mu}$, en la ecuación (2.3) para calcular CME .
 - (v) Sustituya otro valor entre el medio y el menor como $\hat{\mu}$, en la ecuación (2.3) para calcular CME .
 - b. Grafique CME contra $\hat{\mu}$, con los valores calculados de CME en las ordenadas y los valores de $\hat{\mu}$, en las abscisas.
 - c. ¿Cuál es el valor de CME en cada caso? ¿Qué caso proporciona el valor más pequeño de CME ?

10. Una de las suposiciones del modelo lineal usado para describir los datos de los grupos de tratamiento ($y_{ij} = \mu_i + e_{ij}$), es que las observaciones son aleatorias, independientes de la variable aleatoria Y .
- ¿Qué habría que hacer durante la realización del estudio para que las suposiciones de aleatoriedad e independencia fueran razonables si:
 - se trata de un estudio diseñado?
 - se trata de un estudio comparativo por observación?
 - Antes de poder decir que los estimadores de mínimos cuadrados son estimadores de varianza mínima y sin sesgos, ¿qué suposiciones adicionales deben hacerse sobre el modelo?
 - Antes de poder calcular los intervalos de confianza y las pruebas de hipótesis sobre el modelo, ¿qué suposición adicional debe hacerse?
 - Suponga que no se posible hacer las últimas dos suposiciones (las de los incisos b y c). ¿Cómo puede probar la hipótesis de que no hay diferencia entre las medias de los tratamientos?

Apéndice 2A.1 Valores esperados

El valor esperado de una variable aleatoria es su valor promedio. Si una variable aleatoria Y tiene una distribución de probabilidad con una media μ y una varianza σ^2 , el valor esperado de Y se define como $\mu = E(Y)$, donde $E(Y)$ se lee “valor esperado de Y ”.

La varianza de Y se define como $\sigma^2 = E(Y - \mu)^2$, que es el valor esperado del cuadrado de la diferencia entre Y y la media.

Si existen dos variables aleatorias, Y_1 y Y_2 , y $E(Y_1) = \mu_1$, y $E(Y_2) = \mu_2$, entonces la covarianza entre las dos variables Y_1 y Y_2 se define como:

$$\sigma_{12} = E[Y_1 - \mu_1](Y_2 - \mu_2)]$$

La covarianza indica la relación entre Y_1 y Y_2 . Si los valores grandes de Y_1 se asocian con valores grandes de Y_2 , entonces la covarianza es positiva. Si los valores de Y_1 se hacen más pequeños al aumentar los valores de Y_2 o viceversa, entonces la covarianza es negativa. Si los valores de Y_1 y Y_2 son independientes, la covarianza es cero.

El modelo de medias de celdas para un diseño totalmente aleatorizado es:

$$y_{ij} = \mu_i + e_{ij} \quad i = 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad (2A.1)$$

Se supone que los errores experimentales e_{ij} son variables aleatorias independientes con una media igual a cero, varianza σ^2 y covarianza cero entre cualesquiera dos errores.

La media o valor esperado de cualquier e_{ij} es $E(e_{ij}) = 0$, y la varianza de cualquier e_{ij} es $E(e_{ij}^2) = \sigma^2$. Como no hay covarianza entre las e_{ij} , la esperanza de un producto entre cualesquiera dos términos de error en el mismo grupo de tratamiento o en otros es $E(e_{ij} \cdot e_{mk}) = 0$, donde $i \neq m$ o $j \neq k$. Las μ_i son las medias poblacionales y son constantes respecto a los valores esperados. El valor esperado de una constante es la constante misma, o sea, $E(\mu_i) = \mu_i$.

El valor esperado de cualquier observación descrita por las medias de la celda se puede encontrar sustituyendo $\mu_i + e_{ij}$ en lugar de y_{ij} en la esperanza matemática:

$$E(y_{ij}) = E(\mu_i + e_{ij}) = E(\mu_i) + E(e_{ij}) = \mu_i \quad (2A.2)$$

Apéndice 2A.2 Cuadrados medios esperados

Los valores esperados son necesarios para *CME* y *CMT* en el análisis de varianza. Con cualquier número de réplicas de cada tratamiento, el cuadrado de la media para el error experimental se estima como $CME = SCE/(N - t)$, donde:

$$SCE = \sum_{i=1}^t \sum_{j=1}^{r_i} (y_{ij} - \bar{y}_i)^2 \quad (2A.3)$$

La esperanza matemática de *SCE* se encuentra con las sustituciones de $\bar{y}_{ij} = \mu_i + e_{ij}$ y $\bar{y}_i = \mu_i + \bar{e}_i$ en la ecuación (2A.3), donde $\bar{e}_i = (\frac{1}{r_i}) \sum_j e_{ij}$. La expresión que resulta es:

$$\begin{aligned} SCE &= \sum_{i=1}^t \sum_{j=1}^{r_i} [\mu_i + e_{ij} - (\mu_i + \bar{e}_i)]^2 = \sum_{i=1}^t \sum_{j=1}^{r_i} (e_{ij} - \bar{e}_i)^2 \\ &= \sum_{i=1}^t \sum_{j=1}^{r_i} e_{ij}^2 - \sum_{i=1}^t r_i \bar{e}_i^2 \end{aligned}$$

Dado que $E(\bar{e}_i^2) = \frac{1}{r_i^2} E(\sum_{j=1}^{r_i} e_{ij}^2) = \frac{1}{r_i^2} E(e_{i1}^2 + e_{i2}^2 + \dots + e_{ir_i}^2) = \frac{1}{r_i} \sigma^2$, la esperanza matemática de *CME* es:

$$\begin{aligned} E(SCE) &= \sum_{i=1}^t \sum_{j=1}^{r_i} E(e_{ij}^2) - \sum_{i=1}^t r_i E(\bar{e}_i^2) \\ &= N\sigma^2 - t\sigma^2 \\ &= (N - t)\sigma^2 \end{aligned}$$

y

$$E(MCE) = \frac{E(SCE)}{(N - t)} = \sigma^2 \quad (2A.4)$$

El cuadrado de la media para el tratamiento es $CMT = SCT/(t - 1)$, donde:

$$SCT = \sum_{i=1}^t r_i (\bar{y}_i - \bar{y})^2 \quad (2A.5)$$

y la esperanza matemática de CMT requiere las sustituciones de $\bar{y}_i = \mu_i + \bar{e}_i$, y $\bar{y}_{..} = \bar{\mu}_{..} + e_{..}$ en la ecuación (2A.5), donde $\bar{e}_{..} = \frac{1}{N} \sum_{i=1}^t \sum_{j=1}^{r_i} e_{ij}$ y $\bar{\mu}_{..} = \frac{1}{N} \sum_{i=1}^t r_i \mu_i$.

La expresión resultante es:

$$\begin{aligned} CMT &= \sum_{i=1}^t r_i (\mu_i + \bar{e}_i - \bar{\mu}_{..} - \bar{e}_{..})^2 \\ &= \sum_{i=1}^t r_i (\mu_i - \bar{\mu}_{..})^2 + \sum_{i=1}^t r_i (\bar{e}_i - \bar{e}_{..})^2 + 2 \sum_{i=1}^t r_i (\mu_i - \bar{\mu}_{..})(\bar{e}_i - \bar{e}_{..}) \end{aligned}$$

La expansión del segundo término de la última expresión conduce a:

$$\sum_{i=1}^t r_i (\bar{e}_i - \bar{e}_{..})^2 = \sum_{i=1}^t r \bar{e}_i^2 - N \bar{e}_{..}^2$$

Dado que:

$E(\bar{e}_{..}^2) = \frac{1}{N^2} E(\bar{e}^2) = \frac{1}{N^2} E(e_{11}^2 + e_{12}^2 + \dots + e_{rt}^2 + \text{productos cruzados}) = \frac{1}{N} \sigma^2$
y $E(\bar{e}_i^2) = \frac{1}{r_i} \sigma^2$, la esperanza matemática de CMT se encuentra mediante:

$$\begin{aligned} E(CMT) &= \sum_{i=1}^t r_i (\mu_i - \bar{\mu}_{..})^2 + \sum_{i=1}^t r_i E(\bar{e}_i^2) - N E(\bar{e}_{..}^2) \\ &= \sum_{i=1}^t r_i (\mu_i - \bar{\mu}_{..})^2 + t \sigma^2 - \sigma^2 \\ &= \sum_{i=1}^t r_i (\mu_i - \bar{\mu}_{..})^2 + (t - 1) \sigma^2 \end{aligned}$$

La esperanza matemática de CMT es:

$$E(CMT) = \frac{E(CMT)}{(t - 1)} = \sigma^2 + \frac{1}{(t - 1)} \sum_{i=1}^t r_i (\mu_i - \bar{\mu}_{..})^2 \quad (2A.6)$$

El valor esperado en la ecuación (2A.6) se puede expresar en términos de los efectos de los tratamientos con la sustitución $\tau_i = (\mu_i - \bar{\mu}_{..})$.

Si todas las $r_i = r$, entonces $\theta_t^2 = \sum_{i=1}^t (\mu_i - \bar{\mu}_{..})^2 / (t - 1)$ y

$$E(CMT) = \frac{E(SST)}{(t - 1)} = \sigma^2 + r \theta_t^2 \quad (2A.7)$$

3 Comparación de tratamientos

A través de los métodos que se introducen en este capítulo el análisis de varianza y las estimaciones de las medias de los grupos de tratamiento, mediante mínimos cuadrados, proporcionan la información básica necesaria para un análisis profundo de la hipótesis de investigación. Los métodos para el análisis profundo de las respuestas al diseño de tratamientos incluyen contrastes planeados entre los tratamientos, curvas de respuesta por regresión para factores de tratamiento cuantitativos, selección del mejor subconjunto de tratamientos, comparación de tratamientos con el control y comparaciones por pares entre las medias de tratamientos. Todos estos métodos involucran un conjunto de decisiones simultáneas que debe tomar el investigador. Esta inferencia estadística simultánea afecta los errores estadísticos de la misma. En este capítulo se estudian algunos de esos efectos y el control de los errores.

3.1 La comparación de tratamientos responde las preguntas de la investigación

La relación entre los objetivos de investigación y el diseño de tratamientos requiere la identificación de los mismos, en cuanto al papel que desempeñan, para la evaluación de la hipótesis. Cuando se realiza un experimento para contestar a preguntas específicas, los tratamientos se seleccionan de manera que las comparaciones entre ellos contesten esas preguntas. Por ejemplo, es posible obtener respuestas específicas del experimento de almacenamiento de carne del capítulo 2, sobre el efecto de las distintas condiciones de empaque respecto al desarrollo de bacterias mientras está almacenada. Los cuatro tratamientos de este experimento eran 1) empaque comercial, 2) al vacío, 3) mezcla de gases y 4) CO₂ puro. Las estadísticas del experimento se muestran en el cuadro 3.1.

Sobre las condiciones de almacenamiento se pueden hacer las siguientes preguntas:

- Para reducir el desarrollo de bacterias, ¿es más efectiva la creación de una atmósfera artificial que el aire ambiental del empaque comercial?

Cuadro 3.1 Resumen estadístico para el experimento de almacenamiento de carne del ejemplo 2.1

	<i>Tratamiento</i>			
	<i>Comercial</i>	<i>Al vacío</i>	<i>CO, O₂, N</i>	<i>CO₂</i>
$\hat{\mu}_t = \bar{y}_t$	7.48	5.50	7.26	3.36
$t = 4$	$r_t = 3$	$s^2 = CME = 0.116$ con 8 grados de libertad		

- ¿Son más efectivos los gases que el vacío total?
- ¿Es más efectivo el CO₂ que una mezcla de CO, O₂ y N?

3.2 Planeación de comparaciones entre los tratamientos

Las comparaciones o contrastes entre las medias de los tratamientos se pueden construir de manera que se respondan a las preguntas específicas sobre el experimento. Los **contrastos** son formas especiales de las funciones lineales de las observaciones (presentadas en el apéndice 3A). Un contraste entre medias se define como:

$$C = \sum_{i=1}^t k_i \mu_i = k_1 \mu_1 + k_2 \mu_2 + \cdots + k_t \mu_t \quad (3.1)$$

donde $\sum_{i=1}^t k_i = 0$.

Los contrastes adecuados para las tres preguntas específicas del experimento de carnes son:

- empaque comercial *vs.* atmósferas artificiales:

$$C_1 = \mu_1 - \frac{1}{3} (\mu_2 + \mu_3 + \mu_4)$$

- al vacío *vs.* gases:

$$C_2 = \mu_2 - \frac{1}{2} (\mu_3 + \mu_4)$$

- mezcla de gases *vs.* CO₂ puro:

$$C_3 = \mu_3 - \mu_4$$

El primer contraste es la diferencia entre la media del empaque comercial y el promedio de las medias de los otros tratamientos. La suma de los coeficientes (1, $-\frac{1}{3}$, $-\frac{1}{3}$, $-\frac{1}{3}$) es igual a cero, como debe ser en un contraste apropiado. El

segundo contraste es la diferencia entre la media del empaque al vacío y el promedio de las medias de los tratamientos de gas con coeficientes $(0, 1, -\frac{1}{2}, -\frac{1}{2})$, y el tercero es una diferencia entre las medias de los dos tratamientos de gas con coeficientes $(0, 0, 1, -1)$.

Se calculan las estimaciones de los contrastes, los errores estándar, las estimaciones de los intervalos de confianza y las pruebas de hipótesis sobre los contrastes, a partir de los datos observados. Cualquier contraste de las medias del tratamiento, $C = \sum k_i \mu_i$, se estima con el mismo contraste de las medias de los tratamientos observados, como:

$$C = \sum_{i=1}^t k_i \bar{y}_i = k_1 \bar{y}_{1.} + \cdots + k_t \bar{y}_{t.} \quad (3.2)$$

donde \bar{y}_i es la media del i -ésimo grupo de tratamiento.

Las estimaciones para los tres contrastes en el experimento de carnes son:

$$c_1 = \bar{y}_{1.} - \frac{1}{3} (\bar{y}_{2.} + \bar{y}_{3.} + \bar{y}_{4.}) = 7.48 - \frac{1}{2} (5.50 + 7.26 + 3.36) = 2.11$$

$$c_2 = \bar{y}_{2.} - \frac{1}{2} (\bar{y}_{3.} + \bar{y}_{4.}) = 5.50 - \frac{1}{2} (7.26 + 3.36) = 0.19$$

$$c_3 = \bar{y}_{3.} - \bar{y}_{4.} = 7.26 - 3.36 = 3.90$$

Evaluación de las estimaciones de contrastes

Errores estándar de los contrastes

La varianza de una estimación de contraste $c = \sum k_i \bar{y}_i$ se estima mediante:

$$s_c^2 = s^2 \left[\sum_{i=1}^t \frac{k_i^2}{r_i} \right] = s^2 \left[\frac{k_1^2}{r_1} + \frac{k_2^2}{r_2} + \cdots + \frac{k_t^2}{r_t} \right] \quad (3.3)$$

donde $s^2 = CME$ a partir del análisis de varianza. Si las r_i son iguales, el estimador de la varianza se simplifica a:

$$s_c^2 = \frac{s^2}{r} [k_1^2 + k_2^2 + \cdots + k_t^2] \quad (3.4)$$

El estimador del error estándar de un contraste es igual a la raíz cuadrada de la varianza:

$$s_c = \sqrt{s_c^2} \quad (3.5)$$

Para el experimento de almacenamiento de carnes, $CME = 0.116$ y todas las $r_i = 3$. Los estimadores de la varianza y del error estándar para el primer contraste son:

$$s_{c_1}^2 = \frac{0.116}{3} \left[1^2 + \left(-\frac{1}{3} \right)^2 + \left(-\frac{1}{3} \right)^2 + \left(-\frac{1}{3} \right)^2 \right] = 0.052$$

y

$$s_{c_1} = \sqrt{0.052} = 0.228$$

Las estimaciones del error estándar para el segundo y tercer contraste son $s_{c_2} = 0.241$ y $s_{c_3} = 0.278$.

Estimación del intervalo para los contrastes

El estimador del intervalo de confianza $100(1 - \alpha)\%$ para un contraste es:

$$c \pm t_{\alpha/2, (N-t)} (s_c) \quad (3.6)$$

Los grados de libertad correctos para el estadístico t de Student son los grados de libertad para la varianza usados al calcular el error estándar del contraste. El estimador de la varianza en el experimento de carnes es $CME = 0.116$, con $(N-t) = 8$ grados de libertad. La estimación de un intervalo de confianza del 95% para C_1 , usando $t_{.025,8} = 2.306$ es:

$$2.11 \pm (2.306)(0.228)$$

En la tabla 3.1 se muestran los intervalos del 95% estimados para todo el conjunto de contrastes.

Tabla 3.1 Estimaciones, errores estándar e intervalos de confianza estimados del 95% para los tres contrastes del experimento de carnes

Contraste	Estimación	Error	IC 95%	
		estándar	Inferior	Superior
C_1	2.11	0.228	1.58	2.64
C_2	0.19	0.241	-0.37	0.75
C_3	3.90	0.278	3.26	4.54

Una partición de la suma de cuadrados para los contrastes

Con el propósito de indicar qué parte de la variación en los datos se explica por el contraste específico, se calcula una suma de cuadrados para el contraste del tratamiento. La reducción de suma de cuadrados para la estimación de un contraste ($C = \sum k_i \mu_i$) se calcula a partir de las medias de los tratamientos:

$$SCC = \frac{\left(\sum_{i=1}^t k_i \bar{y}_i \right)^2}{\sum_{i=1}^t (k_i^2 / r_i)} \quad (3.6)$$

Si todas las r_i son iguales:

$$SCC = r \frac{\left(\sum_{i=1}^t k_i \bar{y}_i \right)^2}{\sum_{i=1}^t k_i^2} \quad (3.7)$$

Para este contraste existe 1 grado de libertad asociado con la suma de cuadrados. En la tabla 3.2 se muestran las particiones de la suma de cuadrados para los tres contrastes del experimento de carnes. Los cálculos se ilustran en el cuadro 3.2.

Tabla 3.2 Análisis de varianza con contrastes para $\log(\text{No./cm}^2)$ de bacterias sicotrópicas en el experimento de almacenamiento de carne

Fuente	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	11	33.80			
Tratamientos	3	32.87	10.96	94.58	.000
Error	8	0.93	0.12		
Contraste	GL	Contraste SC	Cuadrados medios	F	Pr > F
1 vs. otros	1	9.99	9.99	86.19	.000
2 vs. 3 y 4	1	0.07	0.07	0.62	.453
3 vs. 4	1	22.82	22.82	196.94	.000

Hipótesis sobre contrastes

La hipótesis nula usual establece que el contraste tiene un valor de cero. Por ejemplo, la hipótesis nula para el segundo contraste del experimento de almacenamiento de carne sería:

$$H_0: C_2 = \mu_2 - \frac{1}{2} (\mu_3 + \mu_4) = 0$$

o, de manera equivalente:

$$H_0: C_2 = 2\mu_2 - \mu_3 - \mu_4 = 0$$

Si la hipótesis nula es cierta, $C_2 = 0$, el desarrollo de bacterias promedio para los dos tratamientos con gas es el mismo que para el tratamiento al vacío.

Cuadro 3.2 Cálculo de las sumas de cuadrados para los contrastes del experimento de almacenamiento de carne

$$\bar{y}_1 = 7.48 \quad \bar{y}_2 = 5.50 \quad \bar{y}_3 = 7.26 \quad \bar{y}_4 = 3.36$$

	k_1	k_2	k_3	k_4
$C_1:$	3	-1	-1	-1
$C_2:$	0	2	-1	-1
$C_3:$	0	0	1	-1

$$SCC_1 = 3[3(7.48) - 5.50 - 7.26 - 3.36]^2/[3^2 + (-1)^2 + (-1)^2 + (-1)^2] \\ = 3(6.32)^2/12 = 9.99$$

$$SCC_2 = 3[2(5.50) - 7.26 - 3.36]^2/[2^2 + (-1)^2 + (-1)^2] \\ = 3(0.38)^2/6 = 0.07$$

$$SCC_3 = 3[7.26 - 3.36]^2/[1^2 + (-1)^2] \\ = 3(3.9)^2/2 = 22.82$$

Las dos formas de contrastes establecidas para la hipótesis nula son equivalentes, excepto en que los coeficientes de la última difieren de los de la primera en un múltiplo de 2. La elección es sólo asunto de preferencia personal.

Las hipótesis nulas para los tres contrastes en el experimento de carne son:

$$H_0: C_1 = 3\mu_1 - \mu_2 - \mu_3 - \mu_4 = 0$$

$$H_0: C_2 = 2\mu_2 - \mu_3 - \mu_4 = 0$$

$$H_0: C_3 = \mu_3 - \mu_4 = 0$$

La primera establece que no hay diferencia entre el empaque comercial y los empaques con atmósfera artificial; la segunda determina que no hay diferencia entre el empaque al vacío y los tratamientos con atmósfera de gas; y la tercera establece que no hay diferencia entre el CO₂ puro y la mezcla de CO, O₂ y N₂.

Pruebas sobre los contrastes con la prueba F

Las hipótesis alternas son diferencias distintas de cero, es decir, $H_a: C \neq 0$. La hipótesis nula, $H_0: C = 0$, se prueba con:

$$F_0 = \frac{CMC}{CME} \tag{3.9}$$

donde $CMC = SCC$ es el cuadrado medio con 1 grado de libertad para el contraste. Bajo la hipótesis nula, F_0 tiene distribución F con 1 y $(N - t)$ grados de libertad. La hipótesis nula se rechaza al nivel de significancia α si $F_0 > F_{\alpha, 1, (N-t)}$.

El estadístico de prueba F_0 para cada contraste aparece en la columna “ F ” de la tabla 3.2. La columna “ $Pr > F$ ” muestra la probabilidad de exceder el estadístico F_0 observado. El valor crítico de F_0 a un nivel de significancia $\alpha = .05$ es $F_{.05,1,8} = 5.32$. La hipótesis nula se rechaza para los contrastes C_1 y C_3 , $Pr > F = .000$, y no se rechaza para el contraste C_2 , $Pr > F = .453$.

Como el contraste C_1 tiene una estimación positiva, $c_1 = 2.11$, se puede concluir que el número medio de bacterias en la carne con empaque comercial convencional ($\bar{y}_1 = 7.48$) era mayor que el promedio en la carne empacada en atmósferas artificiales, $\frac{1}{3}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4) = \frac{1}{3}(5.50 + 7.26 + 3.36) = 5.37$. El contraste C_3 tiene una estimación positiva, $c_3 = 3.90$, y se concluye que hay menos bacterias en la carne empacada en CO_2 ($\bar{y}_4 = 3.36$) que en la carne empacada en una atmósfera con mezcla de gases ($\bar{y}_3 = 7.26$).

Pruebas sobre los contrastes con la prueba t de Student

Al igual que con la prueba F , los contrastes entre tratamientos se pueden examinar con la prueba t de Student. La relación entre la distribución t de Student y la distribución F es:

$$t^2 = F \quad (3.10)$$

donde la t de Student tiene v grados de libertad y el estadístico F tiene 1 grado de libertad en el numerador y v grados de libertad en el denominador. La relación siempre es válida cuando hay 1 grado de libertad en el numerador del estadístico F . Los grados de libertad para la t de Student serán los mismos que para el denominador del estadístico F .

Para cualquier estimación del contraste c de un tratamiento, con un error estándar s_c , la razón:

$$t_0 = \frac{c}{s_c} \quad (3.11)$$

tiene la distribución t de Student, bajo la hipótesis nula $H_0: C = 0$.

Por ejemplo, la estimación para el primer contraste del experimento de carnes era $c_1 = 2.11$, con un error estándar $s_{c_1} = 0.228$ (tabla 3.1). La razón:

$$t_0 = \frac{2.11}{0.228} = 9.254$$

prueba la hipótesis nula de que no hay diferencia entre el empaque comercial y los empaques en atmósfera de gas. La hipótesis nula se rechaza al nivel de significancia de .05, ya que $t_0 > t_{.025,8} = 2.306$.

¿Intervalos de confianza o valores P?

El valor P resume en un solo valor la significancia o no significancia de las pruebas de hipótesis sobre los contrastes, sin embargo, no incluye información cuantitativa sobre estos últimos. Un valor P pequeño, que indica significancia, puede

indicar o no que un contraste tiene diferencias representativas con el valor supuesto en la hipótesis.

Por el contrario, los intervalos de confianza proporcionan información cuantitativa de los contrastes, al igual que su estado con respecto a las hipótesis sobre sus valores. Los resultados del estudio de amiodarone del ejemplo 2.2 se pueden usar para ilustrar los beneficios de los intervalos de confianza.

Recuerde que el incremento estimado en la temperatura de las orejas de conejos tratados con amiodarone era 1.20°C , un valor con significado clínico, mientras que las estimaciones respectivas para las soluciones excipiente y salina eran 0.13°C y 0.000°C , que no eran significativas en el sentido clínico.

Si el amiodarone incrementara la temperatura más que la solución excipiente, entonces se presumiría que contribuye a la inflamación de los tejidos. De la misma manera, la comparación de la solución excipiente con la salina proporcionaría información sobre la contribución del excipiente a la inflamación de tejidos.

La siguiente tabla muestra dos estimaciones de contrastes junto con sus errores estándar, intervalos de confianza del 95% y valores P para la prueba t de Student o la prueba F . El caso A muestra los resultados reales del estudio. El caso B muestra una situación en la que los valores P son los mismos pero las estimaciones, los errores estándar y los límites del intervalo de confianza del 95% se han triplicado de manera ficticia.

<i>Contraste</i>	<i>Caso</i>	c	s_c	(L, U)	<i>Valor P</i>
Amiodarone-excipiente	A	1.07	0.25	(0.55, 1.59)	.0036
	B	3.21	0.75	(1.65, 4.77)	.0036
Excipiente-salina	A	0.13	0.25	(−0.39, 0.65)	.6088
	B	0.39	0.75	(−1.17, 1.95)	.6088

La estimación de un intervalo de confianza del 95% para el amiodarone contra el excipiente en el caso A, indica que el amiodarone elevó en forma considerable la temperatura de la oreja en comparación con el excipiente con un límite inferior positivo y un límite superior de 1.59°C , que es clínicamente significativo. En el caso B, la comparación y su error estándar, muestran un incremento al triple, con un incremento del intervalo de confianza a 4.77°C , el triple del límite superior, que es mucho más alto que para el caso A. Sin embargo, el valor P en ambos casos es el mismo: .0036, que sólo indica que el contraste es diferente de 0 y no proporciona información sobre la magnitud y dirección de la diferencia.

La estimación de un intervalo de confianza del 95% para el excipiente contra la solución salina en el caso A, no indica diferencia entre las dos soluciones, los límites incluyen el 0 y no son significativos en el sentido clínico. El caso B de nuevo muestra el incremento al triple en la estimación del contraste, el error estándar y los límites del intervalo de confianza. Sin embargo, en este último caso los límites exceden la significancia clínica. Un valor de .6088 en ambos casos indica que no hay diferencia significativa con cero y no indica nada más. El caso A no tiene significancia clínica, pero es evidente que los límites del intervalo estimado en el caso B exceden la significancia clínica. Así, en el caso B, quizás el investiga-

dor quiera considerar otras opciones para lograr estimaciones más precisas en la comparación de los tratamientos excipiente y salino.

Los contrastes ortogonales expresan información independiente

Cierto tipo de comparaciones, conocidas como **contrastos ortogonales**, tienen propiedades especiales con respecto a la partición de la suma de cuadrados en el análisis de varianza y en cuanto a sus interrelaciones. La *ortogonalidad* implica que un contraste no contiene información sobre otro. El contraste c_2 , entre el empaque al vacío y las dos atmósferas de gases, no proporciona información sobre c_3 , la comparación entre las dos atmósferas de gas.

Supongamos que existen dos contrastes c y d , donde:

$$c = \sum_{i=1}^t k_i \bar{y}_{i\cdot} = k_1 \bar{y}_{1\cdot} + k_2 \bar{y}_{2\cdot} + \cdots + k_t \bar{y}_{t\cdot}$$

y

$$d = \sum_{i=1}^t d_i \bar{y}_{i\cdot} = d_1 \bar{y}_{1\cdot} + d_2 \bar{y}_{2\cdot} + \cdots + d_t \bar{y}_{t\cdot}$$

Los contrastes c y d son ortogonales si:

$$\sum_{i=1}^t \frac{k_i d_i}{r_i} = \frac{k_1 d_1}{r_1} + \frac{k_2 d_2}{r_2} + \cdots + \frac{k_t d_t}{r_t} = 0 \quad (3.12)$$

La suma ponderada del producto cruzado de los coeficientes k_i y d_i debe ser 0. Cuando todas las r_i son iguales, entonces c y d son ortogonales si $\sum k_i d_i = 0$.

Existen $(t - 1)$ contrastes mutuamente ortogonales entre t medias de tratamientos; cada par será ortogonal. Por ejemplo, hay tres posibles contrastes ortogonales en el conjunto de contrastes entre cuatro tratamientos.

Los coeficientes para dos de los contrastes del experimento de almacenamiento de carne, c_2 y c_3 , con igual número de réplicas se muestran en el cuadro 3.3 junto con sus productos cruzados. Los dos contrastes son ortogonales porque sus productos cruzados suman 0. Cálculos análogos muestran que el contraste c_1 , $(3, -1, -1, -1)$, también es ortogonal a c_2 y c_3 .

Cuadro 3.3 Contrastes ortogonales para el experimento de carnes

Contraste	Coeficientes			
	k_1	k_2	k_3	k_4
c_2	0	2	-1	-1
c_3	0	0	1	-1
Suma de productos cruzados = $(0)(0) + (2)(0) + (-1)(1) + (-1)(-1) = 0$				

Otra característica especial de los contrastes ortogonales es que la sumatoria de las sumas de cuadrados de los $(t - 1)$ contrastes ortogonales es igual a la suma de cuadrados de tratamientos del análisis de varianza. La sumatoria de las sumas de cuadrados para los tres contrastes ortogonales del ejemplo (c_1 , c_2 y c_3) es igual a la suma de cuadrados de tratamientos con error de redondeo $SCT = 32.87$, como se ve en la tabla 3.2.

Contrastes entre tratamientos con diferente número de réplicas

El experimento de amiodarone del ejemplo 2.2 tuvo diferente número de réplicas para los tratamientos. El resumen de estadísticas para el experimento se muestran en el cuadro 3.4.

Cuadro 3.4 Resumen de estadísticas para el experimento de amiodarone del ejemplo 2.2

	<i>Tratamiento</i>		
	<i>Amiodarone</i>	<i>Vehículo</i>	<i>Salina</i>
$\hat{\mu}_i = \bar{y}_i$	1.20	0.13	0.00
r_i	9	6	8
$t = 3 \quad N = 23 \quad s^2 = CME = 0.2177$ con 20 grados de libertad			

Dos comparaciones de interés pueden ser 1) el contraste de la media del tratamiento de amiodarone con las medias del excipiente y la solución salina y 2) el contraste entre las medias del excipiente y la solución salina, que se estiman con:

$$c_1 = 2\bar{y}_{1.} - \bar{y}_{2.} - \bar{y}_{3.}$$

y

$$c_2 = \bar{y}_{2.} - \bar{y}_{3.}$$

Los coeficientes de contraste son $(2, -1, -1)$ y $(0, 1, -1)$ con números de réplicas $(9, 6, 8)$. La evaluación de los contrastes mediante la ecuación (3.12),

$$\frac{2(0)}{9} + \frac{-1(1)}{6} + \frac{-1(-1)}{8} \neq 0$$

indica que no son ortogonales.

Para que los contrastes sean ortogonales, tendrían que alterarse los valores de los coeficientes. Por ejemplo, dos contrastes para el experimento de amiodarone que satisfacen el criterio de ortogonalidad de la ecuación (3.12) son:

$$c_1 = 7\bar{y}_{1.} - 3\bar{y}_{2.} - 4\bar{y}_{3.}$$

y

$$c_2 = \bar{y}_{2.} - \bar{y}_{3.}$$

Los coeficientes $(7, -3, -4)$ y $(0, 1, -1)$ satisfacen la condición de un contraste con $\sum k_i = 0$. La evaluación con la ecuación (3.12) es:

$$\frac{7(0)}{9} + \frac{-3(1)}{6} + \frac{-4(-1)}{8} = 0$$

Ahora ambos contrastes son ortogonales y no contienen información de otro.

Sin embargo, considerando el primer contraste, $c_1 = 7\bar{y}_1 - 3\bar{y}_2 - 4\bar{y}_3$, los coeficientes implican una comparación ponderada entre las medias de las poblaciones. A menos que las poblaciones de los tratamientos ocurran en una proporción 7:3:4, la comparación no tiene sentido. Se tiene información desigual en los grupos de tratamiento dentro del experimento; pero no necesariamente ocurre que las poblaciones de los tratamientos tengan una representación desigual en la naturaleza. En el caso del experimento con amiodarone, el contraste con valores relativos iguales para las medias de los tratamientos es una comparación más razonable.

Comentarios sobre ortogonalidad

Para presentar una tabla de análisis de varianza ordenado, la elección de contrastes para el estudio no debe determinarse por su ortogonalidad; de preferencia deben construirse para obtener respuestas a preguntas de investigación específicas. Por ejemplo, en el experimento de carnes, puede ser de interés comparar el tratamiento de control por separado con cada uno de los demás; esto daría un conjunto de comparaciones no ortogonales, $\mu_1 - \mu_2$, $\mu_1 - \mu_3$ y $\mu_1 - \mu_4$. Las comparaciones ortogonales se eligieron a propósito en este experimento para simplificar la presentación. Sin embargo, contestaron algunas preguntas significativas, que en último caso es el objetivo de la investigación. La hipótesis de investigación y el diseño de tratamientos debe dictar la construcción de los contrastes.

3.3 Curvas de respuesta para los factores de tratamiento cuantitativos

Se han realizado muchos estudios para determinar la relación de la tendencia cuantitativa entre dos variables. En los estudios experimentales, es común que una de las variables esté bajo *control* del investigador mientras que la otra es la *respuesta observada*. Los diseños de tratamiento factorial, presentados en la sección 1.4, son ejemplos del diseño de tratamientos estructurados en los que el factor cuantitativo puede tener varios niveles de graduación.

Por ejemplo, se puede diseñar un experimento para medir el crecimiento de animales al aumentar las cantidades de nutrientes en su dieta. El diseño del tratamiento consiste en un factor cuantitativo, la cantidad de nutrientes en una dieta con tres niveles, digamos 0, 500 y 1,000 partes por millón (ppm). El objetivo del estudio será determinar con precisión el crecimiento de los animales como función de la cantidad de nutrientes en la dieta.

Función de respuesta polinomial para las curvas de respuesta

El modelo polinomial que se usa con frecuencia para describir las relaciones de tendencia entre la respuesta medida y y los niveles cuantitativos de un factor x es:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_p x^p + e \quad (3.13)$$

En la figura 3.1 se muestran varios ejemplos de curvas de respuesta. Los polinomios lineal, cuadrático y cúbico proporcionan buenas aproximaciones a muchas relaciones comunes en las ciencias biológicas y físicas.

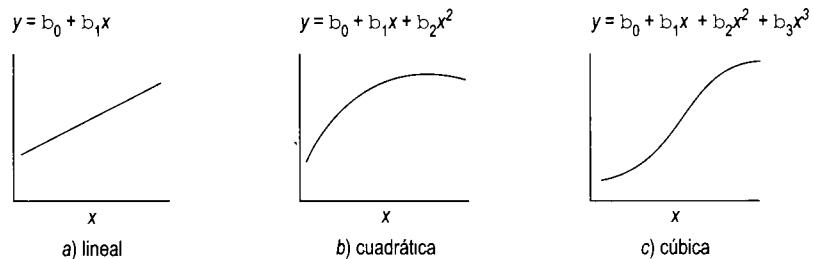


Figura 3.1 Ejemplos de curvas de respuesta polinomial: a) lineal, b) cuadrática y c) cúbica

Ejemplo 3.1 Producción de grano y densidad de la población de plantas

El crecimiento y desarrollo de las plantas, como el de cualquier organismo vivo, depende de la disposición de nutrición suficiente. El objetivo de la mayor parte de los cultivos de grano es maximizar la producción de semilla o grano en el régimen dado de fertilización de tierra o cultivo. Si el productor planta pocas semillas por unidad de área, no tendrá la máxima producción de grano, pues cada planta tiene un potencial genético máximo determinado. Entonces, una mayor producción de grano requiere más plantas por unidad de área. Por otro lado, aunque la razón es completamente distinta, no obtendrá la máxima producción si el número de plantas por unidad de área es excesivo. La disponibilidad de nutrientes por planta queda limitada y su crecimiento y desarrollo se reducen, con la consecuente reducción en la producción de grano.

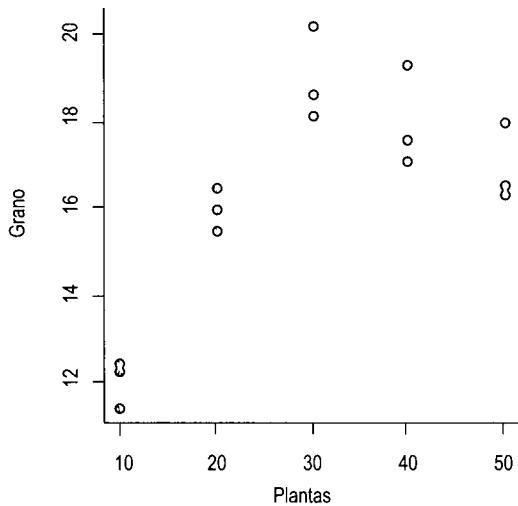
Para un científico, el objetivo en estas circunstancias es determinar la relación entre el número de plantas por unidad de área y la producción de grano, con un régimen de fertilización y cultivo dado.

Se elaboró un experimento para estimar la curva de respuesta polinomial que caracteriza esta relación. El diseño de tratamiento consistió en cinco densidades de plantas (10, 20, 30, 40 y 50). Cada uno de los cinco tratamientos se asignó al azar entre las tres parcelas con un diseño de experimento con muestras totalmente aleatorio. Las cosechas de grano resultantes se muestran en la tabla 3.3.

Tabla 3.3 Cantidad de grano producida por parcela, para cinco densidades de plantas

<i>Densidad de plantas (x)</i>					
	10	20	30	40	50
12.2	16.0	18.6	17.6	18.0	
11.4	15.5	20.2	19.3	16.4	
12.4	16.5	18.2	17.1	16.6	
Medias (\bar{y}_i)	12.0	16.0	19.0	18.0	17.0

Una gráfica de las observaciones en la figura 3.2 sugiere una relación cuadrática entre la cosecha de grano y la densidad de plantas. En este estudio, las respuestas medias de la cosecha (\bar{y}_i) con la densidad de plantas (x) mostrada en la tabla 3.3, también sugiere un polinomio cuadrático de respuesta como modelo aproximado para describir la relación biológica entre la cosecha de grano y la densidad de plantas. El objetivo es determinar la ecuación polinomial de menor orden posible que describa la relación en forma adecuada. Es posible ajustar una ecuación de cuarto grado en x , con cinco valores de x ; sin embargo, puede funcionar igual una ecuación menos compleja de segundo grado.

**Figura 3.2** Cosecha de grano contra densidad de plantas

Simplificación con polinomios ortogonales

El modelo polinomial se puede ajustar a los datos observados con muchos de los programas de regresión disponibles para computadora. Es posible simplificar el análisis de tendencia si se examinan los contrastes ortogonales entre los niveles de los factores de tratamiento que miden los efectos polinomiales lineal, cuadrático y de

orden más alto. Estos contrastes, conocidos como **polinomios ortogonales**, permiten evaluar la importancia de cada componente polinomial en un caso específico.

En el ejemplo 3.1 existen $t = 5$ niveles del factor de densidad de plantas y se pueden estimar $(t - 1) = 4$ contrastes ortogonales. Después de transformar los polinomios de x en polinomios ortogonales, las ecuaciones del modelo polinomial ortogonal completo para la relación entre la densidad de plantas y la cosecha de grano es:

$$y_{ij} = \mu + \alpha_1 P_{1i} + \alpha_2 P_{2i} + \alpha_3 P_{3i} + \alpha_4 P_{4i} + e_{ij} \quad (3.14)$$

donde μ es la gran media y P_{ci} es el polinomio ortogonal de c -ésimo orden, para el i -ésimo nivel del factor de tratamiento.

En el cuadro 3.5 se muestran las transformaciones de las potencias de x en un polinomio ortogonal (P_{ci}) hasta de tercer grado. Los valores tabulados de los polinomios ortogonales para $t = 3$ a 10 se dan en la tabla XI del apéndice. Los valores del cuadro 3.5 o de la tabla XI son válidos para cualquier distancia entre los valores de x (o niveles del factor) siempre que el espaciamiento sea igual entre todos ellos y el número de réplicas sea el mismo para todos los tratamientos. La constante λ_i al principio de cada transformación convierte los valores de P_i en valores enteros.

Cuadro 3.5 Transformación de las potencias de x en polinomios ortogonales

$$\text{Media: } P_0 = 1$$

$$\text{Lineal: } P_1 = \lambda_1 \left[\frac{x - \bar{x}}{d} \right]$$

$$\text{Cuadrática: } P_2 = \lambda_2 \left[\left(\frac{x - \bar{x}}{d} \right)^2 - \left(\frac{t^2 - 1}{12} \right) \right]$$

$$\text{Cúbica: } P_3 = \lambda_3 \left[\left(\frac{x - \bar{x}}{d} \right)^3 - \left(\frac{x - \bar{x}}{d} \right) \left(\frac{3t^2 - 7}{20} \right) \right]$$

t = número de niveles del factor

x = valor del factor

\bar{x} = media de los niveles del factor

d = distancia entre los niveles del factor

Nótese en el cuadro 3.5 que el término $(x - \bar{x})/d$ ocurre constantemente en todas las transformaciones a polinomios ortogonales. La transformación $(x - \bar{x})$ centra los valores de x alrededor de 0, mientras que la división entre el incremento de los valores de x coloca los valores en una escala de incrementos de una unidad. En el ejemplo 3.1 con $\bar{x} = 30$ y $d = 10$ la transformación que se obtiene es:

$x:$	10	20	30	40	50
$(x - 30):$	-20	-10	0	10	20
$(x - 30)/10:$	-2	-1	0	1	2

Algunos paquetes de cómputo para el análisis estadístico pueden producir polinomios ortogonales con espaciamientos iguales o desiguales entre los valores de x . Grandage (1958) proporciona un método manual para derivar polinomios ortogonales con espaciamiento desigual.

La transformación de las densidades de plantas en polinomios ortogonales se muestra en la tabla 3.4 con $d = 10$, $t = 5$, $\bar{x} = 30$ y $x = 10, 20, 30, 40$ y 50 . Cada conjunto de coeficientes de P_1 a P_4 forma una comparación entre los tratamientos, ya que la suma de todos los coeficientes en las columnas es igual a 0. Los contrastes son mutuamente ortogonales.

Tabla 3.4 Cálculos para contrastes con polinomios ortogonales y sumas de cuadrados

Densidad (x)	\bar{y}_t	Coeficientes de los polinomios ortogonales (P_{ci})				
		Media	Lineal	Cuadrática	Cúbica	Cuarta
10	12	1	-2	2	-1	1
20	16	1	-1	-1	2	-4
30	19	1	0	-2	0	6
40	18	1	1	-1	-2	-4
50	17	1	2	2	1	1
λ_c		-	1	1	5/6	35/12
Suma = $\sum P_{ci}\bar{y}_t$	82	12	-14	1	7	
Divisor = $\sum P_{ci}^2$	5	10	14	10	70	
$SCP_c = r(\sum P_{ci}\bar{y}_i)^2 / \sum P_{ci}^2$	-	43.2	42.0	0.3	2.1	
$\hat{\alpha}_C = \sum P_{ci}\bar{y}_i / \sum P_{ci}^2$	16.4	1.2	-1.0	0.1	0.1	

Se puede hacer una partición de la suma de cuadrados de los tratamientos, en un conjunto aditivo de suma de cuadrados con 1 grado de libertad para cada uno de los $(t - 1)$ contrastes polinomiales ortogonales. En consecuencia, es posible examinar de manera secuencial la significancia de los términos lineal, cuadrático, cúbico, etcétera, en este modelo, para determinar que la ecuación polinomial se ajuste mejor.

En la tabla 3.4 se muestran las estimaciones de los coeficientes α_C para la ecuación polinomial ortogonal en la ecuación (3.14) y la suma de cuadrados para cada contraste polinomial ortogonal. La ecuación polinomial ortogonal estimada se encuentra sustituyendo las estimaciones de μ , α_1 , α_2 , α_3 y α_4 de la tabla 3.4 en la ecuación (3.14). La ecuación estimada resultante es:

$$\hat{y}_i = 16.4 + 1.2P_{1i} - 1.0P_{2i} + 0.1P_{3i} + 0.1P_{4i} \quad (3.15)$$

El análisis de varianza para este experimento se muestra en la parte superior de la tabla 3.5. La razón $F_0 = CMT/CME$ examina la hipótesis nula global $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. Para un nivel de significancia de .05, la región crítica es $F_0 > F_{.05,4,10} = 3.48$. La hipótesis nula se rechaza, pues $F_0 = 29.28$ excede el valor crítico.

Las particiones de suma de cuadrados con 1 grado de libertad para cada contraste polinomial ortogonal se resumen en el análisis de varianza de la parte infe-

rior de la tabla 3.5. Observe que las sumas de cuadrados para los cuatro contrastes suman lo mismo que la suma de cuadrados para la densidad: 87.60, con 4 grados de libertad.

Tabla 3.5 Análisis de varianza para la relación del modelo polinomial ortogonal entre la densidad de plantas y la cosecha de grano

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Densidad	4	87.60	21.90	29.28	.000
Error	10	7.48	0.75		
Contraste	GL	SC del contraste	Cuadrados medios	F	Pr > F
Lineal	1	43.20	43.20	57.75	.000
Cuadrática	1	42.00	42.00	56.15	.000
Cúbica	1	.30	.30	.40	.541
Cuarta	1	2.10	2.10	2.81	.125

El interés se centra en la contribución de cada término polinomial del modelo. Una estrategia para determinar la mejor ecuación polinomial es probar la significancia de los términos en la secuencia: lineal, cuadrática, cúbica y así sucesivamente. Primero se estudia el polinomio más sencillo y se va construyendo uno más complicado conforme los datos lo requieran para una descripción adecuada.

La secuencia de las hipótesis es $H_0: \alpha_1 = 0, H_0: \alpha_2 = 0, H_0: \alpha_3 = 0$, etcétera. Se aplica la prueba F respectiva a cada hipótesis sobre los contrastes polinomiales ortogonales. En la tabla 3.5 se dan las razones $F_0 = CMC/CME$ para cada contraste polinomial estimado en el estudio de densidades de plantas. Para cada partición de suma de cuadrados, la hipótesis nula es $H_0: \alpha_C = 0$, con región crítica $F_0 > F_{.05,1,10} = 4.96$.

La hipótesis nula se rechaza para los términos lineal y cuadrático del modelo ($Pr > F = .000$), pero no para los términos cúbico ($Pr > F = .541$) y de cuarto orden ($Pr > F = .125$). Según las pruebas estadísticas el modelo cuadrático es suficiente para describir la relación entre la cosecha de grano y la densidad de plantas.

Cálculo de la curva de respuesta

La curva de respuesta cuadrática estimada sin los términos cúbico y cuarto, $\widehat{\alpha}_3 P_3$ y $\widehat{\alpha}_4 P_4$ de la ecuación (3.15) es:

$$\widehat{y}_i = \bar{y}_{..} + \widehat{\alpha}_1 P_{1i} + \widehat{\alpha}_2 P_{2i} = 16.4 + 1.2P_{1i} - 1.0P_{2i}, \quad (3.16)$$

El valor estimado para una densidad de plantas $x = 10$ se determina sustituyendo $P_1 = -2$ y $P_2 = 2$ en la ecuación (3.16), es decir, $\widehat{y} = 16.4 + 1.2(-2) - 1.0(2) = 12.0$. En la tabla 3.6 se muestran las cosechas de grano observadas y sus estimaciones, a partir de la ecuación (3.16), para todas las densidades de plantas.

Tabla 3.6 Cosechas de grano observadas y sus estimaciones a partir de la ecuación polinomial cuadrática

Densidad	Coeficientes		Estimación	Observación
x	P_1	P_2	\hat{y}_i	\bar{y}_i
10	-2	2	12.0	12
20	-1	-1	16.2	16
30	0	-2	18.4	19
40	1	-1	18.6	18
50	2	2	16.8	17

La relación polinomial expresada como función de y y x en unidades reales de las variables observadas, es más informativa que cuando se expresa en unidades del polinomio ortogonal. Una transformación directa para una ecuación en x , requiere la información contenida en el cuadro 3.5 y en la tabla 3.4. Las cantidades necesarias son $\lambda_1 = 1$, $\lambda_2 = 1$, $d = 10$, $\bar{x} = 30$ y $t = 5$. Entonces, si se sustituyen las P_i , se tiene:

$$\begin{aligned}\hat{y} &= 16.4 + 1.2P_1 - 1.0P_2 \\ &= 16.4 + 1.2(1) \left[\frac{x - 30}{10} \right] - 1.0(1) \left[\left(\frac{x - 30}{10} \right)^2 - \left(\frac{5^2 - 1}{12} \right) \right]\end{aligned}$$

simplificando:

$$\hat{y} = 5.8 + 0.72x - 0.01x^2 \quad (3.17)$$

La curva estimada a partir de la ecuación (3.17) puede graficarse como se muestra en la figura 3.3, junto con los datos observados (cuadrados) y las medias de los tratamientos \bar{y}_i (puntos).

Errores estándar e intervalos de confianza para las curvas de respuesta

Las curvas de respuesta estimadas se componen de las estimaciones de varios parámetros. La ecuación cuadrática elegida como la mejor para la relación entre la producción de grano y la densidad de plantas tiene tres estimaciones: $\bar{y}_{..} = 16.4$, $\hat{\alpha}_1 = 1.2$, y $\hat{\alpha}_2 = -1.0$.

El estimador de la varianza de $\hat{y} = \bar{y}_{..} + \hat{\alpha}_1 P_1 + \hat{\alpha}_2 P_2$ es:

$$s_{\hat{y}}^2 = s_{\bar{y}_{..}}^2 + P_1^2 s_{\hat{\alpha}_1}^2 + P_2^2 s_{\hat{\alpha}_2}^2 \quad (3.18)$$

el estimador de la varianza para un contraste polinomial s_c^2 es:

$$s_c^2 = \frac{s^2}{(r \sum P_{ci}^2)} \quad (3.19)$$

y el estimador del error estándar para el contraste es la raíz cuadrada de la varianza

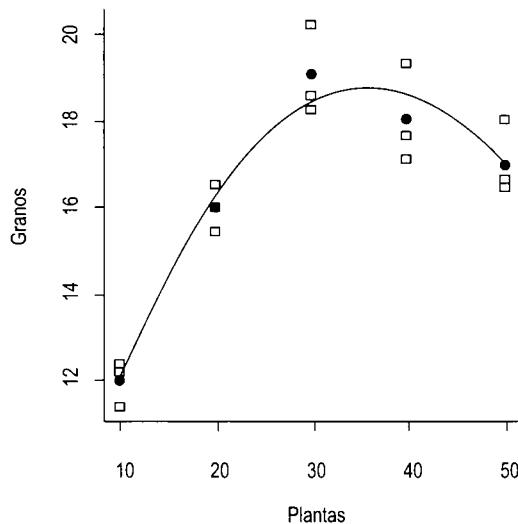


Figura 3.3 Curva de respuesta estimada, $\hat{y} = 5.8 + 0.72x - 0.01x^2$, para la cosecha de grano como función de la densidad de plantas

$$s_c = \sqrt{s_c^2} \quad (3.20)$$

Las estimaciones de la varianza para los coeficientes de la ecuación cuadrática estimada se calculan a partir del análisis de varianza de la tabla 3.5 mediante $s^2 = CME = 0.75$ y de los divisores de la tabla 3.4. La varianza es

$$s_{\bar{y}_{..}}^2 = \frac{0.75}{(3)(5)} = 0.050$$

$$s_{\bar{\alpha}_1}^2 = \frac{0.75}{(3)(10)} = 0.025$$

y

$$s_{\bar{\alpha}_2}^2 = \frac{0.75}{(3)(14)} = 0.018 \quad 3$$

La varianza de un valor estimado es

$$\begin{aligned} s_{\bar{y}}^2 &= s_{\bar{y}_{..}}^2 + s_{\bar{\alpha}_1}^2 P_1^2 + s_{\bar{\alpha}_2}^2 P_2^2 \\ &= .05 + 0.025 P_1^2 + 0.018 P_2^2 \end{aligned} \quad (3.21)$$

Como los valores de P_1 y P_2 difieren para cada valor de densidad de plantas, la varianza del valor estimado también difiere para cada uno de ellos. La varianza de

la cosecha estimada para la densidad de plantas $x = 10$, con $P_1 = -2$ y $P_2 = 2$ en la ecuación (3.21), es:

$$s_{\hat{y}}^2 = 0.05 + 0.025(-2)^2 + 0.018(2)^2 = 0.222$$

y el error estándar es:

$$s_{\hat{y}} = \sqrt{0.222} = 0.471$$

Un intervalo de confianza de $100(1 - \alpha)\%$ para el valor estimado se calcula a partir de:

$$\hat{y} \pm t_{\alpha/2, (N-1)}(s_{\hat{y}})$$

En la tabla 3.7 se muestran los errores estándar y los intervalos de confianza del 95% para los valores estimados de la cosecha de grano con las cinco densidades de plantas, cuando $t_{0.025, 10} = 2.228$.

Tabla 3.7 Cosechas de grano observada y estimada, y errores estándar de las cosechas de grano estimadas para cada densidad de plantas

Densidad	Cosecha de semilla		Error estándar	Intervalo de confianza del 95%
	Observada	Estimada		
10	12	12.0	0.471	(10.95, 13.05)
20	16	16.2	0.305	(15.52, 16.88)
30	19	18.4	0.349	(17.62, 19.18)
40	18	18.6	0.305	(17.92, 19.28)
50	17	16.8	0.471	(15.75, 17.85)

La curva de respuesta estimada (figura 3.3) tienen la ventaja de describir la relación entre y y x para todos los valores de x usados en el experimento. Con este ejemplo es posible describir o estimar la cosecha de grano para cualquier densidad entre 10 y 50 plantas. La descripción de la relación no está restringida, en el análisis de resultados, a las cinco densidades de plantas usadas en el estudio.

3.4 Las comparaciones múltiples afectan las tasas de error

El grupo de contrastes presentados en la sección 3.2 se considera una **comparación múltiple**, puesto que se hizo más de una comparación entre las medias de los tratamientos. Durante el análisis se puede construir el número necesario de estas comparaciones para ayudar a responder a las preguntas de la investigación.

Sin embargo, es posible que las comparaciones múltiples entre las medias de los tratamientos conduzcan al investigador, a veces sin saberlo, a un campo minado estadístico. La gran cantidad de procedimientos disponibles aumenta la dificultad para escoger el método adecuado en una situación específica.

Tasas de error en las comparaciones múltiples

Las dificultades con las comparaciones múltiples residen básicamente en la comprensión de las tasas de error asociadas con las hipótesis de las pruebas múltiples. Las pruebas de hipótesis tienen riesgos asociados con las decisiones de rechazar o no la hipótesis; para cualquier comparación entre medias, el riesgo asociado con declarar que un contraste es real cuando no lo es, es el riesgo de error tipo I. El riesgo de declarar el contraste entre las medias de poblaciones es igual a cero cuando es diferente, es el error tipo II.

Los riesgos de los errores tipo I y II tienen una relación inversa entre sí. El nivel de significancia elegido para la prueba de hipótesis determina el riesgo de error tipo I. El tamaño de la muestra, la varianza y el tamaño de la comparación para las medias poblacionales reales, determina la tasa de error tipo II para una tasa dada de error tipo I.

Una prueba sencilla para la diferencia entre dos medias de tratamientos es la prueba t de Student, con el estadístico calculado como:

$$t_0 = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{s^2 \left[\frac{1}{r_i} + \frac{1}{r_j} \right]}} \quad (3.22)$$

El nivel de significancia o probabilidad de error tipo I para una sola prueba es una tasa de error con *respecto a la comparación*, α_C . Es el riesgo que estamos dispuestos a correr en una sola comparación.

Existen $p(p - 1)/2$ comparaciones por pares entre p medias de los tratamientos. Por ejemplo, cuatro tratamientos (A, B, C, D) tienen $4(3)/2 = 6$ pares posibles: (A, B), (A, C), (A, D), (B, C), (B, D) y (C, D). Si se prueban los seis pares de medias con el estadístico t_0 de la ecuación (3.22), existe la posibilidad de cometer 0, 1, 2, 3, 4, 5 o 6 errores de tipo I, si las seis medias poblacionales son iguales.

Con la posibilidad de hasta seis errores tipo I para seis pruebas, se puede emplear otra forma de error tipo I basada en el riesgo acumulado asociado con la familia de pruebas en estudio. La familia es el conjunto de comparaciones por pares para el ejemplo del párrafo anterior. El riesgo acumulado asociado con una familia de comparaciones se conoce como tasa de error tipo I con *respecto al experimento*, α_E . Es el riesgo de cometer al menos un error tipo I en la familia de comparaciones en el experimento.

Evaluación de la tasa de error máxima

Para una familia de pruebas independientes, se puede evaluar la tasa de error tipo I con respecto al experimento. Sin embargo, no todas las pruebas por pares son independientes con la ecuación (3.22), ya que la s^2 en el denominador de cada estadístico t_0 es la misma, y el numerador de cada prueba contiene la misma media que varios de los otros estadísticos t_0 .

Aunque las pruebas en la familia descrita no son independientes, se puede evaluar el límite superior para el valor del error tipo I con respecto al experimento, si se suponen pruebas independientes. Suponiendo que las hipótesis nulas son ciertas

para cada una de las n pruebas, la probabilidad de un error tipo I para cualquier prueba sola es α_C (la tasa respecto a la comparación) con $(1 - \alpha_C)$ como la probabilidad de una decisión correcta. La probabilidad de cometer x errores tipo I está dada por la distribución binomial como:

$$P(x) = \frac{n!}{x!(n-x)!} \alpha_C^x (1 - \alpha_C)^{n-x} \quad (3.23)$$

para $x = 0, 1, 2, \dots, n$ errores tipo I. La probabilidad de no cometer errores tipo I es:

$$P(x = 0) = (1 - \alpha_C)^n$$

La probabilidad de cometer *al menos un* error tipo I ($x = 1, 2, 3, \dots, n$) es $P(x \geq 1) = 1 - P(x = 0)$, o sea:

$$\alpha_E = 1 - (1 - \alpha_C)^n \quad (3.24)$$

La probabilidad α_E es el riesgo de tener al menos un error tipo I entre las n comparaciones independientes. Es el límite superior de la tasa de error tipo I con respecto al experimento, para n pruebas en un conjunto de medias de tratamientos.

La relación:

$$\alpha_C = 1 - (1 - \alpha_E)^{1/n} \quad (3.25)$$

expresa la tasa de error tipo I con respecto a la comparación como una función de la tasa de error tipo I con respecto al experimento.

En la tabla 3.8 se muestra la relación entre las tasas de error tipo I para algunos valores de n . Si cada prueba se realiza con una tasa de error con respecto a la comparación de $\alpha_C = .05$, el riesgo de al menos un error tipo I se eleva cuando aumenta el número de pruebas. Si se realizan $n = 1$ pruebas, ambos errores tipo I son idénticos, como debe ser, ya que sólo puede cometerse un tipo de error. Cuando $n = 5$, la probabilidad del riesgo de incurrir en al menos un error tipo I en cinco decisiones aumenta a $\alpha_E = .226$, y con $n = 10$, el riesgo se eleva a una probabilidad de .401 de cometer al menos un error tipo I.

La última columna de la tabla 3.8 indica la tasa de error tipo I con respecto a la comparación, para mantener una tasa de error tipo I con respecto al experimento $\alpha_E = .05$. Por ejemplo, cuando se hacen cinco pruebas independientes y se desea mantener el riesgo de cometer al menos un error tipo I en 1 vez de cada 20, la tasa de error con respecto a la comparación de cada una de las $n = 5$ pruebas debe ser $\alpha_C = .01$, y para $n = 10$, debe ser $\alpha_C = .005$.

Tabla 3.8 Relación entre α_C y α_E para $n = 1, 2, 3, 4, 5$ o 10 pruebas independientes

n	α_E cuando $\alpha_C = .05$	α_C cuando $\alpha_E = .05$
1	.050	.050
2	.098	.025
3	.143	.017
4	.185	.013
5	.226	.010
10	.401	.005

¿Qué tasa de error con respecto al experimento?

Existe cierta confusión en cuanto a las tasas de error para inferencias simultáneas, ya que la tasa de error se puede definir según la configuración de las medias de la población, $\mu_1, \mu_2, \dots, \mu_r$. La tasa de error con respecto al experimento se ha definido frecuentemente bajo la configuración $\mu_1 = \mu_2 = \dots = \mu_r$, y bajo esta configuración, la tasa de error respecto a la familia de comparaciones se considera poco controlada (Hochberg y Tamhane, 1987). Es poco probable que la igualdad de las medias sea cierta en la mayoría de las circunstancias. Entonces, cuando ocurren desigualdades, la tasa de error con respecto al experimento en el *sentido débil* ofrece una pobre protección contra decisiones incorrectas. Si no hay restricciones en las relaciones entre las μ_i , entonces la tasa de error con respecto a la familia se controla en el *sentido fuerte* (Hochberg y Tamhane, 1987). Esta definición del *sentido fuerte* se puede interpretar como la probabilidad de tomar al menos una decisión incorrecta sobre todas las configuraciones de los parámetros. Hsu (1996) también presenta un análisis detallado de las tasas de error para la inferencia simultánea.

3.5 Inferencia estadística simultánea

En la sección anterior se consideraron inferencias simultáneas múltiples y las probabilidades de tomar decisiones incorrectas con respecto a ellas. Cuando se probaron las hipótesis de tres contrastes en el experimento de almacenamiento de carne descrito en la sección 3.1, cada uno se probó con una tasa de error tipo I con respecto a la comparación de $a_C = .05$. Si esas afirmaciones han de cumplirse al mismo tiempo para una tasa de error tipo I de $a_E = .05$, las tres pruebas de hipótesis requieren una tasa de error basadas en la familia. Según la tabla 3.8, debe reducirse la tasa de error con respecto a la comparación para que las tres afirmaciones se cumplan de manera simultánea con una tasa de error para la familia de $a_E = .05$. Si los contrastes fueran independientes, entonces $a_C = .017$ sería una tasa de error de la familia adecuada, bajo la hipótesis nula de igualdad entre las medias de los contrastes.

Si un investigador desea tomar decisiones basadas en la información de datos observados con un método estadístico válido, la elección del método depende del *tipo* y de la *fuerza* de la inferencia deseados.

Los **tipos** de contrastes que consideran los investigadores con mayor frecuencia son:

- contrastes planeados
- contrastes polinomiales ortogonales
- comparaciones múltiples con el mejor tratamiento
- comparaciones múltiples con el tratamiento de control
- todas las comparaciones por pares

Los dos primeros métodos se estudiaron en las secciones 3.2 y 3.3. Los otros tres se presentarán más adelante en este capítulo.

La **fuerza** de la inferencia se refiere a cuánto puede decirse sobre una comparación. Las inferencias más fuertes incluyen afirmaciones sobre la magnitud y

dirección de la diferencia. Los métodos de intervalos de confianza simultáneos proporcionan la forma de inferencia más fuerte. El investigador puede asegurar, con un nivel de confianza dado, en qué dirección difieren los efectos de un tratamiento de los de otro y, además, qué tan lejanos están esos efectos entre sí.

Los métodos que proporcionan inferencias de *direcciones confiables* siguen, en fuerza, a los intervalos de confianza simultáneos (Hsu, 1996). Para una comparación dada C_i , estos métodos establecen desigualdad ($C_i > 0$ o $C_i < 0$) para cada comparación con niveles de confianza dados, pero no proporcionan idea de su magnitud.

Las *desigualdades confiables* enuncian la desigualdad de cada comparación, como $C_i \neq 0$, con un nivel de confianza dado, cuyas afirmaciones son correctas (Hsu, 1996). No hay indicación en estas afirmaciones sobre la magnitud o dirección de la desigualdad.

Los métodos más débiles son los métodos de *comparación individual* que no consideran la naturaleza simultánea de la inferencia. La debilidad de estos métodos se demostró en la tabla 3.8 de la sección anterior, donde la probabilidad de error para las afirmaciones simultáneas, α_E , podía aumentar con el número de pruebas.

Métodos de comparación múltiple, tablas de probabilidad y programas de cómputo

Existen métodos bien definidos, con tablas de probabilidad explícitas, para algunos procedimientos de comparación múltiple, que se presentarán más adelante en este capítulo, dado que el estudio se puede modelar como un diseño totalmente aleatorizado con igual número de réplicas para los tratamientos.

Si se usan estos métodos con números de réplicas desiguales o con estructuras de diseño más complejas, deben emplearse ya sea aproximaciones a las tablas de probabilidad o programas de computadora.

Por fortuna, varios programas estadísticos incluyen cálculos exactos para estos métodos; sin embargo, debe verificarse la documentación de los programas para estar seguros de que usan métodos válidos.

Las aproximaciones a las tablas de probabilidad aprovechan varias desigualdades de la teoría de probabilidad. Si no se dispone de rutinas de computadora especiales, estas aproximaciones son una alternativa conservadora. La aproximación que se sugerirá más adelante se basa en la **desigualdad de Bonferroni**, puesto que ya se cuenta con tablas de probabilidad basadas en ella. Esta aproximación también se puede usar para métodos menos estructurados, como pequeños conjuntos de comparaciones y contrastes polinomiales ortogonales, para obtener una protección de la medida del error en la inferencia simultánea.

Estadístico *t* de Bonferroni para la inferencia simultánea

La *desigualdad de Bonferroni* proporciona un medio para obtener una aproximación sencilla a las tasas de error en las comparaciones múltiples. La desigualdad, traducida a nuestro contexto, muestra que la tasa de error para la familia de comparaciones es menor o igual que la suma de las tasas de error de las comparaciones individuales. Cuando se hacen n comparaciones con la misma tasa de error, α_C , la desigualdad de Bonferroni produce la relación:

$$\alpha_E \leq n\alpha_C$$

La igualdad de la relación se cumple cuando las pruebas son independientes. La tasa de error para cada comparación para el estadístico de prueba se determina dividiendo la tasa máxima deseada de error para la familia entre el número de pruebas simultáneas, $\alpha_C = \alpha_E/n$. Por ejemplo, con tres comparaciones en el experimento de carnes, una tasa de error $\alpha_E = .05$ necesita una tasa de error con respecto a la comparación de $\alpha_C = .05/3 = .017$.

Los valores tabulados para el estadístico t de Student, en relación con la t de Bonferroni, se dan en la tabla V del apéndice como $t_{\alpha_E/2,k,v}$ para las pruebas de dos colas, donde k es el número de comparaciones y v son los grados de libertad. Los valores también se pueden obtener con un programa de computadora que calcule probabilidades para la distribución t de Student, con una probabilidad de la cola superior de $\alpha_E/2k$. Por ejemplo, la t de Bonferroni para comparaciones $k = 3$ de dos colas con $v = 8$ grados de libertad y $\alpha_E = .05$ es $t_{.025,3,8} = 3.02$. De manera equivalente, es posible encontrar el valor si se determina el valor de la t de Student para $v = 8$ grados de libertad, excedido con una probabilidad $\alpha_E/2k = .05/2(3) = .00833$.

Estimación de los intervalos de confianza simultáneos (ICS)

Las estimaciones de los intervalos de confianza para las comparaciones son la estimación de los intervalos que cumplen al mismo tiempo con el nivel de confianza $100(1 - \alpha_E)\%$. Los intervalos de confianza simultáneos usan el estadístico t de Bonferroni en lugar del t de Student. El intervalo de confianza $100(1 - \alpha_E)\%$ de dos colas es:

$$c \pm t_{\alpha_E/2,k,v}(s_c) \quad (3.26)$$

Las tres comparaciones para el experimento de almacenamiento de carne requieren $t_{.025,3,8} = 3.02$ para un conjunto de intervalos que cumplen de manera simultánea con una confianza del 95% por lo menos. Los intervalos que resultan se muestran en el cuadro 3.6.

Cuadro 3.6 Intervalos de confianza simultáneos del 95% para tres comparaciones en el experimento de almacenamiento de carne

Contraste	Estimación	Error estándar	ICS 95% (L, U)
$\mu_1 - \frac{1}{3}(\mu_2 + \mu_3 + \mu_4)$	2.11	0.228	(1.42, 2.80)
$\mu_2 - \frac{1}{2}(\mu_3 + \mu_4)$	0.19	0.241	(-0.54, 0.92)
$\mu_3 - \mu_4$	3.90	0.278	(3.06, 4.74)

Los intervalos ICS son más amplios que los calculados como intervalos de confianza del 95% individuales en la tabla 3.1. Existe un trueque entre la fuerza

de la afirmación de confianza y la amplitud del intervalo. El nivel de confianza conjunto es menor para los intervalos, más estrechos, de la tabla 3.1; y mayor para los intervalos ICS del cuadro 3.6, más amplios.

El investigador puede hacer afirmaciones conjuntas, con una tasa de error tipo I con respecto al experimento de .05, como 1) el desarrollo bacterial promedio en atmósferas artificiales es menor que en la carne con empaque comercial, ya que el límite inferior del intervalo es mayor que 0; 2) no hay diferencia entre el empaque al vacío y el promedio de los gases en cuanto al desarrollo de bacterias, pues el intervalo incluye al 0; y 3) el resultado con CO_2 es un menor número de bacterias que con la mezcla de gases, porque el límite inferior del intervalo es mayor que 0.

Las tres comparaciones del experimento de carnes en la tabla 3.1, sus estimaciones, errores estándar y el estadístico t_0 se muestran en el cuadro 3.7. Los estadísticos t_0 calculados para primera y tercera comparaciones excede al valor crítico $t_{.025,3,8} = 3.02$. El estadístico t en el cuadro 3.7 ilustra el uso de la inferencia de *desigualdades confiables* si sólo se evalúa qué tan significativas son su significancia. Las magnitudes y signos de las comparaciones y el estadístico t se pueden usar para deducir la dirección y magnitud de las comparaciones, aunque de forma menos directa que con los intervalos de confianza.

Cuadro 3.7 Prueba t de Bonferroni para las 3 comparaciones en el experimento de almacenamiento de carne

Comparación	Estimación	Error estándar	t_0
C_1 (Comercial vs. artificial)	2.11	0.228	9.25
C_2 (Al vacío vs. gases)	0.19	0.241	0.79
C_3 (CO_2 vs. mezcla de gases)	3.90	0.278	14.03

Prueba de Scheffé para la inferencia simultánea

El estadístico t de Bonferroni se puede usar de manera segura con un número pequeño de comparaciones preplaneadas con la preservación del error con respecto al experimento propuesto. Scheffé (1953) propuso un método para probar todas las comparaciones posibles o construir intervalos de confianza para todas ellas. Este método proporciona la protección señalada del error con respecto al experimento, para cualquier número de comparaciones; en consecuencia, es bastante conservador y, en general, se usa para comparaciones no planeadas o sugeridas por los datos. La prueba de Scheffé se muestra en el cuadro 3.8.

Los intervalos de confianza simultáneos de $100(1 - \alpha_E)\%$ para todas las comparaciones posibles se calcula con el estadístico de Scheffé como:

$$c \pm S(\alpha_E) \quad (3.27)$$

y se tiene una probabilidad $(1 - \alpha_E)$ de que todos los intervalos simultáneos incluyan los valores verdaderos de las respectivas comparaciones.

Cuadro 3.8 Prueba de Scheffé

Considerando cualquier comparación, $c = \sum_{i=1}^t k_i \bar{y}_i$, entre t medias de tratamientos con un error estándar:

$$s_c = \sqrt{s^2 \left[\sum_{i=1}^t \frac{k_i^2}{r_i} \right]}$$

La hipótesis nula para la comparación, $H_0: C = 0$, se rechaza si:

$$|c| > S(\alpha_E) \quad (3.28)$$

$S(\alpha_E)$ es el estadístico de Schaffé:

$$S(\alpha_E) = s_c \sqrt{(t - 1) F_{\alpha_E, (t-1), v}} \quad (3.29)$$

donde $F_{\alpha_E, (t-1), v}$ es el estadístico F con $(t - 1)$ y v grados de libertad excedida con probabilidad α_E . Además, v es el número de grados de libertad para la varianza del error experimental s^2 , que se usa para estimar el error estándar de la comparación, s_c .

3.6 Comparaciones múltiples con el mejor tratamiento

En algunos estudios, los tratamientos no tienen una alta estructura en su interrelación y es difícil identificar las comparaciones estructuradas. Por otro lado, los tratamientos se relacionan entre sí porque todos forman parte de la investigación por su efecto sobre las variables de respuesta y se refieren a algún problema específico de interés para el investigador.

En estas circunstancias, tal vez el investigador quiera “elegir a los ganadores”. El objetivo es seleccionar el conjunto de tratamientos o un solo tratamiento (si es posible) que proporcione el resultado más deseable.

El procedimiento de **comparaciones múltiples con el mejor** (CMM) de Hsu (1984) permite al investigador clasificar los tratamientos de manera que la “mejor” población esté incluida en un subconjunto con un nivel de confianza dado. Los parámetros de interés son:

$$\mu_i - \max_{j \neq i} \mu_j, \text{ para } i = 1, 2, \dots, t$$

donde $\max_{j \neq i} \mu_j$ es la media máxima de los tratamientos sin incluir μ_i . Si $\mu_i - \max_{j \neq i} \mu_j > 0$, entonces el tratamiento i es el mejor. Por otro lado, si $\mu_i - \max_{j \neq i} \mu_j < 0$, entonces el tratamiento i no es el mejor.

Los intervalos de confianza simultáneos (ICS) de las CMM para $\mu_i - \max_{j \neq i} \mu_j$ tienen la restricción de incluir el 0 con la perspectiva de que a la larga dos tratamientos nunca tienen promedios idénticos (Hsu, 1996). El intervalo de confianza CMM *restringido* establece que el tratamiento i es uno de los mejores si el intervalo para $\mu_i - \max_{j \neq i} \mu_j$ incluye el 0 o el 0 es su límite *inferior*. Por el contrario, si el límite *superior* del intervalo para $\mu_i - \max_{j \neq i} \mu_j$ es 0, entonces el tratamiento i no es el mejor. El procedimiento CMM se describe en el cuadro 3.9.

Cuadro 3.9 Procedimiento de comparaciones múltiples con el mejor

Se calcula la diferencia, D_i , entre cada media de tratamiento, \bar{y}_i , y la mayor media de tratamiento de los restantes, $\max_{j \neq i} (\bar{y}_j)$, como sigue:

$$D_i = \bar{y}_i - \max_{j \neq i} (\bar{y}_j), \text{ para } i = 1, 2, \dots, t \quad (3.30)$$

y la cantidad M

$$M = d_{\alpha, k, v} \sqrt{\frac{2s^2}{r}} \quad (3.31)$$

donde $d_{\alpha, k, v}$ es el estadístico tabulado para las comparaciones de un lado en la tabla VI del apéndice, para una tasa de error experimental de α_e , $k = t - 1$ comparaciones y v grados de libertad para la varianza experimental, $s^2 = CME$.

Intervalos de confianza restringidos simultáneos de $100(1 - \alpha)\%$

El límite inferior de confianza para $\mu_i - \max_{j \neq i} (\bar{y}_j)$ es:

$$L = \begin{cases} D_i - M & \text{si } (D_i - M) < 0 \\ 0 & \text{de otra manera} \end{cases}$$

y el límite superior de confianza para $\mu_i - \max_{j \neq i} (\bar{y}_j)$ es:

$$U = \begin{cases} D_i + M & \text{si } (D_i + M) > 0 \\ 0 & \text{de otra manera} \end{cases}$$

Ejemplo 3.2 Tasas de flujo a través de filtros

El procedimiento CMM se ilustra con un experimento realizado para evaluar filtros con diferentes configuraciones de filtrado. Todos se diseñaron para detener partículas mayores a cierto tamaño. Los investigadores querían averiguar qué filtros, si había alguno, permitían la tasa de flujo más alta de lodos a presión constante.

Se tenían 6 configuraciones para las pruebas. Se construyeron 4 réplicas de cada configuración para el experimento; se probaron los 24 filtros en orden aleatorio con un diseño de experimento totalmente aleatorizado. Las tasas de flujo promedio para los 6 filtros fueron

Filtro:	A	B	C	D	E	F
Media:	8.29	7.23	7.54	8.10	8.59	7.10

La estimación de la varianza del error experimental fue $CME = 0.08$, con 18 grados de libertad.

Intervalos de confianza simultáneos CMM

Para ilustrar el procedimiento, se puede calcular un intervalo de confianza del 95% para una comparación del filtro B con el mejor del resto de los filtros. La media del filtro B es $\bar{y}_2 = 7.23$ y el filtro E es el que tiene la media más grande de los demás, de manera que $\max_{j \neq 2} (\bar{y}_j) = \bar{y}_5 = 8.59$. Entonces $D_2 = 7.23 - 8.59 = -1.36$.

El valor de $d_{\alpha,k,v}$ en la ecuación (3.31) se encuentra a partir de la tabla VI del apéndice con $k = 5$, $\alpha_E = .05$ y $v = 18$ grados de libertad para una $CME = 0.08$. El valor adecuado es $d_{.05,5,18} = 2.41$, para que con $r = 4$ réplicas y $CME = 0.08$ se tenga:

$$M = 2.41 \sqrt{\frac{2(0.08)}{4}} = 0.48$$

Las cantidades requeridas son $D_2 - M = -1.36 - 0.48 = -1.84$ y $D_2 + M = -1.36 + .48 = -0.88$. Usando las reglas para los límites superior e inferior del cuadro 3.9, $L = -1.84$ porque $D_2 - M < 0$, y $U = 0$ porque $D_2 + M$ no es mayor que 0. En la tabla 3.9 se muestran los límites superior e inferior para cada comparación.

Cuatro de los filtros (B, C, D y F) tienen límites superiores a 0 y, por lo tanto, no son el “mejor” filtro. De los dos filtros restantes (A y E) no es claro cuál es el mejor, ya que ambos intervalos incluyen el 0, lo que a su vez implica que cada uno es el mejor filtro, con un 95% de confianza. Además, sus límites inferiores, -0.78 para A y -0.18 para E, también son cercanos a 0, lo que indica que el tratamiento es cercano al mejor (Hsu, 1996). Observe que el ICS no sólo proporciona los medios para identificar el o los mejores tratamientos, también indica qué tan lejos están del mejor. Según el límite inferior de los intervalos en la tabla 3.9 es sencillo ver que los filtros B, C y F son los más lejanos al mejor tratamiento.

Si es claro que un tratamiento es el mejor, entonces el límite inferior será 0. Para ejemplificar lo anterior, supongamos que el filtro E tiene una media $\bar{y}_5 = 9.00$ en lugar de 8.59, la media más grande que le sigue es el filtro A con $\bar{y}_1 = 8.29$. El intervalo de confianza restringido para comparar el filtro E con el A sería:

Tabla 3.9 Procedimiento CMM con las medias de la tasa de flujo de seis filtros diferentes

Filtro	\bar{y}_i	$\max_{j \neq i} (\bar{y}_j)$	D_i	$D_i - M$	$D_i + M$	$\frac{ICS del 95\%}{(L, U)}$	¿Elegir?*
A	8.29	8.59	-0.30	-0.78	0.18	(-0.78, 0.18)	Sí
B	7.23	8.59	-1.36	-1.84	-0.88	(-1.84, 0)	No
C	7.54	8.59	-1.05	-1.53	-0.57	(-1.53, 0)	No
D	8.10	8.59	-0.49	-0.97	-0.01	(-0.97, 0)	No
E	8.59	8.29	0.30	-0.18	0.78	(-0.18, 0.78)	Sí
F	7.10	8.59	-1.49	-1.97	-1.01	(-1.97, 0)	No

*Elegir como el “mejor” cuando $D_i + M > 0$.

$$D_5 - M = 9.00 - 8.29 - 0.48 = 0.23$$

$$D_5 + M = 9.00 - 8.29 + 0.48 = 1.19$$

de manera que con $D_5 - M > 0$ y $D_5 + M > 0$ los límites del intervalo restringido son (0, 1.19) y el intervalo con límite inferior de 0 indica que el filtro E es el mejor. De manera similar, el filtro A no sería uno de los mejores en estas circunstancia ya que

$$D_1 - M = 8.29 - 9.00 - 0.48 = -1.19$$

$$D_1 + M = 8.29 - 9.00 + 0.48 = -0.23$$

por lo que con $D_1 - M < 0$ y $D_1 + M < 0$ los límites del intervalo restringido son (-1.19, 0). Así, con un límite superior de 0 el filtro A no sería el mejor tratamiento ni estaría entre los mejores.

Elección del subconjunto con la media más grande

Si sólo interesa determinar el tratamiento o tratamientos que constituyen el “mejor” sin importar cuánto difieran sus efectos entre sí, se puede usar una sencilla regla de selección (Hsu, 1984). La regla CMM selecciona un tratamiento para el subconjunto de los mejores con probabilidad $P(EC) = 1 - \alpha$, de una elección correcta si:

$$D_i + M > 0 \quad (3.32)$$

El procedimiento CMM para elegir los mejores tipos de filtros con $P(EC) = 0.95$ utiliza la información de la columna $D_i + M$, considerando como mejores tratamientos aquellos para los que $D_i + M > 0$. Sólo los filtros A y E tienen valores de $D_i + M > 0$; 0.18 y 0.78, respectivamente. El subconjunto de los mejores incluye estos dos filtros, con una probabilidad $P(EC) = 0.95$ de que el mejor filtro esté en el subconjunto A y E.

Las conclusiones del procedimiento de selección del subconjunto no difieren de los resultados del ICS, pues ambos filtros se colocan en el subconjunto de los mejores. Sin embargo, los ICS contienen más información, pues indican qué tan

cercanas a 0 están los límites inferiores para cada tipo de filtro; lo que a su vez proporciona más información al investigador en cuanto al desempeño de la tasa de flujo de cada filtro en comparación con las demás. El procedimiento de selección del subconjunto limita la información a determinar si el filtro pertenece al subconjunto y nada más.

Comparaciones múltiples con la media más pequeña

En algunos estudios, como en el experimento de almacenamiento de carne, donde el empaque era mejor si tenía menos bacterias, el tratamiento con la media más pequeña es el “mejor”. Es posible realizar comparaciones múltiples con ella, si se hacen algunas sencillas modificaciones a la regla para elegir la media más grande del cuadro 3.9.

Se calcula la diferencia, D_i , entre cada media de tratamiento, \bar{y}_i , y la media más pequeña de los tratamientos restantes, $\min_{j \neq i} (\bar{y}_j)$, como:

$$D_i = \bar{y}_i - \min_{j \neq i} (\bar{y}_j) \text{ para } i = 1, 2, \dots, t \quad (3.33)$$

Los límites inferior (L) y superior (U) de los intervalos de confianza de 100 $(1 - \alpha)\%$ se encuentran como sigue: el límite inferior de confianza para $\mu_i - \min_{j \neq i} \mu_j$ es:

$$L = \begin{cases} D_i - M & \text{si } (D_i - M) < 0 \\ 0 & \text{de otra manera} \end{cases}$$

y el superior para $\mu_i - \min_{j \neq i} \mu_j$ es:

$$U = \begin{cases} D_i + M & \text{si } (D_i + M) > 0 \\ 0 & \text{de otra manera} \end{cases}$$

Los intervalos de confianza restringidos CMM para las comparaciones con la media más pequeña de los tratamientos, se interpreta justo al contrario que en el caso de comparaciones con la media más grande. Si el intervalo para $\mu_i - \min_{j \neq i} \mu_j$ incluye el 0 o tiene un límite *superior* a 0, entonces el tratamiento i es el mejor. Por el contrario, si el límite inferior es cero, i no es el mejor tratamiento.

Las comparaciones múltiples considerando al tratamiento con la media más pequeña como el mejor, se explican con el estudio de almacenamiento de carne y el objetivo de elegir el material de empaque con menor desarrollo de bacterias. La varianza del error experimental para el experimento era $CME = 0.116$ con 8 grados de libertad, y se tenían $r = 3$ réplicas. Las estadísticas necesarias para el valor crítico de M en la ecuación (3.31) son $d_{.05,3,8} = 2.42$ y $\sqrt{2CME/r} = 0.278$, de manera que:

$$M = d_{.05,3,8} \sqrt{\frac{2CME}{r}} = 2.42(0.278) = 0.67$$

Tabla 3.10 Selección del subconjunto de tratamientos con las medias más pequeñas en el experimento de almacenamiento de carne

Tratamiento	\bar{y}_i	$\min_{j \neq i} (\bar{y}_j)$	D_i	$D_i - M$	$D_i + M$	<i>ICS del 95%</i>		¿Elegir?*
						(L, U)		
Comercial	7.48	3.36	4.12	3.45	4.79	(0, 4.79)		No
Al vacío	5.50	3.36	2.14	1.47	2.81	(0, 2.81)		No
Mezcla	7.26	3.36	3.90	3.23	4.57	(0, 4.57)		No
CO_2 puro	3.36	5.50	-2.14	-2.81	-1.47	(-2.81, 0)		Sí

*Elegir como el “mejor” cuando $D_i - M < 0$.

En la tabla 3.10 se muestran las medias de los tratamientos y las estimaciones de los intervalos de confianza simultáneos del 95% para las comparaciones con la media más pequeña. El procedimiento se ilustra para el tratamiento de empaque comercial con una media $\bar{y}_1 = 7.48$ y el tratamiento con la media mínima, CO_2 puro, de manera que $\min_{j \neq i} (\bar{y}_j) = \bar{y}_4 = 3.36$ para obtener $D_1 = 7.48 - 3.36 = 4.12$.

El límite inferior es $L = 0$, ya que $D_1 - M = 4.12 - 0.67 = 3.45$ no es menor que 0; y el límite superior es $U = 4.79$ puesto que $D_1 + M = 4.12 + 0.67 = 4.79$ es mayor que 0. Con un límite inferior de 0, se puede asegurar que el empaque comercial no es el mejor tratamiento. El mejor tratamiento es CO_2 puro, porque es el único en la tabla 3.10 con un límite superior del intervalo de confianza de 0. Los límites superiores del resto de los tratamientos están muy alejados del 0 y es claro que no es posible considerarlos cercanos al mejor.

Elección del subconjunto con la media más pequeña

De nuevo, si el único interés es conocer los tratamientos que constituyen el mejor, sin importar cuánto difieran sus efectos de los demás, entonces se puede usar una sencilla regla de selección (Hsu, 1984). La regla CMM elige un tratamiento para el subconjunto de los mejores con una probabilidad de elección correcta, $P(\text{EC}) = 1 - \alpha$, si:

$$D_i - M < 0 \quad (3.34)$$

El procedimiento CMM para elegir el mejor empaque de carne con $P(\text{EC}) = 0.95$, usa la información de la columna $D_i - M$ dada en la tabla 3.10. Los tratamientos en el subconjunto de los mejores son aquellos para los que $D_i - M < 0$, y la única media de tratamiento que cumple con esta condición es la de CO_2 puro, que se elige como el mejor tratamiento, con el menor crecimiento de bacterias, con una probabilidad de una elección correcta $P(\text{EC}) = .95$.

Réplicas desiguales y modelos complejos

Hsu (1996) indica que los valores críticos se deben calcular por separado para cada comparación con un número de réplicas desigual, en un diseño totalmente aleatorizado. Algunos programas de computadora han incorporado rutinas para calcular estos intervalos de confianza restringidos simultáneos con números de

réplicas desiguales. Hsu (1996) no proporcionó aproximaciones basadas en desigualdades de probabilidad.

Si se usan bloques o diseño de tratamiento más complejo y todas las diferencias, $\mu_i - \mu_j$, tienen la misma varianza, es decir, el diseño tiene una *varianza balanceada*, entonces se puede utilizar el procedimiento CMM normal de esta sección.

Si las varianzas de las diferencias, $\mu_i - \mu_j$, no son las mismas en los diseños más complejos, Hsu (1996) recomienda varias aproximaciones basadas en desigualdades de probabilidad. Una aproximación basada en la *desigualdad de Bonferroni* usa la *t* de Bonferroni para k comparaciones, con v grados de libertad, en el nivel apropiado de α en lugar de $d_{\alpha,k,v}$. Deben utilizarse la estimación de mínimos cuadrados, $\hat{\mu}_i - \hat{\mu}_j$, y su error estándar; éstos pueden obtenerse con la mayor parte de los programas de computadora.

3.7 Comparación de todos los tratamientos con un control

En muchos estudios, uno de los tratamientos actúa como control para algunos o todos los restantes. (En la sección 1.4 se analizan los distintos tipos de tratamientos de control.) En ocasiones es interesante determinar si las respuestas medias de los tratamientos difieren de la del tratamiento de control. Dunnett (1955) introdujo un procedimiento basado en una tasa de error con respecto al experimento con este propósito.

Intervalos de confianza simultáneos

El estadístico tabulado para calcular intervalos de confianza simultáneos de 100 $(1 - \alpha)\%$ para las diferencias entre las medias este tratamiento individuales y la media de control $\mu_i - \mu_c$, utilizando el procedimiento de Dunnett, se basa en el mismo estadístico usado para las comparaciones múltiples con el mejor procedimiento de la sección 3.6. La prueba de Dunnett para comparar cada media de tratamiento, \bar{y}_i , con la media del tratamiento de control \bar{y}_c se describe en el cuadro 3.10.

Supongamos que el filtro F del ejemplo 3.2 sirve de control. Las cinco comparaciones ICS del 95% del control contra las medias de los tratamientos se muestran en la tabla 3.11. Las diferencias de las medias entre F y los otros $k = 5$ filtros aparecen en la tercera columna. El valor crítico del estadístico de Dunnett para una prueba de dos lados, con una tasa de error de $\alpha_E = .05$ es $d_{05,5,18} = 2.76$. El error estándar de la diferencia es $\sqrt{2CME/r} = \sqrt{2(0.08)/4} = 0.2$. El criterio de Dunnett es:

$$D(5, .05) = 2.76(0.2) = 0.55$$

Se ilustra un ejemplo de los cálculos con el filtro A. El límite inferior es:

$$L = \bar{y}_1 - \bar{y}_c - D(5, .05) = 8.29 - 7.10 - 0.55 = 0.64$$

y el límite superior es:

$$U = \bar{y}_1 - \bar{y}_c + D(5, .05) = 8.29 - 7.10 + 0.55 = 1.74$$

El filtro A es superior al de control, F, ya que el límite inferior del intervalo es mayor que 0. Según las estimaciones de los intervalos, los filtros A, D y E son

superiores al F. Observe que el filtro E es el más lejano al filtro F, pues su límite inferior es mayor que el de A o de D. Los filtros B y C no difieren del control porque sus intervalos incluyen el 0.

Cuadro 3.10 Método de Dunnett para comparar todos los tratamientos con un control

100(1 – α)% intervalos de confianza simultáneos para $\mu_i - \mu_c$

El criterio de Dunnett para comparar k tratamientos con el control es:

$$D(k, \alpha_E) = d_{\alpha, k, v} \sqrt{\frac{2s^2}{r}} \quad (3.35)$$

Las estimaciones de los intervalos de confianza simultáneos bilaterales para las diferencias entre las medias de los tratamientos individuales y la media del control $\mu_i - \mu_c$ son:

$$\bar{y}_i - \bar{y}_c \pm D(k, \alpha_E) \quad (3.36)$$

Los límites superiores del intervalo de un lado, si se manifiesta superioridad de la media de tratamiento por ser *mayor* que la media del control, son:

$$\bar{y}_i - \bar{y}_c - D(k, \alpha_E) \quad (3.37)$$

Los límites superiores del intervalo de un lado, si se manifiesta superioridad de la media de un tratamiento por ser *menor* que la media de control, son:

$$\bar{y}_i - \bar{y}_c + D(k, \alpha_E) \quad (3.38)$$

Los valores de $d_{\alpha, k, v}$ para el método de Dunnett de uno o dos lados, se encuentran en la tabla VI del apéndice para k tratamientos, un error tipo I de α_E con respecto al experimento y v grados de libertad para la estimación de la varianza del error experimental.

Tabla 3.11 Resultado de la prueba de Dunnett al comparar la media del filtro de control, F, con las de los demás

Filtro	Media	$\bar{y}_i - \bar{y}_c$	<i>ICS de 95%</i>		<i>¿Diferente del control?*</i>
			(L, U)	$ \bar{y}_i - \bar{y}_c $	
F	$\bar{y}_c = 7.10$	—	—	—	—
A	8.29	1.19	(0.64, 1.74)	1.19	Sí
B	7.23	0.13	(-0.42, 0.68)	0.13	No
C	7.54	0.44	(-0.11, 0.99)	0.44	No
D	8.10	1.00	(0.45, 1.55)	1.00	Sí
E	8.59	1.49	(0.94, 2.04)	1.49	Sí

*Si $|\bar{y}_i - \bar{y}_c|$ excede $D(5, .05) = 0.55$, entonces la media del filtro correspondiente es diferente de la del filtro F (control).

Prueba de hipótesis sobre $\mu_i - \mu_c$

Si el investigador sólo desea saber si la media del tratamiento es significativamente diferente de la media de control con un nivel de significancia α , entonces se pueden realizar las siguientes pruebas. Para la alternativa de dos lados con $H_0: \mu_i = \mu_c$ contra $H_a: \mu_i \neq \mu_c$, se rechaza la hipótesis nula si:

$$|\bar{y}_i - \bar{y}_c| > D(k, \alpha_E) \quad (3.39)$$

Para la alternativa de un lado con $H_0: \mu_i \leq \mu_c$ contra $H_a: \mu_i > \mu_c$, se rechaza la hipótesis nula si:

$$(\bar{y}_i - \bar{y}_c) > D(k, \alpha_E) \quad (3.40)$$

Para la alternativa de un lado con $H_0: \mu_i \geq \mu_c$ contra $H_a: \mu_i < \mu_c$, se rechaza la hipótesis nula si:

$$(\bar{y}_i - \bar{y}_c) < -D(k, \alpha_E) \quad (3.41)$$

En la tabla 3.11 se muestran los resultados de la prueba con la alternativa de dos lados. Las diferencias para los filtros A, D y E exceden $D(5, .05) = 0.55$, y sus tasas de flujo difieren significativamente de las del filtro F. Las pruebas de dos lados ilustran una inferencia de desigualdades confiable. Sólo se asegura, con una tasa de error de .05, que los filtros A, D y E difieren del filtro F, y que los filtros B y C no. Solamente se puede deducir de manera indirecta, a partir del tamaño y signo, cuáles diferencias son positivas o negativas y cuáles son las más grandes comparadas con el ICS, que proporciona información directa del tamaño y la dirección.

Réplicas desiguales y modelos complejos

Hsu (1996) indica que en un diseño totalmente aleatorizado los valores críticos se deben calcular por separado para cada comparación con un número de réplicas desigual. Algunos programas de computadora han incorporado rutinas para calcular los intervalos de confianza restringidos simultáneos con un número de réplicas desigual. Se puede usar una aproximación, basada en la desigualdad de Bonferroni, que sustituye la t de Bonferroni para k comparaciones con v grados de libertad en el nivel adecuado de α , en lugar de $d_{\alpha,k,v}$ y usa $\sqrt{2s^2[1/r_i + 1/r_c]}$ para el error estándar de la diferencia.

Si se emplea un agrupamiento o diseño de tratamiento más complejo y todas las diferencias, $\mu_i - \mu_j$, tienen la misma varianza, es decir, si el diseño tiene una varianza balanceada, entonces se puede utilizar el procedimiento normal descrito en esta sección.

En diseños más complejos si las varianzas de las diferencias, $\mu_i - \mu_c$, no son las mismas, entonces Hsu (1996) presenta varias aproximaciones, que se basan en desigualdades de probabilidad. La aproximación basada en la desigualdad de Bonferroni usa la t para k comparaciones con v grados de libertad al nivel adecuado de α , en lugar de $d_{\alpha,k,v}$. Debe usarse la estimación de mínimos cuadrados, $\hat{\mu}_i - \hat{\mu}_j$, y su error estándar, que se pueden obtener con muchos programas de computadora.

Dunnett (1964) proporciona ajustes a los valores críticos en la tabla VI del apéndice para los números de réplicas diferentes. Un conservador límite superior dado por Feiss (1986) para los valores en la tabla VI del apéndice es $md_{\alpha,k,v}$, donde:

$$m \leq 1 + 0.07 \left(1 - \frac{r_i}{r_c} \right) \quad (3.42)$$

Dunnett demostró que la razón óptima de los números de réplicas es r/r_c , donde r es el número común de réplicas para cada tratamiento y r_c es el número de réplicas para el tratamiento de control.

3.8 Comparaciones en pares de todos los tratamientos

Algunos investigadores comparan cada media de tratamiento con cada una de las otras medias usando las **comparaciones en pares**. Los parámetros de interés son todas las diferencias por pares entre las medias de tratamiento, $\mu_i - \mu_j$, para toda $i \neq j$, que resultan en $t(t - 1)/2$ comparaciones. A menudo, estos métodos se aplican con el objeto de detectar desigualdades significativas, $\mu_i \neq \mu_j$, para toda $i \neq j$.

El uso indiscriminado de procedimientos de comparación por pares para el análisis de los resultados experimentales, puede conducir a la tendencia de dar importancia sólo a la significancia estadística para el procedimiento de inferencia en el análisis de datos. Entonces, es posible perder de vista los objetivos de la investigación y el enfoque del investigador puede desviarse de la búsqueda de la comprensión biológica o física a la de la significancia estadística.

En su forma ideal, los investigadores querrán hacer comparaciones con una tasa de error con respecto al experimento en el sentido fuerte, como se hizo en las comparaciones múltiples con el mejor (sección 3.6) y en las comparaciones múltiples con el control (sección 3.7). A continuación se presentan varios métodos de comparación por pares y se analizan algunas de sus propiedades.

Método de Tukey

Tukey (1949a) desarrolló un procedimiento que proporciona una tasa con respecto al experimento en el sentido fuerte, para las comparaciones en pares de todas las medias de tratamiento, que se usa para obtener intervalos de confianza simultáneos de $100(1 - \alpha)\%$. La prueba se conoce por varios nombres, entre ellos “diferencia honestamente significativa”. El método de Tukey se describe en el cuadro 3.11.

Este método se basa en el estadístico estandarizado:

$$q = \frac{\bar{y}(\text{mayor}) - \bar{y}(\text{menor})}{\sqrt{\frac{s^2}{r}}} \quad (3.43)$$

donde $\bar{y}(\text{mayor})$ es la media más grande de un grupo ordenado de medias en un experimento y $\bar{y}(\text{menor})$ es la más pequeña. La diferencia o separación se divide entre el error estándar de la media del tratamiento, de donde se deriva el nombre de estadístico estandarizado (de Student).

Cuadro 3.11 Método de Tukey para todas las comparaciones por pares

Para un grupo de k medias de tratamiento, se calcula la diferencia honestamente significativa como:

$$\text{DHS}(k; \alpha_E) = q_{\alpha, k, v} \sqrt{\frac{s^2}{r}} \quad (3.44)$$

donde $q_{\alpha, k, v}$ es el estadístico estandarizado de Student para un grupo de k medias de tratamiento en un arreglo ordenado. Los valores críticos de la tasa de error con respecto al experimento, α_E , y los v grados de libertad, se pueden encontrar en la tabla VII del apéndice.

Intervalos de confianza simultáneos de $100(1 - \alpha)\%$

Las estimaciones de los intervalos simultáneos de dos lados para el valor absoluto de todas las diferencias por pares, $\mu_i - \mu_j$ para toda $i < j$ son:

$$|\bar{y}_i - \bar{y}_j| \pm \text{DHS}(k, \alpha_E) \quad (3.45)$$

Prueba de desigualdades $100(1 - \alpha)\%$ confiables

Se establece que dos medias de tratamientos no son iguales, $\mu_i - \mu_j \neq 0$, si:

$$|\bar{y}_i - \bar{y}_j| > \text{DHS}(k, \alpha_E) \quad (3.46)$$

En el cuadro 3.11, se usa la diferencia absoluta $|\bar{y}_i - \bar{y}_j|$ para los intervalos de confianza porque la localización de las dos medias en la diferencia calculada, $\bar{y}_i - \bar{y}_j$, es arbitraria, y su signo depende de si se calcula $\bar{y}_i - \bar{y}_j$ o $\bar{y}_j - \bar{y}_i$. Así, la diferencia absoluta es equivalente a restar siempre la media más pequeña de la más grande. Si se necesita la dirección de una diferencia, entonces se calcula el intervalo considerando el signo de la diferencia; en algunas comparaciones las particularidades del estudio indicarán si es mejor trabajar con diferencias absolutas.

Ejemplo 3.3 Resistencia de soldaduras

Las pruebas de comparación por pares se ilustran con un experimento realizado para comparar la resistencia de soldaduras practicadas con cuatro técnicas distintas. Se soldaron cinco pares de placas de metal con cada técnica, en un diseño totalmente aleatorizado. La fuerza promedio para las cinco soldaduras de cada técnica fueron:

Técnica:	A	B	C	D
Media:	69	83	75	71

La estimación de la varianza del error experimental fue $CME = 15$ con 16 grados de libertad.

Si se aplica el método de Tukey a estos datos, el estadístico estandarizado de Student con una tasa de error con respecto al experimento de $\alpha_E = .05$, se encuentra en la tabla VII del apéndice como $q_{.05,4,16} = 4.05$, donde existen $k = 4$ medias de tratamiento en el arreglo ordenado y $v = 16$ grados de libertad para $CME = s^2$. El error estándar es $\sqrt{CME/r} = \sqrt{15/5} = 1.73$. La DHS es $DHS(4, .05) = 4.05(1.73) = 7.0$. El ICS del 95% y los resultados de las pruebas de desigualdades confiables se muestran en la tabla 3.12.

Tabla 3.12 Resultados del método de Tukey para las diferencias entre las medias de resistencia para cuatro técnicas de soldadura

Comparación	$ \bar{y}_i - \bar{y}_j $	ICS del 95%		¿Diferente de 0?*
		(L, U)		
A vs. B	14	(7, 21)		Sí
A vs. C	6	(-1, 13)		No
A vs. D	2	(-5, 9)		No
B vs. C	8	(1, 15)		Sí
B vs. D	12	(5, 19)		Sí
C vs. D	4	(-3, 11)		No

*La diferencia absoluta excede $DHS(.05) = 7.0$.

Dos medias de tratamiento son distintas con un 95% de confianza, si el ICS del 95% no incluye al 0. La técnica B es muy diferente de las demás, éstas no difieren entre sí. La cantidad en que la técnica B difiere de los otros tres métodos se puede evaluar mediante la magnitud del límite inferior. La técnica B difiere más de la técnica A y menos de la C.

La inferencia basada en las desigualdades de confianza establece dos medias de tratamiento distintas si la diferencia $|\bar{y}_i - \bar{y}_j|$, excede $DHS(4, .05) = 7.0$. La prueba DHS juzga que la fuerza de la soldadura de la técnica B es diferente a la de los otros métodos, pero ninguna otra diferencia es significativa. Sin embargo, no es posible hacer inferencias sobre la magnitud de la desigualdad.

Réplicas desiguales y modelos complejos con el método de Tukey

Para réplicas desiguales en el diseño totalmente aleatorizado, en 1953 Tukey (Tukey, 1994) y Kramer (1956) propusieron intervalos de confianza simultáneos aproximados mediante la sustitución de:

$$\sqrt{\frac{s^2}{2} \left(\frac{1}{r_i} + \frac{1}{r_j} \right)} \quad (3.47)$$

en vez de $\sqrt{s^2/r}$ dentro de la ecuación (3.43). Esta aproximación se conoce como la aproximación de Tukey-Kramer, misma que Hayter (1984) demostró que es conservadora.

Si las varianzas de las diferencias, $\mu_i - \mu_j$, no son iguales en los diseños más complejos, entonces se usan las estimaciones de mínimos cuadrados, $\hat{\mu}_i - \hat{\mu}_j$, y sus varianzas, $s^2_{(\hat{\mu}_i - \hat{\mu}_j)}$, para obtener una aproximación a la forma exacta. Los intervalos simultáneos son:

$$|\hat{\mu}_i - \hat{\mu}_j| \pm q_{\alpha, k, v} \sqrt{\frac{1}{2} s^2_{(\hat{\mu}_i - \hat{\mu}_k)}} \quad (3.48)$$

Si se emplean bloques o diseños de tratamiento más complejo y todas las diferencias, $\mu_i - \mu_j$, tienen la misma varianza, es decir, el diseño tiene *varianza balanceada*, entonces se puede usar el procedimiento normal descrito en esta sección, con las estimaciones de mínimos cuadrados, $\hat{\mu}_i - \hat{\mu}_j$, y su varianza común de las diferencias, $s^2_{(\hat{\mu}_i - \hat{\mu}_j)}$, en la ecuación (3.48). En este caso, la tasa de error con respecto al experimento en el sentido fuerte será exacta.

Pruebas de homogeneidad, $\mu_1 = \mu_2 = \dots = \mu_r$

Muchas de las pruebas de comparación múltiple por pares se han aplicado con tasas de error con respecto al experimento, evaluadas bajo la suposición restringida de que todas las medias de los tratamientos son iguales, o $\mu_1 = \mu_2 = \dots = \mu_r$, lo que proporciona tasas de error con respecto al experimento en el sentido *débil*. Algunas usan un solo criterio para establecer la significancia. Otras pruebas se conocen como *pruebas de intervalo múltiple*, ya que usan varios criterios para establecer la significancia, donde el valor de un criterio para una comparación depende de cuánto están separadas las medias en el arreglo ordenado de todas las medias de tratamiento. La *mínima diferencia significativa MDS(α)* (LSD por sus siglas en inglés), es la prueba más común de homogeneidad que usa un solo criterio. La prueba *Student-Newman-Keuls, SNK(k, α)*, es un ejemplo de prueba de intervalos múltiples para la homogeneidad. A continuación se ilustran ambas.

Mínima diferencia significativa (MDS)

Cada hipótesis $H_0: \mu_i = \mu_j$ contra $H_a: \mu_i \neq \mu_j$ se puede probar con el estadístico t de Student:

$$t_0 = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{s^2 \left[\frac{1}{r_i} + \frac{1}{r_j} \right]}} \quad \text{para toda } i \neq j \quad (3.49)$$

Cuando se establece la probabilidad de error tipo I en algún valor de α y la varianza s^2 tiene v grados de libertad, la hipótesis nula se rechaza para cualquier valor observado de $|\bar{y}_i - \bar{y}_j|$ tal que $|t_0| > t_{\alpha/2, v}$. La **mínima diferencia significativa, MDS(α)**, es un método abreviado para realizar todas la pruebas t por pares, como se muestra en el cuadro 3.12.

Cuadro 3.12 Mínima diferencia significativa para todas las comparaciones por pares

Para cualquier par de medias de tratamiento observadas, \bar{y}_i y \bar{y}_j , la mínima diferencia significativa es:

$$\text{MDS}(\alpha) = t_{\alpha/2,v} \sqrt{s^2 \left[\frac{1}{r_i} + \frac{1}{r_j} \right]} \quad (3.50)$$

La hipótesis nula $H_0: \mu_i = \mu_j$ se rechaza si:

$$|\bar{y}_i - \bar{y}_j| > \text{MDS}(\alpha)$$

Una modificación de Fisher (1960) controla la tasa de error con respecto al experimento en el sentido débil. La MDS se utiliza para probar las comparaciones por pares sólo si la hipótesis nula se rechaza durante la prueba F del análisis de varianza; de no ser así se supone entonces que todas las medias de tratamiento son iguales y no se realizan más pruebas, el procedimiento suele llamarse MDS protegida. Cramer y Swason (1973) demostraron empíricamente de que la tasa de error con respecto al experimento con la MDS protegida, tiene casi el mismo nivel de significancia que la prueba F determinada cuando α se establece en el nivel de significancia .05 para la MDS.

Las investigaciones realizadas por Finner (1990) y Hayter (1986) muestran que la MDS era un método de desigualdades confiable si el número de tratamientos era menor que 3, pero no para $t > 3$.

La prueba de la MDS se ilustra con los datos del ejemplo 3.3. Los cálculos para la MDS requieren el valor crítico de la prueba t de Student, $t_{.025,26} = 2.12$, y el error estándar de la diferencia entre dos medias de tratamiento, $\sqrt{2CME/r} = \sqrt{2(15)/5} = 2.45$. La MDS calculada es $\text{MDS}(.05) = 2.12(2.45) = 5.2$.

La hipótesis nula $H_0: \mu_i = \mu_j$ se rechaza si

$$|\bar{y}_i - \bar{y}_j| > \text{MDS}(.05) = 5.2$$

Un método conveniente para realizar pruebas es construir una tabla de diferencias con las medias en orden creciente, como se muestra en la tabla 3.13. Cada diferencia en la tabla se calcula como la diferencia entre la media en el título de la columna y una media de menor valor en la columna de la izquierda. Un asterisco indica las diferencias que exceden $\text{MDS}(.05) = 5.2$.

El resultado de todas las comparaciones por pares que se muestran en la tabla 3.13 indica que la resistencia de la soldadura con la técnica B excede a las otras y que la técnica C excede a la A.

Prueba Student-Newman-Keuls (SNK) de intervalos múltiples

La prueba SNK es una de muchas pruebas de intervalos múltiples. Se basa en el estadístico de Student en la ecuación (3.43), pero al contrario del método de Tukey,

Tabla 3.13 Resultado de la prueba de MDS de las diferencias entre las medias de resistencia para cuatro técnicas de soldadura

Técnica	Media	Técnica			
		A	D	C	B
A	69	—	2	6*	14*
D	71	—	4	—	12*
C	75	—	—	—	8*
B	83	—	—	—	—

*La diferencia excede MDS (.05) = 5.2.

el resultado es una prueba de homogeneidad con tasas de error con respecto al experimento en el sentido débil (Hsu, 1996). El valor crítico del intervalo de Student con la prueba SNK se basa en la separación de un par específico de medias que se prueban del conjunto completo de medias ordenadas. La prueba fue desarrollada de manera independiente por Newman (1939) y Keuls (1952) y se clasifica como una prueba de intervalos múltiples, ya que se usan dos o más intervalos entre medias como criterio. La prueba SNK se describe en el cuadro 3.13.

Cuadro 3.13 Prueba de intervalos múltiples Student-Newman-Keuls

El criterio SNK es:

$$\text{SNK}(k, \alpha_E) = q_{\alpha, k, v} \sqrt{\frac{s^2}{r}} \text{ for } k = 2, 3, \dots, \quad (3.51)$$

donde $q_{\alpha, k, v}$ es el intervalo estadístico de Student, k el número de medias en el conjunto, v los grados de libertad para la estimación de la varianza del error experimental s^2 y α_E es la tasa de error con respecto al experimento para un conjunto de k medias.

Para las medias mayor y menor en un conjunto de k medias, digamos \bar{y}_i y \bar{y}_j , la hipótesis nula $H_0: \mu_i = \mu_j$ se rechaza si:

$$|\bar{y}_i - \bar{y}_j| > \text{SNK}(k, \alpha_E)$$

La prueba no se realiza si existe un conjunto de medias que contiene a \bar{y}_i y \bar{y}_j con mayor tamaño que k , que no es significativa según el criterio SNK.

La prueba SNK se presenta aquí para demostrar los métodos para las pruebas de intervalos múltiples y porque se dispone de tablas de valores críticos del estadístico de Student para una constante α . Posteriormente se presenta una prueba de intervalos múltiples que proporciona una inferencia más fuerte con desigualdades confiables. Aunque esta otra prueba también usa el intervalo estadístico de Student,

no se usa tanto; quizá porque no se tiene acceso sencillo a sus valores críticos, como es el caso en la prueba SNK.

Se ordenan las medias de los tratamientos de menor a mayor:

$$\bar{y}_{[1]} \leq \bar{y}_{[2]} \leq \bar{y}_{[3]} \leq \dots \leq \bar{y}_{[t]}$$

donde $\bar{y}_{[1]}$ es la media del tratamiento con el valor más pequeño y $\bar{y}_{[t]}$ es la media del tratamiento con el valor más grande. El valor crítico para cada par de medias depende del número de medias entre el par específico que se examina.

La prueba SNK para las medias del ejemplo 3.3, con una tasa de error tipo I con respecto al experimento de $\alpha_E = .05$ requiere tres valores críticos del intervalo estadístico de Student (en la tabla VII del apéndice), con $\alpha_E = .05$, $v = 16$ y $k = 2, 3$ y 4. Con un error estándar $\sqrt{CME/r} = \sqrt{15/5} = 1.73$, el estadístico SNK se calcula a partir de:

$$SNK(k, .05) = q_{.05,k,16}(1.73) \quad \text{para } k = 2, 3, 4$$

Los tres valores críticos del intervalo estadístico de Student y $SNK(k, .05)$ para el ejemplo 3.3 son:

$k:$	2	3	4
$q_{.05,k,16}:$	3.00	3.65	4.05
$SNK(k, .05):$	5.2	6.3	7.0

Los valores críticos de la prueba SNK aumentan con el número de medias en el conjunto. Si la distancia entre las dos medias en el arreglo ordenado aumenta, se requiere una diferencia mayor entre las medias para establecer que las medias son distintas. Para el conjunto mínimo de $k = 2$, $SNK(2, .05)$ es igual a $MDS(.05) = 5.2$, y para el conjunto máximo de $k = 4$, $SNK(4, .05)$ es igual a $DHS(4, .05) = 7.0$. La prueba SNK es más conservadora que la MDS, pero menos conservadora que la DHS en cuanto a las diferencias necesarias para rechazar la hipótesis nula. Los resultados de la prueba $SNK(k, .05)$ para el ejemplo 3.3 se muestran en la tabla 3.14.

Tabla 3.14 Resultados de la $SNK(k, .05)$ con las diferencias entre las medias de resistencia para cuatro técnicas de soldadura

<i>Método</i>	<i>Media</i>	<i>Método</i>					<i>k</i>	$SNK(k, .05)$	
		<i>A</i>	<i>D</i>	<i>C</i>	<i>B</i>				
A	69	—	2	6	14*	—	4	—	7.0
D	71	—	—	4	12*	—	3	—	6.3
C	75	—	—	—	8*	—	2	—	5.2
B	83	—	—	—	—	—	—	—	—

*Las diferencias mostradas son de la mayor menos la menor. Las diferencias exceden $SNK(k, .05)$.

Las diferencias entre las medias de tratamiento para el mismo intervalo k , se encuentran en una diagonal que va de la parte superior izquierda a la inferior derecha en la tabla 3.14. La prueba SNK comienza con una comparación entre las medias máxima y mínima, $k = 4$. Si la diferencia máxima no excede el valor crítico, no se realizan más pruebas. Si esa diferencia es significativa, se prueban las diferencias entre las medias con un intervalo ($k - 1$).

Si cualquier par de medias con intervalo de ($k - 1$) no son significativas, no se realizan más pruebas para los otros pares entre ese par específico. Por definición, ningún subgrupo de medias contenido en un grupo no significativo puede ser significativo. Por ejemplo, si no existe una diferencia significativa entre $\bar{y}_{[1]}$ y $\bar{y}_{[3]}$, entonces no deben realizarse pruebas entre $\bar{y}_{[1]}$ contra $\bar{y}_{[2]}$ o $\bar{y}_{[2]}$ contra $\bar{y}_{[3]}$. La prueba procede hasta que no hay diferencias significativas.

La diferencia máxima, $k = 4$, es 14 para B contra A, como excede SNK(4, .05) = 7.0, la prueba sigue con comparaciones para un intervalo $k = 3$, B contra D y C contra A. La diferencia entre C y A no excede SNK(3, .05) = 6.3. Entonces no se hacen más pruebas entre los pares de medias que se encuentren entre las de C y A; que son D contra A y C. La diferencia para B contra D excede SNK(3, .05) = 6.3, por lo que la prueba continúa dentro del grupo de medias que están entre B y D. La diferencia para B contra C excede SNK(2, .05) = 5.2 y es la única prueba necesaria en el intervalo de $k = 2$ medias. La prueba SNK juzga que el método B es diferente de los otros y que no hay diferencia entre los demás.

Con números de réplicas desiguales para las pruebas DHS y SNK, con frecuencia se usa la media armónica de los números de réplicas de todos los grupos de tratamiento, r_h , para estimar el error estándar de todas las comparaciones, a fin de simplificar los cálculos. Sin embargo, Hsu (1996) demostró que su uso conduce hacia una inferencia estadística inválida en general, ya que reduce de manera considerable los niveles de confianza. La media armónica es:

$$r_h = \left[\frac{1}{t} \sum_{i=1}^t \left(\frac{1}{r_i} \right) \right]^{-1} \quad (3.52)$$

Pruebas de intervalos múltiples para desigualdades de confianza

Se ha desarrollado varias pruebas de intervalos múltiples que proporcionan una inferencia de las desigualdades de confianza más fuerte que la que suministra la prueba SNK. Einot y Gabriel (1975) propusieron una elección de α_k para probar la diferencia entre dos medias con un intervalo k , en un conjunto de t tratamientos. Una modificación sugerida por varios autores coincide en el uso de:

$$\alpha_k = \begin{cases} 1 - (1 - \alpha_E)_t^k & \text{si } k = 2, \dots, t - 2 \\ \alpha & \text{si } k = t - 1, t \end{cases} \quad (3.53)$$

para la tasa de error con respecto al experimento deseada α_E . Se usaría el estadístico estandarizado de Student para k medias con el valor crítico $q_{\alpha_k, k, v}$.

La prueba proporciona la inferencia de desigualdades de confianza si los valores críticos son no decrecientes para valores crecientes de k para el intervalo (Hsu,

1996). La prueba se puede realizar con facilidad si se dispone de una computadora para cuantificar el intervalo estadístico de Student.

A continuación se muestran los valores de α_k y los valores críticos del intervalo estadístico de Student calculados con un programa de cómputo para el ejemplo de la soldadura y pueden compararse con los de la prueba SNK anterior. El estadístico para una decisión en cuanto a la significancia de una diferencia entre dos medias es $EG(k, \alpha_k) = q_{\alpha_k, k, v} \sqrt{s^2/r}$. Recuerde que había $t = 4$ tratamientos, $r = 5$ réplicas y un error estándar de 1.73 con $v = 16$ grados de libertad para el cuadrado medio del error.

$k:$	2	3	4
$\alpha_k:$.025	.05	.05
$q_{\alpha_k, k, 16}:$	3.49	3.65	4.05
$EG(k, .05):$	6.0	6.3	7.0

Observe que los valores del intervalo estadístico de Student para $k = 3$ y 4 son iguales a los de la prueba SNK, pero el valor para $k = 2$ es un poco mayor debido a la elección de diferente α_k . La prueba SNK usa $\alpha_k = \alpha_E$ para todos los valores de k . Las comparaciones significativas con $EG(k, \alpha_k)$ son las mismas que en la prueba SNK; pero la fortaleza de la inferencia aumenta la de las desigualdades confiables a diferencia de la homogeneidad estricta (de la prueba SNK). Las diferencias entre las dos pruebas se acentúan al crecer el número de tratamientos.

3.9 Resumen de comentarios sobre las comparaciones múltiples

Las hipótesis de investigación y los diseños de tratamiento son los motores que impulsan los métodos de análisis de los resultados observados en un estudio. Un conjunto de tratamientos estructurado para examinar cierta hipótesis de investigación conduce de manera natural a las comparaciones planeadas, como un conjunto de contrastes ortogonales (o no ortogonales), regresiones polinomiales y la comparación de todos los tratamientos contra un control.

Procedimiento CMM para estudios de evaluación

Un conjunto de tratamientos no estructurados constituyen un reto para que el investigador elija un protocolo adecuado para tomar decisiones. Estos estudios incluyen experimentos para evaluar conjuntos de parcelas de cultivo, productos industriales, pesticidas, etc. El procedimiento CMM para seleccionar un subconjunto de tratamientos con la respuesta deseada, es la elección lógica si el objetivo es buscar los mejores productos, pesticidas o cultivos. Los estudios subsecuentes pueden abrir la posibilidad de hipótesis de investigación mejor estructuradas y diseños de tratamiento con oportunidades para probar las hipótesis.

Las comparaciones múltiples requieren definir las tasas de error

Cuando se trata de comparaciones múltiples, el investigador debe tomar algunas decisiones relevantes para las tasas de error y la potencia de las pruebas. Las tasas

de error con respecto al experimento son apropiadas si la unidad de interés conceptual es una familia de comparaciones y el investigador quiere reducir la posibilidad de tomar demasiadas decisiones incorrectas en la familia de comparaciones.

Se pueden determinar las probabilidades para los errores tipo II y la potencia de las pruebas con respecto a un experimento, pero en general se expresan como tasas con respecto a las comparaciones. Las tasas de error con respecto a las comparaciones más liberales tendrán como resultado un tasa de error tipo II más baja y una prueba más poderosa si el resto de las condiciones permanecen constantes. Para una familia de comparaciones, el investigador más conservador utilizará las tasas de error con respecto al experimento en el sentido fuerte. En estas circunstancias la familia de comparaciones como unidad conceptual es importante para el investigador. Una familia típica es la comparación de todos los tratamientos con uno de control, mediante el procedimiento de Dunnett.

La mayor parte de los buenos experimentos se diseñan de manera que sean eficientes en el uso de los recursos y tiempo disponibles, y es más eficiente diseñar un experimento que responda a múltiples preguntas relacionadas. En cualquier experimento bien planeado existe una familia de comparaciones relacionadas y no debe contemplarse una comparación totalmente aislada de las demás; así, las pruebas basadas en tasas de error con respecto al experimento son más adecuadas.

Elección de un procedimiento de comparación por pares congruente con la filosofía del investigador

La selección de un método de comparación por pares adecuado es difícil, pues cada uno tiene ventajas y desventajas. La mejor alternativa es elegir una prueba coherente con la filosofía que se siga y usarla en todas las pruebas de comparación por pares. Es preferible una prueba de un valor crítico para cada experimento. Algunos análisis informativos sobre muchos métodos de comparación por pares se pueden encontrar en Hsu (1996) y en las extensas notas sobre comparaciones múltiples de J. W. Tukey escritas en 1953 (Tukey, 1994). Pueden ser útiles algunos estudios de Jones (1984), Carmer y Walker (1985) y Saville (1990) sobre su utilización.

El método de Tukey proporciona la mejor protección contra errores de decisión, junto con inferencias fuertes sobre la magnitud y dirección de las diferencias con el ICS de $100(1 - \alpha)\%$.

Hayter (1990) proporciona un método para los límites inferiores de confianza simultáneos de un lado, $\mu_i - \mu_j > L$ para toda $i > j$, que es útil para la inferencia direccional necesaria en algunos tipos de estudios. Estos límites se calculan con:

$$\widehat{\mu}_i - \widehat{\mu}_j - q_a^* \sqrt{\frac{2s^2}{r}} \quad \text{para toda } i > j \quad (3.54)$$

Y serán más exactos que los proporcionados mediante procedimientos de dos lados, como el método de Tukey. Las tablas de valores críticos para el estadístico q_a^* se pueden encontrar en Hayter y Liu (1996).

Recomendación general

La recomendación general para realizar comparaciones múltiples entre tratamientos, es usar la hipótesis de investigación y el diseño de tratamientos para elegir los procedimientos adecuados, con la fuerza de inferencia deseada en el estudio. Los intervalos de confianza proporcionan la inferencia más poderosa, con la magnitud y la dirección de las diferencias, seguidas por la fuerza de las direcciones confiables y después las desigualdades confiables. Si las comparaciones por pares son la única alternativa que queda, se recomienda un procedimiento de intervalos de confianza, que proporciona magnitud y dirección de la inferencia.

EJERCICIOS

1. Utilizando los datos sobre los retrasos en el tránsito del ejercicio 2.1:
 - a. Realice un análisis de varianza de los datos y estime los siguientes contrastes y sus errores estándar:
 - (i) un contraste entre los semáforos programados y el promedio de los semiactivados y activados.
 - (ii) un contraste entre los semáforos semiactivados y activados.
 - b. Calcule la suma de cuadrados para cada contraste y muestre que su suma es igual a la suma de cuadrados de los tratamientos en el análisis de varianza.
 - c. Pruebe la hipótesis nula para cada contraste, $H_0: C = 0$, con la prueba t de Student a un nivel de significancia de 0.5.
 - d. Pruebe la hipótesis nula en el inciso c con la prueba F a un nivel de significancia de .05.
 - e. ¿Cuál es la relación entre las pruebas de los incisos c y d?
2. Utilizando los datos sobre las concentraciones de suero T3 en los experimentos con gallinas del ejercicio 2.3. Las comparaciones de interés eran las diferencias de concentración de suero T3 en las etapas sucesivas: 1) premuda contra ayuno, 2) ayuno contra 60 g de salvado, 3) 60 contra 80 g de salvado y 4) 80 g de salvado contra mezcla de malta:
 - a. Estime cada contraste y su error estándar.
 - b. Pruebe la hipótesis nula para uno de los contrastes con la prueba t de Student.
 - c. Pruebe la hipótesis nula para uno de los contrastes con la prueba F .
 - d. Suponga que debe probar los cuatro contrastes con una tasa de error con respecto a la comparación de .05. Calcule la tasa de error con respecto al experimento máximo, para esta familia de cuatro comparaciones.
3. Utilizando los datos de cosecha de lechuga del ejercicio 2.2:
 - a. Calcule el análisis de varianza para los datos.
 - b. Determine la mejor función polinomial de respuesta que describe la relación entre la cosecha de lechuga y el fertilizante de nitrógeno a un nivel de significancia de .05.
 - c. Vea la tabla 3.7. Construya una tabla similar para los resultados de este problema.
 - d. Grafique las medias observadas junto con la ecuación estimada.
4. Utilizando los datos del ejercicio 2.1:
 - a. Calcule un ICS para comparaciones múltiples del 95%, con el mejor, donde el “mejor” se define como el semáforo con menor tiempo de retraso.

- b. Seleccione el tipo o tipos de semáforo con menor retraso, con una probabilidad de elección correcta de .95.
5. El laboratorio clínico de un hospital mide con un espectrofotómetro la concentración de colesterol en muestras de suero. Un día específico, el laboratorio analizó muestras de ocho pacientes y se prepararon dos muestras de cada uno. Los siguientes datos son las concentraciones de colesterol (mg/dl).

<i>Paciente</i>	<i>Colesterol (mg/dl)</i>
1	167.3, 166.7
2	186.7, 184.2
3	100.0, 107.9
4	214.5, 215.3
5	148.5, 149.5
6	171.5, 167.3
7	161.5, 159.4
8	243.6, 245.5

- a. Calcule el ICS para comparaciones múltiples del 95% con el mejor, donde “mejor” se define como el paciente con mayor nivel de colesterol.
- b. Elija un subconjunto de pacientes que contenga al paciente con nivel de colesterol más alto, con una probabilidad de elección correcta de .95.
6. Un conjunto de comparaciones de interés en el experimento con gallinas del ejercicio 2.3, era la comparación de suero T3 para cada uno de los otros estados contra el de la etapa de premuda.
- Calcule el ICS del 95% para las otras etapas contra la de premuda, mediante el método de Dunnett.
 - ¿Cuáles son sus conclusiones?
7. En un experimento con cinco réplicas y cuatro tratamientos con un diseño totalmente aleatorizado, se cultivaron secciones de tejido de planta de tomate con diferentes cantidades y tipos de azúcares. El crecimiento de tejidos en cada cultivo se da en la tabla siguiente como mm × 10.

<i>Control</i>	<i>3% de glucosa</i>	<i>3% de fructosa</i>	<i>3% de sacarosa</i>
45	25	28	31
39	28	31	37
40	30	24	35
45	29	28	33
42	33	27	34

- a. Calcule el ICS del 95% de las comparaciones de todos los tratamientos contra el tratamiento de control, mediante el método de Dunnett.
- b. ¿Cuáles son sus conclusiones?

8. En la siguiente tabla se encuentran los coeficientes mostrados por un colega para un conjunto de contrastes entre medias de tratamiento. Él le pide que los verifique.

Tratamiento	A	B	C	D	E
C1	1	3	-1	-1	-1
C2	1	-1	0	-1	1
C3	-1	1	-1	1	-1
C4	0	0	2	-1	-1

- a. ¿Constituye un contraste cada conjunto de coeficientes propuesto? Justifique su respuesta.
 - b. ¿Son ortogonales C1 y C2? Justifique su respuesta.
 - c. Construya un contraste ortogonal para C4, diferente de lo que se muestra.
9. Utilice la prueba de Scheffé con un nivel de significancia de .05 para probar la hipótesis nula sobre los contrastes del ejercicio 3.1.
10. Utilice la prueba *t* de Bonferroni con un nivel de significancia de .05, para probar la hipótesis nula sobre los contrastes del ejercicio 3.1.
11. Usando los datos del suero T3 del ejercicio 2.3:
- a. Realice todas las comparaciones por pares, con el método de Tukey a un nivel de significancia de .05.
 - b. Realice todas las comparaciones por pares, con la diferencia menos significativa a un nivel de significancia de .05.
 - c. Realice todas las comparaciones por pares, con la prueba SNK de intervalos múltiples a un nivel de significancia de .05.
 - d. ¿En qué difieren los resultados de las tres pruebas?
 - e. Explique por qué difieren los resultados.
 - f. Calcule el ICS del 95% para todas las comparaciones por pares, con el método de Tukey.
 - g. ¿Qué información adicional obtuvo con el cálculo realizado en el inciso f?
12. La siguiente es la descripción de un experimento sobre sistemas de trabajo humano. En un sistema de este tipo, como una línea de ensamble, con frecuencia los trabajadores requieren mover objetos a un lugar predeterminado con su mano.

El propósito específico del estudio es determinar la exactitud con la que los individuos alcanzan ese lugar específico en un plano horizontal (por ejemplo, encima de una mesa), si se limita su campo de visión de manera que no incluye el objetivo.

Investigaciones previas condujeron a la hipótesis de que los movimientos “distales” (movimientos que se alejan del cuerpo) son más exactos que los “proximales” (hacia el cuerpo). También se planteó la hipótesis de que los movimientos en direcciones cardinales (hacia enfrente, hacia el cuerpo y laterales) no eran más exactos que en otras direcciones.

Se establecieron los objetivos en una circunferencia con un radio de 10 pulgadas (figura 3.4). El sujeto se sentó de manera que un movimiento de la mano derecha a una posición de 90° desde el punto de partida era un movimiento distal; un movimiento de la mano derecha a la posición de 270° era proximal; los movimientos a las posiciones 0° y 180° eran laterales; los movimientos a las posiciones de 45° y 135° no eran movimientos cardinales, como tampoco lo eran a las posiciones de 225° y 315°. Los objetivos a 0°, 90°, 180° y 270° eran direcciones cardinales.

Se capacitó a 16 individuos para el estudio. En la prueba real el investigador decía en voz alta al azar los ángulos de las posiciones y el sujeto marcaba su intento de alcanzar el objetivo que no podía ver. Se registró la distancia entre la marca del sujeto y el objetivo. Se calculó la distancia promedio entre cada objetivo y las observaciones. Un valor más pequeño del promedio representa mayor exactitud.

Muestre en una tabla los coeficientes necesarios para una comparación entre medias para las siguientes comparaciones de exactitud entre las direcciones de los movimientos.

- C1: Distal contra proximal en general
- C2: Cardinal contra no cardinal
- C3: Lateral contra distal
- C4: Lateral contra proximal
- C5: Distal no cardinal contra proximal no cardinal
- C6: Distal cardinal contra proximal cardinal

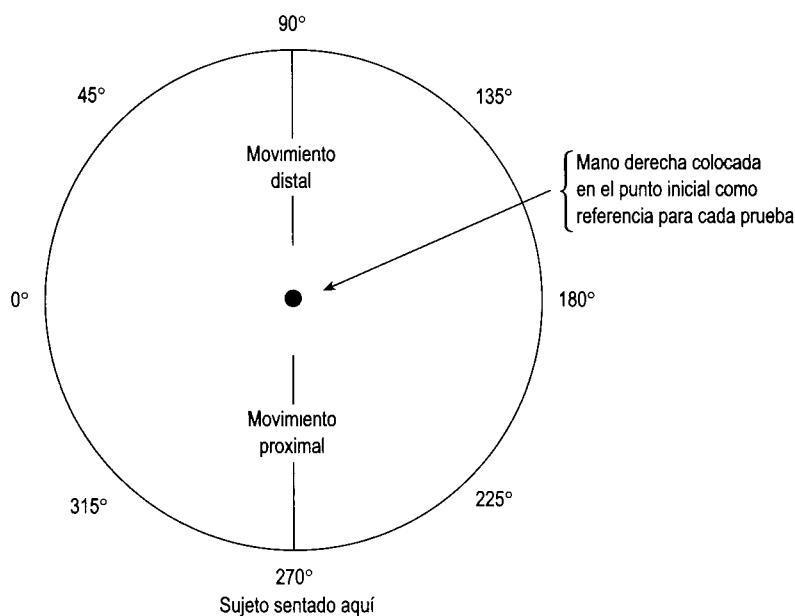


Figura 3.4 Objetivos establecidos en un círculo con radio de 10 pulgadas

Apéndice 3A Funciones lineales de las variables aleatorias

Una función lineal de las variables aleatorias y_1, y_2, \dots, y_n se define como:

$$c = \sum_{i=1}^n k_i y_i = k_1 y_1 + k_2 y_2 + \cdots + k_n y_n$$

Si el valor esperado o media de y_i es $E(y_i) = \mu_i$, entonces el valor esperado o media de c es:

$$\begin{aligned}\mu_c &= E(c) = E\left(\sum_{i=1}^n k_i y_i\right) = \sum_{i=1}^n k_i E(y_i) = \sum_{i=1}^n k_i \mu_i \\ &= k_1 \mu_1 + k_2 \mu_2 + \cdots + k_n \mu_n\end{aligned}$$

La varianza de una función lineal $c = \sum k_i y_i$ es:

$$\sigma_c^2 = \sum_{i=1}^n k_i^2 \sigma_i^2 + 2 \sum_{i < j} k_i k_j \sigma_{ij}$$

Por ejemplo, si $c = y_1 - y_2$ con $k_1 = 1$ y $k_2 = -1$, entonces:

$$\sigma_c^2 = k_1^2 \sigma_1^2 + k_2^2 \sigma_2^2 + 2k_1 k_2 \sigma_{12} = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12}$$

Si las y_i son independientes, entonces $\sigma_{ij} = 0$ y:

$$\sigma_c^2 = \sum_{i=1}^n k_i^2 \sigma_i^2$$

Por lo tanto, si y_1 y y_2 son independientes en $c = y_1 - y_2$, la varianza de c es $\sigma_c^2 = \sigma_1^2 + \sigma_2^2$.

Media muestral

La media de una muestra con r observaciones independientes de una distribución normal con media μ y varianza σ^2 es una función lineal:

$$\bar{y} = \frac{1}{r} y_1 + \frac{1}{r} y_2 + \cdots + \frac{1}{r} y_r$$

donde $k_r = \frac{1}{r}$. El valor esperado de la media muestral es:

$$\mu_{\bar{y}} = E(\bar{y}) = \frac{1}{r} E(y_1) + \cdots + \frac{1}{r} E(y_r) = \frac{1}{r} (r\mu) = \mu$$

Como las varianzas de las observaciones son todas iguales, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_r^2 = \sigma^2$, entonces la varianza de la media es:

$$\sigma_y^2 = \frac{1}{r^2} (r\sigma^2) = \frac{\sigma^2}{r}$$

Función lineal de las medias muestrales

Si t muestras son independientes y r_i es el número de observaciones en la i -ésima muestra, entonces una función lineal de las medias muestrales:

$$c = k_1\bar{y}_1 + k_2\bar{y}_2 + \dots + k_t\bar{y}_t$$

tiene una media:

$$\begin{aligned}\mu_c &= E(c) = k_1E(\bar{y}_1) + k_2E(\bar{y}_2) + \dots + k_tE(\bar{y}_t) \\ &= k_1\mu_1 + k_2\mu_2 + \dots + k_t\mu_t\end{aligned}$$

y una varianza:

$$\sigma_c^2 = k_1^2 \left(\frac{\sigma_1^2}{r_1} \right) + k_2^2 \left(\frac{\sigma_2^2}{r_2} \right) + \dots + k_t^2 \left(\frac{\sigma_t^2}{r_t} \right)$$

Si todas las varianzas de las muestras son iguales, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2 = \sigma^2$, entonces

$$\sigma_c^2 = \sigma^2 \left(\frac{k_1^2}{r_1} + \frac{k_2^2}{r_2} + \dots + \frac{k_t^2}{r_t} \right)$$

4 Diagnóstico de la concordancia entre los datos y el modelo

Si no se satisfacen ciertas suposiciones con respecto a los datos, el análisis de varianza puede conducir a inferencias erróneas. En este capítulo se estudian los métodos de diagnóstico para detectar suposiciones fallidas, junto con las transformaciones de datos que se pueden usar para estudiar los problemas, se sugiere una generalización del modelo lineal para el análisis como alternativa y, además, se introduce un método para evaluar el ajuste de un modelo con los datos.

4.1 Un análisis válido depende de suposiciones válidas

La validez de las estimaciones y pruebas de hipótesis para los análisis derivados del modelo lineal, se apoya en los valores de varias suposiciones clave. Se supone que los errores experimentales aleatorios son independientes, siguen una distribución normal con una media igual a cero y tienen una varianza común (σ^2) para todos los grupos de tratamiento. Cualquier discrepancia entre los datos y una o más de estas suposiciones afecta las estimaciones de las medias de tratamiento y las pruebas de significancia del análisis de varianza.

En Eisenhart (1974) y Cochran (1947) se presenta un resumen de los estudios sobre las suposiciones para el análisis de varianza y los efectos de la desviación de las mismas. Ito (1980) resumió la investigación sobre la validez de los procedimientos de prueba del análisis de varianza con discrepancias respecto a las suposiciones.

4.2 Efectos de la falta de cumplimiento de las suposiciones

Cochran (1947) demostró que si los errores experimentales tienen una correlación positiva, la precisión real de la media del tratamiento es menor que la estimada, pues la estimación usual del error estándar es demasiado pequeña. Por el contrario, la precisión real es mayor que la estimada si las correlaciones de error son negativas. En los experimentos, el mejor seguro contra una correlación excesiva

de las observaciones es la aleatorización de las unidades experimentales al formar los grupos de tratamiento; y el muestreo aleatorio de las poblaciones en los estudios por observación.

Si σ^2 difiere de un grupo de observaciones a otro, en general, los errores estándar de las medias de tratamiento serán mayores que si es constante para todas las observaciones. Los niveles de significancia establecidos para las pruebas F y t pueden ser mayores o menores que los niveles de significancia obtenidos en realidad. Los estudios teóricos de Box (1954a) produjeron resultados acerca de los niveles de significancia reales de la prueba F , llevada a cabo con un nivel de significancia de .05 con varianzas de grupo diferentes, para igual o desigual número de réplicas. Si se tiene una razón 1:3 para la menor:mayor varianza de grupo, los niveles de significancia reales van de .056 a .074 para un igual número de réplicas, pero son de .013 a .14 con número de réplicas diferente. Con una razón de varianza 1:7, el nivel de significancia real era .12 con igual número de réplicas.

Las pruebas F para el análisis de varianza son bastante robustas contra las discrepancias con la distribución normal. Ito (1980) citó los resultados de estudios teóricos y empíricos sobre los efectos de la no normalidad en la que los niveles de significancia reales van de .03 a .06, para las pruebas realizadas con niveles de significancia de .05.

Las condiciones ideales rara vez se cumplen en los estudios reales. Las discrepancias menores de los datos con respecto a la independencia, la distribución normal supuesta y las varianzas homogéneas, generalmente no ocasionan modificaciones sustanciales en la eficiencia de las estimaciones y en los niveles de significancia de las pruebas. Es posible que las discrepancias mayores, en especial una heterogeneidad excesiva de la varianza o alguna heterogeneidad de la varianza con números de réplicas desiguales, afecte en forma importante las inferencias estadísticas. El resto de la presentación se enfoca en estos casos.

4.3 Los residuales son la base de las herramientas de diagnóstico

Los **residuales** observados son la base de muchas de las principales herramientas de diagnóstico que se usan para verificar si las suposiciones del modelo lineal son adecuadas. Los residuales son estimaciones de los errores experimentales, calculados como las diferencias entre las observaciones y las estimaciones de las medias de los tratamientos, es decir:

$$\hat{e}_{ij} = y_{ij} - \hat{\mu}_{i\cdot} = y_{ij} - \bar{y}_i. \quad (4.1)$$

Se recomienda como primer paso en el diagnóstico examinar la magnitud de los residuales y su relación con las otras variables.

Los residuales proporcionan evaluaciones visuales de las suposiciones del análisis de varianza, para varianzas homogéneas y distribuciones normales de los errores experimentales; la suposición de varianza homogénea se evalúa con una gráfica de los residuales contra las medias estimadas de los tratamientos. Y la suposición de distribución normal mediante una gráfica de probabilidad normal.

Las técnicas se demuestran con observaciones de un estudio que no cumple de forma satisfactoria con las suposiciones del modelo lineal.

Ejemplo 4.1 Número de cangrejos ermitaños en hábitats costeros

Un biólogo marino estaba interesado en la relación entre distintos ambientes costeros y las poblaciones de cangrejos ermitaños que habitan el lugar. El biólogo contó los cangrejos en 25 secciones, ubicadas al azar en seis sitios diferentes de la costa. En la tabla 4.1 se encuentra el resumen estadístico de los seis sitios, con el cuadrado de la media del error. Los datos se encuentran en el apéndice 4A.

Deben calcularse 150 residuales para el conjunto de datos resumido en la tabla 4.1. Como ilustración, en el cuadro 4.1 se muestran los primeros cinco para las observaciones en el sitio 1, $\hat{e}_{1j} = y_{1j} - \bar{y}_1$.

Tabla 4.1 Medias, desviaciones estándar y valores máximos y mínimos para la cuenta de cangrejos ermitaños, en seis lugares de la costa

Sitio	Media	Mediana	Desviación estándar	Mínimo	Máximo
1	33.80	17	50.39	0	233
2	68.72	10	125.35	0	466
3	50.64	5	107.44	0	407
4	9.24	2	17.39	0	82
5	10.00	2	19.84	0	94
6	12.64	4	23.01	0	95

$$CME = 5,170 \text{ con } 144 \text{ grados de libertad}$$

Fuente: Department of Ecology and Evolutionary Biology, University of Arizona.

Cuadro 4.1 Observaciones, medias y residuales para el sitio 1 del estudio de cangrejos ermitaños

Sitio	Sección	y_{1j}	\bar{y}_1	\hat{e}_{1j}
1	1	0	33.8	-33.8
1	2	0	33.8	-33.8
1	3	22	33.8	-11.8
1	4	3	33.8	-30.8
1	5	17	33.8	-16.8

Gráfica de probabilidad de los residuales para evaluar la suposición de distribución normal

La media es mucho más grande que la mediana en los seis sitios de la tabla 4.1, lo que indica una distribución sesgada no normal de las observaciones. La gráfica de probabilidad normal de los residuales se usa para evaluar la suposición de distribución normal, comparando de manera visual la distribución acumulada de los

residuales contra la distribución normal estándar. Una gráfica de probabilidad normal organiza los residuales en orden creciente y los grafica contra los *cuantiles*¹ correspondientes de la distribución normal estándar. El residual de orden i -ésimo tiene una frecuencia acumulada de i/N en una muestra de tamaño N .

El cuantil de una variable normal estándar correspondiente se determina para una proporción acumulada² de $f_i = (i - 0.5)/N$. Los valores de los cinco residuales más pequeños y los cinco más grandes y los cuantiles normal estándar correspondientes para las muestras de cangrejo ermitaño se muestran en la tabla 4.2.

Tabla 4.2 Residuales ordenados, probabilidad acumulada (f_i) y cuantiles normal estándar correspondientes para las muestras de cangrejo ermitaño

Orden	Residual	f_i	Cuantiles normales
1	-68.72	0.0033	-2.713
2	-68.72	0.0100	-2.326
3	-68.72	0.0167	-2.128
4	-68.72	0.0233	-1.989
5	-67.72	0.0300	-1.881
:	:	:	:
146	199.20	0.9700	1.881
147	263.36	0.9767	1.989
148	346.28	0.9833	2.128
149	356.36	0.9900	2.326
150	397.28	0.9967	2.713

El cuantil de una variable normal se conoce por una probabilidad acumulada dada, y el valor del cuantil de un residual se conoce por su frecuencia acumulada en el conjunto de datos residuales. Se forman pares con los cuantiles correspondientes de los residuales y la distribución normal estándar y se grafican como valores X y Y en una gráfica bivariada llamada gráfica *cuantil-cuantil*. Si los cuantiles de los residuales son iguales a los de la variable normal para la misma frecuencia acumulada, la gráfica será una línea recta.

En la figura 4.1 se muestra la gráfica de la probabilidad normal de las muestras de cangrejo ermitaño, en ella la recta pasa por los cuartiles inferior y superior (percentiles 25 y 75) de los datos. Los cuantiles de los pares de residuales y sus cuantiles normal estándar correspondientes no están sobre la línea recta que indica

¹ El cuantil de f , $q(f)$, es un valor tal que una proporción aproximada, f , de los datos es menor o igual a $q(f)$.

² El valor de f está diseñado para evitar un valor de $f = 1$, de no evitarse, habría un valor no finito de su desviación estandarizada. En general, cualquier pequeña modificación de $f_i = i/N$ para evitar $f = 1$ es adecuada para estas gráficas. El valor presentado es del programa S-PLUS, usado para las gráficas de esta sección.

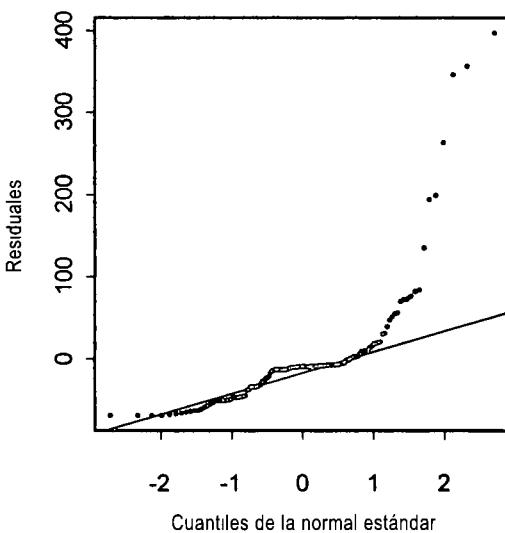


Figura 4.1 Gráfica de la probabilidad normal de los residuales para el estudio de cangrejos ermitaños

la gráfica de probabilidad normal. Esto muestra la gráfica de probabilidad para una distribución sesgada a la derecha con respecto a la distribución normal estándar. Los valores por encima de la recta que se encuentran en la esquina superior derecha, son residuales con valores positivos mayores al esperado para la distribución normal estándar; mientras que la serie de valores por encima de la recta de la esquina inferior izquierda, son residuales con valores negativos menores que el esperado.

Gráfica de residuales para evaluar las suposición de homogeneidad de varianza

Al graficar los residuales contra los valores estimados de las medias de tratamiento, se obtiene una evaluación visual sencilla de la suposición de varianzas iguales en los grupos de tratamiento. Si la variabilidad de las observaciones alrededor de las medias de tratamientos difiere de un grupo a otro, el conjunto de residuales correspondiente reflejará las diferencias en la variación. En la figura 4.2 se muestra una gráfica de los residuales contra las medias estimadas de los sitios para las muestras de cangrejos ermitaños.

La gráfica refleja las diferencias de las desviaciones estándar entre los sitios de la tabla 4.1, la dispersión de los residuales varía mucho entre los seis sitios, pues la variabilidad de los residuales aumenta con el valor de las medias estimadas. Si las varianzas son heterogéneas, la gráfica de residuales contra los valores estimados a menudo tiene la forma de embudo que se muestra en la figura 4.2. La falta de simetría de los residuales alrededor de cero (línea punteada) indica una distribución asimétrica de las observaciones, con una cola larga a la derecha.

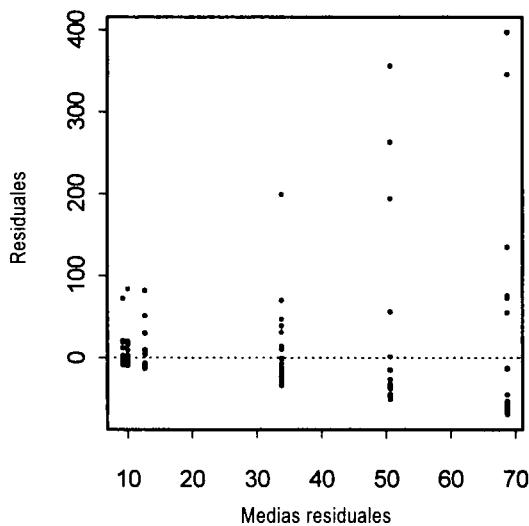


Figura 4.2 Gráfica de residuales contra las medias estimadas de los sitios para las muestras de cangrejo ermitaño

La gráfica de dispersión (*spread-location*) $s-l$, es otra gráfica residual que revela más sobre las varianzas heterogéneas. Las tendencias en la gráfica $s-l$ se pueden usar para revelar las posibles relaciones entre las medias y las varianzas de los grupos en tratamiento.

Las raíces cuadradas de los valores absolutos de los residuales, $\sqrt{|\hat{e}_{ij}|}$, se usan para medir la *dispersión* de los residuales, ya que su magnitud reflejará la dispersión o variación dentro de un grupo de tratamiento. La raíz cuadrada elimina parte de la asimetría en los residuales, mientras las estimaciones de las medias de grupos de tratamientos miden la *localización*. En la figura 4.3 se muestra la gráfica $s-l$ para el estudio de cangrejos ermitaños.

Uniendo, mediante líneas rectas, las medianas de las $\sqrt{|\hat{e}_{ij}|}$ para cada sitio en la gráfica muestran el incremento en su magnitud con un incremento en las medias de los sitios. Esta tendencia del incremento en la magnitud de los valores absolutos de los residuales, representa el aumento en la varianza del sitio cuando crece la media.

Pruebas estadísticas para varianzas homogéneas

Prueba Levene (Med)

Existen muchas pruebas estadísticas formales de homogeneidad de las varianzas para diseños totalmente aleatorizados. Conover, Johnson y Johnson (1981) compararon 56 pruebas y determinaron que la mejor era la prueba de **Levene (Med)**.

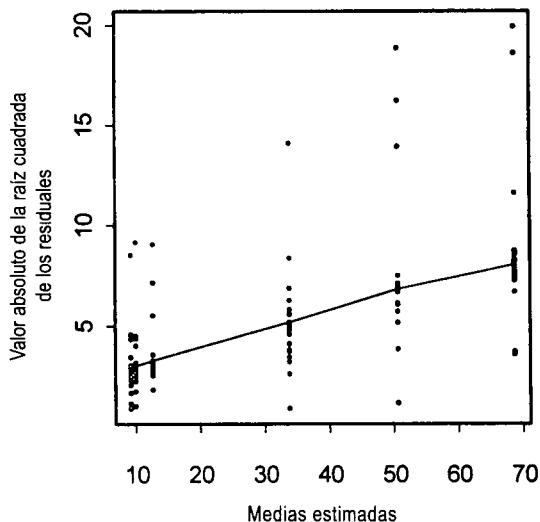


Figura 4.3 Gráfica de separación-localización para los residuales del estudio de cangrejos ermitaños

Sea y_{ij} la j -ésima observación del i -ésimo grupo de tratamiento y \tilde{y}_i su mediana. Sea $z_{ij} = |y_{ij} - \tilde{y}_i|$ el valor absoluto de la diferencia entre una observación y la mediana del grupo de tratamiento i . Para probar la homogeneidad de las varianzas, se calcula el análisis de varianza en un sentido para z_{ij} y del estadístico F_0 :

$$F_0 = \frac{CMT}{CME} = \frac{\sum_{i=1}^t r_i (\bar{z}_{i\cdot} - \bar{z}_{..})^2 / (t-1)}{\sum_{i=1}^t \sum_{j=1}^{r_i} (z_{ij} - \bar{z}_{i\cdot})^2 / (N-t)} \quad (4.2)$$

La hipótesis nula de varianzas homogéneas, $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2$, se rechaza si $F_0 > F_{\alpha/(t-1), (N-t)}$. El estadístico de prueba en la ecuación (4.2) es una modificación de la prueba original introducida por Levene (1960). La modificación sugerida por Brown y Forsythe (1974) fue sustituir la mediana, \tilde{y}_i , con la media, \bar{y}_i , en el cálculo de z_{ij} . Los cálculos de la prueba se ilustran en la tabla 4.3 con cinco observaciones cada uno de los tres primeros sitios del estudio con cangrejos ermitaños. Los valores requeridos por la prueba de Levene (Med) y calculados para todo el conjunto de datos son $CMT = 14,229$; $CME = 4,860$, y $F_0 = 14,229/4,860 = 2.93$. Se concluye que las varianzas son diferentes entre los sitios, ya que se rechaza la hipótesis nula con región crítica $F_0 > F_{.05, 5, 144} = 2.28$.

Tabla 4.3 Ilustración de la prueba de Levene (Med) de homogeneidad de las varianzas con cinco observaciones en tres sitios del estudio con cangrejos ermitaños

Sitio	y_{ij}	\tilde{y}_i	$\frac{z_{ij} =}{ y_{ij} - \tilde{y}_i }$	\bar{z}_i
1	0	3	3	
	0		3	
	22		19	
	3		0	
	17		14	7.8
2	415	14	401	
	466		452	
	6		8	
	14		0	
	12		2	172.6
3	0	4	4	
	0		4	
	4		0	
	13		9	
	5		1	3.6
				$\bar{z}_{..} = 61.3$

$$r_1 = r_2 = r_3 = r = 5, t = 3, N = 15$$

$$SCT = \sum_{i=1}^t r_i (\bar{z}_{i.} - \bar{z}_{..})^2 = 92,896$$

4.4

$$SCE = \sum_{i=1}^t \sum_{j=1}^{r_i} (z_{ij} - \bar{z}_{i.})^2 = 216,539$$

$$F_0 = \frac{CMT}{CME} = \frac{92,896}{2} / \frac{216,539}{12} = 2.57$$

Prueba F Máx

Varias pruebas estadísticas formales son pruebas válidas de la homogeneidad de varianzas en los diseños totalmente aleatorizados cuando los tamaños de las muestras son iguales y las observaciones tienen una distribución normal. Una de las más sencillas es con el estadístico de prueba *F Máx* (Hartley, 1950). La hipótesis nula probada con el estadístico *F Máx* es:

$$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_t^2 \quad (4.3)$$

con la hipótesis alterna de que algunas varianzas difieren.

El estadístico de prueba se calcula como la razón de la mayor varianza observada entre la menor, dentro de los grupos de tratamiento, es decir:

$$F_0 \text{ Máx} = \frac{\max(s_i^2)}{\min(s_i^2)} \quad (4.4)$$

donde $\max(s_i^2)$ y $\min(s_i^2)$ son las varianzas mayor y menor respectivas dentro de los grupos de tratamiento.

La hipótesis nula se rechaza al nivel de significancia α si $F_0 \text{ Máx} > F_\alpha \text{ Máx}$, donde $F_\alpha \text{ Máx}$ es el valor de la variable $F \text{ Máx}$ excedida, con una probabilidad α para t grupos de tratamientos y $v = (r - 1)$ grados de libertad para cada s_i^2 . Los valores críticos del estadístico $F \text{ Máx}$ se encuentran en la tabla VIII del apéndice.

En general, la prueba $F \text{ Máx}$ no se realizaría para el estudio de cangrejos ermitaños, puesto que las muestras no tienen una distribución normal; aquí se efectuó sólo con propósitos ilustrativos. De la tabla 4.1, las desviaciones estándar mayor y menor se encuentran en los sitios 2 y 4 respectivamente; las varianzas de estos grupos son $s_4^2 = 17.39^2 = 302$ y $s_2^2 = 125.35^2 = 15,712$. El valor del estadístico de prueba es:

$$F_0 \text{ Máx} = \frac{15,712}{302} = 52.03$$

El valor crítico para $\alpha = .05$, $t = 6$ y $v = 24$, interpolado a partir de la tabla VIII del apéndice, es $F_{.05} \text{ Máx} = 3.24$; por tanto, la hipótesis nula de varianzas homogéneas se rechaza.

4.4 Búsqueda de residuos inusitados con los residuales

Los valores positivos o negativos demasiado grandes de los residuales estarán muy lejos de la línea recta indicadora de la gráfica normal o de los otros valores en las fronteras superior e inferior de la gráfica de residuales contra medias de tratamiento. Un valor inusitado puede afectar la inferencia estadística porque infla la estimación de la varianza del error experimental e influye en la estimación de la media del tratamiento. Quizá los valores inusitados se deban a errores al reunir o registrar los datos, a errores en la técnica o a una combinación especial de tratamiento y entorno. Antes de descartarlos, debe investigarse qué los causa, para evitar la pérdida de información valiosa.

Residuales estandarizados

Los residuales estandarizados se calculan mediante:

$$\omega_{ij} = \frac{\hat{e}_{ij}}{\sqrt{CME}} \quad (4.5)$$

El *residual estandarizado* (w_{ij}) es útil para verificar rápidamente la presencia de residuos inusitados. La w_{ij} se aproxima a una distribución normal estándar si e_{ij} tiene distribución normal. Un residual se considera inusitado si el valor estandarizado está fuera de los límites ± 3 o ± 4 , ya que la probabilidad de que un valor normal estándar sea mayor que 3 o 4 desviaciones de la media 0 es muy pequeña.

El residual estandarizado más grande para las muestras de cangrejo ermitaño corresponde al residual ordenado 150 en la tabla 4.2 es:

$$w = \frac{397.28}{\sqrt{5,170}} = 5.53$$

Los residuales estandarizados calculados con los residuales ordenados 148 y 149 de la tabla 4.2 son $4.82 = 346.28 / \sqrt{5,170}$ y $4.96 = 356.36 / \sqrt{5,170}$, respectivamente. Los tres residuales mayores están a más de 4 desviaciones estándar de la media 0. La gráfica de probabilidad normal y la gráfica de residuales contra valores esperados indican la posibilidad de que haya otros valores inusitados, derivados de las muestras grandes de cangrejos ermitaños. En este caso, el biólogo desea indagar las condiciones de esas secciones específicas que producen números excepcionalmente grandes.

Se usan residuales estandarizados porque los residuales ordinarios tienen algunas desventajas para el diagnóstico, puesto que los residuales estimados en el mismo grupo de tratamiento no son independientes entre sí, y sus varianzas son heterogéneas de un grupo a otro con diferente número de réplicas. La varianza de un residual estimado es $\sigma_{e_i}^2 = \sigma^2(1 - 1/r_i)$ para el diseño totalmente aleatorizado con r_i réplicas para el i -ésimo tratamiento.

Las gráficas de diagnóstico con réplicas desiguales pueden verse influidas por las varianzas heterogéneas de los residuales ordinarios, también las afectan las correlaciones entre los residuales sin importar el número de réplicas. Las correlaciones entre los residuales tienden a hacer que muestren un mejor ajuste con la distribución normal en la gráfica de probabilidad. En Cook y Weisberg (1982) puede verse una presentación más completa de los residuales, más allá del alcance de este libro.

Residuales de Student

Las varianzas heterogéneas de los residuales ordinarios se pueden corregir utilizando los *residuales de Student*, que en un diseño totalmente aleatorizado son el residual ordinario dividido entre su desviación estándar estimada:

$$\tilde{e}_{ij} = \frac{\hat{e}_{ij}}{\sqrt{CME(1 - 1/r_i)}}$$

Los residuales de Student tienen una varianza constante $\sigma_{\tilde{e}_{ij}}^2 = 1$, pero continúan correlacionados. Se recomienda utilizarlas en lugar de los residuales ordinarios, para las gráficas de residuales en estudios con número de réplicas desiguales.

Las gráficas de residuales y la prueba de Levene (Med) proporcionan una buena evidencia de que las suposiciones de varianzas homogéneas y distribución normal no son adecuadas para los datos del experimento con cangrejos ermitaños. Algunas de sus observaciones son buenos candidatos como residuos inusitados. Si a juicio del investigador, las diferencias con las suposiciones no son demasiado grandes, entonces las consecuencias de ignorarlas no son severas. Cuando existen diferencias grandes con estas suposiciones, como en el caso del estudio con cangrejos ermitaños, deben tomarse algunas decisiones para el siguiente paso en el análisis del estudio; los temas de la siguiente sección tratan algunas de las posibles soluciones.

4.5 Transformaciones estabilizadoras de la varianza para datos con distribuciones conocidas

Las transformaciones se usan para cambiar la escala de las observaciones con el objeto de que éstas cumplan mejor las suposiciones del modelo lineal y ofrezcan inferencias más válidas del análisis de varianza. Las probabilidades de inferencia estadística se aplican sólo a la nueva escala de medición; los niveles de significancia y promedios no se aplican a las medidas originales. Cuando las transformaciones son necesarias, es común realizar el análisis y hacer todas las inferencias en la escala transformada, pero presentar los compendios de medias en la escala de medida original.

Bartlett (1947) resumió muchos aspectos de las transformaciones del análisis de varianza. En esta sección, se estudian varias transformaciones para datos que no tienen distribución normal, pero tienen una distribución de probabilidad conocida. En la sección 4.6 se analiza una transformación basada en una relación empírica entre las medias y las varianzas de las muestras, para datos con distribución no conocida.

Distribución de Poisson

Las observaciones sobre muestras de plantas en cultivos, insectos en plantas, colonias de bacterias en platos, defectos en una superficie y accidentes por unidad de tiempo, pueden tener una distribución Poisson; para la cual la media es igual a la varianza, $\mu_y = \sigma_y^2$. Se recomienda la transformación *raíz cuadrada* para estabilizar las varianzas de las observaciones, esta transformación, $x = \sqrt{y}$, tendrá una varianza constante $\sigma_x^2 = 0.25$, para todos los valores de μ_x . Si la media es pequeña, digamos $\mu_y < 3$, entonces la transformación $x = \sqrt{y + \frac{3}{8}}$ es superior que \sqrt{y} para estabilizar las varianzas (Anscombe, 1948). Si las muestras son grandes no es necesaria la corrección.

Distribución binomial

Las observaciones sobre el número de éxitos en n pruebas independientes siguen una distribución binomial. Los ejemplos incluyen la proporción de artículos de-

fectuosos en lotes de manufactura, de semillas germinadas, de larvas sobrevivientes en estudios de insectos y de plantas que florecen en una sección. La probabilidad binomial estimada es $\hat{\pi} = y/n$, donde y es el número de éxitos en n pruebas independientes con una probabilidad de éxito π . La media y la varianza de la distribución binomial estimada son $\mu = \pi$ y $\sigma^2 = \pi(1 - \pi)/n$, respectivamente, y habrá una relación bien definida entre las proporciones observadas y la varianza de los datos observados.

La transformación *arco seno* o *angular* es recomendable para estabilizar las varianzas de las observaciones con distribución binomial. La transformación arco seno, $x = \text{sen}^{-1}\sqrt{\hat{\pi}}$, tiene una varianza constante, $\sigma_x^2 = 1/4n$, para toda π , si el ángulo se expresa en radianes y $\sigma_x^2 = 821/n$, si el ángulo se expresa en grados. Si n es pequeño, digamos $n < 50$, entonces Anscombe (1948) recomienda sustituir $\hat{\pi}^* = (y + \frac{3}{8})/(n + \frac{3}{4})$ en lugar de $\hat{\pi} = y/n$ en la transformación. Si todas las proporciones observadas en el estudio están entre $\hat{\pi} = 0.3$ y $\hat{\pi} = 0.7$, la varianza binomial tiene una estabilidad relativa y tal vez no sea necesario transformarla.

Probits y logits

Otras dos transformaciones relacionadas con la distribución binomial se usan con mucha frecuencia en estudios biológicos. La transformación *probit* es el valor de la distribución normal estándar que corresponde a la probabilidad acumulada $\hat{\pi} = y/n$. La transformación *logit* es el logaritmo natural de la razón $\hat{\pi}/(1 - \hat{\pi})$, usada en estudios biológicos y en el análisis de datos de supervivencia. Aunque ambas transformaciones constituyen un procedimiento estadístico adecuado para sus propósitos, las varianzas no se estabilizan y deben usarse otros modelos para el análisis. La sección 4.7 contiene una presentación breve de uno de ellos, los detalles para utilizar estas transformaciones se encuentran en Cox (1970), Finney (1978), McCullagh y Nelder (1989) y Collet (1991).

Binomial negativa

Si los incrementos en el número de muestras individuales se pueden relacionar con el número de individuos presentes, la distribución Poisson no puede aplicarse al problema, ya que los individuos contados tienden a presentarse en grupos con una distribución *contagiosa*. Animales infectados con el mismo organismo, plantas o insectos de la misma especie, con frecuencia forman grupos a causa de mecanismos biológicos de reproducción o transmisión de enfermedades. Una distribución de la probabilidad que se usa con frecuencia para datos de esta clase, es la *binomial negativa* con media μ_y y varianza $\sigma_y^2 = \mu_y + \lambda^2\mu_y^2$; donde la varianza aumenta con la media a una tasa mayor que en la distribución Poisson. Una transformación sugerida para estabilizar la varianza es el seno hiperbólico inverso $x = \lambda^{-1}\text{senh}^{-1}\sqrt{y}$; entonces la varianza de las observaciones transformadas es $\sigma_x^2 = 0.25$. La transformación requiere cierto conocimiento de λ . Una transformación $x = \log(y + 1)$, tiene una relación lineal aproximada con la transformación senh^{-1} (Bartlett, 1947).

4.6 Transformaciones con exponentes para estabilizar varianzas

En ocasiones no se puede determinar la distribución de las observaciones con base en las propiedades de la muestra de la variable aleatoria. En estas circunstancias, es posible determinar la transformación según una relación empírica entre la desviación estándar y la media.

Transformación de datos empíricos

Una transformación del exponente altera la simetría o asimetría de la distribución de frecuencia de las observaciones. Las transformaciones se basan en el trabajo de Box y Cox (1964), en el cual se supone que las desviaciones estándar de y son proporcionales a alguna potencia de la media, o sea:

$$\sigma_y \propto \mu^\beta \quad (4.6)$$

Una transformación del exponente de las observaciones:

$$x = y^p \quad (4.7)$$

tiene como resultado una relación proporcional de la desviación estándar con la media:

$$\sigma_x \propto \mu^{p+\beta-1} \quad (4.8)$$

Si $p = 1 - \beta$, entonces la desviación estándar de la variable transformada x será constante, ya que $p + \beta - 1 = 0$ y $\sigma_x \propto \mu^0$ en la ecuación (4.8).

A menudo, las transformaciones se representan como una *escalera de exponentes*, frase que se originó en el análisis exploratorio de datos (Tukey, 1977; Velleman & Hoaglin, 1981). El cuadro 4.2 muestra el orden de la escalera para algunas transformaciones útiles.

Cuadro 4.2 Transformaciones en la escalera de exponentes, $x = y^p$

<i>p</i>	y^p	Nombre	Observaciones
2	y^2	Cuadrada	La más usada
1	y^1	Datos originales	No hay transformación
$\frac{1}{2}$	\sqrt{y}	Raíz cuadrada	Distribución de Poisson
0	$\log(y)$	Logaritmo	“0” en la escalera
$-\frac{1}{2}$	$1/\sqrt{y}$	Raíz cuadrada recíproca	El signo menos preserva el orden de las observaciones
-1	$1/y$	Recíproco	Reexpresa el tiempo como tasa

Los valores de p inferiores a 1 acortarán la cola superior de las observaciones y alargarán la cola izquierda de las observaciones apiladas en una distribución sesgada a la derecha. De manera inversa, si p es mayor que 1, una distribución sesgada a la izquierda cambia a una más simétrica al mover las observaciones que están en el lado izquierdo. La transformación logaritmo se coloca en la posición "0" de la escalera porque su efecto sobre las observaciones cae de manera natural en esa posición.

La transformación recíproca con $p = -1$ puede ser útil en los estudios que necesitan medir el tiempo de ocurrencia de un evento. El recíproco del tiempo se interpreta en términos generales como la tasa a la que un sujeto de investigación llega al evento. Aunque es tentador asignar un valor de 0 a un sujeto para el que nunca ocurre el evento, debe tenerse cuidado, puesto que el evento nunca se observa. La observación puede tratarse mejor como miembro de un conjunto de observaciones truncado, o bien como observaciones faltantes, según las circunstancias.

Estimación empírica para la transformación de exponentes

Es posible determinar una estimación empírica de p si se dispone de estimaciones para la media y la desviación estándar de los grupos de tratamiento. Se expresa la relación entre la desviación estándar y la media del i -ésimo grupo como $\sigma_i = \alpha\mu_i^\beta$, donde α es una constante de proporcionalidad. Se toma el logaritmo de la expresión para obtener:

$$\log(\sigma_i) = \log(\alpha) + \beta \log(\mu_i) \quad (4.9)$$

Una gráfica de $\log(\sigma_i)$ contra $\log(\mu_i)$ es una línea recta con ordenada $\log(\alpha)$ y pendiente β . Las estimaciones de las medias y desviaciones estándar se sustituyen por σ_i y μ_i , y la estimación de β se obtiene de un análisis de regresión lineal simple. El valor de p para una transformación estabilizadora de la varianza puede ser $\hat{p} = 1 - \hat{\beta}$, donde $\hat{\beta}$ es la pendiente estimada para la ecuación (4.9). La determinación de una transformación de exponentes empírica y sus efectos se ilustran con los datos para cangrejos ermitaños del ejemplo 4.1.

En la figura 4.4 se muestra una gráfica de las desviaciones estándar de los seis sitios contra los logaritmos de las seis medias, dados en la tabla 4.1. La estimación de β a partir de los resultados de la regresión mostrada en la gráfica es $\hat{\beta} = 0.99$; el valor empírico estimado de p es $\hat{p} = 1 - 0.99 = 0.01$. Este valor es muy cercano a la posición cero de la escalera de exponentes, lo que implica una transformación tipo logaritmo para los datos del estudio con cangrejos ermitaños.

Cuando hay algunas muestras de cero entre las observaciones, se agrega una pequeña constante c , tal que $0 < c \leq 1$ a la observación y para evitar la evaluación de un logaritmo de 0. Con frecuencia se usan 1 y $\frac{1}{2}$ para c , pero Mosteller y Tukey (1977) sugieren un valor de $c = \frac{1}{6}$.

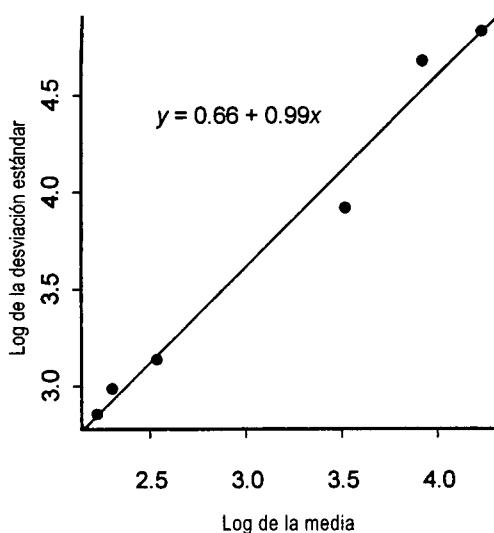


Figura 4.4 Gráfica y estimaciones de regresión de $\log(s_i)$ contra $\log(\bar{y}_{i.})$ para los seis sitios de datos de los cangrejos ermitaños

Los datos de cangrejos ermitaños se transformaron a $x = \log(y + \frac{1}{6})$ porque había observaciones de cero cangrejos en algunas secciones. El resumen estadístico para los datos transformados se muestra en la tabla 4.4; la gráfica de probabilidad normal de los residuales aparece en la figura 4.5; la gráfica de residuales contra las medias de tratamiento estimadas se muestra en la figura 4.6, y la de dispersión s-1 en la figura 4.7.

Tabla 4.4 Medias, medianas, desviaciones estándar, valores mínimos y máximos, y valores residuales estandarizados máximos, w_{ij} , para cada sitio, después de la transformación $x = \log(y + \frac{1}{6})$, para los datos de cangrejos ermitaños

<i>Sitio</i>	<i>Media</i>	<i>Mediana</i>	<i>Desviación estándar</i>	<i>Mínimo</i>	<i>Máximo</i>	w_{ij}
1	0.94	1.24	0.99	-0.78	2.37	2.51
2	1.00	1.01	1.06	-0.78	2.67	2.83
3	0.76	0.71	1.05	-0.78	2.61	2.77
4	0.39	0.34	0.81	-0.78	1.92	2.03
5	0.44	0.34	0.76	-0.78	1.97	2.09
6	0.36	0.62	0.96	-0.78	1.98	2.10

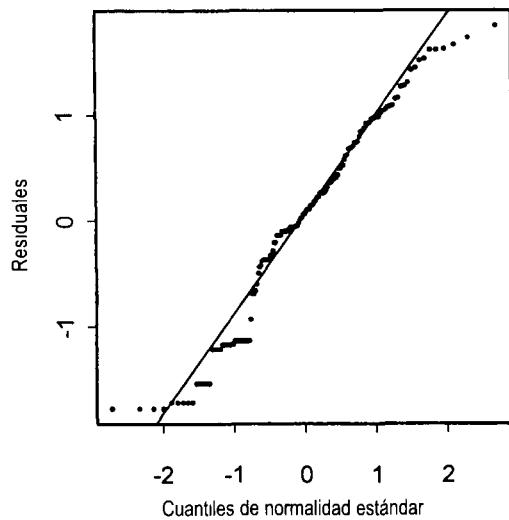


Figura 4.5 Gráfica de probabilidad normal de los residuales para los datos de los cangrejos ermitaños, luego de la transformación $x = \log(y + \frac{1}{6})$

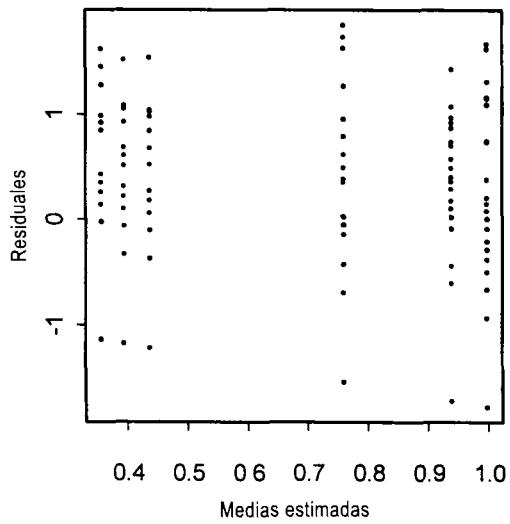


Figura 4.6 Gráfica de los residuales contra las medias de tratamiento estimadas para los datos de los cangrejos ermitaños, después de la transformación $x = \log(y + \frac{1}{6})$

Como todos los valores máximos de los residuales estandarizados son menores a 3 desviaciones estándar de la media (tabla 4.4), no hay evidencia de residuos inusitados luego de la transformación.

La dispersión de los residuales presentada en la figura 4.6 es muy similar para cada uno de los sitios, y la gráfica de dispersión-localización de la figura 4.7 muestra un incremento o decremento no monótono de $\sqrt{|e_{ij}|}$, indicando con ello varianzas relativamente homogéneas. Las varianzas mínima y máxima para los sitios 2 y 5 son $s_2^2 = 1.06^2 = 1.1236$ y $s_5^2 = 0.76^2 = 0.5776$, respectivamente (tabla 4.4). El estadístico calculado para la prueba de Levene fue $CMT = 1.44$ y $CME = 1.91$ con $F_0 = 0.75$; la hipótesis nula de varianzas homogéneas con un valor crítico de $F_{.05,5,144} = 2.28$ es aceptada. Las suposiciones necesarias para el análisis mediante el modelo lineal aparentemente concuerdan bastante bien con las observaciones después de la transformación.

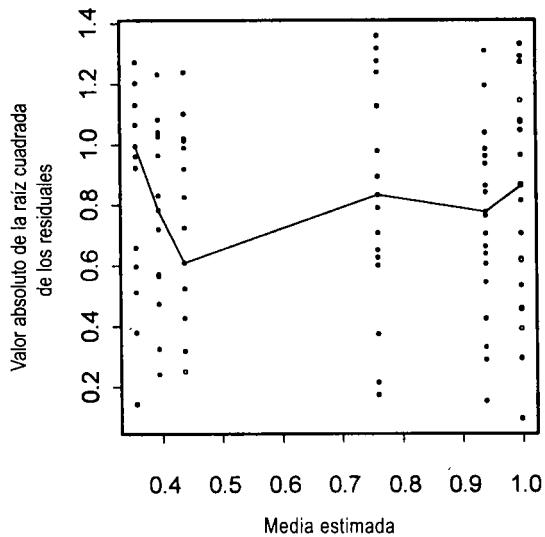


Figura 4.7 Gráfica dispersión s-l de los residuales para los datos de los cangrejos ermitaños, luego de la transformación $x = \log(y + \frac{1}{6})$

Transformación de exponentes de Box-Cox para otros diseños

El método de regresión usado en esta sección para estimar p en la transformación de exponentes $x = y^p$ es eficaz solamente en el diseño totalmente aleatorizado, en el que es posible estimar las medias y desviaciones estándar de los grupos. La estimación de p en los diseños de experimentos más complejos con bloques, estudiados en capítulos posteriores, requiere de un enfoque más riguroso.

La transformación original de Box-Cox es $x = (y^p - 1)/p$ para $p \neq 1$ y $x = \log_e y$ para $p = 0$ (Box y Cox, 1964). La estimación de p se puede calcular al maximizar:

$$L(p) = -\frac{1}{2} \log_e [CME(p)]$$

donde $CME(p)$ es el cuadrado medio del error calculado para el análisis de varianza, usando la transformación $x = (y^p - 1)/p$ para la elección dada de p .

Se puede obtener una solución si se determina $CME(p)$ para un conjunto de valores elegidos de p , a partir de la gráfica de $CME(p)$ contra p , leyendo el valor correspondiente al mínimo de $CME(p)$. Con este método es poco probable que se use el valor exacto de p , es más común que se use la transformación estándar mostrada en el cuadro 4.2, que aproxima los valores estimados de p .

4.7 Generalización del modelo lineal

El modelo lineal usado en este libro supone que los errores experimentales tienen varianza homogénea o constante y que la distribución normal es una buena aproximación a su distribución. Muchas veces, en los estudios se encuentra que un fenómeno natural se comporta de manera lineal dentro del intervalo de interés y tiene errores homogéneos con distribución normal. Si no, es posible que alguna transformación adecuada de la variable de respuesta proporcione la linearidad y la estructura de error necesarias. La posibilidad de usar con tanta frecuencia el modelo lineal y el análisis de varianza los ha hecho populares en las investigaciones científicas. Sin embargo, las transformaciones pueden conducir a escalas no satisfactorias para la interpretación (por ejemplo, el arco seno o sen^{-1} de la raíz cuadrada).

El *modelo lineal generalizado* es una clase genérica de modelos lineales, introducida por Nelder y Wedderburn (1971), que amplía la variedad disponible de modelos probabilísticos para errores experimentales y formas no lineales en el modelo. El modelo lineal usado en este libro, con errores homogéneos y distribución normal, es un subconjunto de la forma generalizada.

El modelo lineal generalizado introduce funciones separadas para permitir varianzas heterogéneas y no linearidad. En lugar de transformar la variable de respuesta, este modelo puede describirse mejor como que reexpresa el modelo. Dobson (1990) proporciona una introducción breve al uso y aplicación de este modelo, y McCullagh y Nelder presentan una cobertura completa de los modelos.

Teniendo en cuenta el modelo lineal, introducido en la sección 2.4, para la respuesta observada y que es una conversión de la variable aleatoria Y con esperanza $E(Y) = \mu$, donde $\mu = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$ es lineal con respecto a los parámetros β . Las x representan variables controladas en el estudio o medidas como covariadas de la respuesta y . Las suposiciones necesarias son una σ^2 constante para Y y la relación lineal entre μ y x_i . Sin embargo, para muchos tipos de datos, un cambio en la media de Y introduce un cambio en la varianza. Los ejemplos incluyen la respuesta binaria (0 o 1) con probabilidad de éxito π para la distribución binomial,

donde $E(Y) = \pi$ y $\sigma^2(Y) = \pi(1 - \pi)/n$, y los datos para la distribución Poisson, en donde la media es igual a la varianza.

El modelo lineal generalizado maneja estos aspectos de manera natural introduciendo una reparametrización que acepta una varianza heterogénea mediante una función de enlace $\eta = g(\mu)$, tal que $\eta = g(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$ es lineal respecto a los parámetros β en lugar de respecto a μ . Por ejemplo, la función de enlace para las respuestas binarias con $\mu = \pi$ es el enlace *logit*:

$$\eta = g(\mu) = \log\left(\frac{\mu}{1 - \mu}\right)$$

de manera que:

$$\mu = \frac{e^\eta}{1 + e^\eta}$$

El dominio de $\mu = \pi$ aún se encuentra dentro del intervalo $[0, 1]$, pero $\eta = g(\mu)$ puede tomar cualquier valor real. La función de enlace logit es la función lineal respecto a los parámetros β y no a $\mu = \pi$.

La función de enlace natural para la distribución Poisson es el enlace *log*, con $\eta = \log(\mu)$ de manera que $\log(\mu)$ tiene una relación lineal con x_i . *Probit* es otro enlace para datos binarios, ampliamente utilizado en estudios biológicos. El enlace *identidad* $\eta = \mu$ no es una reexpresión y tiene como resultado el modelo lineal usual, con varianzas homogéneas y errores con distribución normal.

La disponibilidad actual de software ha popularizado el uso de enlaces específicos, asociados con el modelo lineal generalizado, como la regresión logística con el enlace logit y los modelos lineales log o la regresión Poisson con el enlace log. Collett (1991) presenta en forma exhaustiva el uso y aplicación de enlaces para datos binarios.

La estimación y el análisis con modelos generalizados binarios se basa en métodos de estimación de la máxima posibilidad. El estudio de los modelos de estimación y otros aspectos asociados con los modelos generalizados está fuera del alcance y la intención de este libro. No importa qué modelo se use para la inferencia estadística, un buen diseño estadístico es fundamental para que sea válida.

4.8 Evaluación de modelos por medio de gráficas de residuales vs. valores ajustados

Un método gráfico, la gráfica de *dispersión de residuales vs. valores ajustados* o gráfica *r-f (residual-fitted)*, ayuda a evaluar si los modelos lineales planteados en la hipótesis se ajustan bien a los datos. Este método se ha incluido aquí como una adición al estudio de la evaluación de homogeneidad y normalidad del error experimental a través de residuales de suposiciones. Cleveland (1993) presenta otros métodos gráficos útiles para explorar y confirmar el análisis de datos.

La gráfica *r-f* proporciona una descripción gráfica de la variación relativa o *dispersión* en los errores experimentales y en los *valores ajustados*³ del modelo

lineal. En el caso de un diseño totalmente aleatorizado, los valores ajustados son las medias estimadas de los grupos de tratamiento. La gráfica de dispersión $r-f$ es un complemento visual de la razón $(SCE_r - SCE_f)/SCE_f$, presentada en la sección 2.10 como preludio a la prueba F para las diferencias entre las medias de grupos de tratamiento. Esta razón proporciona un medio para evaluar la mejora relativa del modelo completo sobre el modelo reducido de los datos, donde $(SCE_r - SCE_f)$ refleja la variación atribuible a las componentes estimadas o ajustadas en el modelo, y SCE_f la variación en el error experimental, o variación residual, después de ajustar el modelo a los datos.

La gráfica de dispersión $r-f$ incluye una gráfica de residuales ordenados contra su frecuencia acumulada y otra de los valores ajustados ordenados *menos su media* contra su frecuencia acumulada. La frecuencia acumulada, o *valor f*, es la misma, $f_i = (i - 0.5)/N$, $i = 1, 2, \dots, N$, usada para obtener los cuantiles normales para las gráficas de probabilidad normal contempladas en la sección 4.3.

La razón calculada para el experimento de almacenamiento de carne del capítulo 2, con $SCE_r = 33.7996$ y $SCE_f = 0.9268$ es $(33.7996 - 0.9268)/0.9268 = 35.47$, lo cual indica que la suma de cuadrados para los valores ajustados de las medias de grupos de tratamiento fue 35.47 veces mayor que la suma de cuadrados de los residuales. La gráfica de dispersión $r-f$ para ese experimento se muestra en la figura 4.8.

La mayor dispersión de los valores ajustados con respecto a la dispersión de los residuales (figura 4.8) refleja un valor más grande de la razón de sus respectivas sumas de cuadrados. Es común que la significancia estadística no indique verdaderamente si los tratamientos producen diferencias físicas o biológicas significativas. La gráfica de dispersión $r-f$ ayuda en la evaluación visual de las diferencias físicas entre los grupos de tratamiento respecto a la variabilidad residual restante, que se puede usar como complemento a la prueba F formal en el análisis de varianza. Dada la dispersión diferencial de los valores ajustados y residuales en el experimento de carnes, se puede concluir que, en cuanto a los errores experimentales, algunos tratamientos produjeron diferencias biológicas significativas entre sus respectivas medias.

Al usar la gráfica de dispersión $r-f$, los datos transformados de los cangrejos ermitaños tienen diferencias notables con el experimento de carne. En la tabla 4.5 aparece el análisis de varianza para los datos transformados. Una prueba F rechaza la hipótesis de medias iguales para los sitios con $F_0 = 2.06/0.89 = 2.31$ y un valor P de .046. La razón $(SCE_r - SCE_f)/SCE_f = 10.32/127.99 = 0.08$ indica una pequeña variación entre las medias de tratamiento ajustadas con respecto a los residuales. Esta conclusión se confirma de manera visual con la gráfica de dispersión $r-f$ en la figura 4.9. Aunque la prueba F juzgue que las medias de los sitios son diferentes, es bastante evidente que la variación residual dentro de cada sitio es considerable con respecto a las diferencias entre los promedios de las muestras transformadas de los sitios. Así, la significancia de las diferencias biológicas entre los sitios puede ser despreciable.

³ Valor ajustado viene del concepto de “ajustar” un modelo a los datos y descomponer la observación y_{ij} en dos partes, $y_{ij} = \hat{\mu}_i + \hat{e}_{ij}$ en el diseño totalmente aleatorizado, donde el valor ajustado es $\hat{\mu}_i$, y el residual \hat{e}_{ij}

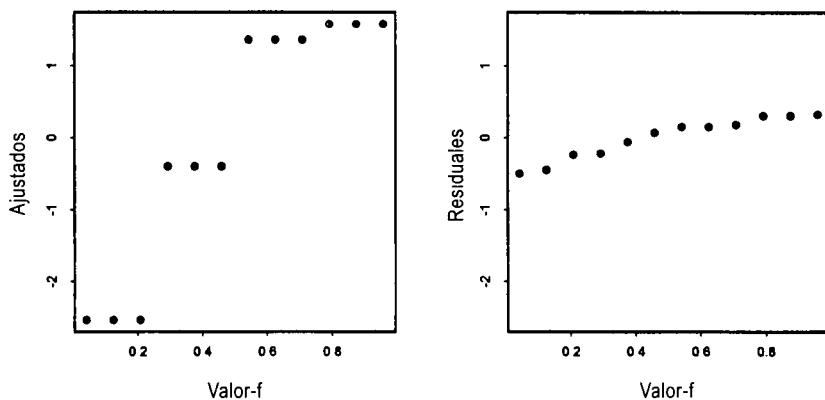


Figura 4.8 Gráfica de dispersión ajustada por residuales para comparar la dispersión de los residuales y los valores ajustados menos sus medias, en el experimento de almacenamiento de carne del capítulo 2

Tabla 4.5 Análisis de varianza para los datos de los cangrejos ermitaños, trasnformados con $x = \log(y + \frac{1}{6})$

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Total	149	138.31	
Sitios	5	10.32	2.06
Error	144	127.99	0.89

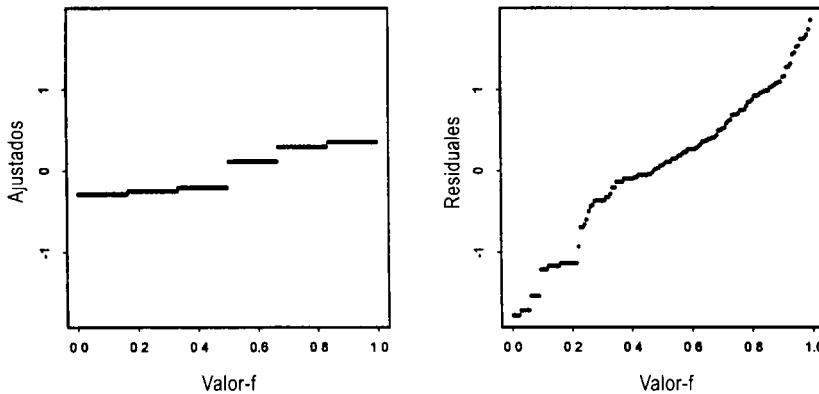


Figura 4.9 Gráfica de dispersión $r-f$ para comparar la dispersión de los residuales y los valores ajustados menos sus medias después de la transformación, $x = \log(y + \frac{1}{6})$, a los datos de los cangrejos ermitaños

EJERCICIOS

1. Se realizó una prueba de la vida útil, a temperatura acelerada, de un tipo de calentador tubular. Se probaron seis calentadores, cada uno a cuatro temperaturas distintas: 1,520°F, 1,620°F, 1,660°F y 1,708°F. Se registró el número de horas transcurridas hasta que se presentó falla en los 24 calentadores utilizados en el estudio.

<i>Temperatura de prueba</i>	<i>Horas hasta la falla</i>
1,520°	1,953, 2,135, 2,471, 4,727, 6,134, 6,314
1,620°	1,190, 1,286, 1,550, 2,125, 2,557, 2,845
1,660°	651, 837, 848, 1,038, 1,361, 1,543
1,708°	511, 651, 651, 652, 688, 729

Fuente: W. Nelson (1972), A short life test for comparing a sample with previous accelerated test results, *Technometrics* 14, 175-185.

- Investigue las suposiciones necesarias para un análisis de varianza de los datos. Incluya una gráfica de probabilidad normal de los residuales y una de dispersión-localización (s-l).
- Determine una transformación razonable de la escalera de exponentes, utilizando la pendiente de la regresión basada en los logaritmos de las desviaciones estándar y las medias de los grupos.
- Determine si la transformación elegida tiene como resultado datos que concuerdan en forma razonable con las suposiciones necesarias para el análisis de varianza.
- Realice un análisis de varianza de los datos transformados, y haga una partición de la suma de los cuadrados de la temperatura en contrastes polinomiales ortogonales, para determinar la mejor relación entre la temperatura y su variable de respuesta. Como las temperaturas de prueba tenían espaciamientos desiguales, use los siguientes coeficientes de contraste:

<i>Temperatura</i>	1,520	1,620	1,660	1,708
Lineal	-0.773	-0.051	0.238	0.585
Cuadrática	0.382	-0.637	-0.328	0.583
Cúbica	-0.078	0.584	-0.765	0.259

2. Un entomólogo contó el número de huevos que pone cada una de las 15 hembras de polillas en días sucesivos, en tres variedades de gusano de tabaco (USDA, campo y resistente). Los siguientes datos son el número de huevos puestos en el tercer día después del apareamiento de cada hembra en cada variedad.

<i>Variedad</i>	<i>Número de huevos por polilla</i>
USDA	448, 906, 28, 227, 634, 48, 369, 137, 29, 522, 319, 242, 261, 566, 734
Campo	211, 276, 415, 787, 18, 118, 1, 151, 0, 253, 61, 0, 275, 0, 153
Resistente	0, 9, 143, 1, 26, 127, 161, 294, 0, 348, 0, 14, 21, 0, 218

Fuente. Dr. T Watson y S. Kelly, Department of Entomology, University of Arizona.

- a. El entomólogo desea realizar un análisis de varianza del número de huevos. ¿Qué distribución de probabilidad es adecuada para los datos?
 - b. ¿Cuál es la transformación sugerida para la distribución de probabilidad seleccionada en el inciso anterior?
 - c. Determine una transformación razonable de la escalera de exponentes, utilizando la pendiente de la recta de regresión basada en los logaritmos de las desviaciones estándar y las medias de los grupos. ¿Concuerda esta transformación con la sugerencia del inciso b?
 - d. Transforme los datos con la transformación elegida y determine si los datos transformados concuerdan con las suposiciones del análisis de varianza.
3. Un criador de plantas evaluó la capacidad de enraizar de nueve clones de pasto bermuda en un experimento de laboratorio. Cultivó dos réplicas de cada clon en una solución oxigenada en un diseño totalmente aleatorizado. La tabla presenta el número de nodos enraizados en los cultivos de cada clon.

Clon	Réplica 1		Réplica 2	
	Enraizado	No enraizado	Enraizado	No enraizado
1	15	49	11	53
2	13	51	11	53
3	13	51	6	58
4	6	42	4	60
5	16	48	12	52
6	14	50	9	55
7	8	56	18	46
8	9	55	10	54
9	8	40	16	48

Fuente: Dr. W Kneebone, Department of Plant Sciences, University of Arizona

- a. El cultivador quiere analizar la proporción de cultivos enraizados o la proporción de nodos enraizados. ¿Qué distribución de probabilidad es adecuada para los datos?
 - b. ¿Cuál es la transformación sugerida para la distribución de probabilidad seleccionada en el inciso a?
 - c. Transforme los datos de la proporción de cultivos (o nodos) enraizados con la transformación adecuada y realice un análisis de varianza con los datos transformados.
 - d. Utilice el mejor procedimiento de comparaciones múltiples para seleccionar un subconjunto de clones con las medias más grandes y $P(EC) = 0.95$.
4. El estudio de Ames sobre la *salmonela*/microsoma se usa para investigar el potencial de sustancias contaminantes, según su capacidad para efectuar cambios en material genético. Se probó el compuesto 4-nitroorto-penilenidiamina (4NoP) con la variedad *salmonela* TA98. Se contó el número de colonias visibles en placas dosificadas con 4NoP. Se usaron cinco niveles de dosis de 4NoP en el estudio. Se muestran las colonias contadas en siete placas de cada nivel.

Dosis ($\mu\text{g/placa}$)	Colonias encontradas
0.0	11, 14, 15, 17, 18, 21, 25
0.3	39, 43, 46, 50, 52, 61, 67
1.0	88, 92, 104, 113, 119, 120, 130
3.0	222, 251, 259, 283, 299, 312, 337
10.0	562, 604, 689, 702, 710, 739, 786

Fuente: B. H. Margolin, B. S. Kim y K. J. Risko (1989), The Ames *Salmonella*/microsome mutagenicity assay: Issues of inference and validation. *Journal of American Statistical Association* 84, 651-661.

- Como los datos incluyen cuentas de colonias de bacterias, ¿puede suponerse con seguridad que los datos tienen una distribución Poisson? Explique su respuesta.
 - Los autores del artículo citado sugieren la distribución binomial negativa como una distribución factible. ¿Está de acuerdo con esta conclusión? Explique su respuesta.
 - Determine una transformación de los datos tal que al transformarlos satisfagan de manera suficiente las suposiciones del análisis de varianza. Realice un análisis de varianza para los datos transformados.
5. Dada la siguiente muestra aleatoria de $N = 15$ observaciones, ordenadas de menor a mayor:
- 14.3, 16.0, 17.3, 17.5, 17.8, 18.7, 18.8, 18.9, 20.0, 20.8, 21.4, 22.7, 23.2, 25.6, 27.8
- Determine los valores f y sus cuantiles normal estándar.
 - Graifique las observaciones contra los cuantiles normal estándar.
 - Interprete la gráfica respecto a la forma de la distribución a partir de la cual se muestrearon las observaciones.
6. Dada la siguiente muestra aleatoria de $N = 16$ observaciones, ordenadas de menor a mayor:
- 2, 3, 4, 5, 10, 28, 34, 35, 39, 63, 87, 97, 112, 156, 188, 253
- Determine los valores f y sus cuantiles normal estándar.
 - Graifique las observaciones contra los cuantiles normal estándar.
 - Interprete la gráfica respecto a la forma de la distribución a partir de la cual se muestrearon las observaciones.

Apéndice 4A: Datos para el ejemplo 4.1

Muestras de cangrejo ermitaño en sitios costeros. Un biólogo marino contó los cangrejos ermitaños en 25 secciones de seis sitios costeros diferentes. Se proporciona el número de cangrejos contados en cada sección.

Sitio	Conteo
1	0, 0, 22, 3, 17, 0, 0, 7, 11, 11, 73, 33, 0, 65, 13, 44, 20, 27, 48, 104, 233, 81, 22, 9, 2
2	415, 463, 6, 14, 12, 0, 3, 1, 16, 55, 142, 10, 2, 145, 6, 4, 5, 124, 24, 204, 0, 0, 56, 0, 8
3	0, 0, 4, 13, 5, 1, 1, 4, 4, 36, 407, 0, 0, 18, 4, 14, 0, 24, 52, 314, 245, 107, 5, 6, 2
4	0, 0, 0, 4, 2, 2, 5, 4, 2, 1, 0, 12, 1, 30, 0, 3, 28, 2, 21, 8, 82, 12, 10, 2, 0
5	0, 1, 1, 2, 2, 1, 2, 29, 2, 2, 0, 13, 0, 19, 1, 3, 26, 30, 5, 4, 94, 1, 9, 3, 0
6	0, 0, 0, 2, 3, 0, 0, 4, 0, 5, 4, 22, 0, 64, 4, 4, 43, 3, 16, 19, 95, 6, 22, 0, 0

5 Experimentos para estudiar las varianzas

En este capítulo se presenta el modelo estadístico para los estudios de investigación sobre las varianzas de poblaciones. El conocimiento sobre las causas asignables de variación es útil para mejorar los procesos de manufactura, mejorar la genética de cultivos o ganado, mejorar el control de calidad en estudios de salud y diseñar estudios de investigación. El objetivo es descomponer la varianza total en componentes identificables.

5.1 Modelos con efectos aleatorios para las varianzas

El experimento de almacenamiento de carne mencionado en los capítulos 2 y 3 incluye cuatro tratamientos específicos y no tiene un interés declarado en otros empaques. Así, toda la población de tratamientos consiste de cuatro métodos de empaque.

Es posible duplicar cada uno de ellos si se repite el experimento. En estas circunstancias, los modelos estadísticos usados para los estudios se conocen como modelos de **efectos fijos**, y las inferencias se restringen al conjunto particular de tratamientos estudiados.

Existen otros tipos de estudios de investigación en los que se desea identificar las fuentes primordiales de variación en un sistema y estimar sus varianzas. Por la naturaleza de 1) los objetivos de investigación, 2) la estructura del tratamiento, 3) los protocolos experimentales y 4) el tipo de inferencias hechas a partir de los resultados observados, los efectos en el modelo se consideran efectos aleatorios y los modelos estadísticos se conocen como modelos de **efectos aleatorios**. El siguiente ejemplo ilustra un sistema cuyo conocimiento de la variabilidad en componentes identificables de un proceso industrial se puede usar para mejorar un proceso.

Ejemplo 5.1 Moldeo de aleaciones a alta temperatura

Una aleación de metal se manufactura en un proceso de moldeado a alta temperatura; cada molde se descompone en barras individuales más pequeñas que se usan en aplicaciones que requieren menores cantidades de la aleación. La fuerza de tensión de la aleación es crítica para su uso futuro, por lo que el proceso de moldeo está diseñado para producir barras con una fuerza de tensión superior en promedio a las especificaciones mínimas. Es aceptable alguna variación en esta fuerza sólo cuando un pequeño número de las barras no cumplen con dichas especificaciones (figura 5.1a). Sin embargo, una variación excesiva tendrá como resultado que una cantidad inaceptable de las barras no las cumplan (figura 5.1b).

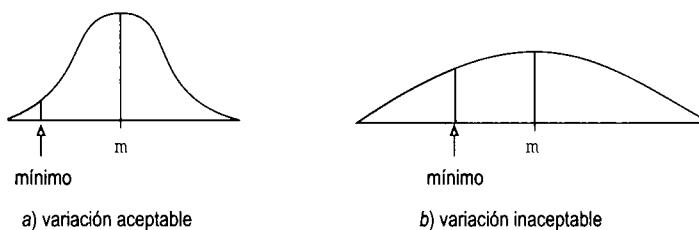


Figura 5.1 Variación a) aceptable y b) inaceptable en la fuerza de tensión

Dos componentes contribuyen a la variación total de la fuerza de tensión de las barras fabricadas: la variabilidad entre los moldes de fabricación y las alteraciones en el proceso de moldeo que afectan a las barras del mismo molde. Para mantener el control sobre la variación se requiere conocer las contribuciones a la variabilidad de cada parte del proceso.

Se planeó un experimento para aislar la variación en la fuerza de tensión debida a los efectos de los distintos moldes de la atribuible a las alteraciones dentro del mismo molde.

Se tomaron los moldes a alta temperatura de la aleación de tres corridas de producción elegidas al azar en la misma fábrica; cada molde se descompuso en barras individuales y se obtuvieron las mediciones destructivas de la fuerza de tensión en una muestra aleatoria de 10 barras de cada uno de los tres moldes. En la tabla 5.1 aparecen, en libras por pulgada cuadrada (psi), los datos de la fuerza de tensión para las 30 barras.

Los tres moldes usados en el estudio representan una muestra de la población potencial de moldes que pueden producirse en esa fábrica; los investigadores estaban interesados en la variación de la fuerza de tensión entre los moldes producidos en ella; la preocupación no está en los tres moldes específicos del experimento. Por ello, consideraron los moldes sólo como una muestra aleatoria de los producidos en la instalación. Los efectos de los moldes serán efectos aleatorios ya que se eligieron al azar de una población potencialmente infinita. Las inferencias se extenderán a la población de moldes que pudieran producirse en esa fábrica. De la

Tabla 5.1 Fuerza de tensión (psi) de las barras de tres moldes distintos de una aleación a alta temperatura

	<i>Molde</i>		
	<i>1</i>	<i>2</i>	<i>3</i>
88.0	85.9	94.2	
88.0	88.6	91.5	
94.8	90.0	92.0	
90.0	87.1	96.5	
93.0	85.6	95.6	
89.0	86.0	93.8	
86.0	91.0	92.5	
92.9	89.6	93.2	
89.0	93.0	96.2	
93.0	87.5	92.5	

Fuente. G. J. Hahn y T. E. Raghunathan (1988), "Combining information from various sources: A prediction problem and other industrial applications", *Technometrics* 30, 41-52.

misma manera, cada barra es una muestra aleatoria de las posibles barras obtenidas de un solo molde, y sus efectos sobre la fuerza de tensión son efectos aleatorios.

La fuerza de tensión observada de una barra en particular (y) difiere de la media del proceso (μ) en un error global $\delta = y - \mu$. Las componentes del error global se ilustran en la figura 5.2. El error global es la suma de dos componentes, $\delta = \delta_c + \delta_b$, donde δ_c es la componente del error para moldes y δ_b es la componente del error de las barras en un molde.

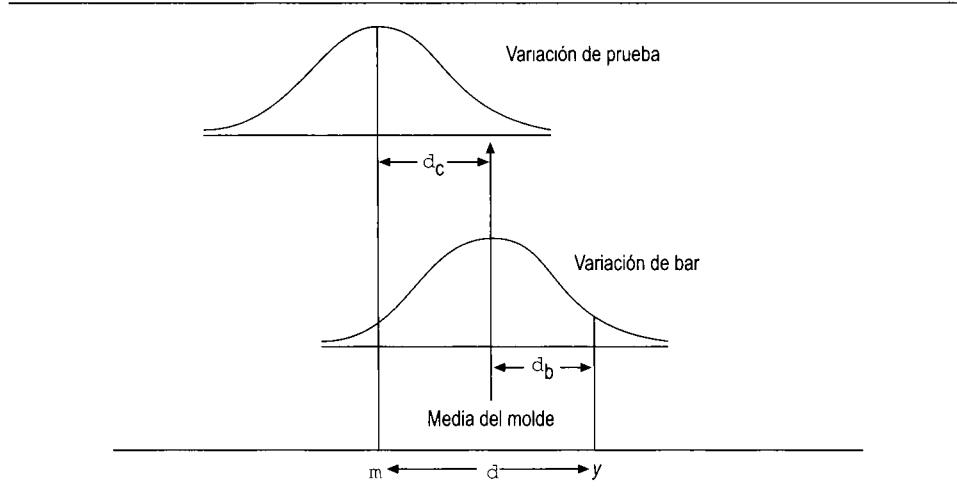


Figura 5.2 Dos componentes del error en el proceso de moldeo de metales

Otros ejemplos: otro estudio típico con efectos aleatorios involucra la herencia de características cuantitativas, como la cantidad de grano cosechada en diversas especies de plantas. En los programas de cultivo de plantas se desarrollan muchas familias genéticas distintas, que representan una muestra aleatoria de las familias que puede desarrollar un agricultor. Los descendientes de cada familia representan una muestra aleatoria de los posibles descendientes de la familia completa. El agricultor quiere descomponer la variación total en contribuciones separadas de las familias y los descendientes.

Los laboratorios clínicos participan por rutina en estudios de intercambio sobre la variabilidad de los resultados que requieren modelos de efectos aleatorios para el análisis estadístico. Regularmente, envían grandes muestras homogéneas de suero a un gran número de laboratorios para su análisis. Los laboratorios participantes y las muestras que reciben representan una muestra aleatoria de las potenciales poblaciones de laboratorios y muestras de suero. Los investigadores desean saber si existe una variación significativa en los resultados entre laboratorios.

5.2 Un modelo estadístico para las componentes de la varianza

Un modelo adecuado para identificar las fuentes de variación de los efectos aleatorios del experimento de moldeo con aleaciones a alta temperatura es:

$$y_{ij} = \mu + a_i + e_{ij} \quad (5.1)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde μ es la media del proceso, las a_i son los efectos aleatorios del moldeo y las e_{ij} son los efectos del error aleatorio para las barras dentro de cada molde. Se supone que los efectos e_{ij} y a_i son independientes entre sí.

También se supone que los efectos del error e_{ij} son una muestra aleatoria de una población con media 0 y varianza σ_e^2 . Y que los efectos aleatorios de los grupos (a_i) son una muestra aleatoria de una población con media 0 y varianza σ_a^2 . Si $\sigma_a^2 = 0$, entonces todos los efectos de los grupos son iguales, pero si $\sigma_a^2 > 0$, existe una variabilidad entre los efectos de los grupos. Como los efectos de grupo en este experimento son sólo una muestra de una población, las diferencias entre las medias de grupo específicas, $\mu + a_i$, no son de particular interés. La varianza de la distribución de efectos de grupo, σ_a^2 , es el centro de atención en cuanto a los efectos aleatorios.

La varianza de una observación, σ_y^2 , se puede expresar como la suma de las dos varianzas, es decir, $\sigma_y^2 = \sigma_a^2 + \sigma_e^2$, en donde σ_a^2 y σ_e^2 se denominan **componentes**, y el modelo de la ecuación (5.1) recibe el nombre de modelo de **componentes de la varianza**. En el estudio de cultivo de plantas, la componente de la varianza entre grupos (σ_a^2) representa la variación genética entre familias, y quizás el agricultor esté interesado en la relación de esta variación genética con la variación total (σ_y^2). Un ingeniero puede usar la estimación de σ_y^2 para calcular los valores

percentiles de la distribución de las fuerzas de tracción de las barras cuando se usen en una aplicación crítica.

En la tabla 5.2 se muestra una descripción del análisis de varianza para las observaciones, con los valores esperados de los cuadrados medios dentro y entre grupos. Para distinguir el modelo de efectos aleatorios del modelo de efectos fijos, se usarán los términos *dentro de grupos* y *entre grupos* en lugar de *tratamientos* y *error de las fuentes de variación*. Los cálculos de las sumas de cuadrados para dentro y entre grupos, son iguales que los de las sumas de cuadrados de tratamientos y error estudiados en el capítulo 2 para el modelo de efectos fijos.

Tabla 5.2 Análisis de varianza para la clasificación en un sentido, con los valores esperados de los cuadrados medios para el modelo de efectos aleatorios

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Esperado del cuadrado medio
Total	$N - 1$	SC total		
Entre grupos	$t - 1$	SCA	CMA	$\sigma_e^2 + r\sigma_a^2$
Dentro de grupos	$N - t$	SCW	CMW	σ_e^2

5.3 Estimaciones puntuales de las componentes de varianza

El método de análisis de varianza se usa para estimar las componentes de la misma. El análisis de varianza se calcula como si se tratara de un modelo de efectos fijos, y los cuadrados medios esperados se obtienen con la suposición del modelo de efectos aleatorios (tabla 5.2). Los cuadrados medios observados son estimaciones de los cuadrados medios esperados, es decir:

$$CMA = \widehat{\sigma}_e^2 + r\widehat{\sigma}_a^2 \quad (5.2)$$

y

$$CMW = \widehat{\sigma}_e^2$$

Los estimadores del análisis de varianza de las componentes de esta última se determinan resolviendo la ecuación (5.2) con dos incógnitas. Las soluciones son:

$$\widehat{\sigma}_e^2 = CMW \quad (5.3)$$

y

$$\widehat{\sigma}_a^2 = \frac{(CMA - CMW)}{r}$$

Los estimadores en la ecuación (5.3) son no sesgados y tienen la menor varianza de todos los estimadores, que son funciones cuadráticas de las observaciones y estimadores no sesgados de σ_e^2 y σ_a^2 .

Se supone que los efectos aleatorios del modelo tienen una distribución normal, y dada esta suposición se puede probar la significancia de las componentes de varianza entre grupos. Las hipótesis nula y alterna son $H_0: \sigma_a^2 = 0$ y $H_a: \sigma_a^2 > 0$, respectivamente. El estadístico de prueba es $F_0 = CMA/CMW$, y la hipótesis nula se rechaza al nivel de significancia α si $F_0 > F_{\alpha,(t-1),(N-t)}$.

En la tabla 5.2 se muestra el análisis de varianza de los datos de fuerza de tensión de la tabla 5.1. La estimación de las componentes de varianza para las barras dentro de los moldes es el cuadrado de la media dentro de grupos, o sea:

$$\hat{\sigma}_e^2 = CMW = 5.82$$

La estimación de la componente de varianza entre grupos es:

$$\hat{\sigma}_a^2 = \frac{(CMA - CMW)}{r} = \frac{(73.94 - 5.82)}{10} = 6.81$$

La varianza total estimada de una observación de la fuerza de tracción es $\hat{\sigma}_y^2 = \hat{\sigma}_a^2 + \hat{\sigma}_e^2 = 6.81 + 5.82 = 12.63$.

La razón F_0 para probar la hipótesis nula $H_0: \sigma_a^2 = 0$ es $F_0 = CMA/CMW = 73.94/5.82 = 12.71$. La hipótesis nula se rechaza con una probabilidad de exceder $F_0 = 12.71$ igual a .000 (tabla 5.3). La variación de los moldes contribuye de manera significativa a la variación en la fuerza de tensión de las aleaciones.

Tabla 5.3 Análisis de varianza para las fuerzas de tensión de las barras de tres moldes para aleación a alta temperatura

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	29	304.99			
Entre grupos	2	147.88	73.94	12.71	.000
Dentro de grupos	27	157.10	5.82		

5.4 Estimaciones de intervalos para las componentes de la varianza

Las estimaciones de los intervalos de confianza se pueden calcular para ambas componentes de la varianza. El estimador exacto del intervalo de confianza $100(1 - \alpha)\%$ para σ_e^2 es:

$$\frac{SCW}{A} < \sigma_e^2 < \frac{SCW}{B} \quad (5.4)$$

donde $A = \chi_{\alpha/2,(N-t)}^2$ y $B = \chi_{(1-\alpha/2),(N-t)}^2$. A y B son valores de la variable ji-cuadrada excedida con probabilidades $\alpha/2$ y $(1 - \alpha/2)$, respectivamente. Los valores de la ji-cuadrada se encuentran en la tabla III del apéndice.

Un intervalo con al menos $100(1 - 2\alpha)\%$ de confianza para σ_a^2 es:

$$\frac{SCA(1 - F_u/F_0)}{rC} < \sigma_a^2 < \frac{SCA(1 - F_l/F_0)}{rD} \quad (5.5)$$

donde $C = \chi_{\alpha/2,(t-1)}^2$, $D = \chi_{(1-\alpha/2),(t-1)}^2$ y $F_0 = CMA/CMW$ es el estadístico F_0 observado (Williams, 1962). Las cantidades $F_u = F_{\alpha/2,(t-1),(N-t)}$ y $F_l = F_{(1-\alpha/2),(t-1),(N-t)}$ son valores de la variable F excedida con probabilidades $\alpha/2$ y $(1 - \alpha/2)$, respectivamente¹.

Dados $SCW = 157.10$, $\chi_{05,27}^2 = 40.1$ y $\chi_{95,27}^2 = 16.2$, la estimación de un intervalo de confianza del 90% de σ_e^2 a partir de la ecuación (5.4) es:

$$\frac{157.10}{40.1} < \sigma_e^2 < \frac{157.10}{16.2}$$

$$3.92 < \sigma_e^2 < 9.70$$

Con $SCA = 147.88$, $r = 10$, $\chi_{025,2}^2 = 7.38$, $\chi_{975,2}^2 = 0.05$, $F_0 = 12.71$, $F_{025,2,27} = 4.24$ y $F_{975,2,27} = 0.025$, la estimación del intervalo de confianza del 90% de σ_a^2 , a partir de la ecuación (5.5) es:

$$147.88 \left[\frac{1 - \frac{4.24}{12.71}}{10(7.38)} \right] < \sigma_a^2 < 147.88 \left[\frac{1 - \frac{0.025}{12.71}}{10(0.05)} \right]$$

$$1.34 < \sigma_a^2 < 295.18$$

Las estimaciones de intervalos para las componentes de varianza serán bastante amplias cuando los cuadrados medios tengan pocos grados de libertad. Más grupos de moldeo proporcionarían una estimación más precisa de intervalo σ_a^2 .

Interpretación de las componentes de la varianza

La media de las fuerzas de tensión para el experimento era $\bar{y}_{..} = 90.9$, con una estimación del error estándar:

$$s_{\bar{y}_{..}} = \sqrt{\frac{MSA}{rt}} = \sqrt{\frac{73.94}{30}} = 1.57 \text{ psi}$$

La varianza estimada de una observación de la fuerza de tensión en las barras es $\widehat{\sigma}_y^2 = \widehat{\sigma}_a^2 + \widehat{\sigma}_e^2 = 12.63$, con una desviación estándar de $\widehat{\sigma}_y = \sqrt{12.63} = 3.55$ psi. Las estimaciones de las componentes de varianza aislaron las distintas fuentes de variación en el proceso de moldeo para barras de aleación: la varianza de los

¹ El valor de $F_{(1-\alpha)}$ no se puede leer directamente en la tabla IV del apéndice, pero es posible determinar su valor a partir de la relación $F_{v_1,v_2,(1-\alpha)} = 1/F_{v_2,v_1,\alpha}$

moldes es responsable del 54% de la variación, y la de las barras es responsable del 46% restante. El ingeniero puede reducir la desviación estándar $\widehat{\sigma}_y = 3.55$ psi al identificar y ajustar los factores que aumentan la variación en el proceso de moldeo. Es posible que la variación entre los moldes sea ocasionada por modificaciones en las mezclas de aleación o por variaciones en la temperatura de un molde a otro, mientras que la variación entre las barras dentro de un molde puede darse por diferencias en las condiciones de enfriamiento o por variaciones en el procedimiento de medición de la fuerza de tensión.

5.5 Cursos de acción con estimaciones de varianzas negativas

Por definición, una componente de varianza es positiva; sin embargo, las estimaciones de σ_a^2 mediante la ecuación (5.3) pueden ser negativas. Existen varios cursos de acción sugeridos en caso de estimaciones negativas (Searle, 1971; Searle, Casella y McCulloch, 1992).

1. Aceptar la estimación como evidencia de un valor verdadero de cero y usar el cero como la estimación, reconociendo que el estimador será sesgado.
2. Conservar la estimación negativa, tomando en cuenta que es posible que los cálculos subsecuentes con los resultados no tengan mucho sentido.
3. Interpretar la estimación de la componente negativa como una indicación de un modelo estadístico incorrecto.
4. Utilizar un método diferente de análisis de varianza para estimar sus componentes.
5. Reunir más datos y analizarlos por separado o junto con los existentes y esperar que el aumento de información conduzca a estimaciones positivas.

Searle (1971, capítulo 9) y Searle *et al.* (1992, capítulo 4) presentan varios métodos sobre la estimación de las componentes de varianza, al igual que otras acciones, con mayor amplitud.

5.6 La correlación intraclases indica similitud dentro de los grupos

El coeficiente de correlación intraclases es una medida de la similitud de las observaciones dentro de grupos con las de entre grupos. Cuando la similitud de las observaciones dentro de grupos es muy alta, σ_e^2 será muy pequeña. En consecuencia, σ_a^2 será una proporción mayor de la variación total ($\sigma_y^2 = \sigma_a^2 + \sigma_e^2$). La correlación intraclases, definida como la razón:

$$\rho_I = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2} \quad (5.6)$$

se usa en varias disciplinas. Se le aplica en estudios de genética con distintas medidas de la herencia de las características cuantitativas, en estudios de confiabilidad para medir la similitud de productos de la misma máquina o proceso, en estudios médicos para medir la repetibilidad de medidas sucesivas en pacientes y en investigaciones de muestreo para medir la similitud de respuesta entre personas entrevistadas por el mismo individuo (Koch, 1983).

R. A. Fisher introdujo el análisis de varianza en la década de los 20, con un modelo de correlación intraclases (Fisher, 1960). Este modelo supone que todas las observaciones (y_{ij}) tienen la misma media (μ) y varianza (σ^2), y que cualesquiera dos miembros del mismo grupo tienen una correlación común (ρ_I). Con este modelo los cuadrados medios esperados para el análisis de varianza de la tabla 5.2 son:

$$\begin{aligned} E(CMA) &= \sigma^2 \{1 + (r - 1)\rho_I\} \\ \text{y} \\ E(CMW) &= \sigma^2(1 - \rho_I) \end{aligned} \tag{5.7}$$

Los estimadores de ρ_I y σ^2 se encuentran igualando los cuadrados de las medias observadas con las expectativas de la ecuación (5.7) y despejando las incógnitas. Las soluciones son:

$$\begin{aligned} \hat{\sigma}^2 &= \frac{\{CMA + (r - 1)CMW\}}{r} \\ \text{y} \\ \hat{\rho}_I &= \frac{(CMA - CMW)}{\{CMA + (r - 1)CMW\}} \end{aligned} \tag{5.8}$$

La estimación de la correlación intraclases puede tener un valor mínimo de $-1/(r - 1)$ y un valor máximo de 1 (Fisher, 1960), porque el valor esperado de CMA debe ser igual o mayor que cero.

El estimador del intervalo de confianza $100(1 - \alpha)\%$ para ρ_I es:

$$\frac{F_0 - F_u}{F_0 + (r - 1)F_u} < \rho_I < \frac{F_0 - F_l}{F_0 + (r - 1)F_l} \tag{5.9}$$

donde $F_u = F_{\alpha/2, (t-1)(N-t)}$, $F_l = F_{(1-\alpha/2), (t-1)(N-t)}$, y $F_0 = CMA/CMW$. El intervalo se puede usar para probar la hipótesis nula $H_0: \rho_I = 0$, donde la hipótesis no se rechaza si el intervalo incluye al cero (Koch, 1983).

La estimación de la correlación intraclases para los moldes de aleación a alta temperatura es:

$$\hat{\rho}_I = \frac{(73.94 - 5.82)}{\{73.94 + 9(5.82)\}} = 0.54$$

Con $F_0 = 12.71$, $F_{.05,2,27} = 3.35$ y $F_{.95,2,27} = 0.051$, la estimación de un intervalo de confianza del 90% es:

$$\frac{(12.71 - 3.35)}{\{12.71 + 9(3.35)\}} < \rho_I < \frac{(12.71 - 0.051)}{\{12.71 + 9(0.051)\}}$$

y

$$0.22 < \rho_I < 0.96$$

El intervalo no incluye el cero y la hipótesis nula de una correlación intraclasses igual a cero entre las barras dentro del molde se rechaza.

Se puede interpretar la correlación intraclasses con base en la razón en la ecuación (5.6). El numerador (σ_a^2) refleja la variación peculiar con las diferencias entre grupos, mientras que la varianza del denominador ($\sigma_a^2 + \sigma_e^2$) pertenece a los individuos de la muestra aleatoria del universo de todos los grupos, sin importar las fronteras entre ellos.

Si la correlación intraclasses es grande, el efecto aleatorio (a_i) común a un grupo afecta de manera similar a todos los individuos que pertenecen a él. Así, la similitud entre los individuos dentro de los grupos será mayor que entre individuos de grupos distintos y σ_e^2 será pequeña comparada con σ_a^2 .

Por otro lado, una correlación pequeña intraclasses indica disimilitud entre los individuos de un grupo, por ejemplo, los recursos nutritivos pueden provocar disparidades de crecimiento dentro de un grupo. Esto puede ocurrir si los individuos más vigorosos y agresivos toman una parte mayor del recurso nutritivo.

5.7 Diferentes números de observaciones en los grupos

El modelo de efectos aleatorios para una clasificación en un sentido con diferentes números de observaciones por grupo es:

$$y_{ij} = \mu + a_i + e_{ij} \quad (5.10)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r_i$$

con las mismas suposiciones e interpretaciones dadas para el modelo aleatorio en la ecuación (5.1). Los cálculos del análisis de varianza son los mismos que para el modelo de efectos fijos con réplicas desiguales. Los cuadrados de la media esperada para entre grupos y dentro de grupos son, respectivamente:

$$E(CMA) = \sigma_e^2 + r_0 \sigma_a^2$$

y

$$E(CME) = \sigma_e^2$$

$$(5.11)$$

donde:

$$r_0 = \frac{1}{t - 1} \left[N - \sum_{i=1}^t \frac{r_i^2}{N} \right] \quad (5.12)$$

Los estimadores del análisis de varianza para determinar las componentes σ_e^2 y σ_a^2 de la misma, son:

$$\widehat{\sigma}_e^2 = CMW \quad (5.13)$$

$$\widehat{\sigma}_a^2 = \frac{(CMA - CMW)}{r_0}$$

Cuando las r_i son diferentes, no se puede usar el estimador del intervalo de confianza para σ_a^2 en la ecuación (5.5). El estimador del intervalo para σ_e^2 en la ecuación (5.4) con $(N - t)$ grados de libertad es válido.

5.8 ¿Cuántas observaciones para estudiar las varianzas?

La hipótesis nula de interés en el modelo de efectos aleatorios, $H_0: \sigma_a^2 = 0$, se prueba con $F = CMA/CMW$, y la potencia de la prueba es:

$$1 - \beta = P(F > F_{\alpha, v1, v2} | H_0 \text{ falso}) = P(F > F_{\alpha, v1, v2} | \sigma_a^2 > 0)$$

Cuando $\sigma_a^2 > 0$, la distribución de F es la distribución central $F_{v1, v2}$ multiplicada por una constante $1/\lambda^2$, donde:

$$\lambda^2 = 1 + \frac{r\sigma_a^2}{\sigma_e^2} \quad (5.14)$$

La potencia de la prueba se puede determinar a partir de la distribución F central como sigue:

$$1 - \beta = P \left[F > \frac{1}{\lambda^2} (F_{\alpha, v1, v2}) \right]$$

Dado el número de grupos (t), un nivel de significancia (α), una potencia deseada ($1 - \beta$) y λ , el número de réplicas necesario se determina con gráficas de las curvas de potencia similares a las del modelo de efectos fijos.

El valor de λ se puede obtener con base en una razón deseada para las componentes de la varianza, σ_a^2/σ_e^2 , o según la desviación estándar de una observación individual, σ_y . Considerando el ejemplo 5.1, en el que el ingeniero fabrica varias barras de una aleación a alta temperatura en cada uno de varios moldes, si no hubiera variación en la resistencia de las barras debida al moldeo, entonces la

desviación estándar de una barra elegida al azar sería $\sigma_y = \sigma_e$. Tal vez el ingeniero quiera detectar un aumento en la variabilidad entre moldes (σ_a^2) que ocasione cierto porcentaje de aumento en σ_y . Supongamos que P es el porcentaje fijo de aumento en σ_y aceptable, y que con uno mayor la hipótesis nula se rechaza. La razón de σ_y entre σ_e , expresada en términos de P cuando la hipótesis nula se rechaza es:

$$\frac{\sigma_y}{\sigma_e} = \frac{\sqrt{\sigma_a^2 + \sigma_e^2}}{\sigma_e} = 1 + 0.01P$$

El valor necesario para la razón σ_a^2/σ_e^2 en la ecuación (5.14) es:

$$\frac{\sigma_a^2}{\sigma_e^2} = (1 + 0.01P)^2 - 1 \quad (5.15)$$

En la tabla X del apéndice se encuentran las curvas de potencia para $\alpha = .05, .01$ y una selección de valores para v_1 y v_2 , con distribución F ; muestran la potencia de la prueba, $1 - \beta$, contra λ , donde λ^2 está dada en la ecuación (5.14).

Ejemplo 5.2 Moldeo de aleaciones a alta temperatura de nuevo

En el ejemplo 5.1 las estimaciones de las componentes de la varianza fueron $\widehat{\sigma}_e^2 = 5.82$ y $\widehat{\sigma}_a^2 = 6.81$. Las desviaciones estándar estimadas son $\widehat{\sigma}_e = 2.41$ y $\widehat{\sigma}_y = 3.55$. Supongamos que el ingeniero produce $t = 5$ moldes y quiere detectar un incremento en la desviación estándar σ_y mayor que σ_e de $P = 35$, con una potencia de al menos .80 a un nivel de significancia de .05. El valor requerido para la razón en la ecuación (5.1) es:

$$\frac{\sigma_a^2}{\sigma_e^2} = [1 + 0.01(35)]^2 - 1 = 0.8225$$

de manera que $\lambda = \sqrt{1 + r(0.8225)}$. Si se elige un valor de $r = 10$, entonces $\lambda = 3$. Al leer la tabla X del apéndice para $v_1 = 4$ y localizar la posición aproximada de la línea para $v_2 = 45$, con $\alpha = .05$, el valor $1 - \beta$ se encuentra entre 0.8 y 0.9. Por lo tanto, el ingeniero puede medir 10 barras en cada uno de los cinco moldes para detectar un incremento del 35% o mayor en la desviación estándar debido al moldeo.

5.9 Submuestras aleatorias para reunir datos para el experimento

En ocasiones, para reunir los datos que requiere el estudio, es necesario o conveniente obtener muestras aleatorias de subunidades de las unidades del experimento; de esta manera, la unidad observada es una submuestra. Por ejemplo, al tomar una muestra de varias plantas en un cultivo para medir la plaga de insectos, otras veces se divide una muestra de suero en dos submuestras antes del análisis espectropométrico, o se extraen varias muestras de pintura de los lotes de cada formulación para probar durabilidad.

Las submuestras introducen otra fuente de variabilidad en las observaciones, además de la que hay entre las unidades experimentales. Es importante distinguir entre la contribución a la variación de las submuestras y la de las unidades experimentales; esta distinción es primordial para la estimación de los errores estándar de las medias de tratamientos y para las pruebas de hipótesis sobre éstos. En el capítulo 1 se presentó una introducción a esta distinción en los ejemplos 1.1 y 1.2 sobre las réplicas.

Las estimaciones de las componentes de la varianza para unidades experimentales y submuestras identifican la cantidad de variación a la que contribuyen las dos fuentes. Esta información se usa en la sección 5.10 para determinar el número relativo de unidades experimentales que minimiza el error estándar de las medias de tratamiento o el costo del experimento.

Ejemplo 5.3 Residuos de pesticida en plantas de algodón

La aplicación de pesticida a menudo es parte de programas de manejo de insectos usados en agronomía u horticultura. Una de las preocupaciones que sigue a la aplicación de pesticidas es la concentración de sus residuos que permanecen en las plantas después de cierto periodo. Los residuos de pesticida se evalúan con análisis químicos de laboratorio en una muestra de plantas de los cultivos tratados con pesticida.

Hipótesis de investigación: para un problema en particular, los investigadores plantearon la hipótesis de que la capacidad para recuperar residuos de pesticida en las hojas de la planta de algodón difiere entre dos métodos químicos normales que se usaban de manera regular para las pruebas residuales.

Diseño del tratamiento: los tratamientos consisten en dos métodos químicos normales, A y B, que se usaban regularmente.

Diseño del experimento: se obtuvieron muestras en campo de seis lotes de plantas, cada uno de un cultivo distinto y se prepararon para el análisis de residuos; se asignaron tres lotes al azar a cada método en un diseño totalmente aleatorizado.

La cantidad de material de las plantas en cada lote de la muestra de campo excedía la cantidad necesaria para las pruebas de laboratorio, entonces se tomaron de cada lote de tejido de planta preparado dos submuestras del tamaño necesario y se analizaron con el método adecuado. En consecuencia, se tenían dos métodos químicos, tres lotes (réplicas) por método y dos submuestras de cada lote. En la tabla 5.4 se presentan los residuos de pesticida determinados para cada submuestra, en microgramos.

Modelo estadístico con submuestras

Cuando se tienen n submuestras de cada una de r unidades experimentales para t tratamientos, el modelo estadístico es:

$$y_{ijk} = \mu + \tau_i + e_{ij} + d_{ijk} \quad (5.16)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad k = 1, 2, \dots, n$$

donde μ es la media general, τ_i es el efecto fijo del i -ésimo tratamiento, e_{ij} es el efecto de error experimental aleatorio para la unidad experimental j del tratamiento i ,

Tabla 5.4 Residuos de pesticida (μg) encontrados en muestras de plantas de algodón

Método A					Método B				
Lote	Muestra	y_{ijk}	$\bar{y}_{ij.}$	$\bar{y}_{i..}$	Lote	Muestra	y_{ijk}	$\bar{y}_{ij.}$	$\bar{y}_{i..}$
1	1	120			4	7	71		
	2	110	115.0			8	71	71.0	
2	3	120			5	9	70		
	4	100	110.0			10	76	73.0	
3	5	140			6	11	63		
	6	130	135.0	120.0		12	68	65.5	69.8
$\bar{y}_{...} = 94.9$									

Fuente G. Ware y B. Estesen, Department of Entomology, University of Arizona.

y d_{ijk} es el efecto aleatorio para la k -ésima submuestra de la unidad experimental j del tratamiento i . Se supone que e_{ij} y d_{ijk} tienen efectos aleatorios independientes con una distribución normal, con medias 0 y varianzas σ_e^2 y σ_d^2 , respectivamente. Si los tratamientos son aleatorios, entonces los efectos fijos de los tratamientos (τ_i) en la ecuación (5.16) se reemplazan por los efectos aleatorios de grupo (a_i), que tienen distribución normal con media 0 y varianza σ_a^2 .

Análisis de varianza con submuestras

Las observaciones, expresadas como desviaciones de la gran media, se pueden describir como la suma de tres desviaciones separadas que representan las fuentes de variación en el experimento:

$$(y_{ijk} - \bar{y}_{...}) = (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{ij.} - \bar{y}_{i..}) + (y_{ijk} - \bar{y}_{ij.}) \quad (5.17)$$

La desviación de cualquier observación con respecto a la gran media mostrada en el lado izquierdo de la ecuación (5.17) es la suma de tres términos, que son:

- desviación del tratamiento ($\bar{y}_{i..} - \bar{y}_{...}$) [p. ej. $(\bar{y}_{1..} - \bar{y}_{...}) = 120.0 - 94.9 = 25.1$]
- error experimental ($\bar{y}_{ij.} - \bar{y}_{i..}$) [p. ej. $(\bar{y}_{11.} - \bar{y}_{1..}) = 115.0 - 120.0 = -5.0$]
- error de muestreo ($y_{ijk} - \bar{y}_{ij.}$) [p. ej. $(y_{111} - \bar{y}_{11.}) = 120.0 - 115.0 = 5.0$]

Al elevar al cuadrado y sumar ambos lados de la ecuación (5.17) se obtiene la suma de cuadrados total en el lado izquierdo, expresada como una suma de las sumas de cuadrados respectivas para tratamientos, error experimental y muestreo. La partición fundamental de la suma de cuadrados total es:

$$SC \text{ total} = SC \text{ tratamiento} + SC \text{ error} + SC \text{ muestreo} \quad (5.18)$$

Las particiones de sumas de cuadrados se resumen en el análisis de varianza de la tabla 5.5, con los cuadrados medios esperados para el modelo con efectos de tratamiento fijos.

Tabla 5.5 Análisis de varianza para el diseño totalmente aleatorizado con submuestras²

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Cuadrados medios esperados
Total	$trn - 1$	SC total		
Tratamientos	$t - 1$	SCT	CMT	$\sigma_d^2 + n\sigma_e^2 + rn\theta_t^2$
Error	$t(r - 1)$	SCE	CME	$\sigma_d^2 + n\sigma_e^2$
Muestra	$tr(n - 1)$	SCM	CMM	σ_d^2

$$\begin{aligned} SC \text{ total} &= \sum_{i=1}^t \sum_{j=1}^r \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 \\ SCT &= SC \text{ tratamiento} = rn \sum_{i=1}^t (\bar{y}_{i..} - \bar{y}_{...})^2 \\ SCE &= SC \text{ error} = n \sum_{i=1}^t \sum_{j=1}^r (\bar{y}_{ij.} - \bar{y}_{i..})^2 \\ SCM &= SC \text{ muestra} = \sum_{i=1}^t \sum_{j=1}^r \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.})^2 \end{aligned}$$

Son dos las fuentes que contribuyen a la variación en las observaciones que constituyen la estimación de una media de tratamiento: la variación entre las réplicas de unidades experimentales que se manejan igual (σ_e^2) y la variación entre las unidades de muestreo dentro de la misma unidad experimental (σ_d^2). En consecuencia, la varianza de una media de tratamiento es:

$$\sigma_{\bar{y}_{i..}}^2 = \frac{\sigma_d^2}{rn} + \frac{\sigma_e^2}{r} \quad (5.19)$$

donde existen n submuestras de cada una de las r réplicas de las unidades experimentales. El error de una media de tratamiento se estima por medio de:

$$s_{\bar{y}_{i..}} = \sqrt{\frac{CME}{rn}} \quad (5.20)$$

En la tabla 5.6 se muestra el análisis de varianza de los datos en la tabla 5.4. La hipótesis nula de que no hay diferencias entre los efectos de los tratamientos, $H_0: \tau_i = 0$, se rechaza si $F_0 = CMT/CME$ excede a $F_{\alpha, (t-1), t(r-1)}$. El estadístico³ F_0 de la tabla 5.6 para probar la hipótesis nula de que no hay diferencia entre las medias de los métodos A y B, $F_0 = 7,550.08/190.8 = 39.72$, se excede con una

²La suma de cuadrados del “error” representa la suma de cuadrados de las unidades experimentales anidadas dentro de los grupos de tratamiento. Un programa de computadora requiere un término en su sintaxis que designe a las unidades experimentales *dentro* de los tratamientos.

³Con frecuencia, es necesario especificar el denominador correcto para el estadístico F_0 en las instrucciones del programa de computadora. Muchos programas tienen predefinido el uso de la suma de cuadrados de la última partición en el denominador de F_0 en todas las líneas de la tabla de análisis de varianza.

probabilidad .003. Como la media del método A, $\bar{y}_{1..} = 120$, excede a la del método B, $\bar{y}_{2..} = 69.8$, se puede concluir que el método A recupera más residuos de pesticida que el método B.

Tabla 5.6 Análisis de varianza para residuos de pesticida en las submuestras de plantas de algodón

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	11	8,640.91			
Métodos	1	7,550.08	7,550.08	39.72	.003
Error	4	760.33	190.08		
Muestra	6	330.50	55.08		

La estimación del error estándar para la media de un método, ecuación (5.20), es:

$$s_{\bar{y}_{1..}} = \sqrt{\frac{190.08}{6}} = 5.63$$

y la estimación del error estándar de la diferencia entre las medias de los dos métodos es:

$$s(\bar{y}_{1..} - \bar{y}_{2..}) = \sqrt{\frac{2(190.08)}{6}} = 7.96$$

5.10 Uso de las estimaciones de la varianza para asignar muestras

La distribución de recursos en la etapa de planeación de un experimento que incluye submuestas, requiere decisiones respecto al número de unidades experimentales y el número de submuestas a tomar de cada unidad experimental. El objetivo es lograr un diseño que tenga como resultado una mayor precisión —una varianza más pequeña para la estimación de la media de un tratamiento ($\sigma_{\bar{y}_{i..}}^2$), por un costo fijo.— Cuando se dispone de estimaciones de las componentes de la varianza y de los costos relativos de las unidades experimentales y del muestreo, es posible una asignación óptima del producto entre las unidades experimentales y las unidades de muestra en el experimento.

Cochran (1965) desarrolló una solución óptima de la asignación, basada en la función de costo $C = c_1 r + c_2 rn$. El valor de C es el costo de un solo tratamiento en un experimento compuesto de r unidades experimentales, cada una con costo c_1 y rn unidades muestrales, cada una con costo c_2 . El objetivo puede expresarse como el logro del costo mínimo (C) para una varianza fija en la ecuación (5.19), o el logro de una varianza mínima para un costo fijo. De cualquier manera, la solución para el número de unidades muestrales (n) es:

$$n = \sqrt{\frac{c_1 \sigma_d^2}{c_2 \sigma_e^2}} \quad (5.21)$$

El valor de r se encuentra despejando r de la ecuación de costo, si el costo es fijo; o la varianza de la ecuación (5.19), si aquélla es fija.

Ejemplo 5.4 Residuos de pesticida otra vez

Para realizar los análisis de residuos se requiere de una asignación óptima para los cultivos de plantas de algodón y las submuestras por cultivo. Las estimaciones de las componentes de la varianza obtenidas de la tabla 5.6 son:

$$\widehat{\sigma}_d^2 = CMM = 55.08$$

y

$$\widehat{\sigma}_e^2 = \frac{(CME - CMM)}{n} = \frac{(190.08 - 55.08)}{2} = 67.50$$

Suponiendo que el costo de un cultivo es $c_1 = 1.0$, mientras que el de preparar y analizar una submuestra, $c_2 = 0.1$. El número estimado de submuestras por cultivo es:

$$n = \sqrt{\frac{1(55.08)}{0.1(67.50)}} = 2.86$$

Si el investigador desea un error estándar de la media del tratamiento de $\sigma_{\bar{y}_{i..}} = 3$ o una varianza de 9, el número de cultivos r requeridos se puede encontrar sustituyendo estas cantidades en la ecuación (5.19):

$$\widehat{\sigma}_{\bar{y}_{i..}}^2 = 9, \widehat{\sigma}_e^2 = 67.5, \widehat{\sigma}_d^2 = 55.08, \text{ y } n = 3$$

de manera que:

$$9 = \frac{55.08}{r \cdot 3} + \frac{67.5}{r}; 9r = 85.86; r = 9.54$$

El investigador deberá usar diez cultivos y tres submuestras de cada uno para tener un error estándar de las medias de tratamiento igual a 3, con costos relativos de 1.0 y 0.1 para los cultivos y las submuestras, respectivamente.

5.11 Números diferentes de réplicas y submuestras

Es posible que se realicen números desiguales de submuestras y réplicas en un estudio, en realidad las tres distintas posibilidades son: 1) números diferentes de

unidades experimentales por tratamiento con números diferentes de submuestras por unidad experimental, 2) igual número de unidades experimentales por tratamiento con números diferentes de submuestras por unidad experimental y 3) números diferentes de unidades experimentales por tratamiento con igual número de submuestras por unidad experimental. Cualquier desbalance en el número de observaciones en la unidad experimental o en la etapa de submuestras, afecta los cálculos del análisis de varianza y los cuadrados medios esperados.

Ejemplo 5.5 Biología del gusano de tabaco

Las poblaciones de insectos suelen desarrollar resistencia a los efectos tóxicos de un insecticida después de una exposición prolongada; cuando esto ocurre el insecticida pierde su efectividad para mantener la población por debajo de los niveles dañinos.

Las poblaciones de gusano de tabaco, una plaga nociva para la planta de algodón, ha desarrollado resistencia a varios insecticidas comunes que son parte del programa global para el control de insectos en los cultivos. Otros componentes del programa de control también dependen de la biología de los insectos en términos de su ciclo de vida reproductivo y sus patrones de desarrollo.

Hipótesis de investigación: los entomólogos plantearon la hipótesis de que el desarrollo de la resistencia al insecticida también puede afectar otros aspectos de la biología de los gusanos. Si esto es cierto, entonces los cambios en la biología del insecto tendrán efectos en los programas de control de los mismos.

Diseño del tratamiento: los tratamientos usados para estudiar la hipótesis de investigación incluyen tres variedades de gusano de tabaco: 1) USDA, una variedad muy susceptible a un insecticida de piretroide; 2) Resistente, una variedad bastante resistente al insecticida, y 3) Campo, un variedad de aparición natural en los cultivos de algodón locales. Las variedades Resistente y USDA fueron poblaciones mantenidas en ambientes artificiales para mantener sus características de resistencia. Cualesquiera diferencias en la biología de las variedades se consideraron un reflejo de cambios asociados con el desarrollo de resistencia al insecticida, mientras que la biología de la variedad Campo sirvió como tratamiento de control para esta suposición. Una de las características medidas para evaluar la biología fue el peso de la larva macho.

Diseño del experimento: se realizaron seis apareamientos aleatorios entre hembras y machos de cada variedad, y las descendencias se criaron en secciones separadas en el laboratorio. Las 18 secciones se colocaron al azar en la instalación. El resultado fue un número de descendientes diferente en los 18 apareamientos, como se muestra en los datos de la tabla 5.7.

Modelo estadístico y análisis

Suponiendo que el experimento tiene r_i unidades experimentales para el i -ésimo grupo de tratamiento y n_{ij} submuestras para la j -ésima unidad experimental. El modelo estadístico para muestras diferentes es:

$$y_{ijk} = \mu + \tau_i + e_{ij} + d_{ijk} \quad (5.22)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r_i \quad k = 1, 2, \dots, n_{ij}$$

Tabla 5.7 Peso de las larvas macho resultado de seis apareamientos en cada una de tres variedades de gusano de tabaco

Variedad	Apareamiento	Peso	n_{ij}	$n_i.$	$\bar{y}_{ij.}$	$\bar{y}_{i..}$
USDA	1	305, 300	2	12	302.5	
	2	376, 363, 389	3		376.0	
	3	282	1		282.0	
	4	309, 321	2		315.0	
	5	354, 308, 327	3		329.7	
	6	330	1		330.0	330.3
Campo	7	280	1	13	280.0	
	8	311, 349, 291, 286	4		309.3	
	9	377, 342	2		359.5	
	10	346, 340, 347	3		344.3	
	11	360	1		360.0	
	12	359, 299	2		329.0	329.8
Resistente	13	273, 276	2	11	274.5	
	14	272, 253	2		262.5	
	15	315, 262, 297	3		291.3	
	16	323	1		323.0	
	17	252	1		252.0	
	18	319, 298	2		308.5	285.5
$\bar{y}_{...} = 316.4$						

Fuente. Dr. T. Watson y S. Kelly, Department of Entomology, University of Arizona.

donde μ es la media general, τ_i es el efecto fijo del tratamiento i , e_{ij} es el efecto del error experimental aleatorio para la unidad experimental j de tratamiento i , y d_{ijk} es el efecto aleatorio de la submuestra k de la unidad experimental j del tratamiento i . Se supone que e_{ij} y d_{ijk} son efectos aleatorios independientes que siguen distribuciones normales, con media 0 y varianzas respectivas σ_e^2 y σ_d^2 . Si los tratamientos son aleatorios, entonces los efectos fijos del tratamiento (τ_i) en la ecuación (5.22) se sustituyen por los efectos aleatorios de grupo (a_i), que tienen una distribución normal con media 0 y varianza σ_a^2 .

En la tabla 5.8 pueden verse las particiones de las sumas de cuadrados para el análisis de varianza y los cuadrados medios esperados para los efectos de tratamiento aleatorios. Si los efectos del tratamiento son fríos se usa $\sum n_{ij} (\mu_i - \bar{\mu})^2 / (t - 1)$ en vez de $c_2 \sigma_a^2$.

Observe que las sumas de las desviaciones de la media del tratamiento en SCT se ponderan según el número de observaciones en los tratamientos $n_i.$, y los cuadrados de las desviaciones del error experimental en SCE se ponderan por el número de submuestras, n_{ij} . Los coeficientes de las componentes de la varianza en los cuadrados medios esperados con efectos de tratamiento aleatorio son:

Tabla 5.8 Análisis varianza para el diseño totalmente aleatorizado con números diferentes de réplicas y submuestras

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrado de la media	Cuadrado de la media esperada
Total	$N - 1$	$SC \text{ total}$		
Tratamientos	$t - 1$	SCT	CMT	$\sigma_d^2 + c_1\sigma_e^2 + c_2\sigma_a^2$
Error	$\sum_{i=1}^t r_i - t$	SCE	CME	$\sigma_d^2 + c_3\sigma_e^2$
Muestra	$N - \sum_{i=1}^t r_i$	SCM	CMM	σ_d^2

$$N = \sum_{i=1}^t \sum_{j=1}^{r_i} n_{ij} \quad n_{i..} = \sum_{j=1}^{r_i} n_{ij}$$

$$SC \text{ total} = \sum_{i=1}^t \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{...})^2$$

$$SCT = \sum_{i=1}^t n_{i..} (\bar{y}_{i..} - \bar{y}_{...})^2$$

$$SCE = \sum_{i=1}^t \sum_{j=1}^{r_i} n_{ij} (\bar{y}_{ij.} - \bar{y}_{i..})^2$$

$$SCM = \sum_{i=1}^t \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij.})^2$$

$$c_1 = \frac{1}{t-1} \left(A - \frac{B}{N} \right), \quad c_2 = \frac{1}{t-1} \left(N - \frac{D}{N} \right), \quad c_3 = \frac{1}{\left(\sum_{i=1}^t r_i - t \right)} (N - A)$$

donde:

(5.23)

$$A = \sum_{i=1}^t \sum_{j=1}^{r_i} \left(\frac{n_{ij}^2}{n_{i..}} \right), \quad B = \sum_{i=1}^t \sum_{j=1}^{r_i} n_{ij}^2, \quad D = \sum_{i=1}^t n_{i..}^2$$

Si el número de submuestras es igual para cada unidad experimental, entonces $n_{ij} = n$ para toda i y j ; los coeficientes son:

$$c_1 = c_3 = n \quad y \quad c_2 = \frac{1}{t-1} \left[N - \frac{D}{N} \right] \quad (5.24)$$

Cuando $c_1 = c_3 = n$, los cuadrados medios esperados para tratamientos y error serán idénticos bajo la hipótesis nula $H_0: \tau_i = 0$ y se usa el estadístico $F_0 = CMT/CME$ para probar la hipótesis.

El análisis de varianza de estos datos se muestra en la tabla 5.9, también se muestra el cuadrado medio esperado del mismo análisis para los efectos aleatorios. Los cálculos de los coeficientes para las componentes de la varianza se dan en el apéndice 5A.

Tabla 5.9 Análisis de varianza para los pesos de larvas macho resultantes de seis apareamientos en cada una de tres variedades de gusano de tabaco

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Cuadrados medios esperados
Total	35	46,516.75		
Tratamientos	2	15,187.05	7,593.52	$\sigma_d^2 + 2.36\sigma_e^2 + 11.97\sigma_a^2$
Error	15	23,082.45	1,538.83	$\sigma_d^2 + 1.93\sigma_e^2$
Muestra	18	8,247.25	458.18	σ_d^2

Las pruebas de hipótesis requieren pruebas F aproximadas

Si el número de submuestras no es igual, c_1 y c_3 pueden tener valores distintos. No existe una prueba exacta de la hipótesis nula para los efectos del tratamiento, porque no hay dos medias cuadráticas que tengan el mismo cuadrado medio esperado bajo la hipótesis nula si c_1 y c_3 tienen valores distintos. Es posible calcular un estadístico F_0 aproximado para probar la hipótesis nula de que no hay efectos de tratamiento cuando $c_1 \neq c_3$; es decir, es necesaria una prueba aproximada para el experimento con gusanos de tabaco, puesto que $c_1 = 2.36$ y $c_3 = 1.93$ en la tabla 5.9.

Se desarrolla un cuadrado medio para el error con un valor esperado igual a la del cuadrado medio de los tratamientos, dada una hipótesis nula con $E(CMT) = \sigma_d^2 + c_1\sigma_e^2$; esto se hace con una función lineal de CMM y CME :

$$M = a_1CME + a_2CMM \quad (5.25)$$

Si $a_1 = c_1/c_3$ y $a_2 = 1 - c_1/c_3$, el valor esperado de M será $\sigma_d^2 + c_1\sigma_e^2$ como lo requiere la prueba de F_0 aproximada.

Aproximación de Satterthwaite para los grados de libertad

Satterthwaite (1946) obtuvo el siguiente resultado para una función lineal de medias cuadráticas. Dada una función lineal M , donde:

$$M = a_1 CM_1 + a_2 CM_2 + \dots + a_k CM_k \quad (5.26)$$

y CM_1, CM_2, \dots, CM_k son los cuadrados medios con v_1, v_2, \dots, v_k , grados de libertad, respectivamente; los grados de libertad de M se aproximan por medio de:

$$v = \frac{M^2}{\sum_{i=1}^k \frac{(a_i CM_i)^2}{v_i}} \quad (5.27)$$

La función lineal de los cuadrados medios necesarios para probar la hipótesis de que no hay diferencia en los pesos medios de las larvas entre las distintas variedades de gusanos requiere:

$$a_1 = \frac{c_1}{c_3} = \frac{2.36}{1.93} = 1.22 \quad y \quad a_2 = 1 - \frac{c_1}{c_3} = 1 - 1.22 = -0.22$$

$$CME = 1,538.83 \text{ con } 15 \text{ g. l.} \quad y \quad CMM = 458.18 \text{ con } 18 \text{ g.l.}$$

De la ecuación (5.25):

$$M = a_1 CME + a_2 CMM = 1.22(1,538.83) - 0.22(458.18) = 1,776.57$$

Los grados de libertad de M de la ecuación (5.27) son:

$$v = \frac{1,776.57^2}{\frac{[1.22(1,538.83)]^2}{15} + \frac{[-0.22(458.18)]^2}{18}} = 13.4$$

Se utiliza el valor truncado a $v = 13$ como los grados de libertad de M , el valor del estadístico de prueba es $F_0 = CMT/M = 7,593.53/1,776.57 = 4.27$, el valor crítico al nivel de significancia de $\alpha = .05$ es $F_{.05, 2, 13} = 3.81$ y la hipótesis nula se rechaza. Existen algunas diferencias entre los pesos medios de larvas de las tres variedades. La prueba es sólo aproximada y debe observarse que la aproximación se degrada un poco si algunos coeficientes (a_i) en la ecuación (5.26) son negativos.

EJERCICIOS

- Un estudio de genética con reses consistió en varios machos apareados con grupos separados de hembras. Cuando nacían terneros, se usaban en un estudio de pesos hereditarios. En la siguiente tabla se presentan los pesos al nacer de ocho terneros de cada uno de los cinco grupos de apareamiento.

<i>Macho</i>	<i>Peso al nacer</i>
177	61, 100, 56, 113, 99, 103, 75, 62
200	75, 102, 95, 103, 98, 115, 98, 94
201	58, 60, 60, 57, 57, 59, 54, 100
202	57, 56, 67, 59, 58, 121, 101, 101
203	59, 46, 120, 115, 115, 93, 105, 75

Fuente. Dr. S. DeNise, Department of Animal Sciences, University of Arizona.

- 5.
- a. Suponga un modelo aleatorio para este estudio. Escriba el modelo lineal, explique cada término, calcule el análisis de varianza y muestre los cuadrados medios esperados.
 - b. Estime las componentes de la varianza para machos y progenie dentro de los machos y determine los intervalos de confianza estimados de 90%.
 - c. Pruebe la hipótesis nula $H_0: \sigma_a^2 = 0$ para los machos.
 - d. Estime los coeficientes de correlación intraclases y obtenga el intervalo de confianza estimado de 90%.
2. Los datos del ejercicio 3.5 corresponden a las concentraciones de colesterol en análisis de laboratorio a 2 muestras de cada uno de 8 pacientes.
- a. Suponga un modelo aleatorio para el estudio. Escriba el modelo lineal, explique cada término, calcule el análisis de varianza y muestre los cuadrados medios esperados.
 - b. Estime las componentes de la varianza para pacientes y muestras y determine intervalos de confianza medios al 90%.
 - c. Estime los coeficientes de correlación intraclases y obtenga el intervalo de confianza estimado de 90%.
 - d. ¿Cuál es la interpretación del coeficiente de correlación intraclases en este estudio?
3. Piense en problemas de investigación en su área de interés donde pueda realizar un muestreo aleatorio de los tratamientos en el estudio de una población grande de tratamientos.
- a. Describa un estudio en particular que pueda realizar.
 - b. Describa cómo realizaría el estudio.
 - c. Escriba el modelo lineal para su estudio; identifique los términos, y elabore la tabla de análisis de varianza mostrando fuentes de variación, grados de libertad y cuadrados medios esperados.
 - d. Explique por qué sería importante conocer la magnitud de las componentes de la varianza entre grupos y dentro de los grupos.
 - e. Describa cómo usaría las estimaciones de las componentes de la varianza.
 - f. ¿Qué suposiciones tendrá que hacer sobre su estudio para tener inferencias válidas a partir de sus estimaciones de las componentes de la varianza?
4. Un patólogo de plantas tomó cuatro muestras, de 3 libras cada una, de lotes de 50 toneladas de semilla de algodón acumulada en varias cosechas durante la temporada de limpia. Las muestras se analizaron en el laboratorio para buscar aflatoxin, que es una toxina producida por organismos asociados con las semillas. A continuación se proporcionan las concentraciones de aflatoxin en partes por billón (mil millones) para las muestras de ocho lotes.
- 6.

Número de lote	Aflatoxin (ppb)
3,469 – 72	39, 57, 63, 66
3,849 – 52	56, 13, 25, 31
3,721 – 24	64, 83, 88, 71
3,477 – 80	29, 55, 21, 51
3,669 – 72	38, 66, 53, 81
3,873 – 76	11, 49, 34, 10
3,777 – 80	23, 0, 5, 20
3,461 – 64	10, 11, 23, 37

Fuente: Dr. T. Russell, Department of Plant Pathology,
University of Arizona.

- a. Suponga que los lotes y sus muestras son efectos aleatorios. Escriba el modelo lineal para el estudio, explique los términos, calcule el análisis de varianza completo y muestre los cuadrados medios esperados.
 - b. Estime las componentes de la varianza para los lotes y para las muestras.
 - c. ¿Cuál es la varianza total estimada ($\widehat{\sigma}_y^2$) para una observación individual?
 - d. ¿Cuál es la proporción de la variación total (σ_y^2) de aflatoxin que se puede atribuir a la variación entre lotes y entre muestras respectivamente?
 - e. ¿Cuál es la estimación de la desviación estándar ($\widehat{\sigma}_y$) para una observación individual?
 - f. Explique cómo se pueden usar las estimaciones de las componentes de la varianza para planear futuros muestreos de contaminación de aflatoxin.
5. Piense en problemas de investigación en su área de interés que requieran muestras (u observaciones) de la unidad experimental debido a que no sea posible medir la unidad en su totalidad.
 - a. Describa un estudio específico que pueda realizar.
 - b. Describa cómo realizaría el estudio.
 - c. Escriba un modelo lineal para su estudio; identifique los términos y bosqueje el análisis de varianza, muestre las fuentes de variación, los grados de libertad y los cuadrados medios esperados.
 - d. ¿Cuáles serían los costos relativos de las unidades experimentales (c_1) y las unidades de muestreo (c_2)?
 6. Se realizó en conjunto un estudio sobre cartuchos para filtrado de partículas de alta energía, usados en respiradores comerciales para protección contra partículas de materia. Una prueba específica incluyó tres filtros elegidos al azar de cada uno de dos fabricantes, se hicieron tres réplicas de prueba independientes de cada filtro, las medidas fueron el porcentaje de penetración por medio de una prueba estándar de aerosol.

Filtro	Fabricante 1			Fabricante 2		
	1	2	3	4	5	6
1.12	0.16	0.15		0.91	0.66	2.17
1.10	0.11	0.12		0.83	0.83	1.52
1.12	0.26	0.12		0.95	0.61	1.58

Fuente: R. J. Beckman y C. J. Nachtsheim (1987), Diagnostics for mixed-model analysis of variance, *Technometrics* 29, 413-426.

- Escriba un modelo lineal para este estudio, explique los términos, calcule el análisis de varianza y muestre los cuadrados medios esperados.
 - Pruebe la hipótesis de que no existen diferencias entre la penetración porcentual promedio de los filtros de los dos fabricantes.
 - Calcule las medias, sus errores estándar y las estimaciones del intervalo de confianza de 95% para las medias de cada fabricante.
 - Suponga que los costos relativos, $c_1:c_2$, para el estudio son 200:1, donde c_1 es el costo de un filtro y c_2 es el costo de una prueba independiente de un filtro. Los ingenieros quieren lograr un error estándar para una media de 0.20. ¿Cuántos filtros y cuántas pruebas por filtro se requieren?
7. Un científico de suelos estudió el crecimiento de plantas de cebada en tres niveles diferentes de salinidad en un medio controlado. Tenía dos contenedores réplica de cada tratamiento, en un diseño totalmente aleatorizado y midió tres plantas de cada réplica. Los pesos en seco de las plantas, en gramos, son los siguientes:

<i>Salinidad</i>	<i>Contenedor</i>	<i>Peso (g)</i>
Control	1	11.29, 11.08, 11.10
	2	7.37, 6.55, 8.50
6 barras	3	5.64, 5.98, 5.69
	4	4.20, 3.34, 4.21
12 barras	5	4.83, 4.77, 5.66
	6	3.28, 2.61, 2.69

Fuente: Dr. T. C. Tucker, Department of Soil and Water Resources,
University of Arizona.

- Escriba un modelo lineal para un análisis de datos, explique los términos, calcule el análisis de varianza y muestre los cuadrados medios esperados.
 - Pruebe la hipótesis de que no hay diferencia entre las medias de los niveles salinos.
 - Calcule el error estándar de una media de nivel salino.
 - Haga una partición de las sumas de cuadrados para la salinidad en dos sumas de cuadrados polinomiales ortogonales (lineal y cuadrática), cada una con 1 grado de libertad y pruebe la hipótesis nula de que no hay regresión lineal o cuadrática.
 - Suponga que los costos relativos, $c_1:c_2$, son 10:0.1, donde c_1 es el costo de preparar y mantener un contenedor réplica y c_2 es el costo de medir los pesos en un contenedor. ¿Cuántos contenedores réplica y plantas por contenedor serán necesarios para lograr un error estándar de 0.75 para la media de un tratamiento?
8. El índice de porosidad es una medida usada por los científicos de suelos para ayudar en la predicción del movimiento, almacenamiento, disponibilidad del agua y las condiciones de oxigenación del subsuelo. Un científico de suelos usó un diseño de muestreo especial para tomar muestras del suelo de una de las granjas experimentales de la universidad para medir el índice de porosidad del suelo. Se hizo una partición de la granja en campos de aproximadamente 4 hectáreas, cada una con 8 secciones. El plan de muestreo incluyó una selección aleatoria de los campos dentro de las secciones. Se usó para el estudio un diseño de muestreo escalonado de Goldsmith y Gaylor (1970). A continuación se presenta el índice de porosidad de cada submuestra:

<i>Campo</i>	<i>Sección</i>	<i>Porosidad</i>	<i>Campo</i>	<i>Sección</i>	<i>Porosidad</i>
1	1	3.846, 3.712	9	17	5.942
	2	5.629, 2.021		18	5.014
2	3	5.087	10	19	5.143
	4	4.621		20	4.061
3	5	4.411	11	21	3.835, 2.964
	6	3.357		22	4.584, 4.398
4	7	3.991	12	23	4.193
	8	5.766		24	4.125
5	9	5.677	13	25	3.074
	10	3.333		26	3.483
6	11	4.355, 6.292	14	27	3.867
	12	4.940, 4.810		28	4.212
7	13	2.983	15	29	6.247
	14	4.396		30	4.730
8	15	5.603			
	16	3.683			

Fuente: Dr A. Warrick y M. Coelho, Department of Soil and Water Resources, University of Arizona.

- Suponga que todos los efectos son aleatorios. Escriba un modelo lineal para el estudio, explique cada término, calcule el análisis de varianza para los datos y muestre los cuadrados medios esperados.
 - Estime las componentes de la varianza para campos, secciones y muestras.
 - Pruebe la hipótesis nula $H_0: \sigma_a^2 = 0$, para la componente de la varianza de los campos.
 - Pruebe la hipótesis nula $H_0: \sigma_e^2 = 0$ para la componente de la varianza de las secciones.
9. Use los datos del ejercicio 5.3 para determinar cuántas muestras tendrá que tomar el patólogo de plantas de cada uno de los cinco lotes de semilla de algodón para detectar una razón $\sigma_a^2/\sigma_e^2 = 2$, a 0.1 de nivel de significancia, con una potencia de .90.

Apéndice 5A: Cálculo de los coeficientes para los cuadrados medios esperados de la tabla 5.9

$$A = \sum_{i=1}^t \sum_{j=1}^{r_i} \left(\frac{n_{ij}^2}{n_{i\cdot}} \right)$$

$$= \frac{2^2 + \dots + 1^2}{12} + \frac{1^2 + \dots + 2^2}{13} + \frac{2^2 + \dots + 2^2}{11} = 7.11655$$

$$B = \sum_{i=1}^t \sum_{j=1}^{r_i} n_{ij}^2 = 2^2 + 3^2 + \dots + 1^2 + 2^2 = 86$$

$$D = \sum_{i=1}^t n_{i\cdot}^2 = 12^2 + 13^2 + 11^2 = 434$$

$$c_1 = \frac{1}{t-1} \left(A - \frac{B}{N} \right) = \frac{1}{2} \left(7.11655 - \frac{86}{36} \right) = 2.36$$

$$c_2 = \frac{1}{t-1} \left(N - \frac{D}{N} \right) = \frac{1}{2} \left(36 - \frac{434}{36} \right) = 11.97$$

$$c_3 = \frac{1}{\left(\sum_{i=1}^t r_i - t \right)} (N - A) = \frac{1}{15} (36 - 7.11655) = 1.93$$

6 Diseños factoriales

Los diseños factoriales que se introdujeron en el capítulo 1 como una manera de investigar las relaciones entre varios tipos de tratamientos, en este capítulo se introducen para un diseño de experimento totalmente aleatorizado y su análisis. Se aplican las comparaciones planeadas y la estimación de las curvas de respuesta, estudiadas en el capítulo 3, al diseño de tratamientos de este tipo, también se estudian algunos métodos para determinar el número necesario de réplicas y para analizar el diseño factorial de tratamientos con una réplica o con un número desigual de réplicas.

6.1 Experimentos eficientes con diseños factoriales

Es posible que las comparaciones entre los tratamientos se vean afectadas de manera sustancial por las condiciones en las que ocurren. Con frecuencia, las interpretaciones claras de los efectos para un factor de tratamiento deben tomar en cuenta los efectos de los otros factores. Para investigar más de un factor a la vez, se desarrolló un tipo especial de diseño de tratamientos, el **diseño factorial**.

Los diseños factoriales producen experimentos más eficientes, pues cada observación proporciona información sobre todos los factores, y es factible ver las respuestas de un factor en diferentes niveles de otro factor en el mismo experimento. La respuesta a cualquier factor observado en diferentes condiciones indica si los factores actúan en las unidades experimentales de manera independiente. La interacción entre factores ocurre cuando su actuación no es independiente.

Ejemplo 6.1 Efectos de la compactación en la durabilidad del concreto asfáltico

Los pavimentos de asfalto sufren deterioros asociados con el agua como grietas, agujeros y desgaste de la superficie; el pavimento se debilita cuando hay una

ruptura en el adhesivo que une el cemento mezclado y el asfáltico que lo conforman. La investigación se dirige a encontrar un pavimento mejorado más resistente al deterioro.

Para desarrollar una mezcla superior de asfalto se requiere de un método confiable para probar la resistencia adhesiva de la mezcla experimental, por lo que se han desarrollado varios métodos para preparar las pruebas de resistencia en especímenes de pavimento asfáltico.

Se sabe que dos factores tienen efecto sobre la resistencia del espécimen: 1) los métodos de compactación de los especímenes durante la construcción y 2) el tipo de mezclado usado al unir el asfalto. Si en las pruebas de resistencia dos métodos de compactación producen los mismos resultados relativos con dos tipos de mezclado diferentes, entonces se puede usar cualquier método para evaluar las mezclas experimentales para cualquier tipo de mezclado. Si los resultados dependen del tipo de mezclado, entonces uno o ambos métodos, pueden ser inadecuados para discriminar entre las mezclas experimentales de asfalto.

El diseño factorial se puede usar para evaluar si la actuación de los dos factores es independiente para la resistencia de los especímenes de prueba. En la tabla 6.1 se ilustra el arreglo factorial para los especímenes de prueba preparados con dos métodos de compactación (estático y amasado) usando dos tipos de mezcla (roca silícea y basalto) en cada uno. Los valores del coeficiente de ruptura de los especímenes de adhesivo se ilustran en la tabla 6.1 para los cuatro tratamientos, en términos de presión (psi) al momento de la falla.

Los **factores** son los tipos de tratamiento tales como el método de compactación y el tipo de mezcla, y las distintas categorías de un factor son sus niveles, es decir, para el método de compactación son estático y amasado y para el tipo de mezcla son roca silícea y basalto. En la tabla 6.1, los factores se identifican con las letras mayúsculas A y B , los niveles de los factores se denotan por $A_1, A_2, \dots; B_1, B_2, \dots$; y así sucesivamente. El arreglo factorial de la tabla 6.1 con dos factores A y B cada uno con dos niveles tiene $2 \times 2 = 4$ combinaciones de tratamiento, A_1B_1, A_1B_2, A_2B_1 y A_2B_2 .

Tabla 6.1 Coeficiente de ruptura de los especímenes de asfalto, en psi

Tipo de mezcla (A)	<i>Método de compactación (B)</i>		Media de la mezcla
	(B_1) Estático	(B_2) Amasado	
Silícea (A_1)	68	60	64.0
Basalto (A_2)	65	97	81.0
Medias de compactación	66.5	78.5	

Fuente: A. M. Al-Marshed (1981), Compaction effects on asphaltic concrete durability. M. S. Thesis, Civil Engineering, University of Arizona.

El diseño factorial consiste en realizar todas las combinaciones posibles de los niveles de varios factores. Con frecuencia, los experimentos con diseños factoriales se conocen como *factoriales* o *experimentos factoriales*.

Los niveles de un factor *cuantitativo* toman valores métricos, mientras que los niveles de un factor *cualitativo* son las categorías del factor. Los dos factores del ejemplo 6.1 son cualitativos y los dos niveles de cada factor son categorías. La temperatura ilustra un factor cuantitativo con niveles de 10°C, 20°C y 30°C, por ejemplo.

6.2 Tres efectos de los factores

El **efecto** de un factor es un cambio en la respuesta medida ocasionado por un cambio en el nivel de ese factor, los tres efectos de interés en un experimento factorial son los *simples*, los *principales* y los *de interacción*; Estos efectos se ilustran con las medias de población para los tratamientos factoriales del ejemplo 6.1.

Las medias de población para un experimento factorial con dos factores, *A* y *B*, se pueden representar con las medias de las celdas μ_{ij} . El término *media de la celda* se deriva de un arreglo tabulado de las medias para cada combinación de tratamientos como se ilustra en la tabla 6.2 para un factorial de 2×2 .

Las medias de las combinaciones de tratamientos, $\mu_{11}, \mu_{12}, \mu_{21}$ y μ_{22} , se localizan en las celdas de tabla, de ahí el nombre de *medias de las celdas*. Las medias en los márgenes de la tabla son los promedios de las medias de las celdas y se conocen como *medias marginales*. La gran media o media global es el promedio de las medias de las celdas, $\bar{\mu}_+ = \frac{1}{4}(\mu_{11} + \mu_{12} + \mu_{21} + \mu_{22})$.

Tabla 6.2 Tabla de medias para un experimento factorial de 2×2

		<i>B</i>		Medias del factor <i>A</i>
<i>A</i>		1	2	
1		μ_{11}	μ_{12}	$\bar{\mu}_1 = \frac{1}{2}(\mu_{11} + \mu_{12})$
	2	μ_{21}	μ_{22}	$\bar{\mu}_2 = \frac{1}{2}(\mu_{21} + \mu_{22})$
Medias del factor <i>B</i>		$\bar{\mu}_1$	$\bar{\mu}_2$	
		$= \frac{1}{2}(\mu_{11} + \mu_{21}) = \frac{1}{2}(\mu_{12} + \mu_{22})$		

Los efectos simples son contrastes

Los **efectos simples** de un factor son las comparaciones entre los niveles de un factor a un solo nivel de otro. El efecto simple (l_1) del tipo de mezclado (*A*) sobre el coeficiente de ruptura con compactación estática (B_1), calculado a partir de las medias de las celdas de la tabla 6.1 es:

$$l_1 = \mu_{21} - \mu_{11} = 65 - 68 = -3$$

y mide la diferencia en el coeficiente de ruptura entre los especímenes de basalto y de roca silícea cuando se usa el método de compactación estático para formar los

especímenes. La fuerza de tensión promedio de los especímenes de roca silícea fue mayor que la de los especímenes de basalto. De manera similar, el efecto simple:

$$l_2 = \bar{\mu}_{22} - \bar{\mu}_{12} = 97 - 60 = 37$$

mide la diferencia del coeficiente de ruptura entre los especímenes de basalto y roca silícea cuando se usa el método de amasado. En este caso, el coeficiente de ruptura promedio de los especímenes de basalto fue mayor que el de los especímenes de roca silícea.

Los efectos principales son los efectos promedio de un factor

Los **efectos principales** de un factor son comparaciones entre los niveles de un factor promediados para todos los niveles de otro factor. El efecto principal del tipo de mezclado sobre el coeficiente de ruptura es la diferencia entre las medias marginales para el tipo de mezcla en la tabla 6.2:

$$l_3 = \bar{\mu}_{2.} - \bar{\mu}_{1.} = 81 - 64 = 17 \quad (6.1)$$

Cuando se promedian sobre ambos métodos de compactación la diferencia del coeficiente de ruptura entre los especímenes de basalto y roca silícea es de 17 psi en favor del basalto.

Después de una inspección cuidadosa, el efecto principal para el tipo de mezclado se puede expresar como el promedio de los dos efectos. Así, de la ecuación (6.1) se tiene:

$$\begin{aligned} l_3 &= \bar{\mu}_{2.} - \bar{\mu}_{1.} = \frac{1}{2} (\mu_{21} + \mu_{22}) - \frac{1}{2} (\mu_{11} + \mu_{12}) \\ &= \frac{1}{2} (\mu_{22} - \mu_{12}) + \frac{1}{2} (\mu_{21} - \mu_{11}) \\ &= \frac{1}{2} (l_1 + l_2) \end{aligned}$$

o $17 = \frac{1}{2}(-3 + 37)$. El contraste de efecto principal, $l_3 = 17$, implica que los especímenes de basalto son más fuertes que los de roca silícea, lo que se corrobora con comparación del efecto simple para la compactación por amasado, $l_2 = 37$; sin embargo, el contraste del efecto simple para la compactación estática, $l_1 = -3$, sugiere la conclusión opuesta. La diferencia entre los dos efectos simples indica que el tipo de mezclado y el método de compactación no actúan en forma independiente en cuanto a su influencia sobre la fuerza del espécimen.

Los efectos de interacción son las diferencias entre los efectos simples

Los efectos de interacción miden las diferencias entre los efectos simples de un factor a diferentes niveles de otro. Consideremos los dos efectos simples del tipo de mezclado sobre el coeficiente de ruptura mostrado en el cuadro 6.1.

La diferencia entre los dos efectos simples, $l_4 = l_2 - l_1$, mide la interacción entre los factores del tipo de mezclado y del método de compactación según afectan el coeficiente de ruptura de los especímenes. La diferencia entre los especímenes de basalto y roca silícea fue de 40 unidades más con el fraguado amasado que con el estático.

Cuadro 6.1 Efecto de la interacción entre el método de compactación y el tipo de mezclado

<i>Método de compactación</i>	<i>Efecto simple del tipo de mezclado</i>
Estático	$I_1 = 65 - 68 = -3$
Amasado	$I_2 = 97 - 60 = 37$
Diferencia	$I_4 = 37 - (-3) = 40$

Así, el efecto del tipo de mezclado sobre las medidas de resistencia del asfalto depende del método empleado para compactar el espécimen de prueba.

El ejemplo ilustra la precaución que debe tenerse al interpretar los efectos principales, puesto que en esta situación no debe basarse en ellos. El efecto del tipo de mezcla sobre la resistencia de los especímenes difiere considerablemente entre los métodos de compactación estático y por amasado, aunque la medida del efecto principal sugiere que los especímenes de basalto son más fuertes, esto sólo es cierto para el fraguado estático.

El contraste de la interacción factorial de 2×2 se deriva de la diferencia entre los dos efectos simples del factor A .

$$\begin{aligned} AB &= (\mu_{22} - \mu_{12}) - (\mu_{21} - \mu_{11}) \\ &= (\mu_{22} - \mu_{21} - \mu_{12} + \mu_{11}) \end{aligned} \quad (6.2)$$

Es posible obtener la misma expresión para los efectos simples del factor B .

En las figuras 6.1 y 6.2 se ilustra de manera gráfica la presencia o ausencia de interacción para un arreglo con dos factores, A y B , cada uno con dos niveles; en la figura 6.1 se muestra la medición gráfica de un solo efecto de cada uno de los factores. La respuesta causada por A se grafica por separado para cada nivel de B , pero cuando no hay interacción las líneas son paralelas, mostrando que existe una respuesta idéntica al factor A en ambos niveles de B . En este caso, los factores son independientes y los efectos principales pueden usarse para interpretar por separado los efectos de ambos factores.

En la figura 6.2 se muestra la presencia de la interacción, entre los 2 factores, aquí las líneas de respuesta no son paralelas. Las diferencias en la magnitud (figura 6.2a) o en la dirección (figura 6.2b) de las repuestas representan la interacción entre los factores, éstos no actúan de forma independiente y las interpretaciones deben basarse en los contrastes de efectos simples.

Ejemplo 6.2 Nuevos efectos del fraguado sobre el concreto asfáltico

Objetivos de investigación: Se sabe que la variación en la resistencia de los especímenes de prueba, según el ejemplo 6.1, está asociada con el método de fraguado y de mezcla usados para construirlos. Un ingeniero civil realizó un experimento para calificar las diferencias entre un conjunto de métodos de fraguado conforme sus efectos sobre la resistencia de los especímenes de prueba y para determinar hasta qué grado el tipo de mezcla afecta las comparaciones entre los métodos de fraguado.

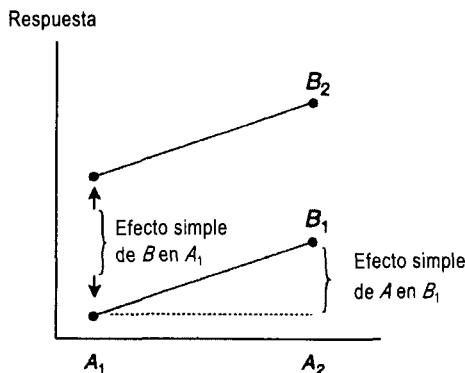


Figura 6.1 Ilustración de no interacción en un arreglo factorial

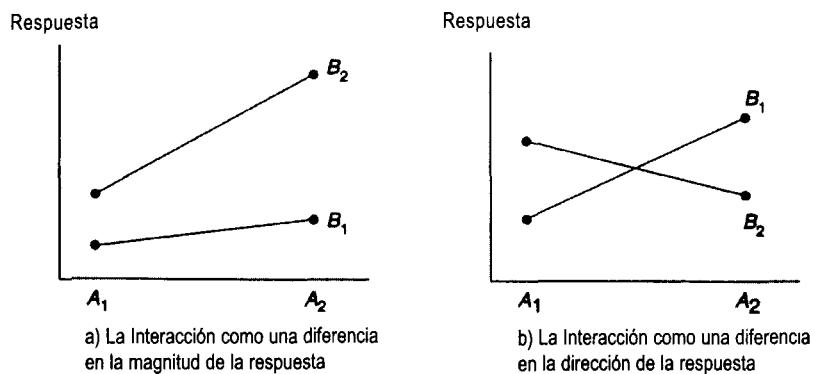


Figura 6.2 Ilustración de interacción en un arreglo factorial

Diseño de tratamientos: Se usó un arreglo factorial con los factores “método de fraguado” y “tipo de mezclado”. Existen dos niveles de tipo de mezcla — A_1 (basalto) y A_2 (roca silícea)— y cuatro niveles de método de fraguado — C_1 (presión estática), C_2 (amasado normal), C_3 (amasado lento) y C_4 (amasado muy lento).

Diseño del experimento: Se construyeron tres réplicas de especímenes y se probaron las ocho combinaciones. Los 24 especímenes se prepararon y probaron en orden aleatorio para un diseño totalmente aleatorizado.

Los datos para las medidas del coeficiente de ruptura (psi) de los 24 especímenes se muestran en la tabla 6.3.

Tabla 6.3 Coeficiente de ruptura (psi) de los especímenes de concreto asfáltico, para dos tipos de mezcla con cada uno de los cuatro métodos de fraguado

Tipo de mezcla	Estático	Método de fraguado		
		Normal	Lento	Muy lento
Basalto	68	126	93	56
	63	128	101	59
	65	133	98	57
Silícea	71	107	63	40
	66	110	60	41
	66	116	59	44

Fuente: A. M. Al Marshed (1981), Compaction effects on asphaltic concrete durability. M. S. Thesis, Civil Engineering, University of Arizona.

6.3 Modelo estadístico para dos factores

Modelo de medias de las celdas

Las observaciones de un experimento factorial con dos factores, A con a niveles y B con b niveles, se pueden representar con el **modelo de medias de celdas**. Este modelo para el factor de $a \times b$ con r réplicas, en un diseño totalmente aleatorizado es:

$$y_{ijk} = \mu_{ij} + e_{ijk} \quad (6.3)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

donde μ_{ij} es la media de la combinación de los tratamientos $A_i B_j$ y e_{ijk} son los errores experimentales aleatorios con media 0 y varianza σ^2 .

Estimaciones de las medias de celdas con el método de mínimos cuadrados

Las medias de las celdas se pueden estimar por el método de mínimos cuadrados descrito en el capítulo 2. La suma de cuadrados para el error experimental es

$$SS\ Error = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r \widehat{e}_{ijk}^2 = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \widehat{\mu}_{ij})^2 \quad (6.4)$$

Los estimadores de mínimos cuadrados para μ_{ij} son las medias de celdas observadas del las combinaciones del tratamiento.

(6.5)

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b$$

Las medias marginales observadas son estimaciones no sesgadas de las medias marginales del factor, de manera que $\hat{\mu}_i = \bar{y}_i$ y $\hat{\mu}_j = \bar{y}_j$; la media global $\bar{\mu}$ se estima con la gran media observada $\bar{y}_{..}$, las medias de celdas y marginales observadas para los especímenes de concreto asfáltico se muestran en la tabla 6.4.

Tabla 6.4 Medias de celdas y marginales estimadas para la resistencia de los especímenes de concreto asfáltico

Tipo de mezcla	Método de fraguado				Medias del tipo de mezcla (\bar{y}_i)	
	Amasado					
	Estático	Normal	Lento	Muy lento		
Basalto	65.3	129.0	97.3	57.3	87.3	
Silícea	67.7	111.0	60.7	41.7	70.3	
Medias del fraguado (\bar{y}_j)	66.5	120.0	79.0	49.5	$\bar{y}_{..} = 78.8$	

Aditividad y efectos de los factores

Las medias de celdas μ_{ij} representan la verdadera respuesta para la combinación del nivel i de A y el nivel j de B . Cuando no hay interacción, las medidas de celdas pueden expresarse como la suma de una media general, $\bar{\mu}$, más un efecto que es contribución de A , digamos α_i , y un efecto contribución de B , digamos β_j , de manera que $\mu_{ij} = \bar{\mu} + \alpha_i + \beta_j$.

El efecto para el nivel i -ésimo del factor A se puede definir como $\alpha_i = \bar{\mu}_i - \bar{\mu}$. El efecto de A es la desviación que tiene la media marginal de la gran media. El efecto para el nivel j -ésimo de B se puede definir como $\beta_j = \bar{\mu}_j - \bar{\mu}$. Los efectos serán **efectos fijos** si los niveles de los factores son reproducibles. En ausencia de interacción, la media de celda es la suma de la gran media y los efectos de los factores en esa celda:

$$\mu_{ij} = \bar{\mu} + (\bar{\mu}_i - \bar{\mu}) + (\bar{\mu}_j - \bar{\mu}) \quad (6.6)$$

El efecto de la ij -ésima combinación, $\tau_{ij} = (\mu_{ij} - \bar{\mu})$, es la suma de los efectos de ambos factores:

$$(\mu_{ij} - \bar{\mu}) = (\bar{\mu}_i - \bar{\mu}) + (\bar{\mu}_j - \bar{\mu}) \quad (6.7)$$

y los efectos de los factores son *aditivos* en ausencia de interacción.

En presencia de interacción, el efecto del tratamiento no será igual a la suma de los efectos individuales de los factores, como se muestra en la ecuación (6.7). Un efecto de interacción, denotado como $(\alpha\beta)_{ij}$, puede definirse como la diferencia entre los dos lados de la ecuación (6.7), es decir:

$$\begin{aligned} (\alpha\beta)_{ij} &= (\mu_{ij} - \bar{\mu}_{..}) - (\bar{\mu}_i - \bar{\mu}_{..}) - (\bar{\mu}_j - \bar{\mu}_{..}) \\ &= \mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \bar{\mu}_{..} \end{aligned} \quad (6.8)$$

El nuevo conjunto de parámetros, α_i , β_j y $(\alpha\beta)_{ij}$, puede usarse para escribir el modelo lineal del factorial como un modelo de *efectos*:

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \\ i &= 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r \end{aligned} \quad (6.9)$$

donde μ es la gran media $\bar{\mu}_{..}$, α_i es el efecto del i -ésimo nivel de A , β_j es el efecto del j -ésimo nivel de B , y $(\alpha\beta)_{ij}$ es el efecto de interacción entre el i -ésimo nivel de A y el j -ésimo de B . Por la naturaleza de sus definiciones, las sumas de los efectos son iguales a cero, esto es,

$$\sum_{i=1}^a \alpha_i = 0, \quad \sum_{j=1}^b \beta_j = 0, \quad \sum_{i=1}^a (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = 0 \quad (6.10)$$

6.4 Análisis para dos factores

Partición de la suma de cuadrados

La partición de la suma de cuadrados total se puede deducir de la ecuación:

$$(y_{ijk} - \bar{y}_{..}) = (\bar{y}_{ij} - \bar{y}_{..}) + (y_{ijk} - \bar{y}_{ij}) \quad (6.11)$$

La desviación que una observación tiene respecto a la gran media $(y_{ijk} - \bar{y}_{..})$ es la suma de dos partes:

- el efecto del tratamiento $(\bar{y}_{ij} - \bar{y}_{..})$
- el error experimental $(y_{ijk} - \bar{y}_{ij})$

Por ejemplo, si se usan las observaciones de la tabla 6.3 y las medias de la tabla 6.4, la desviación total para la primera observación en la tabla 6.3 es:

$$(y_{111} - \bar{y}_{..}) = 68 - 78.8 = -10.8.$$

La desviación del tratamiento es:

$$(\bar{y}_{11.} - \bar{y}_{..}) = (65.3 - 78.8) = -13.5$$

y el error experimental es:

$$(y_{111} - \bar{y}_{11.}) = (68 - 65.3) = 2.7$$

y

$$-10.8 = -13.5 + 2.7$$

La última desviación ($y_{ijk} - \bar{y}_{ij\cdot}$) es una medida del error experimental para la observación en un experimento con réplicas adecuadas. Si ambos lados de la ecuación (6.11) se elevan al cuadrado y sumados para todas las observaciones, el resultado es

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y}_{ij\cdot})^2 &= r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 \\ &\quad + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y}_{ij\cdot})^2 \end{aligned} \quad (6.12)$$

o:

$$SC \text{ total} = SC \text{ tratamiento} + SC \text{ error}$$

La suma de cualquier producto cruzado formado al elevar al cuadrado el lado derecho de la ecuación (6.11) es cero. Existe un total de rab observaciones, de manera que la SC total tiene $(rab - 1)$ grados de libertad; con $t = ab$, la SC tratamiento tiene $(ab - 1)$ grados de libertad y los $ab(r - 1)$ grados de libertad restantes se asocian con SC error.

Suma de cuadrados para los efectos de los factores

El efecto de la ecuación (6.11) se puede expresar como la identidad:

$$(\bar{y}_{ij\cdot} - \bar{y}_{\cdot\cdot\cdot}) = (\bar{y}_i - \bar{y}_{\cdot\cdot\cdot}) + (\bar{y}_j - \bar{y}_{\cdot\cdot\cdot}) + (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdot\cdot\cdot}) \quad (6.13)$$

donde el efecto del tratamiento es una suma de tres efectos:

- el efecto principal del factor A , $(\bar{y}_i - \bar{y}_{\cdot\cdot\cdot})$
- el efecto principal del factor B , $(\bar{y}_j - \bar{y}_{\cdot\cdot\cdot})$
- la interacción $(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdot\cdot\cdot})$

Por ejemplo, el efecto del tratamiento para el basalto con amasado lento de la tabla 6.4 es:

$$(\bar{y}_{13} - \bar{y}_{\cdot\cdot\cdot}) = 97.3 - 78.8 = 18.5$$

Los efectos principal y de interacción son:

- efecto principal del basalto:

$$(\bar{y}_1 - \bar{y}_{\cdot\cdot\cdot}) = 87.3 - 78.8 = 8.5$$

- efecto principal del amasado lento:

$$(\bar{y}_3 - \bar{y}_{\cdot\cdot\cdot}) = 79.0 - 78.8 = 0.2$$

- interacción

$$(\bar{y}_{13.} - \bar{y}_{1.} - \bar{y}_{3.} + \bar{y}_{..}) = 97.3 - 87.3 - 79.0 + 78.8 = 9.8$$

y

$$18.5 = 8.5 + 0.2 + 9.8$$

Si ambos lados de la ecuación (6.13) se elevan al cuadrado y sumados para todas las observaciones, el lado izquierdo es el la tratamiento SC , ya que se hace una partición de la suma de cuadrados del tratamiento en las 3 componentes representadas por los efectos en el lado derecho de la ecuación (6.13). La suma de cualquier producto cruzado formado por el cuadrado del lado derecho es cero.

La suma de cuadrados de la primera componente será la suma de cuadrados entre las medias marginales de A :

$$SCA = rb \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 \quad (6.14)$$

y la segunda, será la suma de cuadrados entre las medias marginales de B

$$SCB = ra \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{...})^2 \quad (6.15)$$

La suma de cuadrados de la tercera componente

$$SC(AB) = r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})^2 \quad (6.16)$$

es la suma de cuadrados de la interacción. Como esa parte de la suma de cuadrados del tratamiento no se explica mediante la suma de cuadrados atribuida a los dos factores SCA y SCB , la partición aditiva de SC tratamiento es:

$$SC \text{ tratamiento} = SCA + SCB + SC(AB)$$

Los $(ab - 1)$ grados de libertad de la suma de cuadrados del tratamiento se asignan a las tres particiones del tratamiento SC . Los factores A y B tienen a y b niveles, respectivamente; los grados de libertad restantes asignados a la suma de cuadrados de la interacción son los grados de libertad de los tratamientos $(ab - 1)$, menos los grados de libertad de las sumas de cuadrados de los factores separados $(a - 1)$ y $(b - 1)$, es decir, $(ab - 1) - (a - 1) - (b - 1) = (a - 1)(b - 1)$. Los grados de libertad para las sumas de cuadrados de la interacción en los factoriales es el producto de los grados de libertad de los factores, incluida la interacción.

Para un arreglo factorial con dos factores, la partición completa de la suma de cuadrados total se resume en el análisis de varianza que se muestra en la tabla 6.5.

La obtención de las particiones de la suma de cuadrados a partir de las soluciones de las ecuaciones normales de mínimos cuadrados del modelo se muestra en el apéndice 6A.

Tabla 6.5 Análisis de varianza para un diseño de tratamientos con dos factores

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Cuadrados medios esperados
Total	$rab - 1$	SC total		
Factor A	$a - 1$	SCA	CMA	$\sigma_e^2 + rb\theta_a^2$
Factor B	$b - 1$	SCB	CMB	$\sigma_e^2 + ra\theta_b^2$
Interacción AB ($(a - 1)(b - 1)$)		$SC(AB)$	$CM(AB)$	$\sigma_e^2 + r\theta_{ab}^2$
Error	$ab(r - 1)$	SCE	CME	σ_e^2

$\theta_a^2 = \sum_{i=1}^a (\bar{\mu}_i - \bar{\mu})^2 / (a - 1)$ $\theta_b^2 = \sum_{j=1}^b (\bar{\mu}_j - \bar{\mu})^2 / (b - 1)$
 $\theta_{ab}^2 = \sum_{i=1}^a \sum_{j=1}^b (\mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \bar{\mu}_{..})^2 / (a - 1)(b - 1)$

Tabla 6.6 Análisis de varianza del coeficiente de ruptura para los especímenes de asfalto, en un arreglo factorial de 4×2

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	23	19 274.50			
Fraguado (C)	3	16 243.50	5 414.50	569.95	.000
Mezcla (A)	1	1 734.00	1 734.00	182.53	.000
Interacción (AC)	3	1 145.00	381.67	40.18	.000
Error	16	152.00	9.50		

el análisis de varianza para los datos de la tabla 6.3 acerca de los especímenes de concreto asfáltico se muestra en la tabla 6.6.

Pruebas de hipótesis sobre el efecto de los factores

Las inferencias sobre los efectos de factores individuales depende de la presencia o ausencia de interacción. Primero se determina la significancia de la interacción antes de establecer cualquier significancia para los efectos principales de los factores.

Si no existe interacción la ecuación (6.8) proporciona $(\alpha\beta)_{ij} = \mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \bar{\mu}_{..} = 0$ y $\theta_{ab}^2 = 0$ es el cuadrado medio esperado para la interacción AB . La hipótesis nula para la interacción:

$$H_0: (\alpha\beta)_{ij} = \mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \bar{\mu}_{..} = 0 \text{ para toda } i, j$$

contra la alternativa:

$$H_a: (\alpha\beta)_{ij} = \mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \bar{\mu}_{..} \neq 0 \text{ para toda } i, j$$

se prueban contra:

$$F_0 = \frac{CM(AB)}{CME} \quad (6.17)$$

con valor crítico $F_{\alpha,(a-1)(b-1),ab(r-1)}$.

Si no hay diferencias entre las medias marginales para A , entonces $\alpha_i = \bar{\mu}_i - \bar{\mu} = 0$ y $\theta_a^2 = 0$ es el cuadrado medio para A . La hipótesis nula de igualdad entre las medias:

$$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_a$$

contra la hipótesis alternativa:

$$H_a: \bar{\mu}_i \neq \bar{\mu}_k = 0 \text{ para alguna } i, k$$

se prueba contra la razón:

$$F_0 = \frac{CMA}{CME} \quad (6.18)$$

con valor crítico $F_{\alpha,(a-1),ab(r-1)}$.

Si no hay diferencias entre las medias marginales para B , entonces $\beta_j = \bar{\mu}_j - \bar{\mu} = 0$ y $\theta_b^2 = 0$ es el cuadrado medio esperado para B . La hipótesis nula de igualdad entre medias:

$$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_b$$

contra la alternativa:

$$H_a: \bar{\mu}_j \neq \bar{\mu}_m \text{ para alguna } j, m$$

se prueba con la razón:

$$F_0 = \frac{CMB}{CME} \quad (6.19)$$

con valor crítico $F_{\alpha,(b-1),ab(r-1)}$.

Pruebas F para los efectos del tipo de mezcla y del método de fraguado

La prueba F para la interacción, $F_0 = CM(AC)/CME = 381.67/9.50 = 40.18$ en la tabla 6.6, indica una interacción significativa entre el tipo de mezcla y el método de fraguado, ya que F_0 excede $F_{0.05,3,16} = 3.24$. Las medias marginales para el tipo de mezcla son significativamente diferentes, puesto que $F_0 = CMA/CME = 1734.00/9.50 = 182.53$ excede $F_{0.05,1,16} = 4.49$. Las medias marginales para el método de fraguado también son diferentes, pues $F_0 = CMC/CME = 5414.50/9.50 = 569.95$ excede $F_{0.05,3,16} = 3.24$. El nivel de significancia para cada prueba se da como $Pr > F = .000$ en la columna derecha de la tabla 6.6.

Una interacción significativa puede modificar cualquier inferencia basada en las diferencias significativas entre las medias marginales de la mezcla y el fraguado.

do, por lo que el compendio de las medias marginales y de celdas que se encuentra en la tabla 6.4 será útil para interpretar los resultados.

Errores estándar e intervalos estimados para las medias

Los errores estándar para las medias estimadas de celdas y marginales del experimento factorial son:

$$\begin{aligned} \text{Mezcla: } s_{\bar{y}_i} &= \sqrt{\frac{CME}{rb}} = \sqrt{\frac{9.5}{(3)(4)}} = 0.89 \\ \text{Fraguado: } s_{\bar{y}_{J_i}} &= \sqrt{\frac{CME}{ra}} = \sqrt{\frac{9.5}{(3)(2)}} = 1.26 \\ \text{Medias de celdas: } s_{\bar{y}_{ij}} &= \sqrt{\frac{CME}{r}} = \sqrt{\frac{9.5}{3}} = 1.78 \end{aligned} \quad (6.20)$$

Se requiere la t de Student con $ab(r - 1)$ grados de libertad para las estimaciones de los intervalos de las medias marginales y de celdas, un intervalo de confianza estimado del 95% requiere $t_{025,16} = 2.12$ y la estimación del intervalo para una media de celda es:

$$\bar{y}_i \pm t_{025,16}(s_{\bar{y}_{ij}}) \quad (6.21)$$

por ejemplo, un intervalo de confianza estimado del 95% para la media de una mezcla de basalto con fraguado estático es:

$$\begin{aligned} 65.3 &\pm 2.12(1.78) \\ &(61.5, 69.1) \end{aligned}$$

Los intervalos estimados se calculan de manera similar para las otras medias después de sustituir en la ecuación (6.20) la media y la estimación del error estándar adecuados.

Las comparaciones múltiples ayudan a interpretar los efectos significativos de la interacción

La interacción significativa entre el tipo de mezcla y el método de fraguado indica que los efectos simples de un factor difieren entre los niveles del otro factor. En consecuencia, las pruebas de hipótesis sobre los factores deben basarse inicialmente en los contrastes de los efectos simples con las medias de celdas.

Las hipótesis de investigación específicas para el estudio determinan los contrastes entre las medias de celdas necesarias para investigar los efectos simples; una pregunta general de investigación para este estudio podría inquirir qué tipo de mezcla tiene como resultado los especímenes más resistentes para cada método de fraguado. Quizá otra hipótesis se refiera al efecto de los métodos de fraguado por amasado en comparación con el fraguado estático.

Se pueden usar las comparaciones entre las medias de celdas de la tabla 6.4, para probar los efectos simples del tipo de mezcla (basalto contra silícea) en cada nivel de fraguado, para encontrar cuál produce los especímenes más resistentes

con cada método de fraguado. Las cuatro comparaciones y sus intervalos del 95% de confianza simultáneos se muestran en el cuadro 6.2. La tasa de error con respecto al experimento para la familia de cuatro pruebas se puede controlar con el estadístico t de Bonferroni, $t_{0.025,4,16} = 2.81$ de la tabla V del apéndice.

Cuadro 6.2 Contrastes estimados para los efectos simples de la mezcla

<i>Método de fraguado</i>	<i>Contraste (basalto-siliceo)</i>	<i>ICS 95%</i>	
		<i>(L, U)</i>	
Estático	$c_1 = \bar{y}_{11} - \bar{y}_{21} = 65.3 - 67.7 = -2.4$		(-9.5, 4.7)
Regular	$c_2 = \bar{y}_{12} - \bar{y}_{22} = 129.0 - 111.0 = 18.0$		(10.9, 25.1)
Lento	$c_3 = \bar{y}_{13} - \bar{y}_{23} = 97.3 - 60.7 = 36.6$		(29.5, 43.7)
Muy lento	$c_4 = \bar{y}_{14} - \bar{y}_{24} = 57.3 - 41.7 = 15.6$		(8.5, 22.7)
Error estándar	$s_c = \sqrt{\frac{CME}{r}} [1^2 + (-1)^2] \sqrt{\frac{9.5}{3}} (2) = 2.52$		

Exposición resumida de los efectos de los factores

Con el fraguado estático, no existe diferencia del coeficiente de ruptura entre los especímenes de los dos tipos de mezcla, ya que el intervalo incluye el 0. Para los métodos de fraguado con amasado, los especímenes construidos con basalto tienen mayor resistencia que los construidos con roca silícea, puesto que los límites inferiores de los 3 intervalos son mayores que 0 y la diferencia más grande entre los tipos de mezcla se encontró con la fraguado por amasado lento.

Las comparaciones a considerar dependen de la naturaleza del problema y de la información requerida por el estudio, por ello las presentadas en los párrafos anteriores se hicieron para ilustrar la interpretación de la interacción. Explotan el concepto de que las inferencias relativas a comparaciones entre los tipos de mezcla o entre los métodos de fraguado dependen de otros factores.

Una comparación entre las medias marginales para un factor pueden proporcionar información cuando todos sus efectos simples tienen dirección y magnitud similares, al resultado de las comparaciones entre medias marginales es una inferencia más general sobre el factor y son más precisas con errores estándar más pequeños; vea la ecuación (6.20). En presencia de interacción, debe tenerse cuidado al generalizar. La diferencia entre las medias marginales para los tipos de mezcla de basalto y roca silícea, $c = (87.3 - 70.3) = 17.0$, con error estándar $s_c = \sqrt{2CME/rb} = \sqrt{2(9.5)/12} = 1.26$, será significativa. La estimación del efecto principal es una diferencia promediada de todos los métodos de fraguado. Los efectos simples mostrados en el resumen 6.2 para los tipos de mezcla van de $c_1 = -2.4$ a $c_3 = 36.6$, según el método de fraguado. Las generalizaciones sobre los tipos de mezcla, basadas en las estimaciones de los efectos principales, estarian equivocadas.

Análisis residual para evaluar las suposiciones

Las suposiciones para el modelo en cuanto a la homogeneidad de las varianzas y la distribución normal, se pueden evaluar con los residuales estudiados en el capítulo 4. Los residuales para el factorial con dos factores se calculan como la desviación de los valores observados de las medias estimadas para cada celda en el arreglo; el residual de cualquier celda es $\hat{e}_{ijk} = y_{ijk} - \bar{y}_{jk} - \bar{y}_i$. Por ejemplo, el residual de la primera observación en la tabla 6.3 es

$$\hat{e}_{111} = y_{111} - \bar{y}_{11\cdot} = 68 - 65.3 = 2.7$$

Las gráficas de la raíz cuadrada de los residuales absolutos contra los valores estimados y la gráfica de la probabilidad normal de los residuales se muestran en las figuras 6.3 a y b, no muestran una evidencia fuerte de varianza heterogénea o no normalidad. La prueba Levene (Med) para la homogeneidad de la varianza se deja como ejercicio para el lector con referencia a los métodos del capítulo 4 para evaluar las suposiciones.

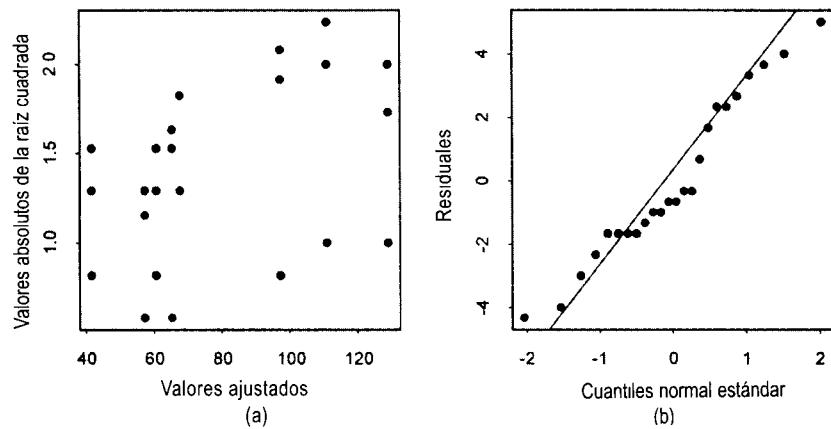


Figura 6.3 Gráficas residuales del análisis de varianza para los datos de resistencia del ejemplo 6.2: a) raíz cuadrada de los residuales absolutos contra valores estimados y b) gráfica de la probabilidad normal de los residuales

6.5 Uso de curvas de respuesta para los factores de tratamiento cuantitativos

Las curvas de tendencia de respuesta para los factores cuantitativos se estimaron en la sección 6.6 con polinomiales ortogonales. Recuerde que la curva de respuesta estimada tiene la ventaja de describir la relación entre la variable de respuesta y el factor de tratamiento en el intervalo numérico del factor usado en el estudio. La evaluación de las curvas de tendencia de respuesta se extiende al experimento con dos factores en esta sección; en principio, se analiza el experimento con un factor

cuantitativo, después se analizan las respuestas para experimentos con dos factores cuantitativos.

Un factor cuantitativo y un factor cualitativo

Asegurar si existe o no interacción entre un factor cuantitativo y uno cualitativo es el primer objetivo del análisis. Cuando dos factores interactúan la respuesta al factor cuantitativo será diferente para los distintos niveles del factor cualitativo, en ese caso, las curvas de respuesta para el factor cuantitativo se puede estimar por separado para cada nivel del factor cualitativo. Cuando no existe interacción, la tendencia de la respuesta al factor cuantitativo será similar en todos los niveles del factor cualitativo y será suficiente una curva de respuesta simple para describir el proceso respecto al factor cuantitativo.

Ejemplo 6.3 Metales pesados en lodos de desagüe

El lodo de desagüe es el residuo seco que resulta de procesar las aguas negras; como contiene nutrientes benéficos para el crecimiento de plantas, se puede usar como fertilizante en la agricultura, siempre que no contenga niveles tóxicos de ciertos elementos como metales pesados. Por regla general, los niveles de metales en los lodos se prueban según el crecimiento de plantas en ambientes que contienen distintas dosis de lodo.

Hipótesis de investigación: un científico de suelos planteó la hipótesis de que la concentración de ciertos metales en los lodos difiere según las áreas metropolitanas de las que se obtuvo el lodo, variación que puede ser el resultado de una gran cantidad de causas, como las distintas bases industriales que rodean el área. Si esto fuera cierto, entonces las recomendaciones de aplicación en cultivos tendrían que ser precedidas por la ubicación de la fuente de material. Se planeó una prueba para determinar si había una variación significativa en las concentraciones de metales pesados entre las diversas áreas metropolitanas.

Diseño del tratamiento: el investigador obtuvo lodos de las plantas de tratamiento localizadas en tres áreas metropolitanas diferentes. Se cultivaron plantas de cebada en un medio de arena al que se agregó el lodo como fertilizante, en tres cantidades diferentes: 0.5, 1.0 y 1.5 toneladas métricas/acre. El arreglo factorial para el diseño del tratamiento consistió en un factor cualitativo, “ciudad”, con tres niveles y un factor cuantitativo, “cantidad”, con tres niveles.

Diseño del experimento: cada una de las nueve combinaciones de tratamientos se asignó a cuatro contenedores en un *diseño totalmente aleatorizado*. Los contenedores se colocaron al azar en una cámara de cultivo; en cierta etapa del crecimiento se determinó el contenido de Zinc, en partes por millón, para las plantas de cebada cultivadas en cada contenedor. Los datos se muestran en la tabla 6.7 y el análisis de varianza en la 6.8; los cálculos manuales para las sumas de cuadrados lineales y cuadráticas se muestran en la tabla 6.9.

Tabla 6.7 Contenido de zinc (ppm) en los cultivos de plantas de cebada en ambientes que contienen tres cantidades diferentes de lodo proveniente de tres áreas metropolitanas

<i>Ciudad y cantidad (Ton/hectárea)</i>								
<i>A</i>			<i>B</i>			<i>C</i>		
0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
26.4	25.2	26.0	30.1	47.7	73.8	19.4	23.2	18.9
23.5	39.2	44.6	31.0	39.1	71.1	19.3	21.3	19.8
25.4	25.5	35.5	30.8	55.3	68.4	18.7	23.2	19.6
22.9	31.9	38.6	32.8	50.7	77.1	19.0	19.9	21.9

Fuente: J. Budzinsky, Department of Soil and Water Science, University of Arizona.

Tabla 6.8 Análisis de varianza para el contenido de zinc en cultivos de plantas de cebada en ambientes que contienen tres cantidades diferentes de lodos de drenaje provenientes de tres áreas metropolitanas

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>Pr > F</i>
Total	35	9993.38			
Cantidad (<i>R</i>)	2	1945.45	972.72	50.71	.000
Cantidad lineal	1	1944.00	1944.00	101.35	.000
Cantidad cuadrática	1	1.45	1.45	0.08	.786
Ciudad (<i>C</i>)	2	5720.67	2860.34	149.13	.000
Cantidad × ciudad (<i>RC</i>)	4	1809.40	452.35	23.58	.000
Cantidad lineal × ciudad (<i>RC</i>)	2	1760.15	880.07	45.88	.000
Cantidad cuadrática × ciudad	2	49.25	24.63	1.28	.293
Error	27	517.86	19.18		

El análisis de varianza presentado en la tabla 6.8 indica una interacción significativa entre “cantidad” y “ciudad”, y los efectos principales significativos para ambos factores ($Pr > F = .000$). Los dos grados de libertad para la suma de cuadrados de la cantidad se dividen en 1 grado de libertad para cada una de las cantidades lineal y cuadrática, la prueba F indica significancia para la partición de regresión lineal ($F_0 = 101.35$) y no significancia ($F_0 = 0.08$) para la partición cuadrática para cantidad.

Los 2 grados de libertad para cada suma de cuadrados de interacción “cantidad por ciudad” indican la variabilidad entre ciudades en cuanto a los coeficientes de regresión lineal y cuadrática para la cantidad. La interacción entre la regresión lineal de la cantidad y la ciudad es significativa ($F_0 = 45.88$), pero la interacción entre la regresión cuadrática de cantidad y ciudad no lo es ($F_0 = 1.28$).

Interpretación de los efectos de factores con contrastes de regresión

La interacción significativa entre ciudad y la regresión lineal de la cantidad de lodo aplicado sugieren que la interpretación debe basarse en distintas líneas de regresión para cada ciudad. Estas líneas estimadas por ciudad se grafican

Tabla 6.9 Cálculo de la partición de las sumas de cuadrados de contraste lineal y cuadrático para la interacción entre cantidad y cantidad \times ciudad

Ciudad	Cantidad (Ton/hectárea)			Lineal	Cuadrática
	0.5	1.0	1.5	$\sum P_{1j}\bar{y}_{ij}$	$\sum P_{2j}\bar{y}_{ij}$
A	24.55	30.45	36.18	11.63	-0.17
B	31.18	48.20	72.60	41.42	7.38
C	19.10	21.90	20.05	0.95	-4.65
Medias (\bar{y}_{ij})	24.94	33.52	42.94	18.00	0.84
Lineal (P_{1j}) - 1	0	1			
Cuadrática (P_{2j})	1	-2	1		

$SC[R \text{ lineal}] = ra[\sum P_{1j}\bar{y}_{ij}]^2/\sum P_{1j}^2 = 12[18]^2/2 = 1944$
 $SC[R \text{ cuadrática}] = ra[\sum P_{2j}\bar{y}_{ij}]^2/\sum P_{2j}^2 = 12[0.84]^2/6 = 1.41$
 $SC[R \text{ lineal} \times C] = r\sum_i [\sum_j P_{1j}\bar{y}_{ij}]^2/\sum P_{1j}^2 - SC[R \text{ lineal}]$
 $= \frac{4[11.63^2 + 41.42^2 + 0.95^2]}{(2)} - SC[R \text{ lineal}]$
 $= 1760$
 $SC[R \text{ cuad} \times C] = r\sum_i [\sum_j P_{2j}\bar{y}_{ij}]^2/\sum P_{2j}^2 - SC[R \text{ cuad}]$
 $= \frac{4[(-0.17)^2 + 7.38^2 + (-4.65)^2]}{(6)} - SC[R \text{ cuad}]$
 $= 49$

en la figura 6.4 junto con las medias de celdas. La gráfica ilustra la cantidad(lineal) por interacción de ciudad. La respuesta a la cantidad es lineal para cada ciudad, la significancia de la interacción entre la ciudad y la partición lineal para la cantidad se muestra en la gráfica como una respuesta lineal del zinc diferente para la cantidad correspondiente a cada ciudad.

Los contrastes de regresión lineal para la cantidad son los efectos simples para las cantidades calculadas de cada ciudad. Se pueden calcular líneas de regresión lineal para cada ciudad a partir de los efectos estimados en la tabla 6.9, o con un programa de computadora para calcularla. La línea de regresión se puede calcular, de acuerdo con el procedimiento mostrado en la sección 3.3, usando las medias de celdas en la tabla 6.9. La estimación del coeficiente polinomial ortogonal para la ciudad A es:

$$\begin{aligned} a_1 &= \sum P_{1j}\bar{y}_{1j}/\sum P_{1j}^2 \\ &= \frac{-1(24.55) + 0(30.45) + 1(36.18)}{[(-1)^2 + 0^2 + 1^2]} = \frac{11.63}{2} = 5.8 \end{aligned}$$

La media de la ciudad A es $\bar{y}_1 = 30.39$. Con $\lambda_1 = 1$, $\bar{R} = 1.0$ y $d = 0.5$ (cuadro 3.5), la transformación a una ecuación en términos de la cantidad (R) es:

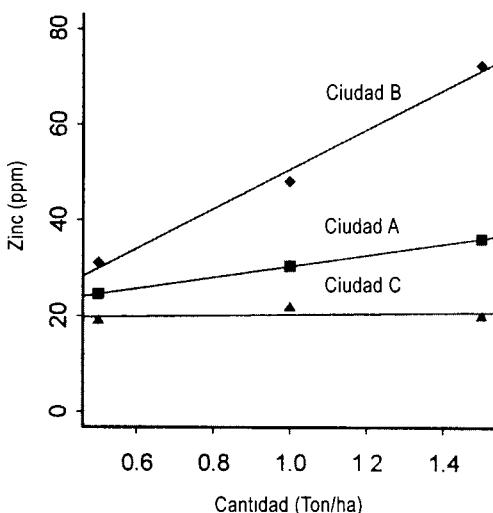


Figura 6.4 Contenido medio de zinc contra cantidad de lodo aplicada para tres ciudades

$$\begin{aligned}\hat{y} &= \bar{y}_1 + a_1 P_1 \\ &= 30.39 + 5.8 \frac{(R - 1.0)}{0.5} = 18.79 + 11.6(R)\end{aligned}$$

El contraste lineal para la ciudad C (0.95 en la tabla 6.9), es mucho más pequeña que las de las otras dos ciudades, 11.63 y 41.42 para A y B, respectivamente. El error estándar del contraste lineal de la cantidad, basado en las medias de celdas de la tabla 6.9, para una ciudad en particular es:

$$s_c = \sqrt{\frac{(CME)[(-1)^2 + 0^2 + 1^2]}{r}} = \sqrt{\frac{(19.18)[2]}{4}} = 3.1$$

Los intervalos de confianza simultáneos del 95% para los tres contrastes lineales requieren una t de Bonferroni, $t_{05,3,27} = 2.55$, de donde los respectivos ICS del 95% para las ciudades A, B y C son (3.73, 19.54), (33.5, 49.33) y (-6.96, 8.86). Las respuestas lineales para las ciudades A y B son significativamente positivas, donde la ciudad B tiene el contraste lineal positivo mayor, en tanto que el intervalo para la ciudad C incluye al 0, por lo que se puede concluir que el zinc se acumulará en mayor grado en los cultivos de cebada fertilizados con cantidades crecientes de lodos de la ciudad B y en menor grado con lodos de A, pero no habrá acumulación significativa con lodos de C.

Dos factores cuantitativos

La respuesta a dos factores cuantitativos se puede representar mediante una ecuación polinomial con dos variables independientes, en la que el grado del polinomio

dependerá del número de niveles de cada factor. Las ecuaciones de primer grado pueden representar un factor con dos niveles, las de segundo grado, uno con tres niveles, etcétera. La representación geométrica de una ecuación polinomial con dos variables independientes, es una superficie de respuesta en tres dimensiones, por ejemplo, si los niveles de los factores A y B se representan con dos variables métricas x_1 y x_2 en un polinomio cuadrático, para efectos de los estudios experimentales la ecuación polinomial de segundo grado

$$y = \beta_0 + \beta_1x_1 + \beta_2x_1^2 + \beta_3x_2 + \beta_4x_2^2 + \beta_5x_1x_2 \quad (6.22)$$

será una función empírica de uso común en la aproximación de una superficie de respuesta. La superficie cuadrática se puede explorar para los niveles del factor que se obtienen en la respuesta óptima o en distintas combinaciones de los niveles con respuestas equivalentes.

El análisis del experimento factorial con dos factores cuantitativos consiste de particiones polinomiales ortogonales de las sumas de cuadrados del efecto principal del factor y la interacción; la naturaleza de la función de respuesta polinomial se puede determinar a partir de estas particiones. Es posible usar una gráfica de respuesta como auxiliar en la interpretación del papel de cada factor en la respuesta.

Ejemplo 6.4 Agua consumida por las plantas de cebada

Los depósitos de sal se acumulan en los suelos irrigados para cultivos agrícolas y de horticultura. Con el paso del tiempo, la creciente salinidad del suelo impide el desarrollo de las plantas y disminuye las cosechas.

Hipótesis de investigación: Un investigador planteó la hipótesis de que la exposición de las plantas a elevadas cantidades de sales en su medio inhibe el consumo de agua y nutrientes de la planta, lo que impide su crecimiento y desarrollo. Se llevó a cabo un experimento con plantas de cebada para medir el efecto del aumento de la salinidad en la cantidad de agua consumida por las plantas.

Diseño del tratamiento: Se usó un arreglo factorial con “salinidad del medio” y “edad de la planta” en días, como los dos factores. Las plantas se cultivaron en soluciones de nutrientes con 3 niveles de salinidad que expresados como unidades de presión osmótica, fueron de 0, 6 y 12 barios y se cosecharon a los 14, 21 y 28 días.

Diseño del experimento: Cada una de las nueve combinaciones de tratamiento, de salinidad con días, se asignó a dos contenedores réplica en un diseño totalmente aleatorizado, los contenedores se colocaron en un cámara de cultivo en un arreglo al azar.

Una de las medidas tomadas al cosechar fue la cantidad de agua consumida por las plantas durante el experimento, expresada en milímetros de agua por cada 100 gramos de peso de la planta seca. Los datos se muestran en la tabla 6.10, el análisis de varianza en la 6.11, y los cálculos manuales de las particiones de sumas de cuadrados en la 6.12.

Notas sobre los cálculos: El arreglo de la tabla 6.12 es conveniente para los cálculos manuales de particiones de sumas de cuadrados a partir de las medias de celdas; en ella se pueden calcular la partición del efecto principal y

Tabla 6.10 Agua consumida (ml/100 g) por plantas de cebada a los 14, 21 y 28 días de crecimiento, en soluciones con niveles salinos de 0, 6 y 12 barios

Salinidad	0 barios			6 barios			12 barios		
Días	14	21	28	14	21	28	14	21	28
	2.2	5.0	13.2	3.7	5.9	9.4	2.8	4.5	7.6
	3.3	5.7	12.4	4.5	7.2	11.0	3.4	5.9	8.3
Medias (\bar{y}_j)	2.75	5.35	12.80	4.10	6.55	10.20	3.10	5.20	7.95

Fuente: Dr. T. C. Tucker, Department of Soil and Water Science, University of Arizona.

Tabla 6.11 Análisis de varianza del consumo de agua por plantas de cebada

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	17	184.73			
Salinidad (S)	2	9.51	4.75	8.52	.008
S lineal	1	7.21	7.21	12.92	.006
S cuadrática	1	2.30	2.30	4.12	.073
Días (D)	2	151.99	75.99	136.24	.000
D lineal	1	147.00	147.00	263.55	.000
D cuadrática	1	4.99	4.99	8.94	.015
Salinidad × días (SD)	4	18.21	4.55	8.16	.005
S lin × D lin	1	13.52	13.52	24.24	.001
S lin × D cuad	1	2.94	2.94	5.27	.047
S cuad × D lin	1	1.21	1.21	2.18	.174
S cuad × D cuad	1	0.53	0.53	0.96	.353
Error	9	5.02	0.56		

de la interacción. Las particiones del efecto principal normalmente se calculan a partir de las medias marginales, pero, en la tabla 6.12 se usan las medias de celdas así, los valores repetidos de los coeficientes polinomiales de contraste para las particiones del efecto principal son necesarias para las medias de celdas que contribuyen a cada una de sus respectivas medias marginales. Por ejemplo, S_l , el contraste lineal de salinidad, requiere un -1 para cada celda del nivel con 0 barios, un 0 para cada celda del nivel con 6 barios y un $+1$ para las del nivel con 12 barios.

Los coeficientes para las particiones de interacción se determinan como el producto de los coeficientes por las componentes de interacción correspondientes. Por ejemplo, los coeficientes de la interacción entre la salinidad(lineal) y el día(lineal) de la tabla 6.12 están formados como los productos de los coeficientes del contraste lineal por los efectos principales de los dos factores. Cada coeficiente para $S_l D_l$ es un producto de los coeficientes correspondientes de S_l y D_l . El cálculo se presenta en el cuadro 6.3.

Tabla 6.12 Cálculos de las particiones de sumas de cuadrados lineales y cuadráticas para el efecto principal de salinidad y día, y las sumas de cuadrados de la interacción

Barios	Días	Medias (\bar{y}_y)	S_l	S_q	D_l	D_q	$S_l D_l$	$S_l D_q$	$S_q D_l$	$S_q D_q$
0	14	2.75	-1	1	-1	1	1	-1	-1	1
	21	5.35	-1	1	0	-2	0	2	0	-2
	28	12.80	-1	1	1	1	-1	-1	1	1
6	14	4.10	0	-2	-1	1	0	0	2	-2
	21	6.55	0	-2	0	-2	0	0	0	4
	28	10.20	0	-2	1	1	0	0	-2	-2
12	14	3.10	1	1	-1	1	-1	1	-1	1
	21	5.20	1	1	0	-2	0	-2	0	-2
	28	7.95	1	1	1	1	1	1	1	1
$\Sigma P_{cy} \bar{y}_y$		-4.7	-4.6	21.0	6.7	-5.2	-4.2	2.7	3.1	
ΣP_{cy}^2		6.0	18.0	6.0	18.0	4.0	12.0	12.0	36.0	
SC^*		7.2	2.3	147.0	5.0	13.5	2.9	1.2	0.5	
Efecto [†]		-0.78	-0.25	3.50	0.37	-1.30	-0.35	0.23	0.09	

* $SC = r(\Sigma P_{cy} \bar{y}_y)^2 / \Sigma P_{cy}^2$; [†]Efecto = $\Sigma P_{cy} \bar{y}_y / \Sigma P_{cy}^2$

Cuadro 6.3 Cálculo de los coeficientes para contraste polinomial ortogonal de la interacción

S_l	-1	-1	-1	0	0	0	1	1	1
D_l :	-1	0	1	-1	0	1	-1	0	1
$S_l D_l$:	1	0	-1	0	0	0	-1	0	1

Interpretaciones de los contrastes de regresión

El estadístico F_0 de la tabla 6.11 indica una interacción significativa entre la (lineal) partición de salinidad y las particiones de día(lineal y cuadrática), también fueron significativas las particiones de los efectos principales para ellas, pero ninguno de los efectos cuadráticos de salinidad fueron significativos al nivel 0.5 de significancia. El punto fundamental de la investigación fue la respuesta del consumo de agua de acuerdo con el nivel salino, una gráfica de perfiles facilita la interpretación de los resultados con interacción importante. En la figura 6.5 se muestra una gráfica de las medias de celdas y la regresión lineal del consumo diario de agua, en relación con el nivel de salinidad.

Las rectas de regresión lineal calculadas por separado se muestran en la figura 6.5 junto con las medias de tratamientos estimadas. Las comparaciones lineales de salinidad por día, calculadas a partir de la columna S_l de la tabla 6.12, se muestran en el cuadro 6.4 con sus intervalos de confianza del 95% con una $t_{05,3,9} = 2.93$ de Bonferroni. El ICS del 95% indica que la salinidad no tiene efecto sobre el consumo de agua de las plantas durante las primeras tres semanas (hasta el día 21), ya que los intervalos para 14 y 21 días incluyen al 0; Sin embargo, al final de las cuatro semanas (el día 28), el consumo de agua decreció al aumentar la salinidad

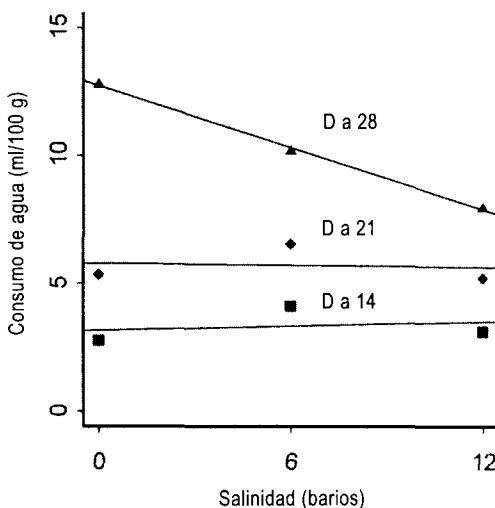


Figura 6.5 Agua consumida por plantas de cebada con tres niveles salinos en 14, 21 y 28 días

Cuadro 6.4 Comparaciones lineales de salinidad por día

Día	Comparaciones lineales	ICS del 95%
		(L, U)
14	$S_l = -1(2.75) + 0(4.10) + 1(3.10) = 0.35$	(-1.85, 2.55)
21	$S_l = -1(5.35) + 0(6.55) + 1(5.20) = -0.15$	(-2.35, 2.05)
28	$S_l = -1(12.80) + 0(10.20) + 1(7.95) = -4.85$	(-7.05, -2.65)
<i>Error estándar</i>		$s_c = \sqrt{CME[(-1)^2 + 0^2 + 1^2]/r} = \sqrt{0.56} = 0.75$

en el medio, pues el límite de confianza superior para la comparación lineal fue -2.65.

De las diferencias entre las respuestas lineales a la salinidad se obtuvieron interacciones significativas entre la partición (lineal) de salinidad y las participaciones de días. La interacción significativa entre los efectos salinidad(lineal) y día(cuadrática) indica que la respuesta lineal a la salinidad cambia de manera cuadrática con los días.

Los valores de comparación S_l disminuyen cuadraticamente cuadráticamente del día 14 al 28. La disminución del valor de comparación del día 14 al 21 es $(-0.15) - 0.35 = -0.50$, mientras que del día 21 al 28 es $(-4.85) - (-0.15) = -4.70$.

Cálculo de la ecuación de superficie de respuesta

La ecuación polinomial que relaciona el consumo de agua con la salinidad y los días, incluirá los términos que se juzguen significativos en las pruebas de análisis de varianza, que en este caso fueron S_l , D_l , D_q , S_lD_l y S_lD_q . La ecuación en términos de S y D será:

$$y = \beta_0 + \beta_1 S + \beta_2 D + \beta_3 D^2 + \beta_4 SD + \beta_5 SD^2 \quad (6.23)$$

Los coeficientes de la ecuación (6.23) pueden estimarse directamente de los datos con un programa de regresión múltiple, o evaluarse a partir del polinomio ortogonal como se presentó en la sección 3.3. Las comparaciones de los coeficientes polinomiales del cuadro 3.5 se usan para las transformaciones. Si P_{cs} y P_{cd} son las comparaciones polinomiales para salinidad y día, la ecuación polinomial ortogonal se puede escribir como:

$$y = \mu + \alpha_1 P_{1s} + \gamma_1 P_{1d} + \gamma_2 P_{2d} + (\alpha\gamma)_{11} P_{1s}P_{1d} + (\alpha\gamma)_{12} P_{1s}P_{2d} \quad (6.24)$$

donde α_1 es el coeficiente polinomial de salinidad, γ_1 es el coeficiente lineal para el día y $(\alpha\gamma)_{12}$ es el coeficiente de interacción para la salinidad lineal por día cuadrático. Las estimaciones de los coeficientes para la ecuación (6.24) se calculan en la tabla 6.12 como: Efecto = $\sum P_{cij}\bar{y}_j / \sum P_{cij}^2$. Por ejemplo, la estimación para $(\alpha\gamma)_{11}$ a partir de la recta S_lD_l en la tabla 6.12 es $-5.2/4 = -1.3$. Por tanto, el término $(\alpha\gamma)_{11} P_{1s}P_{1d}$ en la ecuación (6.24), con $\lambda_1 = 1$, se convierte en:

$$-1.3\lambda_1 \left(\frac{S - 6}{6} \right) \lambda_1 \left(\frac{D - 21}{7} \right) = -0.0310(S - 6)(D - 21)$$

Las estimaciones restantes se calculan de la misma manera con $\hat{\mu} = \bar{y} = 6.44$, $\lambda_1 = 1$ y $\lambda_2 = 3$, y la ecuación resultante es:

$$\begin{aligned} \hat{y} = & 5.70 - 0.0133(-6) + 0.50(D - 21) + 0.0227(D - 21)^2 \\ & - 0.0310(S - 6)(D - 21) - 0.0036(S - 6)(D - 21)^2 \end{aligned}$$

Estas ecuaciones se pueden usar para explorar los máximos o mínimos en las superficies de respuesta, o para determinar los valores de los factores que tienen respuestas equivalentes. En el capítulo 13, Diseño de superficies de respuesta, se presentan herramientas especializadas para estos métodos.

6.6 Tres factores de tratamiento

La inclusión de más factores al diseño de tratamientos aumenta la complejidad de los patrones de interacción entre los factores de tratamiento. El número de combinaciones de tratamientos aumenta tanto como se agregan factores al diseño, es decir, un diseño de tres factores con a niveles de A , b niveles de B y c niveles de C ,

tiene abc combinaciones. Un cuarto factor, D , con d niveles aumenta el número de tratamientos en un múltiplo de d .

El diseño con dos factores permite la investigación de la interacción de *primer orden* o doble, AB . En el diseño con tres factores las dos interacciones de primer orden adicionales, AC y BC , amplían las inferencias del estudio y debe considerarse además una interacción de *segundo orden* o de tres factores (o triple), ABC . Las interacciones de *tercer orden* o de cuatro factores como $ABCD$, introducen mayor complejidad en la estructura de la inferencia de la interacción.

Ejemplo 6.5 Cultivo de camarón en un acuario

El camarón café de California desova en el mar y los huevos se transforman en larva mientras son transportados a la costa, pasada la etapa larvaria entran en los estuarios, donde crecen con rapidez y se convierten en pre-adultos que emigran de nuevo al mar para alcanzar ahí su madurez sexual.

Como resultado de sus migraciones, el camarón encuentra una amplia variación de la temperatura y salinidad durante su ciclo de vida, por lo que es de gran importancia conocer cómo afectan su crecimiento para entender su vida y ecología.

Cuando se realizó este experimento había un gran interés en el cultivo comercial del camarón y, desde el punto de vista de la acuacultura, otro factor importante era la densidad de almacenamiento en los tanques de cultivo, ya que esta afecta la competencia entre los ejemplares.

Objetivo de la investigación: Los investigadores desean conocer como influyen la temperatura y salinidad del agua y la densidad de población del camarón en la tasa de crecimiento de los camarones cultivados en acuarios y si estos factores actúan independientemente sobre la población.

Diseño del tratamiento: Se usó un arreglo factorial con tres factores: “temperatura” (25°C , 35°C); “salinidad” del agua (10%, 25%, 40%), y “densidad” de camarones en el acuario (80 camarones/40 litros, 160 camarones/40 litros). Se consideró que si el factor influía en el crecimiento de los camarones éstos niveles tenían las posibilidades más altas de mostrar efectos.

Diseño del experimento: El diseño del experimento consistió en tres acuarios réplica para cada una de las 12 combinaciones de tratamiento del factorial $2 \times 2 \times 3$, cada combinación se asignó al azar a tres acuarios en un diseño totalmente aleatorizado. y al principio de la prueba se pusieron camarones en etapa postlarvaria en los 36 acuarios. En la tabla 6.13 se muestra el aumento de peso por camarón en cuatro semanas para cada acuario.

Tabla 6.13 Aumento del peso de los camarones cultivados en acuarios con diferentes niveles de temperatura (T), densidad de población (D) y salinidad del agua (S), luego de cuatro semanas

T	D	S	<i>Aumento de peso (mg)</i>
25°C	80	10%	86, 52, 73
		25%	544, 371, 482
		40%	390, 290, 397
	160	10%	53, 73, 86
		25%	393, 398, 208
		40%	249, 265, 243
35°C	80	10%	439, 436, 349
		25%	249, 245, 330
		40%	247, 277, 205
	160	10%	324, 305, 364
		25%	352, 267, 316
		40%	188, 223, 281

Fuente: Dr. J. Hendrickson y K. Dorsey, Department of Ecology and Evolutionary Biology, University od Arizona.

Modelo estadístico para tres factores

El modelo de medias de celdas para un experimento con tres factores y r réplicas de cada una de las abc combinaciones de tratamientos en un diseño totalmente aleatorizado es:

$$y_{ijkl} = \mu_{ijk} + e_{ijkl} \quad (6.25)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, c \quad l = 1, 2, \dots, r$$

La media de celdas μ_{ijk} expresada como una función del factorial de los efectos principales y las interacciones es:

$$\mu_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_j + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} \quad (6.26)$$

donde $\mu = \bar{\mu}$ es la media general y α_i , β_j y γ_k son los efectos principales de A , B y C . Los efectos de la interacción respectivos de dos factores son $(\alpha\beta)_j$, $(\alpha\gamma)_{ik}$ y $(\beta\gamma)_{jk}$, y el efecto de la interacción de tres factores es $(\alpha\beta\gamma)_{ijk}$. Las definiciones de las interacciones de efectos principales y de dos factores se obtienen de los desarrollos dados en las ecuaciones (6.6) a (6.8) para el experimento con dos factores. Los efectos principales son:

$$\alpha_i = (\bar{\mu}_i - \bar{\mu}), \beta_j = (\bar{\mu}_j - \bar{\mu}), \gamma_k = (\bar{\mu}_k - \bar{\mu})$$

y una interacción de dos factores típica es:

$$\begin{aligned} (\beta\gamma)_{jk} &= (\bar{\mu}_{jk} - \bar{\mu}_{..}) - \beta_j - \gamma_k \\ &= \bar{\mu}_{jk} - \bar{\mu}_j - \bar{\mu}_{..k} + \bar{\mu}_{..} \end{aligned} \quad (6.27)$$

La interacción de tres factores se presenta cuando las interacciones del efecto principal y dos factores no logran explicar la variación en las desviaciones de las medias de celdas ($\mu_{ijk} - \bar{\mu}$). La interacción de tres factores es la diferencia entre la desviación de la media de celdas y la suma de los efectos principales y los efectos la interacción de dos factores:

$$\begin{aligned}\alpha\beta\gamma_{ijk} &= (\mu_{ijk} - \bar{\mu}) - [\alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk}] \\ &= \mu_{ijk} - \bar{\mu}_i - \bar{\mu}_{jk} + \bar{\mu}_i + \bar{\mu}_j + \bar{\mu}_k - \bar{\mu}\end{aligned}\quad (6.28)$$

Análisis de tres factores

Se hace una partición de las sumas de cuadrados para los tratamientos en las sumas de cuadrados de los efectos principales y de la interacción, como sigue:

$$\begin{aligned}SC \text{ tratamiento} &= SCA + SCB + SCC + SC(AB) \\ &\quad + SC(AC) + SC(BC) + SC(ABC)\end{aligned}\quad (6.29)$$

Recuerde que los grados de libertad para las sumas de cuadrados de los efectos principales son $(a - 1)$, $(b - 1)$ y $(c - 1)$ para los factores A , B y C , respectivamente, y que los grados de libertad de las interacciones de dos factores son el producto de los grados de libertad del efecto principal para los factores incluidos. De la misma manera, los grados de libertad la interacción de tres factores o más son el producto de los grados de libertad del efecto principal de los factores incluidos, por lo que $SC(ABC)$ tiene $(a - 1)(b - 1)(c - 1)$ grados de libertad.

En la tabla 6.14 se muestran las particiones de sumas de cuadrados y la tabla del análisis de varianza para este experimento con tres factores, en ella el cuadrado medio del error es el denominador del estadístico F_0 para probar la hipótesis nula para cualquier conjunto de efectos factoriales con el modelo de efectos fijos. El estadístico F_0 en la tabla 6.14 conduce al rechazo de la hipótesis nula para la interacción de dos factores TS , la interacción de tres factores TDS , y todos los efectos principales. Las medias de celdas y marginales para todos los factores se muestran en la tabla 6.15.

Tabla 6.14 Análisis de varianza para el aumento de peso en los camarones cultivados

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	35	537 327.01			
Temp (T)	1	15 376.00	15 376.00	5.30	.030
Salinidad (S)	2	96 762.50	48 381.25	16.66	.000
Densidad (D)	1	21 218.78	21 218.78	7.31	.012
TS	2	300 855.17	150 427.58	51.80	.000
TD	1	8 711.11	8 711.11	3.00	.096
SD	2	674.39	337.19	0.12	.891
TDS	2	24 038.39	12 019.19	4.14	.029
Error	24	69 690.67	2 903.78		

Tabla 6.15 Medias de celdas y marginales del aumento de peso en cuatro semanas de camarones cultivados con diferentes niveles de temperatura (T), densidad de población (D) y salinidad del agua (S)

<i>Medias de celdas (\bar{y}_{jk})</i>				<i>Densidad</i>						
<i>Salinidad</i>	<i>80</i>		<i>160</i>		\bar{y}_k					
	<i>Temperatura</i>		<i>Temperatura</i>							
	25°	35°	25°	35°						
10%	70	408	71	331	220					
25%	466	275	333	312	346					
40%	359	243	252	231	271					
<i>Medias $T \times D$</i>	298	309	219	291						
(\bar{y}_y)										
<i>Medias $T \times D$ (\bar{y}_{ik})</i>				<i>Medias $D \times S$ (\bar{y}_{jk})</i>						
<i>S</i>				<i>S</i>						
<i>T</i>	10%	25%	40%	\bar{y}_i	<i>D</i>	10%	25%			
25°	71	399	306	259	80	239	370			
35°	370	293	237	300	160	201	322			
						301	242			
						303	255			

Las interpretaciones deben estar condicionadas a alguna medida de significancia estadística junto con la significancia biológica de las respuestas. Se requieren los errores estándar de las medias de celdas y marginales para cualquier prueba estadística subsecuente de comparaciones específicas; el error estándar para cualquier media es $s_{\bar{y}} = \sqrt{CME/n}$, donde n es el número de observaciones en la media y el error estándar de la diferencia entre cualquier par de medias es $s(\bar{y}_i - \bar{y}_j) = \sqrt{2CME/n}$. Los errores estándar estimados para el experimento de cultivo de camarón se muestran en la tabla 6.16.

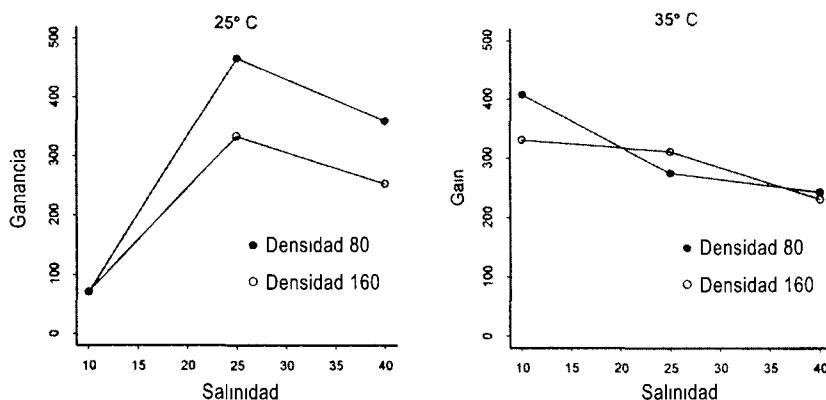
Algunas interpretaciones preliminares acerca de los efectos de los factores

La significancia de la interacción de estos factores indica que temperatura, salinidad y densidad se interrelacionan en cuanto a su efecto sobre el crecimiento del camarón. La interacción significativa de los tres factores implica que la interacción entre dos de ellos no es constante para los niveles del tercer factor. Considere la interacción entre la densidad y la salinidad por separado, a temperaturas de 25°C y 35°C , como se muestra en las gráficas de medias de celdas en la figura 6.6.

Para interpretar los resultados se puede usar una comparación de los efectos simples de la salinidad sobre cada nivel de densidad y temperatura, los efectos simples de la salinidad se estiman mejor como contrastes polinomiales ortogonales lineales y cuadráticas para cada combinación de temperatura y densidad. Las particiones de las sumas de cuadrados calculadas para las tres interacciones del factor TDS son $SC(T \times D \times S \text{ lineal}) = 11\ 051$ y $SC(T \times D \times S \text{ cuadrática}) = 12\ 987$, con los valores $P\ .063$ y $.045$, respectivamente, que indican que el coeficiente cuadrático de salinidad depende de los niveles de temperatura y densidad.

Tabla 6.16 Errores estándar para las medias de celdas y marginales, en un diseño de tratamiento con tres factoresTemperatura: $a = 2$ niveles; densidad: $b = 2$ niveles; salinidad: $c = 3$ niveles

Medias de los factores principales		
Temperatura	Salinidad	Densidad
$\sqrt{\frac{CME}{rbc}} = \sqrt{\frac{2903.78}{18}} = 12.7$	$\sqrt{\frac{CME}{rab}} = \sqrt{\frac{2903.78}{12}} = 15.6$	$\sqrt{\frac{CME}{rac}} = \sqrt{\frac{2903.78}{18}} = 12.7$
Medias marginales de dos factores		
<i>Densidad por temperatura</i> $\sqrt{\frac{CME}{rc}} = \sqrt{\frac{2903.78}{9}} = 18.0$	<i>Densidad por salinidad</i> $\sqrt{\frac{CME}{ra}} = \sqrt{\frac{2903.78}{6}} = 22.0$	
<i>Salinidad por temperatura</i> $\sqrt{\frac{CME}{rb}} = \sqrt{\frac{2903.78}{6}} = 22.0$		<i>Medias de celdas</i> $\sqrt{\frac{CME}{r}} = \sqrt{\frac{2903.78}{3}} = 31.1$

**Figura 6.6** Aumento en el peso de los camarones cultivados en un arreglo factorial de $2 \times 2 \times 3$ de temperatura, densidad y salinidad

Los coeficientes cuadráticos de salinidad se calcularon como contrastes polinomiales ortogonales para las cuatro combinaciones de temperatura y densidad, a partir de las medias de celdas de la tabla 6.15, con un patrón similar al que proporciona la tabla 6.12. Por ejemplo, el coeficiente cuadrático de salinidad a 25°C y densidad de 80 es:

$$\frac{[70 - 2(466) + 359]}{[1^2 + (-2)^2 + 1^2]} = -\frac{503}{6} = -83.8$$

con error estándar $\sqrt{2903.78/6(3)} = 12.7$. Las estimaciones de un ICS del 95% se calcularon para los cuatro coeficientes según la $t_{05.4,24} = 2.70$ de Bonferroni.

Las estimaciones de un ICS del 95% para los coeficientes cuadráticos de una salinidad a 25°C son (-118.1, -49.5) para una densidad de 80 y (-91, -22.9) para una densidad de 160, mientras que las estimaciones respectivas para esas densidades a 35°C son (-17.5, 51.1) y (-44.6, 24.0). Es claro que la cuadratura a 25°C es significativa, ya que el ICS del 95% no incluye al 0; y no significativa a 35°C, pues esos intervalos sí lo incluyen.

6.7 Estimación de la varianza del error con una réplica

En los estudios de investigación surgen situaciones en las que sólo se dispone de una observación para cada celda del arreglo factorial. La varianza del error experimental no puede estimarse con sólo una réplica de las combinaciones de tratamiento, porque las particiones de las sumas de cuadrados para los efectos principales e interacción del factor son iguales a la suma de cuadrados total para las observaciones.

La *aditividad* describe que no hay interacción entre los factores, bajo la aditividad de factores se puede usar la partición de los cuadrados medios de la interacción como una estimación del error experimental. Como la aditividad de los efectos principales o de la ausencia de interacción no está garantizada, son necesarios algunos medios para evaluar la presencia de interacción.

Estimaciones de la varianza del error con dos factores cuantitativos

La aditividad de los factores cuantitativos se puede investigar con las componentes de interacción para las particiones de regresión lineal y, eventualmente, cuadrática (sección 6.5). Por ejemplo, es posible hacer una partición de la suma de cuadrados de la interacción en sumas de cuadrados de interacción lineal \times lineal, con la suposición de que la suma de cuadrados restante para las desviaciones de tal interacción incluirá todos los polinomios de interacción de orden más elevado, como lineal \times cuadrático, etcétera; también se puede utilizar el cuadrado de la media para las desviaciones de la interacción lineal \times lineal como el cuadrado medio del error. El número de términos de interacción con 1 grado de libertad que se partitionan a partir de la interacción es cuestión de juicio, con base en el número de grados de libertad disponibles para una prueba razonablemente potente de la interacción de los efectos principales.

Estimaciones de la varianza del error con un factor cualitativo y otro cuantitativo

Si uno de los factores es cualitativo y el otro es cuantitativo se puede usar el mismo enfoque. En este caso, las sumas de cuadrados para la interacción entre el factor

cualitativo y el efecto lineal del factor cuantitativo se puede separar a partir de la suma de cuadrados de la interacción (sección 6.5). La suma de cuadrados de las desviaciones restantes se puede usar para estimar el error experimental.

Estimaciones de la varianza del error con dos factores cualitativos

Si ambos factores son cualitativos el problema es un poco más difícil, pero Tukey (1949b) elaboró un método para aislar una suma de cuadrados con 1 grado de libertad para probar la no aditividad de efectos entre dos factores con una observación por celda. El término para la no aditividad en el modelo lineal es el producto simple de los efectos principales, $\lambda\alpha_i\beta_j$, donde el parámetro λ representa el parámetro agregado por la no aditividad. El producto de los efectos principales es una forma de interacción multiplicativa y si no hay aditividad en este tipo específico de interacción entre los efectos principales, α_i y β_j , entonces $\lambda \neq 0$. Con este modelo, las medias de celdas son una suma de la media general, los efectos de los factores y el término del producto, es decir:

$$\mu_{ij} = \bar{\mu}_{..} + (\bar{\mu}_i - \bar{\mu}_{..}) + (\bar{\mu}_j - \bar{\mu}_{..}) + \lambda(\bar{\mu}_i - \bar{\mu}_{..})(\bar{\mu}_j - \bar{\mu}_{..})$$

La suma de cuadrados para la no aditividad requiere un cálculo que incluye la desviación que tienen las medias de A y B con respecto la gran media, $(\bar{y}_{..} - \bar{y})$ y $(\bar{y}_j - \bar{y}_{..})$, respectivamente. La técnica se ilustra con el ejemplo 6.6.

Ejemplo 6.6 Niveles auditivos en hombres adultos

Los datos de la tabla 6.17 representan el porcentaje de hombres entre 55 y 64 años con niveles auditivos de 16 decibeles por encima del cero métrico de sonido. Las categorías por renglón son los niveles de sonido en ciclos por segundo (hertz) y las columnas describen siete categorías ocupacionales.

Los cálculos necesarios incluyen:

$$P_j = \sum_{i=1}^a (\bar{y}_{ij} - \bar{y}_{..})y_{ij} \quad (6.30)$$

$$P = \sum_{j=1}^b P_j(\bar{y}_j - \bar{y}_{..}) = 17\,416 \quad (6.31)$$

y

$$\sum_{i=1}^a (\bar{y}_{ij} - \bar{y}_{..})^2 = 6\,944, \quad \sum_{j=1}^b (\bar{y}_j - \bar{y}_{..})^2 = 162 \quad (6.32)$$

La suma de cuadrados con 1 grado de libertad para la no aditividad es:

Tabla 6.17 Porcentaje de hombres con niveles auditivos de 16 decibeles por encima del cero métrico de sonido, clasificados en un arreglo factorial de 7×7 con una observación por celda

A	B							\bar{y}_i	$(\bar{y}_i - \bar{y})$
	1	2	3	4	5	6	7		
1	2.1	6.8	8.4	1.4	14.6	7.9	4.8	6.6	-31.6
2	1.7	8.1	8.4	1.4	12.0	3.7	4.5	5.7	-32.5
3	14.4	14.8	27.0	30.9	36.5	36.4	31.4	27.3	-10.9
4	57.4	62.4	37.4	63.3	65.5	65.6	59.8	58.8	20.6
5	66.2	81.7	53.3	80.7	79.7	80.8	82.4	75.0	36.8
6	75.2	94.0	74.3	87.9	93.3	87.8	80.5	84.7	46.5
7	4.1	10.2	10.7	5.5	18.1	11.4	6.1	9.4	-28.8
\bar{y}_j	31.6	39.7	31.4	38.7	45.7	41.9	38.5	$\bar{y} = 38.2$	
$(\bar{y}_j - \bar{y})$	-6.6	1.5	-6.8	0.5	7.5	3.7	0.3		
P_j	6719	7730	5046	7776	6850	7313	7192		

Fuente: C. Daniel (1978), Patterns in residuals in the two-way layout. *Technometrics* 20, 385-395. Datos originales publicados en J. Roberts y J. Cohrssen (1968), *Hearing Levels of Adults*, Tabla 4, p. 36. U. S. National Center for Health Statistics Publications, Series 11, núm. 31, Rockland, Md.

$$\begin{aligned}
 S(\text{no aditividad}) &= \frac{P^2}{\sum_{i=1}^a (\bar{y}_i - \bar{y})^2 \sum_{j=1}^b (\bar{y}_j - \bar{y})^2} \\
 &= \frac{17\ 416^2}{(6\ 944)(162)} = 269.6
 \end{aligned} \tag{6.33}$$

En la tabla 6.18 se muestra el análisis de varianza para estos datos con una partición de la suma de cuadrados del error en una suma de cuadrados con 1 grado de libertad para las sumas de cuadrados no aditividad y residual.

Tabla 6.18 Partición de 1 grado de libertad para la no aditividad, en el análisis de varianza para un factorial de 7×7 , con una observación por celda

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Renglones	6	48 589.1	8 098.2
Columnas	6	1 141.5	190.2
Error	36	1 444.7	40.1
No aditividad	1	269.6	269.6
Residual	35	1 175.1	33.6

La hipótesis nula de no aditividad se prueba con el estadístico $F_0 = CM(\text{no aditividad})/CM(\text{residual}) = 269.6/33.6 = 8.02$, y se rechaza con una región crítica $F_0 > F_{05,1,35} = 4.12$.

Se han desarrollado varios métodos para comprobar la fuente de no aditividad en una tabla de dos efectos: Daniel (1978) usó un método basado en los residuales en cada celda, $y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}$; Johnson y Graybill (1972), Bradu y Gabriel (1978) y Mandel (1971) presentan discusiones técnicas y ejemplos de otros métodos y modelos de no aditividad.

6.8 ¿Cuántas réplicas se requieren para probar los efectos de un factor?

En la sección 2.14 se describen los procedimientos para estimar el número de réplicas, con base en la prueba de las diferencias entre las medias de tratamiento y el estadístico F_0 . Los valores de Φ [ecuación (2.25)] se pueden aplicar directamente a una prueba de diferencias entre las medias de celdas en el arreglo factorial y la hipótesis nula $H_0: \mu_{11} = \mu_{12} = \dots = \mu_{ab}$; en este caso, la estructura factorial se ignora y el modelo de medias de celda $y_{ijk} = \mu_{ij} + e_{ijk}$ expresado con la forma del modelo de efectos es $y_{ijk} = \mu + \tau_{ij} + e_{ijk}$, donde τ_{ij} es el efecto de la ij -ésima combinación de tratamiento en el arreglo factorial. Entonces:

$$\Phi^2 = \frac{r \sum_{i=1}^a \sum_{j=1}^b \tau_{ij}^2}{ab\sigma^2} \quad (6.34)$$

se usa para estimar el número de réplicas, a partir de gráficas basadas en los valores de la τ_{ij} , necesarias para ser significativas.

Si se requieren números de réplicas basados en los efectos factoriales, los parámetros no centrados son:

$$\lambda_a = br \sum_{i=1}^a \frac{\alpha_i^2}{\sigma^2}, \lambda_b = ar \sum_{j=1}^b \frac{\beta_j^2}{\sigma^2}, \text{ y } \lambda_{ab} = r \sum_{i=1}^a \sum_{j=1}^b \frac{(\alpha\beta)_{ij}^2}{\sigma^2} \quad (6.35)$$

respectivamente, para los efectos principales de A y B y su interacción. Entonces Φ se determina como $\Phi = \sqrt{\lambda}/(v_1 + 1)$, donde v_1 son los grados de libertad para el numerador del estadístico F_0 .)

6.9 Réplicas desiguales en los tratamientos

Es inevitable que falten datos en los estudios de investigación y con un conjunto incompleto de datos, el diseño ya no está balanceado y no se pueden aplicar las fórmulas convencionales. Antes del advenimiento de las computadoras modernas, lo mejor era contar con un conjunto de datos completos ya que permitía usar fórmulas bastante sencillas para los cálculos manuales. Se dedicó un gran esfuerzo para desarrollar los métodos de partición de las sumas de cuadrados en el análisis

de varianza, cuando existía un número desigual de observaciones entre las celdas del arreglo factorial. Las rutinas estadísticas generales, programadas para aceptar la teoría estadística conocida, han eliminado el esfuerzo de cálculo asociado con el análisis de conjuntos incompletos de datos.

La ortogonalidad se pierde al faltar observaciones

Cuando hay números desiguales de observaciones para las combinaciones de tratamientos, la suma de cuadrados de un efecto factorial transmite cierta información sobre otros efectos factoriales y las sumas de cuadrados se calculan de manera usual. Esta relación no ortogonal en las particiones de la suma de cuadrados para el análisis de varianza requiere que se preste mucha atención a las estimaciones usadas como parámetros del modelo y a los estadísticos utilizados para probar las hipótesis críticas al analizar el estudio.

Los contrastes ortogonales se describen en el capítulo 3 como aquellas que no contienen información acerca de las demás; la ortogonalidad tiene el mismo significado en las observaciones de un diseño factorial. Cuando el número de observaciones es igual en cada combinación de tratamientos, las sumas de cuadrados en el análisis de varianza constituyen una partición ortogonal de la suma de cuadrados del tratamiento. En la sección 6.4, la partición aditiva de las de tratamiento *SC* para un experimento de dos factores balanceado era:

$$SC \text{ tratamiento} = SCA + SCB + SC(\mathbf{AB})$$

y la suma de cuadrados para un efecto factorial no transmite información sobre los demás efectos factoriales.

Los datos mostrados en el cuadro 6.5 ilustran las complicaciones que introducen las réplicas desiguales en un arreglo factorial del tratamiento. Los datos en las celdas representan el excedente promedio del límite de velocidad establecido, para automóviles involucrados en 20 accidentes fatales, 10 ocurridos en clima lluvioso y 10 en días claros y los factores para el estudio son *W* (clima) y *R* (tipo de camino). Observe que 8 de los 10 accidentes con clima lluvioso ocurrieron en carreteras interestatales, mientras que 8 de los 10 accidentes en días claros ocurrieron en carreteras de 2 carriles.

La observación de las medias marginales del clima indica que el excedente promedio del límite de velocidad establecido para clima lluvioso es un poco mayor que el del día claro, $\bar{y}_1 = 13$ contra $\bar{y}_2 = 12$. Sin embargo, las medias de celdas observadas indican un resultado muy diferente, las velocidades promedio para accidentes fatales en días claros fueron 5 millas por hora mayores que en días lluviosos, tanto en las carreteras interestatales ($20 - 15$) como en las de dos carriles ($10 - 5$).

Las réplicas desiguales de tratamientos conducen a resultados contradictorios a partir de las medidas de celdas y las marginales, más accidentes ocurrieron en carreteras interestatales en días lluviosos y más accidentes ocurrieron en carreteras de dos carriles en días claros. Así, la media marginal para días claros está sesgada hacia abajo por el exceso de accidentes en carreteras de dos carriles con velocidades globales menores, y la media marginal para clima lluvioso está sesgada

Cuadro 6.5 Réplicas desiguales de tratamientos en un factorial de 2×2 para el excedente de los límites de velocidad establecidos

		<i>Medias de celdas</i>			
		<i>Interestatales</i>	<i>Dos carriles</i>	<i>Sumas</i>	<i>Medias</i>
Lluvioso	15	5		130	13
	$r_{11} = 8$	$r_{12} = 2$		$r_{1\cdot} = 10$	
Claro	20	10		120	12
	$r_{21} = 2$	$r_{22} = 8$		$r_2 = 10$	
Sumas		160	90	250	
Medias		$r_{\cdot 1} = 10$	$r_{\cdot 2} = 10$	$r = 20$	$\bar{y} = 12.5$
		16	9		

hacia arriba por el exceso de accidentes en las interestatales con velocidades globales mayores.

La suma de cuadrados para los tratamientos con números de réplicas desiguales se calcula de manera correcta como sigue:

$$\begin{aligned} SC \text{ tratamiento} &= 8(15 - 12.5)^2 + 2(5 - 12.5)^2 + 2(20 - 12.5)^2 + 8(10 - 12.5)^2 \\ &= 325 \end{aligned}$$

Las sumas de cuadrados para los efectos principales y para la interacción calculados de manera *incorrecta* con los métodos descritos en la sección 6.4 son

$$SCW = 10(13 - 12.5)^2 + 10(12 - 12.5)^2 = 5$$

$$SCR = 10(16 - 12.5)^2 + 10(9 - 12.5)^2 = 245$$

y

$$SC(WR) = SC \text{ tratamiento} - SCW - SCR = 325 - 5 - 245 = 75$$

La inspección de las medias de celdas y las sumas de cuadrados de la interacción indica otra contradicción en los métodos usuales de análisis: la suma de cuadrados calculada, $SC(WR) = 75$, señala que está presente cierta interacción en el estudio mientras que la inspección de las medias de celdas no indica interacción alguna. El efecto simple observado del clima es igual a 5 para ambos tipos de carretera, por tanto, la suma de cuadrados para la interacción sería 0 si el cálculo de las particiones de la suma de cuadrados fuera correcto.

Los principios generales de un análisis correcto de diseños factoriales con réplicas desiguales se ilustran con el experimento de dos factores sobre la durabilidad del concreto asfáltico.

Ejemplo 6.7 Analizando una vez más la durabilidad del concreto asfáltico.

El experimento del coeficiente de ruptura de los especímenes de concreto asfáltico del ejemplo 6.2 se puede usar para describir el análisis de un diseño factorial con réplicas desiguales de tratamiento. Con el propósito de ejemplificar, supongamos que los especímenes se construyeron con mezclas de basalto o rocas siliceas para los tres métodos de fraguado y que algunos se dañaron de las pruebas, por lo que se dispone de un número desigual de especímenes en los tratamientos. Los datos para réplicas desiguales se muestran en la tabla 6.19.

Tabla 6.19 Coeficiente de ruptura (psi) para los especímenes de concreto asfáltico elaborados con dos tipos de mezclado y tres métodos de fraguado

Tipo de mezcla	Método de fraguado			Medias de la mezcla (\bar{y}_t)
	Normal	Lento	Muy lento	
Basalto	106	93	56	
	108	101		
		98		
Medias (\bar{y}_{1j})	107.0	97.3	56	93.7
Silícea	107	63	40	
	110	60	41	
	116		44	
Medias (\bar{y}_{2j})	111.0	61.5	41.7	72.6
Fraguado				
Medias (\bar{y}_j)	109.4	83.0	45.3	

Establecimiento de estimadores con el modelo de medias de celdas

El modelo de medias de celdas se puede usar para establecer los estimadores adecuados de los parámetros de la población y las hipótesis que deben probarse. El modelo de medias de celdas es:

$$y_{ijk} = \mu_{ij} + e_{ijk} \quad (6.36)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r_{ij}$$

donde μ_{ij} es la media de celda del nivel i -ésimo del factor A y el nivel j -ésimo del factor B , e_{ijk} es el error aleatorio experimental independiente con distribución normal con media 0 y varianza σ^2 , y r_{ij} es el número de observaciones réplica en la celda (i, j) . Se supondrá que existe al menos una observación por celda del arreglo factorial, de manera que $r_{ij} > 0$ para toda i y j .

Estimadores de mínimos cuadrados para las medias de celdas y marginales

Es posible usar el método de mínimos cuadrados para estimar las medias de celdas con el procedimiento descrito en el capítulo 2, los estimadores de las medias de celdas son las medias de celdas observadas:

$$\hat{\mu}_y = \frac{1}{r_y} \sum_{k=1}^{r_y} y_{yk} = \bar{y}_y \quad (6.37)$$

y el estimador para la varianza del error experimental es:

$$\hat{\sigma}^2 = s^2 = \frac{1}{N - ab} \sum (y_{yk} - \bar{y}_{y.})^2 \quad (6.38)$$

donde $N = \sum r_y$.

Las estimaciones de las medias de celdas para la resistencia de los especímenes de concreto asfáltico se muestran en la tabla 6.19, y el error experimental estimado es:

$$s^2 = \frac{1}{14 - 6} [(106 - 107.0)^2 + \dots + (44 - 41.7)^2] = \frac{89.83}{8} = 11.23$$

Los estimadores de mínimos cuadrados no sesgados de las medias marginales son:

$$\hat{\mu}_i = \frac{1}{b} \sum_{j=1}^b \hat{\mu}_y \quad y \quad \hat{\mu}_J = \frac{1}{a} \sum_{i=1}^a \hat{\mu}_y \quad (6.39)$$

con errores estándar estimados:

$$\hat{s}_{\hat{\mu}_i} = \sqrt{\frac{s^2}{b^2} \sum_{j=1}^b \frac{1}{r_{ij}}} \quad y \quad \hat{s}_{\hat{\mu}_J} = \sqrt{\frac{s^2}{a^2} \sum_{i=1}^a \frac{1}{r_{ij}}} \quad (6.40)$$

Las estimaciones por mínimos cuadrados de las medias marginales para los especímenes de concreto asfáltico y sus errores estándar estimados se muestran en la tabla 6.20.

Por ejemplo, la estimación por mínimos cuadrados de la media marginal del tipo de mezcla de basalto es:

$$\hat{\mu}_1 = \frac{1}{3} (107.0 + 97.3 + 56.0) = 86.8$$

con un error estándar estimado:

$$\hat{s}_{\hat{\mu}_1} = \sqrt{\frac{11.23}{3^2} \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{1} \right)} = 1.51$$

Las medias marginales observadas, \bar{y}_i y \bar{y}_J mostradas en la tabla 6.19, no tienen el mismo valor que las estimaciones de los mínimos cuadrados de las medias marginales de la tabla 6.20. Las medias marginales observadas estiman funciones ponderadas de las medias poblacionales, donde los valores relativos son proporcionales al número de réplicas en las celdas. Los valores esperados de las medias marginales observadas son:

Tabla 6.20 Estimaciones por mínimos cuadrados de las medias marginales para el coeficiente de ruptura de los especímenes de concreto asfáltico y sus errores estándar

Mezcla	Media	Error estándar
	$\widehat{\mu}_i$	$s_{\widehat{\mu}_i}$
Basalto	86.8	1.51
Silícea	71.4	1.21
<i>Fraguado</i>	$\widehat{\mu}_j$	$s_{\widehat{\mu}_j}$
Normal	109.0	1.53
Lento	79.4	1.53
Muy lento	48.8	1.93

$$E(\bar{y}_i) = \frac{1}{r_i} \sum_{j=1}^b r_{ij} \mu_j \quad (6.41)$$

$$E(\bar{y}_j) = \frac{1}{r_j} \sum_{i=1}^a r_{ij} \mu_i$$

Si el número de observaciones en las celdas de tratamiento del estudio es proporcional a la frecuencia con que ocurren esas combinaciones de tratamiento en la población, entonces las medias marginales observadas proporcionan los estimadores adecuados para las medias marginales de la población. En las investigaciones sencillas es común la relación proporcional de los números observados con las frecuencias poblacionales, sin pero, no se espera que la relación proporcional se cumpla para un experimento diseñado o un estudio comparativo por observación y deben usarse las estimaciones de mínimos cuadrados de la tabla 6.20.

Las hipótesis no cambian ante las réplicas desiguales de tratamiento

Las hipótesis de interés en el diseño factorial de tratamientos con números desiguales de réplicas no son diferentes de las que interesan en un diseño con igual número de réplicas. La pregunta de investigación inicial en el diseño factorial del tratamiento considera la existencia de interacción entre los factores A y B , este efecto mide la diferencia entre los efectos simples de A a diferentes niveles de B . La diferencia entre los niveles i y k de A en los niveles j y m de B es la forma general de interacción:

$$(\mu_{ij} - \mu_{kj}) - (\mu_{im} - \mu_{km}) = \mu_{ij} - \mu_{kj} - \mu_{im} + \mu_{km}$$

La hipótesis nula de no interacción se puede expresar en términos de las medias de celdas como:

$$H_0: \mu_{ij} - \mu_{im} - \mu_{km} + \mu_{im} = 0 \quad \text{para toda } i, j, k, \text{ y } m \quad (6.42)$$

En ausencia de interacción, el efecto de los factores individuales sobre la variable respuesta se puede explorar por separado con las pruebas de hipótesis para las medias marginales. La hipótesis nula de interés para el factor A es la igualdad de las medias marginales, es decir:

$$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \cdots = \bar{\mu}_a \quad (6.43)$$

y la del factor B es:

$$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \cdots = \bar{\mu}_b \quad (6.44)$$

Cuadrados medios ponderados para las pruebas de hipótesis

Entre los numerosos métodos desarrollados para analizar los experimentos factoriales con réplicas desiguales, sólo el método de *cuadrados medios ponderados*, propuesto por Yates (1934), proporciona particiones de la suma de cuadrados para probar las tres hipótesis de las ecuaciones (6.42) a (6.44); Speed, Hocking y Hackney (1978) presentan una descripción de otros métodos y las hipótesis que se pueden probar con ellos. Las pruebas para la igualdad de las medias marginales son de interés sólo en ausencia de interacción.

Cálculo de las sumas de cuadrados para la interacción a partir de modelos completos y reducidos

La partición de las sumas de cuadrados para la interacción se determina a partir del principio de los modelos completos y reducidos introducido en el capítulo 2. El modelo completo, expresado en términos de los efectos factoriales, es:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad (6.45)$$

Se usan las soluciones obtenidas de las ecuaciones de mínimos cuadrados para calcular la suma de cuadrados del error experimental para el modelo completo, como sigue:

$$SCE_f = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{r_{ij}} [y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\hat{\alpha}\hat{\beta})_{ij}]^2 \quad (6.46)$$

Bajo la hipótesis nula de que no hay interacción, el modelo reducido es:

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \quad (6.47)$$

Las soluciones obtenidas con las ecuaciones de mínimos cuadrados se usan para calcular la suma de cuadrados del error experimental para el modelo reducido, como sigue:

$$SCE_r = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{r_{ij}} [y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j]^2 \quad (6.48)$$

La suma de cuadrados para la interacción se calcula como:

$$SC(\mathbf{AB}) = SCE_r - SCE_f \quad (6.49)$$

El cuadrado medio para el error experimental es $CME = (SCE_f)/(N - ab)$ y el cuadrado medio para la interacción es $CM(\mathbf{AB}) = SC(\mathbf{AB})/(a - 1)(b - 1)$. El estadístico F_0 usual, $F_0 = CM(\mathbf{AB})/CME$, prueba la hipótesis nula si no hay interacción en la ecuación (6.42). Los cálculos se ilustran en el apéndice 6A.

Cuadrados medios ponderados para probar la igualdad de los efectos principales en ausencia de interacción

Se pueden realizar pruebas de hipótesis para la igualdad de las medias marginales de los factores, si la prueba de interacción no es significativa y es seguro suponer que no hay interacción. En el cuadro 6.6 se muestran las particiones de la suma de cuadrados correctas para el método de cuadrados medios ponderados para probar la hipótesis nula en las ecuaciones (6.43) y (6.44). El análisis se basa en las sumas de cuadrados medios de celdas designados como las observaciones $x_{ij} = \bar{y}_{ij}$.

Cuadro 6.6 Cuadrados medios ponderados de las particiones de la suma de cuadrados	
Factor A	Factor B
$SCA_w = \sum_i^a w_i (\bar{x}_i - \bar{x}_{[1]})^2$	$SCB_w = \sum_j^b v_j (\bar{x}_j - \bar{x}_{[2]})^2$
$w_i = \left[\frac{1}{b^2} \sum_{j=1}^b \frac{1}{r_{ij}} \right]$	$v_j = \left[\frac{1}{a^2} \sum_{i=1}^a \frac{1}{r_{ij}} \right]$
$\bar{x}_{[1]} = \sum_i^a w_i \bar{x}_i / \sum_i^a w_i$	$\bar{x}_{[2]} = \sum_j^b v_j \bar{x}_j / \sum_j^b v_j$
$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_a$	$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_b$

Las particiones de la suma de cuadrados requeridas por los cuadrados medios ponderados se pueden calcular con numerosos programas estadísticos, pero es posible que tengan varias opciones para el tipo de particiones de las sumas de cuadrados que se desean para el análisis. Es importante que se usen las opciones correctas, de manera que el programa calcule la suma adecuada.¹

¹Los programas usados para el análisis de varianza proporcionan la suma de cuadrados correcta para los cuadrados medios ponderados. Las instrucciones de uso de la mayoría de los programas indican si se dispone de distintos tipos de particiones de sumas de cuadrados y cómo obtenerlas. Las opciones correctas de las sumas de cuadrados para varios programas conocidos son:

Programa	Suma de cuadrados
SAS GLM	Type III
SPSS MANOVA	UNIQUE
MINITAB GLM	Adjusted
BMDP 2V	Default
Splus	summary.aov(...,ssType=3)

En la tabla 6.21 se muestra el análisis de varianza para la resistencia de los especímenes de concreto asfáltico del ejemplo 6.7.

Tabla 6.21 Análisis de varianza para la resistencia de especímenes de concreto asfáltico con réplicas desiguales de tratamientos, con los cuadrados medios de Yates

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	13	10 963.21			
Mezcla	1	710.45	710.45	63.27	.000
Fraguado	2	6 806.45	3 403.23	303.07	.000
Interacción	2	953.45	476.72	42.45	.000
Error	8	89.83	11.23		

Interpretación del ejemplo

La hipótesis nula de que no existe interacción entre el tipo e mezcla y el método de fraguado es:

$$H_0: \mu_{ij} - \mu_{ik} - \mu_{im} + \mu_{km} = 0 \quad \text{para toda } i, j, k, \text{ y } m \quad (6.50)$$

La interacción es significativa puesto que el estadístico $F_0 = CM(\mathbf{AB})/CME = 476.72/11.23 = 42.45$; en la tabla 5.21, es significativo con $Pr > F = .000$.

Con una interacción significativa entre el tipo de mezclado y el método de fraguado, será necesario estudiar los efectos simples de un factor en cada nivel del otro para entender la naturaleza de la interacción. En el cuadro 6.7 se muestran las comparaciones entre las medias de celdas para el tipo de mezclado para cada nivel del método de fraguado.

Cuadro 6.7 Pruebas *t* de Bonferroni para los efectos simples del tipo de mezcla para cada nivel del método de fraguado

Método de fraguado	$\widehat{\mu}_{1j} - \widehat{\mu}_{2j}$	Error estándar	t_0
Normal	$107.0 - 111.0 = -4.0$	$\sqrt{11.23 \left[\frac{1^2}{2} + \frac{(-1)^2}{3} \right]} = 3.1$	-1.29
Lento	$97.3 - 61.5 = 35.8$	$\sqrt{11.23 \left[\frac{1^2}{3} + \frac{(-1)^2}{2} \right]} = 3.1$	11.55
Muy lento	$56.0 - 41.7 = 14.3$	$\sqrt{11.23 \left[\frac{1^2}{1} + \frac{(-1)^2}{3} \right]} = 3.9$	3.67

En este cuadro se calcularon los estadísticos t_0 para la comparación de los tipos de mezcla para cada nivel de los métodos de fraguado. El valor crítico para la t de Bonferroni con tres comparaciones es $|t_0| > t_{025,3,8} = 3.02$. No hay una diferencia significativa entre el coeficiente de ruptura de los especímenes de asfalto y roca silícea elaborados con un amasado normal, pero con amasado lento y muy lento las resistencias del basalto fueron mucho mayores que las de la roca silícea.

Pruebas para las medias marginales

Es común no se tomen en cuenta las pruebas de hipótesis de la igualdad de medias marginales para A y B cuando la interacción es significativa y no se considerarán en este ejemplo. Sin embargo, para complementarlo, se presenta el procedimiento para un caso donde no existe interacción y en el que las pruebas sobre las medias marginales son de interés. Las dos hipótesis a probar en el ejemplo del concreto asfáltico son: 1) no hay diferencias entre las medias marginales para el tipo de mezcla:

$$H_0: \bar{\mu}_1 = \bar{\mu}_2 \quad (6.51)$$

y 2) no hay diferencia entre las medias marginales para el método de fraguado:

$$H_0: \bar{\mu}_1 = \bar{\mu}_2 = \bar{\mu}_3 \quad (6.52)$$

La suma de cuadrados en el método de cuadrados medios ponderados se muestra en la tabla 6.21. El estadístico $F_0 = CMA/CME = 710.45/11.23 = 63.26$ prueba la igualdad de las medias marginales para el tipo de mezcla, el estadístico para la prueba de igualdad entre las medias marginales del método de compactación es $F_0 = CMC/CME = 3403.23/11.23 = 303.05$ y ambos estadísticos son significativos con $Pr > F = .000$ en la tabla 6.21.

Algunos comentarios acerca de los datos y las celdas faltantes

En esta sección se ilustró un método para analizar los datos de un estudio con subclases de réplicas desiguales en un diseño factorial de tratamiento. Este método proporciona los estimadores correctos de las medias poblacionales y pruebas de hipótesis creíbles sobre los factores. Otros métodos para analizar los diseños factoriales de tratamientos no balanceados proporcionan pruebas de hipótesis distintas para los efectos de los factores.

Searle, Speed y Henderson (1981) estudiaron los cinco métodos para calcular las sumas de cuadrados en el análisis de varianza que se usaron en programas de computadora. Estos métodos producen los mismos resultados con datos balanceados, pero pueden ser muy distintos con datos no balanceados. Artículos relacionados de Hocking y Speed (1975), Speed y Hocking (1976) y Speed, Hocking y Hackney (1978) proporcionan más información de los métodos de cálculo y las hipótesis que prueban los diferentes programas de computadora en el análisis. En Searle (1971, 1987) y Milliken y Johnson (1984) se encuentran ejemplos más amplios de algunos métodos, que incluyen los usados en esta sección junto con ciertos antecedentes teóricos.

Todos los métodos presentados en el contexto de particiones de sumas de cuadrados para los efectos principales y la interacción, proporcionan pruebas de hipótesis inadecuadas para los efectos factoriales cuando faltan todas las celdas de un arreglo factorial. En estas circunstancias Urquhart, Weeks y Henderson (1973), Hocking y Speed (1975), Yrquhart y Weeks (1978) y Searle (1987) recomiendan un análisis basado en el modelo de medias de celdas.

EJERCICIOS PARA EL CAPÍTULO 6

- Un proceso de producción química consiste de una primera reacción con un alcohol y una segunda reacción con una base. Se realizó un experimento factorial de 3×2 , con tres alcoholes y dos bases, con cuatro reacciones réplica en un diseño totalmente aleatorizado. Los datos se reunieron como porcentaje de la reacción.

Base	Alcohol					
	1		2		3	
1	91.3	89.9	89.3	88.1	89.5	87.6
	90.7	91.4	90.4	91.4	88.3	90.3
2	87.3	89.4	92.3	91.5	93.1	90.7
	91.5	88.3	90.6	94.7	91.5	89.8

Fuente: P. R. Nelson (1988), "Testing for interactions using analysis of means", *Technometrics* 30, 53-61.

- Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
- Construya una tabla de medias de celdas y marginales, y muestre sus errores estándar respectivos.
- Pruebe la hipótesis nula de que no hay efectos de interacción base \times alcohol. ¿Cuál es su conclusión de la prueba? ¿Qué recomienda como siguiente paso en el análisis?
- Utilice contraste múltiples entre las medias de celdas para ayudar a explicar la interacción. Por ejemplo, compare las dos bases para cada alcohol.
- Realice un análisis residual con un esquema normal y después con un esquema preestablecido; realice también una prueba de Leverne (Med). ¿Cuál es su conclusión?
- Una compañía probó dos métodos químicos para determinar la glucosa en el suero. Se usaron tres recipientes con suero para el experimento, cada uno contenía distintos niveles de glucosa mediante la adición de glucosa al nivel base. Se prepararon tres muestras de suero de cada recipiente independientes del nivel de glucosa, con cada uno de los dos métodos químicos. Se midió la concentración de glucosa (mg/dl) de todas las muestras en una corrida del espectrómetro.
- Se r...
no e...
alfal...
Leon...

Nivel de glucosa	Método 1			Método 2		
	1	2	3	1	2	3
	42.5	138.4	180.9	39.8	132.4	176.8
	43.3	144.4	180.5	40.3	132.4	173.6
	42.9	142.7	183.0	41.2	130.3	174.9

Fuente: Dr. J. Anderson, Beckman Instruments Inc.

- Escriba un modelo lineal para este experimento, explique los términos, realice un análisis de varianza para los datos y calcule los residuales. ¿Es necesaria una transformación de los datos? Explique.
 - Si es necesaria, calcule la transformación de los datos y el análisis de varianza respectivo.
 - Pruebe la hipótesis nula de que no existen efectos de interacción método \times glucosa. ¿Cuál es su conclusión? ¿Debe hacerse una prueba para los efectos principales?
 - Prepare una tabla de medias de celdas y marginales con sus respectivos errores estándar.
 - Pruebe la diferencia entre las medias de los métodos para cada nivel de glucosa e interprete los resultados.
3. Se llevó a cabo un estudio del efecto de la temperatura sobre el porcentaje de encogimiento de telas teñidas, con dos réplicas para cada uno de cuatro tipos de tela en un diseño totalmente aleatorizado. Los datos son el porcentaje de encogimiento de dos réplicas de tela secadas a cuatro temperaturas.

Tela	Temperatura			
	210° F	215° F	220° F	225° F
1	1.8, 2.1	2.0, 2.1	4.6, 5.0	7.5, 7.9
2	2.2, 2.4	4.2, 4.0	5.4, 5.6	9.8, 9.2
3	2.8, 3.2	4.4, 4.8	8.7, 8.4	13.2, 13.0
4	3.2, 3.6	3.3, 3.5	5.7, 5.8	10.9, 11.1

- Escriba un modelo lineal para el experimento, explique los términos y calcule el análisis de varianza.
 - Pruebe la hipótesis nula de que no hay interacción tela \times temperatura.
 - Divida la suma de cuadrados del efecto principal de la temperatura en particiones con 1 grado de libertad para las sumas de cuadrados de regresión lineal y cuadrática, y pruebe la hipótesis nula de que no hay respuesta lineal o cuadrática a la temperatura.
 - Haga una partición de sumas de cuadrados de la interacción temperatura \times tela en sumas de cuadrados de interacción temperatura lineal \times tela y temperatura cuadrática \times tela y pruebe la hipótesis nula de que no hay interacción para las respectivas particiones.
 - Prepare una gráfica de perfiles de las medias de celdas contra la temperatura para cada tela e interprete los resultados. Por ejemplo, puede hacerse las siguientes preguntas: “¿Cómo afecta la temperatura de secado al encogimiento de la tela?” “¿En qué difiere la relación entre el encogimiento y la temperatura entre los tipos de tela?”.
4. Se realizó un experimento de microbiología de suelos para determinar el efecto de la fertilidad del nitrógeno en la fijación de nitrógeno por bacterias *Rizobium*. El experimento se ejecutó con cuatro cosechas; alfalfa, soya, guar y habas. Se inocularon dos plantas con el *Rhizobium* y se cultivaron en un frasco de Leonard con una de las tres siguientes tasas de nitrógeno en el medio: 0, 50 y 100 ppm N. Se usaron cuatro

rélicas de frascos de Leonard para cada una de las 12 combinaciones de tratamiento, los tratamientos se arreglaron en un diseño totalmente aleatorizado en una cámara de cultivo y se midió la reducción de acetileno para cada tratamiento en la etapa de florecimiento de las plantas. La reducción de acetileno refleja la cantidad de nitrógeno que fija la bacteria en la relación simbiótica con la planta.

		<i>Cultivo</i>			
<i>Nitrógeno</i>		<i>Alfalfa</i>	<i>Soya</i>	<i>Guar</i>	<i>Habas</i>
0	2.6, 1.1	6.5, 2.6	0.3, 0.1	0.8, 0.9	
	0.9, 1.2	3.9, 4.3	0.4, 0.4	2.2, 1.2	
50	0.0, 0.0	0.6, 0.6	0.0, 0.1	0.7, 0.4	
	0.0, 0.0	0.3, 0.8	0.0, 0.2	0.3, 0.8	
100	0.0, 0.0	0.0, 0.1	0.0, 0.2	0.3, 0.1	
	0.0, 0.0	0.1, 0.0	0.0, 0.0	0.0, 0.1	

Fuente: Dr. I. Pepper, Department of Soil and Water Science,
University of Arizona.

- Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
- Realice un análisis residual y determine si es necesaria una transformación de los datos. Si lo es, transforme los datos y calcule el análisis de varianza para los datos transformados.
- Pruebe la hipótesis nula de que no hay efectos de interacción de cultivo, nitrógeno o cultivo \times nitrógeno.
- Divida la suma de cuadrados del efecto principal del nitrógeno y de la interacción nitrógeno \times cultivo en particiones con 1 grado de libertad para regresión lineal y cuadrática.
- Pruebe la hipótesis nula de que no hay efectos lineal o cuadrático del nitrógeno y la hipótesis nula de que no hay interacción lineal o cuadrática del nitrógeno con los cultivos.
- Haga una gráfica de los perfiles de medias de celdas contra el nivel de nitrógeno para cada cultivo e interprete el experimento. Por ejemplo, puede preguntar: “¿En qué afecta la adición de nitrógeno en el medio a la fijación de nitrógeno por la *Rhizobium*?” o “¿es igual el efecto de la adición de nitrógeno en la fijación del mismo para cada cultivo?”.
- Observe que dos combinaciones de tratamiento, alfalfa con 50 y 100 ppm N, tienen valor de cero en todas las observaciones. Este fenómeno es posible si algún cambio de nitrógeno en el medio de cultivo inhibe por completo la actividad de la *Rhizobium*. ¿En qué afecta esto la suposición del análisis de varianza en cuanto a la homogeneidad de la varianza? ¿Tiene alguna sugerencia que acepte esta situación en su análisis de datos?
- Un agrónomo realizó un experimento para determinar los efectos combinados de un herbicida y un insecticida en el crecimiento y desarrollo de plantas de algodón (delta de hoja suave). El insecticida y el herbicida se incorporaron al suelo usado en los contenedores de cultivo; se usaron cuatro contenedores cada uno con cinco plantas de algodón, para cada combinación de tratamiento. Los contenedores se arreglaron dentro del invernadero en un diseño totalmente aleatorizado. Se usaron cinco niveles (lb/acre) tanto de insecticida como de herbicida para obtener 25 combinaciones. Los datos que siguen son las medias de celdas para el peso de las raíces secas (gramos/planta) cuando las plantas tenían tres semanas.

Insecticida	Herbicida				
	0	0.5	1.0	1.5	2.0
0	122.0	72.50	52.00	36.25	29.25
20	82.75	84.75	71.50	80.50	72.00
40	65.75	68.75	79.50	65.75	82.50
60	68.00	70.00	68.75	77.25	68.25
80	57.50	60.75	63.00	69.25	73.25

Cuadrado medio del error del experimento = 174 con 75 grados de libertad

Fuente: Dr. K. Hamilton, Department of Plant Science, University of Arizona.

- Calcule las particiones de sumas de cuadrados de regresión con 1 grado de libertad para las sumas de cuadrados de herbicida, insecticida e interacción. Calcule un polinomio de grado no mayor a la regresión cúbica para el herbicida o el insecticida.
 - Pruebe las hipótesis nulas para cada partición y determine la forma de regresión polinomial que sea una descripción adecuada de la respuesta.
 - Trasforme la ecuación polinomial ortogonal en una ecuación en términos de herbicida e insecticida. Use un programa de regresión estándar, o bien las ecuaciones de transformación del capítulo 3.
 - Interprete los resultados de las gráficas de medias de celdas o la ecuación polinomial estimada.
6. Se realizó un experimento sobre la duración de tela recubierta sujeta a pruebas con abrasivos normales. El diseño factorial de $2 \times 2 \times 3$ incluyó dos sustancias distintas (F_1, F_2) en tres proporciones diferentes (25%, 50%, 75%) con y sin tratamiento de superficie (S_1, S_2); se probaron dos especímenes réplica de cada una de las 12 combinaciones en un diseño totalmente aleatorizado. Los datos corresponden a la pérdida de peso (mg) de los especímenes de tela por la prueba de abrasión.

Proporción de sustancia	Tratamientos de superficie y sustancia			
	S_1	S_2	F_1	F_2
25%	194	239	155	137
	208	187	173	160
50%	233	224	198	129
	241	243	177	98
75%	265	243	235	155
	269	226	229	132

Fuente: G. Box (1950), Problems in the analysis of growth and water curves. *Biometrics* 6, 362-289.

- Escriba un modelo lineal para el experimento, explique los términos y calcule el análisis de varianza.
- Prepare una tabla de medias de celdas y marginales con sus respectivos errores estándar.
- Pruebe la hipótesis nula para todos los efectos principales y de interacción.
- Calcule las particiones de regresión de las sumas de cuadrados con 1 grado de libertad para la proporción de sustancia y la interacción entre esa proporción y los otros factores.

- e. Pruebe las hipótesis nulas para las particiones de regresión.
- f. Grafique las celdas de medias contra la proporción de sustancia para las cuatro combinaciones de tratamiento de superficie y tipo de sustancia, e interprete los resultados del análisis.
7. Un científico de suelos realizó un experimento para evaluar una red de resistencias de cuatro electrodos y calcular la electroconducibilidad (EC) del suelo en celdas conductivas de acrílico especiales. El objetivo del estudio era evaluar la relación entre la EC medida y la salinidad del agua en el suelo con diferentes cantidades de agua. Se incluyeron tres texturas básicas de suelo, ya que la EC es específica de la textura; se usaron dos celdas para cada combinación de tratamiento y los tres tipos de suelo fueron arena arcillosa, arcilla y barro. El agua salina, en tres niveles, se basó en la EC del agua a 2, 8, y 16 dS/m (decisiemens/metro) y se establecieron tres niveles de contenido de agua en el suelo, 0%, 5% u 15%. El experimento resultante fue un arreglo factorial de $3 \times 3 \times 3$ con dos réplicas en un diseño totalmente aleatorizado; los siguientes son los valores de EC del suelo determinados con base en las lecturas de la red de cuatro electrodos.

Salinidad del agua	2			8			16		
	0	5	15	0	5	15	0	5	15
Arena	0.60	1.69	3.47	0.05	0.11	0.06	0.07	0.08	0.22
arcillosa	0.48	2.01	3.30	0.12	0.09	0.19	0.06	0.14	0.17
Arcilla	0.98	2.21	5.68	0.15	0.23	0.40	0.07	0.23	0.43
	0.93	2.48	5.11	0.26	0.35	0.75	0.21	0.35	0.35
Barro	1.37	3.31	5.74	0.72	0.78	2.10	0.40	0.72	1.95
	1.50	2.84	5.38	0.51	1.11	1.18	0.57	0.88	2.87

Fuente: H. Bohn y T. Tabbara, Department of Soil and Water Science, University of Arizona.

- a. Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
- b. Prepare una tabla de medias de celdas y marginales con sus respectivos errores estándar.
- c. Pruebe las hipótesis nulas para todos los efectos principales e interacciones.
- d. Calcule las particiones de sumas de cuadrados de regresión polinomial ortogonal lineal y cuadrática para salinidad, agua y sus interacciones, incluya las interacciones con la textura. Observe que los niveles de salinidad y agua tienen un espaciamiento desigual; por lo tanto, no se aplican los coeficientes polinomiales ortogonales estándar dados en la tabla XI del apéndice. Algunos programas estadísticos de computadora calculan de manera automática los coeficientes polinomiales ortogonales, dados los valores para los niveles de los factores (como MANOVA en SPSS y 2V en BMDP). A continuación se dan los coeficientes polinomiales ortogonales que se pueden usar para calcular las particiones ortogonales:

Agua lineal:	-0.617	-0.154	0.772
Agua cuadrática:	0.535	-0.802	0.267
Salinidad lineal:	-0.671	-0.067	0.738
Salinidad cuadrática:	0.465	-0.814	0.349

8. Se colocaron cinco varillas de níquel de 1 mm de diámetro en un sujetador metálico en una suspensión de óxido de aluminio y se aplicó una tensión de 100 volts entre las varillas de níquel y el contenedor con la suspensión de óxido de aluminio. Se registró el grueso de la capa de óxido de aluminio depositada en las varillas a tres posiciones de altura de las varillas; los datos que siguen son el grueso del depósito en micrones.

Altura	Posición del sujetador de varillas de níquel				
	1	2	3	4	5
1	125	130	128	134	143
2	126	150	127	124	118
3	130	155	168	159	138

Fuente: H. Hamaker (1955), Experimental design in industry.
Biometrics 11, 257-286.

- a. Escriba un modelo lineal para este experimento, explique los términos y establezca las suposiciones del modelo.
 - b. ¿Serán válidas las suposiciones del modelo para este experimento? Explique.
 - c. Admitiendo esas suposiciones como razonablemente válidas. Calcule el análisis de varianza para los datos.
 - d. Calcule las sumas de cuadrados con 1 grado de libertad para la no aditividad.
 - e. ¿Es suficiente el modelo aditivo de la posición del sujetador para estos datos?
9. Un entomólogo realizó un experimento sobre la energía consumida por las abejas de miel al beber, para determinar el efecto de la temperatura del ambiente y la viscosidad del líquido en el consumo de energía. Los niveles de temperatura fueron 20, 30 y 40°C, la viscosidad del líquido se controló por las concentraciones de sacarosa, que eran de 20, 40 y 60% del total de sólidos disueltos en el líquido que bebían las abejas de miel. El entomólogo registró la energía gastada por las abejas en joules/ segundo. Los datos que siguen corresponden a tres réplicas de cada uno de los nueve tratamientos en un diseño totalmente aleatorizado.

Temperatura (°C)	Sacarosa (%)		
	20	40	60
20	3.1, 3.7, 4.7	5.5, 6.7, 7.3	7.9, 9.2, 9.3
30	6.0, 6.9, 7.5	11.5, 12.9, 13.4	17.5, 15.8, 14.7
40	7.7, 8.3, 9.5	15.7, 14.3, 15.9	19.1, 18.0, 19.9

Fuente: Dr. S. Buckman, USDA Bee Research Lab, Tucson, Arizona.

- a. Calcule las particiones de regresión de sumas de cuadrados con 1 grado de libertad, las sumas de cuadrados para el % de sacarosa y la interacción.
- b. Pruebe la hipótesis nula para cada partición y determine la forma de la regresión polinomial que sea adecuada para describir la respuesta.
- c. Transforme la ecuación polinomial ortogonal en una ecuación en términos de la temperatura y el porcentaje de sacarosa. Utilice un programa de regresión estándar, o bien las ecuaciones de transformación del capítulo 3.
- d. Construya una gráfica de perfiles como la de la figura 6.5 con las medias de celdas estimadas y la ecuación polinomial estimada.
- e. Interprete los resultados.

10. *Números de réplicas desiguales.* Un biólogo incubó glándulas suprarrenales de ratas *in vitro* bajo estimulación de ACTH y midió su producción de esteroides, las glándulas se tomaron de animales en cuatro etapas distintas de crecimiento y se sujetaron a dos tratamientos diferentes. Se usaron glándulas de cuatro animales para cada combinación de tratamiento, pero varios análisis del laboratorio se invalidaron, por lo que se obtuvieron números de réplicas desiguales para los tratamientos. Los datos que siguen proporcionan la producción de esteroides por cada 100 mg de glándula por hora.

Etapa	Tratamiento	
	1	2
1	6.98, 6.58	8.62, 9.40, 9.20
2	6.07, 7.16, 6.34	9.42, 6.67, 8.64
3	5.38, 7.31, 6.65, 7.44	4.96, 6.80, 7.61
4	7.02, 9.23, 7.32	7.17, 7.65, 6.52, 6.86

Fuente: Dr. R. Chaisson, Department of Veterinary Science, University of Arizona.

- a. Calcule el análisis de varianza a fin de probar la hipótesis global de no interacción entre etapa y tratamiento.
 - b. Calcule las medias de mínimos cuadrados y sus errores estándar para las medias de celdas y marginales.
 - c. Estime la comparación entre los dos tratamientos para los niveles 1, 2, 3 y 4 de las etapas de crecimiento ($\widehat{\mu}_{1j} - \widehat{\mu}_{2j}$, para $j = 1, 2, 3, 4$) y pruebe la hipótesis de que no hay diferencia al nivel 0.5 de significancia entre las dos medias en cada caso.
 - d. ¿Cuál es la diferencia entre las medias de mínimos cuadrados y las medias observadas?
11. *Números de réplicas desiguales.* Suponga que el experimento del proceso de producción químico en el ejercicio 6.1 tenía réplicas desiguales entre las seis combinaciones de tratamiento de dos factores, base y alcohol. Los datos disponibles son los siguientes.

Base	Alcohol		
	1	2	3
1	90.7, 91.4	89.3, 88.1	89.5, 87.6
		90.4	88.3, 90.3
2	87.3, 88.3	94.7	93.1, 90.7
	91.5		91.5

- a. Calcule el análisis de varianza a fin de probar la hipótesis global de no interacción entre base y alcohol.
- b. Calcule las medias de mínimos cuadrados y sus errores estándar para las medias de celdas y marginales.
- c. Estime la comparación entre las dos bases para los niveles 1, 2 y 3 de alcohol ($\widehat{\mu}_{1j} - \widehat{\mu}_{2j}$; $j = 1, 2, 3$) y pruebe la hipótesis de que no hay diferencia al nivel .05 de significancia entre las dos medias en cada caso.
- d. ¿Cuál es la diferencia entre las medias de mínimos cuadrados y las medias observadas?

Apéndice 6A: Mínimos cuadrados para diseños factoriales

Réplicas iguales de tratamiento

Las particiones de sumas de cuadrados para los datos de un diseño factorial de tratamiento se pueden deducir de las soluciones de las ecuaciones de mínimos cuadrados para un modelo factorial. Se usará el modelo completo para un diseño factorial con dos factores para ilustrar el desarrollo.

Por simplicidad en la notación, el modelo completo se escribe como

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \quad (6A.1)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

donde μ es la media general, α_i es el efecto del factor A , β_j es el efecto del factor B , γ_{ij} es el efecto de interacción y e_{ijk} es el error experimental aleatorio independiente. El término de interacción $(\alpha\beta)_{ij}$ usado en el cuerpo principal de este capítulo se ha sustituido por γ_{ij} para simplificar la notación en la presentación de este apéndice.

Las estimaciones de los mínimos cuadrados para los parámetros en el modelo completo son aquellas que minimizan las sumas de cuadrados del error experimental.

$$Q = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r e_{ijk}^2 = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2 \quad (6A.2)$$

Las ecuaciones normales de la minimización incluyen una ecuación para μ y otra para cada efecto factorial, $\alpha_1, \alpha_2, \dots, \alpha_a; \beta_1, \beta_2, \dots, \beta_b$; y $\gamma_{11}, \gamma_{12}, \dots, \gamma_{ab}$. Las ecuaciones normales se obtienen a partir del siguiente conjunto de derivadas:

$$\frac{\partial}{\partial \mu} \sum_{i} \sum_{j} \sum_{k} (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2 = 0$$

$$\frac{\partial}{\partial \alpha_i} \sum_{j} \sum_{k} (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2 = 0 \quad i = 1, 2, \dots, a \quad (6A.3)$$

$$\frac{\partial}{\partial \beta_j} \sum_{i} \sum_{k} (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2 = 0 \quad j = 2, 3, \dots, b$$

$$\frac{\partial}{\partial \gamma_{ij}} \sum_{k} (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2 = 0 \quad i = 1, 2, \dots, a$$

$$j = 1, 2, \dots, b$$

Al simplificar, el conjunto de ecuaciones normales que debe resolverse es:

$$\begin{aligned}
 \mu: \quad abr\widehat{\mu} + br\sum_j \widehat{\alpha}_i + ar\sum_j \widehat{\beta}_j + r\sum_{i,j} \widehat{\gamma}_{ij} &= y \dots \\
 \alpha_i: \quad br\widehat{\mu} + br\widehat{\alpha}_i + r\sum_j \widehat{\beta}_j + r\sum_j \widehat{\gamma}_{ij} &= y_i \quad i = 1, 2, \dots, a \\
 \beta_j: \quad ar\widehat{\mu} + r\sum_i \widehat{\alpha}_i + ar\widehat{\beta}_j + r\sum_i \widehat{\gamma}_{ij} &= y_j \quad j = 1, 2, \dots, b \\
 \gamma_{ij}: \quad r\widehat{\mu} + r\widehat{\alpha}_i + r\widehat{\beta}_j + r\widehat{\gamma}_{ij} &= y_{ij} \quad i = 1, 2, \dots, a \\
 &\quad j = 1, 2, \dots, b
 \end{aligned} \tag{6A.4}$$

Después de una inspección cuidadosa, la suma de las a ecuaciones derivadas para los efectos del factor A es igual a la primera ecuación para μ ; la suma de las b ecuaciones derivadas para los efectos del factor B es igual a la primera ecuación para μ , lo mismo que las ab ecuaciones para la interacción; las γ_{ij} ecuaciones sumadas sobre el subíndice j darán la ecuación α_i y las γ_{ij} ecuaciones sumadas sobre el subíndice i darán la ecuación β_j . Estas dependencias lineales requieren restricciones impuestas sobre las estimaciones para proporcionar una solución única de las ecuaciones, cualesquiera restricciones que conduzcan a una solución serán suficientes. Un conjunto de restricciones de uso común es el de las restricciones de suma cero. Éstas son $\sum_i \widehat{\alpha}_i = 0$, $\sum_j \widehat{\beta}_j = 0$, $\sum_i^a \widehat{\gamma}_{ij} = 0$ ($j = 1, 2, \dots, b$) y $\sum_j^b \widehat{\gamma}_{ij} = 0$ ($i = 1, 2, \dots, a$).

Con las restricciones, las ecuaciones son:

$$\begin{aligned}
 \mu: \quad abr\widehat{\mu} &= y \dots \\
 \alpha_i: \quad br\widehat{\mu} + br\widehat{\alpha}_i &= y_i \quad i = 1, 2, \dots, a \\
 \beta_j: \quad ar\widehat{\mu} + ar\widehat{\beta}_j &= y_j \quad j = 1, 2, \dots, b \\
 \gamma_{ij}: \quad r\widehat{\mu} + r\widehat{\alpha}_i + r\widehat{\beta}_j + r\widehat{\gamma}_{ij} &= y_{ij} \quad i = 1, 2, \dots, a \\
 &\quad j = 1, 2, \dots, b
 \end{aligned} \tag{6A.5}$$

Las soluciones son:

$$\begin{aligned}
 \widehat{\mu} &= \frac{y}{abr} = \bar{y} \\
 \widehat{\alpha}_i &= \frac{y_i}{br} - \widehat{\mu} = \bar{y}_i - \bar{y} \quad i = 1, 2, \dots, a \\
 \widehat{\beta}_j &= \frac{y_j}{ar} - \widehat{\mu} = \bar{y}_j - \bar{y} \quad j = 1, 2, \dots, b \\
 \widehat{\gamma}_{ij} &= \frac{y_{ij}}{r} - \widehat{\mu} - \widehat{\alpha}_i - \widehat{\beta}_j = \bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y} \quad i = 1, 2, \dots, a \\
 &\quad j = 1, 2, \dots, b
 \end{aligned} \tag{6A.6}$$

La estimación de la suma de cuadrados para el error experimental se obtiene con una sustitución de las estimaciones $\hat{\mu}$, $\hat{\alpha}_i$, $\hat{\beta}_j$ y $\hat{\gamma}_{ij}$ en la ecuación (6A.2), como sigue:

$$SCE_f = \sum_{i,j,k} (y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij})^2 = \sum_{i,j,k} (y_{ijk} - \bar{y}_{ij})^2 \quad (6A.7)$$

Las diferencias entre la suma de cuadrados total y SCE_f se conoce como la reducción en la suma de cuadrados debida al ajuste del modelo y en ocasiones se escribe $R(\mu, \alpha, \beta, \gamma)$. Con réplicas iguales para todas las combinaciones de tratamiento la suma de cuadrados se puede deducir de la suma de cuadrados para cada efecto factorial a partir del cálculo usado para $R(\mu, \alpha, \beta, \gamma)$. La reducción en la suma de cuadrados debida al ajuste del modelo completo es:

$$R(\mu, \alpha, \beta, \gamma) = \hat{\mu} y + \sum_{i=1}^a \hat{\alpha}_i y_i + \sum_{j=1}^b \hat{\beta}_j y_j + \sum_{i=1}^a \sum_{j=1}^b \hat{\gamma}_{ij} y_{ij} \quad (6A.8)$$

Para datos balanceados con igual número de réplicas para cada combinación de tratamiento, las particiones de la suma de cuadrados para el análisis de varianza se pueden tomar de los términos individuales de la ecuación (6A.8), es decir:

$$\begin{aligned} CF &= \hat{\mu} y = \frac{(y)^2}{abr} \\ SCA &= \sum_{i=1}^a \hat{\alpha}_i y_i = br \sum_{i=1}^a (\bar{y}_i - \bar{y})^2 \\ SCB &= \sum_{j=1}^b \hat{\beta}_j y_j = ar \sum_{j=1}^b (\bar{y}_j - \bar{y})^2 \\ SC(AB) &= \sum_{i=1}^a \sum_{j=1}^b \hat{\gamma}_{ij} y_{ij} = r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 \end{aligned} \quad (6A.9)$$

Las particiones de sumas de cuadrados mostradas en la ecuación (6A.9) son las mostradas en la sección 6.4 y se pueden deducir de las consideraciones de los modelos completo y reducido. Por ejemplo, la suma de cuadrados de la interacción $SC(AB)$ se puede encontrar como la diferencia entre las sumas de cuadrados del error experimental para el modelo reducido sin los términos de interacción y el modelo completo que incluye los términos de interacción. Los modelos y las sumas de cuadrados son:

$$\text{Modelo completo: } y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \text{ con } SCE_f$$

$$\text{Modelo reducido: } y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \text{ con } SCE_r$$

La suma de cuadrados de la interacción se encuentra mediante $SC(AB) = SCE_r - SCE_f$, que es $\sum_i \sum_j \hat{\gamma}_{ij} y_{ij}$ [lo mismo que el último término de la ecuación (6A.8)].

La equivalencia de $SCE_r - SCE_f$ con $\sum_i \sum_j \hat{\gamma}_{ij} y_{ij}$ se puede demostrar si se resuelve la ecuación normal del modelo reducido y se calcula la reducción en la suma de cuadrados debida al ajuste del modelo reducido a $R(\mu, \alpha, \beta)$. Las

ecuaciones normales para el modelo reducido se obtienen de las del modelo completo en la ecuación (6A.5), eliminando las ecuaciones para γ_j y los términos $\widehat{\gamma}_j$ en las ecuaciones restantes. Las soluciones de $\widehat{\mu}$, $\widehat{\alpha}_i$ y $\widehat{\beta}_j$ son las mostradas en la ecuación (6A.6) y la reducción en la suma de cuadrados debida al ajuste del modelo reducido será:

$$R(\mu, \alpha, \beta) = \widehat{\mu} y + \sum_{i=1}^a \widehat{\alpha}_i y_{i.} + \sum_{j=1}^b \widehat{\beta}_j y_j \quad (6A.10)$$

Se observa que la diferencia entre $R(\mu, \alpha, \beta, \gamma)$ y $R(\mu, \alpha, \beta)$ es $\sum_i \sum_j \widehat{\gamma}_{ij} y_{ij}$, y, por lo tanto, las diferencias entre las sumas de cuadrados para el error experimental de los dos modelos será equivalente a la misma cantidad. Esto es:

$$SC(AB) = SCE_r - SCE_f = R(\mu, \alpha, \beta, \gamma) - R(\mu, \alpha, \beta) \quad (6A.11)$$

Las sumas de cuadrados SCA y SCB se pueden deducir de manera similar para los datos balanceados, o $SCA = R(\mu, \alpha) - R(\mu)$ y $SCB = R(\mu, \beta) - R(\mu)$.

Réplicas desiguales de tratamientos

La obtención de las sumas de cuadrados de la interacción con réplicas desiguales en el diseño factorial del tratamiento se ilustra con el ejemplo sencillo del factorial de 2×2 presentado en la tabla 6A.1.

Tabla 6A.1 Datos de ejemplo para el factorial de 2×2 con réplicas desiguales de tratamientos.

		<i>B</i>		<i>y_i</i>
		1	2	
<i>A</i>	1	6,5,3 <i>r₁₁</i> = 3	2, 4 <i>r₁₂</i>	20 <i>r₁</i> = 5
	2	5,4 <i>r₂₁</i> = 2	3 <i>r₂₂</i> = 1	12 <i>r₂</i> = 3
<i>y_j</i>		23 <i>r₁</i> = 5		32 <i>r₂</i> = 3 <i>r</i> = 8

Modelo completo

Las ecuaciones normales del modelo completo para los datos del ejemplo de la tabla 6A.1 se deducen mediante los métodos presentados al principio de este apéndice. Los coeficientes para los parámetros de las ecuaciones reflejarán números desiguales de réplicas. En general, las ecuaciones serán:

$$\mu: N\widehat{\mu} + \sum_{i=1}^a r_i \widehat{\alpha}_i + \sum_{j=1}^b r_j \widehat{\beta}_j + \sum_{r=1}^a \sum_{j=1}^b r_{ij} \widehat{\gamma}_{ij} = y$$

$$\alpha_i: r_i \widehat{\mu} + r_i \widehat{\alpha}_i + \sum_{j=1}^b r_{ij} \widehat{\beta}_j + \sum_{j=1}^b r_{ij} \widehat{\gamma}_{ij} = y_i$$

$$\beta_j: r_j \widehat{\mu} + \sum_{i=1}^a r_{ij} \widehat{\alpha}_i + r_j \widehat{\beta}_j + \sum_{i=1}^a r_{ij} \widehat{\gamma}_{ij} = y_j$$

$$\gamma_{ij}: r_{ij} \widehat{\mu} + r_{ij} \widehat{\alpha}_i + r_{ij} \widehat{\beta}_j + r_{ij} \widehat{\gamma}_{ij} = y_{ij}$$

Las ecuaciones normales para los datos de la tabla 6A.1 son:

$$\begin{aligned}
 8\widehat{\mu} + 5\widehat{\alpha}_1 + 3\widehat{\alpha}_2 + 5\widehat{\beta}_1 + 3\widehat{\beta}_2 + 3\widehat{\gamma}_{11} + 2\widehat{\gamma}_{12} + 2\widehat{\gamma}_{21} + \widehat{\gamma}_{22} &= 32 \\
 5\widehat{\mu} + 5\widehat{\alpha}_1 + 3\widehat{\beta}_1 + 2\widehat{\beta}_2 + 3\widehat{\gamma}_{11} + 2\widehat{\gamma}_{12} &= 20 \\
 3\widehat{\mu} + 3\widehat{\alpha}_2 + 2\widehat{\beta}_1 + \widehat{\beta}_2 + 2\widehat{\gamma}_{21} + \widehat{\gamma}_{22} &= 12 \\
 5\widehat{\mu} + 3\widehat{\alpha}_1 + 2\widehat{\alpha}_2 + 5\widehat{\beta}_1 + 3\widehat{\gamma}_{11} + 2\widehat{\gamma}_{21} &= 23 \\
 3\widehat{\mu} + 2\widehat{\alpha}_1 + \widehat{\alpha}_2 + 3\widehat{\beta}_2 + 2\widehat{\gamma}_{12} + \widehat{\gamma}_{22} &= 9 \\
 3\widehat{\mu} + 3\widehat{\alpha}_1 + 3\widehat{\beta}_1 + 3\widehat{\gamma}_{11} &= 14 \\
 2\widehat{\mu} + 2\widehat{\alpha}_1 + 2\widehat{\beta}_2 + 2\widehat{\gamma}_{12} &= 6 \\
 2\widehat{\mu} + 2\widehat{\alpha}_2 + 2\widehat{\beta}_1 + 2\widehat{\gamma}_{21} &= 9 \\
 \widehat{\mu} + \widehat{\alpha}_2 + \widehat{\beta}_2 + \widehat{\gamma}_{22} &= 3
 \end{aligned}$$

Como hay dependencias lineales en las ecuaciones, éstas se pueden sujetar a las restricciones de "suma cero" para obtener una solución. Las restricciones son:

$$\widehat{\alpha}_1 + \widehat{\alpha}_2 = 0, \widehat{\beta}_1 + \widehat{\beta}_2 = 0, \widehat{\gamma}_{11} + \widehat{\gamma}_{12} + \widehat{\gamma}_{21} = 0, \widehat{\gamma}_{11} + \widehat{\gamma}_{21} = 0$$

$$\widehat{\gamma}_{21} + \widehat{\gamma}_{22} = 0, \text{ y } \widehat{\gamma}_{12} + \widehat{\gamma}_{22} = 0$$

Después de aplicar las restricciones, la solución de las ecuaciones es

$$\widehat{\mu} = \frac{91}{24}, \widehat{\alpha}_1 = \frac{1}{24}, \widehat{\alpha}_2 = -\frac{1}{24}, \widehat{\beta}_1 = \frac{19}{24}, \widehat{\beta}_2 = -\frac{19}{24}$$

$$\widehat{\gamma}_{11} = \frac{1}{24}, \widehat{\gamma}_{12} = -\frac{1}{24}, \widehat{\gamma}_{21} = -\frac{1}{24}, \widehat{\gamma}_{22} = \frac{1}{24}$$

La suma de cuadrados para el error experimental se determina con:

$$SCE_f = \sum_{i,j,k} y_{ijk}^2 - R(\mu, \alpha, \beta, \gamma)$$

donde:

$$R(\mu, \alpha, \beta, \gamma) = \widehat{\mu}y_{...} + \sum_i \widehat{\alpha}_i y_i + \sum_j \widehat{\beta}_j y_j + \sum_{i,j} \widehat{\gamma}_{ij} y_{ij}$$

El cálculo para $R(\mu, \alpha, \beta, \gamma)$ es

$$\begin{aligned} R(\mu, \alpha, \beta, \gamma) &= \widehat{\mu}y_{...} + \widehat{\alpha}_1 y_1 + \widehat{\alpha}_2 y_2 + \widehat{\beta}_1 y_1 + \widehat{\beta}_2 y_2 \\ &\quad + \widehat{\gamma}_{11} y_{11} + \widehat{\gamma}_{12} y_{12} + \widehat{\gamma}_{21} y_{21} + \widehat{\gamma}_{22} y_{22} \\ &= \frac{1}{24} [91(32) + 1(20) + (-1)(12) + 19(23) + \dots + (-1)(9) + 1(3)] \\ &= 132.833 \end{aligned}$$

Dado $\Sigma y_{ijk}^2 = 140$, la suma de cuadrados para el error experimental a partir del modelo completo es

$$SCE_f = 140 - 132.833 = 7.167$$

Modelo reducido

Las ecuaciones normales para el modelo reducido sin los términos de la interacción γ_{ij} se obtienen eliminando las ecuaciones para γ_{ij} y los términos $\widehat{\gamma}_{ij}$ de las ecuaciones restantes del modelo completo. Las ecuaciones del modelo reducido son:

$$\begin{array}{lclclclcl} 8\widehat{\mu} & + & 5\widehat{\alpha}_1 & + & 3\widehat{\alpha}_2 & + & 5\widehat{\beta}_1 & + & 3\widehat{\beta}_2 = 32 \\ 5\widehat{\mu} & + & 5\widehat{\alpha}_1 & & & + & 3\widehat{\beta}_1 & + & 2\widehat{\beta}_2 = 20 \\ 3\widehat{\mu} & & & + & 3\widehat{\alpha}_2 & + & 2\widehat{\beta}_1 & + & \widehat{\beta}_2 = 12 \\ 5\widehat{\mu} & + & 3\widehat{\alpha}_1 & + & 2\widehat{\alpha}_2 & + & 5\widehat{\beta}_1 & & = 23 \\ 3\widehat{\mu} & + & 2\widehat{\alpha}_1 & + & \widehat{\alpha}_2 & & & + & 3\widehat{\beta}_2 = 9 \end{array}$$

Las restricciones de suma cero son $\widehat{\alpha}_1 + \widehat{\alpha}_2 = 0$ y $\widehat{\beta}_1 + \widehat{\beta}_2 = 0$. Las soluciones son:

$$\widehat{\mu} = \frac{212}{56}, \widehat{\alpha}_1 = \frac{3}{56}, \widehat{\alpha}_2 = -\frac{3}{56}, \widehat{\beta}_1 = \frac{45}{56}, \widehat{\beta}_2 = -\frac{45}{56}$$

La reducción en la suma de cuadrados debida al ajuste del modelo reducido es:

$$\begin{aligned} R(\mu, \alpha, \beta) &= \widehat{\mu}y_{...} + \widehat{\alpha}_1 y_1 + \widehat{\alpha}_2 y_2 + \widehat{\beta}_1 y_1 + \widehat{\beta}_2 y_2 \\ &= \frac{1}{56} [212(32) + 3(20) + (-3)(12) + 45(23) + (-45)(9)] \\ &= 132.821 \end{aligned}$$

La suma de cuadrados del error experimental para el modelo reducido es:

$$SCE_r = \sum_i \sum_j \sum_k y_{ijk}^2 - R(\mu, \alpha, \beta) = 140 - 132.821 = 7.179$$

La suma de cuadrados de la interacción es

$$SC(AB) = R(\mu, \alpha, \beta, \gamma) - R(\mu, \alpha, \beta) = 132.833 - 132.821 = 0.012$$

$$SC(AB) = SCE_f - SCE_r = 7.179 - 7.167 = 0.012$$

Suma de cuadrados del efecto principal

Las pruebas de igualdad para las medias marginales de A y B , $H_0: \bar{\mu}_1 = \bar{\mu}_2$ y $H_0: \bar{\mu}_1 \neq \bar{\mu}_2$, en ausencia de interacción, requieren particiones de sumas de cuadrados a partir del método de cuadrados medios ponderados (Yates, 1934). En la sección 6.9 se indicaron algunos programas de computadora que calculan las sumas de cuadrados requeridas. La hipótesis probada por la partición de sumas de cuadrados para un efecto principal depende mucho de la técnica de cálculo que se use en el proceso de estimación de mínimos cuadrados. Los detalles de los resultados de las distintas técnicas se pueden encontrar en Hocking y Speed (1975), Speed y Hocking (1976), Speed, Hocking y Hackney (1978) y Searle, Speed y Henderson (1981).

7 Diseños factoriales: modelos aleatorios y mixtos

En este capítulo se amplía el análisis de los componentes de la varianza a diseños más complejos, desarrollándose modelos varianza para algunas variaciones del diseño factorial; se introducen modelos de efectos aleatorios y modelos con mezclas de efectos fijos y aleatorios para arreglos factoriales. El concepto de diseño factorial se extiende para incluir experimentos con factores anidados con otros factores, se estudian diseños para experimentos que tienen una combinación de factores anidados y factores cruzados, incluyendo la información para determinar el número de réplicas y se proporcionan reglas para derivar los cuadrados medios esperados para una variedad de experimentos factoriales balanceados.

7.1 Efectos aleatorios para diseños factoriales

Los efectos aleatorios manejados en el capítulo 5 están dirigidos a estudios donde los niveles del factor eran muestras aleatorias de una población de niveles, en ellos el objetivo fue descomponer la varianza total en componentes identificables. La variabilidad causada por una fuente o factor puede depender de las condiciones en las que se evalúa, por que una parte de la varianza total se asocia con la interacción entre dos o más factores. El siguiente ejemplo ilustra la varianza para la interacción entre dos factores.

Ejemplo 7.1 Evaluación del desempeño de maquinaria especializada con los componentes de la varianza.

Un fabricante se encontraba desarrollando un nuevo espectrofotómetro de uso en laboratorios clínicos. El proceso de desarrollo estaba en la etapa piloto de ensam-

ble, luego de la cual debía evaluarse el desempeño de cada máquina en la línea de producción.

Pregunta de investigación: Una componente crítica del desempeño de instrumentos es la uniformidad de las mediciones de un día a otro entre las máquinas. En este caso específico, el científico que desarrolló el instrumento deseaba saber si la variabilidad de las mediciones entre máquinas operadas durante varios días estaba dentro de los estándares aceptables para las aplicaciones clínicas.

Diseño del tratamiento: El científico estableció un diseño factorial de tratamiento con “máquinas” y “días” como factores; debían probarse cuatro máquinas en cuatro días separados con un arreglo de 4×4 .

Diseño del experimento: Se seleccionaron cuatro máquinas al azar de la línea de producción piloto, cada día se prepararon ocho muestras de suero del mismo reactivo y se asignaron al azar dos muestras de suero a cada máquina cada día en un diseño totalmente aleatorizado con dos réplicas; el mismo técnico preparó las muestras de suero y operó las máquinas durante el experimento. En la tabla 7.1 se muestran las observaciones de niveles de triglicéridos (mg/dl) en las muestras de suero.

Tabla 7.1 Nivel de triglicéridos (mg/dl) en las muestras de suero examinadas en cuatro máquinas durante cada uno de los cuatro días

Día	Máquina			
	1	2	3	4
1	142.3, 144.0	148.6, 146.9	142.9, 147.4	133.8, 133.2
2	134.9, 146.3	145.2, 146.3	125.9, 127.6	108.9, 107.5
3	148.6, 156.5	148.6, 153.1	135.5, 138.9	132.1, 149.7
4	152.0, 151.4	149.7, 152.0	142.9, 142.3	141.7, 141.2

Fuente: Dr. J. Anderson, Beckman Instruments, Inc.

Las máquinas eran factores aleatorios porque representaban una muestra aleatoria de una población posible de máquinas que se fabricarían, y los “días” eran una muestra aleatoria de una población de días en los que las máquinas podrían operar. El arreglo factorial permitió evaluar la interacción entre máquinas y días. La uniformidad en el desempeño de las máquinas sería evidente con la ausencia de interacción.

Modelo estadístico para varianzas con dos factores

La variabilidad debida a la interacción de factores aleatorios puede tener un papel importante durante el proceso de inferencia. Un modelo de efectos aleatorios para el experimento con dos factores en un diseño totalmente aleatorizado es:

$$y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk} \quad (7.1)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

Se supone que los efectos aleatorios a_i , b_j y $(ab)_{ij}$ son independientes y tienen distribución normal con media 0 y varianzas σ_a^2 , σ_b^2 y σ_{ab}^2 , respectivamente; que los efectos deben ser independientes entre sí y que los errores aleatorios e_{ijk} son independientes y tienen distribución normal con media 0 y varianza σ^2 .

Las observaciones y_{ijk} en el modelo de efectos aleatorios tienen distribución normal con media μ y varianza:

$$\sigma_y^2 = \sigma^2 + \sigma_a^2 + \sigma_b^2 + \sigma_{ab}^2 \quad (7.2)$$

Las componentes de la varianza en la ecuación (7.2) se convierten en el punto central de cualquier investigación con efectos aleatorios.

En la tabla 7.2 se muestra el análisis de varianza para la concentración de triglicéridos medida en el suero para las pruebas del espectrómetro descritas en el ejemplo 7.1, donde los cálculos del análisis de varianza son los presentados en el capítulo 6 y los cuadrados medios esperados se incluyen en su tabla. En la sección 7.6 se proporciona un conjunto completo de reglas para determinar los cuadrados medios esperados aplicables a varios tipos de modelos factoriales que incluyen los modelos de efectos aleatorios.

Tabla 7.2 Análisis de varianza para las lecturas espectrométricas de cuatro máquinas en cada uno de cuatro días

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrado medio esperado
Día	3	$CMD = 445$	$\sigma^2 + r\sigma_{dm}^2 + rb\sigma_d^2$
Máquina	3	$CMM = 549$	$\sigma^2 + r\sigma_{dm}^2 + ra\sigma_m^2$
Interacción	9	$CM(DM) = 87$	$\sigma^2 + r\sigma_{dm}^2$
Error	16	$CME = 18$	σ^2

Estimadores puntuales para las componentes de la varianza

Con igual número de observaciones para cada combinación de tratamiento, se puede usar el análisis de varianza estudiado en la sección 5.3 para estimar sus componentes. Para el experimento factorial el análisis de varianza se calcula igual que para el modelo de efectos fijos, las estimaciones de las componentes se determinan igualando los cuadrados medios observados con los cuadrados medios esperados correspondientes y se despejan los valores desconocidos de las componentes. Las estimaciones de las cuatro componentes de la varianza en la tabla 7.2 son:

Error: $\widehat{\sigma}^2 = CME = 18.0$

Interacción: $\widehat{\sigma}_{dm}^2 = \frac{CM(\mathbf{DM}) - CME}{r} = \frac{87 - 18}{2} = 34.5$

Máquinas: $\widehat{\sigma}_m^2 = \frac{CMM - CM(\mathbf{DM})}{ra} = \frac{549 - 87}{2(4)} = 57.8$

Días: $\widehat{\sigma}_d^2 = \frac{CDM - CM(\mathbf{DM})}{rb} = \frac{445 - 87}{2(4)} = 44.8$

La estimación de la variación total para una sola observación es:

$$\widehat{\sigma}_y^2 = \widehat{\sigma}^2 + \widehat{\sigma}_d^2 + \widehat{\sigma}_m^2 + \widehat{\sigma}_{dm}^2 = 18.0 + 44.8 + 57.8 + 34.5 = 155.1$$

y la desviación estándar estimada es $\widehat{\sigma}_y = 12.5$.

Pruebas de hipótesis para los componentes de la varianza

La significancia de la contribución de las componentes de máquinas, días e interacción se puede evaluar con la prueba F , siendo el denominador del estadístico F_0 el cuadrado medio con el mismo valor esperado que el cuadrado medio en el numerador bajo la hipótesis nula. La prueba de no interacción, $H_0: \sigma_{dm}^2 = 0$, requiere que se use el cuadrado medio del error en el denominador, ya que tendrá el mismo valor esperado que $CM(\mathbf{DM})$ bajo la hipótesis nula. El estadístico es:

$$F_0 = \frac{CM(\mathbf{DM})}{CME} = \frac{87}{18} = 4.83$$

y se rechaza la hipótesis, nula de no interacción con $F_0 > F_{0.05,9,16} = 2.54$.

La situación es diferente para las pruebas que involucran a los componentes de la varianza para los efectos principales, los cuadrados medios esperados para los efectos principales son iguales a los cuadrados medios esperados para la interacción cuando la hipótesis nula es cierta, por lo tanto, el estadístico F_0 correcto para probar $H_0: \sigma_d^2 = 0$ es:

$$F_0 = \frac{CMD}{CM(\mathbf{DM})} = \frac{445}{87} = 5.11$$

y se rechaza la hipótesis nula para el componente “días” con $F_0 > F_{0.05,3,9} = 3.86$. De la misma manera, el estadístico para probar $H_0: \sigma_m^2 = 0$ es:

$$F_0 = \frac{CMM}{CM(\mathbf{DM})} = \frac{549}{87} = 6.31$$

y la hipótesis nula para el componente “máquinas” se rechaza con $F_0 > F_{0.05,3,9} = 3.86$.

Se puede calcular una estimación de intervalo para σ^2 como se muestra en la sección 5.4, ya que SCE/σ^2 es una variable ji-cuadrada con $ab(r - 1)$ grados de libertad.

Interpretación de las estimaciones de los componentes de la varianza

Cada componente contribuye de manera significativa a la variación de una medida para este modelo específico de espectrofolámetro: el componente error, $\hat{\sigma}^2 = 18.0$, representa la variación en la preparación de las muestras de suero; el componente máquinas, $\hat{\sigma}_m^2 = 57.8$, es la variabilidad del desempeño de la máquina y contribuye con el 37% de la variación; el componente días, $\hat{\sigma}_d^2 = 44.8$, es la variabilidad asociada con una nueva preparación de las máquinas con nuevos reactivos para el análisis de las muestras y otras fuentes de variabilidad que se pueden identificar con las diferencias en la operación diaria y el componente de interacción, $\hat{\sigma}_{dm}^2 = 34.5$, contribuye con 22% de la variación total. La interacción significativa implica que el desempeño relativo de varias máquinas no varía de manera congruente con los cambios en la operación de un día a otro y la falta de uniformidad en la calibración diaria de las máquinas es una posible explicación de la interacción.

El diseño factorial hizo posible identificar varias fuentes de variación en las mediciones que hace este modelo de espectrofotómetro y el investigador, según su experiencia, podrá decidir si alguna de las fuentes que contribuyen a la variabilidad excede un nivel aceptable y, de ser necesario, corregirá las deficiencias en la máquina o en las condiciones de operación.

Componentes de la varianza en estudios de tres factores

En la tabla 7.3 se muestran los cuadrados medios esperados en un experimento de tres factores con factores aleatorios. En la construcción del estadístico F_0 para las pruebas de hipótesis sobre los componentes de la varianza en modelos aleatorios con más de dos factores aparecen algunas complicaciones, pues el cuadrado medio del error se puede usar para probar la hipótesis de que no hay interacción de tres factores, y el cuadrado medio de la interacción de tres factores, $CM(ABC)$, para probar las hipótesis de las componentes de interacción de dos factores, pero al observar con cuidado los cuadrados medios esperados en la tabla 7.3, se puede ver que no hay un cuadrado medio legítimo para el denominador del estadístico F_0 al probar la hipótesis nula sobre las componentes de la varianza que corresponden al efecto principal.

Para algunas hipótesis se requiere una prueba F aproximada

Para probar la significancia de las componentes de la varianza del efecto principal es necesario construir un cuadrado medio para el denominador del estadístico F_0 (La construcción de un estadístico F_0 aproximado se estudia en el capítulo 5 para números desiguales de submuestras). Por ejemplo, para probar la hipótesis $H_0: \hat{\sigma}_b^2 = 0$, existen dos estadísticos posibles que se aproximan a la distribución F .

La primera aproximación se construye con CMB en el numerador y $M = CM(AB) + CM(BC) - CM(ABC)$ en el denominador y la segunda usa $M1 = CMB + CM(ABC)$ en el numerador y $M2 = CM(AB) + CM(BC)$ en el denominador.

Tabla 7.3 Cuadrados medios esperados para un experimento con tres factores aleatorios

Fuente de variación	Grados de libertad	Cuadrado medio esperado
<i>A</i>	$a - 1$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2 + r\sigma_{ac}^2 + r\sigma_{bc}^2$
<i>B</i>	$b - 1$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2 + r\sigma_{bc}^2 + r\sigma_{ac}^2$
<i>C</i>	$c - 1$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2 + r\sigma_{bc}^2 + r\sigma_{ab}^2$
<i>AB</i>	$(a - 1)(b - 1)$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2$
<i>AC</i>	$(a - 1)(c - 1)$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2$
<i>BC</i>	$(b - 1)(c - 1)$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{bc}^2$
<i>ABC</i>	$(a - 1)(b - 1)(c - 1)$	$\sigma^2 + r\sigma_{abc}^2$
Error	$abc(r - 1)$	σ^2

Tal vez sea más sencillo usar el primer cociente porque deben calcularse los grados de libertad aproximados para uno solo de los cuadrados medios mediante el procedimiento de Satterthwaite (ecuación 5.27), pero esto puede conducir a la construcción de un cuadrado medio negativo cuando algunas de las medias utilizadas quedan con signos negativos en la función; Gaylor y Hopper (1969) presentan algunos problemas asociados con aproximar la distribución *F* con combinaciones lineales de los cuadrados medios. El estadístico F_0 recomendado sería el segundo, ya que maneja los cuadrados medios sintetizados, $M1$ y $M2$.

7.2 Modelos mixtos

Muchos experimentos se diseñan para estudiar los efectos de un factor sobre la media de la población y los efectos de otro sobre la varianza de la misma, estos experimentos tienen un mezcla de factores fijos y aleatorios. Los modelos para arreglos factoriales que incluyen factores aleatorios y factores fijos se llaman **modelos mixtos**, pues contienen una mezcla de los dos tipos de efectos.

El modelo y análisis para los efectos mixtos se compone de dos partes porque hay dos tipos de inferencias. Las inferencias para el factor de efectos aleatorios se aplica a la variación en una población de efectos, mientras que las inferencias para los factores de efectos fijos están restringidas a los niveles específicos usados en el experimento. El experimento descrito en el siguiente ejemplo incluye una mezcla de factores aleatorios y fijos.



Ejemplo 7.2 Evaluación de dos métodos químicos en cuatro días.

Con frecuencia se desarrollan nuevos métodos químicos para probar compuestos en el contexto de un laboratorio clínico. Dado que tiene la posibilidad de elegir entre dos o más métodos químicos, el clínico debe evaluar el desempeño relativo de los métodos.

donde μ es la media, α_i es el efecto fijo del factor A , b_j es el efecto aleatorio del factor B , $(ab)_{ij}$ es el efecto de interacción y e_{ijk} es el error experimental aleatorio. Se supone que los efectos aleatorios, b_j y e_{ijk} , son independientes y siguen una distribución normal con media 0 y varianza σ_b^2 y σ_e^2 , respectivamente.

Se asume que los efectos de interacción $(ab)_{ij}$ son efectos aleatorios, independientes y siguen una distribución normal con media 0 y varianza σ_{ab}^2 . Se supone que los efectos de interacción son aleatorios cuando uno de los factores involucrados es efecto aleatorio.

Análisis de experimentos con factores mixtos

Los cuadrados medios esperados para el análisis de varianza de los modelos mixtos son diferentes a las de los completamente fijos o aleatorios. Para el experimento sobre el método químico del ejemplo 7.2 se ilustran los cuadrados medios esperados y el análisis del experimento factorial con un factor fijo y otro aleatorio en la tabla 7.5 se muestra el análisis de varianza para las 16 observaciones, incluyendo los cuadrados medios esperados en la tabla correspondiente. Observe el uso de θ_m^2 para la varianza del efecto fijo según se definió en la tabla 6.5.

Tabla 7.5 Análisis de varianza para un experimento factorial con un factor fijo, método, y otro aleatorio, día

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrados medios esperados
Método	1	$CMM = 329$	$\sigma^2 + r\sigma_{md}^2 + rb\theta_m^2$
Día	3	$CMD = 144$	$\sigma^2 + r\sigma_{md}^2 + ra\sigma_d^2$
Interacción	3	$CM(MD) = 62$	$\sigma^2 + r\sigma_{md}^2$
Error	8	$CME = 14$	σ^2

Fuente: Dr. J. Anderson, Beckman Instruments, Inc.

Pruebas de hipótesis para las componentes de la varianza y las medias

La hipótesis nula de no interacción, $H_0: \sigma_{md}^2 = 0$, se prueba con:

$$F_0 = \frac{CM(MD)}{CME} = \frac{62}{14} = 4.43$$

y la hipótesis nula se rechaza, ya que $F_0 > F_{0.05,3,8} = 4.07$. La presencia de interacción con los días sugiere la posibilidad de diferencias entre los métodos químicos que varían con los días. Se pueden hacer comparaciones entre los métodos químicos en un mismo día con contrastes entre las medias de celdas.

Sólamente en ausencia de interacción, sería adecuada una prueba de la hipótesis nula de que no hay diferencia entre las medias marginales del factor de efecto fijo. Con fines ilustrativos, el estadístico F_0 para los métodos químicos es:

$$F_0 = \frac{CMM}{CM(MD)} = \frac{329}{62} = 5.31$$

y la hipótesis nula no sería rechazada, ya que $F_0 < F_{05,1,3} = 10.13$, pues las medias observadas para los dos métodos químicos fueron $\bar{y}_1 = 147$ y $\bar{y}_2 = 138$ mg/dl.

El estadístico F_0 para probar la significancia de la componente de la varianza para días, σ_d^2 , es:

$$F_0 = \frac{CMD}{CM(DM)} = \frac{144}{62} = 2.32$$

y la hipótesis nula $H_0: \sigma_d^2 = 0$, no se rechazaría pues $F_0 < F_{05,3,3} = 9.28$.

Errores estándar para las medias del factor fijo

El error estándar para la diferencia entre los dos métodos químicos en un día dado:

$$\sigma_{(\bar{y}_j - \bar{y}_k)} = \sqrt{\frac{2(\sigma^2 + r\sigma_{md}^2)}{r}} \quad (7.4)$$

se estima mediante:

$$s_{(\bar{y}_j - \bar{y}_k)} = \sqrt{\frac{2CM(MD)}{r}} = \sqrt{\frac{2(62)}{2(4)}} = 7.9 \quad (7.5)$$

Por regla general, el cuadrado medio usado para el error estándar de una diferencia entre las medias de un factor de efecto fijo, es el cuadrado medio en el denominador de F_0 que se usó para probar la hipótesis nula sobre el efecto fijo. Entonces, el error estándar de la diferencia entre las medias marginales de los dos métodos químicos se estima con:

$$s_{(\bar{y}_i - \bar{y}_k)} = \sqrt{\frac{2CM(MD)}{rb}} = \sqrt{\frac{2(62)}{2(4)}} = 3.9 \quad (7.6)$$

Experimentos de tres factores con factores fijos y aleatorios

En la sección 7.1, se analiza el problema de construir el estadístico F_0 para probar las hipótesis en un experimento de tres factores con tres factores aleatorios. Para modelos de efectos mixtos con dos o más factores aleatorios se presentan diferencias similares. En la tabla 7.6 se muestran los cuadrados medios esperados para experimentos de tres factores que tienen uno o dos factores de efectos aleatorios y el resto de los factores de efecto fijo.

Tabla 7.6 Cuadrados medios esperados para experimentos de tres factores con 1) un factor de efecto fijo y dos de efecto aleatorio, y 2) dos factores de efecto fijo y uno de efecto aleatorio

Fuente de variación	Cuadrado medio esperado	
	A fijo, B y C aleatorios	A y B fijos, C aleatorio
A	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2 + r\sigma_{ac}^2 + rbc\theta_a^2$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2 + rbc\theta_a^2$
B	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2 + r\sigma_{bc}^2 + rac\sigma_b^2$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{bc}^2 + rac\theta_b^2$
C	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2 + r\sigma_{bc}^2 + rab\sigma_c^2$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2 + r\sigma_{bc}^2 + rab\sigma_c^2$
AB	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ab}^2$
AC	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{ac}^2$
BC	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{bc}^2$	$\sigma^2 + r\sigma_{abc}^2 + r\sigma_{bc}^2$
ABC	$\sigma^2 + r\sigma_{abc}^2$	$\sigma^2 + r\sigma_{abc}^2$
Error	σ^2	σ^2

Las dificultades encontradas en la construcción del estadístico F_0 para algunas hipótesis se hacen evidentes de inmediato con la inspección de los cuadrados medios esperados e incluso en algunos casos, no se dispone de una media cuadrática válida para el denominador de F_0 . Las medias cuadráticas para el estadístico F_0 deben sintetizarse como se describe en las secciones 5.11 y 7.1.

Modelos mixtos alternativos con restricciones sobre la interacción

Existen varias versiones de modelos mixtos basados en la definición usada para los efectos de interacción (vea Hocking (1973, 1985) y Searle *et al* (1992) los detalles técnicos). El modelo alternativo, que coloca una restricción sobre los efectos de interacción, es:

$$y_{ijk} = \mu + \alpha_i + g_j + (ag)_{ij} + e_{ijk} \quad (7.7)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

donde μ es la media, α_i es el efecto fijo del factor A, g_j es el efecto aleatorio del factor B, $(ag)_{ij}$ es el efecto de interacción y e_{ijk} es el error experimental aleatorio. Se supone que g_j y e_{ijk} son independientes, y tienen una distribución normal con medias 0 y varianzas σ_g^2 y σ_e^2 ; y que los efectos de interacción $(ab)_{ij}$ son aleatorios y tienen una distribución normal con media 0 y varianza $\frac{(a-1)}{a} \sigma_{ag}^2$.

Como uno de los factores, α_i , es fijo, el modelo alternativo tiene la suma del efecto de la interacción igual a cero para los niveles del factor fijo, de manera que:

$$\sum_{i=1}^a (ag)_{ij} = (ag)_j = 0$$

Con este modelo, la suma de los efectos de interacción para los niveles del factor aleatorio $\sum_j^b (ag)_{ij} = (ag)$, no será igual a cero, porque representan sólo una muestra aleatoria de los efectos de interacción en cada nivel del factor fijo. Sin embargo, en cualquier nivel dado del factor aleatorio existe un conjunto finito de efectos de interacción igual al número de niveles del factor fijo, y la suma $\sum_i^a (ag)_{ij} = (ag)_{\cdot j}$, es igual a cero. En consecuencia, existe una covarianza entre dos efectos de interacción al mismo nivel del efecto aleatorio y diferentes niveles del efecto fijo, que es $-\frac{1}{a} \sigma_{ag}^2$. Por ejemplo, en el experimento del ejemplo 7.2, las observaciones hechas el mismo día con dos modelos químicos estarán correlacionadas, pero las observaciones del mismo método químico en dos días distintos no tendrán correlación.

Así, la distinción primordial para el modelo alternativo es la presencia de correlación entre los efectos de interacción y en consecuencia, los valores esperados de algunas cuadrados medios son diferentes en él. Por ejemplo, con dos factores, A fijo y B aleatorio, las medias cuadráticas esperadas son:

$$\begin{aligned} E(CMA) &= \sigma^2 + r\sigma_{ag}^2 + rb\theta_a^2 \\ E(CMB) &= \sigma^2 + ra\sigma_g^2 \\ E[CM(AB)] &= \sigma^2 + r\sigma_{ag}^2 \\ E(CME) &= \sigma^2 \end{aligned} \tag{7.8}$$

El cuadrado medio esperado para el factor de efecto principal aleatorio no incluye el componente de interacción en la ecuación (7.8), mientras que previamente, en la tabla 7.5, se encontraba presente. Esta diferencia en el cuadrado medio esperado para el efecto principal aleatorio puede tener un impacto considerable en la inferencia estadística, por ejemplo, si se usara el modelo restringido para el ejemplo 7.2, una prueba de la hipótesis $H_0: \sigma_d^2 = 0$ requeriría el estadístico de prueba $F_0 = CMD/CME = 144/14 = 10.29$ y la hipótesis nula se rechazaría con $F_0 > F_{0.05,3,8} = 4.07$, que es justamente la conclusión opuesta a la obtenida con la prueba original para el modelo sin restricciones en la interacción.

Hocking (1973), estudió la relación entre los dos modelos y muestra que la relación entre los componentes de la varianza es:

$$\begin{aligned} \sigma_g^2 &= \sigma_b^2 + \frac{1}{a} \sigma_{ab}^2 \\ \sigma_{ag}^2 &= \sigma_{ab}^2 \end{aligned} \tag{7.9}$$

El modelo original sin restricciones sobre los términos de interacción supuso que los efectos de interacción aleatorios $(ab)_{ij}$ eran no correlacionados, con media 0 y varianza σ_{ab}^2 ; también que b_i y e_{ijk} son efectos aleatorios no correlacionados, con varianza σ_b^2 y σ^2 respectivamente, pero no se hizo la suposición de que la suma de $(ab)_{ij}$ sobre los niveles del factor fijo, $\sum_i^a (ab)_{ij} = (ab)_{\cdot j}$, fuera igual a cero.

Por lo tanto, debe hacerse una elección razonable del modelo más apropiado para la situación experimental. El modelo sin restricciones sobre la interacción tiene una ventaja importante: los cuadrados medios esperados para los datos no balanceados son congruentes con el modelo no restringido (Hartley y Searle, 1969). El modelo restringido no se considera en el caso desbalanceado. Si existe una posibilidad de correlación entre los efectos de un factor fijo para un nivel dado del efecto aleatorio, y los datos están balanceados, entonces el modelo restringido de interacción puede ser adecuado. Si no es así, o si los datos son no balanceados, entonces lo más apropiado será el modelo sin correlación entre los efectos de interacción expuesto en esta sección.

7.3 Diseños de factores anidados: una variación del tema

El diseño estándar factorial tiene dos características notables: cada nivel de cada factor ocurre con todos los niveles de los demás factores, y es posible examinar la interacción entre ellos.

En cierto tipo de estudios, los niveles de un factor, B , no serán idénticos en todos los niveles de otro factor, A . Cada nivel del factor A contendrá diferentes niveles del factor B .

Se dice que los niveles del factor B están anidados dentro de los niveles del factor A . En este caso, reciben el nombre de **diseños de factor anidado**, también se conocen como *diseños jerárquicos*. El siguiente ejemplo ilustra un diseño para estudiar las componentes de la varianza de factores anidados.

Ejemplo 7.3 Estándares de glucosa en química clínica.

Los laboratorios clínicos realizan análisis de suero humano que son críticos para el diagnóstico médico correcto. Los laboratorios mantienen programas de control de calidad para supervisar el desempeño de los análisis y cerciorarse de que el médico recibe información exacta para el diagnóstico.

Las fuentes de variación importantes en estos análisis son los días en los que se realizan, las corridas repetidas cada día y la preparación de las réplicas de muestras de suero en las corridas. El programa de control de calidad requiere que se pruebe un espectrofotómetro con varias corridas cada día, usando suero normal para las corridas de control. Las preparaciones de las réplicas de suero se evalúan dentro de cada corrida.

Los datos de la tabla 7.7 son las observaciones de un diseño usado para controlar la calidad de los análisis de glucosa. Los sueros con glucosa normal se guardan en el laboratorio justo para las corridas de control de calidad. Se hicieron $c = 3$ réplicas del estándar preparado para cada una de las $b = 2$ corridas en cada uno de los $a = 3$ días.

Se trata de un diseño anidado con dos corridas independientes y únicas cada día. El anidado de las corridas ocurre porque una corrida de cualquier día no tiene relación con la de otro; por ejemplo, la primera corrida en día 1 no tiene nada en común con las primeras corridas de los días 2 y 3. En la tabla 7.7 las corridas se numeraron del 1 al 6 para reflejar su independencia mutua entre los diferentes días.

De la misma manera, las preparaciones de las réplicas de suero están anidadas dentro de las corridas. Los días, corridas y réplicas representan factores en el primero, segundo y tercer nivel de la jerarquía.

Tabla 7.7 Glucosa (mg/dl) en los estándares de control de calidad

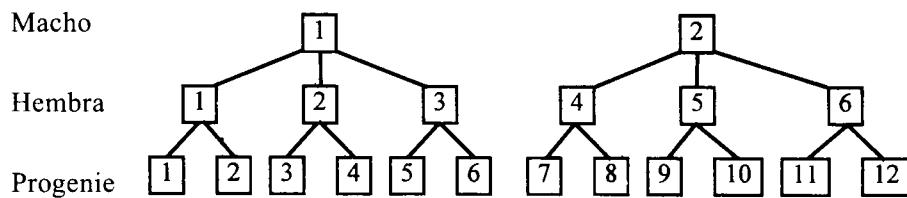
<i>Día 1</i>		<i>Día 2</i>		<i>Día 3</i>	
<i>Corrida 1</i>	<i>Corrida 2</i>	<i>Corrida 3</i>	<i>Corrida 4</i>	<i>Corrida 5</i>	<i>Corrida 6</i>
42.5	42.2	48.0	42.0	41.7	40.6
43.3	41.4	44.6	42.8	43.4	41.8
42.9	41.8	43.7	42.8	42.5	41.8
Media diaria (\bar{y}_i)	42.4	44.0		42.0	

Fuente: Dr. J. Anderson, Beckman Instruments, Inc.

Otros ejemplos con factores anidados

Consideremos un estudio de genética en animales donde cada macho (padre) se cruza con una muestra aleatoria de hembras (madres), y cada cruce produce una camada de varios descendientes como se muestra en el cuadro 7.1. Como los machos, las hembras y los descendientes representan los factores de interés en el estudio, los descendientes o progenie de una hembra son diferentes de los de otras, la progenie está anidada dentro de las hembras y los factores en el diseño anidado forman una jerarquía. Las jerarquías o anidado de los factores se ilustran en el cuadro 7.1. El nivel más alto de la jerarquía representa los machos, seguido de las hembras y la progenie en el segundo y tercer nivel, respectivamente.

Cuadro 7.1 Diseño de factores anidados en un estudio de genética.



Se tienen $b = 3$ hembras distintas cruzadas con cada uno de los $a = 2$ machos que dan un total de $ab = 6$ hembras, y se tienen $c = 2$ descendientes por hembra, con los que se obtiene un total de $abc = 12$ descendientes.

Los diseños anidados suelen presentarse en aquellas investigaciones educativas que utilizan varias escuelas elementales. Los salones de clase están anidados dentro de las escuelas y los estudiantes están anidados dentro de los salones de clase.

Un experimento sobre fórmulas para teñido de telas requiere que varias réplicas de cada fórmula se mezclen de manera independiente y que cada lote se pruebe con varios especímenes de tela común. Los lotes réplica están anidados dentro de las fórmulas de teñido y los especímenes de tela están anidados dentro de los lotes.

Modelos estadísticos para los factores anidados

Los factores en la jerarquía del diseño anidado pueden ser fijos o aleatorios. El diseño para las normas de glucosa en el ejemplo 7.3 tendría todos los factores aleatorios si días, corridas y réplicas de preparaciones de suero se consideraran muestras aleatorias de sus respectivas poblaciones.

El modelo lineal para un diseño anidado con tres factores aleatorios anidados, A , B dentro de A y C dentro de B , es

$$y_{ijk} = \mu + \alpha_i + b_{j(i)} + c_{k(j)} \quad (7.10)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, c$$

donde a_i es el efecto del factor A , $b_{j(i)}$ es el efecto del factor B anidado en A , y $c_{k(j)}$ es el efecto del factor C anidado en B y el subíndice $j(i)$ se refiere al factor representado por el subíndice j anidado en el factor representado por el subíndice i . Se supone que los efectos a_i , $b_{j(i)}$ y $c_{k(j)}$ son aleatorios e independientes, con medias 0 y varianzas σ^2_a , $\sigma^2_{b(a)}$ y $\sigma^2_{c(b)}$, respectivamente.

En el estudio de genética del cuadro 7.1, el factor A es un efecto fijo si sólo se dispone de dos machos para el estudio y el investigador desea restringir a ellos los resultados genéticos, y los efectos de la hembra son aleatorios si representan seis hembras elegidas al azar de una población potencial de hembras. Si los efectos del factor A son fijos, entonces el efecto aleatorio a_i mostrado en la ecuación (7.10) se sustituye por la notación de efecto fijo α_i .

En un estudio, ambos factores pueden tener efectos fijos. Supongamos que el departamento de bomberos de un área metropolitana grande desea evaluar el efecto de dos tipos de rotación de brigadas sobre su eficiencia en seis áreas, un tipo de rotación se prueba en tres de los distritos elegidos al azar y el otro en los tres restantes. En cada áreas se prueba una muestra aleatoria de brigadas después del periodo de evaluación. El factor rotación, A , es fijo porque sólo dos tipos de rotación de brigadas están bajo consideración; también el factor área, B , es fijo porque sólo se estudian las seis áreas que existen en el área metropolitana. Si los factores A y B son fijos, entonces los efectos aleatorios a_i y $b_{j(i)}$ se sustituyen en la ecuación (7.10) por los efectos fijos α_i y $\beta_{j(i)}$.

Análisis de factores anidados aleatorios

Los objetivos de los estudios que utilizan el diseño anidado de efectos aleatorios dependen del tema del estudio, pero comprenden la estimación de los componentes de la varianza y las pruebas de hipótesis sobre éstas. Las particiones de las sumas de cuadrados del análisis de varianza se utilizan para los procedimientos de

estimación y de prueba los cálculos del análisis de varianza son idénticos a los de las submuestras que se muestran en la tabla 5.5 Los cuadrados medios esperados con todos los efectos aleatorios están dadas en el análisis de varianza abreviado descrito en la tabla 7.8.

Tabla 7.8 Cuadrados medios esperados para el análisis de varianza de un diseño anidado con tres factores aleatorios, A , B y C

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
Total	$abc - 1$		
A	$a - 1$	CMA	$\sigma_{c(b)}^2 + c\sigma_{b(a)}^2 + bc\sigma_a^2$
B dentro de A	$a(b - 1)$	$CM(B/A)$	$\sigma_{c(b)}^2 + c\sigma_{b(a)}^2$
C dentro de B	$ab(c - 1)$	$CM(C/B)$	$\sigma_{c(b)}^2$

Estimadores puntuales para los componentes de la varianza

Las estimaciones varianza de los componentes de la varianza se encuentran igualando las medias cuadráticas observadas con los cuadrados medios esperados y despejando. Los tres estimadores para el análisis de la tabla 7.8 son:

$$\begin{aligned}\widehat{\sigma}_{c(b)}^2 &= CM(C/B) \\ \widehat{\sigma}_{b(a)}^2 &= \frac{[CM(B/A) - CM(C/B)]}{c} \end{aligned}\quad (7.11)$$

y

$$\widehat{\sigma}_a^2 = \frac{[CMA - CM(B/A)]}{bc}$$

Las hipótesis nulas de interés son $H_0: \sigma_a^2 = 0$ si las a_i son efectos aleatorios, y $H_0: \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_t$ para efectos fijos con el estadístico $F_0 = MCA/MC(B/A)$. La hipótesis $H_0: \sigma_{b(a)}^2 = 0$ se prueba con $F_0 = CM(B/A)/CM(C/B)$.

Normas de control de calidad para el análisis de glucosa

El análisis de varianza para los estándares de glucosa del ejemplo 7.3 se muestra en la tabla 7.9. Muchos programas de computadora imprimen una tabla de los cuadrados medios esperadas y las estimaciones de las componentes de la varianza si se les solicita. El coeficiente para la componente día (σ_a^2) es $bc = 6$, mientras que el coeficiente para la componente corrida/día es $c = 3$.

Tabla 7.9 Análisis de varianza para las normas de control de calidad sobre la glucosa

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Valor F	Pr > F
Total	17	43.88			
Día	2	13.76	6.88	1.26	.400
Corrida/día	3	16.36	5.45	4.75	.021
Réps/corrida	12	13.76	1.15		

Las estimaciones de los componentes de la varianza son:

$$\text{Réps: } \widehat{\sigma}_{c(b)}^2 = CM(\text{Rep}/\text{Corrida}) = 1.15$$

$$\text{Corrida: } \widehat{\sigma}_{b(a)}^2 = \left[\frac{CM(\text{Corrida/Día}) - CM(\text{Rep}/\text{Corrida})}{c} \right] = \frac{(5.45 - 1.15)}{3} = 1.43$$

$$\text{Día: } \widehat{\sigma}_a^2 = \frac{[CM\text{Día} - CM(\text{Corrida/Día})]}{bc} = \frac{(6.88 - 5.45)}{6} = 0.24$$

La estimación de la varianza total para un análisis de glucosa estándar es:

$$\widehat{\sigma}_y^2 = \widehat{\sigma}_a^2 + \widehat{\sigma}_{b(a)}^2 + \widehat{\sigma}_{c(b)}^2 = 0.24 + 1.43 + 1.15 = 2.82$$

con desviación estándar $\widehat{\sigma}_y = 1.68$. Cerca del 9% de la variación se atribuye a la variación de un día a otro, $\widehat{\sigma}_a^2$; 51% a la variación de las corridas de un día, $\widehat{\sigma}_{b(a)}^2$, y 41% a la variación entre las réplicas de una corrida, $\widehat{\sigma}_{c(b)}^2$. La gran media del estudio fue $\bar{y} = 42.8$, y el porcentaje del coeficiente de variación de los estándares de glucosa en este conjunto de corridas fue $\%CV = (100)(\widehat{\sigma}_y/\bar{y}) = (100)(1.68/42.8) = 3.9\%$.

Errores estándar para medias

Las varianzas para la gran media del estudio \bar{y} y una media del día \bar{y}_i son:

$$\sigma_{\bar{y}}^2 = \frac{\sigma_{c(b)}^2}{abc} + \frac{\sigma_{b(a)}^2}{ab} + \frac{\sigma_a^2}{a} \quad \text{y} \quad \sigma_{\bar{y}_i}^2 = \frac{\sigma_{c(b)}^2}{bc} + \frac{\sigma_{b(a)}^2}{b} \quad (7.12)$$

respectivamente. Las estimaciones son:

$$s_{\bar{y}}^2 = \frac{CM \text{ Día}}{abc} = \frac{6.88}{(3)(2)(3)} = 0.38$$

y

$$s_{\bar{y}_{1.}}^2 = \frac{CM(\text{Corrida/Día})}{bc} = \frac{5.45}{(2)(3)} = 0.91$$

Las estimaciones de los errores estándar son $s_{\bar{y}} = \sqrt{0.38} = 0.62$ y $s_{\bar{y}_{1.}} = \sqrt{0.91} = 0.95$.

Las varianzas, desviaciones estándar, errores estándar, coeficientes de variación y medias, son estadísticos útiles para la supervisión del control de calidad. Los registros acumulados de estos estadísticos con frecuencia se almacenen en un laboratorio y se inspeccionan con regularidad, si los valores se desvían de algunas normas establecidas, el analista debe investigar el proceso en busca de la fuente o causa de desviación.

Pruebas de hipótesis sobre varianzas

La prueba de la hipótesis nula para la componente de varianza “día” es:

$$F_0 = \frac{CM \text{ Día}}{CM(\text{Corrida/Día})} = \frac{6.88}{5.45} = 1.26$$

y no es significativa con $Pr > F = .400$ (tabla 7.9). La prueba de la hipótesis nula para la componente “corrida en un día” es:

$$F_0 = \frac{CM(\text{Corrida/Día})}{CM(\text{Rep/Corrida})} = \frac{5.45}{1.15} = 4.75$$

y es significativa con $Pr > F = .021$ (tabla 7.9). Estos resultados indican que el analista puede aumentar la precisión de manera efectiva si concentra sus esfuerzos en reducir la variación entre las corridas del mismo día y las réplicas dentro de las corridas para obtener estimaciones más precisas.

Análisis para efectos de factor fijo

Los estadísticos F_0 para las pruebas de hipótesis con efectos fijos en el diseño anidado se pueden determinar a partir de los cuadrados medios esperados que se muestran en la tabla 7.10. Los efectos de factor fijo se definen de manera que $\sum_i \alpha_i = 0$ y $\sum_j \beta_{j(i)} = 0$ para $i = 1, 2, \dots, a$.

Tabla 7.10 Cuadrados medios esperados para el análisis de varianza de un diseño anidado con factores fijos y mixtos para A y B .

Cuadrados medios	A y B fijos	A fijo y B aleatorio
CMA	$\sigma_{c(b)}^2 + bc \sum \alpha_i^2 / (a - 1)$	$\sigma_{c(b)}^2 + c\sigma_{b(a)}^2 + bc \sum \alpha_i^2 / (a - 1)$
$CM(B/A)$	$\sigma_{c(b)}^2 + c \sum \beta_{i(i)}^2 / a(b - 1)$	$\sigma_{c(b)}^2 + c\sigma_{b(a)}^2$
$CM(C/B)$	$\sigma_{c(b)}^2$	$\sigma_{c(b)}^2$

Con A fijo y B aleatorio, la hipótesis nula para los efectos del factor A , $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_a$, se prueba con $F_0 = CMA/CM(B/A)$. Cuando ambos factores

son fijos, la hipótesis nula para los efectos del factor A se prueba con $F_0 = CMA/CM(C/B)$. La hipótesis nula para los efectos del factor B con ambos factores fijos es $H_0: \beta_{1(i)} = \beta_{2(i)} = \dots = \beta_{b(i)}$ para toda i , y se prueba con $F_0 = CM(B/A)/CM(C/B)$.

Diseño anidados desequilibrados para igualar información sobre las varianzas

El diseño de factores anidados contiene más información de los factores en niveles bajos de la jerarquía del diseño que en los niveles altos. En estudios grandes, las discrepancias en los grados de libertad entre las fuentes de variación pueden ser considerables, por lo que se desarrollaron los *diseños anidados desequilibrados* para igualar los grados de libertad correspondientes a los cuadrados medios en cada nivel de la jerarquía.

Los diseños desequilibrados tienen números de niveles diferentes para factores que están anidados dentro de otros, los niveles para el factor B anidado en el factor A varían de un nivel a otro del factor A de manera que los grados de libertad para CMA y $CM(B/A)$ se parecen más. De la misma manera, los niveles del factor C anidados en el factor B pueden variar con los niveles de B para lograr grados de libertad para $CM(C/B)$ similares a los de otros cuadrados medios.

Anderson (1960) y Bainbridge (1963) proporcionaron los primeros resultados sobre el uso de diseños desequilibrados, Smith y Beverly (1981) proporcionan una discusión general sobre el uso y análisis de diseño desequilibrados y Goldsmith y Gaylor (1970) enumeraron 61 diseños desequilibrados para tres etapas de factores (A , B y C), tales que las componentes de la varianza se pueden estimar mediante el método de análisis. Los diseños enumerados por Goldsmith y Gaylor incluyen dos o tres de las cinco posibles estructuras fundamentales mostradas en el cuadro 7.2.

Cuadro 7.2 Cinco estructuras fundamentales para diseños anidados desequilibrados de tres etapas (los niveles de B dentro de A son b_i y los niveles de C dentro de B son c_{ij})

Etapa	Estructuras				
	1	2	3	4	5
A					
B/A					
C/B					
b_i	2	2	2	1	1
c_{ij}	2	2, 1	1	2	1

El análisis de los diseños desequilibrados de tres etapas siguen el patrón establecido para el análisis con muestras y réplicas desiguales en la tabla 5.8. Leone *et al* (1968), proporcionaron una fórmula computacional para los diseños anidados de cuatro etapas con niveles desiguales en todas las etapas, también establecieron fórmulas para las sumas de cuadrados del análisis de varianza, los cuadrados medios esperados y las estimaciones de las varianzas para las medias en las diversas etapas

del diseño ($\bar{y}_{...}$, $\bar{y}_{i...}$, $\bar{y}_{ij..}$ y $\bar{y}_{ijk..}$); en tanto que Gates y Shiue (1962) proporcionaron fórmulas para calcular las sumas de cuadrados del análisis de varianza y los cuadradoss medios esperados para una clasificación jerárquica general de S etapas.

Ejemplo 7.4 Un diseño desequilibrado para muestras de suelo

Los datos del ejercicio 5.8 fueron generados a partir de un diseño anidado desequilibrado para estimar los componentes de la varianza para las características de las muestras de suelo. Los tres factores en el diseño fueron campos (F), secciones (C) y localidades dentro de las secciones (L). Supongamos que se muestrean $a = 12$ campos y $b_i = 2$ secciones por campo y que en los campos 1 a 6 se muestrean $c_{ij} = 2$ localidades por sección y en los campos 7 a 12 se muestrean $c_{ij} = 1$ localidades por sección. Este diseño desequilibrado tiene seis réplicas en las estructuras 1 a 3 del cuadro 7.2 como se muestra en el 7.3. El diseño es uno de los enumerados por Goldsmith y Gaylor (1970).

Cuadro 7.3 Diseño anidado desequilibrado para muestras de suelo

	<i>Estructura 1</i>	<i>Estructura 3</i>
	<i>Campos 1 a 6</i>	<i>Campos 7 a 12</i>
Campo	1	7
Sección	1 2	13 14
Ubicación	1 2 1 2	1 1 1
b_i	2	2
c_{ij}	2	1
c_{ij}	4	2

El análisis de varianza descrito para este diseño, mostrado en la tabla 7.11, sigue el formato mostrado en la tabla 5.8 y utiliza la ecuación (5.23) para calcular los coeficientes de los componentes de la varianza en los cuadradoss medios esperados, los grados de libertad casi iguales para los cuadradoss medios proporcionan información casi igual en cada fuente de variación y las componentes de la varianza se estiman mediante el siguiente método de análisis:

$$\widehat{\sigma}_{c(b)}^2 = CM(L/S), \quad \widehat{\sigma}_{b(a)}^2 = \frac{[CM(S/F) - CM(L/S)]}{1.50} \quad (7.13)$$

$$\widehat{\sigma}_a^2 = \frac{[1.50CMF - 1.48CM(S/F) - (1.50 - 1.48)CM(L/S)]}{(2.97)(1.50)}$$

Tabla 7.11 Descripción del análisis de varianza con cuadrados medios esperados para el diseño anidado desequilibrado

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrados medios esperados
Campo	$a - 1 = 11$	CMF	$\sigma_{c(b)}^2 + 1.48\sigma_{b(a)}^2 + 2.97\sigma_a^2$
Secciones	$\sum_{i=1}^a b_i - a = 12$	$CM(S/F)$	$\sigma_{c(b)}^2 + 1.50\sigma_{b(a)}^2$
Ubicación	$N - \sum_{i=1}^a b_i = 12$	$CM(L/S)$	$\sigma_{c(b)}^2$

7.4 Diseños de factores cruzados y anidados

Ciertas condiciones experimentales dan lugar a arreglos factoriales que contienen factores cruzados y anidados. En este caso, algunos factores están cruzados con el arreglo factorial de niveles usual, mientras que otros están anidados dentro de las celdas del arreglo factorial o dentro de los niveles de al menos uno de los factores. En ocasiones estos diseños reciben el nombre de *diseño factorial anidado* (Anderson y McLean, 1974; Hicks, 1973; Smith y Beverly, 1981).

Ejemplo 7.5 Evaluación del espectrofotómetro

Las compañías que fabrican máquinas e instrumentos tienen departamentos de investigación y desarrollo para producir nuevos instrumentos o mejorar los actuales. Las pruebas de desempeño forman parte de la etapa de desarrollo de cualquier máquina, en ellas se prueban las funciones mecánicas o eléctricas de las máquinas, su exactitud en el desempeño de las funciones asignadas y otros aspectos.

Problema de investigación: En este contexto, un investigador desarrollaba un nuevo espectrofotómetro para uso en laboratorios médicos, se había construido un modelo según el diseño propuesto y estaba listo para la evaluación de sus capacidades en un laboratorio. Era necesario determinar si ese diseño en particular establecía las propiedades del espectro en el intervalo requerido por los estándares de glucosa en suero, por lo que el investigador debía establecer si la variabilidad y uniformidad de los resultados en múltiples corridas diarias quedaban dentro de las especificaciones.

Diseño del tratamiento: Se usó un diseño de tratamiento factorial con “concentraciones” de glucosa y “días” como factores y se elaboraron muestras de suero con diferentes niveles de glucosa para cubrir el intervalo de concentraciones que debía analizar el instrumento. Las tres concentraciones se analizaron cada día, de manera que las concentraciones se cruzaron con los días en un arreglo factorial

de 3×3 se realizaron dos corridas del instrumento cada día, con lo que las corridas quedaron anidadas dentro de los días.

Diseño del experimento: Se prepararon cuatro muestras réplica de suero con cada concentración de glucosa cada día, se asignaron al azar dos muestras de cada concentración a cada corrida diaria y las seis muestras se analizaron en orden aleatorio en cada corrida. El mismo técnico preparó las muestras y operó el instrumento durante el experimento.

El diseño tenía factores cruzados y anidados con $a = 3$ concentraciones cruzadas, con $b = 3$ días, con $c = 2$ corridas anidadas en cada día y $r = 2$ réplicas de suero preparadas para cada concentración de cada corrida. Las concentraciones de glucosa observadas en el espectrómetro se muestran en la tabla 7.12.

Tabla 7.12 Concentraciones de glucosa (mg/dl) para tres muestras de concentraciones estándar de dos corridas de un espectrofotómetro en cada uno de tres días

Concentración	Día 1		Día 2		Día 3	
	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5	Corrida 6
1	41.2	41.2	39.8	41.5	41.9	45.5
	42.6	41.4	40.3	43.0	42.7	44.7
2	135.7	143.0	132.4	134.4	137.4	141.1
	136.8	143.3	130.3	130.0	135.2	139.1
3	163.2	181.4	173.6	174.9	166.6	175.0
	163.3	180.3	173.9	175.6	165.5	172.0

Fuente: Dr. J. Anderson, Beckman Instruments, Inc.

Modelo estadístico para factores cruzados y anidados

El modelo estadístico para este experimento en particular es:

$$y_{ijkl} = \mu + \alpha_i + b_j + c_{k(j)} + (ab)_{ij} + (ac)_{ik(j)} + e_{ijkl} \quad (7.14)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, c \quad l = 1, 2, \dots, r$$

donde α_i es el efecto fijo para la concentración, b_j es el efecto aleatorio del día, $c_{k(j)}$ es el efecto aleatorio de las corridas anidadas en el día, $(ab)_{ij}$ es el efecto aleatorio para la concentración por día, y e_{ijkl} es el error experimental aleatorio. Las suposiciones para los efectos son congruentes con las indicadas en la sección anterior para modelos aleatorios, mixtos y anidados.

Los efectos del modelo para los dos factores cruzados, concentración y día, siguen la convención usual con los efectos principales y de interacción; el efecto del modelo para las corridas dentro de los días sigue la convención usual para un factor anidado en otro factor y el efecto de interacción para la concentración por corrida anidada en el día es la característica nueva en este modelo. Como cada

concentración se evalúa en cada corrida, los dos factores constituyen un arreglo factorial completo cada día. Por lo que, la concentración por corrida puede evaluarse por día y anidarse o unirse a través de los días, ya que las corridas anidadas dentro de los días son corridas únicas en cada uno.

Cuadrados medios esperados

El modelo mostrado en la ecuación (7.14) es un modelo mixto con uno de los factores cruzados aleatorio y el otro fijo. Los cuadrados medios esperados para el análisis de varianza están afectadas por el modelo supuesto para los factores cruzados y la manera en la que otros factores se anidan en el experimento; en consecuencia, es posible que exista una variedad de patrones en los cuadrados medios esperados. Los cuadrados medios esperados para los modelos con A y B fijos o con A fijo y B aleatorio se muestran en la tabla 7.13.

Tabla 7.13 Cuadrados medios esperados para un factorial anidado con A y B cruzados y C anidado dentro de B

Fuente de variación	Grados de libertad	Cuadrado medio esperado	
		A fijo, B y C aleatorios	A y B fijos, C aleatorio
Total	$abcr - 1$		
A	$a - 1$	$\sigma^2 + r\sigma_{ac(b)}^2 + cr\sigma_{ab}^2 + bcr\theta_a^2$	$\sigma^2 + r\sigma_{ac(b)}^2 + bcr\theta_a^2$
B	$b - 1$	$\sigma^2 + r\sigma_{ac(b)}^2 + ar\sigma_{c(b)}^2 + cr\sigma_{ab}^2 + acr\sigma_b^2$	$\sigma^2 + r\sigma_{ac(b)}^2 + ar\sigma_{c(b)}^2 + acr\theta_b^2$
AB	$(a - 1)(b - 1)$	$\sigma^2 + r\sigma_{ac(b)}^2 + cr\sigma_{ab}^2$	$\sigma^2 + r\sigma_{ac(b)}^2 + cr\theta_{ab}^2$
C/B	$b(c - 1)$	$\sigma^2 + r\sigma_{ac(b)}^2 + ar\sigma_{c(b)}^2$	$\sigma^2 + r\sigma_{ac(b)}^2 + ar\sigma_{c(b)}^2$
AC/B	$b(a - 1)(c - 1)$	$\sigma^2 + r\sigma_{ac(b)}^2$	$\sigma^2 + r\sigma_{ac(b)}^2$
Error	$abc(r - 1)$	σ^2	σ^2

Grados de libertad

Los grados de libertad de las fuentes de variación en el análisis de varianza siguen las convenciones usuales para los factores cruzados, A y B (sección 6.4); los grados de libertad de C anidado dentro de B siguen la convención para un factor anidado dentro de otro (sección 7.3); los grados de libertad para la interacción AC anidada dentro de B siguen la convención de grados de libertad anidados, sólo en este caso la interacción medida en cada nivel de B tiene $(a - 1)(c - 1)$ grados de libertad; entonces, al anidar en b niveles de B se tienen $b(a - 1)(c - 1)$ grados de libertad y los grados de libertad para el error son lo de cada celda, $r - 1$, agrupados sobre las abc celdas del experimento.

Análisis de factores cruzados y anidados

En la tabla 7.1 se muestra el análisis de varianza para las observaciones de la evaluación del espectrofotómetro del ejemplo 7.5. Se requieren cuatro cuadrados medios diferentes como denominadores del estadístico F_0 para probar las hipóte-

sis sobre los efectos o los componentes de la varianza en el modelo con un factor fijo y dos aleatorios (vea las tablas 7.13 y 7.14). El cuadrado medio del error es el denominador de F_0 para probar concentración \times corridas en un día, CR/D ; El cuadrado medio para CR/D es el denominador de F_0 necesario para probar las corridas en un día, R/D , y la interacción concentración \times día, CD ; el cuadrado medio CD es el denominador de F_0 requerido para probar las diferencias entre las medias de concentración, C , los cuadrados medios sintetizados para F_0 necesarias para probar la variación de un día, D , son:

$$CMN = CM(D) + CM(CR/D) = 42.4$$

para el numerador y:

$$CM = CM(CD) + CM(R/D) = 131.8$$

para el denominador. Es obvio que $F_0 = CMN/CM = 42.4/131.8 = 0.32$ no será significativo. Muchos programas usan en forma automática el cuadrado medio obtenido para el error experimental en todos los estadísticos F_0 , lo cual no siempre es el denominador correcto; esos programas requieren instrucciones especiales para calcular el estadístico F_0 correcto si tienen la capacidad.

Tabla 7.14 Análisis de varianza para las mediciones de glucosa del espectrofotómetro, a partir de un factorial con factores cruzados y anidados.

Fuente [†] de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	35	108,934.1			
C	2	108,263.6	54,131.8	1,227.48	.000
D	2	24.9	12.4	*	*
CD	4	176.4	44.1	1.47	.321
R/D	3	263.1	87.7	2.92	.122
CR/D	6	180.2	30.0	21.43	.000
Error	18	25.8	1.4		

* Prueba con cuadrado medio sintetizado

[†] C = concentración, D = día, R/D = corrida anidada en días

Como era de esperarse, las diferencias de concentración fueron significativas, $F_0 = 1227.48$ con $Pr > F = .000$. La interacción concentración \times día no fue significativa: $Pr > F = .321$, lo que indica un desempeño bastante congruente del instrumento de un día a otro respecto a la medición de concentraciones, pero concentración \times corrida en un día fue significativa, $F_0 = 21.43$ y $Pr > F = .000$, mientras las corridas en los días con $F_0 = 2.29$ y $Pr > F = .122$ no lo fueron. Debe revisarse la uniformidad entre las corridas del instrumento a través de las concentraciones, las discrepancias pueden deberse a la operación del instrumento o a diferencias en la preparación de las muestras para cada concentración de una corrida a otra.

7.5 ¿Cuántas réplicas?

Modelos aleatorios

El número de réplicas para detectar las contribuciones significativas deseadas de una componente de la varianza requiere un valor de la constante λ , donde el estadístico F_0 tiene una distribución central F_{v_1, v_2} multiplicada por λ^2 (vea la sección 5.8). En general, el valor de λ^2 puede evaluarse como sigue: sea F_0 el cociente de los cuadrados medios $F_0 = CMN/CMD$, donde CMN y CMD denotan las medias respectivas para el numerador y el denominador de F_0 ; la constante λ^2 es el cociente de los cuadrados medios esperados, es decir, $\lambda^2 = E(CMN)/E(CMD)$ (Los detalles técnicos se pueden encontrar en Graybill (1961)). Las gráficas de la tabla X del apéndice se usan como se describe en la sección 5.8.

Considerando el modelo de dos factores aleatorios de la sección 7.1. Una prueba de hipótesis $H_0: \sigma_a^2 = 0$ requiere $F_0 = CMA/CM(AB)$, de manera que:

$$\lambda^2 = \frac{E(CMA)}{E[CM(AB)]} = \frac{\sigma^2 + r\sigma_{ab}^2 + rb\sigma_a^2}{\sigma^2 + r\sigma_{ab}^2} = 1 + \frac{rb\sigma_a^2}{\sigma^2 + r\sigma_{ab}^2} \quad (7.15)$$

Modelos mixtos

La detección de efectos de factores fijos prescritos, digamos A , para el experimento del modelo mixto de dos factores requiere:

$$\Phi^2 = \frac{br \sum_{i=1}^t \alpha_i^2}{a(\sigma^2 + r\sigma_{ab}^2)} \quad (7.16)$$

para las gráficas de la tabla IX del apéndice para factores fijos. El valor de la constante λ para pruebas sobre σ_b^2 y σ_{ab}^2 se puede determinar como se mostró para el modelo aleatorio.

7.6 Reglas para los cuadrados medios esperados

Las reglas para los cuadrados medios esperados expuestas en esta sección se aplican a la mayor parte de los diseños balanceados con igual número de réplicas. El número de niveles para cualquier factor no varía en los diseños balanceados, incluyendo factoriales cruzados, factoriales anidados y mezclas de factores cruzados y anidados. Las reglas se adaptaron de las proporcionadas en varias publicaciones como Bennett y Franklin (1954) y Mason, Gunst y Hess (1989). Muchos programas de computadora tienen instrucciones para producir los cuadrados medios esperados para el análisis de varianza.

Reglas ilustradas con el modelo mixto no restringido

Las reglas se ejemplifican con un experimento con modelo mixto de dos factores, A fijo y B aleatorio, con r réplicas para cada combinación del tratamiento.

1. Se escribe el modelo lineal para el diseño:

$$y_{ijk} = \mu + \alpha_i + b_j + (ab)_{ij} + e_{k(ij)}$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

Observe el subíndice de la réplica k está anidado dentro de la ij -ésima combinación de tratamiento.

2. Se construye una tabla de dos factores con:
 - Un renglón para cada término del modelo, excluyendo a μ , etiquetada con el término del modelo
 - Una columna para cada subíndice usado en el modelo.
3. En la parte superior de cada subíndice de columna se escribe el número de niveles del factor correspondiente y escriba “R” si el factor es aleatorio y “F” si es fijo.
4. Se añade una columna cuyos elementos son la componente fija o aleatoria de la varianza para el efecto representado por ese renglón de la tabla.

<i>Fuente</i>	<i>F</i>	<i>R</i>	<i>R</i>	<i>Componente</i>
	<i>a</i>	<i>b</i>	<i>r</i>	
	<i>i</i>	<i>j</i>	<i>k</i>	
<i>A</i>	α_i			θ_a^2
<i>B</i>		b_j		σ_b^2
<i>AB</i>		$(ab)_{ij}$		σ_{ab}^2
Error		$e_{k(ij)}$		σ^2

5. Para cada renglón, si no aparecen los subíndices de la columna en el efecto del renglón, se introduce el número de niveles correspondientes al subíndice.

<i>Fuente</i>	<i>F</i>	<i>R</i>	<i>R</i>	<i>Componente</i>
	<i>a</i>	<i>b</i>	<i>r</i>	
	<i>i</i>	<i>j</i>	<i>k</i>	
<i>A</i>	α_i	<i>b</i>	<i>r</i>	θ_a^2
<i>B</i>	b_j	<i>a</i>	<i>r</i>	σ_b^2
<i>AB</i>		$(ab)_{ij}$	<i>r</i>	σ_{ab}^2
Error		$e_{k(ij)}$		σ^2

6. Si un subíndice está entre paréntesis en el efecto del renglón, se coloca un 1 en celdas bajo esos subíndices en el paréntesis.

Fuente	<i>F</i>	<i>R</i>	<i>R</i>	<i>Componente</i>
	<i>a</i>	<i>b</i>	<i>r</i>	
<i>i</i>	<i>j</i>	<i>k</i>		
<i>A</i>	α_i	<i>b</i>	<i>r</i>	θ_a^2
<i>B</i>	b_j	<i>a</i>	<i>r</i>	σ_b^2
<i>AB</i>	$(ab)_{ij}$		<i>r</i>	σ_{ab}^2
Error	$e_{k(y)}$	1	1	σ^2

7. a. Para cada renglón, si ahora el subíndice concuerda con el subíndice de la columna, se introduce un 0 si la columna representa un factor fijo F y existe una componente fina de la varianza para el efecto representado por el renglón.
- b. Se coloca un 1 en el resto de las celdas.

Fuente	<i>F</i>	<i>R</i>	<i>R</i>	<i>Componente</i>
	<i>a</i>	<i>b</i>	<i>r</i>	
<i>i</i>	<i>j</i>	<i>k</i>		
<i>A</i>	α_i	0	<i>b</i>	θ_a^2
<i>B</i>	b_j	<i>a</i>	1	σ_b^2
<i>AB</i>	$(ab)_{ij}$	1	1	σ_{ab}^2
Error	$e_{k(y)}$	1	1	σ^2

8. Para determinar el cuadrado medio esperado para una fuente específica de variación:
- Se incluye σ^2 con un coeficiente de 1 en todas los cuadrados medio esperados.
 - De las componentes de la varianza restantes se incluyen sólo aquellas cuyos términos correspondientes en el modelo incluyen subíndices del efecto en consideración. Para $E(CMB)$ el efecto b_j , se incluyen σ_{ab}^2 y σ_b^2 además de σ^2 .
 - Se recorren las columnas que contienen subíndices sin paréntesis para el efecto en consideración. Para α_i se recorre *i* y para $e_{k(y)}$ se recorre *k*.
 - El coeficiente para cada componente en la $E(CM)$ es el producto de las columnas restantes del renglón de ese efecto. Para $E(CMB)$ se recorre la columna con *j* de manera que sólo los valores en las columnas *i* y *k* quedan visibles. Para el renglón $(ab)_{ij}$, los valores visibles son 1 y *r* de manera que el coeficiente para σ_{ab}^2 es $1 \cdot r = r$. Para el renglón de b_j , los valores visibles son *a* y *r* por lo que el coeficiente para σ_b^2 es $a \cdot r$.

Fuente	<i>F</i>	<i>R</i>	<i>R</i>	<i>Componente</i>	$E(CM)$
	<i>a</i>	<i>b</i>	<i>r</i>		
<i>i</i>	<i>j</i>	<i>k</i>			
<i>A</i>	α_i	0	<i>b</i>	θ_a^2	$\sigma^2 + r\sigma_{ab}^2 + br\theta_a^2$
<i>B</i>	b_j	<i>a</i>	1	σ_b^2	$\sigma^2 + r\sigma_{ab}^2 + ar\sigma_b^2$
<i>AB</i>	$(ab)_{ij}$	1	1	σ_{ab}^2	$\sigma^2 + r\sigma_{ab}^2$
Error	$e_{k(y)}$	1	1	σ^2	σ^2

Ilustración: A continuación se presenta la tabla completa para determinar los cuadrados medios esperados de un factorial de modelo mixto con factores cruzados y anidados. Los factores A y B son fijos y cruzados y el factor C es aleatorio y anidado dentro de B a través de A . El modelo es

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + c_{k(j)} + (ac)_{ik(j)} + e_{l(ijk)}$$

Fuente	<i>F</i>	<i>F</i>	<i>R</i>	<i>R</i>	<i>Componente</i>	<i>E(CM)</i>
	<i>a</i>	<i>b</i>	<i>c</i>	<i>r</i>		
	<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>		
<i>A</i>	α_i	0	<i>b</i>	<i>c</i>	θ_a^2	$\sigma^2 + r\sigma_{ac(b)}^2 + bcr\theta_a^2$
<i>B</i>	β_j	<i>a</i>	0	<i>c</i>	θ_b^2	$\sigma^2 + r\sigma_{ac(b)}^2 + ar\sigma_{c(b)}^2 + acr\theta_b^2$
<i>AB</i>	$(\alpha\beta)_{ij}$	0	0	<i>r</i>	θ_{ab}^2	$\sigma^2 + r\sigma_{ac(b)}^2 + cr\theta_{ab}^2$
<i>C/B</i>	$c_{k(j)}$	<i>a</i>	1	1	$\sigma_{c(b)}^2$	$\sigma^2 + r\sigma_{ac(b)}^2 + ar\sigma_{c(b)}^2$
<i>(AC)/B</i>	$(ac)_{ik(j)}$	1	1	1	$\sigma_{ac(b)}^2$	$\sigma^2 + r\sigma_{ac(b)}^2$
Error	$e_{l(ijk)}$	1	1	1	σ^2	σ^2

Alteración de las reglas para los modelos mixtos restringidos

Si se usa el modelo mixto restringido, se requiere una alteración del método en el paso 7, es decir, el modelo en el que los efectos de interacción están correlacionados y su suma sobre los subíndices de los efectos fijos está restringida a cero. El paso 7 a) será “*Para cada renglón, si cualquier subíndice de renglón concuerda con el subíndice de la columna, se coloca un 0 si el subíndice de la columna representa un factor fijo F*”. El paso 7 b) permanece igual.

Otros métodos de estimación para los componentes de la varianza

En los capítulos 5 y 7 sólo se consideraron estimadores del análisis de varianza para la estimación de sus componentes. La estimación de los componentes de la varianza mediante el método de análisis de varianza es bastante directa con datos balanceados (esto es, todas las celdas de datos contienen el mismo número de observaciones), estimación que se dificulta cuando los datos no son balanceados. Henderson (1953) realizó el primer avance importante en la estimación de las componentes de la varianza; él presentó tres adaptaciones distintas del método para estimar las componentes de la varianza con datos no balanceados y modelos aleatorios o mixtos.

Desde entonces, Hartley y Rao (1967) desarrollaron otros métodos de estimación, incluyendo la estimación de la máxima posibilidad (MP), mientras que Thompson (1962) y Patterson y Thompson (1971) establecieron una modificación de la máxima posibilidad conocida como la máxima posibilidad restringida. El método MINQUE para encontrar estimadores cuadráticos no sesgados de la varianza mínima se puede considerar como presentado por una variedad de autores. Por lo

general, los métodos de estimación, se encuentran disponibles en muchos paquetes estadísticos completos; un tratamiento global de modelos y estimadores de las componentes de la varianza se encuentra en Searle *et al.* (1992).

EJERCICIOS PARA EL CAPÍTULO 7

1. Se midió el colesterol en muestras de suero de cinco pacientes seleccionados al azar entre un conjunto y se prepararon dos tubos réplica independientes de cada paciente para cuatro corridas en un espectrofotómetro. El objetivo del estudio era determinar si las mediciones relativas de colesterol de los pacientes eran uniformes de una corrida a otra en la clínica. Los datos son mg/dl de colesterol en las muestras réplica de cada paciente en cada corrida.

Corrida	Paciente				
	1	2	3	4	5
1	167.3	186.7	100.0	214.5	148.5
	166.7	184.2	107.9	215.3	148.5
2	179.6	193.8	111.6	228.9	158.6
	175.3	198.9	114.4	220.4	154.7
3	169.4	179.4	105.9	208.2	144.7
	165.9	177.6	104.1	207.1	145.9
4	177.7	190.4	113.4	221.0	156.1
	177.1	192.4	114.6	219.7	151.0

Fuente: Dr. J. Anderson, Beckman Instruments, Inc.

- a. Escriba un modelo lineal para el experimento, suponiendo que los pacientes y las corridas son efectos aleatorios, explique los términos y realice el análisis de varianza.
 - b. Muestre los cuadrados medios esperadas para el análisis de varianza.
 - c. Estime las componentes de la varianza para corridas, pacientes e interacción.
 - d. Establezca las hipótesis nula y alternativa para los efectos principales y de interacción, pruebe cada una de las hipótesis nulas e interprete sus resultados.
2. Un científico de animales realizó un experimento para estudiar el efecto de la calidad del agua en el alimento de terneras, las fuentes de agua se designaron como normal (N) y salina (S), el agua salina se formuló de manera que se aproxime a las concentraciones minerales en algunas fuentes de agua en el subsuelo utilizadas como agua para ganado y las cuatro combinaciones de agua usadas en dos períodos consecutivos de 56 días fueron N-N, N-S, S-N y S-S. La prueba de alimentación consistió en los cuatro tratamientos de agua con dos corrales réplica de animales para cada tratamiento, en un diseño totalmente aleatorizado y se realizó en dos ocasiones (dos veranos consecutivos). El diseño resultante es un arreglo factorial de cuatro tratamientos de agua y dos veranos y los tratamientos de agua se consideran efectos fijos y los veranos efectos aleatorios, de manera que es adecuado un modelo mixto para el estudio. Los datos de la ganancia de peso promedio diarias para los 16 corrales de terneras son:

Verano	Agua			
	N-N	N-S	S-N	S-S
1	2.65	2.46	2.56	2.43
	2.53	2.36	2.38	2.50
2	2.25	1.95	2.01	2.14
	2.20	2.25	1.98	2.37

Fuente: Dr. D. Ray, Department of Animal Sciences,
University of Arizona.

- Escriba un modelo lineal para el experimento, explique los términos y realice el análisis de varianza para los datos.
 - Prepare una tabla de medias de celdas y marginales y sus respectivos errores estándar.
 - Muestre las cuadrados medios esperados para el análisis de varianza.
 - Pruebe las hipótesis nulas para los efectos principales y la interacción e interprete sus resultados.
 - Los tratamientos de agua constituyen un arreglo factorial de 2×2 , el primer factor (A) es agua salina o normal en el primer periodo de 56 días, y el segundo factor (B) es agua salina o normal en el segundo periodo. Escriba un modelo lineal para el experimento con este arreglo, considerando los veranos como efectos aleatorios y los factores A y B como efectos fijos. Repita los incisos a) a d) de este ejercicio con el nuevo modelo.
3. Se prepararon tres fórmulas de aleación con cuatro colados distintos para cada fórmula y se probó el coeficiente de ruptura en dos barras de cada lote. Los datos son la fuerza de las barras individualmente y existen cuatro colados anidados dentro de cada aleación.

Aleación	Moldes			
	1	2	3	4
A	13.2	15.2	14.8	14.6
	15.5	15.0	14.2	15.1
B	17.1	16.5	16.1	17.4
	16.7	17.3	15.4	16.8
C	14.1	13.2	14.5	13.8
	14.8	13.9	14.7	13.5

- Escriba un modelo lineal para el experimento, suponga que las aleaciones son efectos fijos y que los colados dentro de las aleaciones y las barras dentro ellos son efectos aleatorios. Explique los términos y calcule el análisis de varianza.
 - Muestre los cuadrados medios esperadas para el análisis.
 - Pruebe la hipótesis nula para los efectos de la aleación e interprete sus resultados.
 - Calcule las medias estimadas y los intervalos de confianza estimados del 95% para las medias de cada aleación.
 - Estime las componentes de la varianza para los moldes y la barras.
4. Se realizó un estudio de ingeniería de tránsito para evaluar los efectos de tres tipos de semáforo sobre el retraso del tránsito en las intersecciones. Además, el estudio se diseñó para evaluar dos métodos para la medición

de los retrasos. Los tres tipos de semáforos fueron programados, semiactivos y activos; Los dos métodos, muestreo puntual y rastreo de trayectoria, estimaron el tiempo de detención por vehículo en una intersección.

Se usaron dos intersecciones para cada tipo de señal, se hicieron mediciones durante horas pico y no pico los tres factores cruzados en el estudio fueron los tipos de semáforo, el método y la hora del día (pico o no pico). Las intersecciones estaban anidadas en los tipos de semáforo y cruzadas con el método y la hora, ya que ambos métodos se usaron en la misma intersección durante los dos periodos. Los siguientes datos son los retrasos del tránsito medidos en segundos por vehículo:

Señal	Intersección	Muestra puntual		Trayectoria	
		Pico	No pico	Pico	No pico
Programado	1	61.7	57.4	53.1	36.5
	2	35.8	18.5	35.5	15.9
Semiactivo	3	20.0	24.6	17.0	21.0
	4	2.7	3.1	1.5	1.1
Activo	5	35.7	26.8	35.4	20.7
	6	24.3	25.9	27.5	23.3

Fuente: W. Reilly, C. Gardner, J. Kell (1976), A technique for measurement of delay at intersections. *Technical Report*, FHWA-RD-76-135, Federal Highway Administration, Office of R&D, Washington, D.C.

- Considere sólo un método para medir el retraso por detención (muestra puntual o trayectoria). Escriba un modelo lineal para el estudio y suponga que el tipo de semáforo y la hora del día son efectos fijos. Bosqueje la tabla del análisis de varianza, incluyendo la fuente de variación, los grados de libertad y los cuadrados medios esperados.
- Suponga que sospecha una interacción intersección × hora del día, ¿Puede probar una hipótesis nula sobre la interacción?
- Suponga que existe interacción entre la intersección y la hora del día. ¿Qué hipótesis se puede probar a partir del análisis de varianza?
- Calcule el análisis de varianza para los datos puntuales. Establezca sus suposiciones sobre el modelo y realice las pruebas de hipótesis posibles con el modelo establecido.
- Ahora suponga que desea incluir el factor del método de medición en el análisis, es decir, la muestra puntual contra la trayectoria como un factor de efecto fijo. Escriba un modelo lineal para el análisis y bosqueje la tabla del análisis de varianza, incluyendo las fuentes de variación, los grados de libertad y los cuadrados medios esperados.
- ¿Qué suposiciones son necesarias sobre la interacción entre la intersección y los otros factores para probar algunas hipótesis del análisis de varianza?
- Calcule el análisis de varianza para todo el conjunto de datos. Establezca las suposiciones sobre el modelo y realice las pruebas de hipótesis posibles sobre los efectos del modelo.
- Se realizó un experimento para comparar la exactitud de dos espectrómetros de masa en la medición de las proporciones de ^{14}N a ^{15}N , en él se tomaron dos muestras de suelo de cada una de tres parcelas tratadas con ^{15}N y se analizaron dos submuestras de cada muestra en cada una de dos máquinas. El diseño resultante tiene máquinas cruzadas con parcelas y muestras, pero las muestras están anidadas dentro de las parcelas. Los siguientes datos son las proporciones ^{14}N a ^{15}N (multiplicadas por 1000).

Muestra	1		2		3	
	Parcela	1	2	3	4	5
Máquina A	3.833	3.819	3.756	3.882	3.720	3.729
	3.866	3.853	3.757	3.871	3.720	3.768
Máquina B	3.932	3.884	3.832	3.917	3.776	3.833
	3.943	3.888	3.829	3.915	3.777	3.827

Fuente: D. Robinson (1987), Estimation and use of variance components. *The Statistician* 36, 3-14.

- Escriba un modelo lineal para el experimento, suponga que las máquinas tienen efectos fijos y las parcelas y muestras, efectos aleatorios; explique los términos y calcule el análisis de varianza para los datos.
 - Muestre los cuadrados medios esperados.
 - Pruebe la hipótesis nula de que no hay diferencia entre las medias para las dos máquinas.
6. Utilice las reglas dadas en la sección 7.6 para derivar los cuadrados medios esperados para los siguientes estudios o modelos:
- el estudio de colesterol en el ejercicio 7.1
 - la prueba de alimentación de ganado en el ejercicio 7.2
 - el experimento de colados de aleación en el ejercicio 7.3
 - el estudio de tránsito en el ejercicio 7.4
 - el estudio de suelos en el ejercicio 7.5
 - el diseño anidado de cuatro etapas con el modelo:

$$y_{ijkl} = \mu + a_i + b_{j(i)} + c_{k(j)} + d_{l(k)}$$

$$i = 1, 2, 3, 4 \quad j = 1, 2, 3 \quad k = 1, 2 \quad l = 1, 2, 3$$

- g. un modelo con factores anidados y cruzados escrito como:

$$y_{ijklm} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + c_{k(j)} + \delta_l + (\alpha\delta)_{il} + (\beta\delta)_{jl}$$

$$+ (\alpha\beta\delta)_{ijl} + (cd)_{kl(j)} + e_{m(ykl)}$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, c \quad l = 1, 2, \dots, d \quad m = 1, 2, \dots, r$$

donde α_i , β_j y δ_k , y sus interacciones son efectos fijos y $c_{k(j)}$, $(cd)_{kl(j)}$ y $e_{m(ykl)}$ son efectos aleatorios.

7. ¿En qué cambiaría su inferencia estadística si se usara el modelo con restricciones para el ejemplo 7.5?

8 Diseños de bloques completos

El diseño de experimentos para mejorar la precisión de los resultados obtenidos de los estudios de investigación es el tema de discusión de este capítulo y otros subsecuentes. La bloquización, explicado en el capítulo 1 como un método para reducir la variación del error experimental, agrupa las unidades experimentales en bloques para comparar tratamientos en un medio más homogéneo. Los diseños en este capítulo usan ya sea un criterio de agrupamiento en un diseño de bloques completos aleatorizados, o dos criterios de agrupamiento en arreglos de cuadros latinos. Aquí se estudian las características, la aleatorización, el análisis y la evaluación de estos diseños; las extensiones de los diseños incluyen diseños factoriales, unidades experimentales múltiples por tratamiento en cada bloque y submuestras; también se estudia el análisis cuando faltan algunas observaciones. El tema de la última sección de este capítulo es combinar los resultados a partir de varias repeticiones del mismo experimento en diferentes lugares o varias veces.

8.1 Uso de bloques para aumentar la precisión

El objetivo es tener comparaciones precisas entre los tratamientos de los estudios de investigación. La bloquización es un medio para reducir y controlar la varianza del error experimental con el fin de lograr una mayor precisión.

Los capítulos anteriores se concentraron en los diseños de tratamiento y sus métodos estadísticos asociados para realizar un análisis eficiente de las hipótesis de investigación, en todas las situaciones se utilizan diseños totalmente aleatorizados, pero fuera de la selección de la unidad experimental apropiada y las buenas técnicas de investigación, el diseño totalmente aleatorizado no controla la varianza del error experimental. Se supone que las unidades experimentales en esta clase de diseños son relativamente homogéneas respecto a las variables de respuesta medidas; pero algunas veces no existe un número suficiente de unidades homogéneas para un experimento completo con estos diseños.

Cualquier factor que afecta la variable de respuesta y que varía entre las unidades experimentales aumenta la varianza del error experimental y disminuye la precisión de los resultados del experimento. Factores tales como la edad o peso de los animales, lotes distintos de reactivos o material fabricado, género de sujetos humanos y separaciones físicas de parcelas son ejemplos de variables externas al tratamiento que pueden aumentar la variación entre las observaciones de las variables de respuesta.

El uso de bloques estratifica las unidades experimentales en grupos homogéneos, o unidades parecidas. Una buena elección de los criterios de bloquización disminuye la variación entre las unidades dentro de los bloques en comparación con las unidades de diferentes bloques, las categorías generales de buenos criterios son 1) proximidad (parcelas adyacentes), 2) características físicas (edad o peso), 3) tiempo y 4) administración de tareas en el experimento.

Un grupo de parcelas adyacentes forma un bloque en los experimentos de agro-nomía, los animales agrupados según su peso o etapa de lactancia forman bloques de unidades experimentales homogéneas, el ingeniero usa un solo lote de material fabricado para formar un bloque o grupo homogéneo de unidades experimentales para el tratamiento, los experimentos de laboratorio usan técnicos como factor de bloque para eliminar la variación entre ellos y cada técnico prepara una réplica del tratamiento como un bloque.

8.2 Los diseños de bloques completos aleatorizados usan un criterio de bloqueo

El diseño de bloques completos aleatorizados es el más sencillo de este tipo de diseños utilizados para controlar y reducir el error experimental, en él las unidades experimentales quedan estratificadas en bloques de unidades homogéneas, cada tratamiento se asigna al azar a un número igual (por lo general uno) de unidades experimentales en cada bloque y es posible hacer comparaciones más precisas entre los tratamientos dentro del conjunto homogéneo de unidades experimentales en un bloque. El uso de bloques fue muy provechoso en el siguiente estudio.

Ejemplo 8.1 El momento de fertilizar el trigo con nitrógeno

Las recomendaciones actuales para fertilizar el trigo con nitrógeno incluyen la aplicación de cantidades específicas en etapas establecidas del crecimiento de la planta. Las recomendaciones se desarrollaron a través de un análisis periódico del contenido de nitratos en los tejidos de la espiga, se pensó que el análisis del tejido era un medio efectivo para supervisar la cantidad de nitrógeno en la cosecha y tener una base para predecir el nitrógeno necesario para una producción óptima.

Objetivo de investigación: en ciertas situaciones, las pruebas de nitrato en los tejidos de la espiga predecían una mayor cantidad de nitrógeno, en consecuencia, el investigador quería evaluar el efecto de varios programas de fertilización sobre esas cantidades de nitrógeno y sobre la producción de trigo, para refinar las recomendaciones del procedimiento.

Diseño del tratamiento: el diseño del tratamiento incluyó seis programas diferentes de aplicación de nitrógeno que podían proporcionar el intervalo de condiciones necesarias para evaluar el proceso. Para la comparación se incluyó un tratamiento sin nitrógeno al igual que la recomendación normal vigente.

Diseño del experimento: el experimento se llevó a cabo en un campo irrigado, con un gradiente de agua en dirección del área de parcelas experimentales. Como las respuestas de las plantas dependían de la variabilidad en la humedad disponible, las parcelas se agruparon en bloques de seis de manera que cada bloque se encontraba en partes con el mismo gradiente de agua, de modo que cualesquiera diferencias en las respuestas de las plantas causada por el gradiente de agua podía asociarse con los bloques. El diseño de experimento resultante fue un diseño de bloques completo aleatorizado, con cuatro bloques de seis parcelas a las que se asignaron al azar los tratamientos de nitrógeno.

La distribución de las parcelas experimentales en el campo se muestra en el cuadro 8.1, donde se proporciona el contenido de nitrógeno observado ($\text{ppm} \times 10^{-2}$) en una muestra de espigas de trigo para cada parcela junto con los números de tratamiento, que aparecen en el recuadro pequeño para cada parcela.

Cuadro 8.1 Arreglo de las parcelas experimentales para el experimento de trigo en un diseño de bloques completos aleatorizado

							<i>Gradiente de irrigación</i>
							↓
<i>Bloque 1</i>	2	5	4	1	6	3	
	40.89	37.99	37.18	34.98	34.89	42.07	
<i>Bloque 2</i>	1	3	4	6	5	2	
	41.22	49.42	45.85	50.15	41.99	46.69	
<i>Bloque 3</i>	6	3	5	1	2	4	
	44.57	52.68	37.61	36.94	46.65	40.23	
<i>Bloque 4</i>	2	4	6	5	3	1	
	41.90	39.20	43.29	40.45	42.91	39.97	

Fuente: Dr. T. Doerge, Department of Soil and Water Science, University of Arizona.

Cómo aleatorizar el diseño

La asignación aleatoria de tratamientos a las unidades experimentales está restringida en un diseño de bloques completos aleatorizado de manera que cada trata-

miento debe presentarse el mismo número de veces (una o más) dentro de cada bloque. La aleatorización se ilustra con el experimento de trigo del ejemplo 8.1.

Una permutación aleatoria del orden en el que se colocan los tratamientos en las unidades de cada bloque proporciona una asignación aleatoria de los tratamientos a las unidades, existen $6 = 720$ permutaciones posibles en la asignación aleatoria de tratamientos a unidades. Se selecciona una permutación al azar para cada bloque ya que se requiere una aleatorización separada para cada uno de ellos.

Se asigna una etiqueta de tratamiento, como A, B, C, D, E, F, a los valores enteros respectivos 1, 2, 3, 4, 5, 6 y se obtiene una permutación aleatoria de los valores enteros con un programa de computadora o una tabla de permutaciones como la que se encuentra en la tabla XIII del apéndice (una de esas permutaciones se muestra en el cuadro 8.2). Dada la permutación 2, 5, 4, 1, 6, 3, se asigna el tratamiento B a la unidad 1, el E a la 2, el D a la 3, y así sucesivamente en el bloque 1.

Cuadro 8.2 Asignación de tratamientos a las unidades experimentales en un bloque completo

<i>Permutación</i>	2	5	4	1	6	3
<i>Tratamiento</i>	B	E	D	A	F	C
<i>Unidad experimental</i>	1	2	3	4	5	6

Se requieren permutaciones aleatorias separadas para cada uno de los tres bloques restantes. Dadas las permutaciones aleatorias adicionales, digamos (1, 3, 4, 6, 5, 2), (6, 3, 5, 1, 2, 4) y (2, 4, 6, 4, 3, 1), la asignación final de tratamientos a las unidades dentro de cada bloque se muestra en la figura 8.1.

Modelo estadístico y análisis para el diseño de bloques completos aleatorizados

El modelo lineal para un experimento en un diseño de bloques completos aleatorizado requiere un término que represente la variación identificable en las observaciones como consecuencia de los bloques. La respuesta de la unidad con el i -ésimo tratamiento en el j -ésimo bloque se escribe como:

$$y_{ij} = \mu + \tau_i + \rho_j + e_{ij} \quad (8.1)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde μ es la media general, τ_i es el efecto del tratamiento y e_{ij} es el error experimental. El efecto del bloque ρ_j representa la desviación promedio de las unidades en el bloque j a partir de la media general. Se supone que los efectos del tratamiento y del bloque son aditivos, *Aditividad* significa que no existe interacción entre tratamientos y bloques; también se supone que los errores experimentales son independientes, con medias cero y varianza común σ^2 . La suposición de independencia se justifica a través de la asignación aleatoria de los tratamientos a las unidades experimentales.

<i>Bloque 1</i>	<table border="1"> <tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr> <tr><td>B</td><td>E</td><td>D</td><td>A</td><td>F</td><td>C</td></tr> </table>	1	2	3	4	5	6	B	E	D	A	F	C	<i>Gradiente de irrigación</i>
1	2	3	4	5	6									
B	E	D	A	F	C									
<i>Bloque 2</i>	<table border="1"> <tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr> <tr><td>A</td><td>C</td><td>D</td><td>F</td><td>E</td><td>B</td></tr> </table>	1	2	3	4	5	6	A	C	D	F	E	B	↓
1	2	3	4	5	6									
A	C	D	F	E	B									
<i>Bloque 3</i>	<table border="1"> <tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr> <tr><td>F</td><td>C</td><td>E</td><td>A</td><td>B</td><td>D</td></tr> </table>	1	2	3	4	5	6	F	C	E	A	B	D	
1	2	3	4	5	6									
F	C	E	A	B	D									
<i>Bloque 4</i>	<table border="1"> <tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr> <tr><td>B</td><td>D</td><td>F</td><td>E</td><td>C</td><td>A</td></tr> </table>	1	2	3	4	5	6	B	D	F	E	C	A	
1	2	3	4	5	6									
B	D	F	E	C	A									

Figura 8.1 Asignación aleatoria de tratamientos en un diseño de bloques completos aleatorizado

Los bloques añaden una partición de suma de cuadrados al análisis de varianza

En la tabla 8.1 se muestran los datos básicos para un diseño de bloques completos aleatorizado; la desviación de cualquier observación de la gran media en la tabla 8.1, $y_{ij} - \bar{y}_{..}$, se puede escribir como la identidad algebraica:

$$y_{ij} - \bar{y}_{..} = (\bar{y}_i - \bar{y}_{..}) + (\bar{y}_j - \bar{y}_{..}) + (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}) \quad (8.2)$$

Tabla 8.1 Tabla de datos para un diseño de bloques completo aleatorizado

<i>Tratamiento</i>	<i>Bloque</i>				<i>Medias de tratamiento</i>
	1	2	...	<i>r</i>	
1	y_{11}	y_{12}	...	y_{1r}	\bar{y}_1
2	y_{21}	y_{22}	...	y_{2r}	\bar{y}_2
⋮	⋮	⋮		⋮	⋮
<i>t</i>	y_{t1}	y_{t2}	...	y_{tr}	\bar{y}_t
Medias de bloque	\bar{y}_1	\bar{y}_2	...	\bar{y}_r	\bar{y}

Los términos en el lado derecho de la ecuación (8.2) son:

- una desviación de tratamiento ($\bar{y}_i - \bar{y}_{..}$)
- una desviación de bloque ($\bar{y}_j - \bar{y}_{..}$)
- error experimental ($y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}$)

Por ejemplo, las medias para el experimento de trigo en el ejemplo 8.1 (mostradas en la tabla 8.3) se usan para ilustrar las desviaciones del tratamiento 1 en el bloque 2, y_{12} , como sigue:

- la desviación del tratamiento 1: $\bar{y}_1 - \bar{y}_{..} = 38.28 - 42.07 = -3.79$
- la desviación del bloque 2: $\bar{y}_2 - \bar{y}_{..} = 45.89 - 42.07 = 3.82$
- el error experimental para y_{12} :

$$y_{12} - \bar{y}_1 - \bar{y}_{..} + \bar{y}_{..} = 41.22 - 38.28 - 45.89 + 42.07 = -0.88$$

y la suma de las tres desviaciones, $-3.79 + 3.82 - 0.88 = -0.85$, es igual a la desviación total $y_{12} - \bar{y}_{..} = 41.22 - 42.07 = -0.85$, como debe ser.

Al observar con cuidado los dos últimos términos de la ecuación (8.2) forman una identidad algebraica para la desviación de las observaciones de la media del tratamiento:

$$y_{ij} - \bar{y}_i = (\bar{y}_j - \bar{y}_{..}) + (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}) \quad (8.3)$$

Se hace una partición de la desviación del error experimental del diseño totalmente aleatorizado, $y_{ij} - \bar{y}_i$, en dos componentes, el primer término se identifica con el criterio de bloques como $\bar{y}_j - \bar{y}_{..}$, el segundo término se identifica sólo como un residual o error experimental, $y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}$.

Elevando al cuadrado y sumando ambos lados de la ecuación (8.2) se tiene:

$$\begin{aligned} \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_{..})^2 &= r \sum_{i=1}^t (\bar{y}_i - \bar{y}_{..})^2 + t \sum_{j=1}^r (\bar{y}_j - \bar{y}_{..})^2 \\ &\quad + \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..})^2 \end{aligned} \quad (8.4)$$

o

$$SC \text{ total} = SC \text{ tratamiento} + SC \text{ bloques} + SC \text{ error}$$

La suma de cualesquiera productos cruzados en el lado derecho es cero. La tabla 8.2 resume la partición de la suma de cuadrados.

Tabla 8.2 Análisis de varianza para un experimento con un diseño de bloques completo aleatorizado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Cuadrados medios esperados
Total	$rt - 1$	$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2$		
Bloques	$r - 1$	$t \sum_j (\bar{y}_j - \bar{y}_{..})^2$	CMB	
Tratamientos	$t - 1$	$r \sum_i (\bar{y}_i - \bar{y}_{..})^2$	CMT	$\sigma^2 + r\theta_t^2$
Error	$(r - 1)(t - 1)$	$SC \text{ error}$	CME	σ^2

Las observaciones de nitrogenación en espigas junto con las medias de tratamientos y bloques para el estudio de fertilizante para trigo se muestran en la tabla 8.3. El análisis de varianza para los datos de nitrato de las espigas se muestra en la tabla 8.4.

Tabla 8.3 Nitrogenación de las espigas de trigo ($\text{ppm} \times 10^{-2}$) en seis tratamientos programados de nitrógeno para cada uno de cuatro bloques con un diseño de bloques completos aleatorizado

Programa de tiempos para nitrógeno	Bloque				Medias de tratamiento (\bar{y}_i)
	1	2	3	4	
Control	34.98	41.22	36.94	39.97	38.28
2	40.89	46.69	46.65	41.90	44.03
3	42.07	49.42	52.68	42.91	46.77
4	37.18	45.85	40.23	39.20	40.62
5	37.99	41.99	37.61	40.45	39.51
6	34.89	50.15	44.57	43.29	43.23
Medias de bloque (\bar{y}_j)	38.00	45.89	43.11	41.29	$\bar{y}_{..} = 42.07$

Tabla 8.4 Análisis de varianza para la nitrogenación de las espigas de trigo

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	23	506.33			
Bloque	3	197.00	65.67	9.12	
Nitrógeno	5	201.32	40.26	5.59	.004
Error	15	108.01	7.20		

Como consecuencia de la bloquización, se partitiona una suma de cuadrados de lo que hubiera sido la suma de cuadrados del error experimental con el diseño totalmente aleatorizado. Los diseños por bloques mejorarán mucho la precisión de las estimaciones de las medias de tratamiento si la reducción del *SC* error con el uso de bloques es sustancial, reducción que se puede eliminar si se reducen los grados de libertad, ya que $r - 1$ de los grados de libertad deben asignarse al *SC* bloques. Es necesaria una medida de la eficiencia relativa, que se muestra más adelante, para evaluar el beneficio completo de la bloquización.

Errores estándar para las medias de tratamiento

La estimación del error estándar para una media de tratamiento es:

$$s_{\bar{y}_i} = \sqrt{\frac{CME}{r}} = \sqrt{\frac{7.20}{4}} = 1.34 \quad (8.5)$$

y una estimación del intervalo de confianza del 95% de cualquier media de tratamiento en la tabla 8.3 es $\bar{y}_i \pm t_{025,15}(s_{\bar{y}_i})$, donde $t_{025,15} = 2.131$. El error estándar de una diferencia entre cualesquiera dos tratamientos se estima mediante:

$$s_{(\bar{y}_i - \bar{y}_j)} = \sqrt{\frac{2 CME}{r}} = \sqrt{\frac{2(7.20)}{4}} = 1.90 \quad (8.6)$$

Pruebas de hipótesis sobre medias de tratamientos

El estadístico F_0 para probar la hipótesis nula de que no hay diferencias entre tratamientos es:

$$F_0 = \frac{CMT}{CME} = \frac{40.26}{7.20} = 5.59 \quad (8.7)$$

que excede el valor crítico de $F_{.05,5,15} = 2.90$. El nivel de significancia observado es $Pr>F = .004$ (tabla 8.4), por lo que en esta etapa del desarrollo de la planta, hay diferencias significativas entre los tratamientos de nitrógeno y la nitrogenación de las espigas.

Existe poco interés en la inferencia formal sobre los efectos de bloques, por lo que no suele calcularse el estadístico F_0 para este propósito, aun cuando aparezca entre los resultados de un programa de computadora. La media en la que el uso de bloques aumente la eficiencia del diseño para utilizar los recursos se estudia más adelante en esta sección.

Interpretación con comparaciones múltiples

El programa 4 era la recomendación estándar de fertilizante para trigo; la nitración de la espiga de trigo medida en toda la temporada de crecimiento se usa para evaluar los requerimientos de nitrógeno para optimizar las cosechas. El investigador estaría interesado en las diferencias entre los programas de tratamiento y la recomendación actual durante cada etapa de crecimiento, para lo que es posible usar el método Dunnett (capítulo 3) normal incluso en el caso de control sin nitrógeno. El tratamiento de control permite evaluar el nitrógeno disponible sin fertilizante en estas parcelas. (Vea el uso de tratamientos de control, capítulo 1.)

Los intervalos de confianza simultáneos del 95% de Dunnett requieren el error estándar de la diferencia, $s(\bar{y}_i - \bar{y}_j) = 1.90$, y el estadístico de Dunnett, a partir de la tabla VI del apéndice, $d_{05,5,15} = 2.82$, para una comparación de dos colas.

El ICS del 95% para la diferencia entre la media de cualquier otro programa para la aplicación de nitrógeno y el programa 4 requiere que $D(5, .05) = d_{05,5,15}$ [$s(\bar{y}_i - \bar{y}_4)] = 2.82(1.90) = 5.36$ y se calcula con $\bar{y}_i - \bar{y}_4 \pm 5.36$. Los intervalos se encuentran en la tabla 8.5, junto con los resultados de la prueba de desigualdades de confianza, en la cual la diferencia absoluta se declara con una diferencia significativa de 0 si excede $D(5, .05) = 5.36$. El programa 3 es el único tratamiento que tiene un nivel medio de nitrógeno significativamente diferente del programa 4 actual, tiene un contenido de nitrato mayor, ya que la cota inferior del ICS es mayor que 0. El ICS para las demás comparaciones de tratamientos incluye 0 y tiene cotas superior e inferior muy diferentes de 0.

Tabla 8.5 Resultados del método de Dunnett para tratamientos contra control (ejemplo 8.1)

Programa	Media	$\bar{y}_i - \bar{y}_c$	ICS del 95%	
			(L, U)	—
4	$\bar{y}_c = 40.62$	—	—	—
1	38.28	-2.34	(-7.70, 3.02)	
2	44.03	3.41	(-1.95, 8.77)	
3	46.77	6.15*	(0.79, 11.51)	
5	39.51	-1.11	(-6.47, 4.25)	
6	43.23	2.61	(-2.75, 7.97)	

* $|\bar{y}_i - \bar{y}_c|$ excede a $D(5, .05) = 5.36$ y es significativamente diferente del programa 4.

Si los tratamientos interactúan con los bloques

Suponer que no haya interacción tratamiento \times bloques implica que los efectos de tratamiento y bloque son aditivos. Se supone que las diferencias entre los tratamientos son más o menos constantes de un bloque a otro como consecuencia de los efectos aditivos de tratamientos y bloques, aunque al usar bloques puedan obtenerse mayores diferencias entre las respuestas de unidades de bloques distintos. Se puede realizar la prueba de no aditividad, explicada en el capítulo 6 (Tukey, 1949b), para detectar la no aditividad de la forma multiplicativa $\lambda\tau_i\rho_j$. Si hay suficientes diferencias entre las condiciones ambientales de uno o más bloques, pueden afectar el desempeño relativo de los tratamientos; por ejemplo, si la base nutriente residual en el suelo es muy diferente de un bloque a otro en una prueba de fertilizante de cosechas, quizás no haya respuesta o sea muy pequeña a los fertilizantes en algunos bloques, mientras que las respuestas pueden ser considerables sustanciales en otros.

Unidades experimentales múltiples por tratamiento en cada bloque

La no aditividad más general se representa por el término de interacción general $(\tau\rho)_{ij}$. Para probar la existencia de la interacción, el experimento debe tener más de una unidad experimental para cada tratamiento dentro de cada bloque. El modelo lineal para un experimento con u unidades experimentales con cada tratamiento en cada uno de los r bloques es:

$$y_{ijk} = \mu + \tau_i + \rho_j + (\tau\rho)_{ij} + e_{ijk} \quad (8.8)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad k = 1, 2, \dots, u$$

donde e_{ijk} son los errores experimentales aleatorios independientes con media 0 y varianza σ^2 , esta última es la variabilidad entre las unidades experimentales dentro de un bloque que recibió el mismo tratamiento. Los cálculos de las particiones de sumas de cuadrados y la prueba de interacción son los mismos que para un arreglo factorial con dos factores presentado en el capítulo 6.

Submuestras de unidades experimentales

Existen situaciones en las que se requieren dos o más muestras de las unidades experimentales al recolectar los datos (en el capítulo 5 se estudiaron situaciones que requieren submuestras de las unidades). El modelo lineal para un experimento en un diseño de bloques completo aleatorizado con n submuestras de las unidades experimentales es:

$$\begin{aligned} y_{ijk} &= \mu + \tau_i + \rho_j + e_{ij} + d_{ijk} & (8.9) \\ i &= 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad k = 1, 2, \dots, n \end{aligned}$$

donde las d_{ijk} son los efectos aleatorios para las submuestras con media 0 y varianza σ^2_d , y los demás términos son los descritos en la ecuación (8.1) para el diseño de bloques completos aleatorizado. Los ajustes al análisis son los presentados en el capítulo 5.

Gates (1995) proporciona detalles sobre la estimación del error experimental en el diseño de bloques con configuraciones diferentes de las unidades experimentales y de muestreo.

Análisis de residuales para evaluar las suposiciones

Las suposiciones relativas a los errores experimentales del modelo lineal para el diseño de bloques completo aleatorizado se puede evaluar con un análisis de residuales (estudiado en el capítulo 4). Se calculan los residuales a partir de la componente de la desviación del error experimental mostrada en la ecuación (8.2) como $\widehat{e}_{ij} = y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}$. Por ejemplo, el residual para el tratamiento de control en el bloque 1 del ejemplo 8.1 es:

$$\begin{aligned} \widehat{e}_{11} &= y_{11} - \bar{y}_{1\cdot} - \bar{y}_{\cdot 1} + \bar{y}_{\cdot\cdot} \\ &= 34.98 - 38.28 - 38.00 + 42.07 = 0.77 \end{aligned}$$

La gráfica de residuales contra los valores esperados y la de probabilidad normal de los residuales se muestran en las figuras 8.2a y 8.2b. Parece que las suposiciones de varianza homogénea (figura 8.2a) y de distribución normal (figura 8.2b) se cumplen para estos datos.

¿Aumentó la precisión con el uso de bloques?

La expectativa de mayor precisión en las estimaciones de las medias de tratamiento es la motivación para usar el diseño de bloques completos aleatorizado, pues aunque planear y realizar un experimento con este diseño exige un mayor esfuerzo relativo en comparación con el diseño totalmente aleatorizado, la medida de eficiencia relativa (presentada en el capítulo 1) evalúa los beneficios de los bloques para un experimento específico.

Comparando la eficiencia de un diseño de bloques completo aleatorizado con la del diseño totalmente aleatorizado, la estimación de σ^2 , digamos s^2_{rcb} , es el cuadrado medio

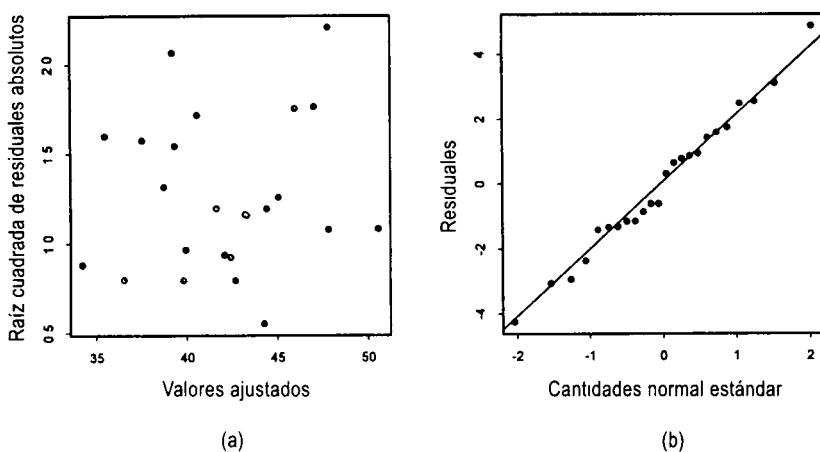


Figura 8.2 Gráficas residuales del análisis de varianza para los datos sobre nitrato de nitrógeno en el ejemplo 8.1: a) raíz cuadrada de los residuales absolutos vs. valores estimados y b) gráfica de probabilidad normal de los residuales

para el error experimental a partir del análisis de varianza para este experimento con un diseño de bloques completos aleatorizado. Se requiere una estimación de σ^2 a partir de un diseño totalmente aleatorizado, digamos s_{cr}^2 , para medir la eficiencia relativa. El cálculo de s_{cr}^2 a partir de la información en el análisis de varianza de bloques completos aleatorizados (Cochran y Cox, 1957; Kempthorne, 1952) es:

$$s_{cr}^2 = \frac{SC \text{ bloques} + r(t-1)CME}{rt-1} \quad (8.10)$$

La σ_{cr}^2 estimada para el estudio de fertilización de trigo es:

$$s_{cr}^2 = \frac{197.0 + 20(7.2)}{23} = 14.8$$

y la estimación de σ_{rcb}^2 es $s_{rcb}^2 = CME = 7.2$.

La estimación de la eficiencia relativa ER, sin corrección por grados de libertad para las estimaciones de σ^2 , es:

$$ER = \frac{s_{cr}^2}{s_{rcb}^2} = \frac{14.8}{7.2} = 2.06 \quad (8.11)$$

La corrección para la estimación de σ^2 mediante s^2 es:

$$\frac{(f_{rcb} + 1)(f_{cr} + 3)}{(f_{rcb} + 3)(f_{cr} + 1)} = \frac{16 \cdot 21}{18 \cdot 19} = 0.98$$

donde $f_{rcb} = 15$ y $f_{cr} = 18$ son los grados de libertad respectivos para el bloqueo completo aleatorizado y el diseño totalmente aleatorizado. La corrección reduce la ER a $0.98(2.06) = 2.02$, misma que tiene poco efecto con grados de libertad moderados para las estimaciones de la varianza del error experimental.

Se estima que el diseño de bloques completo aleatorizado para el experimento de trigo tiene un poco más del doble de la eficacia que un diseño totalmente aleatorizado. En el diseño totalmente aleatorizado se requieren ocho réplicas para tener varianzas equivalentes de las medias de tratamiento bajo las mismas condiciones experimentales en ese campo. En este caso el uso de bloques en el gradiente de irrigación fue efectivo como medida de control y para reducir la estimación de la varianza del error experimental. Lo más probable es que los futuros experimentos de esta naturaleza se beneficien con la práctica del bloqueo.

Una verificación rápida de la efectividad del uso de bloques

Lentner, Arnold y Hinkelmann (1989) analizan una relación entre la eficiencia relativa y la razón $H = CMB/CME$; señalan que, aunque la razón H es equivalente a un estadístico F_0 , no existe una prueba válida para los efectos de bloques. De las ecuaciones (8.10) y (8.11), la medida de la eficiencia relativa se puede expresar como:

$$ER = \frac{s_{cr}^2}{s_{rcb}^2} = \frac{(r - 1)CMB + r(t - 1)CME}{(rt - 1)CME} \quad (8.12)$$

Reescribiéndola, la expresión se convierte en:

$$ER = k + (1 - k)H \quad (8.13)$$

donde $H = CMB/CME$ y $k = r(t - 1)/(rt - 1)$. Las siguientes relaciones para ER y H se pueden determinar a partir de la ecuación (8.13):

$ER < 1$	si y sólo si	$H < 1$
$ER = 1$	si y sólo si	$H = 1$
$ER > 1$	si y sólo si	$H > 1$

H se puede usar para evaluar la efectividad del uso de bloques, aunque no es un estadístico F válido para probar los efectos de bloques. Por ejemplo, si $ER > 1$, entonces $H > 1$; el uso de bloques ha sido efectivo en cuanto a reducir el error experimental, por lo que con un diseño de bloques completo se requieren menos réplicas que con el diseño totalmente aleatorizado. El valor de H no proporciona información completa sobre la eficiencia relativa, solamente contiene lo relacionado con una eficiencia mayor, menor o igual; H es una verificación rápida de la efectividad del uso de bloques.

Bloques aleatorios

Los bloques de unidades con frecuencia constituyen una muestra aleatoria de los bloques disponibles para el investigador. Los sitios usados como bloques en estudios ecológicos, forestales o de vida silvestre pueden ser muestras aleatorias de muchos sitios disponibles para el estudio, es posible establecer parcelas en cada uno de los sitios para los tratamientos. Los lotes de material fabricado (como tela, asfalto o producto químico) usados como bloques de tratamientos experimentales son lotes aleatorios, el lote se divide en unidades experimentales más pequeñas a las que se administra el tratamiento. Las escuelas usadas como criterios de bloque en estudios de educación son representantes aleatorios de las escuelas existentes en el área y los salones de clase dentro de ellas sirven de unidades experimentales para los tratamientos.

. La base de inferencia para los tratamientos en un estudio con bloques aleatorios se extiende a la población de bloques a partir de la cual se obtuvo la muestra aleatoria de bloques. Como consecuencia de los bloques aleatorios, los errores estándar para las medias de tratamientos serán diferentes de los de un experimento con bloques fijos. El modelo lineal con efectos de bloques aleatorios es:

$$y_{ij} = \mu + \tau_i + b_j + e_{ij} \quad (8.14)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde μ es la media general, τ_i es el efecto fijo del tratamiento, b_j es el efecto aleatorio de bloque con media 0 y varianza σ_b^2 , y e_{ij} es el error experimental con media 0 y varianza σ^2 . Con el modelo mixto, una observación tiene un valor esperado $E(y_{ij}) = \mu + \tau_i$ y varianza $\sigma^2 + \sigma_b^2$; también existe una covarianza de σ_b^2 entre cualesquiera dos observaciones en el mismo bloque con bloques aleatorios.

La varianza de una media de tratamiento con bloques aleatorios en un diseño totalmente aleatorizado es:

$$\sigma_{\bar{y}_i}^2 = \frac{1}{r} (\sigma^2 + \sigma_b^2) \quad (8.15)$$

La varianza de una media de tratamiento con bloques aleatorios incluye la componente de la varianza para los bloques, σ_b^2 , y será mayor que la varianza con efectos de bloques fijos. Los efectos de bloques aleatorios no afectan la varianza de la diferencia entre dos medias de tratamiento que será la mostrada para los efectos fijos en la ecuación (8.6).

8.3 Los diseños de cuadrado latino usan dos criterios de bloqueo

Pueden ser necesarios dos factores de bloqueo

Reconocer un factor, diferente de los tratamientos planeados, que influyera en la variable de respuesta era importante en el experimento con trigo del ejemplo 8.1; los bloques de parcelas experimentales según el gradiente de irrigación duplicó la eficiencia del experimento.

En algunas situaciones experimentales pueden influir dos factores, distintos de los tratamientos, en la variable de respuesta y es posible lograr aún más precisión si se bloquizan las unidades de acuerdo con estos factores; si un segundo factor es candidato como criterio de bloque, se puede usar el arreglo de cuadrado latino para diseñar el experimento.

El arreglo del cuadrado latino se deriva de las letras del alfabeto latino A, B, C, ... dispuestas en un arreglo cuadrado de manera que cada letra aparece una vez en cada columna y una vez en cada renglón del cuadrado, en sus aplicaciones a experimentos, los renglones y las columnas del arreglo se identifican con los dos criterios de bloque y las letras latinas con los tratamientos. Una de estas aplicaciones ocurrió en el siguiente ejemplo:

Ejemplo 8.2 Relación entre la cosecha de trigo y la tasa de siembra

Las prácticas del cultivo de trigo como la cantidad de semillas plantadas, el espaciamiento de filas y la fecha de siembra tiene efectos directos en la cosecha. Las prácticas de cultivo para optimizar la producción se establecen con experimentos en cultivos nuevos.

Objetivo de investigación: En cierto caso, un investigador deseaba determinar la tasa de siembra óptima para una nueva especie de trigo, con alto contenido de extracto de semolina, importante para la elaboración de pastas.

Diseño del tratamiento: Se usaron cinco tasas de siembra (30, 80, 130, 180 y 230 lb/acre) para el diseño del tratamiento. Con base en otros cultivos comunes al área, estas tasas de siembra deben incluir la tasa de producción óptima.

Diseño del experimento: El experimento se llevó a cabo en un campo irrigado con un gradiente de agua en una dirección del área experimental. Además, se sabía que los campos experimentales de la granja tienen diferencias en el suelo creadas por la pendiente requerida para la irrigación. En general, estas diferencias de suelo eran perpendiculares a los canales de irrigación.

El investigador hizo un bloque de las parcelas con un arreglo de renglones y columnas para controlar los gradientes de suelo y agua en dos direcciones sobre el campo experimental. Los tratamientos de la tasa de siembra se asignaron de manera aleatoria a las parcelas en un arreglo de cuadrado latino de 5×5 .

En el cuadro 8.3 se muestra la distribución de las parcelas experimentales en el diseño de cuadrado latino después de la aleatorización, la cosecha de grano de cada parcela se muestra en quintales (100 lb) por acre junto con la letra que indica el tipo de tratamiento.

Los bloques por fila del campo coinciden con el gradiente de irrigación, y los bloques por renglón-columna corresponden al gradiente perpendicular del suelo en relación con el gradiente de irrigación. Los tratamientos tienen un arreglo de cuadrado latino donde cada uno aparece una vez en cada bloque de renglón y una vez en cada bloque de columna.

Cuadro 8.3 Arreglo de las parcelas experimentales para el experimento con trigo en un diseño de cuadrado latino

Fila en el campo	Columna 1	Columna 2	Columna 3	Columna 4	Columna 5	Gradiente irrigación
1	59.45(E)	47.28(A)	54.44(C)	50.14(B)	59.45(D)	↓
2	55.16(C)	60.89(D)	56.59(B)	60.17(E)	48.71(A)	
3	44.41(B)	53.72(C)	55.87(D)	47.99(A)	59.45(E)	
4	42.26(A)	50.14(B)	55.87(E)	58.74(D)	55.87(C)	
5	60.89(D)	59.45(E)	49.43(A)	59.45(C)	57.31(B)	
<i>Pendiente del suelo</i>		→				

Fuente: Dr. M. Ottman, Department od Plant Science, University of Arizona.

Otras aplicaciones del cuadrado latino

Un experimento para probar tratamientos en llantas de automóvil es un ejemplo clásico usado para ilustrar el diseño de cuadrado latino, este experimento prueba cuatro llantas (A, B, C, D) en cuatro automóviles y cada llanta aparece en una de las cuatro posiciones de cada automóvil. Los criterios de bloque por renglón y columna son los automóviles y las posiciones de las llantas, respectivamente, para el diseño mostrado en el cuadro 8.4.

Cuadro 8.4 Arreglo de cuadrado latino para tratamientos de llantas de automóvil

Auto	Posición de la llanta			
	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

Cada tratamiento (A, B, C o D) aparece una vez en cada renglón (auto) y una vez en cada columna (posición de la llanta); el razonamiento para los criterios de bloque es que el desgaste de las llantas puede diferir entre los automóviles y las posiciones en las que se montan.

El arreglo del bloques no tiene que ser rectangular para se reduzca la varianza del error. Un arreglo lineal de los tratamientos en un experimento de invernadero o

un arreglo de tratamientos procesado en el tiempo se puede ordenar de acuerdo con una asignación de cuadrado latino. El cuadro 8.5 ilustra un arreglo lineal de cuatro tratamientos con este tipo de asignación.

Cuadro 8.5 Asignación de cuadrado latino con un arreglo lineal de las unidades experimentales

Renglón	1	2	3	4
Columna	A B C D 1 2 3 4	B C D A 1 2 3 4	C D A B 1 2 3 4	D A B C 1 2 3 4
Gradiente				
en tiempo o espacio				

Uso del cuadrado latino estándar para generar los diseños

Todos los cuadrados latinos de un tamaño específico se pueden generar a partir de los *cuadrados estándar*. Un cuadrado estándar tiene los símbolos de tratamiento (A, B, C, ...) en orden alfabético en el primer renglón y en la primera columna del arreglo, cada símbolo de tratamiento ocurre una vez en cada columna y una vez en cada renglón del arreglo. Sólo existe un cuadrado estándar para $t = 2$ o 3 tratamientos, existen 4 cuadrados estándar para $t = 4$ tratamientos y 56 para $t = 5$; el número de cuadrados estándar aumenta en forma drástica con el número de tratamientos, ya que existen 9408 cuadrados estándar con 6 tratamientos.

En el apéndice 8A se muestran los cuadrados estándar para $t = 2, 3$ y 4 tratamientos y ejemplos de cuadrados hasta con 10 tratamientos. Fisher y Yates (1963) publicaron el conjunto completo de cuadrados para $t = 4$ a 6 tratamientos junto con muestras de hasta $t = 12$ tratamientos.

El cuadrado estándar de cualquier tamaño se puede generar escribiendo el primer renglón de letras en orden alfabético, el segundo renglón se obtiene a partir del primero si se corre una letra a la izquierda y se mueve la letra A al extremo derecho, el tercer renglón se obtiene al correr el segundo una letra a la izquierda y colocando la B en la posición de la extrema derecha, este proceso continúa para los renglones restantes. Un cuadrado latino de 6×6 construido de esta manera es

A	B	C	D	E	F
B	C	D	E	F	A
C	D	E	F	A	B
D	E	F	A	B	C
E	F	A	B	C	D
F	A	B	C	D	E

El cuadrado latino es un diseño restrictivo porque requiere que el número de tratamientos, de renglones y de columnas sea igual, requisito difícil de satisfacer en algunas situaciones experimentales que requieren dos criterios de bloqueo. Los cuadrados latinos con $t = 4$ o menos tratamientos tienen pocos grados de libertad para estimar la varianza del error experimental; entonces, su valor puede limitarse con pequeños experimentos a menos que sea posible tener dos o más repeticiones del diseño. Con números de tratamiento mayores que $t = 8$ a 10, el número requerido de unidades experimentales puede ser prohibitivo según las circunstancias de los experimentos contemplados. El tamaño adecuado para muchos experimentos con el arreglo del cuadrado latino es $t = 5$ a 7 tratamientos.

Preece (1983) proporciona un antecedente histórico del cuadrado latino y un análisis general de las variaciones sobre los diseños del cuadrado latino sugeridos para trabajo experimental con diseño de bloques renglón-columna.

Cómo aleatorizar el diseño

Si se dispone de todos los cuadrados latinos de tamaño $t \times t$, la aleatorización se logra con los siguientes pasos:

- Paso 1.* Seleccionar al azar uno de los cuadrados estándar.
- Paso 2.* Ordenar al azar todos menos el primer renglón.
- Paso 3.* Ordenar al azar todas las columnas.
- Paso 4.* Asignar al azar los tratamientos a las letras.

Se pueden generar todas las aleatorizaciones posibles sin incluir el primer renglón en el paso 2 si se selecciona de manera aleatoria un cuadrado estándar. Si no se dispone de uno para elegir, entonces se recomienda que en el paso 2 se incluyan todos los renglones en la aleatorización. No todos los cuadrados latinos se generan de esta manera pero el número de posibilidades aumenta en forma considerable. Suponiendo que el cuadrado estándar seleccionado en el paso 3 para el experimento del cuadrado latino de 4×4 con las llantas de automóvil es

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Paso 2. Se obtiene una permutación aleatoria de número para ordenar los últimos tres renglones:

Permutación	Renglón original
3	2
1	3
2	4

La colocación de los renglones para el cuadrado latino con el renglón 1 en su posición original es:

<i>Renglón original</i>				
1	A	B	C	D
3	C	D	A	B
4	D	A	B	C
2	B	C	D	A

Paso 3. Se obtiene una permutación aleatoria de números para ordenar las cuatro columnas del paso 2:

<i>Permutación</i>	<i>Columna original</i>
1	1
4	2
3	3
2	4

La colocación de las columnas para el cuadrado estándar es:

<i>Columna original</i>			
<i>I</i>	4	3	2
A	D	C	B
C	B	A	D
D	C	B	A
B	A	D	C

Paso 4. Se obtiene una permutación aleatoria para asignar tratamientos a las letras. Esta asignación no es necesaria si el cuadrado latino se seleccionó al azar entre todos los cuadrados posibles. El método de asignación se muestra como ilustración, supongamos que las etiquetas del tratamiento son W, X, Y y Z:

<i>Permutación</i>	<i>Tratamiento</i>
4 = D	W
2 = B	X
3 = C	Y
1 = A	Z

Las etiquetas de tratamiento W, X, Y, Z sustituyen a las letras en el cuadrado latino en el orden D, B, C, A en el arreglo aleatorizado. La colocación final de los tratamientos de llantas sobre las posiciones de automóviles y llantas es:

Auto	Posición de llantas			
	1	2	3	4
1	Z	W	Y	X
2	Y	X	Z	W
3	W	Y	X	Z
4	X	Z	W	Y

Modelo estadístico y análisis para diseños de cuadrado latino

El modelo estadístico lineal para un experimento con t tratamientos en un diseño de cuadrado latino de $t \times t$ es:

$$y_{ij} = \mu + \rho_i + \gamma_j + \tau_k + e_{ij} \quad (8.16)$$

$$i, j, k = 1, 2, \dots, t$$

donde y_{ij} es la observación de la unidad experimental en el i -ésimo renglón y la j -ésima columna del diseño. Los efectos respectivos de renglón y columna son ρ_i y γ_j ; τ_k es el efecto del k -ésimo tratamiento, y las e_{ij} son errores experimentales independientes aleatorios con media 0 y varianza σ^2 . Se supone que no hay interacción entre los tratamientos y las columnas y renglones.

La notación para los totales y las medias de las observaciones para renglones y columnas siguen la convención usual con $y_{..} = \sum_j y_{ij}$ para el total de un renglón y $y_{.j} = \sum_i y_{ij}$ para el total de una columna. El total del tratamiento estará representado como y_k , que implica una suma de observaciones sobre las t unidades experimentales que reciben el tratamiento k . De la misma manera, \bar{y}_k representará la media de las observaciones en el k -ésimo tratamiento.

Dos particiones de sumas de cuadrados para bloques

Las particiones de sumas de cuadrados se pueden derivar de la identidad algebraica:

$$y_{ij} - \bar{y}_{..} = (\bar{y}_{i..} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (\bar{y}_k - \bar{y}_{..}) \\ + (y_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_k - 2\bar{y}_{..}) \quad (8.17)$$

La desviación de una observación de la gran media $y_{ij} - \bar{y}_{..}$ se expresa como una suma aditiva de:

- una desviación de renglón ($\bar{y}_{i..} - \bar{y}_{..}$)
- una desviación de columna ($\bar{y}_{.j} - \bar{y}_{..}$)
- una desviación de tratamiento ($\bar{y}_k - \bar{y}_{..}$)
- error experimental ($y_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_k - 2\bar{y}_{..}$)

Por ejemplo, las medias para el experimento de trigo del ejemplo 8.2 (mostradas en la tabla 8.7) se usan para ilustrar las desviaciones para la observación en el renglón 1 y la columna 1 con el tratamiento 5 = E, como:

- la desviación del renglón 1: $\bar{y}_1 - \bar{y} = 54.15 - 54.53 = -0.38$
- la desviación de la columna 1: $\bar{y}_1 - \bar{y} = 52.43 - 54.53 = -2.10$
- la desviación del tratamiento E: $\bar{y}_5 - \bar{y} = 58.88 - 54.53 = 4.35$
- el error experimental para y_{11} :

$$y_{11} - \bar{y}_1 - \bar{y}_1 - \bar{y}_5 + 2\bar{y} = 59.45 - 54.15 - 52.43 - 58.88 + 2(54.53) = 3.05$$

La suma de las cuatro desviaciones es $-0.38 - 2.10 + 4.35 + 3.05 = 4.92$, que es igual a la desviación $y_{11} - \bar{y} = 59.45 - 54.53 = 4.92$.

Al elevar al cuadrado ambos lados de la ecuación (8.17) y sumar los términos se llega a una partición aditiva de:

$$SC \text{ total} = SC \text{ renglones} + SC \text{ columnas} + SC \text{ tratamiento} + SC \text{ error}$$

La tabla 8.6 resume las sumas de cuadrados en un análisis de varianza con los cuadrados medios esperados para los efectos de tratamientos fijos.

Tabla 8.6 Análisis de varianza para experimentos con un diseño de cuadrado latino

Fuentes de variación	Grados de libertad	Sumas de cuadrados	Cuadrados medios	Cuadrados medios esperados
Total	$t^2 - 1$	$\sum_i \sum_j (y_{ij} - \bar{y})^2$		
Renglones	$t - 1$	$t \sum_i (\bar{y}_i - \bar{y})^2$	<i>CMR</i>	
Columnas	$t - 1$	$t \sum_j (\bar{y}_j - \bar{y})^2$	<i>CMC</i>	
Tratamientos	$t - 1$	$t \sum_k (\bar{y}_k - \bar{y})^2$	<i>CMT</i>	$\sigma^2 + t\theta_t^2$
Error	$(t - 1)(t - 2)$	<i>SC error</i>	<i>CME</i>	σ^2

La suma de cuadrados para el error experimental se ha reducido respecto a la del diseño de bloque totalmente aleatorizado en una cantidad igual a *SC* renglones o *SC* columnas, con un costo de $t - 1$ grados de libertad. Los cuadrados medios para el error experimental como estimación de σ^2 tiene muy pocos grado de libertad con un pequeño número de tratamientos. Se pierde mucho poder en las pruebas de hipótesis para comparaciones entre tratamientos a menos que la reducción en la suma de cuadrados del error, debida al uso de bloques según los criterios de renglón y columna sea sustancial.

La efectividad del uso de bloques por otros criterios evaluados con la medida de eficiencia relativa se demuestra en el análisis de cosecha de grano a partir del experimento de tasa de siembra del ejemplo 8.2.

En la tabla 8.7 se muestran las observaciones junto con las medias de renglón, columna y tratamiento en un arreglo de cuadrado latino. Los datos se refieren a la cosecha de grano para cada gráfica en cientos de libras por acre. El análisis de la varianza se muestra en la tabla 8.8.

Tabla 8.7 Cosecha de grano de una variedad de trigo para cinco tasas de siembra en un diseño de cuadrado latino [etiqueta de tratamiento (**A**, **B**, **C**, **D** o **E**) en los paréntesis enseguida del valor de la cosecha]

Medias de renglón	Columna de campo					Renglón de campo (\bar{y}_i)
	1	2	3	4	5	
1	59.45(E)	47.28(A)	54.44(C)	50.14(B)	59.45(D)	54.15
2	55.16(C)	60.89(D)	56.59(B)	60.17(E)	48.71(A)	56.30
3	44.41(B)	53.72(C)	55.87(D)	47.99(A)	59.45(E)	52.29
4	42.26(A)	50.14(B)	55.87(E)	58.74(D)	55.87(C)	52.58
5	60.89(D)	59.45(E)	49.43(A)	59.45(C)	57.31(B)	57.31
Medias de columna (\bar{y}_j)	52.43	54.30	54.44	55.30	56.16	$\bar{y} =$ 54.53
Tratamiento		A	B	C	D	E
Tasa de siembra		30	80	130	180	230
Media (\bar{y}_k)		47.13	51.72	55.73	59.17	58.88

Tabla 8.8 Análisis de varianza para la cosecha de grano de una variedad de trigo, con cuatro tasas de siembra en un diseño de cuadrado latino de 5×5

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	24	716.61			
Renglón	4	99.20	24.80		
Columna	4	38.48	9.62		
Tasa de siembra	4	522.30	130.57	27.67	.000
Error	12	56.63	4.72		

Errores estándar para las medias de tratamiento

La estimación del error estándar para una media de tratamiento es:

$$s_{\bar{y}_k} = \sqrt{\frac{CME}{t}} = \sqrt{\frac{4.72}{5}} = 0.97 \quad (8.18)$$

y la estimación del error estándar para la diferencia entre dos medias de tratamiento es:

$$s_{(\bar{y}_k - \bar{y}_m)} = \sqrt{\frac{2 \text{ CME}}{t}} = \sqrt{\frac{2(4.72)}{5}} = 1.37 \quad (8.19)$$

Pruebas de hipótesis sobre medias de tratamientos

El estadístico F_0 para probar la hipótesis nula de que no existe diferencia entre las medias de tratamientos:

$$F_0 = \frac{CMT}{CME} = \frac{130.57}{4.72} = 27.67 \quad (8.20)$$

excede el valor crítico $F_{05,4,12} = 3.26$ con un nivel de significancia observado de $P_r > F = .000$ (tabla 8.8).

Interpretaciones del factor de tratamiento cuantitativo con contrastes de regresión

El factor de tratamiento para este experimento es un factor cuantitativo con cinco niveles. Un análisis para la regresión de la cosecha de trigo sobre las tasas de siembra con contrastes polinomiales ortogonales para la tasa de siembra proporcionará una buena descripción del efecto de la tasa de siembra sobre la cosecha de grano. Los coeficientes de los contrastes polinomiales ortogonales se encuentran en la tabla XI del apéndice. Los coeficientes ortogonales y las sumas de cuadrados para los contrastes de regresión polinomial lineal y cuadrática se muestran en la tabla 8.9.

Tabla 8.9 Particiones de sumas de cuadrados para la regresión lineal y cuadrática de las tasas de siembra (ejemplo 8.2)

Tasa de siembra	30	80	130	180	230	SC*
Media (\bar{y}_k)	47.13	51.72	55.73	59.17	58.88	
Lineal (P_{1k})	-2	-1	0	1	2	478.95
Cuadrática (P_{2k})	2	-1	-2	-1	2	38.11

* $SC = r[\sum P_{ck}\bar{y}_k]^2 / \sum P_{ck}^2$

Los estadísticos F_0 respectivos para probar las hipótesis nulas de los contrastes lineal y cuadrático son $F_0 = 478.95/4.72 = 101.47$ y $F_0 = 38.11/4.72 = 8.07$. Ambas razones exceden el valor crítico $F_{05,1,12} = 4.75$. La suma de cuadrados para las desviaciones de la regresión lineal y cuadrática es:

$$\begin{aligned} SC \text{ semilla(desviación)} &= SC \text{ semilla} - SC \text{ semilla(lineal)} - SC \text{ semilla(cuadrática)} \\ &= 522.30 - 478.95 - 38.11 = 5.24 \end{aligned}$$

con 2 grados de libertad; una prueba de hipótesis indicaría que no hay desviaciones significativas de la ecuación cuadrática.

La ecuación de regresión polinomial cuadrática calculada para estimar la cosecha de grano (\hat{y}) a partir de la tasa de siembra (R) mediante las técnicas descritas en el capítulo 3 es:

$$\hat{y} = 43 + 0.14R - 0.0003R^2$$

La figura 8.3 muestra una gráfica de esta ecuación, aquí es posible observar que el rendimiento máximo estimado ocurre con la tasa de siembra $R = 233$ libras por acre, que es una extrapolación hacia arriba de la tasa más alta usada en el experimento. Se requerirá otro experimento con un tasa de siembra de 230 libras por acre para estimar con seguridad la tasa máxima de siembra.

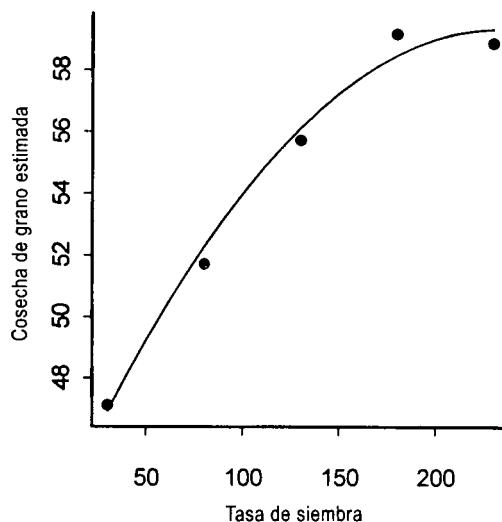


Figura 8.3 Respuesta estimada entre la cosecha y la tasa de siembra

Análisis de residuales para evaluar las suposiciones

Los residuales se puede usar para evaluar las suposiciones del modelo (como se describe en el capítulo 4); de la ecuación (8.17), el residual de las observaciones del k -ésimo tratamiento en el i -ésimo renglón y la j -ésima columna es:

$$\hat{e}_{ij} = y_{ij} - \bar{y}_i - \bar{y}_j - \bar{y}_k + 2\bar{y}$$

Las gráficas de residuales se dejan como ejercicio para el lector.

Si los tratamientos interactúan con los bloques

Tukey (1955) desarrolló una prueba para la suposición de efectos aditivos de tratamiento, renglón o columna. De la ecuación (8.17), la aditividad de los efectos del tratamiento, renglón y columna da como resultado el valor estimado de una observación:

$$\hat{y}_{ij} = \bar{y} + (\bar{y}_i - \bar{y}) + (\bar{y}_j - \bar{y}) + (\bar{y}_k - \bar{y}) = \bar{y}_i + \bar{y}_j + \bar{y}_k - 2\bar{y}$$

El cálculo de la suma de cuadrados de 1 grado de libertad para la no aditividad requiere los valores de \hat{y}_{ij} y el residual, $\hat{e}_{ij} = y_{ij} - \hat{y}_{ij}$, para cada unidad experimental. La suma de cuadrados para la no aditividad es:

$$SC(\text{no aditividad}) = \frac{\left[\sum_{i=1}^t \sum_{j=1}^t \hat{y}_{ij}^2 \hat{e}_{ij} \right]^2}{SC} \quad (8.21)$$

donde SC es la suma de cuadrados del error obtenido con un análisis de varianza del cuadrado latino sobre \hat{y}_{ij}^2 . El estadístico F_0 para probar la no aditividad es:

$$F_0 = \frac{SC(\text{no aditividad})}{CM \text{ residual}} \quad (8.22)$$

donde

$$CM \text{ residual} = \frac{[SCE - SC(\text{no aditividad})]}{v}$$

tiene $v = (t - 1)(t - 2) - 1$ grados de libertad y SCE es la suma de cuadrados para el error del análisis de varianza del cuadrado latino para las observaciones originales y_{ij} . Se recomienda que los valores de \hat{y}_{ij} se codifiquen según $k(\hat{y}_{ij} - \bar{y}_{..})$, donde k es una constante para cambiar la escala de valores por conveniencia para los cálculos.

¿Aumentaron la precisión los dos factores de bloque?

La eficiencia del diseño del cuadrado latino con dos criterios de bloque se determina en relación con el diseño de bloques completos aleatorizado con sólo un criterio de bloque. Las medidas de eficiencia relativa se puede calcular por separado para los criterios de bloque de renglón y columna en el cuadrado latino.

Eficiencia relativa del bloque por columna

Si sólo se usara el criterio de bloque de los renglones con un diseño de bloques completos aleatorizado, la estimación del cuadrado medio del error es:

$$s_{rcb}^2 = \frac{CM \text{ columnas} + (t - 1) CME}{t} \quad (8.23)$$

donde CME es el cuadrado medio del error para el error correspondiente al análisis de varianza del cuadrado latino.

El valor estimado para el ejemplo 8.2 es $s_{rcb}^2 = [9.62 + 4(4.72)]/5 = 5.70$ y $s_{ls}^2 = CME = 4.72$. La eficiencia relativa del bloque por columna para el experimento es:

$$RE_{col} = \frac{s_{rcb}^2}{s_{ls}^2} = \frac{5.70}{4.72} = 1.21$$

Hay una ganancia del 21% en eficiencia sobre el diseño de bloques completos aleatorizado, en el que sólo se usa el criterio de bloque correspondiente a los renglones del cuadrado latino. Así, los bloques de las columnas para los gradientes del suelo en el campo reducen la varianza del error en un 21% de manera efectiva. El diseño de bloques aleatorizado sin los bloques de las columnas para los gradientes de suelo requeriría $1.21(5) = 6$ réplicas para obtener una varianza estimada de la media del tratamiento igual a la del diseño del cuadrado latino.

Eficiencia relativa del bloque por renglón

Si sólo se usa el criterio de las columnas para el bloque con diseño de bloques completos aleatorizado, la media cuadrática estimada del error es:

$$s_{rcb}^2 = \frac{CM \text{ renglón} + (t - 1)CME}{t} \quad (8.24)$$

Para el ejemplo 8.2, $s_{rcb}^2 = [24.80 + 4(4.72)]/5 = 8.74$ y $RE_{col} = 8.74/4.72 = 1.85$. Existe una ganancia del 85% en la eficiencia con el bloque por renglón para el gradiente de irrigación en el experimento. Sin el bloque por el gradiente de irrigación el experimento requeriría $1.85(5) = 9.25$ o 10 réplicas de cada tratamiento en el diseño de bloques completos aleatorizado para obtener una varianza estimada del tratamiento igual a la del diseño del cuadrado latino.

Corrección para estimar σ^2

La corrección para estimar σ^2 mediante s^2 es:

$$\frac{(f_{ls} + 1)(f_{rcb} + 3)}{(f_{ls} + 3)(f_{rcb} + 1)} = \frac{13 \cdot 19}{15 \cdot 17} = 0.97$$

donde $f_{ls} = 12$ y $f_{rcb} = 16$ son los grados de libertad del error respectivos del cuadrado latino y el diseño de bloques completos aleatorizado. La corrección reduce la ER de 1.21 a $0.97(1.21) = 1.17$ para el bloque por columna y de 1.85 a $0.97(1.85) = 1.79$ para el bloque por renglón y tiene un efecto pequeño en la eficiencia de las estimaciones.

Rectángulos y cuadrados latinos múltiples

El diseño del cuadrado latino con dos criterios de bloque y cuatro tratamientos o menos es muy restrictivo y proporciona dos grados de libertad menos para una estimación efectiva de la varianza del error experimental. En estas circunstancias, es común repetir el experimento con más de un cuadrado.

Existen dos formas para diseñar cuadrados múltiples, la primera tiene formas de identificación distintas para renglones y columnas para cada cuadrado en el experimento y la segunda forma tiene una identificación común en todos los cuadrados para los renglones o para las columnas. Un ejemplo de la primera forma ocurre en las pruebas de campo agrícolas que usan dos cuadrados latinos en áreas separadas del campo de investigación. La segunda se ilustra con el experimento de las llantas para automóvil en las que se utilizan dos grupos de cuatro autos en arreglos de cuadrados latinos, donde los ocho automóviles representan el bloque por columna y las cuatro posiciones comunes a ambos cuadrados representan el bloque por renglón, esta última forma de diseño se conoce como *rectángulo latino*. Las dos formas se ilustran en la figura 8.4.

La figura 8.4a es representativa de experimentos agrícolas con cuadrados latinos únicos múltiples y la figura 8.4b lo es de experimentos como el de las llantas para automóvil con las posiciones en los renglones y los autos en las columnas en un rectángulo latino.

Renglón	Columna							
	1	2	3	4	5	6	7	8
1	A	B	C	D				
2	B	C	D	A				
3	C	D	A	B				
a)	D	A	B	C				
5					A	B	C	D
6					B	A	D	C
7					C	D	B	A
8					D	C	A	B

Renglón	Columna							
	1	2	3	4	5	6	7	8
1	A	B	C	D	A	B	C	D
b)	B	C	D	A	B	A	D	C
2	C	D	A	B	C	D	B	A
3	D	A	B	C	D	C	A	B

Figura 8.4 Cuadrados latinos múltiples con a) renglones y columnas únicas para cada cuadrado y b) rectángulo latino con renglones comunes a ambos cuadrados y columnas únicas en cada uno

La aleatorización se realiza por separado para cada cuadrado cuando el bloque por renglón y columna es único para cada uno de los s cuadrados. Con el rectángulo latino también es posible considerar único cada cuadrado y aleatorizar. Cuando el criterio por renglón es uniforme a través de las columnas, sólo es necesario que cada tratamiento ocurra s veces en cada renglón y una vez en cada columna, luego se realiza una aleatorización de los t renglones y otra separada del conjunto completo de las st columnas.

El modelo lineal para el rectángulo latino con criterios de bloque por renglón y columna únicas en cada cuadrado latino es:

$$y_{yl} = \mu + k_l + \rho_i(l) + \gamma_j(l) + \tau_k + e_{yl}$$

$$l = 1, 2, \dots, s \quad i, j, k = 1, 2, \dots, t \quad (8.25)$$

donde k_l es el efecto del cuadrado, y $\rho_{i(l)}$ y $\gamma_{j(l)}$ son los efectos respectivos de renglón y columna anidados en los cuadrados. Quizá sea necesario considerar una componente de interacción cuadrado \times tratamiento en el modelo si se sospecha que las comparaciones entre tratamientos pueden diferir de un cuadrado a otro. En la sección 8.6 se presenta un experimento por interacción entre tratamientos. En la tabla 8.10 se proporciona una descripción del análisis de varianza con bloques por renglón y columna únicos.

Tabla 8.10 Análisis de varianza para un experimento repetido con s arreglos de cuadrado latinos únicos

Fuente de variación	Grados de libertad	Sumas de cuadrados
Total	$st^2 - 1$	$\sum_l \sum_i \sum_j (y_{yl} - \bar{y}_{..})^2$
Cuadrados	$s - 1$	$t^2 \sum_l (\bar{y}_l - \bar{y}_{..})^2$
Renglones en los cuadrados	$s(t - 1)$	$t \sum_l \sum_i (\bar{y}_{il} - \bar{y}_{..l})^2$
Columnas en los cuadrados	$s(t - 1)$	$t \sum_l \sum_j (\bar{y}_{jl} - \bar{y}_{..l})^2$
Tratamientos	$t - 1$	$st \sum_k (\bar{y}_k - \bar{y}_{..})^2$
Error	$(st - s - 1)(t - 1)$	Restando

El modelo lineal para un rectángulo latino que tiene criterios comunes de bloque por renglón para s cuadrados latinos completos y bloque por columna único es:

$$y_y = \mu + \rho_i + \gamma_j + \tau_k + e_y$$

$$j = 1, 2, \dots, st \quad i, k = 1, 2, \dots, t. \quad (8.26)$$

En la tabla 8.11 se muestra una descripción del análisis de varianza para un bloque común por renglón y único por columna. Las estimaciones del error estándar para las medias de tratamiento y las diferencias entre las dos medias, respectivamente, son $\sqrt{CME/st}$ y $\sqrt{2CME/st}$.

8.4 Experimentos factoriales en diseño de bloques completos

El diseño de tratamientos usado para referirse a la hipótesis de investigación se puede ajustar a cualquier diseño de experimento compatible. Es posible calcular las particiones de sumas de cuadrados adecuadas del análisis de varianza siempre que se satisfagan las restricciones de aleatorización para el diseño de experimento en cuestión.

Tabla 8.11 Análisis de varianza para un experimento con t tratamientos en un rectángulo latino con t renglones y st columnas

Fuente de variación	Grados de libertad	Sumas de cuadrados
Total	$st^2 - 1$	$\sum_i \sum_j (y_{ij} - \bar{y})^2$
Renglones	$t - 1$	$st \sum_i (\bar{y}_i - \bar{y}_{..})^2$
Columnas	$st - 1$	$t \sum_j (\bar{y}_j - \bar{y})^2$
Tratamientos	$t - 1$	$st \sum_k (\bar{y}_k - \bar{y})^2$
Error	$(st - 2)(t - 1)$	Restando

El factorial de dos factores para un diseño de bloques completos aleatorizado tiene las combinaciones ab de tratamiento un número igual de veces en cada bloque. En el diseños de cuadrado latino, cada una de las ab combinaciones aparece una vez en cada renglón y una en cada columna. Un diseño de cuadrado latino con un criterio de renglones común para s cuadrados requiere que cada combinación de tratamiento aparezca una vez en cada columna y s veces en cada renglón.

El modelo lineal para un factorial de dos factores, el factor A con a niveles y el B con b niveles, en un diseño de bloques completos aleatorizado con r bloques es:

$$y_{ijk} = \mu + \rho_k + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad (8.27)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

En la tabla 8.12 se ilustran las particiones de sumas de cuadrados para un factorial de dos factores en un diseño de bloques completos aleatorizado. La suma de cuadrados del tratamiento con $(t - 1) = (ab - 1)$ grados de libertad se divide en sumas de cuadrados para los efectos principales de los factores A y B y los efectos de interacción como se describe en la tabla 6.5; la suma de cuadrados para los bloques y el error experimental son análogos a los mostrados en la tabla 8.2 para el diseño de bloques completo aleatorizado y los errores estándar se calculan de acuerdo a las convenciones descritas en el capítulo 6 para el experimento factorial en un diseño totalmente aleatorizado.

Para el cuadrado y el rectángulo latinos se sigue un patrón similar de análisis en el que se hace una partición de las sumas de cuadrados del cuadrado, los renglones y las columnas, según el análisis en las tablas 8.6, 8.10 y 8.11 con $t = ab$. Las particiones de sumas de cuadrados para los efectos principales del factorial y las interacciones se calculan como se describe en el capítulo 6.

Tabla 8.12 Análisis de varianza para un diseño de tratamientos de dos factores en un diseño experimental de bloques completos aleatorizado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Cuadrados medios esperados
Total	$rab - 1$	SC total		
Bloques	$r - 1$	SC bloque		
A	$a - 1$	SCA	CMA	$\sigma^2 + rb\theta_a^2$
B	$b - 1$	SCB	CMB	$\sigma^2 + ra\theta_b^2$
AB	$(a - 1)(b - 1)$	SC(AB)	CM(AB)	$\sigma^2 + r\theta_{ab}^2$
Error	$(ab - 1)(r - 1)$	SCE	CME	σ^2

* Vea en el capítulo 6 las fórmulas para calcular SCA, SCB y SC(AB). La SCE se obtiene con una resta. La SC bloques = $ab \sum (\bar{y}_k - \bar{y})^2$.

8.5 Datos faltantes en diseños por bloques

En el capítulo 6 se presentan los datos faltantes en estudios de investigación con diseños factoriales, las observaciones que faltan afectan las relaciones entre los bloques y los tratamientos al igual que las relaciones entre los factores en los factoriales.

Los efectos de tratamientos y bloques no son ortogonales si faltan datos y un contraste para un conjunto de efectos contiene cierta información sobre el otro conjunto de efectos. Así, los efectos de los bloques deben tomarse en cuenta al calcular la partición de la suma de cuadrados para los tratamientos.

Los modelos alternativos completo y reducido se usan en el diseño de bloques para 1) calcular particiones de sumas de cuadrados no sesgadas para los tratamientos y el error experimental, y 2) calcular estimaciones de mínimos cuadrados no sesgadas de las medias de tratamiento y sus errores estándar a partir de los datos disponibles. Los procedimientos para las soluciones de mínimos cuadrados de las ecuaciones normales y las sumas de cuadrados de los modelos completo y reducido se describen en el capítulo 6; muchos programas de computadora para el análisis de varianza incluyen la ejecución de este análisis.

Análisis con datos faltantes en un diseño de bloques completos

El análisis con datos faltantes para el diseño de bloques completos aleatorizado difiere un poco del diseño factorial descrito en el capítulo 6 porque el primero supone que no hay efectos de interacción. Entonces, el análisis omite cualquier prueba de interacción previa a la estimación de los efectos del tratamiento.

Las soluciones de las ecuaciones normales para el modelo completo, $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$, se usan para calcular las estimaciones, $\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\rho}_j$, y la suma de cuadrados del error para el modelo completo:

$$SCE_f = \sum (y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\rho}_j)^2$$

Las soluciones de las ecuaciones normales para el modelo reducido, $y_{ij} = \mu + \rho_j + e_{ij}$ se usan para calcular las estimaciones, $\hat{y}_{ij} = \hat{\mu} + \hat{\rho}_j$, y la suma de cuadrados del error experimental para el modelo reducido:

$$SCE_r = \sum (y_{ij} - \hat{\mu} - \hat{\rho}_j)^2$$

La suma de cuadrados del tratamiento se calcula como:

$$SC \text{ Tratamiento (ajustado)} = SCE_r - SCE_f \quad (8.28)$$

y representa la reducción en la suma de cuadrados como resultado de incluir τ_i en el modelo completo. Se conoce como una suma de cuadrados de tratamiento ajustada, lo que implica que también se consideran los efectos de bloque al estimar los efectos del tratamiento en el modelo completo.

Las particiones de sumas de cuadrados ortogonales con m observaciones faltantes se muestran en la tabla 8.13. El modelo reducido ignora la clasificación del tratamiento y la SC bloques (no ajustada) del modelo reducido es la suma de cuadrados debida a las diferencias entre las medias de bloques ignorando los tratamientos.

La aplicación de un análisis con observaciones faltantes se deja como ejercicio al final del capítulo.

Tabla 8.13 Análisis de varianza para un diseño de bloques completos aleatorizado con m observaciones faltantes

Fuente de variación	Grados de libertad	Sumas de cuadrados
Total	$tr - m - 1$	$\sum_i \sum_j (y_{ij} - \bar{y})^2$
Bloques (no ajustados)*	$r - 1$	$\sum_j n_j (\bar{Y}_j - \bar{y})^2$
Tratamiento (ajustado)	$t - 1$	$SCE_r - SCE_f$
Error	$(r - 1)(t - 1) - m$	SCE_f

* SC bloques (no ajustado) = SC total - SCE_f ; n_j = número de observaciones en el j -ésimo bloque

8.6 Experimentos realizados varias veces

Los experimentos se repiten en varios lugares o en distintas ocasiones por muchas razones, pues la repetición a través del tiempo o espacio proporciona una forma de réplica para incrementar la precisión de las estimaciones de las medias de tratamiento o aumentar los grados de libertad de las estimaciones del error experimental. Los experimentos repetidos pueden proporcionar una base de inferencia más amplia para evaluar los tratamientos sobre un conjunto más grande de condiciones. En otros casos, se espera que la magnitud de las comparaciones de tratamientos difiera entre lugares o tiempos. Las series de experimentos se usan para examinar la variación en la diferencias de tratamiento relativas a los cambios ambientales.

Cualquiera que sea la razón para llevar a cabo el mismo experimento en una serie de lugares o tiempos, debe tenerse cierta cautela antes de combinar los datos

de las series para realizar un solo análisis global. El siguiente ejemplo aclarará algunos detalles a través de una serie de experimentos.

Ejemplo 8.3 Eficiencia del uso de agua en el pasto Bermuda

El pasto Bermuda se usa mucho en jardines, parques y campos de golf en climas calientes y secos. En áreas secas su mantenimiento requiere irrigación constante. Los cuidadores de césped desean especies de plantas que usen el agua de manera eficiente para reducir los costos de mantenimiento y conservar el agua. Existe una considerable variación en el uso eficiente de agua entre las especies y dentro de ellas.

Objetivo de investigación: Un agricultor quería determinar la cantidad de variabilidad en la eficiencia con la que el pasto Bermuda utiliza agua de manera que pudiera atribuirla a diferencias genéticas. Dada una variación genética suficiente, podría iniciar un programa de cultivo para desarrollar un pasto Bermuda híbrido con un uso eficiente de agua.

Diseño del tratamiento: El agricultor realizó todas las cruzas posibles (30) de híbridos entre seis cultivos de pasto Bermuda en lo que se conoce como *diseño de crusa bimodal*. Este diseño permite evaluar el potencial genético de cultivos específicos o de la población representada por los cultivos en él.

Diseño de experimento: Se cultivó la progenie y los seis cultivos padres en un diseño de bloques completos aleatorizado, con dos bloques en cada uno de cuatro sitios separados en parcelas. Dos de las cruzas no se reprodujeron, de manera que se tenían 34 parcelas en cada bloque de cada experimento.

El agricultor midió la producción total de materia seca de las plantas en cada parcela y la cantidad de agua utilizada en ellas para producir el material. La medida que usó para su análisis fue el cociente del agua utilizada entre la producción total de materia seca en cada parcela. El análisis de varianza con su fuente de variación, grados de libertad y cuadrados medios para los cuatro experimentos se muestra en la tabla 8.14.

Tabla 8.14 Análisis de varianza para el consumo de agua en cuatro experimentos de un diseño de crusa bimodal con seis cultivos de pasto Bermuda

Fuente de variación	Grados de libertad	Cuadrados medios por sitio			
		1	2	3	4
Bloque	1	2.80	20.17	2.53	1.81
Genotipos	33	1.08	17.85	1.92	1.39
Error	33	1.61	10.56	1.07	0.74

Fuente: Dr. W. Kneebone, Department of Plant Sciences, University of Arizona.

El sitio 2 es diferente

La característica más notable del análisis en la tabla 8.14 es que los cuadrados medios para todas las fuentes de variación en el sitio 2 son mucho más grandes

que las de los otros sitios, sobre todo, la varianza del error en el sitio 2, $CME = 10.56$, es mucho mayor que la correspondiente a las otras localizaciones. La variabilidad más grande en el sitio 2 indica que quizás las condiciones experimentales fueron diferentes de las de otros sitios.

Se requieren varianzas homogéneas al combinar los resultados de varios experimentos

Se realizaron cuatro experimentos, cada uno con dos réplicas, es decir, un total de ocho réplicas. Si se combinaran los cuatro experimentos en un análisis, las estimaciones de las medias de genotipos serían mucho más precisas.

Es necesaria la homogeneidad en las varianzas del error para hacer un análisis combinado de los cuatro experimentos, una prueba de tal homogeneidad en las varianzas del error con la prueba F_{Max} (capítulo 4) rechaza la hipótesis de iguales varianzas del error en los cuatro sitios. Una solución posible es una transformación a una escala logarítmica para lograr varianzas homogéneas antes del análisis combinado; el análisis para cada sitio después de la transformación logarítmica se muestra en la tabla 8.15.

Tabla 8.15 Análisis de varianza para $10[\log_{10}(\text{uso de agua})]$ en cuatro experimentos de un diseño de crusa bimodal con seis cultivos de pasto Bermuda

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Cuadrados medios por sitio</i>			
		<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Bloque	1	1.68	3.28	1.42	1.01
Genotipos	33	0.56	2.51	1.09	1.11
Error	33	0.82	1.54	0.56	0.51

La transformación logarítmica reduce de manera considerable la disparidad en las varianzas del error, pero los cuadrados medios observados para todas las fuentes de variación todavía son mayores en el sitio 2. Todavía debe tenerse cuidado al hacer un análisis combinado; deben interpretarse los resultados iniciales de cada experimento separado. El estadístico F_0 en cada sitio para probar las hipótesis de que no hay diferencias entre los genotipos, $F_0 = CM \text{ genotipo}/CM \text{ error}$, rechaza la hipótesis nula a un nivel de significancia de .05 para los sitios 3 y 4, y no la rechaza en los sitios 1 y 2. Se llega a las mismas conclusiones utilizando el estadístico F_0 con los datos de la tabla 8.14.

Modelo estadístico y análisis de varianza para el análisis combinado de experimentos

Los resultados de las pruebas F indican un desempeño diferente entre los genotipos en los cuatro sitios. El análisis combinado debe incluir la posibilidad de interacción genotipo \times sitio. El modelo estadístico para el análisis combinado con efectos aleatorios de sitio y genotipo es:

$$y_{ijk} = \mu + \rho_i + b_{j(i)} + g_k + (gp)_{ik} + e_{ijk} \quad (8.29)$$

donde μ es la media general, ρ_i es el efecto aleatorio del sitio, $b_{j(i)}$ es el efecto aleatorio de bloqueo dentro del sitio, g_k es el efecto aleatorio de genotipo, $(gp)_{ik}$ es el efecto aleatorio de la interacción genotipo \times sitio, y e_{ijk} es el efecto aleatorio del error. El análisis de varianza combinado para los datos transformados de los cuatro experimentos se muestra en la tabla 8.16 con los cuadrados medios esperados para los efectos aleatorios de genotipo y los efectos fijos o aleatorios del sitio.

Tabla 8.16 Análisis de varianza para 10 [\log_{10} (uso de agua)] para los experimentos combinados de un diseño de cruzas bimodal de seis cultivos de pasto Bermuda

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Cuadrados medios esperados
Sitios (L)	3	433.93	144.54	
Bloques/ L	4	7.39	1.85	
Genotipos (G)	33	81.11	2.46	$\sigma^2 + r\sigma_{gl}^2 + rl\sigma_g^2$
$G \times L$	99	93.01	0.94	$\sigma^2 + r\sigma_{gl}^2$
Error combinado	132	113.19	0.86	σ^2

El cuadrado medio del error en el análisis combinado es el error combinado de los cuatro experimentos (promedio de los cuadrados medios para el error de los cuatro experimentos), los bloques son únicos para cada experimento y constituyen un efecto de factor anidado para el análisis de varianza y la suma de cuadrados para los bloques dentro de los sitios es la suma de las sumas de cuadrados de los bloques en el análisis de varianza individual.

Los cuadrados medios esperados para el modelo aleatorio o mixto se derivaron siguiendo las partes expuestas en el capítulo 7. Si los genotipos (tratamientos) tienen niveles fijos, entonces σ_g^2 se sustituye por la componente equivalente de los efectos fijos, θ_g^2 . Cuando tanto sitios como genotipos (o tratamientos) son fijos, entonces σ_{gl}^2 también se sustituye por la componente equivalente para efectos fijos, θ_{gl}^2 en la tabla 8.16.

El hecho de que las repeticiones de experimentos en el tiempo y los lugares sean aleatorias depende del objetivo de la repetición, si las repeticiones se eligen para investigar las respuestas del tratamiento a cambios deliberados en el medio, entonces parece apropiado un modelo de efectos fijos para los sitios o el tiempo, pero si las repeticiones se justifican como representantes legítimos aleatorios de los lugares o el tiempo, entonces se puede usar el modelo de efectos aleatorios. Quizá lo más difícil sea considerar repeticiones aleatorias si sólo se dispone de un número limitado de lugares para el experimento o de semanas, meses o años sucesivos que representen la repetición en el tiempo.

Pruebas de hipótesis en el análisis combinado

La hipótesis de que no hay interacción genotipo \times sitio se realiza con el estadístico F_0 de la tabla 8.16 como $F_0 = CM(G \times L)/CME = 0.94/0.86 = 1.09$, y el valor

crítico $F_{05,99,132} = 1.36$ y no se rechaza la hipótesis nula de no interacción genotipo \times sitio. Se prueba la hipótesis de que no hay efectos de genotipo con el estadístico $F_0 = CMG/CME = 2.46/0.94 = 2.62$ con valor crítico $F_{.05,33,99} = 1.55$ y se rechaza la hipótesis nula de que no hay efectos del genotipo.

Análisis de la interacción tratamiento \times experimento

La no significancia de la interacción genotipo \times sitio debe estudiarse con cuidado, ya que el análisis de los experimentos separados reveló diferencias significativas al nivel .05 de significancia en dos sitios. Es muy posible que ciertas comparaciones de tratamientos interactúen con el entorno mientras que otras sean relativamente constantes en distintos entornos. Se mencionan tres conjuntos de comparaciones importantes en el diseño de cruzas bimodales del pasto Bermuda:

- comparaciones entre los seis cultivos padres
- comparaciones entre las 28 cruzas
- un contraste entre la media de los seis cultivos padres y la de las 28 cruzas

Un análisis de varianza separado para los cultivos padres proporciona las sumas de cuadrados para el primer conjunto de comparaciones y otro para las cruzas obtiene las sumas de cuadrados para el segundo conjunto de comparaciones. En el análisis se usa el modelo lineal de la ecuación (8.29), que es el mismo modelo usado al incluir todos los genotipos en el análisis. En la tabla 8.17 se muestran sólo la sumas de cuadrados requeridas por las comparaciones de interés, no se muestran las sumas de cuadrados para sitios y bloques dentro cada sitio.

Tabla 8.17 Análisis de varianza separado para padres y de un diseño de cruzas bimodal para seis cultivos de pasto Bermuda

1) *Análisis de padres*

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Sumas de cuadrados</i>	<i>Cuadrados medios</i>
Padres	5	3.45	0.69
Padres \times sitios	15	4.87	0.32
Error (P)	20	10.04	0.50

2) *Análisis de cruzas*

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Sumas de cuadrados</i>	<i>Cuadrados medios</i>
Cruzas	27	46.06	1.71
Cruzas \times sitios	81	79.56	0.98
Error (C)	108	101.22	0.94

Las sumas de cuadrados para el contraste entre las medias de los padres y las cruzas y la suma de cuadrados para la interacción entre el contraste y el sitio se pueden calcular restando. Se usan las sumas de cuadrados de las tabla 8.16 y 8.17, que son:

- 1) Padres contra cruzas

$$\begin{aligned} SC(P \text{ contra } C) &= SCG - SCP - SCC = 81.11 - 3.45 - 46.06 \\ &= 31.60 \end{aligned}$$

- 2) (Padres contra cruzas) \times sitio

$$\begin{aligned} SC[(P \text{ contra } C) \times L] &= SC(G \times L) - SC(P \times L) - SC(C \times L) \\ &= 93.01 - 4.87 - 79.56 \\ &= 8.58 \end{aligned}$$

El análisis de varianza separado para los cultivos padres y las cruzas también proporciona sumas de cuadrados separadas para el error experimental que se pueden identificar con cada conjunto de comparaciones. Una partición de la suma de cuadrados del error puede ser útil, puesto que los errores experimentales pueden diferir de manera considerable entre las comparaciones. La tabla 8.17 muestra el cuadrado medio del error experimental para padres y cruzas, Error(*P*) y Error(*C*), de los análisis separados.

La suma de cuadrados para el error experimental asociado con el contraste entre padres y cruzas se encuentra a partir de las sumas de cuadrados del error de las tablas 8.16 y 8.17 como sigue:

$$\begin{aligned} SC \text{ error}(P \text{ contra } C) &= SC \text{ error} - SC \text{ error}(P) - SC \text{ error}(C) \\ &= 113.19 - 10.04 - 101.22 \\ &= 1.93 \end{aligned}$$

La tabla 8.18 contiene un resumen del análisis de varianza con todas las particiones de sumas de cuadrados; las tres cuadrados medios para la partición del error experimental mostradas al final de la tabla tienen valores similares. Es posible usar el error combinado con 132 grados de libertad con cierta confianza de que las varianzas del error de los tres grupos de comparaciones son homogéneas.

Combinar o no combinar varianzas

La decisión de usar el término de la partición del error o el del error combinado puede afectar la prueba de hipótesis. Las pruebas con el error combinado tienen más grados de libertad y, por lo tanto, mayor capacidad (poder) para detectar diferencias que los cuadrados medios del error de las particiones, este efecto se puede observar en la tabla 8.18 con una prueba de la interacción (*P* contra *C*) \times *L*. El estadístico de prueba con el error combinado es $F_0 = 2.86 / 0.86 = 3.33$ con valor crítico $F_{0.05,3,132} = 2.67$, el estadístico de prueba con la partición del cuadrado medio del error es $F_0 = 2.86 / 0.48 = 5.96$ con valor crítico $F_{0.05,3,4} = 6.59$; en este último caso, con menos grados de libertad para el cuadrado medio del error, la hipótesis nula no se rechaza. La hipótesis se rechaza con el

Tabla 8.18 Análisis de varianza para $10[\log_{10}(\text{uso de agua})]$ para experimentos combinados de un diseño de cruzas bimodal de seis cultivos de pasto Bermuda

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Sitios (L)	3	433.63	144.54
Bloques/ L	4	7.39	1.85
Genotipos (G)	33	81.11	2.46
Padres (P)	5	3.45	0.69
Cruzas (C)	27	46.06	1.71
P vs. C	1	31.60	31.60
$G \times L$	99	93.01	0.94
$P \times L$	15	4.87	0.32
$C \times L$	81	79.56	0.98
$(P$ vs. $C) \times L$	3	8.58	2.86
Error combinado	132	113.19	0.86
Error(P)	20	10.04	0.50
Error(C)	108	101.22	0.94
Error(P vs. C)	4	1.93	0.48

cuadrado medio del error combinado. Debe observarse que la hipótesis nula no se rechazaría para ninguna de las pruebas a un nivel de significancia de 0.01, en esta situación la diferencia entre las dos pruebas no es grande, pero es necesario tener en mente las dos posibilidades.

En este ejemplo específico, el uso del error de la partición o combinado no afecta los resultados de otras pruebas F , ninguna de las otras particiones de interacción $G \times L$ es significativa aunque la interacción global genotipo \times sitio no era significativa, existe cierta evidencia de que una componente de la interacción fue significativa. Este resultado puede explicar las diferencias en los niveles significativos de las pruebas de genotipos a partir de los análisis separados de los sitios que aparecen en la tabla 8.18.

Más detalles que pueden encontrarse

McIntosh (1983) proporcionó tablas de análisis de varianza con fuentes de variación, grados de libertad y los estadísticos F_0 adecuados para probar las hipótesis para un grupo extenso de experimentos con modelos fijos aleatorios y mixtos combinados a través del tiempo y lugar o una combinación de ellos. Carmer, Nyquist y Walker (1989) proporcionaron fórmulas para la estimación de varianzas en diferencias medias combinadas a partir de experimentos combinados con diseños de tratamiento de dos y tres factores. Los experimentos tienen diseños de bloques completos aleatorizados con efectos de tratamiento fijos y efectos de tiempo y lugar aleatorios.

EJERCICIOS PARA EL CAPÍTULO 8

1. En un naranjal en Valencia se llevó a cabo un experimento de irrigación con un diseño de bloques completos aleatorizado en el que se usaron seis tratamientos de irrigación en ocho bloques de árboles. Los siguientes datos se refieren al peso en libras de la fruta cosechada en cada parcela.

<i>Método</i>	<i>Bloque</i>							
	1	2	3	4	5	6	7	8
Goteo	450	469	249	125	280	352	221	251
En canal	358	512	281	58	352	293	283	186
Rocío	331	402	183	70	258	281	219	46
Aspersor	317	423	379	63	289	239	269	357
Aspersor + rocío	479	341	404	115	182	349	276	182
Anegado	245	380	263	62	336	282	171	98

Fuente: Dr. R. Roth y Dr. B. Gardner, Department of Soil and Water Science, University of Arizona.

- Escriba un modelo lineal para el experimento, explique los términos y calcule el análisis de varianza.
 - ¿Cuáles son las suposiciones necesarias para que el análisis de varianza sea válido? ¿Cómo se relacionan con el experimento?
 - Calcule el error estándar estimado para una media de tratamiento de irrigación y la diferencia entre dos medias de tratamiento de irrigación.
 - Considere que el método de anegado es la práctica normal. Use el método de Dunnett para probar la diferencia entre el anegado y cada uno de los otros métodos.
 - Calcule la eficiencia relativa de este diseño respecto a un diseño totalmente aleatorizado. ¿Cuáles son sus conclusiones?
 - Obtenga las gráficas residuales del análisis e interprételas.
2. Un científico realizó una prueba de fertilizante en un pastizal, con un diseño de bloques completos aleatorizado. Asignó al azar cinco tratamientos de fertilizante a las parcelas de cada cinco bloques; los siguientes datos son $100 \times$ (porcentaje de fósforo) en una muestra de tejido de planta de cada parcela.

<i>Tratamiento</i>	<i>Bloque</i>				
	1	2	3	4	5
Sin fertilizante	7.6	8.1	7.3	7.9	9.4
50 lb de nitrógeno	7.3	7.7	7.7	7.7	8.2
100 lb de nitrógeno	6.9	6.0	5.6	7.4	7.0
50 lb de nitrógeno (Pentóxido de fósforo)					
+ 75 lb P ₂ O ₅	10.8	11.2	9.0	12.9	11.6
100 lb de nitrógeno + 75 lb P ₂ O ₅	9.6	9.3	12.0	10.6	10.4

Fuente: Dr. P. Ogden, Range Management, Univestiy of Arizona.

- a. Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
 - b. Calcule la suma de cuadrados con 1 grado de libertad para cada uno de los siguientes contrastes y pruebe la hipótesis nula en cada caso. Los cuatro contrastes ortogonales entre los cinco tratamientos son 1) sin fertilizante contra los cuatro tratamientos de fertilizante, 2) el efecto principal del nitrógeno, 3) el efecto principal de P_2O_5 y 4) la interacción entre el nitrógeno y P_2O_5 .
 - c. Calcule el error estándar para cada contraste del inciso b).
 - d. Calcule la eficiencia relativa del diseño de bloques completos aleatorizado.
 - e. Obtenga las gráficas residuales del análisis e interprételas.
3. Se midió la autoinductancia de bobinas con núcleos de óxido de hierro con diferentes condiciones del puente de medición. La temperatura de la bobina se mantuvo constante, se usaron cinco bobinas en el experimento se midió la autoinductancia para cuatro temperaturas (22° , 23° , 24° y 25°) del puente de medición; las temperaturas se utilizaron en orden aleatorio para cada bobina. Los siguientes datos son las desviaciones porcentuales de una medida estándar.

Tempertaura	Bobina				
	1	2	3	4	5
22	1.400	0.264	0.478	1.010	0.629
23	1.400	0.235	0.467	0.990	0.620
24	1.375	0.212	0.444	0.968	0.495
25	1.370	0.208	0.440	0.967	0.495

Fuente: H. Hamaker (1955), Experimental design in industry, *Biometrics* 11, 257-286.

- a. Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
 - b. ¿Cuáles son las suposiciones necesarias para un análisis de varianza válido? ¿Cómo se relacionan con el experimento?
 - c. Calcule los contrastes de regresión polinomial ortogonal para la temperatura y sus sumas de cuadrados. Determine la ecuación que mejor se ajuste a los datos.
 - d. Calcule la eficiencia relativa de usar las bobinas como bloques.
 - e. Obtenga las gráficas residuales del análisis e interprételas.
4. Un ingeniero de tránsito realizó un estudio para comparar el tiempo sin uso de la luz roja para cinco secuencias distintas de semáforo, el experimento se llevó a cabo con un diseño de cuadrado latino en el que los dos factores de bloque eran 1) cinco cruceros elegidos al azar y 2) cinco periodos. En la tabla de datos, las cinco secuencias de tratamiento se muestran entre paréntesis como A, B, C, D, E y los valores numéricos son el tiempo de luz roja sin uso expresado en minutos.

Crucero	Periodo				
	1	2	3	4	5
1	15.2(A)	33.8(B)	13.5(C)	27.4(D)	29.1(E)
2	16.5(B)	26.5(C)	19.2(D)	25.8(E)	22.7(A)
3	12.1(C)	31.4(D)	17.0(E)	31.5(A)	30.2(B)
4	10.7(D)	34.2(E)	19.5(A)	27.2(B)	21.6(C)
5	14.6(E)	31.7(A)	16.7(B)	26.3(C)	23.8(D)

Fuente: Mason, Gunst y Hess (1989), 393.

- Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
 - Calcule el error estándar de la media de una secuencia de tratamiento y para la diferencia entre dos medias.
 - Use “comparaciones múltiples” con el mejor procedimiento para seleccionar el conjunto de secuencias de semáforo con el menor tiempo desperdiciado.
 - ¿Cuál es la eficiencia relativa del bloque por periodo?
 - Obtenga las gráficas residuales del análisis e interpretelas.
5. Un ingeniero investigador estudió la eficiencia en tiempo de cuatro métodos de fabricación (A, B, C, D) de un componente electrónico, se eligieron cuatro técnicos para el estudio, pero como el proceso de fabricación produce fatiga de manera que el tiempo requerido por el técnico aumenta al cambiar de un método a otro sin importar el orden, el ingeniero usó un diseño de cuadrado latino con los “técnicos” en las columnas y los “periodos” en los renglones. Los métodos de fabricación se asignaron al azar a los técnicos y los períodos, de acuerdo con el arreglo del cuadrado latino. Los valores son los tiempos de fabricación en minutos requeridos para el componente con el método indicado entre paréntesis.

Periodo de tiempo	Técnico			
	1	2	3	4
1	90(C)	96(D)	84(A)	88(B)
2	90(B)	91(C)	96(D)	88(A)
3	89(A)	97(B)	98(C)	98(D)
4	104(D)	100(A)	104(B)	106(C)

- Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
 - Calcule el error estándar de la media de un método de fabricación y la diferencia entre dos medias.
 - Use el método de Tukey para todas las comparaciones por pares entre las medias de los tiempos de fabricación.
 - Use la medida de eficiencia relativa para determinar si el periodo fue un factor de bloque crítico para reducir la varianza del error experimental.
6. El experimento en el ejemplo 8.2 sobre la relación entre la cosecha de trigo y la tasa de siembra se realizó en un diseño de cuadrado latino de 5×5 . Se realizó una réplica del experimento en un terreno adyacente.

te, por lo que se tienen dos repeticiones del experimento con bloques únicos por renglón y columna para cada uno, con las tasas de siembra de 30, 80, 30, 180 y 230 para A, B, C, D y E, respectivamente. A continuación se proporcionan los datos del segundo experimento, con la cosecha correspondiente a cada parcela expresada en quintales (100 lb) por acre.

Renglón de campo	Columna de campo				
	1	2	3	4	5
1	26.88(A)	38.40(D)	35.33(E)	34.56(C)	24.57(B)
2	37.63(E)	24.57(A)	36.09(C)	23.81(B)	32.25(D)
3	29.95(C)	29.18(B)	33.02(D)	22.27(A)	33.02(E)
4	32.25(D)	31.49(C)	21.50(B)	33.02(E)	18.43(A)
5	26.11(B)	36.09(E)	23.81(A)	29.95(D)	29.95(C)

Fuente: Dr. M. Ottman, Department of Plant Sciences, University of Arizona.

- Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza.
 - Calcule el error estándar estimado para la media de una tasa de siembra y la diferencia entre dos medias de tasas de siembra.
 - Calcule la eficiencia relativa del bloque por renglón para este experimento. ¿Se necesitarán más réplicas para un diseño de bloques completos aleatorizado si se usan sólo las columnas como bloques? Si es así, ¿Cuántas más recomendaría?
 - Calcule las particiones de sumas de cuadrados de regresión polinomial lineal y ortogonal para la tasa de siembra y pruebe sus hipótesis nulas. ¿Son significativas las desviaciones de una relación lineal o una cuadrática?
 - ¿Piensa que es razonable realizar un análisis de varianza para los dos experimentos combinados?, ¿por qué?
 - Calcule el análisis de varianza para los dos experimentos combinados.
 - Calcule las estimaciones de los errores estándar para una media de tasa de siembra y la diferencia entre dos tasas de los experimentos combinados.
 - ¿Existe interacción entre los contrastes lineales o cuadráticos de las tasas de siembra y los experimentos?
7. Un horticultor realizó un experimento con fertilizante de nitrógeno para lechuga en un diseño de bloques completos aleatorizado, asignó al azar cinco tasas de tratamiento de nitrato de amonio (0, 50, 100, 150 y 250 lb/acre) a cada dos parcelas en cada dos bloques para obtener un total de cuatro parcelas de cada nivel de nitrógeno y cada bloque estaba formado por diez parcelas, dos para cada tratamiento en cada bloque. Los siguientes datos son el número de lechugas en cada parcela.

Nitrógeno	Bloque 1	Bloque 2
0	104	114
50	134	130
100	146	142
150	147	160
200	133	146
	156	163
	161	

Fuente: Dr. W. D. Pew, Department of Plant Sciences, University of Arizona.

- a. Escriba un modelo lineal para el experimento, explique los términos y calcule el análisis de varianza. Observe que existen varias parcelas con cada tratamiento en cada bloque, ¿en qué afecta esto sus estimaciones del error experimental a partir del análisis de varianza?
- b. Pruebe la suposición de no interacción bloque \times tratamiento.
- c. Calcule las particiones de las sumas de cuadrados de regresión polinomial lineal y cuadrática para el nitrógeno y pruebe las hipótesis nulas. Interprete los resultados.
- d. ¿Son significativas las desviaciones cúbicas?
8. Se realizó un experimento de parcelas para evaluar la interacción entre los tiempos de aplicación de nitrógeno al suelo (pronto, óptimo, tarde) y dos niveles de un inhibidor de nitrificación (ninguno, .5 lb/acre). El inhibidor retrasa la conversión de las formas de amonio del nitrógeno en formas de nitrato móviles para reducir las pérdidas de nitratos derivados. El nitrógeno se suministró mediante impulsos marcados de ^{15}N a través de un sistema de irrigación por goteo en tres tiempos: pronto, óptimo y tarde. Los siguientes datos son el porcentaje de ^{15}N que absorbieron las plantas de maíz dulce cultivadas en las parcelas.

<i>Inhibidor de nitrógeno</i>						
<i>Bloque</i>	<i>Ninguno</i>			<i>.5 lb/acre</i>		
	<i>Pronto</i>	<i>Óptimo</i>	<i>Tarde</i>	<i>Pronto</i>	<i>Óptimo</i>	<i>Tarde</i>
1	21.4	50.8	53.2	54.8	56.9	57.7
2	11.3	42.7	44.8	47.9	46.8	54.0
3	34.9	61.8	57.8	40.1	57.9	62.0

Fuente: Dr. T. Doerge. Department of Soil and Water Science, University of Arizona.

- a. Escriba un modelo lineal para el experimento, explique los términos y calcule el análisis de varianza.
- b. Calcule las estimaciones del error estándar para las medias marginales de inhibidor de nitrógeno y los tiempos de aplicación de nitrógeno, y las medias de celdas.
- c. Pruebe las hipótesis nulas de no efectos de interacción y de no efectos principales para los dos factores.
- d. Calcule la eficiencia relativa del diseño de bloques completos aleatorizado.
- e. Obtenga las gráficas residuales a partir del análisis e interprételas.
9. Utilice los datos del ejercicio 8.1 sobre el experimento de irrigación en los naranjales de Valencia. Suponga que las parcelas de irrigación por goteo en el bloque 1 (450) y las de anegado en el bloque 5 (336) se perdieron en el experimento.
- a. Use un programa de computadora adecuado para calcular el análisis de varianza ortogonal según las siguientes particiones:

$$\text{Fuente de variación: } SC \text{ bloque (no ajustado)} = 432\ 384$$

$$SC \text{ tratamientos de irrigación (ajustado por bloques)} = 51\ 923$$

$$SC \text{ error} = 130\ 402$$

- b. Pruebe las hipótesis de que no hay diferencias entre tratamientos.

- c. Muestre las estimaciones de mínimos cuadrados de las medias de tratamiento y las de sus errores estándar, si el programa de computadora puede producirlas. Para las medias de tratamiento sin observación faltantes, los errores estándar deben ser los mismos que los calculados en forma usual.
- d. ¿En qué afectaron las parcelas perdidas a los errores estándar de las medias de tratamiento para el goteo y el anegado?
- e. Si el programa de computadora puede hacerlo, calcule el error estándar de la diferencia entre:
- las medias de irrigación por goteo y por anegado.
 - las medias de irrigación por goteo y por canal.
 - la medias de irrigación por anegado y por aspersión.
 - las medias de irrigación por canal y por aspersión.
- f. ¿En qué afectó a los errores estándar del inciso e) la pérdida de las parcelas? ¿Qué efecto tuvo esto en las pruebas asociadas de las diferencias entre los pares de medias de tratamiento?
10. Un científico de animales realizó una prueba de alimentación para reses con cuatro tratamientos compuestos de diferentes calidades de agua potable para los animales, en un diseño totalmente aleatorizado con dos réplicas. El experimento se llevó a cabo en los meses de primavera y de invierno dos años seguidos y cada una de las cuatro pruebas duró 112 días. Los datos que siguen son las ganancias de peso promedio diarias en cada corral de animales para cada prueba.

Tratamiento	Año 1		Año 2	
	Primavera	Invierno	Primavera	Invierno
1	1.81	2.14	2.06	2.17
	1.88	2.32	1.91	2.55
2	1.77	2.27	1.57	2.06
	1.60	2.02	1.32	2.20
3	1.85	2.13	1.51	2.25
	1.59	1.93	1.49	1.94
4	1.51	1.85	1.31	1.83
	1.56	1.95	1.20	2.15

Fuente: Dr. D. Ray, Department of Animal Sciences, University of Arizona.

- a. Calcule el análisis de varianza para cada una de las cuatro pruebas como un diseño totalmente aleatorizado.
- b. Determine si las varianzas del error experimental son homogéneas entre los experimentos.
- c. Calcule el análisis de varianza combinado para las cuatro pruebas con año, estación, tratamiento y todos los efectos de interacción en el modelo. El error experimental es el error combinado de los cuatro análisis de varianza a partir de las pruebas separadas.
- d. Suponga que los años son efectos aleatorios y que las estaciones y tratamientos son efectos fijos. ¿Cuáles son sus conclusiones?
11. Utilice los datos del ejercicio 8.4, el estudio de semáforos. Suponga que faltan las observaciones del crucero 1 durante el periodo 2 (33.8) y del crucero 4 en el periodo 5 (21.6).
- a. Utilice un programa de computadora adecuado para calcular el análisis de varianza ortogonal de acuerdo con las siguientes particiones:

Fuente de variación

-
- Cruceros (no ajustados)
 Periodos (ajustados por crucero)
 Secuencia de luces (ajustadas por crucero y periodo)
 Error
-

- b. Pruebe la hipótesis de que no hay diferencia por la secuencia.
 - c. Muestre las estimaciones de mínimos cuadrados de las medias de secuencias y su errores estándar estimados si el programa de computadora puede producirlas.
 - d. ¿En qué afectan los datos perdidos a los errores estándar de las medias de tratamiento para las secuencias **B** y **C**.
 - e. Si el programa puede hacerlo, calcule el error estándar de la diferencia entre:
 - i) las medias de las secuencias **B** y **C**.
 - ii) las medias de las secuencias **B** y **A**.
 - iii) las medias de las secuencias **C** y **D**.
 - iv) las medias de las secuencias **A** y **D**.
 - f. ¿De qué manera afectaron los datos perdidos a los errores estándar del inciso e)? ¿Qué efecto tienen estos datos faltantes en las pruebas asociadas de las diferencias entre los pares de medias de tratamientos?
12. Debe realizarse un experimento con un diseño de bloques completos aleatorizado con $t = 6$ tratamientos en $r = 4$ bloques.
- a. Proporcione una asignación aleatoria de seis tratamientos a las unidades experimentales con un diseño de bloques completos aleatorizado, muestre los detalles de su procedimiento de aleatorización.
 - b. ¿Cuántos arreglos distintos de tratamientos son posibles en cada bloque?
 - c. ¿Cuántos arreglos distintos son posibles para todo el experimento?
13. Debe realizarse un experimento en un arreglo de cuadrado latino con $t = 5$ tratamientos. Seleccione uno de los cuadrados latinos normales del apéndice 8A y aleatorice los cinco tratamientos a las unidades experimentales en el arreglo y muestre los detalles de su procedimiento de aleatorización.
14. Construya un arreglo de cuadrado latino de 7×7 normal.
- a. Dé una asignación aleatoria de los siete tratamientos a las unidades experimentales en el arreglo, muestre los detalles de su construcción y aleatorización del diseño.
 - b. ¿De cuántas maneras pueden arreglarse las columnas del cuadrado latino en bloques de columna?
 - c. ¿De cuántas maneras pueden arreglarse los renglones del cuadrado latino en bloques de renglón?
 - d. ¿De cuántas maneras pueden arreglarse las letras en los bloques de tratamiento?
 - e. ¿Cuántos arreglos de columnas, renglones y tratamientos son posibles para el experimento completo?
15. Debe realizarse un experimento en un arreglo de rectángulo latino de 4×8 con cuatro tratamientos. Seleccione dos cuadrados latinos estándar del apéndice 8A y realice una asignación aleatoria de los cuatro tratamientos a las unidades experimentales en el arreglo, muestre los detalles de su procedimiento de aleatorización.
16. Debe realizarse un experimento sobre pruebas de falla acelerada con motores eléctricos pequeños a cinco temperaturas diferentes, pero se puede realizar un máximo de cinco pruebas en un día y se dispone de 20

motores para las pruebas. Diseñe un experimento con una asignación aleatoria de las cinco temperaturas a los motores y muestre un bosquejo del conjunto final de pruebas que se harán.

17. Se le pide que realice un diseño para investigar la relación planta-suelo en un bosque mixto de cedros, el factor a estudiar es el porcentaje de cedros en el bosque mixto. Ha identificado tres sitios réplica adecuados para cada uno de los siguientes porcentajes: 1) 0%, 2) 20-30%, 3) 45-55%, 4) 70-80% y 5) 100% y debe recolectar muestras de hojarasca del suelo de cada bosques y realizar análisis químicos de laboratorio de las muestras. Tendrá que dividir el trabajo por partes iguales entre tres personas debido a la cantidad de trabajo en el campo y el laboratorio que esto implica.
- Diseñe el estudio para controlar la variación del error experimental con los 15 sitios elegidos entre usted y otras dos personas.
 - Bosqueje el análisis de varianza para los datos del estudio, incluya la fuente de variación y los grados de libertad para cada partición de sumas de cuadrados.
 - Suponga que toma dos muestras de cada sitio y repita el inciso b).
18. Usted lleva a cabo una prueba *in-vitro* de digestión de alimentos, en matraces que deben inocularse con CO₂ y microorganismos rumiados obtenidos de un novillo justo antes de inocular los matraces. El oxígeno y las temperaturas menores a 37°C pueden afectar a los microorganismos de manera adversa, pero dado el tiempo requerido para agregar el CO₂ y los microorganismos a los matraces, aun en las mejores condiciones, los primeros matraces reciben microorganismos sanos, pero los posteriores reciben microorganismos con actividad reducida.
- Suponiendo que tiene cinco tratamientos y 25 matraces que inocular en serie, establezca un diseño de bloques completos que controle la variación causada por la actividad reducida de los microorganismos por la exposición.
 - Bosqueje el análisis de varianza para los datos del estudio, incluyendo la fuente de variación y los grados de libertad para cada partición de sumas de cuadrados.
 - Suponga que dispone de 50 matraces para los cinco tratamientos, ¿cómo podría diseñar el estudio? Repita el inciso b).
19. Usted realizará un estudio para determinar la contaminación de los ríos debida a la actividad humana en un bosque, para lo que ha localizado cuatro ríos, cada uno con una pequeña comunidad permanente cercana con una planta de procesamiento de desperdicio en el cauce del río. Además, cada río tiene un campamento recreativo grande localizado cinco o diez millas río abajo de la comunidad.
- Usted debe tomar muestras del agua de los cuatro ríos: una muestra río arriba de la comunidad, otra muestra 1 milla río abajo y otras dos justo antes y después del campamento recreativo.
- También debe tomar una muestra en cada uno de cuatro días de la semana: viernes, domingo, lunes y miércoles, pero como tiene recursos limitados sólo puede tomar 4 muestras de agua de cada río para obtener un total de 16 muestras para todo el estudio.
- Establezca un diseño de bloques completos para adquirir las muestras de agua, con la “localización en el río” como factor de tratamiento.
 - Bosqueje el análisis de varianza para los datos, incluyendo la fuente de variación y los grados de libertad de cada partición de sumas de cuadrados.
 - Suponga que toma dos muestras de agua cada vez que toma una muestra de la localización en el río y repita el inciso b).

Apéndice 8A Selección de cuadrados latinos

4 × 4

A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
B	A	D	C	B	C	D	A	B	D	A	C	B	A	D	C
C	D	B	A	C	D	A	B	C	A	D	B	C	D	A	B
D	C	A	B	D	A	B	C	D	C	B	A	D	C	B	A

5 × 5

A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
B	A	E	C	D	B	A	D	E	C	B	A	D	E	C
C	D	A	E	B	C	E	B	A	D	C	D	E	A	B
D	E	B	A	C	D	C	E	B	A	D	E	B	C	A
E	C	D	B	A	E	D	A	C	B	E	C	A	B	D

A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
B	C	D	E	A	B	C	E	A	D	B	C	A	E	D
C	E	A	B	D	C	A	D	E	B	C	E	D	A	B
D	A	E	C	B	D	E	B	C	A	D	A	E	B	C
E	D	B	A	C	E	D	A	B	C	E	D	B	C	A

6 × 6

A	B	C	D	E	F	A	B	C	D	E	F
B	F	D	C	A	E	B	A	F	E	C	D
C	D	E	F	B	A	C	F	B	A	D	E
D	A	F	E	C	B	D	C	E	B	F	A
E	C	A	B	F	D	E	D	A	F	B	C
F	E	B	A	D	C	F	E	D	C	A	B

A	B	C	D	E	F	A	B	C	D	E	F
B	C	F	A	D	E	B	A	E	C	F	D
C	F	B	E	A	D	C	F	B	A	D	E
D	E	A	B	F	C	D	E	F	B	C	A
E	A	D	F	C	B	E	D	A	F	B	C
F	D	E	C	B	A	F	C	D	E	A	B

7×7

A	B	C	D	E	F	G	A	B	C	D	E	F	G
B	E	A	G	F	D	C	B	E	A	G	F	D	C
C	F	G	B	D	A	E	C	F	G	B	D	A	E
D	G	E	F	C	B	A	D	G	E	F	B	C	A
E	D	B	C	A	G	F	E	D	B	C	A	G	F
F	C	D	A	G	E	B	F	C	D	A	G	E	B
G	A	F	E	B	C	D	G	A	F	E	C	B	D

 8×8

A	B	C	D	E	F	G	H	A	B	C	D	E	F	G	H
B	C	D	E	F	G	H	A	B	C	A	E	F	D	H	G
C	D	E	F	G	H	A	B	C	A	D	G	H	E	F	B
D	E	F	G	H	A	B	C	D	F	G	C	A	H	B	E
E	F	G	H	A	B	C	D	E	H	B	F	G	C	A	D
F	G	H	A	B	C	D	E	F	D	H	A	B	G	E	C
G	H	A	B	C	D	E	F	G	E	F	H	C	B	D	A
H	A	B	C	D	E	F	G	H	G	E	B	D	A	C	F

 9×9

A	B	C	D	E	F	G	H	I
B	C	D	E	F	G	H	I	A
C	D	E	F	G	H	I	A	B
D	E	F	G	H	I	A	B	C
E	F	G	H	I	A	B	C	D
F	G	H	I	A	B	C	D	E
G	H	I	A	B	C	D	E	F
H	I	A	B	C	D	E	F	G
I	A	B	C	D	E	F	G	H

10 × 10

A	B	C	D	E	F	G	H	I	J
B	C	D	E	F	G	H	I	J	A
C	D	E	F	G	H	I	J	A	B
D	E	F	G	H	I	J	A	B	C
E	F	G	H	I	J	A	B	C	D
F	G	H	I	J	A	B	C	D	E
G	H	I	J	A	B	C	D	E	F
H	I	J	A	B	C	D	E	F	G
I	J	A	B	C	D	E	F	G	H
J	A	B	C	D	E	F	G	H	I

9 Diseños de bloques incompletos: Introducción

En ocasiones es necesario bloquizar unidades experimentales en grupos más pequeños que una réplica completa de todos los tratamientos que se usaría con bloques completos aleatorizado o un diseño de cuadrado latino, como los que se ilustran en el capítulo 8. El diseño de bloques incompletos se usa para disminuir la varianza del error experimental y proporcionar comparaciones más precisas entre tratamientos de lo que es posible con el diseño de bloques completos. En este capítulo se presenta una descripción general de algunos grupos importantes de diseños de bloques incompletos, se muestran el método de aleatorización y los métodos básicos de análisis para diseños de bloques incompletos balanceados y parcialmente balanceados, tomando en cuenta también la eficiencia de los diseños.

9.1 Bloques incompletos de tratamientos para reducir el tamaño de los bloques

Por alguna razón los experimentos pueden requerir una reducción del tamaño de bloques. Los diseños de bloques completos pueden reducir las varianzas del error experimental estimadas, pero esta reducción puede ser insuficiente, pues el número de tratamientos puede ser tan grande que resultan imprácticos para reducirla. Además, con el agrupamiento natural de las unidades experimentales en bloques quizás no se obtengan las unidades por bloque necesarias para el número de tratamientos de un diseño de bloques completos. En el siguiente ejemplo, el número limitado de cámaras de control ambiental evita la ejecución de una réplica completa de todos los tratamientos en una corrida de las cámaras disponibles.

Ejemplo 9.1 Germinación de semillas de tomate a alta temperatura constante

Es usual que los tomates se produzcan durante los meses de invierno en las regiones áridas tropicales, la producción invernal se siembra a finales del verano cuando las temperaturas del suelo pueden exceder 40°C, lo que sobrepasa los 35°C, temperatura máxima sugerida para germinación.

Objetivo de investigación: Un científico de plantas deseaba determinar a qué intervalos de temperatura podía esperar la inhibición de la germinación de las semillas de tomate para un grupo de cultivos.

Diseño del tratamiento: Se eligieron cuatro temperaturas para representar un intervalo común para el área de cultivo en consideración: 25°C, 30°C, 35°C y 40°C. La semilla de tomate se sometió a una temperatura constante en cámaras de ambiente controlado.

Diseño experimental: Una cámara sería una unidad experimental, pues la réplica verdadera de cualquier tratamiento de temperatura requería una corrida independiente del tratamiento en una cámara. Cualquier número de factores puede contribuir a la variación de la respuesta entre corridas, ya que las condiciones de todo el experimento debían repetirse para una corrida réplica, por lo que se consideró esencial bloquear las corridas.

El bloque completo y la réplica del experimento requerían cuatro cámaras; sin embargo, el científico sólo disponía de tres. Como el bloque natural de una corrida tenía menos cámaras (unidades experimentales) que tratamientos, construyó un diseño de bloques incompletos.

El cuadro 9.1 muestra un diagrama del diseño, en el que se probaron tres temperaturas diferentes en cada una de las cuatro corridas. Las corridas representan bloques incompletos de los tres tratamientos de temperatura y los tratamientos se asignaron al azar a las cámaras para cada corrida. Algunas características especiales de este diseño se analizan en la siguiente sección.

Cuadro 9.1 Diseño de bloques incompleto con cuatro tratamientos en bloques de tres unidades

Cámara			Cámara				
1	2	3	1	2	3		
Corrida 1	25°	30°	40°	Corrida 2	35°	30°	25°
Corrida 3	40°	25°	35°	Corrida 4	40°	30°	35°

Fuente: Dr. J. Coons, Department of Botany, Eastern Illinois University.

Otros ejemplos: Los lotes de material para la investigación industrial sirven de bloques, pero puede no haber material suficiente en uno solo para todos los tratamientos experimentales; los criterios para agrupar sujetos pueden provocar que el número de sujetos parecidos sea insuficiente en cada grupo para asignar los tratamientos planeados para el estudio, la variedad agrícola de las pruebas con frecuencia contiene un gran número de especies y los bloques completos no son prácticos para reducir la varianza del error; los diseños de bloques incompletos son la elección adecuada para los experimentos del tipo de estos ejemplos.

La guía principal para el tamaño del bloque es tener un conjunto homogéneo de unidades experimentales para obtener comparaciones precisas entre los tratamientos. Los diseños de bloques incompletos fueron introducidos por Yates (1936a, 1936b) para experimentos en los que el número de unidades experimentales por bloque es menor que el número de tratamientos. Los diseños se desarrollaron ante la necesidad de experimentos que incluyeran el conjunto relevante de tratamientos para estudiar la hipótesis de investigación aun cuando estaban restringidos a tamaños de bloque lógicos.

Es posible hacer una clasificación del diseño de bloques incompleto en dos grandes grupos: aquellos arreglados como bloques incompletos aleatorizados con un criterio de bloque y aquellos con arreglos basados en el cuadrado latino con dos criterios de bloque. Los diseños también pueden ser balanceados, donde cada tratamiento se aparea un número igual de veces con los demás tratamientos en un mismo bloque, considerando todos los bloques en el experimento. Un diseño parcialmente balanceado ocurre cuando se tiene diferente número de pares de tratamientos en el mismo bloque o cuando algún par de tratamientos nunca ocurre en un mismo bloque. En este capítulo se presenta una descripción general de los diseños de bloques incompletos y una introducción al análisis de los datos correspondientes a estos diseños.

9.2 Diseños de bloques incompletos balanceados (BIB)

El diseño BIB compara todos los tratamientos con igual precisión

El **diseño de bloques incompleto balanceado** es un arreglo tal que todos los tratamientos tienen igual número de réplicas y cada par de tratamientos se presenta en el mismo bloque un número igual de veces en algún lugar del diseño, el balance obtenido con el mismo número de ocurrencias de todos los pares de tratamientos en el mismo bloque tiene como resultado una precisión igual en todas las comparaciones entre los pares de medias de tratamiento.

El diseño de bloque incompleto tiene r réplicas de t tratamientos en b bloques de k unidades experimentales con $k < t$ y el número total de unidades experimentales en $N = rt = bk$, en el cuadro 9.1 por ejemplo, el experimento con tomates descrito tiene $b = 4$ bloques de $k = 3$ unidades experimentales, cada uno de los $t = 4$ tratamientos tiene $r = 3$ réplicas y existe un total de $N = bk = 4 \cdot 3 = 12$, o sea, $N = rt = 3 \cdot 4 = 12$ unidades experimentales. Por inspección se observa que cada par de tratamientos ocurre dos veces en los bloques, el par $(25^\circ, 30^\circ)$ está en los bloques 1 y 2 y el par $(30^\circ, 35^\circ)$ en los bloques 2 y 4.

El número de bloques donde ocurre cada par de tratamientos es $\lambda = r(k-1)/(t-1)$, donde $\lambda < r < b$. El valor entero λ se deriva del hecho de que cada tratamiento está apareado con los otros $t-1$ tratamientos en algún lugar del diseño λ veces. Existen $\lambda(t-1)$ pares para un tratamiento específico en el experimento, el mismo tratamiento aparece en r bloques con $k-1$ de los otros tratamientos, y cada tratamiento aparece en $r(k-1)$ pares. Por lo tanto:

$$\lambda(t-1) = r(k-1) \quad o \quad \lambda = r(k-1)/(t-1)$$

Es decir, en el diseño del experimento con tomates del cuadro 9.1, $\lambda = 3(3-1)/(4-1) = 2$.

Se puede construir un diseño de bloques incompleto balanceado mediante la asignación de las combinaciones adecuadas de tratamientos a cada uno de los $b = \binom{t}{k}$ bloques; con frecuencia es posible obtener el balance con menos de $\binom{t}{k}$ bloques.

No existe un método único para construir todas las clases de diseños de bloques incompletos pero existen métodos para construir algunos. La construcción de estos diseños ha sido tema de diversas investigaciones matemáticas que han arrojado un vasto arreglo de diseños de bloques balanceados y parcialmente balanceados.

El apéndice 9A.1 proporciona algunos planes para pocos tratamientos y se pueden encontrar otras tablas con diseños útiles para muchas situaciones prácticas en Cochran y Cox (1957) y en Fisher y Yates (1963). En el capítulo 10 se ilustrarán con más detalle varias categorías de diseño de bloques incompletos balanceados tradicionales, en tanto este capítulo presenta una introducción a su estructura básica con ejemplos de su aplicación y análisis.

9.3 Cómo aleatorizar los diseños de bloques incompletos

Una vez construido el diseño básico con los números de código de los tratamientos, los pasos para la aleatorización son los siguientes:

- Paso 1.* Aleatorizar el arreglo de los bloques conformados por grupos de los números de código asignados a los tratamientos.
- Paso 2.* Aleatorizar el arreglo de los números de tratamiento dentro de cada bloque.
- Paso 3.* Aleatorizar la asignación de los tratamientos a los números de código en el plan.

A continuación se ilustra la aleatorización con el plan de diseño básico para $t = 4$ tratamientos en $b = 4$ bloques de $k = 3$ unidades experimentales cada uno. Antes de la aleatorización, el plan es:

Bloque

1	1	2	3
2	1	2	4
3	1	3	4
4	2	3	4

Paso 1. Suponiendo que los bloques para el experimento son las corridas de las tres cámaras de cultivo con tres tratamientos de temperatura usadas en el ejemplo 9.1, los grupos de tratamiento (1, 2, 3), (1, 2, 4), (1, 3, 4) y (2, 3, 4) deben asignarse de manera aleatoria a las corridas, se elige una permutación aleatoria de los números 1 a 4 y se asignan los cuatro bloques a las cuatro corridas; con la permutación 2, 4, 1, 3 la asignación es:

Corrida				Bloque original
1	1	2	4	2
2	2	3	4	4
3	1	2	3	1
4	1	3	4	3

Paso 2. Se asignan al azar los números de código de tratamiento a las cámaras de cultivo en cada corrida. Se elige una permutación de los números 1 a 4 para cada cámara y se omite el número de tratamiento ausente en la corrida. Las siguientes son las cuatro permutaciones junto con la asignación a cada cámara (A, B, C) en cada corrida:

Corrida	Cámara			Permutación
	A	B	C	
1	2	4	1	2 4 3 1
2	3	4	2	3 4 1 2
3	1	2	3	4 1 2 3
4	1	4	3	1 4 3 2

Paso 3. Suponiendo también que los tratamientos son las temperaturas 25°, 30°, 35° y 40°C, una permutación aleatoria 2, 4, 3, 1 produce una asignación de las temperaturas para los números de código de tratamiento con la sustitución 2 → 25°C, 4 → 30°C, 3 → 35°C y 1 → 40°C en la tabla anterior.

Corrida	Cámara		
	A	B	C
1	25°C	30°C	40°C
2	35°C	30°C	25°C
3	40°C	25°C	35°C
4	40°C	30°C	35°C

9.4 Análisis de diseños BIB

Modelo estadístico para los diseños BIB

El modelo lineal estadístico para un diseño de bloques incompleto balanceado es:

$$\begin{aligned} y_{ij} &= \mu + \tau_i + \rho_j + e_{ij} \\ i &= 1, 2, \dots, t \quad j = 1, 2, \dots, b \end{aligned} \quad (9.1)$$

donde μ es la media general, τ_i es el efecto fijo del i -ésimo tratamiento, ρ_j es el efecto fijo del j -ésimo bloque, y las e_{ij} son errores experimentales aleatorios independientes con media 0 y varianza σ^2 .

Recordando que se tienen r réplicas de los t tratamientos en b bloques incompletos de k unidades experimentales, el número total de observaciones es $N = rt = bk$, donde cada par de tratamientos aparece junto en $\lambda = r(k-1)/(t-1)$ bloques del experimento.

Los efectos de tratamientos y bloques no son ortogonales en el diseño de bloques incompletos porque no aparecen todos los tratamientos en cada bloque; por lo tanto, no sería correcto para el diseño incompleto calcular la partición de suma de cuadrados para los tratamientos igual que para los diseño de bloque completos; tampoco las medias de tratamiento observadas proporcionarían estimaciones no sesgadas de $\mu_i = \mu + \tau_i$. Los parámetros estimados y las sumas de cuadrados de los tratamientos para los diseños de bloque incompletos balanceados se calculan con fórmulas relativamente directas, el desarrollo de las estimaciones de mínimos cuadrados se encuentra en el apéndice 9A.3.

Particiones de sumas de cuadrados para diseños BIB

Las particiones de sumas de cuadrados se pueden derivar al considerar modelos alternativos completos y reducidos para el diseño. Las soluciones de las ecuaciones normales se obtienen para el modelo completo, $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$, con las estimaciones, $\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\rho}_j$, para calcular la suma de cuadrados del error experimental para el modelo completo:

$$SCE_f = \sum_i \sum_j (y_{ij} - \hat{y}_{ij})^2 = \sum_i \sum_j (y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\rho}_j)^2 \quad (9.2)$$

Las soluciones a las ecuaciones normales para el modelo reducido, $y_{ij} = \mu + \rho_j + e_{ij}$, con estimaciones, $\hat{y}_{ij} = \hat{\mu} + \hat{\rho}_j$, se usan para calcular las sumas de cuadrados del error experimental para el modelo reducido;

$$SCE_r = \sum_i \sum_j (y_{ij} - \hat{y}_{ij})^2 = \sum_i \sum_j (y_{ij} - \hat{\mu} - \hat{\rho}_j)^2 \quad (9.3)$$

La diferencia $SCE_f - SCE_r$ es la reducción en la suma de cuadrados que resulta al incluir τ_i en el modelo completo, se trata de la suma de cuadrados debida a los tratamientos después de considerar los efectos de bloque en el modelo, esto se conoce como SC tratamiento (ajustada) e implica que los efectos de bloque también se toman en cuenta al estimar los efectos de tratamiento en el modelo comple-

to. Para los diseños de bloque incompletos balanceados, la suma de cuadrados de los tratamientos ajustada, $SCE_r - SCE_f$, se puede calcular directamente como:

$$SCT(\text{ajustada}) = \frac{k \sum_{i=1}^t Q_i^2}{\lambda_t} \quad (9.4)$$

con $t - 1$ grados de libertad. La cantidad Q_i es un total de tratamiento ajustado que se calcula con:

$$Qi = y_{i.} - \frac{1}{k} B_i \quad (9.5)$$

donde $B_i = \sum_j^n n_{ij} y_j$ es la suma de todos los totales de bloques que incluyen el i -ésimo tratamiento y $n_{ij} = 1$ si el tratamiento i aparece en el bloque j y $n_{ij} = 0$ de otra manera. Esta corrección al total del tratamiento tiene el efecto neto de eliminar los efectos de bloque de ese total.

La suma de cuadrados para los bloques se deriva del modelo reducido al ignorar los tratamientos en el modelo con $SCB(\text{ajustada}) = SC \text{ total} - SCE_r$. Los efectos de tratamiento no se consideran cuando se estiman los efectos de bloque, y la suma de cuadrados recibe el nombre de suma de cuadrados *no ajustada*.

La partición aditiva de SC total es:

$$SC \text{ total} = SCB(\text{no ajustada}) + SCT(\text{ajustada}) + SCE_f \quad (9.6)$$

La tabla 9.1 muestra una descripción del análisis de varianza para las particiones de suma de cuadrados. Muchos programas de computadora tienen la capacidad de calcular las particiones correctas de suma de cuadrados, las estimaciones de mínimos cuadrados de las medias de tratamiento, los contrastes entre las medias de mínimos cuadrados y sus errores estándar, para diseño de bloques incompletos.

Tabla 9.1 Análisis de varianza para un diseño de bloque incompleto balanceado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Total	$N - 1$	$\sum (y_{ij} - \bar{y})^2$	
Bloques	$b - 1$	$k \sum (\bar{y}_j - \bar{y})^2$	CMB (no adj.)
Tratamientos	$t - 1$	$\frac{k \sum Q_i^2}{\lambda_t}$	CMT (aj.)
Error	$N - t - b + 1$	Restando	CME

Ejemplo 9.2 Vinilación de metilglucósidos

Se ha encontrado que el resultado de la adición de acetileno a un metilglucósido en presencia de una base, a alta presión, es la producción de varios éteres monovinílicos, proceso que se conoce como *vinilación*. Los éteres monovinílicos son convenientes para la polimerización en muchas aplicaciones industriales.

Objetivo de investigación: Los químicos deseaban obtener más información específica sobre el efecto de la presión en el porcentaje de conversión del metilglucósido en isómeros de monovinil.

Diseño de tratamiento: Con base en trabajos anteriores, se seleccionaron presiones dentro del intervalo que producía la conversión máxima. Se eligieron cinco presiones para estimar una ecuación de respuesta: 250, 325, 400, 475 y 550 psi.

Diseño del experimento: Como sólo se disponía de tres cámaras de alta presión para una corrida de las condiciones experimentales, fue necesario bloquizar las corridas porque no podía haber una variación sustancial de una corrida a otra producida por nuevas preparaciones de las cámaras para el experimento. Los químicos establecieron un diseño de bloques incompleto balanceado (corridas), cada uno con tres unidades experimentales (cámaras presurizadas) y se usaron tres presiones diferentes en cada corrida; el diseño obtenido tenía seis réplicas de cada tratamiento de presión.

Las presiones usadas en cada corrida y las conversiones porcentuales a isómeros de monovinil se muestran en la tabla 9.2 y la suma aditiva de la partición de suma de cuadrados del metilglucósido se muestra en la tabla 9.3.

Tabla 9.2 Conversión porcentual del metilglucosidos con acetileno a acetileno a alta presión, en un diseño de bloques incompleto balanceado

Corrida	Presión (psi)					y_j
	250	325	400	475	550	
1	16	18	—	32	—	66
2	19	—	—	46	45	110
3	—	26	39	—	61	126
4	—	—	21	35	55	111
5	—	19	—	47	48	114
6	20	—	33	31	—	84
7	13	13	34	—	—	60
8	21	—	30	—	52	103
9	24	10	—	—	50	84
10	—	24	31	37	—	92
$y_{1\cdot}$	113	110	188	228	311	950
* $B_{1\cdot}$	507	542	576	577	648	
$\dagger Q_1$	-56.0	-70.7	-4.0	35.7	95.0	

*Ejemplo: $B_1 = y_1 + y_2 + y_6 + y_7 + y_8 + y_9$

$\dagger Q_1 = y_{1\cdot} - \frac{1}{3}B_{1\cdot} = 113 - \frac{1}{3}(507) = -56.0$

Fuente: Drs. J. Berry y A. Deutschman, University of Arizona.

Inferencias para las medias de tratamiento

Las estimaciones de mínimos cuadrados de las medias de tratamiento y sus errores estándar estimados se muestran en la tabla 9.4 y la estimación de mínimos cuadrados para una media de tratamiento μ_i es $\hat{\mu}_i = \bar{\mu} + \hat{\tau}_i$, donde:

Tabla 9.3 Análisis de varianza para la conversión porcentual de metilglucósidos, en un diseño de bloques incompleto balanceado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr < F
Total	29	5576.67			
Bloques (no adj.)	9	1394.67	154.96		
Presión (aj.)	4	3688.58	922.14	29.90	.000
Error	16	493.42	30.84		

$$\widehat{\mu} = y \quad y \quad \widehat{\tau}_t = \frac{kQ_t}{\lambda_t} \quad (9.7)$$

Por ejemplo, de la tabla 9.2, $Q_1 = -56.00$ y $\widehat{\mu} = \bar{y} = 950/30 = 31.67$, de manera que:

$$\widehat{\tau}_1 = \frac{kQ_1}{\lambda_t} = \frac{3}{(3)(5)} (-56.00) = -11.20 \quad y \quad \widehat{\mu}_1 = 31.67 - 11.20 = 20.47$$

Errores estándar para las medias de tratamiento

El error estándar para una estimación de la media de tratamiento es:

$$s_{\widehat{\mu}_i} = \sqrt{\frac{CME}{rt} \left(1 + \frac{kr(t-1)}{\lambda_t} \right)} = \sqrt{\frac{30.84}{(6)(5)} \left(1 + \frac{(3)(6)(4)}{(3)(5)} \right)} = 2.44 \quad (9.8)$$

Una estimación de un intervalo de confianza del 95% para una media de tratamiento en la tabla 9.4 es $\widehat{\mu}_i \pm t_{025,16}(s_{\widehat{\mu}_i})$, donde $t_{025,16} = 2.120$. El error estándar de la diferencia estimada entre dos medias de tratamiento, $\widehat{\mu}_i - \widehat{\mu}_j$, es

$$s_{(\widehat{\mu}_i - \widehat{\mu}_j)} = \sqrt{\frac{2kCME}{\lambda_t}} = \sqrt{\frac{(2)(3)30.84}{(3)(5)}} = 3.51 \quad (9.9)$$

Tabla 9.4 Estimaciones de mínimos cuadrados de las medias de presión para la conversión porcentual de metilglucósidos, en un diseño de bloques incompletos balanceado

Presión (psi)	$\widehat{\mu}_i$	$s_{\widehat{\mu}_i}$
250	20.47	2.44
325	17.53	2.44
400	30.87	2.44
475	38.80	2.44
550	50.67	2.44

Pruebas de hipótesis sobre las medias de tratamientos

El estadístico F_0 para probar la hipótesis nula de que no hay diferencias entre las medias de tratamiento es:

$$F_0 = \frac{CMT(\text{adj.})}{CME} = \frac{922.14}{30.84} = 29.90 \quad (9.10)$$

con $Pr > F = .000$ (tabla 9.3). El valor crítico para un nivel de significancia de .05 es $F_{.05,4,16} = 3.01$; por lo tanto, existen diferencias significativas entre las presiones con respecto a la conversión de metilglucósidos en productos de la vinilación.

Contrastes entre las medias de tratamientos

Los contrastes se calculan con las estimaciones de mínimos cuadrados de las medias de tratamiento como sigue:

$$c = \sum_{i=1}^t d_i \hat{\mu}_i \quad (9.11)$$

con error estándar estimado de:

$$s_c = \sqrt{\frac{kCME}{\lambda_f} \left(\sum_{i=1}^t d_i^2 \right)} \quad (9.12)$$

El estadístico $t_0 = c/s_c$ se puede usar para probar la hipótesis nula $H_0: C = 0$ con valor crítico basado en el estadístico t de Student con $N - t - b + 1$ grados de libertad.

La suma de cuadrados de 1 grado de libertad para el *contraste* se puede calcular con las medias de mínimos cuadrados $\hat{\mu}_i$ como:

$$SCC = \frac{\lambda_f (\sum d_i \hat{\mu}_i)^2}{k \sum d_i^2} \quad (9.13)$$

El estadístico $F_0 = SCC/CME$ se usa para probar la hipótesis nula $H_0: C = 0$ con valor crítico $F_{\alpha,1,(N-t-b+1)}$.

El tratamiento de presión para este experimento es un factor cuantitativo con cinco niveles y la regresión de la conversión porcentual de metilglucósido a presión usando contrastes polinomiales ortogonales para la presión describe el efecto de la presión sobre la tasa de conversión. El análisis de regresión se deja como ejercicio al final del capítulo.

Recuperación de la información de tratamientos a partir de la comparación de bloques

El análisis del diseño de bloques incompletos ilustrado hasta aquí estima los efectos de tratamientos con base en la información del tratamiento contenida en los bloques; esto se conoce como análisis *intrabloques*. Los diseños de bloques incompletos son no ortogonales

debido a que no todos los tratamientos aparecen en todos los bloques, y las comparaciones entre bloques contienen cierta información sobre las comparaciones de tratamientos. Yates (1940a) demostró que esta información *interbloques* se puede recuperar con un análisis interbloques y combinarse con la información del análisis intrabloque.

Los contrastes de bloques contienen contrastes de tratamientos

La información sobre la comparaciones de tratamientos contenida en una comparación entre bloques se puede ilustrar con las dos primeras corridas del experimento descrito en el ejemplo 9.2. Los datos de la conversión porcentual de metilglucósidos en las dos primeras corridas son:

Corrida	Presión (psi)						Media
	250	325	400	475	550		
1	16	18	–	32	–	22	
2	19	–	–	46	45	37	

Un contraste entre las medias de las dos corridas, 37 y 22, también es un contraste entre dos de los tratamientos de presión, 550 psi y 325 psi. Todas las comparaciones de los bloques contienen comparaciones similares de los tratamientos. El objetivo es recuperar esta información entre bloques, o interbloques, sobre los tratamientos y combinarla con la información dentro de los bloques, o intrabloques.

El tema de recuperación de información interbloques se menciona aquí sólo para indicar la disponibilidad del método, puede encontrarse un estudio exhaustivo en Kempthorne (1952), Cochran y Cox (1957), John (1971), John (1987) y John y Williams (1995).

La información del análisis interbloques se incorpora al análisis intrabloques con un estimador de los efectos de tratamiento que combina los estimadores intra e interbloques de los efectos del tratamiento; si el bloque ha sido eficaz para reducir el error experimental, la estimación interbloques contribuye sólo con una cantidad pequeña de información a las estimaciones combinadas, pero si los efectos del bloque son pequeños, la información recuperada con la estimación interbloques puede ser sustancial. Si el bloque no es efectivo, entonces el análisis con la recuperación de la información interbloques se reduce casi al análisis normal sin el ajuste con bloques.

9.5 Diseño renglón-columna para dos criterios de bloque

Cuando existe la necesidad de controlar la variación con más de un criterio de bloque, el diseño de cuadrado latino es un diseño de bloques completos usado para controlar la variación entre las unidades experimentales con dos factores de bloque; pero puede ser impráctico en algunas situaciones, pues el número de unidades experimentales que requiere, $N = t^2$, puede exceder las restricciones del material experimental o el número de tratamientos puede exceder el tamaño de los bloques disponibles.

Cuando se requieren dos criterios de bloque para el experimento es posible usar diseños renglón-columna con los renglones, las columnas o ambos como bloques incompletos; estos diseños se arreglan en p renglones y q columnas de unidades experimentales. Consideremos el experimento clásico para probar cuatro llantas de automóvil en las cuatro posiciones de cuatro automóviles con un diseño de cuadrado latino y suponiendo que el equipo de investigación quiere evaluar $t = 7$ tratamientos de llantas, todavía se tiene la necesidad de controlar la variación debida a la posición de la llanta y al automóvil, pero los autos sólo tienen cuatro posiciones para probar siete llantas. Para este experimento se puede usar el diseño renglón-columna con un conjunto incompleto de tratamientos en cada columna, mostrado en el cuadro 9.2.

Cuadro 9.2 Diseño de bloques renglón-columna incompleto balanceado de 4×7

Posición	Automóvil						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	3	4	5	6	7	1	2
(2)	5	6	7	1	2	3	4
(3)	6	7	1	2	3	4	5
(4)	7	1	2	3	4	5	6

Los autos se usan como bloques incompletos con cuatro tratamientos evaluados en sus cuatro posiciones y las posiciones son bloques completos ya que cada tratamiento se evalúa en cada posición. Después de una inspección cuidadosa, se observa que cada par de tratamientos ocurre en un automóvil dos veces en algún lugar del experimento y, por lo tanto, el diseño de bloques incompleto está balanceado. El diseño tiene un balance natural respecto a las posiciones porque constituyen bloques completos. Este ejemplo es un caso en el que claramente era necesario un diseño de bloque incompleto porque había un número insuficiente de posiciones disponibles para probar todos los tratamientos a la vez.

Los diseños ortogonales por renglón tienen una réplica completa en cada uno

Dado que el diseño renglón-columna del cuadro 9.2 es un diseño de bloques completos para los renglones e incompleto balanceado para las columnas, el diseño se conoce como un diseño ortogonal por renglón (John, 1987). Como cada tratamiento se presenta en cada renglón, los tratamientos son ortogonales a los renglones.

Youden (1937, 1940) desarrolló arreglos de cuadrados latinos incompletos, conocidos ahora como *cuadrados de Youden*, omitiendo dos o más renglones del diseño de cuadrado latino. Los parámetros del diseño son $t = b$, $r = k$ y $\lambda = k(k - 1)/(t - 1)$. En el apéndice 9A.2 se muestran algunos planes de algunos cuadrados de Youden para experimentos pequeños; otros se pueden encontrar en Cochran y Cox (1957) y Peterson (1985).

El diseño ortogonal por renglones del cuadro 9.2 es un cuadrado de Youden que tiene $r = k = 4$ renglones como réplicas, $b = 7$ columnas y $t = 7$ tratamientos,

con $\lambda = 2$ para los bloques de columna incompletos, las columnas son bloques incompletos de $k = r = 4$ unidades y los renglones son bloques completos que contienen cada uno de los $t = 7$ tratamientos.

La aleatorización en los diseños de renglón-columna se logra de la misma manera descrita para los diseños de cuadrado latino del capítulo 8. Existe una permutación distinta para cada grupo de renglón y columna de tratamientos en los bloques reales, y los tratamientos se asignan al azar a las etiquetas de tratamiento del diseño.

Descripción del análisis de diseños de renglón-columna

El modelo lineal para el diseño renglón-columna es:

$$y_{ijm} = \mu + \tau_i + \rho_j + \gamma_m + e_{ijm} \quad (9.14)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, k \quad m = 1, 2, \dots, b$$

donde μ es la media general, τ_i es el efecto del tratamiento, ρ_j es el efecto del renglón, γ_m es el efecto de la columna y e_{ijm} es el error experimental aleatorio.

Cada renglón contiene una réplica completa de todos los tratamientos, y los tratamientos son ortogonales a los renglones; las columnas también son ortogonales a los renglones y los totales de tratamiento se ajustan sólo para los bloques de columna incompletos para proporcionar estimaciones no sesgadas de medias de tratamiento y una prueba F válida para los efectos de tratamiento. El análisis de varianza intrabloques para los diseños ortogonales por renglón sólo difiere del de diseños de bloques incompletos balanceados en cuanto a la suma de las particiones de sumas de cuadrados para los renglones, el resto de los aspectos del análisis permanece igual.

Tabla 9.5 Análisis intrabloques para un diseño de bloques renglón-columna incompleto balanceado con tratamientos ortogonales por renglones

Fuente de variación	Grados de libertad	Suma de cuadrados
Total	$N - 1$	$\sum(y_{ijm} - \bar{y})^2$
Renglones (réplicas)	$k - 1$	$t\sum(\bar{y}_j - \bar{\bar{y}})^2$
Columnas (no ajustadas)	$b - 1$	$k\sum(\bar{y}_m - \bar{\bar{y}})^2$
Tratamientos (ajustados)*	$t - 1$	$k\sum Q_i^2/\lambda t$
Error	$(t - 1)(k - 1)(b - 1)$	Restando

* $Q_i = y_i - (B_i/k)$ donde B_i es la suma de totales de bloque de aquellas columnas que incluyen el tratamiento i .

9.6 Reducción del tamaño del experimento con diseños parcialmente balanceados (BIPB)

No es posible construir diseños balanceados para todas las situaciones experimentales que requieren bloques incompletos, en algunos casos el número de réplicas necesario puede ser prohibitivo; por tanto, con frecuencia se construyen diseños

parcialmente balanceados que requieren menos réplicas. El número mínimo de réplicas requerido para el diseño balanceado es $r = \lambda(t - 1)/(k - 1)$: suponiendo un experimento con tratamiento $t = 6$, se requieren bloques de tamaño $k = 4$. El diseño balanceado mostrado en el apéndice 9A.1 requiere 10 réplicas o $rt = 60$ unidades experimentales y es muy posible que no se disponga de 60 unidades experimentales o que el costo del experimento con 60 unidades sea demasiado elevado.

Ocurrencias desiguales de pares de tratamientos

Bose y Nair (1939) propusieron el diseño de bloques incompleto parcialmente balanceado, este diseño tiene algunos pares de tratamientos que aparecen en más bloques que otros pares, por lo que algunas comparaciones de tratamientos tendrán mayor precisión que otras. Es más sencillo usar un diseño de bloques balanceado que proporcione la misma precisión para todas las comparaciones entre tratamientos, pero si los recursos están limitados y no se pueden obtener las réplicas suficientes, el diseño parcialmente balanceado es una alternativa atractiva cuando el diseño balanceado requiere un número excesivo de unidades experimentales.

Examinando el diseño de bloques incompleto parcialmente balanceado para seis tratamientos en bloques de cuatro unidades mostrado en el cuadro 9.3, es posible observar que si bien el diseño balanceado del apéndice 9A.1 requiere diez réplicas para obtener un balance con $\lambda = 6$, que el diseño parcialmente balanceado tiene dos réplicas y requiere doce unidades experimentales en tres bloques de cuatro unidades.

Cuadro 9.3 Diseño de bloques incompleto parcialmente balanceado con seis tratamientos en bloques de cuatro unidades

Bloque 1	1	4	2	5
Bloque 2	2	5	3	6
Bloque 3	3	6	1	4

Algunos pares de tratamientos están en dos bloques, mientras que otros pares sólo en uno: los pares de tratamientos (1,4), (2,5) y (3,6) están en dos de los bloques mientras que el resto aparecen en uno solo y por tanto los pares de tratamientos que se presentan juntos en dos bloques se comparan con una precisión un poco mayor que los que están en un solo bloque. Las diferentes precisiones para las comparaciones de tratamiento es el sacrificio pagado por un experimento más pequeño, pero la diferencia en precisión no es tan grande como para evitar el uso de diseño de bloques parcialmente balanceados.

Al aumentar el número de réplicas puede aumentar la precisión de las comparaciones de tratamientos; si las dos réplicas que proporciona el experimento inicial no son suficientes, otra repetición del mismo experimento proporcionará cua-

tro réplicas que, de resultar suficientes, todavía representan ganancia en términos de reducción de costos en comparación con un diseño balanceado completo.

Clases asociadas para las ocurrencias de pares de tratamientos

En un diseño parcialmente balanceado cada tratamiento es miembro de dos o más clases asociadas; una *clase asociada* es un grupo de tratamientos donde cada par de tratamientos ocurre en λ_i bloques, los pares de tratamientos que ocurren en λ_i bloques se conocen como los i -ésimos asociados.

El diseño del cuadro 9.3 tiene dos clases asociadas, los pares de tratamientos (1,4), (2,5) y (3,6) son asociados primeros con $\lambda_1 = 2$, cada par ocurre en dos bloques y cada tratamiento tiene $n_1 = 1$ primeros asociados. El resto de los pares de tratamientos son segundos asociados con $\lambda_2 = 1$, existen $n_2 = 4$ segundos asociados para cada tratamiento; por ejemplo, los segundos asociados del tratamiento 1 son los tratamientos 2, 3, 5 y 6, porque están con el tratamiento 1 en algún bloque del diseño. En el cuadro 9.4 se muestran todos los conjuntos de asociados.

Cuadro 9.4 Primeros y segundos asociados para los seis tratamientos en el diseño de bloques parcialmente balanceado del cuadro 9.3

<i>Tratamiento</i>	<i>Asociados</i>	
	<i>Primeros</i>	<i>Segundos</i>
1	4	2, 3, 5, 6
2	5	1, 3, 4, 6
3	6	1, 2, 4, 5
4	1	2, 3, 5, 6
5	2	1, 3, 4, 6
6	3	1, 2, 4, 5

Un catálogo de algunos grupos principales de diseños parcialmente balanceados con dos clases asociadas se puede encontrar en Bose, Clatworthy y Shrikhande (1954) y en Clatworthy (1973).

Notas sobre el análisis de diseños BIPB

Una ventaja de los diseño balanceados sobre los parcialmente balanceados es que los cálculos manuales del análisis de varianza son un poco más sencillos. Antes del advenimiento de los programas estadísticos modernos, era imperativo disponer de fórmulas de cálculo manual para el análisis de datos, pero ahora es posible calcular las particiones de sumas de cuadrados adecuadas para el análisis de varianza, al igual que estimaciones no sesgadas de las medias de tratamiento y sus errores estándar para la mayoría de los diseños con los programas disponibles. Por lo anterior, el diseño parcialmente balanceado es una alternativa real siempre que existan comparaciones razonablemente precisas de todos los pares de tratamientos.

Particiones de sumas de cuadrados para diseños BIPB

El modelo lineal para el diseño de bloques incompletos parcialmente balanceado es el mismo que para el diseño de bloques incompleto balanceado mostrado en la ecuación (9.1). Como los tratamientos no son ortogonales a los bloques, las particiones ortogonales de sumas de cuadrados de nuevo se derivan ajustando el modelo completo, $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$, para obtener SCE_f y el modelo reducido alternativo sin efectos de tratamiento, $y_{ij} = \mu + \rho_j + e_{ij}$, para obtener SCE_r .

Las sumas de cuadrados para tratamientos ajustados por los bloques se derivan como SC tratamiento(ajustada) = $SCE_r - SCE_f$ y la suma de cuadrados de bloque no ajustada para los tratamientos es la misma que la mostrada en la tabla 9.1 para el diseño balanceado.

Las fórmulas para calcular las estimaciones de mínimos cuadrados de los efectos de tratamiento, las medias de tratamiento y las suma de cuadrados de tratamiento ajustada no son directas como las de los diseño balanceados, ya que existe más de una clase asociada para cada tratamiento y deben hacerse ajustes más complejos a los totales de tratamiento por fortuna, muchos programas estadísticos disponibles pueden realizar los cálculos y no es necesario usar las fórmulas para los tediosos cálculos manuales de un análisis exhaustivo de los datos. Los detalles de los cálculos manuales se pueden encontrar en Cochran y Cox (1957) y los detalles del desarrollo se pueden encontrar en John (1971) y John (1987).

9.7 Eficiencia de los diseños de bloques incompletos

La eficiencia de un diseño con respecto a otro se mide mediante la comparación de varianzas para las estimaciones de las diferencias entre medias de tratamiento en los dos diseños. Por ejemplo, la varianza de la diferencia entre dos medias de tratamiento para el diseño de *bloques completo aleatorizado* (BCA) es $2\sigma_{rcb}^2/r$, ¿Es esa varianza menor que su contraparte del diseño totalmente aleatorizado $2\sigma_{cr}^2/r$? La eficiencia relativa del diseño BCA con respecto al diseño totalmente aleatorizado se puede determinar porque se puede obtener una estimación de σ^2 para este último a partir de los datos en el diseño BCA. Entonces es posible evaluar la efectividad del bloque para reducir la varianza del error experimental.

No existe la misma facilidad si se quiere determinar la eficiencia de un diseño de bloques incompleto balanceado (BIB) con respecto al diseño BCA porque no es posible calcular una estimación de σ^2 para el diseño BCA; sería bueno determinar si un tamaño de bloque más pequeño del diseño BIB daría como resultado una menor varianza del error experimental. Todavía se usa la razón de las varianzas para una diferencia entre dos medias de tratamiento para comparar el diseño BIB con el diseño BCA, la diferencia es que mide sólo la eficiencia *potencial* del diseño BIB porque no es posible estimar σ^2 para el diseño BCA.

Factor de eficiencia para los diseños de bloque incompletos

La varianza de la diferencia entre dos medias de tratamiento en el diseño BIB es $2k\sigma_{bib}^2/\lambda t$. Si este diseño y el BCA tienen el *mismo número de tratamientos y réplicas*, la eficiencia del diseño BIB respecto al diseño BCA es la razón de las varianzas:

$$\text{Eficiencia} = \frac{(2\sigma_{rcb}^2/r)}{(2k\sigma_{bib}^2/\lambda t)} = \frac{\sigma_{rcb}^2}{\sigma_{bib}^2} \cdot \frac{\lambda t}{rk} \quad (9.15)$$

La cantidad $E = \lambda t/rk$ se llama *factor de eficiencia* para el diseño de bloques completo balanceado y proporciona una indicación de la pérdida causada al usar un diseño de bloque incompleto sin reducir σ^2 . El diseño BIB es más preciso en la comparación de dos medias de tratamiento que el diseño BCA si

$$\frac{\sigma_{bib}^2}{\sigma_{rcb}^2} < \frac{\lambda t}{rk} \quad (9.16)$$

El factor de eficiencia es un límite inferior de la eficiencia del diseño BIB con respecto al diseño BCA para un experimento con el mismo número de réplicas y la misma varianza del error σ^2 . Dados un diseño BIB con $t = 4$, $r = 3$, $k = 3$ y $\lambda = 2$ y un diseño BCA con $t = 4$ y $r = 3$, ambos requieren el mismo número de unidades experimentales, pero tienen diferentes arreglos de bloque. El diseño BIB es más preciso que el diseño BCA si:

$$\frac{\sigma_{bib}^2}{\sigma_{rcb}^2} < \frac{2(4)}{3(3)} = 0.89$$

En otras palabras, σ_{bib}^2 tendría que ser cerca de un 11% más pequeño que σ_{rcb}^2 para que el diseño BIB tuviera la misma precisión que el diseño BCA con el mismo número de réplicas.

La intención de un diseño de bloques incompleto es reducir la varianza del error e incrementar la precisión de las comparaciones entre las medias de tratamiento, la meta sería reducir σ^2 para los diseños BIB de manera que se logre la desigualdad de la ecuación (9.16). Un diseño de bloques incompleto exitoso reducirá la varianza del error y σ_{bib}^2 será menor que σ_{rcb}^2 . En el siguiente capítulo se estudiarán algunos elementos que forman parte de una estrategia de bloquización exitosa.

EJERCICIOS PARA EL CAPÍTULO 9

- Un horticultor experimentó con la germinación de semillas de tomate a cuatro temperaturas diferentes (25°C , 30°C , 35°C y 40°C) en un diseño de bloques incompleto balanceado porque sólo disponía de dos cámaras de cultivo para el estudio, cada corrida del experimento fue un bloque que consistía en dos cámaras de cultivo como unidades experimentales ($k = 2$) y se asignaron al azar dos temperaturas a las cámaras para cada corrida. Los siguientes datos son las tasas de germinación de las semillas de tomate:

<i>Corrida</i>	25°C	30°C	35°C	40°C
1	24.65	—	—	1.34
2	—	24.38	—	2.24
3	29.17	21.25	—	—
4	—	—	5.90	1.83
5	28.90	—	18.27	—
6	—	25.53	8.42	—

Fuente: Dr. J. Coons, Department of Botany, Eastern Illinois University.

- ¿Cuántas veces se presentó cada par de tratamientos en los bloques?
 - ¿Cuál es el factor de eficiencia para este diseño?
 - Escriba un modelo lineal para el experimento, explique los términos, calcule el análisis de varianza intrabloques y pruebe la hipótesis nula para los efectos de la temperatura.
 - Calcule las estimaciones de mínimos cuadrados de las medias de la temperatura y sus errores estándar.
 - Calcule el error estándar de la diferencia entre dos estimaciones de mínimos cuadrados de las medias de temperatura.
 - Calcule las particiones de sumas de cuadrados con 1 grado de libertad para la regresión polinomial ortogonal lineal y cuadrática y sus hipótesis nulas.
 - ¿Son significativas las desviaciones de la regresión cuadrática?
- Una compañía está en proceso de desarrollar un compuesto aislante y realiza una prueba acelerada de vida útil para determinar el tiempo transcurrido en minutos hasta la ruptura después de sujetarlo a voltajes elevados. Se probaron cuatro voltajes en una corrida dada de la pruebas, se usó un diseño de bloque incompleto balanceado para probar el compuesto con $t = 7$ voltajes en $b = 7$ corridas de $k = 4$ pruebas y se registraron los minutos hasta la rotura del compuesto para cada prueba.

<i>Bloque</i>	<i>Voltajes (kv)</i>						
	24	28	32	36	40	44	48
1	—	—	38.19	—	5.44	1.96	0.55
2	220.22	—	—	7.66	—	2.54	0.67
3	270.85	200.67	—	—	6.24	—	0.76
4	360.14	170.52	45.43	—	—	3.22	—
5	—	220.12	56.74	9.32	—	—	0.61
6	300.66	—	55.34	10.41	7.19	—	—
7	—	190.78	—	8.74	6.92	2.21	—

- ¿Cuántas veces apareció cada par de tratamientos en los bloques?
- ¿Cuál es el factor de eficiencia para este diseño?
- Escriba un modelo lineal para el experimento, explique los términos y calcule el análisis de varianza intrabloques.
- Calcule los residuales de mínimos cuadrados para cada observación y analícelos según los procedimientos presentados en el capítulo 4. Una transformación típica utilizada para datos de esta naturaleza es el logaritmo natural, ¿sería esto adecuado?

- e. Transforme los datos, calcule las particiones de sumas de cuadrados aditivas y resuma los en una tabla de análisis de varianza. Describa cómo la transformación cambió la naturaleza de los residuales.
- f. Calcule las estimaciones de mínimos cuadrados de las medias del voltaje y sus errores estándar.
- g. Calcule el error estándar de la diferencia entre las estimaciones de mínimos cuadrados de las medias de voltaje.
- h. Calcule la partición de suma de cuadrados con 1 grado de libertad para la regresión lineal del tiempo hasta la ruptura (o una transformación adecuada) con el voltaje y pruebe las hipótesis nulas.
- i. ¿Son significativas las desviaciones de la regresión lineal?
3. Se realizó un estudio para evaluar un método para medir el retraso en cruceros urbanos en siete ciudades, los siete tipos de intersecciones para el estudio se eligieron con base en su geometría y las configuraciones de los semáforos se midieron cuatro de los tipos de crucero en cada ciudad con un diseño de bloques incompleto. Las mediciones fueron realizadas por cuatro observadores, uno en cada crucero durante un periodo específico en las horas pico del tránsito. El diseño se estableció como un cuadrado de Youden con $t = 7$ tratamientos de tipos de crucero, el bloque renglón-columna consistió en $k = 4$ observadores en los renglones y $b = 7$ ciudades como columnas, el orden en que se visitaron las ciudades y la secuencia de tratamientos del tipo de crucero que midió cada observador se asignaron de manera aleatoria según el procedimiento del cuadrado latino. Los tipos de intersección se asignaron al azar a las etiquetas de tratamiento de un cuadrado de Youden básico y se muestran entre paréntesis en la tabla de datos del tiempo en cola (segundo por vehículo):

Observador	Ciudad						
	1	2	3	4	5	6	7
1	(4)	(1)	(3)	(6)	(2)	(7)	(5)
	45.8	66.4	27.0	92.6	32.7	34.6	44.1
2	(2)	(6)	(1)	(4)	(7)	(5)	(3)
	28.6	94.6	64.6	39.9	34.5	45.7	23.7
3	(7)	(4)	(6)	(2)	(5)	(3)	(1)
	32.3	40.7	82.6	29.3	47.6	31.2	74.7
4	(3)	(7)	(2)	(5)	(1)	(6)	(4)
	28.0	31.1	25.5	41.9	68.7	68.1	38.7

- a. Escriba un modelo lineal para este estudio, describa los términos y realice el análisis de varianza intrabloques para este experimento.
- b. Calcule el error estándar de la diferencia entre las medias de mínimos cuadrados de dos intersecciones.
- c. ¿Cuál es el factor de eficiencia para este diseño?
- d. Utilice comparaciones múltiples con el procedimiento del mejor para seleccionar el conjunto de tipos de intersección con el retraso mínimo.
4. (Nota: para este ejercicio se ha supuesto que puede usar un programa para calcular el análisis de varianza, las estimaciones de mínimos cuadrados de las medias de tratamiento con sus errores estándar y los contrastes con sus errores estándar.) Se realizó un estudio con un diseño de bloques incompleto parcialmente balanceado para evaluar $t = 9$ raciones de alimento sobre el balance de nitrógeno en rumiantes el

estudio requirió dispositivos y equipo costosos de digestión y sólo se contaba con tres de ellos, el diseño consistió en $r = 3$ réplicas de 9 tratamientos en $b = 9$ bloques de $k = 3$ animales y había $n_1 = 6$ primeros asociados para cada tratamiento con $\lambda_1 = 1$ y $n_2 = 2$ segundos asociados para cada tratamiento con $\lambda_2 = 0$. Las respuestas observadas se muestran en la tabla con los números de tratamientos entre paréntesis:

<i>Bloque</i>			
1	33.72 (1)	37.80 (2)	42.25 (3)
2	38.58 (1)	45.39 (4)	47.75 (6)
3	34.55 (1)	34.82 (5)	38.29 (7)
4	42.95 (2)	45.35 (4)	48.84 (9)
5	45.12 (2)	36.36 (5)	40.58 (8)
6	43.04 (3)	45.22 (6)	37.26 (8)
7	40.64 (3)	30.49 (7)	36.34 (9)
8	38.53 (4)	34.58 (7)	40.81 (8)
9	36.40 (5)	33.28 (6)	38.46 (9)

Fuente: J. L. Gill (1978), *Design and analysis of experiments in the animal and medical science*, vol 2, Ames Iowa: Iowa State University Press.

- Construya una tabla de primeros y segundo asociados para cada tratamiento.
- Escriba un modelo lineal para el experimento y explique los términos.
- Utilice un programa de computadora y calcule el análisis de varianza para obtener el cuadrado medio ajustado del tratamiento y la varianza del error experimental. Pruebe la hipótesis nula para las razones.
- Calcule las estimaciones de mínimos cuadrados para el experimento y las medias de las raciones con sus errores estándar.
- Calcule el error estándar de la diferencia entre dos estimaciones de mínimos cuadrados de las medias de la ración para los primeros y segundos asociados.
- Con frecuencia se usa un error estándar promedio ponderado para todas las comparaciones entre dos medias de tratamiento en un diseño parcialmente balanceado. Con $n_1 = 6$ primeros asociados y $n_2 = 2$ segundos asociados, la varianza promedio de la diferencia entre dos medias de tratamiento para este estudio es:

$$\text{varianza promedio} = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$$

donde s_1^2 y s_2^2 son las varianzas estimadas de las diferencias entre primeros y segundos asociados respectivamente. Calcule un error estándar ponderado de la diferencia para este experimento.

- Un diseño de bloques incompleto se compone del siguiente arreglo de bloques (1, 2, 3, 4, 5) y tratamientos (A, B, C, D, E).

<i>Bloque</i>	1	(B, C, D, E)
	2	(A, B, D, E)
	3	(A, C, D, E)
	4	(A, B, C, D)
	5	(A, B, C, E)

- a. ¿Cuáles son los parámetros de diseño t , r , k y b ?
 b. Verifique que el diseño esté balanceado.
6. Un diseño de bloques incompleto está constituido por el siguiente arreglo de bloques 1 a 6 y tratamientos (A, B, C, D, E, F).

<i>Bloque</i>	1	(A, B, C)	<i>Bloque</i>	4	(A, B, D)
	2	(C, D, E)		5	(C, D, F)
	3	(B, E, F)		6	(A, E, F)

- a. ¿Cuáles son los parámetros de diseño t , r , k y b ?
 b. ¿Está balanceado el diseño? Explique.
7. Suponga que debe probar 6 combustibles de automóvil y tiene $k = 3$ motores disponibles para las pruebas, por tanto es necesario usar un diseño de bloques incompleto con las corridas de los tres motores como bloques. Encuentre el diseño de bloques incompleto balanceado adecuado en el apéndice 9A.1 para realizar el experimento, asigne aleatoriamente los grupos de tratamiento a los bloques de corridas y de los combustibles a los motores.
8. Debe realizarse un estudio de comparaciones por pares para evaluar lugares para acampar en un parque nacional, existen $t = 10$ diseños de campamento, que van de los primitivos a los de instalaciones completas. Los visitantes al parque nacional se elegirán al azar durante el mes de junio y se les mostrarán fotografías de $k = 2$ de los diez campamentos y ellos indicarán su preferencia por uno de los dos. Construya un diseño de bloques incompleto balanceado para el estudio, ¿cuántos visitantes se requerirán?, ¿cuántas réplicas de cada campamento habrá en el estudio?, ¿cuántas veces será visto un par de fotografías por un visitante?, ¿De qué manera cambiaría el diseño del estudio si se pidiera a cada visitante que mirara $k = 3$ fotografías de campamentos?

Apéndice 9A.1 Algunos diseños de bloques incompletos balanceados

Plan 9A.1 $t = 5, k = 3, r = 6, b = 10, \lambda = 3, E = 0.83$
 (1,2,3), (1,2,5), (1,4,5), (2,3,4), (3,4,5,) reps 1,2,3
 (1,2,4), (1,3,4), (1,3,5), (2,3,5,), (2,4,5) reps 4,5,6

Plan 9A.1 $t = 6, k = 3, r = 5, b = 10, \lambda = 2, E = 0.80$
 (1,2,5), (1,2,6), (1,3,4), (1,3,6), (1,4,5),
 (2,3,4), (2,3,5), (2,4,6), (3,5,6), (4,5,6)

Plan 9A.3 $t = 6, k = 4, r = 10, b = 15, \lambda = 6, E = 0.90$
 (1,2,3,4), (1,4,5,6), (2,3,5,6) reps 1,2
 (1,2,3,5), (1,2,4,6), (3,4,5,6) reps 3,4
 (1,2,3,6), (1,3,4,5), (2,4,5,6) reps 5,6
 (1,2,4,5), (1,3,5,6), (2,3,4,6) reps 7,8
 (1,2,5,6), (1,3,4,6), (2,3,4,5) reps 9,10

Plan 9A.4 $t = 7, k = 3, r = 3, b = 7, \lambda = 1, E = 0.78$
 (1,2,4), (2,3,5), (3,4,6), (4,5,7), (1,5,6),
 (2,6,7), (1,3,7)

Plan 9A.5 $t = 7, k = 4, r = 4, b = 7, \lambda = 2, E = 0.88$
 Sustituya cada bloque en el plan 9A.4 por un bloque que contenga el resto de los tratamientos.

Plan 9A.6 $t = 8, k = 4, r = 7, b = 14, \lambda = 3, E = 0.86$
 (1,2,3,4), (5,6,7,8) rep 1 (1,2,7,8), (3,4,5,6) rep 2
 (1,3,6,8), (2,4,5,7) rep 3 (1,4,6,7), (2,3,5,8) rep 4
 (1,2,5,6), (3,4,7,8) rep 5 (1,3,5,7), (2,4,6,8) rep 6
 (1,4,5,8), (2,3,6,7) rep 7

Plan 9A.7 $t = 9, k = 3, r = 4, b = 12, \lambda = 1, E = 0.75$
 (1,2,3), (1,4,7), (1,5,9), (1,6,8), (2,4,9), (2,5,8)
 (2,6,7), (3,4,8), (3,5,7), (3,6,9), (4,5,6), (7,8,9)

Plan 9A.8 $t = 9, k = 4, r = 8, b = 18, \lambda = 1, E = 0.84$
 (1,2,3,4), (1,3,8,9), (1,4,6,7), (1,5,7,8), (2,3,6,7),
(2,4,5,8), (2,6,8,9), (3,5,7,9), (4,5,6,9), reps 1,2,3,4
(1,2,4,9), (1,2,5,7), (1,3,6,8), (1,5,6,9), (2,3,5,6),
(2,7,8,9), (3,4,5,8), (3,4,7,9), (4,6,7,8), reps 5,6,7,8

Plan 9A.9 $t = 9, k = 5, r = 10, b = 18, \lambda = 5, E = 0.90$
 Sustituya cada bloque en el plan 9A.8 por un bloque que contenga el resto de los tratamientos.

Plan 9A.10 $t = 9, k = 6, r = 8, b = 12, \lambda = 5, E = 0.94$
 Sustituya cada bloque en el plan 9A.7 por un bloque que contenga el resto de los tratamientos.

Plan 9A.11 $t = 10, k = 3, r = 9, b = 30, \lambda = 2, E = 0.74$
 (1,2,3), (1,4,6), (1,7,9), (2,5,8), (2,8,10)
(3,4,7), (3,9,10), (4,6,9), (5,6,10), (5,7,8) reps 1,2,3
(1,2,4), (1,5,7), (1,8,10), (2,3,6), (2,5,9),
(3,4,8), (3,7,10), (4,5,9), (6,7,10), (6,8,9) reps 4,5,6
(1,3,5), (1,6,8), (1,9,10), (2,4,10), (2,6,7),
(2,7,9), (3,5,6), (3,8,9), (4,5,10), (4,7,8) reps 7,8,9

Plan 9A.12 $t = 10, k = 4, r = 6, b = 15, \lambda = 2, E = 0.83$
 (1,2,3,4), (1,2,5,6), (1,3,7,8), (1,4,9,10), (1,5,7,9),
 (1,6,8,10), (2,3,6,9), (2,4,7,10), (2,5,8,10), (2,7,8,9),
 (3,5,9,10), (3,6,7,10), (3,4,5,8), (4,5,6,7), (4,6,8,9)

Plan 9A.13 $t = 10, k = 5, r = 9, b = 18, \lambda = 4, E = 0.89$
 (1,2,3,4,5), (1,2,3,6,7), (1,2,4,6,9), (1,2,5,7,8), (1,3,6,8,9), (1,3,7,8,10),

(1,4,5,6,10), (1,4,8,9,10), (1,5,7,9,10), (2,3,4,8,10), (2,3,5,9,10), (2,4,7,8,9),
 (2,5,6,8,10), (2,6,7,9,10), (3,4,6,7,10), (3,4,5,7,9), (3,5,6,8,9), (4,5,6,7,8)

Plan 9A.14 $t = 10, k = 6, r = 9, b = 15, \lambda = 5, E = 0.93$

Sustituya cada bloque en el plan 9A.12 por un bloque que contenga el resto de los tratamientos.

Plan 9A.15 $t = 11, k = 5, r = 5, b = 11, \lambda = 2, E = 0.88$

(1,2,3,5,8), (1,2,4,7,11), (1,3,6,10,11), (1,4,8,9,10),
 (1,5,6,7,9), (2,3,4,6,9), (2,5,9,10,11), (2,6,7,8,10),
 (3,4,5,7,10), (3,7,8,9,11), (4,5,6,8,11)

Plan 9A.16 $t = 11, k = 6, r = 6, b = 11, \lambda = 3, E = 0.92$

Sustituya cada bloque en el plan 9A.15 por un bloque que contenga el resto de los tratamientos.

Apéndice 9A.2 Algunos diseños de cuadrados latinos incompletos

Tabla 9A Planes para diseños de cuadrados latinos incompletos derivados de cuadrados latinos completos.

<i>t</i>	<i>k</i>	<i>r</i>	<i>b</i>	λ	<i>E</i>	<i>Plan</i>
4	3	3	4	2	.89	*
		3	8	4	.89	**
		9	12	6	.89	**
5	4	4	5	3	.94	*
		8	10	6	.94	**
6	5	5	6	4	.96	*
		10	12	8	.96	**
7	6	6	7	5	.97	*
8	7	7	8	6	.98	*
9	8	8	9	7	.98	*
10	9	9	10	8	.99	*
11	10	10	11	9	.99	*

*Construido a partir de un cuadrado latino de $t \times t$ por omisión de la última columna

**Por repetición del plan para $r = t - 1$, que se construyó con un cuadrado latino de $t \times t$ sin la última columna

Planes para otros diseños de cuadrados latinos incompletos:

Plan 9B.1 $t = 5, k = 2, r = 4, b = 10, \lambda = 1, E = 0.63$

<i>Reps I y II</i>					<i>Reps III y IV</i>				
1	2	3	4	5	1	2	3	4	5
2	5	4	1	3	3	4	2	5	1

Plan 9B.2 $t = 5, k = 3, r = 6, b = 10, \lambda = 3, E = 0.83$

Reps I,II, y III					Reps IV,V, y VI				
1	2	3	4	5	1	2	3	4	5
2	1	4	5	3	2	3	4	5	1
3	5	2	1	4	4	5	1	2	3

Plan 9B.3 $t = 7, k = 2, r = 6, b = 21, \lambda = 1, E = 0.58$

Reps I y II							Reps III y IV							Reps V y VI						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
2	6	4	7	1	5	3	3	4	5	6	7	1	2	4	3	6	5	2	7	1

Plan 9B.4 $t = 7, k = 3, r = 3, b = 7, \lambda = 1, E = 0.78$

7	1	2	3	4	5	6
1	2	3	4	5	6	7
3	4	5	6	7	1	2

Plan 9B.5 $t = 7, k = 4, r = 4, b = 7, \lambda = 2, E = 0.88$

3	4	5	6	7	1	2
5	6	7	1	2	3	4
6	7	1	2	3	4	5
7	1	2	3	4	5	6

Plan 9B.6 $t = 9, k = 2, r = 8, b = 36, \lambda = 1, E = 0.56$

Reps I y II									Reps V y VI								
1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
2	8	4	7	6	1	3	9	5	4	6	2	5	7	8	9	1	3

Reps III y IV									Reps VII y VIII								
1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
3	5	6	9	8	7	1	4	2	5	4	8	6	3	9	2	7	1

Plan 9B.7 $t = 9, k = 4, r = 8, b = 18, \lambda = 3, E = 0.84$

Reps I, II, III y IV									Reps V, VI, VII y VIII								
1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
4	6	8	1	7	9	3	2	5	2	3	4	9	1	8	6	5	7
6	8	9	3	1	4	2	5	7	5	6	7	2	9	1	4	3	8
7	9	1	2	8	5	6	4	3	7	5	9	1	6	3	8	4	2

Plan 9B.8 $t = 9, k = 5, r = 10, b = 18, \lambda = 5, E = 0.90$

Reps I, II, III y IV									Reps V, VI, VII y VIII								
1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
2	6	8	3	1	4	9	5	7	2	6	5	3	7	8	4	9	1
3	8	5	9	7	2	1	4	6	3	5	1	2	9	7	8	4	6
7	4	9	2	3	5	6	1	8	5	1	4	8	2	3	9	6	7
8	1	2	6	4	7	3	9	5	9	8	6	7	4	5	1	3	2

Plan 9B.9 $t = 10, k = 3, r = 9, b = 30, \lambda = 2, E = 0.74$

Reps I, II, y III									Reps IV, V, y VI									Reps VII, VIII y IX										
1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	
2	5	7	1	8	4	9	10	3	2	5	7	8	10	1	2	10	5	6	3	4	6	7	8	9	10	11	12	
3	8	4	6	7	9	1	2	10	4	6	8	5	1	9	3	10	2	7	5	6	7	8	9	10	11	12	13	
4	6	8	5	1	9	3	10	2	4	6	8	2	7	1	3	10	2	7	1	5	6	7	8	9	10	11	12	13
5	6	9	10	3	8	2	7	1	5	7	8	4	10	1	2	11	4	10	5	6	7	8	9	10	11	12	13	14

Plan 9B.10 $t = 11, k = 2, r = 10, b = 55, \lambda = 1, E = 0.55$

Reps I, y II											Reps III, y IV											Reps V y VI											Reps VII y VIII											Reps IX y X										
1	2	3	4	5	6	7	8	9	10	11	2	11	10	5	6	7	1	3	4	9	10	11	1	2	3	4	5	6	7	8	9	10	11	12																				
3	6	5	10	9	8	2	1	7	11	4	5	6	10	1	2	11	7	11	4	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12																				
4	3	7	6	10	9	11	2	1	3	12	1	2	11	10	9	8	7	1	3	4	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12																			
5	9	6	2	7	10	8	4	11	1	3	12	1	2	11	10	9	8	7	1	3	4	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12																		
6	5	4	7	8	11	10	9	3	2	1	12	1	2	11	10	9	8	7	1	3	4	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12																		

Plan 9B.11 $t = 11, k = 5, r = 5, b = 11, \lambda = 2, E = 0.88$

1	7	9	11	10	8	2	6	3	5	4
2	1	8	9	11	7	6	3	4	10	5
3	6	1	7	5	2	4	11	10	9	8
4	10	6	1	8	3	11	5	9	2	7
5	3	2	4	1	11	10	9	8	7	6

Plan 9B.12 $t = 11, k = 6, r = 6, b = 11, \lambda = 3, E = 0.92$

6	5	4	3	2	1	9	8	7	11	10
7	8	5	10	3	6	1	2	11	4	9
8	4	7	2	9	10	3	1	5	6	11
9	11	3	6	7	4	5	10	1	8	2
10	2	11	5	4	9	8	7	6	1	3
11	9	10	8	6	5	7	4	2	3	1

Plan 9B.13 $t = 13, k = 3, r = 6, b = 26, \lambda = 1, E = 0.72$

<i>Reps I, II, y III</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
	3	4	5	6	7	8	9	10	11	12	13	1	2
	9	10	11	12	13	1	2	3	4	5	6	7	8

<i>Reps IV, V, y VI</i>	2	3	4	5	6	7	8	9	10	11	12	13	1
	6	7	8	9	10	11	12	13	1	2	3	4	5
	5	6	7	8	9	10	11	12	13	1	2	3	4

Plan 9B.14 $t = 13, k = 4, r = 4, b = 13, \lambda = 1, E = 0.81$

13	1	2	3	4	5	6	7	8	9	10	11	12
1	2	3	4	5	6	7	8	9	10	11	12	13
3	4	5	6	7	8	9	10	11	12	13	1	2
9	10	11	12	13	1	2	3	4	5	6	7	8

Plan 9B.15 $t = 13, k = 9, r = 9, b = 13, \lambda = 6, E = 0.96$

2	3	4	5	6	7	8	9	10	11	12	13	1
5	6	7	8	9	10	11	12	13	1	2	3	4
6	7	8	9	10	11	12	13	1	2	3	4	5
7	8	9	10	11	12	13	1	2	3	4	5	6
9	10	11	12	13	1	2	3	4	5	6	7	8
10	11	12	13	1	2	3	4	5	6	7	8	9
11	12	13	1	2	3	4	5	6	7	8	9	10
12	13	1	2	3	4	5	6	7	8	9	10	11
13	1	2	3	4	5	6	7	8	9	10	11	12

Plan 9B.16 $t = 15, k = 7, r = 7, b = 15, \lambda = 3, E = 0.92$

13	5	15	12	4	11	1	2	8	10	9	14	7	3	6
8	14	12	11	5	9	2	3	6	4	13	7	15	1	10
12	10	11	6	8	7	3	1	4	5	14	13	9	2	15
6	7	5	9	1	2	4	13	15	11	10	3	12	8	14
7	12	8	2	14	13	5	15	10	1	6	4	3	9	11
1	2	3	4	9	15	6	14	13	12	5	8	10	11	7
9	8	6	14	15	5	7	12	2	13	3	11	4	10	1

Plan 9B.17 $t = 15, k = 8, r = 8, b = 15, \lambda = 4, E = 0.94$

11	4	9	15	7	6	12	10	5	8	1	2	13	14	3
4	1	2	3	13	10	9	11	14	15	7	6	8	5	12
2	3	14	1	10	12	13	9	7	6	11	15	5	4	8
5	15	4	10	12	1	14	7	3	2	8	9	11	6	13
10	13	7	8	11	4	15	6	1	3	2	5	14	12	9
3	11	1	13	2	14	10	8	9	7	4	12	6	15	5
14	6	10	7	3	8	11	5	12	9	15	1	2	13	4
15	9	13	5	6	3	8	4	11	14	12	10	1	7	2

Apéndice 9A.3 Estimaciones de mínimos cuadrados para diseño BIB

El modelo lineal para el diseño de bloque incompleto balanceado es

$$y_{ij} = \mu + \tau_i + \rho_j + e_{ij} \quad (9A.1)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, b$$

donde μ es la media general, τ_i es el efecto de tratamiento, ρ_j es el efecto de bloque y las e_{ij} son errores experimentales aleatorios independientes con media 0 y varianza σ^2 .

Sea n_{ij} una variable indicadora con $n_{ij} = 1$ si el i -ésimo tratamiento está en el j -ésimo bloque y $n_{ij} = 0$ de otra manera.

Las estimaciones de mínimos cuadrados para los parámetros en el modelo completo minimizan la suma de cuadrados para el error experimental:

$$Q = \sum_{i=1}^t \sum_{j=1}^b n_{ij} e_{ij}^2 = \sum_{i=1}^t \sum_{j=1}^b n_{ij} (y_{ij} - \mu - \tau_i - \rho_j)^2 \quad (9A.2)$$

Las ecuaciones normales para el diseño de bloques incompleto balanceado son:

$$\mu: N\hat{\mu} + r \sum_{i=1}^t \hat{\tau}_i + k \sum_{j=1}^b \hat{\rho}_j = y \quad (9A.3)$$

$$\tau_i: r\hat{\mu} + r\hat{\tau}_i + \sum_{j=1}^b n_{ij}\hat{\rho}_j = y_i \quad i = 1, 2, \dots, t \quad (9A.4)$$

$$\rho_j: k\hat{\mu} + \sum_{i=1}^t n_{ij}\hat{\tau}_i + k\hat{\rho}_j = y_j \quad j = 1, 2, \dots, b \quad (9A.5)$$

Cada una de las sumas del conjunto de t ecuaciones para las τ_i y del conjunto de b ecuaciones para las ρ_j es igual a la ecuación para μ , esto tiene como resultado dependencias lineales entre las ecuaciones. Las restricciones $\sum_i \hat{\tau}_i = 0$ y $\sum_j \hat{\rho}_j = 0$ se pueden usar para proporcionar una solución única.

Al imponer las restricciones, la estimación de μ es $\hat{\mu} = y/N$ a partir de la ecuación (9A.3). Las ecuaciones para ρ_j , ecuaciones (9A.5), se usan para eliminar las $\hat{\rho}_j$ de las ecuaciones para τ_i en la ecuación 9A.4. Despues de eliminar las ecuaciones con $\hat{\tau}_i$, son:

$$rk\hat{\tau}_i - r\hat{\tau}_i - \sum_{j=1}^b \sum_{p=1, p \neq i}^t n_{ij}n_{pj}\hat{\tau}_p = ky_i - \sum_{j=1}^b n_{ij}y_j \quad (9A.6)$$

El lado derecho de la ecuación (9A.6) es kQ_i , donde Q_i es el total del tratamiento ajustado que se usa para calcular la suma de cuadrados de tratamiento ajustada. Como:

$$\sum_{j=1}^b n_{ij}n_{pj} = \lambda, \quad \text{para } p \neq i \quad (9A.7)$$

y $n_{pj}^2 = n_{pj}$ ($n_{pj} = 0$ o 1), la ecuación (9A.6) se puede escribir como:

$$r(k-1)\hat{\tau}_i - \lambda \sum_{p=1, p \neq i}^t \hat{\tau}_p = kQ_i \quad (9A.8)$$

Como se impuso la condición $\sum_i \widehat{\tau}_i$ sobre la solución, la sustitución:

$$-\widehat{\tau}_i = \sum_{\substack{p=1 \\ p \neq i}}^t \widehat{\tau}_p$$

se puede convertir en la ecuación (9A.8). Con la igualdad $\lambda(t-1) = r(k-1)$, la ecuación para $\widehat{\tau}_i$ es:

$$\begin{aligned} \lambda t \widehat{\tau}_i &= k Q_i & (9A.9) \\ \tau_i &= \frac{k Q_i}{\lambda_t} \quad i = 1, 2, \dots, t \end{aligned}$$

10 Diseño de bloques incompletos: Diseños resolubles y cíclicos

En el capítulo 9 se presenta la descripción general y el análisis de los diseños de bloques incompletos. En este capítulo se explican varias importantes clases de diseños de bloques incompletos, incluyendo los diseños resolubles con bloques agrupados en réplicas completas de tratamientos. También se presentan los diseños cíclicos que se pueden construir sin usar extensos planes sistemáticos y los diseños α , que amplían el número de diseños resolubles aprovechables para los experimentos.

10.1 Diseños resolubles para ayudar a manejar el experimento

Los **diseños resolubles** tienen bloques agrupados de manera que cada grupo constituye una réplica completa de los tratamientos, el agrupamiento en réplicas completas es eficaz para administrar un experimento.

Una de las primeras aplicaciones de los diseños resolubles fue con pruebas de cultivo con plantas colocadas en las parcelas de una granja experimental, los labradores querían probar un gran número de líneas genéticas y hacer todas las comparaciones por pares entre ellas con la misma precisión. Fue necesario ordenar las parcelas en bloques más pequeños que la réplica completa para reducir la varianza del error experimental aún más de lo que es posible con los diseños de bloques completos. El diseño de bloques incompleto resoluble era atractivo porque además de reducir el tamaño de los bloques para obtener mayor precisión, también permitía al investigador controlar estos grandes estudios en campo basándose en las réplicas.

Los diseños resolubles son útiles en la práctica cuando no es posible realizar el experimento completo al mismo tiempo, pues los experimentos con diseño resoluble pueden llevarse a cabo en etapas, con una o más réplicas terminadas en cada

etapa. Además, si por alguna razón el experimento termina antes de tiempo, habrá un número igual de réplicas en todos los tratamientos.

Los diseños resolubles son arreglos en r grupos de réplicas de s bloques con k unidades por bloque. En los diseños de bloques incompletos balanceados resolubles, el número de tratamientos es un múltiplo del número de unidades por bloque, $t = sk$, y el número total de bloques satisface la relación $b = rs \geq t + r - 1$.

Ejemplo 10.1 Un diseño resoluble para pruebas con productos alimenticios

Examinemos un experimento para probar nueve variedades de un producto alimenticio en el mismo día con sujetos humanos como jueces. Varios factores asociados con el experimento hacen que un diseño de bloques incompleto sea atractivo, se puede esperar que un juez discrimine de manera adecuada hasta cuatro o cinco muestras de productos en cualquier momento, pero por problemas de programación, no es posible tener un número suficiente de jueces disponibles el mismo día para las réplicas de los tratamientos; por lo tanto, deben prepararse de nuevo los alimentos cada día de pruebas y es necesario garantizar que la variación en la preparación de un día a otro no interfiera con las comparaciones de tratamientos en el experimento.

El cuadro 10.1 muestra una prueba diseñada para tres jueces en la que cada uno evaluará tres variedades del producto en un día dado. Se programan tres jueces durante cuatro días de pruebas, ellos constituyen un bloque incompleto de tres tratamientos y cada uno se asigna al azar a uno de los bloques de tres tratamientos y prueba tres productos presentados al azar para su evaluación.

El diseño es resoluble con una réplica completa del experimento realizado un día dado y está balanceado, pues cada tratamiento ocurre una vez en el mismo bloque con cada uno de los otros tratamientos en el experimento. Los jueces sólo evalúan tres productos a la vez, lo que reduce considerablemente la variación dentro del bloque en comparación con un juez que evalúa nueve productos a la vez.

Cuadro 10.1 Diseño de bloques incompleto balanceado resoluble para evaluar productos alimenticios

	<i>Día I</i>	<i>Día II</i>	<i>Día III</i>	<i>Día IV</i>	
<i>Juez</i>	<i>Juez</i>	<i>Juez</i>	<i>Juez</i>		
1	(1, 2, 3)	4	(1, 4, 7)	7	(1, 5, 9)
2	(4, 5, 6)	5	(2, 5, 8)	8	(2, 6, 7)
3	(7, 8, 9)	6	(3, 6, 9)	9	(3, 4, 8)
				10 (1, 6, 8)	
				11 (2, 4, 9)	
				12 (3, 5, 7)	

Diseños de retícula balanceados

Los **diseños de retícula (lattice) balanceados** son un grupo muy conocido de diseños resolubles propuestos por Yates (1936b). El número disponible de estos cuadrados de retícula está limitado debido a que el número de tratamientos debe ser un cuadrado exacto $t = k^2$, los diseños requieren $r = (k + 1)$ réplicas y $b = k(k + 1)$ bloques para un balance completo con $\lambda = 1$, cada grupo de réplicas contiene $s = k$ bloques de k unidades experimentales cada uno y el número k de unidades por bloque debe ser un número primo o una potencia de un número primo. Se pueden encontrar planes para diseños de retícula balanceados con $t = 9, 16, 25, 49, 64$ y 81 en Cochran y Cox (1957) o Peterson (1985).

El diseño del cuadro 10.1 es una retícula balanceada con nueve tratamientos en bloques de tres unidades que tiene cuatro grupos de réplicas con un total de 12 bloques. Observe que cada par de tratamientos ocurre en un bloque $\lambda = 1$ vez en alguna parte del diseño balanceado.

El balance completo requiere $(k + 1)$ réplicas con los diseños de retícula y, como son diseños resolubles, uno o más grupos de réplicas se pueden eliminar para producir un diseño de retícula parcialmente balanceado. Los diseños con $r = 2, 3$ o 4 grupos de réplicas respectivos se conocen como retículas *simples*, *triples* o *cuádruples*.

El factor de eficiencia promedio para un diseño de retícula balanceado es:

$$E = \frac{(k + 1)(r - 1)}{r(k + 2) + (k + 1)}$$

y para las retículas simple y triple, los factores de eficiencia promedio respectivos son $(k + 1)/(k + 3)$ y $(2k + 2)/(2k + 5)$.

Diseños de retícula rectangulares

La restricción en el número de tratamientos o variedades de las retículas cuadradas es importante. Los **diseños de retícula rectangulares** desarrollados por Harshbarger (1949) proporcionaron diseños resolubles con números de tratamientos intermedios a los proporcionados por los diseños de retícula cuadrados.

Los diseños de retícula rectangulares son arreglos de $t = s(s - 1)$ tratamientos en bloques con $k = (s - 1)$ unidades y los números de tratamientos están cerca del punto medio entre los que proporcionan las retículas cuadradas. Cochran y Cox (1957) construyeron planes para retículas rectangulares con $t = 12, 20, 30, 42, 56, 72$ y 90 tratamientos. Los diseños con $r = 2$ y $r = 3$ réplicas se conocen como retículas rectangulares *simple* y *triple*, respectivamente.

Existen otros diseños de bloques incompletos balanceados resolubles sin las restricciones de las retículas cuadradas o rectangulares; sin embargo, son para un conjunto limitado de tamaños de bloque y números de tratamientos. Algunos diseños para $t \leq 11$ tratamientos se encuentran en el apéndice 9A.1, otros diseños resolubles para $t > 11$ están dados por Cochran y Cox (1957).

Descripción del análisis de diseños resolubles

El modelo lineal para el diseño resoluble es:

$$y_{ijm} = \mu + \tau_i + \gamma_j + \rho_{m(j)} + e_{ijm} \quad (10.1)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad m = 1, 2, \dots, s$$

donde μ es la media general, τ_i es el efecto del tratamiento, γ_j es el efecto del grupo de réplicas, $\rho_{m(j)}$ es el efecto del bloque anidado dentro de la réplica y e_{ijm} es el error experimental aleatorio.

El ajuste secuencial de modelos alternativos descrito en el capítulo 9 proporciona una partición de suma de cuadrados ortogonal, pero en este caso se requiere una partición de suma de cuadrados adicional para los grupos de réplicas. Los bloques están anidados dentro de los grupos de réplicas y las sumas de cuadrados para los bloques de los grupos de réplicas se unen como $SC(\text{bloques/réplicas})$. La tabla 10.1 contiene una descripción del análisis de varianza intrabloques.

Tabla 10.1 Análisis intrabloques para un diseño de bloques incompleto balanceado resoluble

Fuente de variación	Grados de libertad	Sumas de cuadrados
Total	$N - 1$	$\sum (y_{ijm} - \bar{y})^2$
Réplicas	$r - 1$	$sk \sum (\bar{y}_j - \bar{y})^2$
Bloques (no ajustados) dentro de réplicas	$r(s - 1)$	$k \sum (\bar{y}_{jm} - \bar{y})^2$
Tratamientos (ajustados)	$t - 1$	$k \sum Q_t^2$
		λ_t
Error	$N - t - rs + 1$	Restando

$Q_t = y_t - (B_t/k)$, y B_t es la suma de totales de bloques que incluyen el i -ésimo tratamiento.

Los tratamientos son ortogonales a los grupos de réplicas ya que cada tratamiento aparece una vez en cada grupo de réplica y los totales de tratamiento ajustados sólo por los efectos de bloque proporcionan estimaciones no sesgadas de las medias de tratamiento y una prueba F válida para los efectos de tratamiento.

10.2 Diseños resolubles renglón-columna para dos criterios de bloque

Algunas situaciones de investigación requieren dos criterios de bloque y no permiten bloques completos de tratamientos para los renglones o las columnas necesarios para los cuadrados ortogonales de Youden. Los diseños renglón-columna tienen t tratamientos colocados en bloques de $k = pq$ unidades y los tratamientos de cada bloque se arreglan en un diseño renglón-columna de $p \times q$.

Existe un número limitado de diseños de bloques incompletos renglón-columna balanceados para los experimentos que requieren que los tamaños de bloque para renglones y columnas sean menores que el número de tratamientos, estos diseños necesitan un gran número de bloques para cumplir con este requerimiento de balance. Vea en John (1971) una bibliografía amplia al respecto.

Ejemplo 10.2 Diseño renglón-columna anidado para muestras de población de insectos

Examinemos un estudio epidemiológico sobre la transmisión de un parásito micróbico a los humanos a través de un insecto en un área agrícola tropical. El parásito se transmite a los humanos mediante la picadura del insecto, pero si el insecto no está infectado por el parásito, puede adquirirlo cuando pica a un humano infectado. Un proyecto de salud pública es la planeación de pruebas de efectividad de una droga que se sabe que varía el ciclo de vida del parásito en los humanos, un entomólogo que trabaja en el proyecto piensa obtener muestras de las poblaciones de insectos para supervisar el efecto que el programa tiene en la reducción de la tasa de infección.

Se programaron nueve tipos de hábitats dentro y cerca de las cuatro plantaciones agrícolas para supervisar el proyecto, los entornos son lugares como corrientes de ríos, pueblos cercanos a las plantaciones, campos agrícolas, aguas estancadas, etcétera, y el entomólogo sólo recolectará muestras de insectos en tres de ellos en un día. Se estableció un diseño de muestreo renglón-columna para controlar la variación potencial causada por el muestreo en distintos días y horas del día en las cuatro plantaciones, el diseño se muestra en el resumen 10.2 tiene nueve tratamientos de entorno con muestreo en un diseño renglón-columna de 3×3 en cuatro plantaciones. Los tres renglones del arreglo de 3×3 corresponden a la hora del día en que se toman las muestras, y las tres columnas se refieren a los tres días de recolección en cada plantación.

Por inspección de cada par de tratamientos o entornos se puede verificar que en el diseño se toma una muestra a la vez de cada uno en el mismo día, cada par de tratamientos debe aparecer una vez en el mismo renglón y una vez en la misma columna en alguna parte tener un diseño renglón-columna balanceado. El diseño se puede resolver con las plantaciones como réplicas y una réplica completa de los nueve tipos de hábitat en cada plantación; es un diseño anidado en donde los renglones y las columnas se anidan dentro de los grupos de réplicas.

Diseños de retículas cuadradas balanceadas

Estos diseños, también propuestos por Yates (1940b), son diseños renglón-columna anidados resolubles, que están entre los primeros desarrollos de diseño de bloques incompletos; el ejemplo del cuadro 10.2 es uno de ellos. La **retícula cuadrada balanceada** tiene $t = k^2$ tratamientos dispuestos en un diseño renglón-columna de $k \times k$, un arreglo de $k \times k$ consiste en una de las réplicas completas de tratamientos con $p = q = k$. El diseño balanceado requiere $r = (k + 1)$ réplicas para el balance sea completo, de manera que cada par de tratamientos se encuentre una vez en el mismo renglón y una vez en la misma columna.

Cuadro 10.2 Diseño de bloques incompleto renglón columna anidado de 3×3 para las muestras de la población de insectos

Hora	Día			Hora	Día		
	(1)	(2)	(3)		(4)	(5)	(6)
(1)	1	2	3	(4)	1	4	7
(2)	4	5	6	(5)	2	5	8
(3)	7	8	9	(6)	3	6	9
	Plantación I				Plantación II		
Hora	Día			Hora	Día		
	(7)	(8)	(9)		(10)	(11)	(12)
(7)	1	6	8	(10)	1	9	5
(8)	9	2	4	(11)	6	2	7
(9)	5	7	3	(12)	8	4	3
	Plantación III				Plantación IV		

en algún momento del experimento. Cuando k es impar, un valor de repetición de $r = \frac{1}{2}(k + 1)$ proporcionará un semibalance tal que cualquier par de tratamientos esté una vez en el mismo renglón o una vez en la misma columna.

En Cochran y Cox (1957) y en Peterson (1985) se pueden encontrar planos para diseños cuadrados de retícula con $t = 9, 16, 25, 49, 64, 81, 121$ y 169. Una presentación más amplia de estos diseños se encuentra en Kempthorne (1952) y Federer (1955).

Descripción del análisis para diseños renglón-columna anidados

El modelo lineal para cualquier diseño renglón-columna anidado resoluble es:

$$y_{ijlm} = \mu + \beta_m + \rho_{j(m)} + \gamma_{l(m)} + \tau_i + e_{ijlm} \quad (10.2)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, p \quad l = 1, 2, \dots, q \quad m = 1, 2, \dots, r$$

donde μ es la media general, τ_i es el efecto del tratamiento, β_m es el efecto del grupo de réplica, $\rho_{j(m)}$ y $\gamma_{l(m)}$ son el renglón y la columna respectivos anidados en los efectos de grupos de réplicas y e_{ijlm} es el error experimental aleatorio.

Cada tratamiento está en cada grupo de réplicas y son ortogonales a ellos, los renglones y columnas también son ortogonales a los grupos de réplicas. Las particiones de sumas de cuadrados para las réplicas, los renglones dentro de las réplicas y las columnas dentro de las réplicas se pueden calcular de manera usual, como se muestra en el análisis de varianza descrito en la tabla 10.2. Los tratamientos no son ortogonales a los renglones o columnas de los grupos de réplicas y deben ajustarse por los efectos de renglón y de columna.

Los totales de tratamiento ajustado necesarios para las estimaciones de mínimos cuadrados de los efectos de tratamientos y las sumas de cuadrados de tratamiento ajustado son:

Tabla 10.2 Análisis intrabloques para un diseño de bloques incompleto renglón-columna anidado

Fuente de variación	Grados de libertad	Sumas de cuadrados
Total	$N - 1$	$\sum (y_{jlm} - \bar{y})^2$
Grupos (réplicas)	$r - 1$	$t \sum (\bar{y}_m - \bar{y})^2$
Columnas (no ajustadas) dentro de grupos	$r(q - 1)$	$p \sum (\bar{y}_{lm} - \bar{y}_m)^2$
Renglones (no ajustados) dentro de grupos	$r(p - 1)$	$q \sum (\bar{y}_{jm} - \bar{y}_m)^2$
Tratamientos (ajustados)	$t - 1$	$\frac{pq(t - 1)}{rt(p - 1)(q - 1)} \sum Q_i^2$
Error	$r(p - 1)(q - 1) - (t - 1)$	Restando

$$Q_i = y_i + r\bar{y} - \frac{1}{p} R_i - \frac{1}{q} C_i \quad (10.3)$$

donde y_i es el total para el i -ésimo tratamiento \bar{y} es la gran media para el experimento, R_i es la suma de totales de renglones que incluyen al i -ésimo tratamiento y C_i es la suma de totales de columnas que incluyen al i -ésimo tratamiento.

La estimación de los mínimos cuadrados del efecto de tratamiento es :

$$\hat{\tau}_i = \frac{pq(t - 1)Q_i}{rt(p - 1)(q - 1)} \quad (10.4)$$

y la estimación de los mínimos cuadrados de la media de tratamiento es $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i$, donde $\hat{\mu} = \bar{y}$ es la estimación de la media general.

El factor de eficiencia para el diseño renglón-columna anidado es:

$$E = \frac{t(p - 1)(q - 1)}{pq(t - 1)} \quad (10.5)$$

donde el factor de eficiencia para la retícula cuadrada balanceada es $E = (k - 1)/(k + 1)$, con $p = q = k$ y $t = k^2$; por ejemplo, con $k = 3$ del cuadro 10.2, el factor de eficiencia es $E = 2/4 = .50$. La retícula cuadrada balanceada tendría que reducir la varianza del error experimental en un 50% para estimar las comparaciones entre las medias de tratamiento de manera tan precisa como un diseño de bloques completos del mismo tamaño.

10.3 Los diseños cílicos simplifican la construcción del diseño

Los diseños de bloques incompletos complejos pueden añadir más complicaciones al proceso si hay necesidad de hacer consultar con frecuencia la distribución del diseño, por lo que debe contarse con tablas amplias de planes de diseño para usar los diseños de bloques incompletos que se han presentado hasta ahora, establecer

un diseño en la instalación física necesaria para el experimento exige una vigilancia constante, debe tenerse cuidado de evitar errores cuando se asignan los tratamientos a las unidades experimentales y al registrar los datos.

Los diseños cílicos ofrecen cierto grado de simplificación en la construcción del diseño y en el establecimiento del proceso experimental y como se generan a partir de un bloque inicial de tratamientos, es necesario un esquema completo de los planes de distribución del experimento. Una vez que se conoce el bloque inicial, se genera el plan del experimento y se asignan al azar los tratamientos a las etiquetas numéricas en el diseño.

Cómo construir un diseño cílico

La designación *cílica* se refiere al método usado para construir los diseños, al igual que al esquema de asociación entre tratamientos. Los tratamientos se asignan a los bloques a través de una sustitución cílica de etiquetas de tratamiento de un bloque generador inicial.

El método requiere etiquetar los tratamientos como $(0, 1, 2, \dots, t - 1)$, el ciclo comienza con un bloque de tamaño k requerido por el experimento y los bloques subsecuentes se obtienen sumando 1 a la etiqueta de tratamiento del bloque anterior, si una etiqueta de tratamiento excede $(t - 1)$, se reduce mediante el módulo t ; el ciclo continúa hasta que regresa a la configuración del bloque inicial.

Considerando el diseño para seis tratamientos en bloques de tres unidades generadas por sustitución cílica del cuadro 10.3, se etiquetan los seis tratamientos $(0, 1, 2, 3, 4, 5)$ para el diseño, que se construye con el bloque inicial $(0, 1, 3)$; el segundo bloque se construye sumando 1 a cada etiqueta de tratamiento del primer bloque, resultando en $(1, 2, 4)$ como configuración del segundo bloque; se suma 1 a cada etiqueta del segundo bloque y se obtiene $(2, 3, 5)$ como los tratamientos para el tercer bloque.

Cuadro 10.3 Diseño cílico para seis tratamientos en bloques de tres unidades

Bloque

1	$(0, 1, 3)$ <i>bloque inicial</i>
2	$(1, 2, 4)$
3	$(2, 3, 5)$
4	$(3, 4, 0)$
5	$(4, 5, 1)$
6	$(5, 0, 2)$

La transición del bloque 3 al bloque 4 en el ciclo requiere reducir una etiqueta en el bloque 4 mediante el módulo $t = 6$, porque si se suma 1 a cada etiqueta en el bloque 3, la configuración del bloque 4 será $(3, 4, 6)$ donde $5 + 1 = 6$ excede $(t - 1) = 5$; por lo tanto, 6 se debe reducir según el módulo 6, que es $6 \equiv 0 \pmod{6}$,

y 0 sustituye a 6 en la configuración del bloque 4, es decir, (3, 4, 0); en la trasición del bloque 5 al 6 se realiza una sustitución similar.

Dentro del ciclo completo mostrado en el cuadro, se construyó un conjunto completo de $b = 6$ bloques, los números forman un ciclo completo de etiquetas de tratamiento hacia abajo de las columnas, las etiquetas de tratamiento en la primera posición del bloque tienen un ciclo de 0 a 5 mientras que en la segunda posición el ciclo comienza con la etiqueta 1 y la tercera posición con la etiqueta 3, entonces el diseño está completo con los bloques 1 a 6 pues el séptimo miembro del ciclo sería la configuración del bloque inicial (0, 1, 3).

Construcción de diseños renglón-columna con el método cíclico

Los diseños cílicos se pueden usar para diseños renglón-columna si el experimento requiere dos criterios de bloqueo, el único requisito es que el número de réplicas sea un múltiplo del tamaño de bloque, o sea, $r = ik$, donde i es un valor entero. El diseño del cuadro 10.3, con, $t = 6$ y $r = k = 3$ generado a partir del bloque inicial (0, 1, 3), se puede usar como diseño renglón-columna con seis renglones y tres columnas.

Se trata de un diseño ortogonal por columna, ya que cada tratamiento aparece en cada columna; si se mantiene la integridad del renglón y la columna al aleatorizar, entonces la suma de cuadrados para columnas y renglones es una partición de la suma de cuadrados total y en el análisis los tratamientos se ajustan por los bloques de los renglones.

Tablas para construir diseños cílicos

En la tabla 10A.1 del apéndice se puede encontrar la tabla compacta de bloques iniciales necesaria para diseños cílicos de relativa eficiencia con $4 \leq t \leq 15$ tratamientos, tomada de John (1987), quién también elaboró otra tabla para $16 \leq t \leq 30$ tratamientos; pueden encontrarse más tablas en John (1966).

Para un tamaño de bloque k , los primeros $k - 1$ tratamientos del bloque inicial son iguales para cualquier número t de tratamientos. El último tratamiento del bloque se toma como el valor correspondiente en la columna para t en la tabla 10A.1 del apéndice; por ejemplo, con $k = 3$, $r = 3$, el bloque inicial para $t = 5$ es (0, 1, 2) y es (0, 1, 3) para $t = 6$.

Si se requiere un diseño con más réplicas ($r > k$), la tabla proporciona un bloque generador de segundo o tercer nivel, en ella se construye un conjunto de bloques a partir de cada uno mediante la sustitución cíclica. El diseño con $t = 6$ tratamientos en bloques de tamaño $k = 3$ generados con el bloque inicial (0, 1, 3) tenían $r = 3$ réplicas, el diseño con $r = 6$ réplicas se construye agregando el bloque inicial (0, 2, 1) que se encuentra en la tabla 10A.1 en la línea con $k = 3$ y $r = 6$; el diseño obtenido tiene $b = 12$ bloques, la mitad de ellos generados por cada uno de los bloques iniciales como se muestra en el cuadro 10.4. Este diseño está parcialmente balanceado con $\lambda_1 = 3$ y $\lambda_2 = 2$ para las dos clases asociadas.

$$E = \frac{(t-1)(r-1)}{(t-1)(r-1) + r(s-1)} \quad (10.6)$$

Cómo construir los diseños α

La construcción de un diseño para $t = sk$ tratamientos con r réplicas comienza con un arreglo de $k \times r$ designado como el arreglo *generador* α . Cada columna del arreglo generador α se usa para construir $s - 1$ columnas adicionales por sustitución cíclica y el nuevo arreglo de $k \times rs$ recibe el nombre de arreglo *intermedio* α^* .

Consideremos un arreglo generador α de 4×3 para $t = 12$ tratamientos, $k = 4$ unidades por bloque, $r = 3$ réplicas y $s = 3$ bloques por réplica:

Arreglos generador α

Columna	1	2	3
0	0	0	
0	0	2	
0	2	1	
0	1	1	

Cada columna se usa para generar $s - 1 = 2$ columnas adicionales por sustitución cíclica con $\text{mod}(s) = \text{mod}(3)$ para derivar el arreglo intermedio α^* ; el resultado será $r = 3$ arreglos de tamaño $k \times s$. En este ejemplo, habrá tres arreglos de 4×3 , los arreglos intermedios α^* generados de cada una de las tres columnas del arreglo α son:

*Arreglos intermedios α^**

Arreglo generado a partir de		
Columna 1	Columna 2	Columna 3
0 1 2	0 1 2	0 1 2
0 1 2	0 1 2	2 0 1
0 1 2	2 0 1	1 2 0
0 1 2	1 2 0	1 2 0

Por último, las etiquetas se obtienen sumando $s = 3$ a cada elemento del segundo renglón del arreglo intermedio α^* , $2s = 6$ para cada elemento del tercer renglón, y $3s = 9$ al cuarto renglón. Las columnas de los arreglos son los bloques del diseño y cada conjunto de s columnas generadas con una columna arreglo generador α constituye una réplica completa; en el cuadro 10.5 se muestra el diseño $\alpha(0, 1, 2)$ obtenido.

Las etiquetas de tratamiento para el diseño son $(0, 1, 2, \dots, t-1)$ y los tratamientos reales se asignan al azar a las etiquetas. Además, la aleatorización procede con una asignación aleatoria de los bloques del diseño a los bloques reales dentro de

Cuadro 10.5 Diseño α para 12 tratamientos en tres grupos de réplicas

<i>Réplica</i>	I			II			III			
	<i>Bloque</i>	1	2	3	1	2	3	1	2	3
	0	1	2		0	1	2	0	1	2
	3	4	5		3	4	5	5	3	4
	6	7	8		8	6	7	7	8	6
	9	10	11		10	11	9	10	11	9

los grupos de réplicas y una asignación aleatoria de los tratamientos a las unidades experimentales reales dentro de los bloques físicos.

Una tabla de 11 arreglos α básicos reproducida de Paterson, Williams y Hunter (1978) se muestra en la tabla 10A.2 del apéndice. Se puede generar un total de 147 diseños α a partir de 11 arreglos. Los arreglos se pueden usar para generar diseños con $20 \leq t \leq 100$, números de tratamientos, con $r = 2, 3$ o 4 réplicas y la restricción usual de $t = sk$. Los arreglos producirán diseños con tamaños de bloques $k \geq 4$ con la condición $k \leq s$.

Observando el primer arreglo que aparece en la tabla 10A.2 del apéndice para $s = k = 5$, vemos que el arreglo tiene $k = 5$ renglones y $r = 4$ columnas.

0	0	0	0
0	1	4	2
0	2	3	4
0	3	2	1
0	4	1	3

Este arreglo en particular se puede usar para generar un diseño α con $r = 2$ réplicas si se usan sólo sus dos primeras columnas; de la misma manera, se pueden generar diseño de tres o cuatro réplicas con las primeras tres o cuatro columnas, respectivamente. Un diseño con tamaño de bloque $k = 5$ se obtiene con los cinco renglones del arreglo, un diseño con tamaño de bloque $k = 4$ se construye con los primeros cuatro renglones, con $s = 5$ se puede construir un diseño para $t = 25$ tratamientos en bloques de $k = 5$ unidades con $r = 2, 3$ o 4 réplicas. Además, con $s = 5$ es posible construir un diseño para $t = 20$ tratamientos en bloques de $k = 4$ unidades con $r = 2, 3$ o 4 réplicas.

Diseños de bloque incompletos resolubles latinizados

Los diseños resolubles presentados hasta ahora tienen estructuras de bloque anidadas. La aleatorización de los bloques incompletos es independiente dentro de cada réplica, ya sea que el diseño tenga un solo criterio de bloque o dos en los renglones y columnas.

En algunos casos, los bloques de una réplica pueden tener alguna relación con los bloques de otra, estas relaciones existen en experimentos donde los bloques en diferentes réplicas son contiguos en su espacio, como se ilustra en el cuadro 10.6.

Cuadro 10.6 Ilustración de bloques contiguos en las réplicas

	<i>Bloque</i>				
	1	2	3	4	5
<i>Réplica I</i>					
<i>Réplica II</i>					
<i>Réplica III</i>					

Si la distribución del espacio es la mostrada en el cuadro, entonces cualquier bloque, como el bloque 1, en cada réplica se encuentra en la misma línea de espacio como bloque 1 en todas las demás réplicas. El resultado es la ocurrencia de un bloque “largo” en dirección vertical de la distribución del cuadro 10.6.

Si los bloques se aleatorizan por separado en cada réplica, existe la posibilidad de que todas las réplicas de uno o más tratamientos ocurran en el mismo bloque “largo”, lo que producirá una circunstancia indeseable en el caso de perturbaciones no previstas que negaran el uso del bloque “largo”. Para evitar esta posibilidad, Harshbarger y Davis (1952) propusieron los diseños de retícula *latinizados*, en los que la retícula rectangular se estructuró de manera que cada tratamiento ocurra sólo una vez en cada bloque largo.

Williams (1986) presentó los resultados en diseños más generales donde los tratamientos no aparecen más de una vez en los bloques largos y diseños latinizados para incluir el bloque renglón-columna. Los diseños α se pueden usar para generar los diseños latinizados más generales como lo ilustra John y Williams (1995); en el cuadro 10.7 se muestra un diseño renglón-columna latinizado, observe que los tratamientos aparecen no más de una vez en cada bloque largo (columnas). John y Williams (1955) hacen un análisis exhaustivo de los diseños resolubles y cílicos.

Cuadro 10.7 Diseño renglón-columna latinizado

		Bloque			
		1	2	3	4
<i>Réplica I</i>	0	1	2	3	
	8	5	10	7	
	4	9	6	11	
<i>Réplica II</i>	6	0	3	5	
	1	7	4	2	
	11	10	9	8	
<i>Réplica III</i>	9	8	0	1	
	7	6	5	4	
	2	3	11	10	

10.4 Elección de diseños de bloques incompletos

Un experimento exitoso incluye el diseño de tratamiento correcto para contestar las preguntas de investigación específicas y proporcionar estimaciones precisas de las medias de tratamiento y los contrastes de interés para el investigador. Con frecuencia se necesita un diseño de bloques incompletos para obtener las mejores condiciones posibles para las comparaciones entre tratamientos con grupos tan homogéneos de unidades experimentales como sea posible.

El conocimiento del investigador sobre las condiciones del experimento y el material es indispensable para construir los bloques del tamaño y composición adecuados. Cada experimento presenta sus propios retos y condiciones, ningún conjunto de reglas proporciona el diseño correcto para todo experimento, pero un investigador conocedor puede trazar un experimento exitoso con la aplicación inteligente de los principios de diseño para el material experimental.

Las pautas que guían la selección o construcción de un diseño para un problema de investigación específico es hacer todo lo posible por elegir un diseño adecuado para el problema de investigación. En la mayoría de los casos, se dispone de diseños apropiados para las necesidades de los experimentos, ya que se cuenta con un gran número de planes publicados para los diseños de bloques incompletos balanceados, producto de la investigación sobre la construcción de diseños.

Los métodos sistemáticos para construir diseños de bloque incompletos balanceados requiere matemáticas que están fuera del alcance de este libro, pero se pueden encontrar exposiciones de las investigaciones sobre la existencia y construcción de diseños en John (1971), John (1987) y en John y Williams (1995).

Dónde encontrar planes de diseños

Los programas de computadora proporcionan el método más conveniente para desarrollar un plan de diseño de bloque incompleto; CycDesiGN (Whitaker, Williams y John, 1998) y ALPHA+ (Williams y Talbot, 1993) son dos ejemplos de programas que generan los diseños presentados en los capítulos 9 y 10.

En el apéndice 9A.1 se encuentran tablas de planes para experimentos pequeños con $t \leq 11$ tratamientos, y los planes para otros números de tratamientos están en Cochran y Cox (1957), Cox (1958), Fisher y Yates (1963), Box, Hunter y Hunter (1978) y en Mason Gunst y Hess (1989). Cochran y Cox (1957) y Patersen (1985) presentan planes para retículas balanceadas y parcialmente balanceadas. Algunos planes para diseños con bloques renglón-columna como cuadrados latinos incompletos ortogonales se encuentran en Youden (1940), Cochran y Cox (1957), Cox (1958) y Patersen (1985). Algunos planes para diseños ortogonales por renglón están en el apéndice 9A.2. Un amplio número de diseños de bloques incompletos parcialmente balanceados se encuentran en Clatworthy (1973) y en Bose, Clatworthy y Shrikande (1954).

Un enfoque informal para la construcción de diseños

Cuando los diseños estándar no se ajustan al problema de investigación, quizá sea necesario encontrar una solución de diseño innovadora, con un enfoque más informal para el diseño de experimentos (Mead, 1988). La razón más importante para pensar en un diseño informal es evitar la trampa de cambiar el diseño del tratamiento a fin de ajustarse a un diseño existente en particular, porque en ninguna circunstancia sería inteligente cambiar el problema de investigación para ajustarse a un diseño experimental. Sin embargo, tiene ventajas entender algunas propiedades formales del diseño de experimentos para apreciar los enfoques innovadores y evitar obstáculos posibles del enfoque informal, el investigador debe estar familiarizado con las condiciones experimentales y el material para construir un diseño adecuado para el estudio.

La construcción puede ser bastante informal si se tiene mucha atención al colocar los tratamientos en las unidades experimentales para evitar contratiempos, debe prestarse atención especial a la cantidad de información disponible para las distintas comparaciones entre tratamientos al construir el diseño. Las diferencias relativas en las varianzas de las comparaciones de tratamientos pertinentes debe evaluarse antes de la experimentación real. Mead (1988) apoya este tipo de enfoque.

El factor de eficiencia aumenta con el tamaño de bloque

Se espera que la varianza del error experimental disminuya conforme lo hace el tamaño del bloque, la disminución en σ^2 con tamaños de bloque más pequeños depende del éxito logrado al colocar las unidades experimentales en el arreglo correcto de bloques, pero sin embargo, una inspección cuidadosa del factor de eficiencia sugiere adoptar un tamaño de bloque tan grande como sea posible. La medida de eficiencia para un diseño de bloque incompleto balanceado se puede expresar como:

$$E = \frac{\lambda t}{rk} = \frac{t}{t-1} \left(1 - \frac{1}{k}\right) \quad (10.7)$$

El valor de E aumenta con el tamaño de bloque k y un examen de la lista de diseños de bloque incompletos balanceados en el apéndice 9A.1 confirmará esta relación. Una relación similar se cumple para los diseños de bloque incompletos parcialmente balanceados.

Los recursos determinan el tamaño de bloque

Los recursos disponibles determinan la elección del tamaño de bloque porque en algunos experimentos el tamaño de bloque está restringido por los instrumentos disponibles como ocurrió en el experimento de vinilación de metilglucósidos del ejemplo 9.2. El investigador disponía sólo de tres cámaras para probar cinco tratamientos de presión, con $t = 5$, $r = 6$, $k = 3$ y $\lambda = 3$, el factor de eficiencia para el diseño de bloques incompleto balanceado era $E = 3(5)/6(3) = 0.83$. Una precisión equivalente para un diseño de bloques incompletos con cinco cámaras de presión requeriría una reducción del 17% en σ^2 para el diseño, pero la única elección disponible para el investigador era el diseño de bloques incompletos con tres tratamientos por bloque.

Algunas condiciones experimentales permiten una mayor flexibilidad al seleccionar el tamaño de bloque, como las pruebas de variedad de cosecha que se colocan en parcelas de una granja experimental. En muchas situaciones los bloques pueden tener cualquier tamaño que convenga, el principio a seguir es tener un tamaño de bloque razonable que proporcione parcelas lo más homogéneas posible, lo que es necesario para las comparaciones precisas entre las variedades. En este caso, el tamaño de bloque debe ser lo más grande que se pueda mientras mantenga la precisión necesaria para las comparaciones de variedad.

Los diseños parcialmente balanceados aumentan la eficiencia en el uso de recursos

Los diseños parcialmente balanceados se pueden usar en lugar de los diseños balanceados para optimizar el uso de los recursos. Examinando el diseño parcialmente balanceado del cuadro 9.3, los tamaños de bloque se restringieron a cuatro unidades experimentales, en tanto que un diseño balanceado habría requerido diez réplicas de los 6 tratamientos y 15 bloques. Los 3 bloques del diseño parcialmente balanceado contienen dos réplicas, pero si se necesitan más, se puede repetir este diseño. A menos que se requieran diez réplicas, el diseño parcialmente balanceado exige menos recursos que el balanceado y el investigador tenía que hacer algún sacrificio para tener un experimento más pequeño. El diseño parcialmente balanceado proporciona más precisión en algunas comparaciones que en otras debido a que no todos los pares de tratamientos ocurren en el mismo número de bloques; sin embargo, con una elección prudente del diseño es posible tener un diseño parcialmente balanceado que proporcione casi un equilibrio en todas las comparaciones.

En el diseño parcialmente balanceado del cuadro 9.3, que tenía algunos pares de tratamientos en $\lambda_1 = 2$ bloques y otros en $\lambda_2 = 1$ bloque, se puede demostrar que la

eficiencia de las comparaciones respectivas entre primero y segundos asociados es $E_1 = 1.00$ y $E_2 = 0.86$, con una eficiencia promedio de $E = 0.88$ (Clatworthy, 1973). Existe una precisión casi igual para ambos conjuntos de asociados y sólo se sacrifica una pequeña parte de esa precisión con el diseño parcialmente balanceado.

En experimentos grandes los diseños α liberan restricciones

Los diseños de bloque incompletos tuvieron su origen en pruebas experimentales que incluyeron un gran número de tratamientos, como en las pruebas de variedades. El objetivo principal de los diseños de bloque incompletos era reducir el tamaño de bloque para reducir la varianza del error experimental y tener experimentos que se pudieran manejar con base en grupos de réplicas; en las pruebas agrícolas, los tamaños de bloque eran flexibles siempre que su reducción lograra disminuir la varianza del error. Los diseños de retícula cuadrada y rectangular originales proporcionaron cierto alivio a estas situaciones, el número de variedades de estas pruebas todavía era una restricción aunque a menudo era posible que el investigador ajustara el número de variedades incluidas sin consecuencias graves.

Los diseños α relajaron mucho la restricción del número de tratamientos para los diseños de bloque incompletos resolubles, aunque todavía requieren que el número de tratamientos sea producto del número de unidades por bloque y el número de bloques en el grupo de réplicas, pero ahora es posible tener un diseño resoluble para una gran variedad de tamaños de tratamientos, bloques y réplicas. Además, los diseños α se pueden generar a partir de un conjunto de tablas relativamente sencillas.

Los diseños de tratamiento factoriales requieren atención especial

Los diseños de tratamiento estructurados, como los factoriales, requieren que se tenga cuidado al construir diseños de bloques incompletos. En el capítulo 11 se estudiará, la construcción de diseños de bloque incompletos con efectos factoriales que se confunden con los bloques, estos diseños son útiles en particular para las series de 2ⁿ factoriales que tienen muchos factores, porque los resultados no muestran una gran diferencia al confundir por completo las interacciones de orden mayor con los bloques.

Los diseños cíclicos ayudan en los diseños de tratamiento no estructurados

Los diseños balanceados o los resolubles pueden, en algunos casos, agotar los recursos disponibles debido al número de réplicas requerido, es posible usar los diseños parcialmente balanceados más flexibles en lugar de los diseños balanceados con poca pérdida de precisión, pero tal vez no se disponga de tablas de diseños parcialmente balanceados porque muchas aparecen en publicaciones antiguas, muchas de las cuales se pueden construir mediante la sustitución cíclica. Quizá los diseños cíclicos sean la mejor alternativa para producir un diseño eficiente, pues proporcionan diseño de muchos tamaños y su construcción es sencilla a partir de uno o dos bloques iniciales de tratamientos. Las tablas de bloques iniciales para los diseños cíclicos se encuentran en publicaciones un poco más recientes.

EJERCICIOS PARA EL CAPÍTULO 10

1. Un horticultor llevó a cabo una prueba de campo con $t = 8$ variedades de brócoli, utilizando un diseño de bloques incompletos balanceados, el diseño tenía $k = 4$ parcelas por bloque y $s = 2$ bloques por réplica y había $r = 7$ grupos de réplicas. Los siguientes datos son las libras de brócoli cosechadas por parcela (R = réplica, B = bloque dentro de la réplica, T = variedad).

<i>R</i>	<i>B</i>	<i>T</i>	Cosecha	<i>R</i>	<i>B</i>	<i>T</i>	Cosecha	<i>R</i>	<i>B</i>	<i>T</i>	Cosecha
1	1	1	46.5	3	2	2	52.7	6	1	1	45.6
1	1	2	55.7	3	2	4	52.5	6	1	3	37.0
1	1	3	37.7	3	2	5	46.4	6	1	5	49.5
1	1	4	50.3	3	2	7	46.3	6	1	7	45.4
1	2	5	43.1	4	1	1	40.7	6	2	2	48.2
1	2	6	47.6	4	1	4	53.0	6	2	8	54.0
1	2	7	35.5	4	1	6	45.0	6	2	4	47.4
1	2	8	45.9	4	1	7	38.0	6	2	6	53.8
2	1	1	40.1	4	2	2	56.1	7	1	1	44.6
2	1	2	55.8	4	2	3	39.0	7	1	4	52.4
2	1	7	39.7	4	2	5	54.7	7	1	5	50.2
2	1	8	51.7	4	2	8	48.5	7	1	8	52.0
2	2	3	41.2	5	1	1	44.1	7	2	2	56.8
2	2	4	61.7	5	1	2	56.6	7	2	3	37.8
2	2	5	49.8	5	1	5	44.7	7	2	6	45.7
2	2	6	53.6	5	1	6	51.7	7	2	7	42.6
3	1	1	42.3	5	2	3	39.0				
3	1	3	43.8	5	2	4	47.8				
3	1	6	45.6	5	2	7	41.6				
3	1	8	51.0	5	2	8	49.4				

- a. Escriba un modelo lineal para este experimento, describa los términos y realice el análisis intrabloques.
 - b. Calcule la estimación del error estándar de la diferencia entre las estimaciones de mínimos cuadrados de dos medias de tratamiento.
 - c. ¿Cuál es el factor de eficiencia para este diseño?
 - d. Escriba un resumen de los resultado de su análisis y la evaluación del diseño.
2. Un agrónomo realizó una prueba de variedad de alfalfa con un diseño de retícula cuádruple, tenía 25 cultivos de variedades en cuatro grupos de réplicas y cada uno tenía cinco bloques de cinco variedades en una retícula cuadrada. Los datos de la cosecha que siguen son libras de alfalfa cosechadas por parcela, los números de las variedades aparecen entre paréntesis.

Bloque	Réplica I									
1	19.1	(2)	20.4	(1)	23.2	(4)	19.3	(3)	21.4	(5)
2	18.6	(8)	18.3	(6)	21.3	(10)	12.0	(9)	19.3	(7)
3	20.8	(15)	19.5	(11)	20.8	(13)	20.3	(12)	19.0	(14)
4	21.0	(17)	19.4	(16)	19.7	(19)	17.5	(18)	20.2	(20)
5	19.6	(24)	19.0	(23)	19.4	(21)	20.6	(25)	20.3	(22)

Bloque	Réplica II									
1	19.4	(11)	19.8	(21)	18.4	(6)	21.5	(1)	19.7	(16)
2	18.6	(17)	20.8	(2)	20.9	(12)	19.7	(7)	20.0	(22)
3	16.6	(18)	19.8	(3)	18.2	(23)	19.1	(8)	20.7	(13)
4	21.8	(4)	20.6	(14)	20.0	(19)	16.8	(9)	19.5	(24)
5	21.3	(15)	19.6	(20)	20.1	(25)	20.4	(10)	20.7	(5)

Bloque	Réplica III									
1	21.1	(1)	20.1	(13)	20.3	(7)	20.8	(25)	22.2	(19)
2	19.5	(21)	19.1	(14)	20.5	(20)	20.5	(2)	20.3	(8)
3	13.5	(9)	23.1	(22)	20.5	(3)	21.2	(15)	18.6	(16)
4	22.6	(10)	17.2	(23)	22.2	(4)	17.0	(17)	17.8	(11)
5	20.0	(24)	17.7	(18)	18.7	(5)	19.8	(12)	16.4	(6)

Bloque	Réplica IV									
1	20.2	(12)	22.6	(1)	19.5	(23)	22.4	(20)	13.6	(9)
2	22.5	(24)	21.9	(16)	22.5	(10)	21.1	(13)	19.3	(2)
3	19.2	(25)	20.0	(3)	17.1	(6)	21.0	(17)	18.9	(14)
4	21.7	(15)	20.0	(21)	18.2	(7)	17.3	(18)	20.9	(4)
5	17.3	(8)	19.3	(22)	20.7	(5)	19.2	(11)	16.9	(19)

Fuente: Dr. M. Ottman, Department of Plant Sciences, University of Arizona.

- Escriba un modelo lineal para el experimento, explique los términos y realice el análisis intrabloques para un diseño de retícula simple, triple o cuádruple (según indique el profesor).
 - Calcule las estimaciones del error estándar de la diferencia entre las estimaciones intrabloques de dos medias de variedades que sean primeros asociados y segundos asociados. También calcule un error estándar promedio.
 - ¿Cuál es el factor de eficiencia para el diseño de retícula o triple? Interprételo.
 - Escriba un resumen de los resultados de su análisis y la evaluación del diseño.
3. Un agrónomo realizó una prueba de variedades de trigo con un diseño de retícula cuadrada balanceado, había $t = 9$ variedades en una retícula cuadrada balanceada de 3×3 en $r = 4$ grupos de réplicas. Las cosechas de trigo se muestran en la tabla siguiente en arreglos renglón-columna para cada grupo de réplicas y los números de las variedades aparecen entre paréntesis.

<i>Réplica I</i>			<i>Réplica II</i>		
53.5 (6)	53.2 (4)	57.7 (5)	53.7 (4)	53.6 (2)	57.8 (9)
53.1 (3)	58.6 (1)	53.9 (2)	54.5 (3)	52.8 (7)	53.3 (5)
57.2 (9)	55.0 (7)	51.5 (8)	48.9 (8)	53.5 (6)	56.7 (1)
<i>Réplica III</i>			<i>Réplica IV</i>		
49.4 (8)	54.7 (4)	55.6 (3)	54.0 (7)	57.2 (1)	53.2 (4)
54.4 (6)	54.2 (2)	54.4 (7)	56.9 (9)	54.8 (3)	55.4 (6)
59.7 (1)	55.7 (9)	54.1 (5)	48.9 (8)	53.4 (2)	55.9 (5)

- Escriba un modelo lineal para este experimento, describa los términos y realice el análisis intrabloques para este experimento.
 - Calcule la estimación del error estándar de la diferencia entre las estimaciones de mínimos cuadrados de las dos medias de tratamiento.
 - ¿Cuál es el factor de eficiencia para este diseño?
 - Escriba un resumen de los resultados del análisis y la evaluación del diseño.
4. Se realizó una prueba de variedades con un diseño $\alpha(0, 1, 2)$ resoluble, se tenían $t = 18$ variedades en $r = 4$ grupos de réplicas y había $s = 3$ bloques de $k = 6$ variedades en cada réplica. Los siguientes son los datos de la cosecha con los números de variedad entre paréntesis. Las variedades 1 y 5 fueron las variedades de control.

<i>Bloque</i>	<i>Réplica I</i>											
1	88.2	(5)	82.5	(10)	84.3	(15)	87.0	(6)	84.5	(12)	88.9	(8)
2	82.4	(1)	82.9	(14)	83.1	(3)	84.7	(13)	83.3	(16)	89.0	(4)
3	93.1	(2)	82.7	(11)	88.9	(17)	88.6	(18)	84.1	(9)	87.5	(7)

<i>Bloque</i>	<i>Réplica II</i>											
1	85.4	(4)	73.0	(11)	84.2	(7)	80.3	(14)	79.6	(10)	86.0	(6)
2	87.9	(8)	85.1	(9)	79.4	(18)	80.7	(13)	89.3	(5)	81.5	(3)
3	82.4	(1)	88.5	(2)	87.0	(12)	85.4	(17)	85.9	(15)	79.1	(16)

<i>Bloque</i>	<i>Réplica III</i>											
1	83.6	(6)	79.4	(17)	81.3	(4)	80.5	(9)	80.9	(8)	79.3	(1)
2	80.4	(7)	88.2	(5)	82.3	(14)	88.0	(12)	90.0	(2)	83.6	(3)
3	81.4	(18)	84.8	(15)	81.0	(10)	81.2	(13)	79.1	(11)	83.8	(16)

<i>Bloque</i>	<i>Réplica IV</i>											
1	80.5	(16)	77.1	(11)	84.4	(17)	90.4	(6)	82.9	(14)	83.0	(12)
2	87.9	(8)	78.9	(18)	81.4	(1)	83.5	(2)	82.2	(15)	79.0	(3)
3	84.2	(7)	83.0	(10)	87.6	(9)	81.7	(13)	91.3	(5)	87.4	(4)

Fuente: P. Seeger, Department of Statistics, The Swedish University of Agricultural Sciences.

- a. Identifique los primeros, segundos y terceros asociados de la variedad 1, donde $\lambda_1 = 0$, $\lambda_2 = 1$ y $\lambda_3 = 2$. ¿Cuántas réplicas de cada variedad hay y qué variedades están en cada grupo de asociados?
 - b. Escriba un modelo lineal para el experimento, explique los términos y realice el análisis intrabloques para obtener las estimaciones intrabloques de las medias de las variedades y sus errores estándar estimados.
 - c. Los diseños α tienen una multiplicidad de errores estándar para las diferencias entre los pares de medias de tratamiento estimadas; si su programa puede hacerlo, calcule los errores estándar estimados de la diferencia entre las estimaciones intrabloques de las medias de variedad de control, las variedades 1 y 5, y las otras variedades; el error estándar promedio para los primeros, segundos y terceros asociados; y un error estándar promedio global.
 - d. Escriba un resumen de los resultados de su análisis y la evaluación del diseño.
5. Debe realizarse un estudio de factores humanos sobre la velocidad de percepción relativa a la forma del objeto con gráficas estadísticas, se usarán ocho formas como tratamientos del estudio, y como se piensa que un sujeto no debe evaluar más de cuatro formas en una corrida, los investigadores han decidido usar un diseño de bloques completos con los sujetos como bloques.
- a. Construya un diseño cíclico con $r = 4$ réplicas de $t = 8$ tratamientos de formas de objetos. Use los sujetos como bloques con $k = 4$ tratamientos por sujeto.
 - b. Haga una asignación aleatoria de las etiquetas de los tratamientos a los sujetos y de las formas reales de los objetos a las etiquetas de tratamiento.
 - c. Suponga que el orden de presentación a los sujetos puede ser una fuente importante de variación. Construya el diseño como uno de renglón-columna y haga la aleatorización correspondiente.
 - d. Suponga que se requirieron ocho réplicas para el estudio; construya el diseño cíclico con ocho réplicas.
6. Genere un diseño α para tres réplicas de 35 tratamientos en bloques de cinco parcelas.
- a. Muestre el arreglo generador α , los arreglos intermedios α y el diseño final alfa con etiquetas de tratamientos.
 - b. Haga una asignación aleatoria de los bloques incompletos de etiquetas de tratamiento a los bloques reales y de las etiquetas de tratamiento a la unidades dentro de los bloques.
 - c. Asigne al azar los tratamientos reales a las etiquetas de tratamiento.

Apéndice 10A.1 Planes para diseños cíclicos

Tabla 10A.1 Bloques iniciales para generar diseños de bloques incompletos parcialmente balanceados para $4 \leq t \leq 15$, $r \leq 10$

<i>k</i>	<i>r</i>	<i>Primeros k - 1 tratamientos</i>	<i>k</i> -ésimo tratamiento, <i>t</i> =											
			4	5	6	7	8	9	10	11	12	13	14	15
2	2	0	1	1	1	1	1	1	1	1	1	1	1	1
	4	0		2	2	2	3	3	3	3	3	3	5	4
	6	0		1	3	3	2	2	2	2	5	5	2	6
	8	0		1	4	5	4	4	4	4	2	2	4	3
	10	0		2	1	4	5	5	5	5	4	4	3	5
	3	0 1		2	2	3	3	3	3	4	4	4	4	4
3	6	0 2		1	3	1	3	7	6	7	7	5	7	7
	9	0 1		3	2	3	3	4	5	3	3	6	4	6
4	4	0 1 3		—	2	2	6	7	7	6	7	7	9	7
	8	0 1 4		—	2	2	6	7	8	2	6	6	6	6
	5	0 1 2 4		—	—	5	5	7	7	7	7	7	9	10
	10	{ 0 2 3 4 0 2 3 6 }		—	—	5	5	7	8	9	8		7	11
5	6	0 1 2 3 6		—	—	—	5	5	5	5	10	10	10	10
	7	0 1 2 3 4 7		—	—	—	—	5	5	9	9	9	9	10
8	8	0 1 2 3 4 6 8		—	—	—	—	—	5	9	9	9	9	11
	9	0 1 2 3 4 5 7 9		—	—	—	—	—	—	8	8	8	10	10
10	10	0 1 2 3 4 5 6 9 10		—	—	—	—	—	—	7	7	7	12	12

Reproducido con permiso de J. A. John (1981), "Efficient Cyclic Designs", *Journal of the Royal Statistical Society, B*, 43, 76-80.

Apéndice 10A.2 Arreglos generadores para diseños α

Tabla 10A.2 Arreglos generadores de $(k \times r)$ para diseños α

$s = k = 5$				$s = k = 6$				$s = k = 7$			
0	0	0	0	0	0	0	0	0	0	0	0
0	1	4	2	0	1	5	4	0	1	3	2
0	2	3	4	0	3	2	5	0	2	6	4
0	3	2	1	0	2	3	1	0	4	5	1
0	4	1	3	0	4	1	2	0	3	2	6
				0	5	1	3	0	5	1	3
								0	6	4	5
$s = k = 8$				$s = k = 9$				$s = k = 10$			
0	0	0	0	0	0	0	0	0	0	0	0
0	1	2	6	0	1	8	7	0	1	9	5
0	3	7	1	0	3	6	4	0	3	6	9
0	5	3	4	0	7	2	3	0	5	7	2
0	2	5	3	0	2	3	5	0	4	5	6
0	4	1	6	0	4	1	6	0	6	3	1
0	6	0	2	0	5	7	2	0	7	2	4
0	7	6	5	0	6	5	1	0	8	4	7
				0	8	4	7	0	9	8	2
								0	2	6	3
$s = 11, k = 9$				$s = 12, k = 8$				$s = 13, k = 7$			
0	0	0	0	0	0	0	0	0	0	0	0
0	1	6	7	0	1	2	3	0	1	4	10
0	4	8	1	0	7	5	1	0	3	8	11
0	9	7	5	0	9	6	4	0	9	2	1
0	2	3	6	0	4	11	8	0	12	10	6
0	5	1	3	0	11	3	10	0	8	5	12
0	6	5	10	0	10	4	7	0	6	7	8
0	3	9	4	0	5	1	6				
0	7	4	1								
$s = 14, k = 7$				$s = 15, k = 6$							
0	0	0	0	0	0	0	0				
0	1	8	10	0	1	8	7				
0	9	10	7	0	3	12	14				
0	11	13	2	0	7	2	5				
0	2	6	1	0	10	13	11				
0	5	1	12	0	14	3	8				
0	3	1	11								

Reproducido de H. D. Patterson, E. R. Williams y E. A. Hunter (1978), “Block designs for variety trials”, *Journal of Agricultural Science* 90, tabla 2, p. 399, con permiso de Cambridge University Press.

11 Diseños de bloques incompletos: Tratamientos factoriales

En los capítulos 11 y 12 se explora la versatilidad de los diseños de tratamientos factoriales, haciendo hincapié en los factoriales 2^n y 3^n . En el presente capítulo se utilizan los efectos y notación específicos de los tratamientos factoriales a fin de desarrollar métodos para construir diseños de bloque incompletos orientados a ellos.

11.1 Aprovechamiento de los diseños de tratamientos factoriales

El diseño de tratamientos factoriales se explica en los capítulos 6 y 7 como un medio para investigar los efectos de varios factores de tratamiento en el mismo experimento. La ventaja principal de los arreglos factoriales estriba en la posibilidad de determinar si los factores actúan de manera independiente o interactúan entre ellos al afectar las unidades experimentales.

Los diseños factoriales 2^n y 3^n son de gran importancia práctica y se usan ampliamente en los estudios de investigación. Los factoriales 2^n tienen n factores a dos niveles y los factoriales 3^n tienen n factores a tres niveles; conforme el número de factores aumenta, el número de combinaciones de tratamientos aumenta con rapidez, de manera que se requiere que los diseños de bloque incompletos controlen el error experimental. Los arreglos factoriales en cada diseño se pueden aprovechar para la investigación de los efectos del factor y para facilitar el análisis de esos efectos.

11.2 Factoriales 2ⁿ para evaluar muchos factores

Las necesidades de un programa de investigación pueden exigir que se investiguen muchos factores y sus interrelaciones en cuanto a cómo afectan los resultados del proceso. Considerando un proceso de producción continua en una planta de manufactura química que no alcanza el rendimiento proyectado, los factores de los diferentes pasos del proceso afectan la cantidad obtenida de producto final. El problema es identificar esos factores y determinar el nivel que optimiza el rendimiento, como la concentración de catalíticos, las concentraciones de reactivos en solventes y la razón de un reactivo a otro, al igual que la temperatura, presión, tasa de agitación y tiempo de reposo en las cámaras de reacción.

Es claro que los ingenieros que intentan mejorar el desempeño de la planta se enfrentan a una tarea colosal, tienen muchos factores que investigar con la posibilidad de que sólo unos cuantos sean de importancia. Deben estudiar los factores para determinar cuales estudiar más a fondo y al mismo tiempo, controlar el costo de los experimentos.

El factorial 2ⁿ con muchos factores, cada uno a nivel “bajo” o “alto”, se puede usar para detectar los factores importantes en el proceso con un mínimo de unidades experimentales; es posible detectar las tendencias principales con factores de dos niveles para identificar los factores potencialmente importantes. En consecuencia, los factoriales de 2ⁿ se usan con frecuencia en las primeras etapas de experimentación para detectar los factores que son candidatos potenciales para una investigación más detallada. El siguiente ejemplo con tres factores explica las características de los diseños factoriales 2ⁿ.

Ejemplo 11.1 Fabricación de muelles de camión

Un experimento descrito por Pignatello y Ramberg (1985) se diseñó para investigar los efectos de factores sobre el proceso de manufactura para muelles de camión. Los muelles ensamblados pasan por un horno a alta temperatura y luego son colocados en una máquina de formación que induce una curvatura en ellos, sosteniéndolos en una prensa de alta presión durante un periodo corto.

El tratamiento factorial diseñado para el experimento consistió en tres factores, con dos niveles cada uno, que eran temperatura del horno (*A*), tiempo de calentamiento en el horno (*B*) y tiempo de transferencia entre el horno y el maquinado (*C*). En la tabla 11.1 se muestran las ocho combinaciones para los tres factores junto con la medida observada de calidad del producto (*y*) para una réplica de cada tratamiento.

Nuevas etiquetas de tratamiento

En general, el nivel bajo de un factor cuantitativo se denota con “0” y el nivel alto con “1”. De manera equivalente, las dos categorías del factor cuantitativo se pueden codificar como “0” o “1”. Otra notación útil para las combina-

Tabla 11.1 Observaciones de la calidad de los muelles para camión en un experimento factorial 2^n

<i>A</i>	<i>B</i>	<i>C</i>	<i>y</i>			
				Temperatura de horno ($^{\circ}$ F)	Tiempo de calentamiento (seg)	Tiempo de transferencia (seg)
1840	23	10	32			
1880	23	10	35			
1840	25	10	28			
1880	25	10	31			
1840	23	12	48			
1880	23	12	39			
1840	25	12	28			
1880	25	12	29			

ciones de tratamientos en factoriales 2^n se ilustra con un factorial 2^3 en el cuadro 11.1.

Las letras mayúsculas *A*, *B* y *C* representan los factores y la etiqueta de “tratamiento” usa las letras minúsculas correspondientes *a*, *b* y *c*. La letra minúscula está presente si el factor está en el nivel 1, de los contrario, el factor tiene nivel 0. La etiqueta de tratamiento es “(1)” si todos los factores están en el nivel 0. La correspondencia entre las etiquetas de tratamiento y las designaciones (0, 1) para los niveles del factor se muestran en el cuadro 11.1 junto con los niveles reales de cada factor en el experimento de muelles para camión.

Evaluación de los efectos de factoriales 2^n

El *efecto* de un factor en un factorial 2^n corresponde al cambio en la respuesta cuando se pasa del nivel bajo al alto del factor. (Los efectos simples, principales y de interacción para los factoriales se encuentran en el capítulo 6.) Para explicarlo,

Cuadro 11.1 Etiquetas de tratamiento para el experimento con muelles de camión

Tratamiento	<i>A</i>	<i>B</i>	<i>C</i>	Temperatura del horno ($^{\circ}$ F)	Tiempo de calentamiento (seg)	Tiempo de transferencia (seg)	<i>y</i>
(1)	0	0	0	1840	23	10	32
<i>a</i>	1	0	0	1880	23	10	35
<i>b</i>	0	1	0	1840	25	10	28
<i>ab</i>	1	1	0	1880	25	10	31
<i>c</i>	0	0	1	1840	23	12	48
<i>ac</i>	1	0	1	1880	23	12	39
<i>bc</i>	0	1	1	1840	25	12	28
<i>abc</i>	1	1	1	1880	25	12	29

analizamos las observaciones del experimento de muelles de camión. En la tabla 11.2 se muestran los efectos simples de la temperatura del horno cuando cambia de 1840° a 1880°F, manteniendo constantes los factores de tiempo de calentamiento y tiempo de transferencia.

Tabla 11.2 Efectos simples de la temperatura del horno a niveles constantes de tiempo de calentamiento y tiempo de transferencia para el experimento con muelles de camión

Tiempo de calentamiento (B)	Tiempo de transferencia (C)	Temperatura del horno (A)			B	C
		1840	1880	Efecto simple		
23	10	32	35	35 – 32 = 3	0	0
25	10	28	31	31 – 28 = 3	1	0
23	12	48	39	39 – 48 = - 9	0	1
25	12	28	29	29 – 28 = 1	1	1

El efecto principal de un factor es el efecto promedio consecuencia de cambiarlo del nivel 0 al nivel 1. Así, El efecto principal de la temperatura del horno es el promedio de sus efectos simples:

$$A = \frac{1}{4}[3 + 3 + (-9) + 1] = - 0.5 \quad (11.1)$$

Cálculos equivalentes producen los efectos principales $B = -9.5$ y $C = 4.5$.

Dos factores, digamos A y B , interactúan si el efecto de A es diferente para los dos niveles de B . Cuando el tiempo de calentamiento es 23 segundos, $B = 0$ y el efecto de la temperatura del horno es:

$$(A | B = 0) = \frac{1}{2}[3 + (-9)] = - 3$$

Sin embargo, cuando el tiempo de calentamiento es 25 segundos, $B = 1$ y el efecto de la temperatura del horno es:

$$(A | B = 1) = \frac{1}{2}(3 + 1) - 2$$

Si la temperatura del horno aumenta de 1840° a 1880°F, la calidad del producto se *reduce* en tres unidades si el tiempo de calentamiento es 23 segundos. No obstante, La calidad del producto *aumenta* en dos unidades si el tiempo de calentamiento es 25 segundos. La respuesta a la temperatura del horno es diferente para los dos tiempos de calentamiento, lo que implica la posible interacción entre la temperatura del horno y el tiempo de calentamiento.

La interacción entre dos factores, digamos A y B , se define como la mitad de la diferencia entre el efecto de A para $B = 1$ y $B = 0$, entonces la estimación de la interacción entre la temperatura del horno y el tiempo de calentamiento es:

$$AB = \frac{1}{2} \{(A | B = 1) - (A | B = 0)\} = \frac{1}{2}[2 - (-3)] = 2.5$$

Si se intercambian los papeles de los factores A y B , el valor de interacción que se obtiene no cambia; los cálculos para las otras interacciones de dos factores producen $AC = -3.5$ y $BC = -5.5$.

La interacción de tres factores ABC surge si una interacción de dos factores, como AB , es diferente para $C = 0$ y $C = 1$; la interacción AB es la mitad de la diferencia entre el efecto de A para $B = 1$ y $B = 0$. Cuando el tiempo de transferencia es 10 segundos, $C = 0$ y la estimación de la interacción entre la temperatura del horno y el tiempo de calentamiento es:

$$(AB | C = 0) = \frac{1}{2}(3 - 3) = 0$$

La respuesta para A es de tres unidades sin importar el nivel de B . El valor cero para la interacción AB cuando $C = 0$ se presenta en la figura 11.1a como dos rectas de respuesta paralelas.

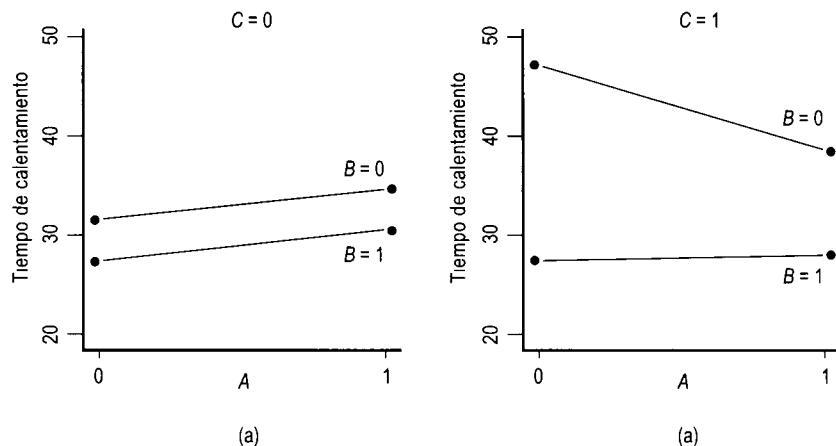


Figura 11.1 Representación gráfica de la interacción ABC descrita por interacciones diferentes de AB para $C = 0$ y $C = 1$

Cuando el tiempo de transferencia es 12 segundos, $C = 1$ y la interacción entre la temperatura del horno y el tiempo de calentamiento es:

$$(AB | C = 1) = \frac{1}{2}[1 - (-9)] = 5$$

La respuesta de A es diferente para $B = 0$ y $B = 1$. La medida positiva de la interacción AB cuando $C = 1$ se representa con las dos rectas de respuesta diferentes de la figura 11.1b.

La interacción entre la temperatura del horno y el tiempo de calentamiento difiere con el tiempo de transferencia. La estimación de la interacción AB es 0 cuando el tiempo de transferencia es de 10 segundos y 5 cuando ese tiempo es 12 segundos.

La mitad de la diferencia entre estas dos evaluaciones de la interacción AB es la interacción de tres factores. La estimación de la interacción de tres factores es:

$$ABC = \frac{1}{2} \{ (AB | C = 1) - (AB | C = 0) \} = \frac{1}{2}(5 - 0) = 2.5$$

La existencia de una interacción de tres factores se traduce en las dos representaciones gráficas de la interacción AB que se encuentran en la figura 11.1. Se logra el mismo valor para la interacción ABC si se evalúa la interacción AC con $B = 0$ y $B = 1$ o si la interacción BC se evalúa para $A = 0$ y $A = 1$.

Tabla de contrastes para resumir los efectos

Los efectos para factoriales 2ⁿ se pueden definir como contrastes con coeficientes +1 o -1 para cada combinación de tratamiento. Una tabla de signos + y - determina el coeficiente correcto para cualquier combinación de tratamiento en el contraste y proporciona un método sistemático para calcular los efectos factoriales.

El complemento total de los signos + y - necesarios para los contrastes en el factorial 2³ para el experimento con muelles de camión se muestra en la tabla 11.3. La tabla de signos + y - se genera de la siguiente manera:

Tabla 11.3 Coeficientes para contrastes en un diseño de tratamiento factorial 2³

Tratamiento	Efectos factoriales								
	I	A	B	C	AB	AC	BC	ABC	y
(1)	+	-	-	-	+	+	+	-	32
a	+	+	-	-	-	-	+	+	35
b	+	-	+	-	-	+	-	+	28
ab	+	+	+	-	+	-	-	-	31
c	+	-	-	+	+	-	-	+	48
ac	+	+	-	+	-	+	-	-	39
bc	+	-	+	+	-	-	+	-	28
abc	+	+	+	+	+	+	+	+	29
Divisor	8	4	4	4	4	4	4	4	
Efecto	33.8	- 0.5	- 9.5	4.5	2.5	- 3.5	- 5.5	2.5	
SC		0.5	180.5	40.5	12.5	24.5	60.5	12.5	

- La columna I, que contiene sólo signos +, se usa para estimar la gran media con un divisor de 2ⁿ.
- Las siguientes tres columnas, etiquetadas con los factores A, B y C, tienen signos + y - en *orden estándar*. Este orden tiene los niveles de los factores

La interacción entre la temperatura del horno y el tiempo de calentamiento difiere con el tiempo de transferencia. La estimación de la interacción AB es 0 cuando el tiempo de transferencia es de 10 segundos y 5 cuando ese tiempo es 12 segundos.

La mitad de la diferencia entre estas dos evaluaciones de la interacción AB es la interacción de tres factores. La estimación de la interacción de tres factores es:

$$ABC = \frac{1}{2} \left\{ (AB | C = 1) - (AB | C = 0) \right\} = \frac{1}{2}(5 - 0) = 2.5$$

La existencia de una interacción de tres factores se traduce en las dos representaciones gráficas de la interacción AB que se encuentran en la figura 11.1. Se logra el mismo valor para la interacción ABC si se evalúa la interacción AC con $B = 0$ y $B = 1$ o si la interacción BC se evalúa para $A = 0$ y $A = 1$.

Tabla de contrastes para resumir los efectos

Los efectos para factoriales 2ⁿ se pueden definir como contrastes con coeficientes +1 o -1 para cada combinación de tratamiento. Una tabla de signos + y - determina el coeficiente correcto para cualquier combinación de tratamiento en el contraste y proporciona un método sistemático para calcular los efectos factoriales.

El complemento total de los signos + y - necesarios para los contrastes en el factorial 2³ para el experimento con muelles de camión se muestra en la tabla 11.3. La tabla de signos + y - se genera de la siguiente manera:

Tabla 11.3 Coeficientes para contrastes en un diseño de tratamiento factorial 2³

Tratamiento	Efectos factoriales								
	I	A	B	C	AB	AC	BC	ABC	y
(1)	+	-	-	-	+	+	+	-	32
a	+	+	-	-	-	-	+	+	35
b	+	-	+	-	-	+	-	+	28
ab	+	+	+	-	+	-	-	-	31
c	+	-	-	+	+	-	-	+	48
ac	+	+	-	+	-	+	-	-	39
bc	+	-	+	+	-	-	+	-	28
abc	+	+	+	+	+	+	+	+	29
Divisor	8	4	4	4	4	4	4	4	
Efecto	33.8	- 0.5	- 9.5	4.5	2.5	- 3.5	- 5.5	2.5	
SC	0.5	180.5	40.5	12.5	24.5	60.5	12.5		

- La columna I, que contiene sólo signos +, se usa para estimar la gran media con un divisor de 2ⁿ.
- Las siguientes tres columnas, etiquetadas con los factores A, B y C, tienen signos + y - en *orden estándar*. Este orden tiene los niveles de los factores

dispuestos de manera que la columna A tiene pares sucesivos de signos $-$ y $+$, la columna B tiene pares de signos $-$ seguidos por pares de signos $+$ y la columna C tiene cuatro signos $-$ seguidos por cuatro signos $+$. En general la columna k tiene 2^{k-1} signos $-$ seguidos por el mismo número de signos $+$.

- Los coeficientes para cualquier interacción entre dos factores se obtienen mediante el producto de las columnas de coeficientes para los efectos principales correspondientes. Por ejemplo, los coeficientes para la columna AB son los productos de los elementos correspondientes en la columna A y la columna B .
- Los coeficientes para la interacción triple ABC se obtienen del producto de los coeficientes para cualquier conjunto de columnas cuyo símbolo es ABC . Es decir, los coeficientes para ABC se pueden obtener de cualquier producto de símbolos $A \times B \times C, AB \times C, AC \times B$ o $BC \times A$.

La columna I suele recibir el nombre de columna *identidad*. Cada columna, excepto I , tiene igual número de signos $+$ y $-$.

Las estimaciones de los efectos en la tabla 11.3 se calculan mediante el producto del signo de la columna correspondiente por la respuesta y en cada renglón de la tabla, la suma de los productos y la división de la suma entre el divisor adecuado. Los cálculos del efecto principal para la temperatura del horno en la ecuación (11.1) es el resultado de usar los coeficientes $+$ y $-$ de la columna A de la tabla 11.3 para el valor correspondiente de la respuesta y y dividir entre 4, es decir:

$$A = \frac{1}{4} (-32 + 35 - 28 + 31 - 48 + 39 - 28 + 29) = -0.5$$

De manera similar se usan los signos $+$ y $-$ de la columna B para el efecto principal del tiempo de calentamiento:

$$B = \frac{1}{4} (-32 - 35 + 28 + 31 - 48 - 39 + 28 + 29) = -9.5$$

El efecto de interacción AB se calcula usando los signos $+$ y $-$ de la columna AB para obtener:

$$AB = \frac{1}{4} (32 - 35 - 28 + 31 + 48 - 39 - 28 + 29) = 2.5$$

En general, el contraste entre medias de tratamiento para cualquier efecto principal o interacción es:

$$l_{AB} = \sum_i k_i \bar{y}_i \quad (11.2)$$

donde los coeficientes para los contrastes son $k_i = \pm 1$. Siguiendo la convención para calcular los contrastes, los errores estándar y las sumas de cuadrados dada en el capítulo 3, la estimación del efecto para cualquier contraste entre medias de tratamiento en un factorial 2^n completo se puede expresar como:

$$AB\dots = \frac{1}{2^{n-1}} (l_{AB\dots}) \quad (11.3)$$

La estimación del error estándar para una estimación de efecto es:

$$s_{AB\dots} = \sqrt{\frac{4\sigma^2}{r2^n}} \quad (11.4)$$

donde r es el número de réplicas para cada tratamiento. La suma de cuadrados con 1 grado de libertad para los efectos mostrados al final de la tabla 11.3 se obtiene con:

$$SC(AB\dots) = \frac{r}{2^n} (l_{AB\dots})^2 \quad (11.5)$$

Para el ejemplo, en la tabla 11.3, $r = 1$ porque cada combinación de tratamientos sólo ocurre una vez. La suma de cuadrados con 1 grado de libertad para A es:

$$SCA = \frac{1}{8}(-2)^2 = 0.5$$

mientras que para B y AB son:

$$SCB = \frac{1}{8}(-38)^2 = 180.5 \quad \text{y} \quad SC(AB) = \frac{1}{8}(10)^2 = 12.5$$

11.3 Diseños de bloque incompletos para factoriales 2ⁿ

El uso de diseños de bloque incompletos para reducir la varianza del error experimental se explica en el capítulo 9, pero puede no ser posible realizar una réplica completa para los factoriales 2ⁿ con muchos factores en un bloque completo. Si la materia prima no es suficiente en un lote de producción para manejar todos los tratamientos, se puede usar cada lote de materia prima como un bloque incompleto. Si el error experimental es demasiado grande con los diseños de bloque completos en experimentos agrícolas, la variación entre las parcelas se puede controlar con tamaños de bloque reducidos para tener grupos más homogéneos de parcelas experimentales. Se pueden elaborar bloques de tamaño reducido para factoriales 2ⁿ aprovechando la construcción de los contrastes de sus efectos.

Se sacrifica información de los tratamientos para aumentar la precisión

Los diseños de bloque incompletos para factoriales de 2ⁿ se construyen de manera que uno o más contrastes de tratamiento sean idénticos a los contrastes de bloque. Se dice que el efecto del tratamiento se **confunde** por completo con los bloques y no es posible distinguirlo del efecto de esos bloques.

En general, se elige el efecto de interacción de mayor orden en el factorial 2^n para confundirlo con los bloques. En el caso de experimentos con muchos factores, los efectos principales, las interacciones de dos factores y otras interacciones de menor orden son los efectos de mayor interés. Al confundir la interacción de orden más alto los demás efectos se estiman sin penalización.

La construcción de un diseño de bloque incompletos se ilustra con un factorial 2^3 . Un diseño de bloques completos requiere ocho unidades experimentales por bloque. La mitad de los tratamientos tiene coeficiente + y la mitad - para todo efecto. Se puede construir un diseño de bloque incompleto resoluble con dos bloques de cuatro unidades cada uno por réplica usando el contraste l_{ABC} .

A partir de la tabla 11.3, la interacción ABC se estima con la comparación:

$$l_{ABC} = abc + a + b + c - ab - ac - bc - (1)$$

Las combinaciones de tratamientos con coeficiente +, abc , a , b y c , se colocan en un bloque y las combinaciones con coeficiente -, ab , ac , bc y (1), en otro. En el cuadro 11.2 se muestran los dos bloques incompletos.

Cuadro 11.2 Interacción ABC confundida en dos bloques de cuatro unidades experimentales cada uno

<i>Bloque</i>	
<i>1</i>	<i>2</i>
<i>abc</i>	<i>ab</i>
<i>a</i>	<i>ac</i>
<i>b</i>	<i>bc</i>
<i>c</i>	(1)

<i>ABC</i>	+ 1	- 1
------------	-----	-----

La comparación (bloque 1 – bloque 2) es el contraste necesario para estimar la interacción ABC ; en consecuencia, la estimación del efecto de interacción ABC se confunde por completo con una comparación entre los bloques. No será posible estimar la interacción de tres factores independiente de los efectos de bloque.

Por otro lado, los otros seis efectos factoriales no se confunden con los bloques y se pueden estimar en la forma usual; por ejemplo, el efecto principal del factor A estimado con:

$$l_A = (abc + a - b - c) + (ab + ac - bc - (1))$$

contiene dos coeficientes + y dos - de las cuatro unidades en cada bloque. Cualquier diferencia entre los bloques no afectará la estimación.

Construcción de bloques con la regla pares-impares

Como la mitad de los tratamientos tiene coeficiente + y la mitad coeficiente – para todo efecto en los factoriales 2ⁿ, los tratamientos se pueden dividir en dos grupos con la regla *pares-impares*.

Cualquier combinación de tratamientos que tenga un número par de letras del efecto factorial recibe uno de los coeficientes (+ o –), mientras los tratamientos con un número impar de letras del efecto factorial reciben el otro coeficiente.

Si la interacción *ABC* tiene asociadas las letras *a*, *b* y *c*, los tratamientos (1), *ab*, *ac* y *bc* tienen un número par de letras del efecto factorial *ABC*, el tratamiento (1) con cero letras tiene un número par de letras, ya que cero se considera un número par y los tratamientos *a*, *b*, *c* y *abc* tienen un número impar de letras. Las combinaciones de tratamiento con coeficiente + para el efecto de interacción *ABC* son:

$$a, b, c, \text{ y } abc$$

y las combinaciones de tratamiento con coeficiente – para el efecto de interacción *ABC* son:

$$(1), ab, ac, \text{ y } bc$$

La combinación de tratamiento que contiene todas las letras de factores, *abc* en este caso, siempre tendrá coeficiente + para cualquier efecto factorial y todas las combinaciones de tratamiento en ese grupo también tendrán coeficiente +, mientras que las combinaciones del otro grupo tendrán coeficiente –. Esta asignación de coeficientes + y – para los contrastes coincide con la que se muestra en la tabla 11.3.

Suponiendo que la interacción *ABCD* de un factorial 2⁴ se confunde con los bloques en un experimento con dos bloques de ocho unidades experimentales cada uno en cada réplica; las letras asociadas con los efectos de interacción de cuatro factores *ABCD* son *a*, *b*, *c* y *d*. Entonces, las combinaciones de tratamientos con un número par de letras son:

$$(1), ab, ac, ad, bc, bd, cd, \text{ y } abcd$$

y las combinaciones de tratamientos con un número impar de letras son:

$$a, b, c, d, abc, abd, acd, \text{ y } bcd$$

Los tratamientos en el grupo que contiene la combinación *abcd* reciben un coeficiente +, ya que esa combinación contiene las letras de los cuatro factores. El otro grupo de tratamientos tiene coeficiente – para el efecto de interacción *ABCD* y se colocan en un bloque separado del primer grupo.

Descripción del análisis de varianza para un diseño confundido por completo

Las sumas de cuadrados se calculan de la manera usual para el análisis de varianza excepto que se excluye la partición de la suma de cuadrados para el efecto de interacción confundido con los bloques. La suma de cuadrados del bloque incluirá el efecto factorial confundido.

Las fuentes de variación y los grados de libertad para el análisis de varianza para el factorial 2^3 con $b = 2$ bloques incompletos en cada una de $r = 2$ réplicas se describen en la tabla 11.4. Como ABC se confunde con los bloques, la suma de cuadrados para los bloques incluye el efecto ABC .

Tabla 11.4 Análisis de varianza para un factorial 2^3 con $b = 2$ bloques incompletos en cada una de $r = 2$ grupos de réplicas

Fuente de variación	Grados de libertad
Réplicas	$r - 1 = 1$
Bloques dentro de las réplicas	$r(b - 1) = 2$
Tratamientos	6
<i>A</i>	1
<i>B</i>	1
<i>C</i>	1
<i>AB</i>	1
<i>AC</i>	1
<i>BC</i>	1
Error	6
Total	15

Se conserva parte de la información de tratamientos con confusión parcial

En la sección anterior se redujo el tamaño de bloque mediante la confusión la interacción de orden más alto con los bloques, pero la ganancia obtenida al reducir el error experimental tiene un precio. En los diseños anteriores, la pérdida de toda la información sobre el efecto factorial confundido fue el costo de una posible reducción en la varianza del error experimental.

Como cualquier efecto factorial se puede confundir con los bloques, para evitar la pérdida de toda la información sobre un efecto factorial se puede confundir un efecto diferente en cada grupo de réplicas del diseño resoluble. De esta manera, un efecto factorial sólo se confunde en una réplica, por lo que se dice que está *parcialmente confundido* con los bloques.

El ejemplo 11.2 usa un factorial 2^3 con un efecto distinto confundido en cada una de las tres réplicas para ilustrar el principio.

Ejemplo 11.2 Confusión parcial en un factorial 2^3

Se pensó que la pureza de un producto químico tenía influencia de tres factores, tasa de agitación (*A*), concentración del compuesto base (*B*) y concentración del reactivo (*C*). El químico estableció un experimento con un diseño factorial con factores a dos niveles para un arreglo factorial 2^3 .

Diseño del experimento: El químico deseaba tres réplicas del experimento, pero sólo podía realizar cuatro corridas del proceso químico en un día. Por lo tanto, cada réplica debía correrse en dos bloques incompletos (días).

Un contraste de efecto para un factorial de 2³ consiste en cuatro combinaciones de tratamientos con coeficiente + y cuatro con -. Así, fue posible construir el diseño de bloque incompleto confundiendo un efecto del factorial 2³ con los bloques. Para evitar la total confusión de un efecto, el químico confundió una interacción de dos factores distinta en cada réplica. Las combinaciones de tratamiento necesarias para cada bloque se muestran en el diagrama de la tabla 11.5.

Tabla 11.5 Pureza observada de un producto químico en un factorial 2³ parcialmente confundido

<i>BC Confundido</i>		<i>AC Confundido</i>		<i>AB Confundido</i>			
+ 1	- 1	+ 1	- 1	+ 1	- 1		
(1) 25	ab 43	abc 39	bc 38	(1) 26	a 43		
bc 34	c 30	b 29	a 37	c 32	b 34		
abc 42	ac 40	(1) 27	ab 46	ab 52	ac 40		
a 25	b 33	ac 40	c 34	abc 51	bc 36		
Bloque 1		Bloque 2		Bloque 3			
Réplica I		Bloque 4		Bloque 5			
Réplica II			Bloque 6				
Réplica III							

Cada una de las interacciones *BC*, *AC* y *AB* se confundieron en una réplica del experimento, la interacción de dos factores confundida con los dos bloques en cada réplica se muestran encima de los bloques y las tasas de pureza observadas están a la derecha de las combinaciones de tratamiento en cada bloque.

Las combinaciones de tratamiento necesarias en cada uno de los bloques incompletos se pueden determinar con la regla de pares-impares. En la réplica I la interacción *BC* se confunde al colocar combinaciones de tratamiento con un número par de las letras *b* y *c* juntas en el bloque 1, que son (1), *a*, *bc* y *abc*; los tratamientos con un número impar de las letras *b* y *c* se colocan en el bloque 2, éstos son *c*, *b*, *ab* y *ac*; los tratamientos en el bloque 1 tienen coeficiente +1 para el efecto de interacción *BC* porque la combinación de tratamiento *abc* contiene todas las letras correspondientes a los factores y los tratamientos en el bloque 2 tienen coeficiente -1 para el efecto de interacción *BC*. En consecuencia, la diferencia entre las observaciones de los bloques 1 y 2 tendrá la estimación del efecto de interacción *BC* confundido con la diferencia entre los efectos de esos bloques. Las combinaciones de tratamientos en cada bloque de réplicas II y III se puede determinar de manera similar con el efecto *AC* confundido en la réplica II y el efecto *AB* confundido en la réplica III.

Cálculo de sumas de cuadrados para diseños parcialmente confundidos

Las sumas de cuadrados para los efectos no confundidos con bloques, A , B , C y ABC , se pueden calcular de manera normal mediante la ecuación (11.5), utilizando los coeficientes de los contrastes de la tabla 11.3. Los efectos se pueden estimar en todas las réplicas, de esta manera la suma de cuadrados es $SC = r(\sum k_i \bar{y}_i)^2 / 2^n$, donde $r = 3$, $2^n = 8$ y $k_i = \pm 1$. Los cálculos se muestran en la tabla 11.6.

Tabla 11.6 Cálculos para los efectos factoriales no confundidos con bloques

Medias:	(1)	a	b	ab	c	ac	bc	abc	$\sum k_i \bar{y}_i$	2^n	SC
	26	35	32	47	32	40	36	44			
A	—	+	—	+	—	+	—	+	40	8	600.00
B	—	—	+	+	—	—	+	+	26	8	253.50
C	—	—	—	—	+	+	+	+	12	8	54.00
ABC	—	+	+	—	+	—	—	+	- 6	8	13.50

La suma de cuadrados para cada interacción de dos factores parcialmente confundida debe calcularse a partir de la réplica en la que no está confundida. La suma de cuadrados para AB debe calcularse con las réplicas I y II, para AC con las réplicas I y II y para BC con las réplicas II y III; para estos cálculos se usan los totales en lugar de las medias.

La suma de cuadrados se puede calcular de manera conveniente si se determina el efecto del contraste a partir de todas las observaciones y se resta el valor del contraste, representado por la diferencia de los totales de bloques para las réplicas en las que el efecto está confundido. Los tres contrastes confundidos, AB , AC y BC , calculados a con los totales de todas las observaciones de tratamientos son:

Totales de tratamiento	(1)	a	b	ab	c	ac	bc	abc	$\sum k_i y_i$
	78	105	95	141	96	120	108	132	
AB	+	—	—	+	+	—	—	+	18
AC	+	—	+	—	—	+	—	+	- 24
BC	+	+	—	—	—	—	+	+	- 30

Los totales y la diferencia calculada para cada par de bloques son:

	<i>Réplica I</i>		<i>Réplica II</i>		<i>Réplica III</i>	
	<i>BC confundido</i>		<i>AC confundido</i>		<i>AB confundido</i>	
	<i>Bloque 1</i>	<i>Bloque 2</i>	<i>Bloque 3</i>	<i>Bloque 4</i>	<i>Bloque 5</i>	<i>Bloque 6</i>
Totales	126	146	135	155	161	153
Diferencia	- 20		- 20		8	

Para estimar el efecto del contraste AB sólo a partir de las réplicas I y II, la diferencia entre los bloques 5 y 6 (8) se resta del contraste AB obtenido de los totales de las observaciones (18). La estimación corregida del contraste es:

$$l_{AB} = \frac{18 - 8}{2} = 5$$

donde el divisor es $r = 2$ para poner el resultado como una media. La suma de cuadrados para la interacción AB es:

$$SC(AB) = \frac{2(5)^2}{8} = 6.25$$

De manera similar, la estimación del contraste AC es:

$$l_{AC} = \frac{-24 - (-20)}{2} = -2$$

con suma de cuadrados:

$$SC(AC) = \frac{2(-2)^2}{8} = 1$$

Por último:

$$l_{BC} = \frac{-30 - (-20)}{2} = -5$$

y

$$SC(BC) = \frac{2(-5)^2}{8} = 6.25$$

El análisis de varianza se muestra en la tabla 11.7. Se trata de un diseño de bloques incompletos resoluble con bloques de tratamientos en grupos de réplicas completas. La suma de cuadrados total y las sumas de cuadrados para las réplicas y los bloques dentro de las réplicas se calculan de manera normal, en tanto que la suma de cuadrados para el error experimental se encuentra con las sumas de cuadrados para todos los efectos de tratamiento, réplicas y bloques dentro de las réplicas restadas de la suma de cuadrados total.

Pruebas de hipótesis sobre los efectos de un factor

El valor crítico para una prueba de hipótesis para cualquier efecto factorial es $F_{05,1,11} = 4.84$; por tanto, los efectos principales de la tasa de agitación (A) y la concentración del compuesto base (B) son significativos, ninguno de los demás lo fueron en cuanto a la pureza del producto químico.

De las ecuaciones (11.3) y (11.4), el efecto estimado de la tasa de agitación es:

$$A = \frac{(40)}{4} = 10$$

Tabla 11.7 Análisis de varianza para la pureza de un producto químico en un factorial 2^3 parcialmente confundido

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F_0
Réplicas	2	111.00	55.50	
Bloques dentro de las réplicas	3	108.00	36.00	
<i>A</i>	1	600.00	600.00	40.6
<i>B</i>	1	253.50	253.50	17.2
<i>C</i>	1	54.00	54.00	3.7
<i>ABC</i>	1	13.50	13.50	< 1
<i>AB</i> (Réps I, II)	1	6.25	6.25	< 1
<i>AC</i> (Réps I, III)	1	1.00	1.00	< 1
<i>BC</i> (Réps II, III)	1	6.25	6.25	< 1
Error	11	162.50	14.77	
Total	23	1,316.00		

con error estándar:

$$s_A = \sqrt{\frac{4(14.77)}{3(8)}} = 1.57$$

El efecto estimado de la concentración del compuesto base es:

$$B = \frac{26}{4} = 6.5$$

con un error estándar $s_B = 1.57$. La pureza del producto químico aumenta en 10 unidades si se incrementa la tasa de agitación del nivel bajo al alto y en 6.5 unidades si lo hace la concentración del compuesto base. La concentración del agente C no tiene efecto significativo sobre la pureza del producto y no hubo interacción significativa entre el resto de los factores.

Confusión de otra interacción para reducir aún más el tamaño de bloque

El número de unidades experimentales por bloque se puede reducir más si se elige un segundo efecto factorial para confundirlo con los bloques. El diseño tendrá cuatro bloques por réplica si se confunden dos efectos factoriales.

A continuación se explica la técnica con el factorial 2^3 para producir cuatro bloques de dos unidades experimentales. Si se eligen los efectos de interacción *AB* y *AC* como efectos confundidos, la asignación de coeficientes + y - a los tratamientos según la regla pares-impares es:

	(1)	abc	c	ab	b	ac	a	bc
AB	+	+	+	+	-	-	-	-
AC	+	+	-	-	+	+	-	-

Se forman cuatro bloques de dos unidades experimentales con pares de tratamientos que tienen la misma configuración de coeficientes + y - para las interacciones AB y AC .

Existen cuatro configuraciones de coeficientes + y - para los contrastes de interacción conjuntos AB y AC , cada una contiene dos tratamientos que se asignan a un bloque distinto. Las configuraciones y las asignaciones de tratamientos a los bloques son:

AB	AC	Tratamiento	Bloque
+	+	(1) abc	1
+	-	c ab	2
-	+	b ac	3
-	-	a bc	4

Un tercer efecto se confunde de manera automática

Si se confunden dos efectos con bloques de un factorial 2ⁿ, entonces se confunde un tercer efecto. Se tienen 3 grados de libertad para las comparaciones de bloques y se confunden tres tratamientos con los bloques. La interacción AB se confunde con el contraste de bloque:

$$AB = l_1 = B_1 + B_2 - B_3 - B_4$$

donde B_1 , B_2 , B_3 y B_4 representan las medias de los bloques. La interacción AC se confunde con el contraste de bloque:

$$AC = l_2 = B_1 - B_2 + B_3 - B_4$$

El tercer contraste entre los bloques que es ortogonal a $AB = l_1$ y $AC = l_2$ es:

$$\begin{aligned} l_3 &= B_1 - B_2 - B_3 + B_4 \\ &= (1) + abc - c - ab - b - ac + a + bc \end{aligned}$$

Si se verifican los contrastes de la tabla 11.3 se puede observar que el tercero es igual en la interacción BC .

El tercer efecto confundido es una interacción generalizada

El tercer efecto confundido se conoce como la **interacción generalizada** de los dos primeros efectos confundidos, ésta se obtiene formando un producto de símbolos de los dos efectos y tachando la letras que aparecen dos veces en el producto.

Por ejemplo, AB y AC se eligieron como efectos confundidos y su producto simbólico es $ABAC$; si se tacha la letra A , que aparece dos veces, el resultado es $ABAC = \cancel{A}BAC = BC$ la interacción generalizada es BC . Si se eligen ABC y AB como efectos confundidos, la interacción generalizada es $ABCAB = \cancel{ABC}\cancel{A}B = C$ se observa que el efecto principal C está confundido con los bloques junto con ABC y AB . Debe evitarse confundir efectos que son de interés particular en el estudio, en especial los efectos principales.

En el apéndice 11A se proporciona una tabla que ayuda en la construcción de diseños de bloque incompletos útiles para factoriales 2^n . Un método general de confusión que se puede usar con los sistemas factoriales distintos a la serie 2^n se presenta en la siguiente sección.

11.4 Método general para crear bloques incompletos

La asignación de combinaciones de tratamientos a bloques incompletos para los factoriales 2^n se ha logrado hasta ahora con base en los signos \pm de los efectos confundidos con bloques. Sin embargo, el método llega a ser enfadoso cuando se tienen muchos factores, debido al gran número de combinaciones de tratamientos existentes al definir los contrastes. Además, el uso de los signos \pm en los sistemas confundidos no se amplía a otros sistemas como el de factoriales 3^n .

Un método general para construir diseños de bloques incompletos donde se eligen efectos factoriales confundidos con bloques usa las matemáticas de la función *residual módulo m* o *residuo mod m*. Para un entero k el residuo mod m es el residuo de dividir k entre m ; el residuo r para el entero k mod m se escribe como $k = r(\text{mod } m)$.

Con factoriales 2^n se trabaja con los residuos de $(\text{mod } 2)$, porque cualquier entero dividido entre $m = 2$ deja un residuo de 0 o 1; es decir, los valores de los residuos $(\text{mod } 2)$ son 0 y 1: los enteros pares $(\text{mod } 2)$ tienen un residuo 0 y los enteros impares $(\text{mod } 2)$ tienen 1. Por ejemplo, $7 = 1(\text{mod } 2)$ porque 7 entre 2 es 3 con residuo 1; también, $4 = 0(\text{mod } 2)$ ya que 4 entre 2 es 2 y el residuo es 0.

Los niveles de un factor (0 o 1) se usan como valores de una variable x_i que representa al i -ésimo tratamiento, una combinación de tratamiento para el factorial 2^n se presenta como una secuencia $x_1x_2x_3\dots x_n$ y, para un factorial 2^3 , los ocho tratamientos escritos en orden estándar son:

x_1	x_2	x_3	Tratamiento
0	0	0	(1)
1	0	0	a
0	1	0	b
1	1	0	ab
0	0	1	c
1	0	1	ac
0	1	1	bc
1	1	1	abc

Definición de contrastes para dos bloques

Un método general para determinar la asignación de contrastes de tratamiento a los bloques incompletos se logra mediante una función lineal:

$$L = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \quad (11.6)$$

donde L es el *contraste de definición*, o el contraste confundido con los bloques; el valor de α_i es 1 si en el contraste de definición está presente el factor i -ésimo y 0 si no lo está, la función del contraste de definición L , se evalúa para cada combinación de tratamiento; el valor de x_i es el nivel del i -ésimo factor (0 o 1) en cualquier combinación de tratamiento considerando para asignarla a un bloque incompleto.

Suponiendo que el contraste de definición para un factorial 2^3 en dos bloques de cuatro unidades cada uno es la interacción de dos factores AB . Los factores A y B están presentes en el contraste de definición de manera que $\alpha_1 = 1$ y $\alpha_2 = 1$, pero como el factor C no está en el contraste, $\alpha_3 = 0$. La función contraste de definición para la asignación de tratamientos es:

$$L = x_1 + x_2 \quad (11.7)$$

Los valores de x_1 y x_2 para cada combinación de tratamientos se sustituye en lugar de L y se determina el residuo para L módulo 2, $r(\text{mod } 2)$. Por ejemplo, el valor de $L = x_1 + x_2$ para la combinación de tratamientos bc , $x_1x_2x_3 = (011)$, es $L = 0 + 1$ y el residuo es $1(\text{mod } 2)$. Los valores de L y los residuos, $L = r(\text{mod } 2)$, son:

Tratamiento	x_1	x_2	$L = x_1 + x_2$	$r(\text{mod } 2)$
(1)	0	0	0	0
a	1	0	1	1
b	0	1	1	1
ab	1	1	2	0
c	0	0	0	0
ac	1	0	1	1
bc	0	1	1	1
abc	1	1	2	0

Las combinaciones de tratamientos con $L = 0 (\text{mod } 2)$ se asignan a un bloque y las que tienen $L = 1(\text{mod } 2)$ al otro, de manera que las asignaciones de tratamientos son:

Bloque 1	(1) ab c abc	con $L = 0 (\text{mod } 2)$
Bloque 2	a b ac bc	con $L = 1 (\text{mod } 2)$

Sean B_1 y B_2 los totales de bloque. El bloque 1, con residuo 0 para L , contiene las combinaciones de tratamientos con signo + para el contraste AB , mientras que el bloque 2, con residuo 1 para L , contiene las combinaciones de tratamientos con signo -. Así, el contraste 1 entre los totales de bloques, que equivale al contraste AB , es:

$$l_{AB} = B_1 - B_2$$

Uso de dos contrastes de definición para cuatro bloques

Hasta ahora, los diseños se han construido para factoriales 2^n con dos bloques incompletos de 2^{n-1} unidades experimentales, cada una con un contraste de definición confundido con los bloques. La reducción del tamaño al usar dos bloques con un efecto factorial confundido puede no ser suficiente para un factorial 2^n cuando se tienen muchos factores, digamos $n \geq 4$. Por ejemplo, un factorial 2^5 con 32 tratamientos puede requerir bloques con menos de ocho unidades y con cuatro bloques por réplica. Se puede lograr una mayor reducción si se confunde con los bloques un contraste de definición adicional.

Suponga que en un factorial 2^4 se necesita tener las 16 combinaciones de tratamiento colocadas en cuatro bloques de $2^{4-2} = 2^2 = 4$ unidades experimentales cada una, se requieren dos contrastes de definición para construir los cuatro bloques incompletos. Si se eligen AB y CD para confundirse con los bloques, entonces los contrastes de definición son:

$$L_1 = x_1 + x_2 \quad (11.8)$$

$$L_2 = x_3 + x_4$$

donde L_1 representa a AB y L_2 a CD . Cada combinación de tratamiento proporcionará un par de residuos módulo 2 para el par (L_1, L_2) , pero como existen cuatro pares de residuos posibles, $(0,0)$, $(0,1)$, $(1,0)$ y $(1,1)$, las combinaciones de tratamientos con los mismos valores para un par de residuos módulo 2 se colocan en el mismo bloque incompleto. Los valores de los contrastes de definición, L_1 y L_2 en la ecuación (11.8) y los pares de residuos para cada una de las 16 combinaciones de tratamientos se muestran en el cuadro 11.3 junto con las asignaciones de bloques.

Los 3 grados de libertad entre los bloques representan tres contrastes ortogonales entre bloques, se sabe que dos de ellos incluyen las interacciones de dos factores, AB y CD , elegidas para confundirse con los bloques. El contraste para la interacción AB es la diferencia entre los bloques con residuo 0 y 1 para L_1 . Sean B_1 , B_2 , B_3 y B_4 los totales de bloque, entonces el contraste para AB es:

$$l_{AB} = B_1 + B_2 - B_3 - B_4$$

porque los bloques 1 y 2 tienen residuo 0 y los bloques 3 y 4 tienen residuo 1 para L_1 . De manera similar, los bloques con residuos 0 y 1 para L_2 definen el contraste para la interacción CD , cuyo contraste de bloque es:

$$l_{CD} = B_1 - B_2 + B_3 - B_4$$

porque los bloques 1 y 3 tienen residuo 0 y los bloques 2 y 4 tienen residuo 1 para L_2 .

Cuadro 11.3 Diseño de bloques incompletos para un factorial 2⁴ en cuatro bloques de cuatro unidades con contrastes de definición AB y CD

Bloque	Tratamiento	x_1	x_2	x_3	x_4	L_1	L_2	Residuo
1	(1)	0	0	0	0	0	0	(0, 0)
	ab	1	1	0	0	2	0	(0, 0)
	cd	0	0	1	1	0	2	(0, 0)
	$abcd$	1	1	1	1	2	2	(0, 0)
2	c	0	0	1	0	0	1	(0, 1)
	d	0	0	0	1	0	1	(0, 1)
	abc	1	1	1	0	2	1	(0, 1)
	abd	1	1	0	1	2	1	(0, 1)
3	a	1	0	0	0	1	0	(1, 0)
	b	0	1	0	0	1	0	(1, 0)
	acd	1	0	1	1	1	2	(1, 0)
	bcd	0	1	1	1	1	2	(1, 0)
4	ac	1	0	1	0	1	1	(1, 1)
	bc	0	1	1	0	1	1	(1, 1)
	ad	1	0	0	1	1	1	(1, 1)
	bd	0	1	0	1	1	1	(1, 1)

$$L_1 = x_1 + x_2 \quad L_2 = x_3 + x_4$$

Determinación de la interacción generalizada

La interacción $ABCD$ es la *interacción generalizada* confundida como consecuencia de confundir deliberadamente AB y CD con los bloques. Las interacciones generalizadas se pueden determinar mediante reglas algebraicas más formales que las descritas en la sección anterior. Se forma el producto de los símbolos para los contrastes de definición con el exponente de cualquier símbolo reducido por módulo 2. El producto de AB y CD es:

$$AB \times CD = ABCD$$

La interacción generalizada de AB y CD determinada por el producto de símbolos es $ABCD$, ya que todos los exponentes del producto son 1(mod 2).

Suponiendo que se eligieron los contrastes ABC y BCD para los contrastes de definición originales, el producto es $ABC \times BCD = ABBCCD = AB^2C^2D = AD$, donde B^2 y C^2 no aparecen en la interacción generalizada porque sus exponentes son 0(mod 2). Por lo tanto, la interacción generalizada es AD cuando los contrastes de definición son ABC y BCD .

Como los efectos principales y las interacciones de dos factores son de mayor interés en los estudios preliminares de los efectos de un factor, se intenta evitar confundirlos con los bloques. El número de unidades experimentales en un diseño de bloque incompleto para factoriales 2^n es igual a las potencias de 2, es decir, 2,

4, 8, etcétera, hasta 2^{n-1} . En general, si se requieren bloques de tamaño $k = 2^q$, habrá $2^n/2^q = 2^{n-q}$ bloques en una réplica completa; en consecuencia, deben elegirse $n - q$ contrastes de definición.

Considerando el factorial 2^7 con 128 combinaciones de tratamiento, un diseño con $2^4 = 16$ unidades experimentales por bloque requiere $2^{7-4} = 2^3 = 8$ bloques y $(7 - 4) = 3$ contrastes de definición. Si se eligen ABG , CDE y EFG como contrastes de definición, el primero, ABG , se usa para reducir los tamaños de bloques a 64 unidades, el segundo, CDE , se usa para reducir los tamaños de bloque a 32 unidades, y el tercero, EFG para reducirlos a 16 unidades. Habrá cuatro interacciones generalizadas también confundidas con los ocho bloques, que son:

$$ABG \times CDE = ABCDEG$$

$$ABG \times EFG = ABEFG^2 = ABEF$$

$$CDE \times EFG = CDE^2 FG = CDFG$$

y

$$ABG \times CDE \times EFG = ABCDE^2FG^2 = ABCDF$$

Como antes, las combinaciones de tratamiento se asignan a los bloques según los residuos de los contrastes de definición:

$$\begin{array}{rcl} ABG & \longrightarrow & L_1 = x_1 + x_2 + x_7 \\ CDE & \longrightarrow & L_2 = x_3 + x_4 + x_5 \\ EFG & \longrightarrow & L_3 = x_5 + x_6 + x_7 \end{array}$$

Cada contraste de tratamiento tendrá una terna de valores (L_1, L_2, L_3) y habrá ocho ternas de residuos, $(0, 0, 0)$, $(0, 0, 1)$, $(0, 1, 0)$, $(1, 0, 0)$, $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ y $(1, 1, 1)$, donde cada una representa la terna de residuo para una asignación de bloque a las combinaciones de tratamientos. Una combinación de tratamiento con la terna de residuo $(0, 0, 0)$ se asigna al primer bloque, una con la terna de residuo $(0, 0, 1)$ se asigna al segundo bloque, y así sucesivamente.

La tabla 11A.1 del apéndice enumera los contrastes de definición y sus interacciones generalizadas para construir los diseño de bloques incompletos para factoriales 2^n con tamaños de bloque $k \geq 4$ para $n = 4, 5, 6$ y 7 factores. Observe que algunas interacciones de dos factores se confundirán con los bloques cuando éstos tengan cuatro unidades o menos. Siempre es posible tener diseños con sólo tres factores y más interacciones confundidas cuando el bloque tiene ocho unidades o más. Si se eligen los contrastes adecuados será posible evitar confundir cualesquiera interacciones de dos factores o efectos principales.

11.5 Diseños de bloques incompletos para factoriales 3ⁿ

Los factoriales 2ⁿ son diseños útiles para detectar los factores con efectos principales en las respuestas medidas en un experimento. Los factoriales 3ⁿ tienen tres niveles para cada factor, lo que hace posible estimar las tendencias lineales y cuadráticas de los factores cuantitativos y proporcionar descripciones más detalladas de los efectos de factores cualitativos. Sin embargo, el número de unidades experimentales requerido por los factoriales 3ⁿ aumenta en potencias de 3 conforme se agregan factores, es entonces cuando los diseños de bloques incompletos pueden ser útiles para estos diseños de tratamientos. La construcción de diseños de bloques incompletos para factoriales 3ⁿ se estudia brevemente en esta sección.

Algunas bases para 3ⁿ

Notación para factoriales 3ⁿ

Los niveles de un factor se representan por $x_i = 0, 1, 2$. Por ejemplo, las nueve combinaciones de tratamientos para un factorial de 3² con factores A y B son:

		<i>A</i>			
		0	1	2	
		0	00	10	20
<i>B</i>		1	01	11	21
		2	02	12	22

Las tres columnas del arreglo representan las combinaciones de tratamiento para los niveles del factor A y los tres renglones representan las combinaciones de tratamiento para los niveles del factor B .

Se requieren tres bloques incompletos para factoriales confundidos 3ⁿ

La construcción de diseños de bloques incompletos para factoriales 3ⁿ requiere tres bloques para tener bloques de tamaños iguales. Habrá 2 grados de libertad entre bloques y el efecto del tratamiento con 2 grados de libertad debe confundirse con los bloques.

Los factoriales 3ⁿ tienen 2 grados de libertad para los efectos principales, 2² grados de libertad para las interacciones de dos factores, etcétera. Como no se desea confundir los efectos principales con los bloques, se hace una partición de las interacciones en dos componentes ortogonales, cada una con 2 grados de libertad. La componente ortogonal puede entonces usarse para el contraste de definición en la construcción del diseño de bloques incompletos.

Construcción de bloques incompletos con contrastes de definición

La asignación de combinaciones de tratamiento a los bloques utiliza la función de contrastes de definición $L = \alpha_1x_1 + \alpha_2x_2 + \cdots + \alpha_nx_n$ descrita en la sección 11.4. El valor de x_i para el factorial 3ⁿ es $x_i = 0, 1, 2$. Por conveniencia matemática, se

usan los valores de $x_i = 0, 1, 2$ para los factores incluidos en las interacciones con 2 grados de libertad que se usaron como contrastes de definición. Por ejemplo, con dos factores, A y B , la interacción AB con 4 grados de libertad se divide en contrastes de definición etiquetados AB y AB^2 , cada uno con 2 grados de libertad. Por convención, la primer letra de la expresión de interacción siempre tiene potencia 1 de manera que A^2B o A^2B^2 no se usan como expresiones de contrastes de definición. Las potencias de A y B son los coeficientes α_1 y α_2 en la función de contrastes de definición $L = \alpha_1x_1 + \alpha_2x_2$.

Los residuos módulo 3 de $L = \alpha_1x_1 + \alpha_2x_2 + \dots + \alpha_nx_n$ se determinan para factoriales 3^n , ya que existen tres niveles para cada factor. Los residuos $r(\text{mod } 3)$ son 0, 1 y 2. La función de contrastes de definición para AB con $\alpha_1 = \alpha_2 = 1$ es:

$$L_1 = x_1 + x_2$$

y para AB^2 con $\alpha_1 = 1$ y $\alpha_2 = 2$:

$$L_2 = x_1 + 2x_2$$

Los tratamientos se asignan a los bloques de acuerdo con el valor de los residuos para L_1 o L_2 , de manera que los tratamientos con $L_i = 0 \pmod{3}$ se asignan al primer bloque, aquellos con $L_i = 1 \pmod{3}$ se asignan al segundo bloque y los que tienen $L_i = 2 \pmod{3}$ al tercero. La práctica usual es asignar la mitad de las réplicas de acuerdo con cada una las funciones del contraste de definición. Las asignaciones de tratamientos para L_1 y L_2 se muestran en el cuadro 11.4.

Cuadro 11.4 Diseño de bloques incompleto para un factorial 3^2 con las componentes AB o AB^2 confundidas con los bloques

AB confundido

$$L_1 = x_1 + x_2$$

Bloque 1

Bloque 2

Bloque 3

00	12	21
----	----	----

01	10	22
----	----	----

02	11	20
----	----	----

$$L_1 = 0 \pmod{3}$$

$$L_1 = 1 \pmod{3}$$

$$L_1 = 2 \pmod{3}$$

AB² confundido

$$L_2 = x_1 + 2x_2$$

Bloque 1

Bloque 2

Bloque 3

00	11	22
----	----	----

02	10	21
----	----	----

01	12	20
----	----	----

$$L_2 = 0 \pmod{3}$$

$$L_2 = 1 \pmod{3}$$

$$L_2 = 2 \pmod{3}$$

Los residuos de L_1 y L_2 para las nueve combinaciones de tratamientos son:

x_1	x_2	L_1	$r(\text{mod } 3)$	L_2	$r(\text{mod } 3)$
0	0	0	0	0	0
0	1	1	1	2	2
0	2	2	2	4	1
1	0	1	1	1	1
1	1	2	2	3	0
1	2	3	0	5	2
2	0	2	2	2	2
2	1	3	0	4	1
2	2	4	1	6	0

Confusión con tres factores o más

El factorial 3³ requiere de 27 unidades experimentales para una sola réplica. Un diseño de bloques incompleto con tres bloques de nueve unidades experimentales se puede construir si se confunde con los bloques una componente de interacción de tres factores. La interacción de tres factores con 8 grados de libertad tiene cuatro componentes, cada una con 2 grados de libertad. Con el propósito de obtener un contraste de definición, las cuatro componentes se denominan ABC , ABC^2 , AB^2C y AB^2C^2 , el contraste de definición para cada una es:

$$\begin{array}{lll} ABC & \longrightarrow & L = x_1 + x_2 + x_3 \\ ABC^2 & \longrightarrow & L = x_1 + x_2 + 2x_3 \\ AB^2C & \longrightarrow & L = x_1 + 2x_2 + x_3 \\ AB^2C^2 & \longrightarrow & L = x_1 + 2x_2 + 2x_3 \end{array}$$

Se puede usar cualquier componente de la interacción de tres factores para generar tres bloques de nueve unidades experimentales cada uno. Si se usa la componente AB^2C , el contraste de definición es:

$$L = x_1 + 2x_2 + x_3$$

Los tres bloques se construyen con combinaciones de tratamientos $x_1x_2x_3$ que proporcionan los residuos para el contraste de definición de $L = 0 \pmod{3}$, $L = 1 \pmod{3}$ y $L = 2 \pmod{3}$, respectivamente.

Los residuos se determinan para cada una de las 27 combinaciones de tratamientos, 000, 001, 002, 010, ..., 222; por ejemplo, la combinación 021 tiene residuo $L = 0 + 2(2) + 1 = 2 \pmod{3}$. Cada una se coloca en su respectivo bloque con otros tratamientos que tienen los mismos residuos para el contraste de definición y se puede confundir una componente diferente de la interacción de tres factores en cada réplica del experimento.

Las componentes de interacción de cuatro factores se confunden con bloques en los factoriales 3^4 . Las ocho componentes, cada una con 2 grados de libertad, son $ABCD$, AB^2CD , ABC^2D , $ABCD^2$, AB^2C^2D , AB^2CD^2 , ABC^2D^2 y $AB^2C^2D^2$ y cualquiera de ellas se puede usar para bloquizar las 81 unidades experimentales en tres bloques de 27 unidades cada uno.

Reducir más el tamaño del bloque produce dos interacciones generalizadas

Después de reducir los tamaños de los bloques el número de bloques incompletos se conserva como un múltiplo de 3 en los factoriales 3^n un segundo contraste de definición se confunde con los bloques para reducir su tamaño de nueve a tres unidades experimentales. En el caso de factoriales 3^n , existen dos interacciones generalizadas para cada par de contrastes de definición.

Si se eligen dos contrastes de definición, digamos X y Y , entonces las interacciones generalizadas están dadas por los productos simbólicos XY y XY^2 . Los exponentes de estos productos se reducen con módulo 3 para obtener las interacciones generalizadas, suponiendo que se usan las componentes de interacción de dos factores $X = AB$ y $Y = AC^2$ como contrastes de definición.

El producto de símbolos XY es:

$$XY = AB \times AC^2 = A^2BC^2$$

Como el exponente del primer término, A , del producto de símbolos no es la unidad, después de la reducción módulo 3, el producto se eleva al cuadrado y los exponentes se reducen módulo 3 como sigue:

$$(A^2BC^2)^2 = A^4B^2C^4 = AB^2C$$

donde el último término se obtiene al reducir los exponentes del término anterior módulo 3.

El producto de símbolos XY^2 es:

$$XY^2 = AB \times (AC^2)^2 = AB \times A^2C^4 = A^3BC^4$$

y después de reducir los exponentes del último término con módulo 3, el resultado es:

$$XY^2 = BC$$

Al elegir los dos contrastes de definición, AB y AC^2 , también se confunden dos interacciones generalizadas, AB^2C y BC , con los nueve bloques de tres unidades experimentales cada uno. Debe tenerse cuidado al elegir los contrastes de definición para evitar confundir un efecto principal con los bloques. Por ejemplo, la elección de dos componentes de interacción de tres factores, ABC y AB^2C^2 , tiene como resultado una interacción generalizada de:

$$XY = ABC \times AB^2C^2 = A^2B^3C^3 \longrightarrow (A^2B^3C^3)^2 = A^4B^6C^6$$

Cuando los exponentes del cuadrado del producto simbólico se reducen con módulo 3, la interacción generalizada es el efecto principal de A .

Suponiendo que se seleccionan dos componentes, AB y AC^2 , para construir nueve bloques de tres unidades cada uno para el factorial de 3^3 , los contrastes de definición son:

$$L_1 = x_1 + x_2$$

y

$$L_2 = x_1 + 2x_3$$

Se obtienen nueve pares de residuos módulo 3 a partir de los dos contrastes de definición y se asocian tres combinaciones de tratamiento con cada par de residuos. Por ejemplo, los residuos $(L_1, L_2) = (2, 2)$ se presentan con las combinaciones (021), (112) y (200) y estas tres combinaciones se colocan en el mismo bloque.

11.6 Observaciones finales

Como se ha visto en este capítulo, los factoriales son diseños de tratamiento versátiles que se pueden adaptar a una variedad de condiciones experimentales de bloque. El resultado de usar diseños de bloques incompletos es que algunos efectos factoriales se confunden con los bloques, ya sea completa o parcialmente, y aumenta la complejidad del análisis de los resultados con los diseños parcialmente confundidos. El objeto de reducir el tamaño de los bloques es reducir la varianza del error experimental, si la reducción en la estimación de la varianza del error experimental es suficiente para superar la pérdida de información sobre los efectos de los factores confundidos, entonces vale la pena usar el diseño de bloques incompletos. Cochran y Cox (1957) presentan tablas de diseño de bloques incompletos para factoriales simétricos y asimétricos, y proporcionan algunos ejemplos del análisis con varios de ellos. En Kempthorne (1952), John (1987) y John y Williams (1995) se puede encontrar información detallada sobre los principios que fundamentan la construcción de diseños.

EJERCICIOS PARA EL CAPÍTULO 11

- Construya una réplica de un diseño de bloques incompletos para un factorial 2^4 en dos bloques de ocho unidades experimentales con $ABCD$ confundido con los bloques. Asigne al azar los tratamientos a las ocho unidades experimentales en cada bloque.
- Suponga un diseño de bloques incompletos para un factorial 2^5 donde los bloques no pueden tener más de ocho unidades experimentales.
 - ¿Cuántos bloques habrá?
 - ¿Cuántos contrastes de definición se requieren para una réplica del experimento?
 - ¿Cuántos efectos adicionales, interacciones generalizadas, se confundirán con los bloques?

- d. Elija algún(os) contraste(s) de definición para construir una réplica del diseño y determine qué otra(s) interacción(es), si la(s) hay, también se confundirán con los bloques.
3. Se requieren cuatro réplicas de un experimento factorial 2^4 . El diseño debe realizarse con bloques incompletos de ocho unidades experimentales. Construya un diseño en el que ningún efecto se confunda por completo con los bloques.
4. Debe realizarse un experimento factorial 2^7 en ocho bloques de 16 unidades experimentales que usan $ABCD$, $ABEF$ y EFG como contrastes de definición. ¿Qué otros efectos se confunden con los bloques?
5. Suponga que se realizó un experimento factorial 2^6 en ocho bloques de ocho unidades experimentales cada uno, y que se usaron $BCDE$, $ABDE$ y ADE como contrastes de definición.
- ¿Qué otros efectos se confundieron con los bloques?
 - ¿Pudo haber una mejor elección de los contrastes de definición para este diseño? Explique.
 - Si decidió que había una mejor elección, entonces seleccione un conjunto que mejore el diseño y defiéndalo.
6. Un científico de animales realizó un estudio de los efectos del estrés por calor y la dieta de proteína y agua salina en ratones de laboratorio; los tres factores se usaron en dos niveles con un arreglo factorial 2^3 . Los niveles de los factores fueron (A) proteína (alta, baja); (B) agua (normal, salina); (C) estrés por calor (temperatura normal, calor estresante). Se usó un diseño de bloques incompletos con cuatro ratones de una camada en cada bloque, cada ratón se colocó en una jaula individual y se le asignó un tratamiento. Una réplica del experimento consistió en dos camadas de ratones. En el siguiente cuadro se muestra el aumento de peso (gramos) para los ratones junto con la identificación del tratamiento.

<i>Camada 1</i>	<i>Camada 2</i>	<i>Camada 3</i>	<i>Camada 4</i>	<i>Camada 5</i>	<i>Camada 6</i>
(1) 27.5	<i>ab</i> 24.3	<i>bc</i> 19.5	<i>abc</i> 19.7	(1) 24.5	<i>a</i> 33.1
<i>bc</i> 20.6	<i>c</i> 24.3	<i>a</i> 24.1	<i>b</i> 19.5	<i>c</i> 23.0	<i>b</i> 20.5
<i>abc</i> 22.0	<i>ac</i> 22.8	<i>ab</i> 22.4	(1) 22.5	<i>ab</i> 23.4	<i>ac</i> 19.8
<i>a</i> 28.6	<i>b</i> 24.6	<i>c</i> 22.0	<i>ac</i> 18.8	<i>abc</i> 21.7	<i>bc</i> 18.5

Réplica I

Réplica II

Réplica III

- ¿Qué efectos de tratamiento se confunden con los bloques (camadas)?
 - Estime los efectos de factor y las interacciones y sus errores estándar; calcule el análisis de varianza para estos datos.
 - Interprete los resultados.
7. Se realizó un experimento para investigar los efectos de cuatro factores sobre la operación de un torno para metales, cada uno en dos niveles: (A) velocidad de rotación (60, 75); (B) ángulo de corte (30, 45); (C) frecuencia de lubricación (10 seg, 30 seg); (D) aleación para la punta de corte (1, 2), que se usaron en un arreglo factorial 2^4 para el experimento. Sólo podían realizarse ocho pruebas de corte por día, por lo que se estableció un diseño de bloque incompleto con dos bloques (días) en cada una de las dos réplicas, con $ABCD$ como contraste de definición. En el siguiente cuadro se muestra el desgaste de la punta de corte para cada tratamiento en cada bloque. Los días 1 y 2 constituyen la primera réplica y los días 3 y 4 la segunda.

Desgaste de la punta de corte

<i>Día 1</i>	<i>Día 2</i>	<i>Día 3</i>	<i>Día 4</i>
(1) 40	<i>a</i> 24	(1) 43	<i>a</i> 28
<i>ab</i> 33	<i>b</i> 31	<i>ab</i> 30	<i>b</i> 35
<i>ac</i> 31	<i>c</i> 27	<i>ac</i> 30	<i>c</i> 28
<i>bc</i> 38	<i>abc</i> 23	<i>bc</i> 32	<i>abc</i> 20
<i>ad</i> 22	<i>d</i> 48	<i>ad</i> 26	<i>d</i> 44
<i>bd</i> 37	<i>abd</i> 35	<i>bd</i> 33	<i>abd</i> 36
<i>cd</i> 49	<i>acd</i> 29	<i>cd</i> 40	<i>acd</i> 25
<i>abcd</i> 30	<i>bcd</i> 37	<i>abcd</i> 31	<i>bcd</i> 34

- a. Calcule el análisis de varianza para los datos.
- b. Calcule los efectos de factor y sus errores estándar.
- c. Interprete los resultados.
8. Proponga un diseño de bloques incompleto para un factorial 33 que tiene nueve unidades experimentales por bloque en cuatro réplicas de tres bloques cada una. Confunda una componente diferente de las interacciones de tres factores en cada réplica.
9. Un investigador necesita un diseño de bloques incompleto para un factorial 33 que tiene tres unidades experimentales en cuatro réplicas de tres bloques cada una. Considere los dos pares de contrastes de definición para construir el diseño:
- a. ABC, AB^2C
 b. AB, AC^2
- ¿Cuál es el mejor par de contrastes de definición para construir el diseño? Explique.

Apéndice 11A Planes para diseños de bloques incompletos para factoriales 2^n

Tabla 11A.1 Número de factores y bloques, tamaños de bloques, contrastes de definición e interacciones generalizadas para construir diseños de bloques incompletos con factoriales 2^n

Factores <i>n</i>	Bloques 2^{n-q}	Tamaño del bloque $k = 2^q$			
			Contrastes de definición	Interacciones generalizadas	
4	2	8	<i>ABCD</i>		
		4	<i>ABC, ABD</i>	<i>CD</i>	
5	2	16	<i>ABCDE</i>		
		8	<i>ABC, CDE</i>	<i>ABDE</i>	
		4	<i>ABC, ACD,</i> <i>ADE</i>	<i>BD, CE, ABE, BCDE</i>	
6	2	32	<i>ABCDEF</i>		
		16	<i>ABCD, CDEF</i>	<i>ABEF</i>	
	4	8	<i>ACE, ABEF,</i> <i>ABCD</i>	<i>ADF, BCF, BED, CDEF</i>	
		4	<i>ABF, ACF,</i> <i>CDF, DEF</i>	<i>AD, CE, BC, BE, AEF,</i> <i>BDF, ABCD, ABDE,</i> <i>ACDE, ABCEF, BCDEF</i>	
	16	4			
7		64	<i>ABCDEFG</i>		
		32	<i>ABCDE,</i> <i>ABEFG</i>	<i>CDFG</i>	
		16	<i>ABG, CDE,</i> <i>EFG</i>	<i>ABEF, CDFG</i>	
		8	<i>ABC, ADG,</i> <i>CDE, DEFG</i>	<i>ABCDF, ABCDEG</i>	
		4	<i>ABG, BCG,</i> <i>CDG, DEG,</i> <i>EFG</i>	<i>AEF, BDF, BEG, CFG,</i> <i>ABDE, ABFG, ACDF,</i> <i>ACEG, BCDG, BCEF,</i> <i>ABCDEFG</i>	
32	2	64		<i>AC, BD, BF, DG, CE,</i> <i>ADG, AFG, BEG, CFG,</i> <i>ABCD, ABEF, ABFG, ACDF,</i> <i>ABDE, ACEG, ADEF, BCDE</i>	
		32		<i>BCEF, CDEF, ABCEG,</i> <i>ABDFG, ACDEG, ACEFG,</i> <i>BCDFG, BDEFG, ABCDEFG</i>	

12 Diseños factoriales fraccionarios

En este capítulo se continúa estudiando la versatilidad de los diseños de tratamiento factoriales 2^n , se proponen experimentos que utilizan sólo una fracción del arreglo factorial como medio eficaz para obtener información de los factores durante las primeras etapas de investigación, se describen los métodos para construir los diseños y se explica el análisis para estimar y probar la significancia de los efectos en grandes experimentos factoriales 2^n sin réplicas completas.

12.1 Reducción del número de corridas experimentales con diseños fraccionarios

Los diseños de tratamiento factoriales 2^n son útiles para realizar estudios preliminares con muchos factores para identificar los más importantes y sus interacciones. Sin embargo, el número de unidades experimentales aumenta geométricamente con el número de factores en el estudio.

Los diseños **factoriales fraccionarios** usan sólo la mitad, la cuarta parte o incluso una fracción menor de las 2^n combinaciones de tratamientos y se usan por una o varias de las siguientes razones:

- El número de tratamientos necesario excede a los recursos
- Sólo se requiere información sobre los efectos principales y las interacciones de bajo orden
- Se necesitan estudios exploratorios para muchos factores
- Se hace la suposición de que sólo unos cuantos efectos son importantes

Algunos estudios industriales y de desarrollo pueden exceder la capacidad de las instalaciones de investigación si se incluyen en los experimentos todas las combinaciones de tratamientos para un factorial 2^n ; por ejemplo, una réplica completa

de un estudio 2^7 requiere 128 corridas de un proceso. Las observaciones de las 128 corridas proporcionan estimaciones de 7 efectos principales, 21 interacciones de dos factores, 35 interacciones de tres factores y 64 interacciones que incluyen cuatro factores o más. Las 128 combinaciones de tratamientos proporcionan una gran cantidad de información, quizás más de la necesaria para las interacciones de mayor orden.

Razones para usar factoriales fraccionados

Redundancia del diseño en ausencia de interacciones de alto orden

Las experiencias con estudios que involucran muchos factores han conducido a la observación de que, en algún punto, las interacciones de alto orden tienden a convertirse en despreciables y pueden ignorarse en el esquema general de las investigaciones preliminares. En la práctica, los efectos principales tienden a ser mayores que las interacciones de dos factores, que a su vez tienden a ser mayores que las interacciones de tres factores y así sucesivamente.

El factorial completo tiene cierto grado de redundancia si el investigador puede estar más o menos seguro de que las interacciones de alto orden son despreciables. En estas circunstancias, las estimaciones de efectos principales e interacciones de bajo orden se pueden obtener a partir de una fracción del diseño completo.

Hipótesis de diversidad de factores

El uso de diseños factoriales fraccionados en la investigación industrial (Diamond, 1989) o en biotecnología (Haaland, 1989) se basa primordialmente en la *hipótesis de diversidad de factores* (Box y Meyer, 1986). La diversidad de factores supone que una pequeña fracción de los efectos de factores es significativa para un proceso, mientras que el resto de los efectos es inerte para propósitos prácticos. Así, una gran parte de la variación se asocia con sólo unos cuantos factores.

Estudios exploratorios con factoriales fraccionados

Los factoriales fraccionados también son útiles en los estudios exploratorios que incluyen factores “principales” y “secundarios”. Los efectos e interacciones asociados con los factores principales son el objetivo primordial del estudio, pero el estudio puede incluir varios factores secundarios cuyos efectos deben revisarse aun cuando se espere que la mayoría o todos sean despreciables. En este capítulo se exploran las características, construcción y análisis de los factoriales fraccionados, comenzando en la siguiente sección con la mitad del experimento.

12.2 Fracción de un medio del factorial 2^n

El diseño con una fracción de un medio se conoce como diseño factorial fraccionado 2^{n-1} porque $\frac{1}{2}2^n = 2^{n-1}$; la notación indica que el diseño incluye n factores, cada uno en dos niveles que usan sólo 2^{n-1} unidades experimentales.

En el capítulo 11, cuando se colocó una réplica de un factorial 2^n en dos bloques incompletos, se usó un contraste de definición para separar las combinaciones de tratamiento en dos conjuntos y cada conjunto se colocó en uno de los bloques incompletos según los coeficientes + y - de las combinaciones de tratamiento; cada bloque era la mitad de una réplica completa de tratamiento y aunque el contraste de definición se confundía con los bloques, era posible estimar el resto de los efectos.

El mismo principio se usa para construir los diseños de factoriales fraccionados; media réplica del diseño consiste en todas las combinaciones de tratamientos con el coeficiente + en el contraste de definición.

Ejemplo 12.1 Revisión a la fabricación de muelles de camión

Recuerde que en el ejemplo 11.1 un muelle ensamblado de camión pasaba por un horno de alta temperatura, luego se ponía en una máquina de formado para producirle una curvatura manteniéndolo a alta presión por un periodo corto. Los factores y niveles para el factorial 2^3 eran A) temperatura del horno a 1840° o 1880°F , B) tiempo de calentamiento de 23 y 25 segundos y C) tiempo de transferencia de 10 o 12 segundos.

Las ocho combinaciones de tratamiento, las observaciones de la calidad del muelle y los coeficientes para estimar los efectos factoriales se muestran en la tabla 12.1. Observe que se dividieron los ocho tratamientos en dos grupos de cuatro usando el contraste de definición basado en la interacción de tres factores ABC , esta división de tratamientos pudo usarse para construir un diseño de bloque incompleto con la interacción ABC confundida con los bloques.

Tabla 12.1 Combinaciones de tratamientos necesarias para un diseño factorial 2^3

Tratamiento	Efecto factorial								y
	I	A	B	C	AB	AC	BC	ABC	
a	+	+	-	-	-	-	+	+	35
b	+	-	+	-	-	+	-	+	28
c	+	-	-	+	+	-	-	+	48
abc	+	+	+	+	+	+	+	+	29
(1)	+	-	-	-	+	+	+	-	32
ab	+	+	+	-	+	-	-	-	31
ac	+	+	-	+	-	+	-	-	39
bc	+	-	+	+	-	-	+	-	28

Si los ingenieros buscan construir un diseño factorial fraccionado de media réplica, usarían los cuatro tratamientos con coeficiente + de la mitad superior de la tabla 12.1 para el efecto factorial ABC .

Ventajas y desventajas de utilizar un factorial fraccionado.

Cada efecto tiene un alias

La ventaja de reducir el tamaño del experimento tiene un precio: la pérdida de parte de la información sobre algunos efectos. Si se usa la mitad de una réplica del factorial 2^n , se pierde la capacidad de estimar la interacción de tres factores y cada efecto principal se confunde con una interacción de dos factores. Los contrastes para los efectos principales para A , B y C son:

$$l_A = a - b - c + abc = 35 - 28 - 48 + 29 = -12$$

$$l_B = -a + b - c + abc = -35 + 28 - 48 + 29 = -26$$

$$l_C = -a - b - c + abc - 35 - 28 + 48 + 29 = 14$$

y los contrastes para los efectos de interacción BC , AC y AB son:

$$l_{BC} = a - b - c + abc = 35 - 28 - 48 + 29 = -12$$

$$l_{AC} = -a + b - c + abc = -35 + 28 - 48 + 29 = -26$$

$$l_{AB} = -a - b + c + abc - 35 - 28 + 48 + 29 = 14$$

Observe que $l_A = l_{BC}$ y que el contraste $l = (a - b - c + abc)$ estima el efecto combinado $A + BC$, por lo que es imposible diferenciar entre el efecto principal de la temperatura (A) y la interacción entre el tiempo de calentamiento y el tiempo de transferencia (BC). Dos efectos estimados por el mismo contraste se conocen como *alias*.

Existen relaciones similares en los contrastes lineales para los efectos principales de tiempo de calentamiento (B) y el tiempo de transferencia (C), $l_B = l_{AC}$ y $l_C = l_{AB}$. Por lo tanto, B y AC son alias y C y AB también. Las relaciones alias son:

$$A = BC$$

$$B = AC$$

$$C = AB$$

La interacción ABC usada como contraste de definición se conoce como *generador del diseño* y no se puede estimar en media réplica porque en la tabla 12.1 tiene un sólo coeficiente (+) que las estimaciones de la media o la columna identidad I . La relación de identidad:

$$I = ABC$$

se conoce como *relación de definición* para el diseño.

El diseño se genera con cualquiera de las mitades del generador

Es posible usar cualquier mitad del contraste ABC como generador del diseño; si se usan como media réplica, las combinaciones de tratamientos con coeficiente $-$ para la interacción ABC de la parte inferior de la tabla 12.1, la relación de definición para el diseño es $I = -ABC$. El contraste para el efecto principal de la temperatura (A) es:

$$l_A = [- (1) + ab + ac - bc]$$

El contraste para la interacción entre el tiempo de calentamiento y el de transferencia (BC) es:

$$l_{BC} = [(1) - ab - ac + bc]$$

Así, $l_A = -l_{BC}$; el efecto principal de la temperatura es el negativo del efecto de interacción entre el tiempo de calentamiento y el de transferencia.

De manera similar, los efectos principales del tiempo de calentamiento (B) y el tiempo de transferencia (C) son los negativos de las interacciones AC y AB , respectivamente, las relaciones entre los efectos principales y las interacciones de dos factores son $A = -BC$, $B = -AC$ y $C = -AB$. Sin importar qué mitad de diseño se use, los efectos principales son alias de las interacciones de dos factores.

Los alias de cualquier efecto se pueden determinar a partir de su interacción generalizada con el contraste de definición ABC (La determinación de las interacciones generalizadas se explica en el capítulo 11), mientras que los alias de los efectos principales se determinan a partir de sus respectivos productos con el contraste de definición, como sigue:

$$A \times ABC = A^2BC = BC$$

$$B \times ABC = AB^2C = AC$$

$$C \times ABC = ABC^2 = AB$$

Si se usara la relación de definición $I = -ABC$, los alias serían:

$$A \times (-ABC) = -A^2BC = -BC$$

$$B \times (-ABC) = -AB^2C = -AC$$

$$C \times (-ABC) = -ABC^2 = -AB$$

Redundancia en los factoriales 2ⁿ

Cuando no hay interacciones de dos o tres factores en el factorial 2³, los efectos principales para A , B y C se pueden estimar apartir de media réplica del diseño definida por $I = ABC$, o bien, por $I = -ABC$. Así, en ausencia de interacción existe redundancia en el diseño completo de manera que los efectos principales se

pueden estimar con dos contrastes diferentes, cada uno a partir de una mitad distinta del diseño. La fracción media derivada de $I = ABC$ suele recibir el nombre de *fracción principal*, mientras que la fracción media derivada de $I = -ABC$ es la *fracción complementaria*.

Las dos fracciones medias forman un diseño 2^3 completo si cada una se corre por separado, el diseño obtenido es un diseño de bloque incompleto con dos bloques de cuatro tratamientos cada uno. La interacción ABC se confunde con los bloques, pero se pueden estimar todos los efectos principales y las interacciones de dos factores.

Cómo construir diseños 2^{n-1} de media réplica

El diseño fraccionado de un medio se construye con la interacción de mayor orden como generador del diseño y las combinaciones de tratamientos se identifican como sigue:

- Se escriben los coeficientes + y - en el orden estándar para los factores en un factorial 2^{n-1} .
- Se identifican los coeficientes \pm para el n -ésimo factor igualándolos con los coeficientes para la interacción de mayor orden en el factorial de 2^{n-1} .

La construcción de un diseño 2^{4-1} se describe en la tabla 12.2, los coeficientes + y - de los efectos principales de A , B y C para el factorial $2^{4-1} = 2^3$ se escriben en el orden estándar, los coeficientes para la interacción ABC se toman como el nivel del cuarto factor, D , que se combinará con los niveles de los otros tres factores y los coeficientes del contraste de interacción ABC se obtienen mediante el producto de los coeficientes para los tres efectos principales, como se muestra en la cuarta columna.

Tabla 12.2 Construcción de un diseño factorial fraccionado 2^{4-1}

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i> = ABC	Tratamiento
-	-	-	-	(1)
+	-	-	+	<i>ad</i>
-	+	-	+	<i>bd</i>
+	+	-	-	<i>ab</i>
-	-	+	+	<i>cd</i>
+	-	+	-	<i>ac</i>
-	+	+	-	<i>bc</i>
+	+	+	+	<i>abcd</i>

Por ejemplo, el producto de los coeficientes de A , B y C en el primer renglón es un signo -, de esta manera el primer renglón de coeficientes es (- - -) para la combinación de tratamientos (1). El segundo renglón de coeficientes es (+ - - +) para la combinación de tratamientos *ad*.

Construcción de media réplica con la regla de pares-impares

De acuerdo con la regla pares-impares, la interacción de mayor orden es la generadora del diseño. Las combinaciones de tratamientos resultantes necesarias para la media réplica se muestran en la columna derecha de la tabla 12.2, se trata de las combinaciones que reciben un coeficiente + en la interacción $ABCD$ mediante la regla pares-impares, ya que cada una tiene un número par de letras de efectos.

En el apéndice se presenta una tabla de generadores de diseños para los factoriales fraccionarios, el generador para el diseño 2^{4-1} de la tabla 12.2 se encuentra en la tabla 12A.1 como $D = ABC$, para indicar cómo se determinan los niveles del cuarto factor a partir de los niveles de otros factores.

El diseño de la tabla 12.2 es un factorial completo para A , B y C si se omite el factor D . No importa cuál factor se omita en un diseño de fracción media, el diseño resultante es un factorial completo del resto de los efectos.

Uso de la interacción de mayor orden como generador de diseño

Debido a que no es posible estimar el contraste de definición elegido como generador de diseño, por regla general se utiliza la interacción de mayor orden y menor interés para generar la media réplica. Suponiendo que media réplica de un factorial 2^5 se genera con el contraste $ABCDE$, el diseño requiere $2^{5-1} = 16$ unidades experimentales. El conjunto completo de alias se muestra en la tabla 12.3.

Tabla 12.3 Alias para los efectos principales e interacciones de dos factores en una fracción media del diseño factorial 2^5

<i>Efectos principales</i>		<i>Interacciones de dos factores</i>	
	<i>Alias</i>		<i>Alias</i>
<i>A</i>	<i>BCDE</i>	<i>AB</i>	<i>CDE</i>
<i>B</i>	<i>ACDE</i>	<i>AC</i>	<i>BDE</i>
<i>C</i>	<i>ABDE</i>	<i>AD</i>	<i>BCE</i>
<i>D</i>	<i>ABCE</i>	<i>AE</i>	<i>BCD</i>
<i>E</i>	<i>ABCD</i>	<i>BC</i>	<i>ADE</i>
		<i>BD</i>	<i>ACE</i>
		<i>BE</i>	<i>ACD</i>
		<i>CD</i>	<i>ABE</i>
		<i>CE</i>	<i>ABD</i>
		<i>DE</i>	<i>ABC</i>

Cada efecto principal tiene como alias cuatro interacciones de cuatro factores y cada interacción de dos factores tiene una interacción de tres factores. Por ejemplo, el alias de B se encuentra como la interacción generalizada $B \times ABCDE = AB^2CDE = ACDE$, y el alias de DE es $DE \times ABCDE = ABCD^2E^2 = ABC$.

Uso del factorial fraccionado para guiar el proceso experimental

Una estrategia experimental razonable en la investigación industrial utilizará diseños factoriales fraccionarios. Por ejemplo, media réplica del factorial 2^5 requiere 16 corridas de un proceso experimental; suponiendo que se completan las primeras 16 corridas necesarias, la segunda fracción de 16 corridas se puede realizar después en un segundo bloque para tener el factorial completo, de ser necesario, es posible que haya información suficiente sobre los efectos principales y las interacciones de dos factores en la primera fracción, lo que indica la necesidad de cambiar los niveles de uno o varios factores o permite eliminar algunos factores inertes y el investigador puede pasar a la siguiente etapa de experimentos, permitiendo que los recursos necesarios para el segundo conjunto de 16 corridas se asignen, con mejores resultados, a los nuevos experimentos.

Si los tratamientos han de probarse en secuencia, la aleatorización con factoriales fraccionados se obtiene al realizar las corridas reales de las combinaciones de tratamientos en orden aleatorio. Si se prueban todos los tratamientos en un conjunto de unidades experimentales físicas, se asignan al azar a esas unidades.

12.3 Resolución del diseño relacionada con la aleatorización

Los diseños factoriales fraccionados se agrupan en clases de acuerdo con las relaciones de aleatorización que existen en el diseño, estos grupos se identifican por su **resolución**. Los diseños más comunes son aquellos con resolución III, IV y V.

Resolución III: diseño en el que ningún efecto principal se confunde con otro efecto principal, pero los efectos principales se confunden con interacciones de dos factores y éstas con otras interacciones de dos factores.

Resolución IV: diseño en el que ningún efecto principal se confunde con otro efecto principal o con una interacción de dos factores, pero éstos se confunden entre sí.

Resolución V: diseño en el que ningún efecto principal ni interacción de dos factores se confunde con otro efecto principal o interacción de dos factores, pero estas últimas se confunden con las interacciones de tres factores.

La resolución de un diseño se determina por medio del menor número de símbolos que aparece en el generador del diseño. El diseño 2^{3-1} , generado a partir del contraste *ABC* (con tres símbolos), es un diseño de resolución III, porque los efectos principales no se confundieron entre sí, pero lo hicieron con las interacciones de dos factores. El diseño 2^{5-1} , generado mediante el contraste *ABCDE* (con cinco símbolos), es un diseño de resolución V porque no hubo confusión entre los efectos principales o las interacciones de dos factores, pero las interacciones de dos factores se confundieron con las de tres. El diseño 2^{4-1} , generado con la interacción *ABCD* (con cuatro símbolos), tiene resolución IV.

En general, cualquier diseño factorial fraccionado de resolución R no tiene interacciones de q factores confundidas con efectos que tengan menos de $R - q$ factores. La notación usada para identificar un diseño junto con su resolución

contiene un subíndice romano para indicar la resolución del diseño; de esta manera, la notación 2_v^{5-1} identifica un factorial fraccionario 2^{5-1} con resolución V.

12.4 Análisis de diseños 2^{n-1} de media réplica

En el análisis de experimentos factoriales con réplicas, estudiados y explicados en los capítulos 6, 7 y 11, los diseños replicados proporcionaron una estimación de la varianza del error experimental a partir del análisis de varianza. Sin embargo, con la réplica fraccionaria de los experimentos no se dispone de una estimación directa del error experimental para evaluar la significancia de los efectos e interacciones de los factores. En el ejemplo 12.2 se explica una estrategia común para estimar el error experimental a partir de réplicas fraccionadas de experimentos 2^n con un experimento factorial fraccionario 2_v^{5-1} .

Ejemplo 12.2 Recubrimiento de polímero para gabinetes de aluminio

Objetivo de la investigación: Una compañía ganó un contrato para fabricar gabinetes de aluminio con recubrimiento de un polímero plástico para tranceptores. El equipo de investigación que tenía la responsabilidad de desarrollar un proceso para adherir la capa de polímero al aluminio, identificó cinco factores en el proceso que posiblemente afectaran la adhesión. Éstos eran: el tipo de aleación de aluminio (A), el tipo de solvente usado para limpiar el aluminio (S), la estructura molecular del recubrimiento de polímero (M), el porcentaje de catalítico usado en el proceso de adhesión (C) y la temperatura de curado en el proceso (T).

Diseño del tratamiento: Se eligió un diseño factorial 2^5 para aprovechar la estructura del tratamiento factorial al evaluar las interacciones importantes junto con los efectos principales. Usaron dos aleaciones de aluminio, dos tipos de solvente, dos estructuras moleculares para el polímero, 10 y 15% de catalizador y 150 o 175° para la temperatura de curado.

Diseño del experimento: Con este experimento preliminar, el equipo buscaba identificar cualesquiera efectos principales y de interacción de dos factores que fueran importantes para el proceso, además de intentar reducir el número de unidades experimentales y el tiempo necesario para el primer experimento.

Decidieron usar un factorial fraccionario 2_v^{5-1} que requería sólo 16 unidades experimentales. La interacción de cinco factores $ASMCT$ fue el contraste de definición utilizado para generar el diseño, así podrían correr la segunda mitad del diseño, si fuera necesario completar el factorial. De la misma manera, la fracción media podía proporcionar suficiente información para indicar la necesidad de alterar los niveles de algunos factores o eliminar aquellos sin efectos.

En la tabla 12.4 se muestran las combinaciones de tratamientos, los coeficientes de los efectos principales y de los contrastes de interacción de dos factores, y las respuestas de las corridas experimentales. La fuerza necesaria para remover el recubrimiento plástico del gabinete de aluminio se usó como variable de respuesta.

Tabla 12.4 Fuerza necesaria para remover el recubrimiento plástico de una superficie de aluminio, para tratamientos en un experimento factorial fraccionario 2^{5-1}

Diseño										Tratamiento					Fuerza*
A	S	M	C	T	AS	AM	AC	AT	SM	SC	ST	MC	MT	CT	
-	-	-	-	+	+	+	+	-	+	+	-	+	-	-	t 41.5
+	-	-	-	-	-	-	-	-	+	+	+	+	+	+	a 39.6
-	+	-	-	-	-	+	+	+	-	-	-	+	+	+	s 43.9
+	+	-	-	+	+	-	-	+	-	-	+	+	-	-	ast 38.8
-	-	+	-	-	+	-	+	+	-	+	+	-	-	+	m 48.7
+	-	+	-	+	-	+	-	+	-	+	-	-	+	-	amt 52.0
-	+	+	-	+	-	-	+	-	+	-	+	-	+	-	smt 55.8
+	+	+	-	-	+	+	-	-	+	-	-	-	-	+	asm 43.2
-	-	-	+	-	+	+	-	+	+	-	+	-	+	-	c 39.5
+	-	-	+	+	-	-	+	+	-	-	-	-	-	+	act 42.6
-	+	-	+	+	-	+	-	-	-	+	+	-	-	+	sct 44.0
+	+	-	+	-	+	-	+	-	-	+	-	-	+	-	asc 33.8
-	-	+	+	+	+	-	-	-	-	-	-	+	+	+	mct 53.6
+	-	+	+	-	-	+	+	-	-	-	+	+	-	-	amc 48.1
-	+	+	+	-	-	-	-	-	-	-	+	-	-	-	smc 51.3
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	asmct 48.7

*Ejemplo del cálculo para las estimaciones de los efectos [vea la ecuación (12.1)]:

$$A = 2(-41.5 + 39.6 - 43.9 + 38.8 - 48.7 + 52.0 - 55.8 + 43.2 - 39.5 + 42.6 - 44.0 + 33.8 - 53.6 + 48.1 - 51.3 + 48.7)/16 = 2(-31.5)/16 = -3.94$$

En la tabla 12.4, los coeficientes + y - para los primeros cuatro factores (A , S , M y C) de un factorial $2^{5-1} = 2^4$ están escritos en orden normal; los coeficientes para la interacción $ASMC$ se usan para identificar los niveles del quinto factor (T) que se usarán en las combinaciones de tratamientos las 16 combinaciones de tratamientos mostradas tienen coeficiente + en la interacción de cinco factores, $ASMCT$, porque tienen un número impar de letras de efectos.

El diseño factorial fraccionado 2^{5-1} no tiene efecto principal ni interacciones de dos factores que sean alias de otro efecto principal o interacción de dos factores (vea la tabla 12.3), mientras que los efectos principales son alias de interacciones de cuatro factores (por ejemplo, $I_A = I_{SMCT}$), y las interacciones de dos factores son alias de interacciones de tres factores (por ejemplo, $I_{AS} = I_{MCT}$). Si se supone que las interacciones de tres y cuatro factores son despreciables, las estimaciones de los efectos principales se pueden usar para identificar los factores candidatos para una investigación más profunda en estudios sucesivos.

El efecto de un factor en un factorial fraccionado se estima como sigue:

$$AB\dots = \frac{2(l_{AB\dots})}{N} \quad (12.1)$$

donde $I_{AB\dots}$ es el contraste de las combinaciones de tratamientos y N es el número total de observaciones en el experimento. La suma de cuadrados con 1 grado de libertad para un efecto en un factorial fraccionario de 2^{n-1} es:

$$SC(AB\dots) = \frac{1}{2^{n-1}} (I_{AB\dots})^2 \quad (12.2)$$

En la tabla 12.5 se enumeran las estimaciones de los efectos principales y de interacción de dos factores en orden creciente de su valor desde -1.94 hasta $+9.71$. Un ejemplo del cálculo para estimar el efecto del factor A se encuentra al pie de la tabla 12.4 y las estimaciones de los demás se pueden verificar con la ecuación (12.1).

Tabla 12.5 Estimaciones de los efectos principales y de las interacciones de dos factores con la suma de cuadrados para los datos del factorial fraccionado de la tabla 12.4

Efecto	Estimación	Cuantil normal	Suma de cuadrados
A	-3.94	-1.74	62.02
AS	-3.69	-1.24	54.39
S	-0.76	-0.94	2.33
SC	-0.74	-0.71	2.18
AM	-0.41	-0.51	0.68
C	-0.24	-0.33	0.23
SM	-0.09	-0.16	0.03
AC	0.14	0.00	0.08
ST	0.16	0.16	0.11
CT	0.44	0.33	0.77
AT	0.74	0.61	2.18
MC	0.74	0.61	2.18
MT	1.09	0.94	4.73
T	3.61	1.24	52.20
M	9.71	1.74	377.33

Examen de los efectos con una gráfica de probabilidad normal

Se supone que las observaciones del experimento tienen una distribución normal con varianza constante σ^2 , pues representan la variación normal aleatoria alrededor de una media fija cuando los cambios en los niveles de los factores no tienen efecto real sobre la respuesta. Así, los efectos factoriales estimados a partir de estas observaciones tienen una distribución normal alrededor de una media 0 con varianza σ^2 cuando los factores no tienen efecto y la gráfica de los efectos estimados será una línea recta en la gráfica de probabilidad normal (vea el capítulo 4), pero los efectos estimados que no se ajustan a una línea recta no se pueden explicar como variación aleatoria.

El cuantil normal correspondiente calculado para cada efecto se muestra en la tabla 12.5. Recuerde que el *cuantil normal* es el valor esperado para una variable con distribución normal, con media 0 y varianza 1 encontrada en esa posición de los efectos ordenados (vea el capítulo 4).

En la figura 12.1 se muestra la gráfica de probabilidad normal de los efectos principales y de interacción de dos factores; los efectos principales que no se encuentran en la línea recta de la gráfica de probabilidad normal son los efectos de la aleación de aluminio (A), la estructura molecular del polímero (M) y la temperatura de curado (T); el efecto estimado para la interacción entre la aleación de aluminio y el tipo de solvente (AS) se desvía mucho de la recta y los demás efectos parecen representativos del error experimental aleatorio.

Interpretación de los efectos más grandes

En la figura 12.2 se muestran las gráficas de los efectos; en la figura 12.2a, la estimación positiva para T , 3.61, indica que un alto nivel de temperatura de curado produce una adhesión más fuerte. De igual manera, en la figura 12.2b, la estimación positiva para M , 9.71 indica que la estructura molecular representada por el nivel “alto” produce una adhesión más fuerte.

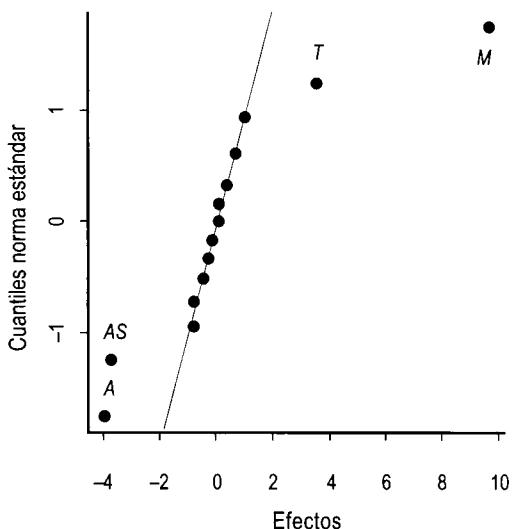


Figura 12.1 Gráfica de probabilidad normal de los efectos de factor estimados y las interacciones del factorial fraccionario 2^{5-1} en el ejemplo 12.2

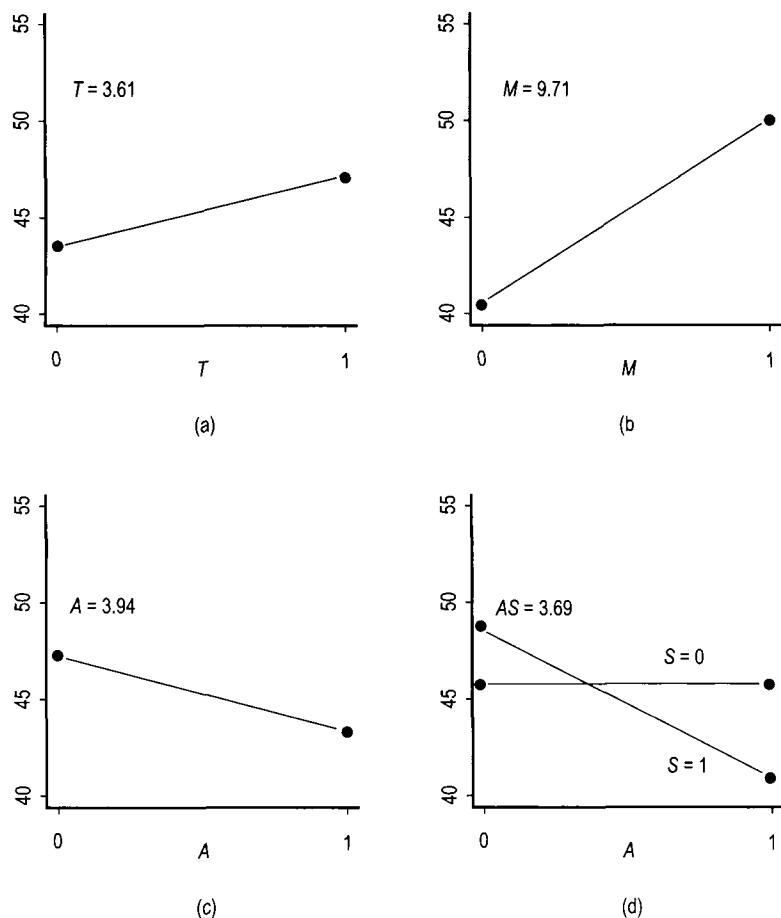


Figura 12.2 Gráficas para los efectos principales de T , M y A y para la interacción AS

En la figura 12.2c, la estimación negativa para A , -3.94 , sugiere que el nivel “bajo” del factor aleación de aluminio produce una adhesión más fuerte. Sin embargo, la estimación mayor para la interacción AS , -3.69 , puede alterar la inferencia sobre el efecto principal de A , la aleación de aluminio.

Una investigación de la interacción entre la aleación de aluminio (A) y el solvente para limpiarlo (S) requiere los promedios de las cuatro combinaciones de tratamientos que contienen a A y S . De la tabla 12.4, los valores son:

Tratamiento	A	S	Total	Media
$t + m + c + mct$	0	0	183.3	45.8
$a + amt + act + amc$	1	0	182.3	45.6
$s + smt + sct + smc$	0	1	195.0	48.8
$ast + asm + asc + asmct$	1	1	164.5	41.1

El efecto de A cuando $S = 0$ es $45.6 - 45.8 = -0.2$, mientras que el efecto cuando $S = 1$ es $41.1 - 48.8 = -7.7$. Por lo tanto, la diferencia entre las dos aleaciones con el segundo solvente ($S = 1$) es mucho más grande que con el primer solvente ($S = 0$), como se muestra en la figura 12.2d.

Algunas advertencias sobre las suposiciones

Es importante recordar que, debido a los alias, los efectos estimados que se juzgan importantes para el proceso son en realidad $A + SMCT$, $M + ASCT$, $T + ASMC$ y $AS + MCT$. Así, la suposición de despreciable para las interacciones de tres factores o más es crucial en las decisiones respecto a la importancia de los tres efectos principales y una interacción de dos factores para el proceso de adhesión.

Daniel (1959) señaló algunas posibles interpretaciones equivocadas con una gráfica de probabilidad de los efectos estimados en factoriales sin réplica. El considerar que una línea suave en la gráfica de los efectos estimados en una parte de la gráfica normal no indica que haya efectos de factor puede llevar a una sobreestimación de la varianza del error. Una línea en extremo irregular puede conducir al error opuesto de juzgar que hay efectos reales cuando de hecho sólo representan el error experimental aleatorio. La gráfica es muy eficaz cuando sólo una pequeña proporción de los efectos factoriales son importantes.

Un método para estimar la varianza del error experimental

Una prueba de la significancia de estos efectos requiere una estimación de la varianza del error experimental. Una estrategia general para obtener tal estimación incluye varios pasos, el primero selecciona aquellos efectos que parecen ser despreciables en la gráfica de probabilidad normal y une sus sumas de cuadrados para estimar el error experimental; según la figura 12.1, parece que sólo los efectos A , M , T y AS son significativos, el resto son despreciables y representativos de la variación debida al error experimental. La tabla 12.6 muestra el análisis de varianza para un modelo que contiene efectos nada más para A , M , T , S y AS ; se incluyó el efecto principal para S en el análisis porque el efecto de interacción AS está definido para los efectos principales en el modelo (capítulo 6) y debe incluirse para obtener los valores correctos de las sumas de cuadrados. Observe

Tabla 12.6 Análisis de varianza para un modelo reducido con efectos A, M, T y AS es un factorial fraccionario 2^{5-1} (Ejemplo 12.2)

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	15	561.41			
A	1	62.02	62.02	44.11	0.000
M	1	377.33	377.33	268.36	0.000
T	1	52.20	52.20	37.12	0.000
S	1	2.33	2.33	1.77	0.213
AS	1	54.39	54.39	38.68	0.000
Error	10	13.14	1.31		

que S no es significativo y podría unirse a la suma de cuadrados del error siguiendo la estrategia general. Análisis de varianza de la 12.6, la suma de cuadrados combinada para los efectos que parecen despreciables es $SCE = SC(\text{error}) + SCS = 13.14 + 2.33 = 15.47$, a partir del análisis de varianza de la tabla 12.6 con 11 grados de libertad. Así, una estimación de la varianza del error experimental es $s^2 = 15.7/11 = 1.41$.

No hay sustituto para las réplicas

Un experimento con réplicas habría proporcionado una estimación legítima de σ^2 , pues la validez de cualquier prueba de significancia es cuestionable cuando la estimación de la varianza se basa en juicios subjetivos y el método descrito para estimar la varianza del error experimental a partir de un experimento sin réplicas es muy subjetivo.

Se han desarrollado varios procedimientos más formales para resolver el problema de estimación de la varianza en factoriales fraccionados, pero cada uno requiere cierto grado de juicio subjetivo por parte del investigador. Ninguno de ellos se estudia en este texto, pero se pueden consultar las obras citadas a continuación.

Berk y Picard (1991) revisaron y evaluaron algunos métodos, incluyendo uno de su invención; Box y Meyer (1986) y Zahn (1975a, 1975b) propusieron procedimientos de cómputo más completos. Los métodos de Zahn son extensiones de los métodos de media gráfica normal originales de Daniel (1959). Algunos procedimientos (Lenth, 1989; Voss 1988) proponen métodos más sencillos basados en un análisis normal de los procedimientos de varianza.

Con base en estudios de simulación, Berk y Picard (1991) concluyeron que la mayoría de los métodos producían tasas de error casi idénticas y la mayoría tiende a declarar que los efectos nulos son reales.

Los peligros inherentes a los factoriales fraccionados sin réplicas no se pueden ignorar, porque sólo las réplicas pueden proteger contra la elección de efectos engañosos como legítimos. Los factoriales fraccionados son estudios legítimos sólo si se reconocen como revisiones preliminares cuyo objetivo es preparar el camino para experimentos con réplicas más rigurosos.

Error estándar y pruebas de hipótesis sobre los efectos

La varianza de un efecto estimado en un factorial fraccionado es:

$$\text{varianza} = \frac{4\sigma^2}{N} \quad (12.3)$$

donde σ^2 es la varianza del error experimental y N es el número de observaciones en el experimento. La varianza del error experimental es $s^2 = 1.41$ con 11 grados de libertad, de donde se puede calcular una estimación del error para cada efecto como se muestra en la tabla 12.5. La estimación de la varianza para cualquier efecto es:

$$\frac{4s^2}{N} = \frac{4(1.41)}{16} = 0.35$$

y el error estándar estimado de las estimaciones de los efectos es $\sqrt{0.35} = 0.59$. La significancia de cada efecto se puede determinar con la prueba t de Student

como una razón de los efectos estimados entre el error estándar, o con la prueba F como se muestra en la tabla 16.6. Todos los efectos supuesto son significativos según las pruebas F de la tabla 12.6.

Gráficas de residuales para evaluar el modelo

La estimación de la varianza del error experimental se encontró al suponer que todos los efectos distintos a los del modelo eran despreciables. Si estos efectos son despreciables, las gráficas residuales para el modelo ajustado deben confirmar que el modelo es correcto; los residuales calculados para el modelo incluyeron sólo los efectos A, M, T, S y AS . En la figura 12.3 se muestra la gráfica de residuales contra los valores pronosticados y la gráfica de probabilidad normal de los residuales; en la primera, los residuales parecen estar en una línea razonablemente recta en la gráfica de probabilidad normal y no parece haber razón para sospechar varianzas heterogénea en ella y parece que no faltan efectos principales en el modelo, ya que no hay evidencia de puntos disparados.

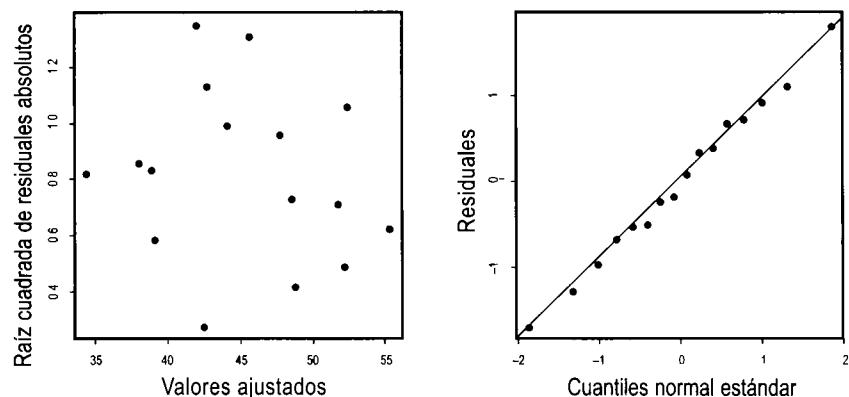


Figura 12.3 Gráficas residuales para el modelo con efectos A, M, T, S y AS del ejemplo 12.2

No se detecta que el catalizador, factor C , tenga efecto y parece ser inerte; de hecho, el diseño fue una réplica de un factorial 2^4 completo para los factores A, S, M y T . El equipo de investigación piensa ahora en un estudio más minucioso de estos factores y sus interacciones.

12.5 Fracción de un cuarto de factoriales 2^n

Las fracciones más pequeñas para las corridas iniciales del factorial 2^n pueden proporcionar suficiente información para tomar decisiones críticas sobre la efectividad de los factores, en especial cuando se tiene un número más o menos grande

de factores o cuando las corridas individuales son costosas. Un cuarto de réplica del diseño 2ⁿ se designa como un factorial fraccionario de $\frac{1}{2} 2^n = 2^{n-2}$.

Un cuarto de réplica de un diseño 2⁶ es un factorial fraccionario 2⁶⁻² que requiere de sólo 16 de las 64 corridas necesarias para el diseño completo 2⁶ y corresponde a un bloque del diseño de bloque incompleto para un diseño 2⁶ en cuatro bloques de 16 unidades (vea el capítulo 11). Se puede usar cualquiera de los cuatro bloques para el diseño 2⁶⁻².

Se usan dos contrastes de definición para los diseños fraccionados a un cuarto

Se requieren dos contrastes de definición para construir los cuatro bloques incompletos que, junto con su interacción generalizada se confunden con los bloques con estos mismos contrastes de definición se construye el diseño 2⁶⁻². Suponiendo que se seleccionan *ABDE* y *ABCF* como generadores del diseño, la interacción generalizada del generador del diseño es:

$$ABDE \times ABCF = A^2B^2CDEF = CDEF$$

Los generadores de diseño junto con su interacción generalizada completan la relación de definición para el diseño, como sigue:

$$I = ABDE = ABCF = CDEF$$

Las combinaciones de tratamientos necesarias para el diseño se pueden identificar en forma análoga a la mostrada para el diseño de media réplica en la tabla 12.2. Los coeficientes + y - los efectos principales de *A*, *B*, *C* y *D* se escriben en orden normal para un factorial completo 2⁶⁻² = 2⁴ como se muestra en la tabla 12.7.

Los alias de *E* y *F* que contienen alguna de las letras de factores *A* a *D* se usan para determinar los coeficientes de *E* y *F* en las combinaciones de tratamientos. Dados los generadores del diseño *ABDE*, *ABCF* y su interacción generalizada *CDEF*, los alias respectivos de *E* y *F* son:

$$E = ABD = CDF = ABCEF$$

y

$$F = ABC = CDE = ABDEF$$

Las equivalencias *E* = *ABD* y *F* = *ABC* se pueden usar para determinar los coeficientes de *E* y *F*; es decir, los coeficientes de *ABD* se toman como coeficientes para *E* y los de *ABC* para *F*. En la tabla 12.7 se muestran los tratamientos necesarios para el diseño.

Los contrastes de definición para el diseño de un cuarto de réplica tienen cuatro letras, de manera que el diseño obtenido tiene resolución IV, o un factorial fraccionado 2⁶⁻²_{IV}. Ningún efecto principal es alias de otro efecto principal o de interacciones de dos factores, pero las interacciones de dos factores son alias de una o más de las otras interacciones de dos factores; en la tabla 12.8 se muestra el conjunto completo de alias.

Tabla 12.7 Construcción de un diseño factorial fraccionario de 2_{IV}^{6-2}

Diseño						Tratamiento
A	B	C	D	E = ABD	F = ABC	
-	-	-	-	-	-	(1)
+	-	-	-	+	+	aef
-	+	-	-	+	+	bef
+	+	-	-	-	-	ab
-	-	+	-	-	+	cf
+	-	+	-	+	-	ace
-	+	+	-	+	-	bce
+	+	+	-	-	+	abcf
-	-	-	+	+	-	de
+	-	-	+	-	+	adf
-	+	-	+	-	+	bdf
+	+	-	+	+	-	abde
-	-	+	+	+	+	cdef
+	-	+	+	-	-	acd
-	+	+	+	-	-	bcd
+	+	+	+	+	+	abcdef

Tabla 12.8 Relaciones de alias para el factorial fraccionario 2_{IV}^{6-2} con relaciones de definición $I = ABDE = ABCF = CDEF$

Efecto	Alias de		
	ABDE	ABCF	CDEF
A	BDE	BCF	ACDEF
B	ADE	ACF	BCDEF
C	ABCDE	ABF	DEF
D	ABE	ABCDF	CEF
E	ABD	ABCEF	CDF
F	ABDEF	ABC	CDE
AB	DE	CF	ABCDEF
AC	BCDE	BF	ADEF
AD	BE	BCDF	ACEF
AE	BD	BCEF	ACDF
BC	ACDE	AF	BDEF
CD	ABCE	ABDF	EF
CE	ABCD	ABEF	DF
ACD	BCE	BDF	AEF
ACE	BCD	BEF	ADF

Una fracción $1/2^p$ de un diseño 2^n se denomina como un factorial fraccionario de 2^{n-p} ; en general, una fracción de este tipo requerirá p generadores de diseño

con $2^p - p - 1$ interacciones generalizadas cada efecto tendrá $2^p - 1$ alias. La fracción $1/2^2$ o diseño de un cuarto 2_{IV}^{6-2} en la tabla 12.7 requiere $p = 2$ diseños generadores que dan $2^2 - 2 - 1 = 1$ interacciones generalizadas se obtuvieron $2^2 - 1 = 3$ alias para cada efecto.

Se eligen generadores de diseño para evitar una mala estructura de alias

Debe tenerse cuidado al seleccionar los generadores de diseño para evitar que los efectos de interés sean alias de otros en el estudio. Así, es importante estudiar la estructura de alias antes de asignar los factores a las letras cuando se establece un experimento.

Observe que el diseño 2_{IV}^{6-2} en la tabla 12.7 consiste de una réplica completa del diseño 2^4 en A, B, C y D ; también es una réplica completa de un diseño 2^4 en cualesquiera cuatro factores que no se combinen para formar las relaciones de definición para el diseño; cualquier combinación de factores distintos a $ABDE$, $ABCF$ y $CDEF$ producirá un factorial completo 2^4 . Una inspección cuidadosa del diseño en la tabla 12.7 muestra que consiste en dos réplicas de un diseño experimental 2^{4-1} , compuestos de cuatro factores usados en cualquiera de las relaciones de definición consideradas por separado.

El diseño 2^{n-p} incluirá un factorial completo para cualesquier $n - p$ factores que no forman una relación de definición. Un factorial fraccionario se puede formar con algún subconjunto de $n - p$ factores, pero si se sospecha que algunos factores tendrán efectos despreciables, entonces el factorial fraccionario original 2^{n-p} se puede establecer de manera que exista un factorial completo para los factores de mayor interés o para los que tengan efectos reales.

12.6 Construcción de diseños 2^{n-p} con resolución III y IV

Diseño de resolución III con N unidades experimentales para $N - 1$ factores

Un diseño de resolución III se puede construir para investigar $n = N - 1$ factores con N unidades experimentales, donde N es un múltiplo de 4. Por ejemplo, la fracción media del factorial 2^3 que se encuentra en la tabla 12.1 es un factorial fraccionario de 2_{III}^{3-1} , que tiene $N - 1 = 3$ factores investigados con $N = 4$ unidades experimentales.

En los diseños con resolución III los efectos principales no son alias entre sí, pero lo son de las interacciones de dos factores. Un diseño 2^{n-p} con resolución III requiere p generadores de diseño. Suponiendo que se desea un diseño de resolución III para $N - 1 = 7$ factores en $N = 8$ corridas, estas últimas serían la fracción (1/16) de las $2^7 = 128$ necesarias para el factorial completo. Así, el diseño es un factorial fraccionario 2_{III}^{7-4} que requiere $p = 4$ generadores de diseño.

La construcción del diseño se inicia escribiendo los coeficientes + y - para los $n - p$ efectos principales en orden normal. En la tabla 12.9 se muestran los coeficientes + y - para $n - p = 3$ efectos principales, A, B y C , en orden normal para un factorial completo de $2^{7-4} = 2^3$.

Tabla 12.9 Un factorial fraccionario 2^{7-4}_{III} con generadores de diseño ABD , ACE , BCF y $ABCG$

<i>Unidad</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D = AB</i>	<i>E = AC</i>	<i>F = BC</i>	<i>G = ABC</i>	<i>Tratamiento</i>
1	—	—	—	—	+	+	—	<i>def</i>
2	+	—	—	—	—	+	+	<i>afg</i>
3	—	+	—	—	+	—	+	<i>beg</i>
4	+	+	—	+	—	—	—	<i>abd</i>
5	—	—	+	+	—	—	+	<i>cdf</i>
6	+	—	+	—	+	—	—	<i>ace</i>
7	—	+	+	—	—	+	—	<i>bcf</i>
8	+	+	+	+	+	+	+	<i>abcdefg</i>

Los coeficientes necesarios para el resto de los factores, D , E , F y G se asocian con los coeficientes de todas las columnas de interacción de A , B y C , de este modo, las columnas que indican los niveles de los factores D , E , F y G se generan a partir de las siguientes relaciones de equivalencia:

$$D = AB \quad E = AC \quad F = BC \quad y \quad G = ABC$$

entonces, los $p = 4$ generadores de diseño son:

$$D \times AB = ABD$$

$$E \times AC = ACE$$

$$F \times BC = BCF$$

y

$$G \times ABC = ABCG$$

La relación de definición se encuentra con la multiplicación de los cuatro generadores en todas las formas posibles. La relación de definición completa es:

$$\begin{aligned} I &= ABD = ACE = AFG = BCF = BEG = CDG = DEF = ABCG \\ &= ABEG = ACDF = ADEG = BCDE = BDGF = CEFG \\ &= ABCDEFG \end{aligned}$$

La estructura de los alias del diseño para los efectos principales se encuentra multiplicando cada uno de estos por las interacciones incluidas en la relación de definición. Las interacciones de dos factores que son alias de los efectos principales son:

$$A = BD = CE = FG$$

$$B = AD = CF = EF$$

$$C = AE = BF = DG$$

$$D = AB = CG = EF$$

$$E = AC = BG = DF$$

$$F = AG = BC = DE$$

$$G = AF = BE = CD$$

El diseño de la tabla 12.9 contiene todos los contrastes disponibles del factorial 2^{n-p} asociados con los efectos principales y sus alias; este se conoce como un *diseño saturado*. El diseño dispone de 7 grados de libertad para estimar cada uno de los 7 efectos principales y no tiene grados de libertad para estimar el error experimental. Los diseños de resolución III se pueden usar en experimentos exploratorios con muchos factores para identificar los claramente dominantes sin desperdicio de tiempo y recursos.

Diseños Plackett-Burman

Como se muestra en los párrafos anteriores; se pueden generar diseños saturados de resolución III con $n = N - 1$ factores y que requieran $N = 2^{n-p}$ unidades experimentales a partir del arreglo factorial básico 2^{n-p} ; el resultado es un diseño con el número de unidades experimentales igual a alguna potencia de 2 (experimentos con 4, 8, 16, ..., 2^{n-p} unidades).

Una clase de diseños con resolución III, desarrollada por Plackett y Burman (1946), que requiere un número de unidades experimentales igual a un múltiplo de 4, se ha utilizado de manera exhaustiva en experimentos exploratorios dentro de la investigación industrial. Ellos establecen diseños para valores intermedios de N que no sean una potencia de 2 y para $N \leq 100$, excepto 92. En la tabla 12.10 se encuentra una muestra de los diseños más prácticos y fáciles de construir encuentra para $N = 12, 16, 20, 24$ y 32 corridas experimentales; los diseños se pueden generar a partir de los renglones con signos + y -.

En esta tabla, cada renglón de generadores tiene $N - 1$ coeficientes, que son los coeficientes + y - necesarios para la primera corrida de los $N - 1$ factores. Los coeficientes del factor para la segunda corrida se generan a partir de la primera, tomando el primer coeficiente de la primera corrida y colocándolo en la última posición de la segunda corrida y moviendo los coeficientes de la primera una posición a la izquierda; las corridas sucesivas se generan de la misma manera. Al completar el ciclo de reemplazo de coeficientes para $N - 1$ corridas, se agrega una corrida final con coeficientes - para los $N - 1$ factores. En la tabla 12.11 se muestra un diseño para $N - 1 = 11$ factores en $N = 12$ corridas.

Tabla 12.10 Generadores para diseños Plackett-Burman con $N = 12, 16, 20, 24$ y 32 corridas

<i>N</i>	
12	++-+++-+--+-
16	+---++-++-++-++
20	++--++++-++-+-+--
24	+++++-++-++-++-+--
32	-+-+-+--+++-++-++-+--+-+

Tabla 12.11 Diseño Plackett-Burman para $N - 1 = 11$ factores en $N = 12$ corridas

<i>Corrida</i>	<i>Factor</i>										
	1	2	3	4	5	6	7	8	9	10	11
1	+	+	-	+	+	+	-	-	-	+	-
2	+	-	+	+	+	-	-	-	+	-	+
3	-	+	+	+	-	-	-	+	-	+	+
4	+	+	+	-	-	-	+	-	+	+	-
5	+	+	-	-	-	+	-	+	+	-	+
6	+	-	-	-	+	-	+	+	-	+	+
7	-	-	-	+	-	+	+	-	+	+	+
8	-	-	+	-	+	+	-	+	+	+	-
9	-	+	-	+	+	-	+	+	+	-	-
10	+	-	+	+	-	+	+	+	-	-	-
11	-	+	+	-	+	+	+	-	-	-	+
12	-	-	-	-	-	-	-	-	-	-	-

Aunque este diseño de 12 corridas se puede usar para revisar 11 factores, se recomienda que se use para 6 factores menos que el número de corridas (Mason, Gunst y Hess, 1989). Después de quitar 6 factores del diseño, y, una vez estimados los efectos principales a partir del diseño y quedan 5 grados de libertad para estimar la varianza del error experimental. Por supuesto, esta práctica todavía supone que los efectos principales son los dominantes en el experimento; por ejemplo, un diseño de 12 corridas para 6 factores se obtiene con sólo las primeras seis columnas del diseño descrito en la tabla 12.11 y hay en total 11 grados de libertad en el diseño, de los cuales 6 se usan para estimar los efectos principales y 5 para estimar la varianza del error experimental.

Los diseños de resolución IV requieren $2N$ unidades experimentales para N factores

Los diseños de 2^{n-p} de resolución IV no tienen efectos principales que sean alias de otros o de interacciones de dos factores y se pueden estimar sin que se hayan confundido con interacciones de dos factores, lo que no era el caso con los diseños de resolución III. Un diseño de 2_{IV}^{n-p} con $N = n$ debe tener al menos $2n$ unidades experimentales o corridas.

Un diseño de resolución IV con n factores se puede generar a partir de un diseño de resolución III con $n - 1$ factores mediante una técnica de *doblez*, que consiste en agregar una segunda fracción a un diseño de resolución III para $n - 1$ factores en n corridas, todos los signos en la segunda fracción son los inversos de la primera: La columna I en la primera fracción tiene coeficientes +, mientras que en la segunda fracción la columna I tiene sólo coeficientes -, esta columna de coeficientes se usa para un factor agregado el diseño ahora tiene n factores con $2n$ corridas.

En la tabla 12.12 se explica el procedimiento con la generación de un diseño 2_{IV}^{4-1} a partir de un diseño de 2_{III}^{3-1} ; la primera fracción es un diseño 2^{3-1} con resolución III, en el que los coeficientes para el factor C se derivan de los coeficientes de la interacción AB y la segunda fracción se deriva de la primera con signos opuestos para el diseño completo 2^{4-1} . El diseño se puede ver como de resolución IV, ya que es una fracción media de un diseño 2^4 que requiere un contraste de definición. Se trata de un factorial 2^3 completo en A , B y D con los coeficientes de C obtenidos de $C = ABD$, de manera que el contraste de definición es $C \times ABD = ABCD$. La resolución del diseño es igual al número de letras en el contraste de definición $ABCD$, o resolución IV.

Tabla 12.12 Generación de un diseño 2^{4-1} con resolución IV mediante doblez de un diseño 2^{3-1} con resolución III

$I = D$	A	B	$C = AB$	
+	-	-	+1	La primera fracción es un diseño de 2_{III}^{3-1}
+	+	-	-	
+	-	+	-	
+	+	+	+	
-	+	+	-	Segunda fracción tiene signos opuestos
-	-	+	+	
-	+	-	+	
-	-	-	-	

12.7 Genichi Taguchi y la mejora de la calidad

Por último, se observa una aplicación especial de los diseños factoriales fraccionados debida a los recientes desarrollos de aplicaciones de la estadística en la manufactura, parte de los cuales tratan el aspecto de diseños estadísticos.

Los diseños factoriales fraccionados se usan en forma amplia en experimentos fuera de la línea para mejorar la calidad del producto. Las investigaciones fuera de la línea integran los principios de diseño de ingeniería y de diseño estadístico para mejorar la calidad de productos e incrementar la productividad. En particular, la metodología de Taguchi (Taguchi, 1986) ha tenido un impacto importante sobre el mejoramiento del diseño de productos y procesos en la manufactura.

El **diseño de parámetros robustos** es la parte de la metodología de Taguchi que involucra diseños de tratamientos factoriales, el diseño consiste en factores que se

pueden controlar durante el proceso de manufactura y los que no son controlables. En la terminología de Taguchi los factores controlables se llaman *parámetros*; los que no se controlan se llaman factores de *ruido*. Las variables o factores de ruido son los más sensibles a los cambios en las condiciones del entorno durante la producción y, por lo tanto, transmiten la variabilidad a las respuestas de interés en el proceso.

Ambos tipos de factores, controlables o de ruido, se pueden controlar en los experimentos fuera de la línea, lo que es una consideración importante en la metodología de Taguchi. Un objetivo es determinar qué combinación de factores controlables es la menos sensible a los cambios en las variables de ruido, de este concepto se deriva el nombre de *diseño de parámetros robustos*. La mejor elección de los niveles de factores controlables conduce a un proceso de manufactura que tiene como resultado el producto deseado y es *robusto* respecto a las fluctuaciones de los factores de ruido.

El concepto de diseño se explica con un experimento hipotético fuera de la línea con tres factores controlables A , B y C , cada uno con dos niveles y un factor de ruido, D , con tres niveles. Un arreglo ortogonal para los factores de control se cruza con un arreglo ortogonal para los factores de ruido. Para el ejemplo con cuatro factores, el factorial 2^3 para A , B y C constituye un *arreglo interno* en la terminología de Taguchi, y el 3^1 para D sería un *arreglo externo*.

Los tres factores de control se configuran en la tabla 12.13 como una fracción $\frac{1}{2}$ de un factorial 2^3 con 4 corridas, o un factorial fraccionario 2^{3-1}_{III} , en el arreglo interior. Se usan tres niveles para el factor de ruido, D , en el arreglo externo para cada corrida en el diseño 2^{3-1}_{III} , y se obtienen 12 corridas del experimento.

Tabla 12.13 Experimento hipotético de Taguchi

A	B	C	D_0	D_1	D_2	\bar{y}	s
–	–	+	54	56	52	54	2.0
+	–	–	69	70	71	70	1.0
–	+	–	58	55	49	54	4.6
+	+	+	58	65	69	64	5.6

Suponiendo que un producto manufacturado tiene como meta una respuesta de 54; en la tabla 12.13 se muestran la media, \bar{y} , y la desviación estándar, s , de las observaciones para las corridas en los tres niveles del factor de ruido, D , para cada corrida de los factores de control.

Dos tratamientos para los factores de control, $(-, -, +)$ y $(-, +, -)$, cumplen la respuesta deseada de 54, pero el tratamiento $(-, -, +)$, con una desviación estándar de 2.0, es menos sensible a los cambios en los niveles de ruido del factor (D) que el tratamiento $(-, +, -)$ con desviación estándar de 4.6. Así, el producto obtenido con el tratamiento $(-, +, -)$ se considera el *más robusto*, ya que las fluctuaciones en el factor de ruido (D) transmiten una variabilidad menor a la respuesta del producto.

El método de análisis de Taguchi se concentra en maximizar una razón señal ruido (S/R) específica para las metas del proceso. Las metas principales pueden

ser: 1) minimizar la respuesta, 2) maximizar la respuesta y 3) lograr una respuesta nominal diferente de la mínima o la máxima.

Este método procede en forma sencilla con el análisis de varianza para la S/R que determina los factores de control que afectan esta razón, luego continúa con el análisis de varianza para la \bar{y} , que comprueba qué factores de control tienen impacto en la respuesta media y, por lo tanto, los niveles de los factores que se deben establecer para lograr la respuesta buscada. Myres y Montgomery (1995) explican otros detalles del análisis para los lectores interesados.

Pignatello y Ramberg (1991) estudian las ventajas y desventajas más importantes del enfoque de Taguchi. Myers y Montgomery (1995), Montgomery (1997) y Vining (1998) presentan con más detalle descripciones, ejemplos y críticas a la metodología de Taguchi. Una contribución importante del método de Taguchi fue la validación de la experimentación y del diseño estadístico como parte del proceso de mejoramiento de la calidad. No obstante, los diseños de Taguchi han despertado gran controversia en cuanto al uso de diseños especiales y su instauración.

Aunque se recomiendan muchos diseños factoriales fraccionarios estándar 2^{n-p} en el método de Taguchi, muchos son diseños Plackett-Burman saturados o casi saturados, o fraccionados con tres niveles para detectar la curvatura de la respuesta y con frecuencia es necesario suponer la no interacción entre los factores de control para tener éxito al usarlos. El cruce del diseño ortogonal de los factores de control con el factor de ruido facilita la estimación de las interacciones entre los factores de control y de ruido, pero sacrifica la posibilidad de estimar las interacciones importantes entre los factores de control. Myers y Montgomery (1995) sugieren varias estrategias de diseño que pueden ser más económicas en las corridas experimentales y aun así proporcionar la manera de estimar las interacciones de interés para el investigador. Hunter (1985) presenta un estudio detallado de los diseños fraccionados de tres niveles de Taguchi con base en cuadrados latinos.

Hunter (1989) y Pignatello y Rambierg (1991) también han sugerido mejores secuencias de las estrategias de diseño; por ejemplo, la secuencia puede comenzar con factoriales fraccionados 2^{n-p} para detectar las interacciones con los niveles de factores agregados al centro del diseño para detectar la curvatura. Otros refinamientos sugeridos al establecer las características del proceso incluyen el uso de factoriales completos y diseños de superficie de respuesta (vea al capítulo 3).

Por último, la aleatorización no se respalda en el enfoque de Taguchi; pues puede llevar a que el diseño no considere efectos importantes confundidos con las variaciones externas.

12.8 Observaciones finales

Como se vio en los capítulos 11 y 12, los factoriales son versátiles diseños de tratamiento que pueden adaptarse a diversas condiciones experimentales, como réplicas sencillas, réplicas de factoriales y experimentos que requieren diseños de bloques incompletos.

Sólo se han presentado los principios básicos de la construcción y análisis de estos diseños, pero existe mucha literatura sobre diseños factoriales fraccionados para el trabajo experimental en la industria. Se pueden encontrar estudios amplios sobre diseño y análisis de factoriales fraccionados en diferentes campos de aplicación en Daniel (1976), Box, Hunter y Hunter (1978), Diamond (1989), Haaland (1989) y Mason, Gunst y Hess (1989); algunas de estas referencias contienen tablas de diseño específicos derivados de los principios básicos que se presentan en este capítulo. Diamond (1989) proporciona un bibliografía amplia de la literatura publicada sobre factoriales fraccionarios.

También se ha desarrollado un gran número de productos comerciales de software para producir diseños y análisis para cualquier experimento que requiere diseños de factoriales fraccionados.

Diseños fraccionados para factoriales de 3ⁿ

Las réplicas fraccionadas se pueden construir para factoriales 3ⁿ y otras series de factoriales con los mismos principios que los de la serie 2ⁿ. Es posible desarrollar factoriales fraccionados de un tercio o un noveno con las series 3ⁿ; los alias para los efectos factoriales se determinan según los productos de los símbolos por los contrastes de definición y sus cuadrados. Una amplia selección de diseños factoriales fraccionados para las series 3ⁿ se encuentran en Conar y Zelen (1959), los detalles de la construcción de factoriales fraccionados en series distintas a 3ⁿ se pueden ver en Kempthorne (1952) y en John (1987) y algunos diseños a manera de ejemplo están en Cochran y Cox (1957), Johnson y Leone (1977) y Montgomery (1997).

EJERCICIOS PARA EL CAPÍTULO 12

1. En seguida se muestra una lista de diseños:

i.	2^{5-1}	v.	2^{7-2}	ix.	2^{9-5}
ii.	2^{6-1}	vi.	2^{7-3}	x.	2^{10-6}
iii.	2^{6-2}	vii.	2^{7-4}	xi.	2^{11-6}
iv.	2^{6-3}	viii.	2^{8-4}	xii.	2^{15-11}

Indique las siguientes cantidades para cada diseño:

- a. La fracción del diseño completo.
 - b. El número de generadores requerido.
 - c. El número de interacciones generalizadas.
 - d. El número de alias para cada efecto.
 - e. El número de unidades experimentales o de corridas necesarias.
2. Construya un diseño factorial fraccionario 2⁵⁻¹ usando $I = ABCDE$ como relación de definición.
- a. ¿Cuál es la resolución del diseño?
 - b. Muestre la estructura de alias para los efectos principales y las interacciones de dos factores.

3. Construya un diseño factorial fraccionado 2^{5-2} usando $ABCD$ y $CDEF$ como generadores de diseño.
 - a. ¿Cuál es la interacción generalizada?
 - b. ¿Cuál es la resolución del diseño?
 - c. Muestre la estructura de alias para todos los efectos principales y las interacciones de dos factores.

4. Construya un diseño de resolución IV que sea $\frac{1}{8}$ de un factorial de 2^7 .

5. Si hace el doblez para un factorial fraccionado, el diseño que obtiene, ¿es un factorial fraccionario 2^{6-2} ? Explique.

6. Se propone un factorial fraccionario 2^{5-2} con dos relaciones generadoras posibles:
 - i. $I = ABCD = BCE$
 - ii. $I = ABCDE = ABCD$
 - a. ¿Qué fracción del diseño de 2^5 será?
 - b. ¿Cuál es la mejor relación de definición para el diseño? Explique.

7. Se diseña un factorial fraccionado de 2^{5-2}_{III} con una relación de definición $I = ACE = BCDE$. Se sabe que los factores A , B y C no interactúan entre sí y que los factores C , D y E no interactúan entre sí.
 - a. ¿Qué efectos se pueden estimar si se ignoran las interacciones de tres factores o más?
 - b. ¿Es posible tener un mejor diseño de 2^{5-2}_{III} para esta situación? Explique.

8. Un investigador realiza un experimento con ocho corridas usando un factorial fraccionario 2^{5-2} construido con las equivalencias $D = AB$ y $E = ABC$; luego construye otras ocho corridas con los mismos cinco factores pero invierte los signos usados en las ocho corridas originales.
 - a. ¿Cuál es la resolución del diseño con las ocho corridas originales?
 - b. ¿Cuál es la resolución del diseño con 16 corridas?
 - c. Suponga que se sabe desde el principio que se van a usar las 16 corridas para el experimento. ¿Se puede construir un diseño mejor de 16 corridas? Si es así, constrúyalo e indique sus propiedades.

9. Se genera un diseño factorial fraccionario de 2^{6-3}_{III} con las equivalencias $D = AB$, $E = AC$ y $F = ABC$. ¿Cuál es el diseño resultante si se eliminan los factores D y E del diseño?

10. Un proceso en estudio involucra un tratamiento de temperatura en muelles para camión. El muelle pasa por un horno a alta temperatura, se transfiere a una maquinado donde se le da forma al resorte y después se sumerge en aceite. La medición de importancia fue la altura libre del resorte sin carga, misma que se fijó como objetivo en 8 pulgadas. Las desviaciones de este valor se consideraron indeseables. Se realizó un experimento para evaluar el efecto de cuatro factores sobre la desviación de la altura libre del valor meta en el proceso de manufactura, los factores considerados fueron A) temperatura del horno; B) tiempo de calentamiento; C) tiempo de transferencia o tiempo para transferir el ensamblaje del resorte del horno al maquinado; D) tiempo de maquinado o tiempo que la máquina que da la curvatura presiona al resorte

caliente. Además, los ingenieros se interesaron en las interacciones de dos factores AB , AC y BC y en los efectos principales de cada factor. Se usó un factorial fraccionado 2^{4-1} para el experimento, con la relación de definición $I = ABCD$. En la tabla aparecen las ocho combinaciones de tratamientos con la razón señal a ruido (Z) para seis resortes construidos con cada tratamiento. La razón señal ruido es una medida de la desviación de los resortes del valor meta.

Corrida	A	B	C	D	Z
1	+	-	-	+	29.46
2	-	-	+	+	28.11
3	-	-	-	-	28.00
4	+	-	+	-	30.59
5	+	+	+	+	35.31
6	+	+	-	-	38.68
7	-	+	+	-	31.55
8	-	+	-	+	47.70

Fuente: J. J. Pignatell y Ij. S. Romberg (1985), *Journal of Quality Technology* 17, 198-206. Discusión sobre el artículo de R. N. Kackar (1985), Off-line quality control, parameter design, and Taguchi method, *Journal of Quality Technology* 17, 176-188.

- a. ¿Cuál es la resolución de este diseño?
 - b. Muestre cómo se determinaron los signos + y - para el factor D.
 - c. Muestre la estructura de alias para el diseño.
 - d. ¿Qué debe suponerse para obtener estimaciones de las interacciones de dos factores AB , AC y BC ?
 - e. Estime los efectos principales y las interacciones de dos factores de interés y sus errores estándar.
 - f. Interprete los resultados.
11. Un fabricante de sopas instantáneas desea producir paquetes con mezclas de sopas secas con una variación mínima en el peso de los paquetes. Como se identificaron cinco factores que pueden influir en la variación del proceso de llenado, se realizó un experimento fraccionario 2^{5-1} con la relación de definición $I = -ABCDE$ para evaluar los efectos de los factores y sus interacciones. Los factores y niveles fueron: A) el número de puertos en las mezcladoras por donde se agrega aceite vegetal (1 y 3); B) la temperatura que rodea la mezcladora (- = enfriado o + = temperatura ambiente); C) el tiempo de mezclado (60 y 80 segundos); D) el peso del lote (1500 y 2000 lb); E) los días de retraso entre mezclar y empacar (1 y 7). El tamaño de la muestra fue entre 125 y 150 paquetes de sopa de una corrida de producción de ocho horas por cada combinación de tratamientos. Se calculó la desviación estándar del peso de los paquetes como una medida de la variación en el proceso de llenado y se usó como variable de respuesta y. Los niveles de los factores y la desviación estándar del proceso para cada una de las 16 combinaciones de tratamiento son las siguientes:

<i>Corrida</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>y</i>
1	-1	-1	-1	1	1	0.78
2	1	-1	1	1	1	1.10
3	1	1	-1	-1	-1	1.70
4	1	-1	1	-1	-1	1.28
5	-1	1	-1	-1	1	0.97
6	-1	-1	1	-1	1	1.47
7	-1	1	-1	1	-1	1.85
8	1	1	1	1	-1	2.10
9	-1	1	1	1	1	0.76
10	1	1	-1	1	1	0.62
11	-1	-1	1	1	-1	1.09
12	-1	-1	-1	-1	-1	1.13
13	1	-1	-1	-1	1	1.25
14	1	1	1	-1	1	0.98
15	1	-1	-1	1	-1	1.36
16	-1	1	1	-1	-1	1.18

Fuente: L. B. Hare (1988), "In the soup: A case study to identify contributors to filling variability", *Journal of Quality Technology* 20, 36-43.

- a. ¿Cuál es la resolución de este diseño?
 - b. Muestre cómo se determinaron los signos + y - para los niveles del factor *E* en el diseño.
 - c. Muestre la estructura de alias para el diseño.
 - d. ¿Qué suposiciones deben hacerse para estimar los efectos principales y las interacciones de dos factores libres de otros efectos?
 - e. Estime los efectos principales y las interacciones de dos factores que sean de interés y sus errores estándar.
 - f. Interprete los resultados.
 - g. Construya un diseño fraccionado de un cuarto a partir del diseño actual.
12. Se usó un diseño factorial fraccionado Plackett-Burman 2^{7-4} , o de ocho corridas, para estudiar los efectos de siete factores en el consumo de aceite por motores diesel. Se considera que 0.195 gramos/caballo de fuerza-hora (g/hp-h) es un buen control del consumo de aceite. Los siete factores, cada uno con dos

<i>Corrida</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>y</i>
1	-	-	-	+	+	-	+	0.204
2	-	-	+	+	-	+	-	0.662
3	-	+	-	-	+	+	-	1.075
4	-	+	+	-	-	-	+	0.404
5	+	-	-	-	-	+	+	0.445
6	+	-	+	-	+	-	-	1.297
7	+	+	-	+	-	-	-	0.386
8	+	+	+	+	+	+	+	1.157

Fuente: P. R. Shepler (1975), "Fractional factorial plans for diesel engine oil control and for seals", *Proceedings of the 31st National Conference on Fluid Power* 204-234.

niveles, fueron: *A*) ajuste del anillo superior, *B*) cuerda del anillo superior, *C*) cara del anillo intermedio, *D*) tipo de anillo intermedio, *E*) espacio para el pistón de corona, *F*) ancho del anillo de aceite y *G*) uso de empaques. El consumo de aceite en gramos/caballos de fuerza-hora (*y*) se muestra para las ocho corridas:

- ¿Cuál es la resolución de este diseño? ¿Qué implica la resolución de este diseño respecto a la estimación de los efectos factoriales?
- ¿Qué fracción del factorial 2^7 corresponde a este diseño?
- Determine los generadores del diseño.
- Muestre la estructura de alias para los efectos principales del diseño.
- Estime los efectos principales.
- Interprete los resultados.

Como el objetivo era encontrar la mejor configuración de los siete factores para reducir el consumo de aceite, los ingenieros determinaron que la mejor prueba se obtenía con la corrida 1 (0.24 g/hp-h); también consideraron una razón del promedio de respuesta de los niveles – y los niveles + (promedio mayor/promedio menor) para cada factor y la llamaron “razón de aceptación”. Si el promedio mayor se obtuvo con los niveles –, se asigna un signo – a la razón; si el mayor promedio se obtuvo de los niveles +, se asigna un signo + a la razón. Los signos de la razón de aceptación fueron los mismos que para la corrida 1 excepto por el signo del factor *E*, espacio para el pistón. Para verificar los resultados de las pruebas anteriores realizaron corridas adicionales, la primera, con etiqueta “verificar”, incluyó niveles de factor iguales a los de la razón de aceptación excepto por los factores *A* y *G*; la segunda, etiquetada “probar”, incluyó los mismos niveles de factores que los de la razón de aceptación en todos los casos y la tercera fue una repetición de la primera con el factor *E* a nivel +. La experiencia anterior indicó que el espacio para el pistón (nivel +) casi siempre tenía mejor control del aceite. Los resultados de las tres corridas adicionales se muestran junto con la razón de aceptación.

Corrida	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>y</i>
Verificar	+	–	–	+	–	–	–	0.435
Probar	–	–	+	+	–	–	+	0.662
Repetición de 1	–	+	–	–	+	+	–	1.075
Razón de aceptación	–	–	–	+	–	–	–	

- ¿Fueron congruentes las conclusiones de los ingenieros acerca de la razón de aceptación con las estimaciones de los efectos a partir de las corridas de prueba? Explique. ¿En qué circunstancias sería congruente la razón de aceptación con las estimaciones de los efectos principales?
- ¿Fueron las sospechas de los ingenieros acerca del factor *E* correctas? Explique.

Apéndice 12A Planes de diseño factoriales fraccionarios

Tabla 12A.1 Diseños seleccionados de factoriales fraccionados 2^{n-p}

Número de factores	Unidades experimentales	Fracción	Resolución del diseño	Generador de diseño*
3	4	$\frac{1}{2}$	III	$C = AB$
4	8	$\frac{1}{2}$	IV	$D = ABC$
5	8	$\frac{1}{4}$	III	$D = AB$ $E = AC$
	16	$\frac{1}{2}$	V	$E = ABCD$
6	16	$\frac{1}{4}$	IV	$E = ABC$ $F = BCD$
	8	$\frac{1}{8}$	III	$D = AB$ $E = AC$ $F = BC$
7	32	$\frac{1}{4}$	IV	$F = ABCD$ $G = ABDE$
	16	$\frac{1}{8}$	IV	$E = ABC$ $F = BCD$ $G = ACD$
	8	$\frac{1}{16}$	III	$D = AB$ $E = AC$ $F = BC$ $G = ABC$
8	64	$\frac{1}{4}$	V	$G = ABCD$ $H = ABEF$
	32	$\frac{1}{8}$	IV	$F = ABC$ $G = ABD$ $H = BCDE$
	16	$\frac{1}{16}$	IV	$E = BCD$ $F = ACD$ $G = ABC$ $H = ABD$

Tabla 12A.1 (continuación)

<i>Número de factores</i>	<i>Unidades experimentales</i>	<i>Fracción</i>	<i>Resolución del diseño</i>	<i>Generador de diseño*</i>
9	16	$\frac{1}{32}$	III	$E = ABC$ $F = BCD$ $G = ACD$ $H = ABD$ $J = ABCD$
	64	$\frac{1}{8}$	IV	$G = ABCD$ $H = ACEF$ $J = CDEF$
	32	$\frac{1}{16}$	IV	$F = BCDE$ $G = ACDE$ $H = ABDE$ $J = ABCE$
10	16	$\frac{1}{64}$	III	$E = ABC$ $F = BCD$ $G = ACD$ $H = ABD$ $J = ABCD$ $K = AB$
	32	$\frac{1}{32}$	IV	$F = ABCD$ $G = ABCE$ $H = ABDE$ $J = ACDE$ $K = BCDE$
	64	$\frac{1}{16}$	V	$G = BCDF$ $H = AGDB$ $J = ABDE$ $K = ABCE$
	128	$\frac{1}{8}$	V	$H = ABCG$ $J = BCDE$ $K = ACDF$

*Es posible usar la mitad positiva o la negativa de los generadores de diseño para construir el diseño fraccionario.

13 Diseños con superficie de respuesta

El tema central en este capítulo es construir diseños para estimar con eficiencia superficies de respuesta a partir de diseños experimentales con factores cuantitativos. Aquí se estudia la naturaleza de las superficies de respuesta lineales y cuadráticas, se describen los diseños desarrollados especialmente para experimentos con superficie de respuesta, incluyendo la estimación de las ecuaciones para la superficie de respuesta y los métodos para explorar las superficies. También se exponen diseños especiales para experimentos con factores que son ingredientes de mezclas.

13.1 Descripción de respuestas con ecuaciones y gráficas

El objetivo de todos los experimentos incluye describir la respuesta a los factores de tratamiento. En este libro, cuando los factores tuvieron niveles cuantitativos, se determinó la respuesta (y) con respecto al factor (x) mediante la ecuación de polinomio de regresión; por ejemplo, en el capítulo 3 se usa una ecuación polinomial para estimar la relación entre la producción de semillas de plantas, y , y la densidad de las plantas en las parcelas, x . La ecuación de regresión cuadrática estimada se graficó como una curva y se pudo visualizar la respuesta de la producción de semillas con respecto a la densidad de plantas para todas las densidades incluidas en el experimento. Una de las mayores ventajas de la curva de respuesta es que con ella se puede visualizar las respuestas para todos los niveles de los factores en el experimento.

Gráficas de superficie de respuesta para dos factores de tratamiento

La ecuación de respuesta se puede mostrar como una superficie cuando los experimentos investigan el efecto de dos factores cuantitativos, como los efectos de temperatura y presión en la tasa de una reacción química. En el capítulo 6, se estima un polinomio cuadrático para dos factores cuantitativos, salinidad del medio y número de días, para determinar la respuesta de la planta.

La ecuación de respuesta cuadrática se representa como una superficie sólida en tres dimensiones (figura 13.1a), en la figura 13.1b la ecuación aparece como curvas de nivel donde las líneas con valores de respuesta iguales son similares a las elevaciones mostradas en los mapas topográficos.

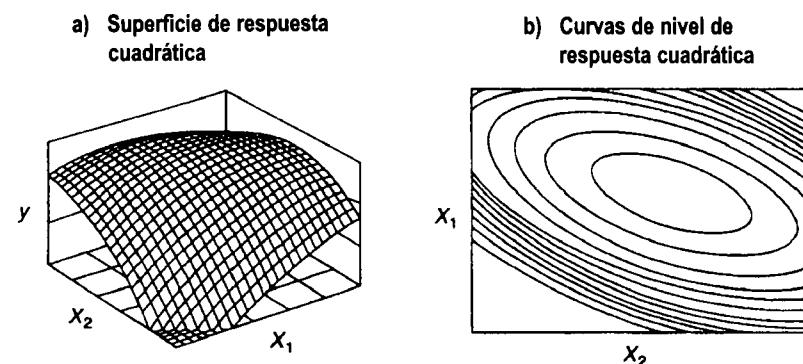


Figura 13.1 Ecuación de respuesta para dos factores, graficada como a) superficie de respuesta y b) curvas con el mismo nivel de respuesta

La superficie de respuesta permite que el investigador inspeccione, de manera visual, la respuesta para cierta zona de los niveles de los factores de interés y evaluar su sensibilidad a los factores de tratamiento. En ciertas aplicaciones industriales, las superficies de respuesta se exploran para determinar la combinación de niveles de los factores que proporcionan una condición operativa óptima, como la combinación de temperatura y tiempo que maximiza la producción química. En otras aplicaciones, se exploran para encontrar la combinación de niveles de los factores que proporcionan mejoras económicas a las respuestas de las condiciones operativas urgentes, si no es demasiado costoso lograr esas condiciones óptimas.

Las superficies de respuesta también se pueden utilizar en estudios analíticos de procesos fundamentales; por ejemplo, suelen usarse en las ciencias biológicas para investigar la influencia mutua de los factores sobre la variable de respuesta, como la interacción entre el nitrógeno y el fósforo en el crecimiento de las plantas.

Modelos polinomiales aproximados a la respuesta real

El diseño de superficies de respuesta y la estrategia de análisis supone que la media de la variable de respuesta μ_y está en función de los niveles de factores cuantitativos representados por las variables x_1, x_2, \dots, x_k . Los modelos polinomiales se usan como aproximaciones prácticas a la función de respuesta verdadera o real. En general, la función real se desconoce y las funciones polinomiales con frecuencia proporcionan buenas aproximaciones en zonas relativamente pequeñas de los niveles de los factores cuantitativos.

Los modelos polinomiales comúnmente empleados para el análisis de superficies de respuesta son el modelo lineal o de *primer orden* y el modelo cuadrático, o de *segundo orden*. El modelo de primer orden para dos factores es:

$$\mu_y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (13.1)$$

y el modelo de segundo orden es:

$$\mu_y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 \quad (13.2)$$

Las gráficas de las curvas de nivel para los modelos de primer orden tienen una serie de líneas paralelas que representan las coordenadas de los niveles del factor que producen los mismos valores de respuesta; para los modelos cuadráticos son más complejas y tienen varios patrones de curvas posibles, uno de ellos, mostrado en la figura 13.1b, es un contorno en forma de montículo con respuesta máxima en el nivel central. En la figura 13.2 se muestran otros cuatro patrones. La figura 13.2a muestra una gráfica de curvas de nivel con la respuesta mínima en el nivel central, lo que indica una superficie simétrica con una depresión en el centro. La figura 13.2b describe una cresta que se eleva hasta un máximo que se encuentra fuera de la región experimental. La figura 13.2c contiene una cresta estacionaria en el centro de la gráfica, con respuestas decrecientes hacia la derecha o la izquierda de la línea central de máxima respuesta. En la figura 13.2d se observa una gráfica con forma de silla de montar o minimax, en donde la respuesta puede aumentar o disminuir a partir del centro de la región, según la dirección en que se aleja del centro.

Experimentos secuenciales para el análisis de superficies de respuesta

Box y Wilson (1951) aceleraron la promoción del análisis de superficies de respuesta en aplicaciones industriales, su tema principal fue el uso de la experimentación secuencial con el propósito de determinar las condiciones operativas óptimas para un proceso industrial.

El enfoque general comienza con diseños de tratamiento factoriales de 2ⁿ para identificar los factores que influyen en el proceso, experimentos subsecuentes usan las combinaciones de tratamientos para localizar un área en el espacio del factor que tenga oportunidad de producir respuestas óptimas, y por último, en esta región se añade un arreglo factorial de 2ⁿ con combinaciones de tratamientos, para determinar la superficie de respuesta con polinomios cuadráticos.

Las escalas de tiempo pueden impedir experimentos secuenciales eficaces

En algunos campos de aplicación, las escalas de tiempo para la terminación de los experimentos no permiten la experimentación secuencial: muchos estudios biológicos pueden requerir meses para completar un solo experimento, pero con frecuencia la información acumulada de estudios previos permite al investigador identificar las regiones de respuesta óptima y diseñar experimentos para explorar la superficie de respuesta en esas regiones.

El objetivo de este capítulo es presentar algunos de los diseños básicos y métodos de análisis usados para identificar las condiciones óptimas en

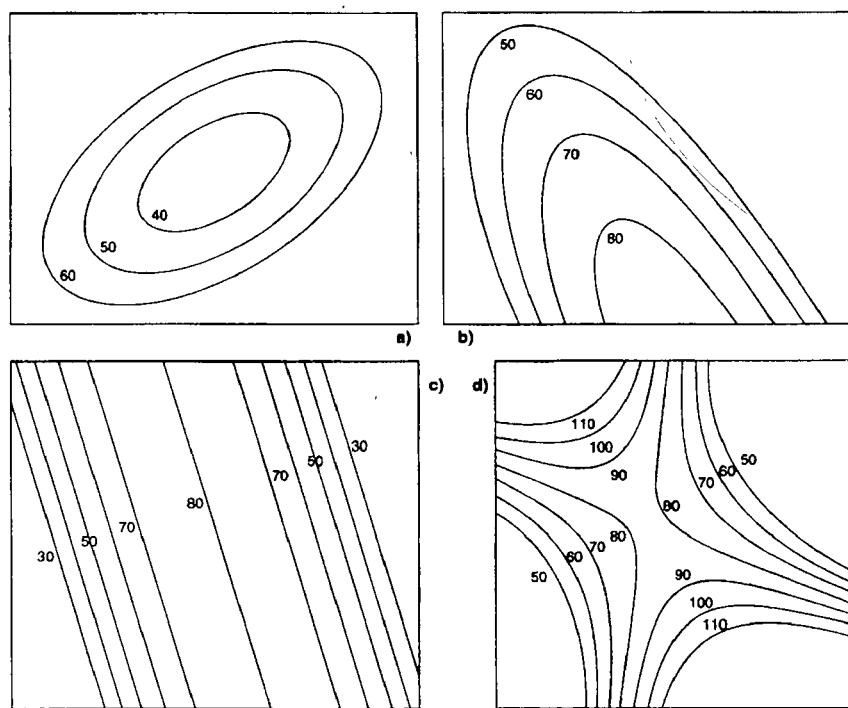


Figura 13.2 Curvas de nivel para a) superficie con un mínimo, b) cresta elevada, c) cresta estacionaria y d) silla de montar

la experimentación secuencial y diseños que permitan la estimación eficaz de las ecuaciones de la superficie de respuesta cuando se identifica la región de respuesta óptima.

13.2 Identificación de los factores significativos con factoriales 2^n

En el estudio de superficies de respuesta, los experimentos factoriales completos o fraccionarios se usan como experimentos iniciales. Cuando la región de respuesta óptima se desconoce, se usan factoriales 2^n o fracciones de ellos para identificar los factores que afectan la variable de respuesta. (Estos diseños se estudian en los capítulos 11 y 12.)

Estimación de las respuestas lineales a los factores

Los factoriales 2^n son diseños adecuados para estimar las respuestas medias para el modelo *lineal*, o de *primer orden* de la ecuación (13.1). La inclusión de dos o más observaciones en el nivel medio de todos los factores suele ser el procedi-

miento recomendado para estimar el error experimental y proporcionar un mecanismo para evaluar si el modelo lineal de superficie de respuesta es apropiado.

Un estudio extenso de los métodos para identificar factores significativos y niveles de factores en la región de condiciones de respuesta óptima se puede encontrar en Box y Draper (1987) y en Myers y Montgomery (1995). Aquí se presenta una breve introducción para identificar los factores importantes en un experimento químico industrial para evaluar los factores que afectan la vinilación de metil glucósidos.

Ejemplo 13.1 Vinilación de metil glucósidos

La vinilación de los glucósidos se presenta cuando se les agrega acetileno a alta presión y alta temperatura, en presencia de una base para producir éteres de monovinil. Los productos de monovinil éter son útiles en varios procesos industriales de síntesis. Los resultados de un estudio de vinilación de metil glucósidos realizado por Marvel *et al.* (1969) se usan para explicar los métodos de identificación y evaluación de los factores significativos para especificar las superficies de respuesta. La meta final del proyecto fue determinar qué condiciones producían una conversión máxima de metil glucósidos para diversos isómeros de monovinil.

Algunos métodos para identificar los factores significativos con superficies de respuesta de primer orden se ilustran con dos factores, presión y temperatura. El diseño de tratamientos fue un factorial 2² con la “temperatura” a 130° y 160°C y la “presión” a 325 y 475 psi como factores, se realizaron cuatro réplicas en el centro de la zona experimental a una temperatura de 145°C y una presión de 400 psi para proporcionar una estimación de la varianza del error experimental y evaluar si el modelo de respuesta lineal era adecuado. Las combinaciones de tratamientos y el porcentaje de conversión del metil glucósido se muestran en la tabla 13.1.

Tabla 13.1 Vinilación de metil glucósidos en un factorial 2ⁿ más cuatro réplicas en el centro del diseño, con temperatura y presión como factores

<i>Factores originales</i>		<i>Factores codificados</i>		<i>% de conversión</i>
<i>Temperatura</i>	<i>Presión</i>	<i>x₁</i>	<i>x₂</i>	
130	325	- 1	- 1	8
160	325	+ 1	- 1	24
130	475	- 1	+ 1	16
160	475	+ 1	+ 1	32
145	400	0	0	21
145	400	0	0	23
145	400	0	0	20
145	400	0	0	24

Niveles de factores codificados por conveniencia

Los niveles de factores codificados proporcionan un marco de trabajo uniforme para investigar los efectos de los factores en cualquier contexto experimental, ya que los valores reales de los niveles dependen de los factores específicos en el estudio. Los niveles codificados de los factores de un diseño factorial 2ⁿ son:

$$x_i = \frac{(A_i - \bar{A})}{D} \quad (13.3)$$

donde A_i es el i -ésimo nivel del factor A ; \bar{A} es el nivel promedio del factor A , y D es $\frac{1}{2}(A_2 - A_1)$. Los niveles codificados de temperatura (T) y presión (P) en la tabla 13.1 son:

$$x_1 = \frac{T - 145}{15} \quad \text{y} \quad x_2 = \frac{P - 400}{15} \quad (13.4)$$

Estimaciones de las respuestas lineales

Las estimaciones de los coeficientes para el modelo de primer orden de la ecuación (13.1) son:

$$\hat{\beta}_0 = \bar{y} = \frac{1}{4}(8 + 24 + 16 + 32) = 20$$

$$\hat{\beta}_1 = \frac{1}{2}T = \frac{1}{4}(-8 + 24 - 16 + 32) = 8$$

$$\hat{\beta}_2 = \frac{1}{2}P = \frac{1}{4}(-8 - 24 + 16 + 32) = 4$$

Las estimaciones de los coeficientes lineales, $\hat{\beta}_1$ y $\hat{\beta}_2$, son la mitad de las estimaciones del efecto del tratamiento factorial para un factorial 2ⁿ (vea el capítulo 11).

La varianza de las cuatro observaciones en el centro del diseño es $s^2 = 3.33$ y una estimación del error estándar para las estimaciones de los coeficientes es:

$$s_{\hat{\beta}} = \sqrt{\frac{4}{16}(3.33)} = 0.91$$

Puede ser importante el hecho de que la varianza del error experimental tenga una estimación adecuada, con réplicas sólo en el centro de los niveles de los factores del diseño. Si la varianza de la respuesta depende en alguna forma del nivel de factor, entonces se recomienda la réplica del diseño con las combinaciones a niveles alto y bajo del factor para detectar cualquier variabilidad heterogénea entre las combinaciones del mismo.

Las estimaciones de los coeficientes de regresión indican que el incremento de la temperatura o la presión, aumentará la vinilación del metil glucósido. La ecuación estimada del modelo de primer orden es:

$$\hat{y} = 20 + 8x_1 + 4x_2$$

La interacción entre temperatura y presión TP mide la falta de ajuste con el modelo lineal y se representa mediante el término $\beta_{12}x_1x_2$ en el modelo cuadrático de la ecuación (13.2). La estimación del coeficiente β_{12} es un medio de la interacción TP , es decir:

$$\hat{\beta}_{12} = \frac{1}{2}TP = \frac{1}{4}(8 - 24 - 16 + 32) = 0$$

El error estándar de $\hat{\beta}_{12}$ es 0.91, el mismo que para los coeficientes lineales. La componente de interacción estimada de 0 indica que temperatura y presión actúan de manera independiente sobre la conversión.

Puntos centrales del diseño para evaluar la curvatura de la superficie

Las observaciones réplica en el centro del diseño no sólo proporcionan una estimación del error experimental, también proporcionan un mecanismo para medir el grado de curvatura en la región experimental. Sea \bar{y}_f la media de las cuatro combinaciones de tratamientos para el factorial 2² y \bar{y}_c la media de los puntos centrales; existe evidencia de la curvatura en la superficie de respuesta si la respuesta promedio en las coordenadas del centro del diseño, \bar{y}_c , es mayor o menor que la respuesta promedio en los niveles extremos de los factores, \bar{y}_f . La diferencia ($\bar{y}_f - \bar{y}_c$) es una estimación de $\beta_{11} + \beta_{22}$, donde β_{11} y β_{22} son los coeficientes de regresión cuadrática en la ecuación (13.2). Las medias observadas son $\bar{y}_f = 20$ y $\bar{y}_c = 22$, con una diferencia de $\bar{y}_f - \bar{y}_c = -2$. El error estándar de la diferencia se estima como $\sqrt{3.33(\frac{1}{4} + \frac{1}{4})} = 1.29$; la respuesta lineal parece describir de manera adecuada la superficie de la zona.

En la figura 13.3 se muestra la gráfica de las curvas de nivel para la ecuación de respuesta lineal estimada. Los valores de los contornos ascienden conforme aumentan los niveles de temperatura y presión, las curvas de nivel crecientes indican que puede existir una combinación de temperatura y presión para maximizar la conversión en una dirección perpendicular a las curvas.

Trayectoria de la mayor pendiente hacia una respuesta óptima

Por último, el investigador querrá determinar la zona de respuesta óptima; para hacerlo, se requiere localizar la región de niveles de los factores que producen las condiciones óptimas. El **método de la mayor pendiente** es un procedimiento desarrollado para llevar la región experimental de la respuesta variable en una dirección de cambio máximo hacia el óptimo.

Con base en la ecuación lineal estimada $\hat{y} = 20 + 8x_1 - 4x_2$, la trayectoria de mayor pendiente, perpendicular a las curvas de igual respuesta, traslada 4 unidades en la dirección de x_2 por cada 8 unidades en la dirección de x_1 . De manera equivalente, la trayectoria tiene un movimiento de $4/8 = 0.5$ unidades en x_2 por cada unidad de movimiento en x_1 .

La trayectoria de mayor pendiente inicia en el centro del diseño con $(x_1, x_2) = (0, 0)$. En la figura 13.3, el centro del diseño para los valores de temperatura y presión es $(T, P) = (145, 400)$. Un cambio de $\Delta x_1 = 1$ unidad en la dirección x_1 es un cambio de 15°C en la temperatura y $\Delta x_2 = 0.5$ unidades en la dirección x_2 es un cambio de 37.5 psi en la presión.

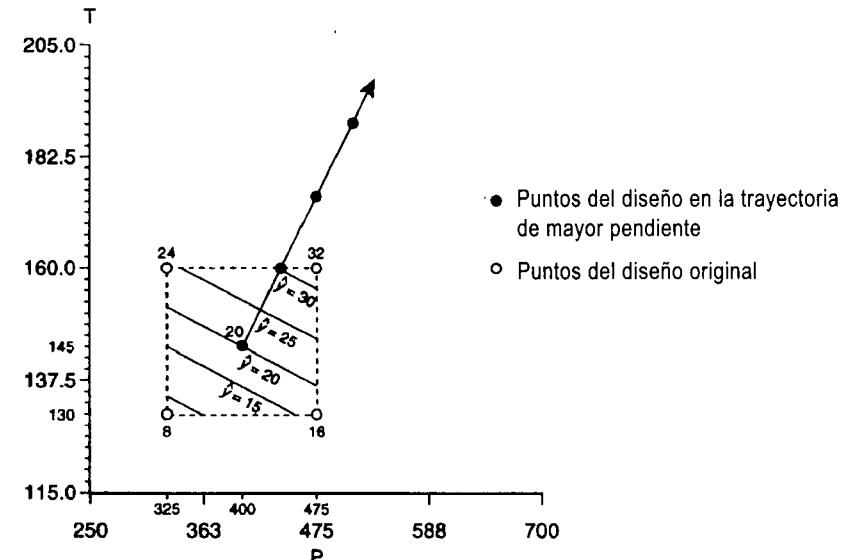


Figura 13.3 Gráfica de curvas de nivel para la respuesta lineal de la vinilación de metil glucósidos, % de conversión para temperatura (T) y presión (P)

El objetivo es moverse a lo largo de la trayectoria de mayor pendiente hasta que se observe una respuesta máxima. El químico realizará experimentos con las combinaciones de temperatura y presión a lo largo de la trayectoria de mayor pendiente. En la tabla 13.2 se muestran, los niveles de temperatura y presión, a partir de $(T, P) = (145, 400)$, centro del diseño, con cambios de 1 unidad en x_1 y de media en x_2 , en el supuesto de que el químico desea realizar los cambios relacionados con la modificación de una unidad en x_1 .

Tabla 13.2 Trayectoria de mayor pendiente para buscar la región de respuesta máxima en la vinilación de metil glucósidos

Paso	x_1	x_2	T	P
0	0	0	145	400.0
1	1	0.5	160	437.5
2	2	1.0	175	475.0
3	3	1.5	190	512.5
4	4	2.0	205	550.0
:	:	:	:	:

Conforme el químico avanza por la trayectoria de mayor pendiente, el aumento en la respuesta es menor hasta que observa una disminución real en ella, lo que indicará que la región de respuesta máxima está en la proximidad de esas condiciones de

temperatura y presión. En este punto del proceso, se puede diseñar un experimento para estimar una ecuación polinomial cuadrática que aproxime la superficie de respuesta.

13.3 Diseños para estimar superficies de respuesta de segundo orden

Una vez que se identifica la región de respuesta óptima, debe diseñarse un nuevo experimento para caracterizar la superficie de respuesta. En general, la superficie se approxima por medio de una ecuación cuadrática para determinar la curvatura de la superficie.

Los factoriales 2^n o sus fracciones son diseños útiles para identificar los factores significativas y las regiones de respuesta óptima. Sin embargo, en la región de respuesta óptima, estos diseños no proporcionan información suficiente para estimar las ecuaciones de respuesta cuadrática, pues se requieren al menos tres niveles para cada factor y el diseño debe tener $1 + 2n + n(n - 1)/2$ puntos distintos para estimar los parámetros con un modelo de regresión cuadrática para aproximar la superficie curva.

Las propiedades de los diseños experimentales convenientes para la estimación de superficies de respuesta incluyen la capacidad para estimar el error experimental y tener en cuenta una prueba de la falta de ajuste del modelo. Los diseños también deben proporcionar estimaciones eficientes de los coeficientes del modelo y predecir las respuestas.

En esta sección se estudian varias clases de diseños desarrollados con las propiedades convenientes para la aproximación de la superficie de respuesta de segundo orden.

Factoriales 3^n para estimar superficies cuadráticas

Los factoriales 3^n se pueden usar para estimar las ecuaciones polinomiales cuadráticas, pero el número de combinaciones de tratamientos que requieren produce un experimento poco práctico de gran tamaño; pues mientras los diseños 3^n con dos factores requieren sólo 9 combinaciones de tratamientos, un diseño con tres factores requiere 27, y uno con cuatro factores requiere 81.

Los diseños centrales compuestos son una alternativa a los factoriales 3^n

Box y Wilson (1951) propusieron **diseños centrales compuestos**, que requieren menos combinaciones de tratamientos que los factoriales 3^n , para estimar las ecuaciones de la superficie de respuesta cuadrática. Los diseños centrales compuestos son diseños de tratamientos factoriales 2^n con $2n$ combinaciones adicionales, llamadas *puntos axiales*, a lo largo de los ejes coordinados de los niveles de factor codificados. Las coordenadas de los puntos axiales de los ejes del factor codificado son $(\pm \alpha, 0, 0, \dots, 0)$, $(0, \pm \alpha, 0, \dots, 0)$, ..., $(0, 0, 0, \dots, \pm \alpha)$. En general, se agregan m réplicas al centro del diseño en las coordenadas $(0, 0, \dots, 0)$.

Los diseños centrales compuestos se usan para aprovechar la experimentación secuencial, el primer paso de la secuencia consiste de una serie de pruebas realizadas a lo largo de la trayectoria de mayor pendiente, como la mostrada en la tabla 13.2. En

algún momento, las pruebas conducen hacia un conjunto de niveles de factores que proporciona un máximo aparente en la trayectoria. Por ejemplo, suponiendo que las respuestas en la trayectoria de mayor pendiente son las que se muestran en la figura 13.4, con una respuesta máxima de 36 observada en $T = 190^{\circ}\text{C}$ y $P = 512.5 \text{ psi}$.

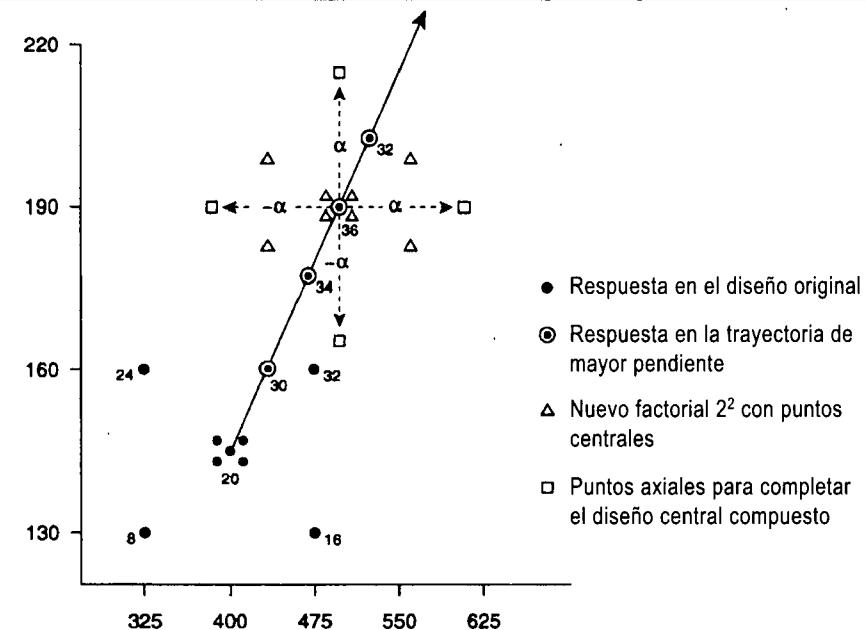


Figura 13.4 Trayectoria de mayor pendiente y un diseño central compuesto

Como segundo paso en la secuencia, el químico realiza un nuevo experimento factorial 2^2 con varias réplicas al centro del diseño ($T, P = 190, 512.5$).

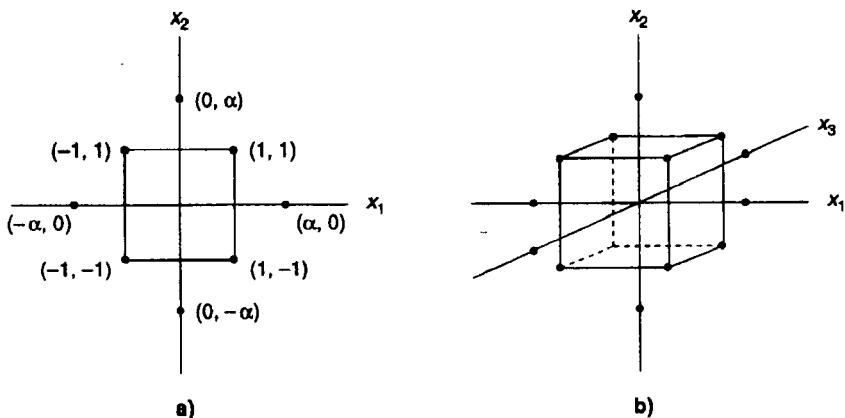
Suponiendo que la diferencia ($\bar{y}_f - \bar{y}_c$) calculada en el nuevo experimento indica un alto grado de curvatura en la superficie, el tercer paso en el experimento secuencial consiste en corridas adicionales del experimento en los puntos axiales ($\pm \alpha, 0$) y $(0, \pm \alpha)$, mostradas con cuadros en la figura 13.4. Este último conjunto de combinaciones de tratamientos en los ejes, junto con el factorial 2^2 y los puntos centrales, constituye un diseño central compuesto como resultado de la experimentación secuencial.

Una réplica de un diseño central compuesto consiste de $N_f = 2^n$ combinaciones de tratamientos del factorial 2^n , $N_a = 2n$ combinaciones de tratamientos en los puntos axiales del diseño y m réplicas en el centro para obtener un total de $N = N_f + N_a + m$ observaciones.

Las coordenadas en los ejes codificados x_1 y x_2 para el diseño central compuesto con dos factores se muestra en el cuadro 13.1, y la figura 13.5 describe la gráfica de la localización de las coordenadas para los niveles de factores codificados del diseño central compuesto de dos y tres factores. Debido a que cada factor tiene cinco niveles, se puede estimar una ecuación cuadrática a partir de este diseño.

Cuadro 13.1 Coordenadas del diseño central compuesto

Diseño 2^2		Axial		Central*	
x_1	x_2	x_1	x_2	x_1	x_2
- 1	- 1	- α	0	0	0
+ 1	- 1	+ α	0		
- 1	+ 1	0	- α		
+ 1	+ 1	0	+ α		

* m réplicas**Figura 13.5** Diseños centrales compuestos para a) dos factores y b) tres factores

Además, como se verá en la siguiente sección, se puede evaluar cualquier desviación significativa de la aproximación cuadrática.

Las $N = 2^n + 2n + m$ unidades experimentales necesarias para el diseño central compuesto con n factores son menos que las requeridas por los factoriales 3^n con tres factores o más. Así, los diseños centrales compuestos son más económicos en cuanto al uso de recursos experimentales y proporcionan la capacidad de estimar las ecuaciones de respuesta. Se pueden usar fracciones de los diseños 2^n con interacciones de orden mayor con alias como base del diseño 2^n cuando hay más factores en el estudio.

Diseños rotatorios para mejorar las exploraciones de la superficie de respuesta

Una propiedad deseable al establecer cualquier diseño es la misma precisión para todas las estimaciones de las medias. Sin embargo, la precisión de los valores estimados sobre la superficie de respuesta basados en la ecuación de regresión estimada no será constante en toda la región experimental. Una propiedad

rotatoria desarrollada para los diseños centrales compuestos requiere que la varianza de los valores estimados sea constante en puntos equidistantes del centro del diseño con coordenadas codificadas $(0, 0, \dots, 0)$.

Lo rotatorio de un diseño cobra importancia en la exploración de una superficie de respuesta porque la precisión de la superficie estimada no depende de la orientación del diseño con respecto a la superficie de respuesta real o a la dirección de la búsqueda de las condiciones óptimas. Los factoriales 2^n usados como diseño de primer orden para aplicar el método de búsqueda de la mayor pendiente en zonas de respuestas óptimas son diseños rotatorios. Así, la orientación del diseño no dificulta el método de búsqueda de la mayor pendiente porque algunas respuestas se estiman con menor precisión que otras.

El diseño central compuesto se puede hacer rotatorio estableciendo los valores de los puntos axiales como $\alpha = (2^n)^{1/4}$. El valor de α para un diseño de dos factores es $\alpha = (4)^{1/4} = \sqrt{2} = 1.414$, y para un diseño de tres factores $\alpha = (8)^{1/4} = 1.682$. Si hay r_f réplicas del factorial 2^n y r_a réplicas de las combinaciones de tratamiento axiales, una forma más general para α es $\alpha = (r_f 2^n / r_a)^{1/4}$, si se usa un factorial fraccionario 2^{n-p} como base para el diseño central compuesto, entonces $\alpha = (r_f 2^{n-p} / r_a)^{1/4}$.

Ejemplo 13.2 Diseño rotatorio para la vinilación de metil glucósidos

Suponiendo que la trayectoria de la mayor pendiente para el estudio de vinilación en la tabla 13.2 proporcionó una respuesta máxima en $T = 190^\circ\text{C}$ y $P = 512.5$ psi y debe construirse un diseño central compuesto con centro en $(T, P) = (190, 512.5)$; y que la relación entre las coordenadas del diseño (x_1, x_2) y los niveles de temperatura y presión (T, P) se conservan como antes, donde un cambio de una unidad en x_1 es 15°C y un cambio de una unidad en x_2 es 75 psi. Con $\alpha = \sqrt{2}$, las coordenadas del diseño y la temperatura y presión requeridas serán:

	<i>Axial</i>				<i>Central</i>	<i>Diseño 2^n</i>			
x_1	$-\sqrt{2}$	$+\sqrt{2}$	0	0	0	-1	+1	-1	+1
x_2	0	0	$-\sqrt{2}$	$+\sqrt{2}$	0	-1	-1	+1	+1
T	169*	211	190	190	190	175	205	175	205
P	512.5	512.5	406.4	618.6	512.5	437.5	437.5	587.5	587.5

*Ejemplo de cálculo, $169^\circ = 190^\circ - \sqrt{2}(15^\circ)$

Diseños para obtener una precisión uniforme en el centro del diseño

Como se estableció antes, la varianza de la superficie estimada no es constante para toda la superficie. Box y Hunter (1957) mostraron que el número de puntos centrales en los diseños centrales compuestos rotatorios puede elegirse de manera que proporcione un diseño con precisión uniforme para la superficie estimada de una unidad alrededor de las coordenadas del centro del diseño en la escala codificada. Su razonamiento fue que el investigador está más interesado en la superficie de respuesta cerca del centro del diseño cuando un punto estacionario de la superficie se localiza cerca del centro; el punto estacionario es un punto de respuesta

máxima, mínima o con forma de silla, como se muestra en la figura 13.2d. En la tabla 13.3 se muestran algunos diseños centrales rotatorios compuestos con precisión uniforme.

Tabla 13.3 Diseños centrales compuestos rotatorios con precisión uniforme

Número de factores	2	3	4	5	5	6	6
Factor 2^n	1	1	1	1	$\frac{1}{2}$	1	$\frac{1}{2}$
α	1.414	1.682	2.000	2.378	2.000	2.828	2.378
N_f	4	8	16	32	16	64	32
N_a	4	6	8	10	10	12	12
m	5	6	7	10	6	15	9
N	13	20	31	52	32	91	53

Los diseños centrales compuestos requieren cinco niveles de cada factor, codificados como $-\alpha, -1, 0, 1, \alpha$. En algunos casos, preparar cinco niveles para algunos factores puede ser difícil, costoso y tardado. El *diseño de cubo con cara centrada* es una variación del diseño central compuesto con $\alpha = 1$ que requiere sólo tres niveles de cada factor. Si se sustituye $\alpha = 1$ en el cuadro 13.1, el diseño de dos factores se convierte en un factorial 3^2 , diseño más atractivo cuando la región de interés tiene forma de cubo producida por este diseño en lugar de la región esférica producida por el diseño central compuesto.

El diseño no es rotatorio, pero la ausencia de esta propiedad se compensa por el deseo de poder hacer inferencias cuboidales y por el ahorro en recursos experimentales. El cubo con cara centrada requiere menos corridas en el punto central que el diseño central compuesto para lograr una varianza estable de los valores estimados en la región del diseño, pero debe recordarse que se necesitan corridas réplica en algún punto o puntos del diseño para estimar la varianza del error experimental. Un diseño de cubo con cara centrada para tres factores o más requiere menos combinaciones de tratamientos que los factoriales 3^n ; entonces, ésta es otra alternativa para los diseños 3^n que requiere menos unidades experimentales.

Diseños Box-Behnken, otra alternativa para los factoriales 3^n

Box y Behnken (1960) propusieron una clase de diseños de tres niveles para estimar las superficies de respuesta de segundo orden. Los diseños son rotatorios, o casi rotatorios, con menor número de unidades experimentales en comparación con los diseños 3^n . Se forman con la combinación de diseños 2^n y diseños de bloques incompletos; los detalles de la construcción se encuentran en Box y Draper (1987) y los niveles de factores codificados para las combinaciones de tratamientos necesarios en un diseño para tres factores se muestran en el cuadro 13.2, donde se presenta un conjunto completo de las combinaciones de tratamientos para un factorial 2^n para cada par de factores acompañados por el nivel 0 de los factores restantes. Se incluyen varias réplicas del centro del diseño (0, 0, ..., 0).

Cuadro 13.2 Diseño coordinado Box-Nehnken para un diseño de tres factores

Factor	A	B	C
Nivel codificado	x_1	x_2	x_3
	-1	-1	0
Factorial 2^2	+1	-1	0
para A y B	-1	+1	0
	+1	+1	0
	-1	0	-1
Factorial 2^2	+1	0	-1
para A y C	-1	0	+1
	+1	0	+1
	0	-1	-1
Factorial 2^2	0	+1	-1
para B y C	0	-1	+1
	0	+1	+1
	0	0	0
Centro del diseño	0	0	0
	0	0	0

Estos diseños son esféricos más que cuboidales puesto que los puntos del diseño se encuentran en las orillas de un cubo en lugar de en las esquinas como los del diseño de cubo con cara centrada. El diseño de Box-Behnken sólo debe usarse si no se tiene interés en predecir las respuestas en las esquinas de la región cuboidal.

Diseños de bloques incompletos para aumentar la precisión

Los diseños de bloques incompletos son útiles para reducir la varianza del error experimental cuando el número de tratamientos es grande o cuando las condiciones experimentales impiden la ejecución de réplicas completas bajo circunstancias constantes.

Box y Hunter (1957) presentaron las condiciones para bloquizar los diseños de superficie de respuesta de segundo orden, de manera que los efectos de los bloques no afectan las estimaciones de los parámetros para la ecuación de la superficie de respuesta. Mostraron que deben satisfacerse dos condiciones para que los bloques sean ortogonales a las estimaciones de los parámetros de la ecuación de la superficie de respuesta. Sea n_b el número de tratamientos en el b -ésimo bloque; las dos condiciones necesarias son:

1. Cada bloque debe ser un diseño ortogonal de primer orden. Para cada bloque debe cumplirse la siguiente relación para cada par de variables de diseño x_i y x_j :

$$\sum_{k=1}^{n_b} x_{ik} x_{jk} = 0 \quad i \neq j = 0, 1, 2, \dots, n \quad (13.5)$$

2. La fracción de la suma de cuadrados total para cada variable de diseño que aporta cada bloque debe ser igual a la fracción de las observaciones totales colocadas en el bloque. Entonces, debe cumplirse la siguiente relación entre las variables de diseño y el número de observaciones para cada bloque:

$$\frac{\sum_{k=1}^{n_b} x_{ik}^2}{\sum_{k=1}^N x_{ik}^2} = \frac{n_b}{N} \quad i = 1, 2, \dots, n \quad (13.6)$$

Una estrategia sugerida para los bloques de diseños centrales compuestos coloca los tratamientos N_f para el diseño 2^n y m_f puntos centrales en un bloque, y los N_a tratamientos axiales con m_a puntos centrales en un segundo bloque. Este arreglo de bloques satisface la primera condición (ecuación 13.5).

El diseño rotatorio central compuesto para dos factores dispuestos en dos bloques se muestra en el cuadro 13.3. El primer bloque se compone de $N_f = 4$ combinaciones de tratamientos del factorial de 2^n más $m_f = 2$ puntos centrales, y el segundo bloque consiste en $N_a = 4$ combinaciones de tratamientos axiales más $m_a = 2$ puntos centrales. Los cálculos necesarios para evaluar la primera condición de un diseño de bloques ortogonal son las sumas de los productos cruzados entre x_1 y x_2 en cada bloque. Es sencillo verificar que $\sum x_1 x_2 = 0$ en ambos bloques.

Cuadro 13.3 Diseño central compuesto rotatorio para dos factores en dos bloques incompletos

Factor Nivel codificado	A		B
	x_1	x_2	
Bloque 1	-1	-1	
	+1	-1	
	-1	+1	
	+1	+1	
	0	0	
	0	0	
Bloque 2	1.414	0	
	-1.414	0	
	0	1.414	
	0	-1.414	
	0	0	
	0	0	

Para el diseño completo:

$$\sum_{k=1}^{12} x_{1k}^2 = \sum_{k=1}^{12} x_{2k}^2 = 8$$

y tanto para el bloque 1 como el bloque 2:

$$\sum_{k=1}^6 x_{1k}^2 = \sum_{k=1}^6 x_{2k}^2 = 4$$

El número de observaciones del tratamiento en los bloques 1 y 2 es $n_1 = n_2 = 6$ y el número total de observaciones es $N = 12$, con una razón $n_i/N = 6/12 = 1/2$. La segunda condición, la ecuación (13.6), requiere de la razón de las sumas de cuadrados de x_1 y x_2 en cada bloque para que todo el experimento sea igual a n_i/N . Para ambos bloques la razón de la suma de cuadrados es $4/8 = 1/2$, que es equivalente a la razón para n_i/N ; por tanto, el diseño es ortogonal.

Para que se satisfaga la segunda condición, debe cumplirse la siguiente relación:

$$\alpha^2 = n \left[\frac{1 + p_a}{1 + p_f} \right] \quad (13.7)$$

donde $p_a = m_a/N_a$ y $p_f = m_f/N_f$. Para que el diseño satisfaga las dos condiciones y sea rotatorio $\alpha = (2^n r_f/r_a)^{1/4}$. No siempre es posible encontrar un diseño que cumpla con exactitud la ecuación (13.7) con $\alpha = (2^n r_f/r_a)^{1/4}$, pero en la práctica, los valores del número de observaciones del diseño se pueden determinar de forma que se obtengan diseños con bloques casi ortogonales y rotatorios. Box y Draper (1987) ofrecen las proporciones relativas de r_f y r_a necesarias para los diseños rotatorios y bloques ortogonales cuando $p_a = p_f$.

Para el diseño del cuadro 13.3, la fracción de observaciones centrales en cada bloque es $p_a = p_f = 1/2$ y $\alpha = \sqrt{2}$.

Al evaluar la condición de rotabilidad y ortogonalidad en la ecuación (13.7) se tiene:

$$\alpha^2 = n \left[\frac{1 + p_a}{1 + p_f} \right] = 2 \left[\frac{1 + 1/2}{1 + 1/2} \right] = 2$$

y $\alpha = \sqrt{2}$, como lo requiere la rotabilidad.

Los diseños centrales compuestos rotatorios enumerados en la tabla 13.3 se pueden colocar en diseños de bloques incompletos útiles para obtener diseños centrales compuestos casi ortogonales y rotatorios. El factorial 2^n o el factorial fraccionario 2^{n-p} se coloca en uno o más bloques incompletos y las combinaciones de tratamiento axiales se colocan en un bloque separado. En la tabla 13.4 se muestran el número de bloques y el número de puntos centrales sugeridos en cada bloque para el factorial 2^n o el factorial fraccionario.

Tabla 13.4 Diseños de bloques incompletos para obtener diseños centrales compuestos casi ortogonales y rotatorios

<i>Número de factores</i>	2	3	4	5	5	6	6
<i>Fracción de 2^n</i>	1	1	1	1	$\frac{1}{2}$	1	$\frac{1}{2}$
N_f	4	8	16	32	16	64	32
m_f	2	2	2	4	2	2	2
<i>Número de bloques*</i>	1	2	2	4	1	8	2
α	1.414	1.682	2.000	2.378	2.000	2.828	2.378
N_a	4	6	8	10	10	12	12
m_a	2	2	1	1	4	1	4

*Vea en el apéndice 11A los contrastes de definición para bloquear la porción del diseño factorial 2^n o del factorial fraccionario.

Reducción del número de puntos de diseño

El costo, la dificultad o el tiempo consumido con cierto tipo de experimentos pueden obligar a reducir el tamaño del experimento, pero tal reducción está limitada por el modelo estadístico que estima la superficie de respuesta. La ecuación de la superficie de respuesta de segundo orden para n factores tiene un término constante, n términos lineales, n términos cuadráticos y $n(n - 1)/2$ términos de interacción, con un total de $(n + 1)(n + 2)/2$ términos. Así, el número mínimo de puntos de un diseño para estimar la superficie de respuesta de segundo orden es $(n + 1)(n + 2)/2$.

Se han desarrollado diseños que tienen números de puntos cercanos al mínimo para estimar una superficie de respuesta de segundo orden; las tablas de estos diseños o los métodos para construirlos se pueden encontrar en Box y Draper (1974), Roquemore (1976), Notz (1982), Draper (1985), Draper y Lin (1990) y Myers y Montgomery (1995).

La mayoría de los diseños se basan en factoriales fraccionarios 2^{n-p} incrementados con puntos de diseño para estimar los modelos de superficie de respuesta de segundo orden. En muchos casos los diseños se saturan con puntos de diseño con pocas o ninguna réplica y se requiere una estimación independiente del error experimental para probar la eficacia del modelo de la superficie de respuesta, a menos que el diseño tenga réplicas. Además, los diseños saturados no permiten probar la falta de ajuste del modelo hipotético de la superficie de respuesta de segundo orden.

Evaluación de los diseños de superficie de respuesta

Myers et al. (1992) usaron la predicción de la varianza de la ecuación de la superficie de respuesta de segundo orden para evaluar muchos de los diseños conocidos de esta superficie, considerando que un diseño era superior si la varianza de los valores pronosticados era menor que la de los otros diseños.

Los diseños centrales compuestos fueron superiores en general para superficies esféricas cubiertas por puntos de diseño (vea la figura 13.5). Cuando los diseños se restringieron a las regiones cuboidales ($\alpha = 1$ en la figura 13.5), el diseño

de cubo con cara centrada formado por el diseño central compuesto, en general, era superior que el diseño de Box-Behnken en la región cuboidal.

Entre los diseños saturados, los diseños de Roquemore (1976), Notz (1982) y Box y Draper (1974) fueron bastante eficientes en relación con los demás.

Myers y Montgomery (1995) presentaron las eficiencias de los diseños para estimar los coeficientes del modelo y para predecir las varianzas en una región esférica. Sus conclusiones generales fueron que los diseños central compuesto y de Box-Behnken eran bastante eficientes lo mismo que algunos diseños saturados de Roquemore (1976).

13.4 Estimación de la superficie de respuesta cuadrática

Cuando se ha identificado la supuesta región de respuesta óptima mediante el método de la mayor pendiente o algún otro método de experimentación, suele ser necesario determinar la superficie de respuesta en esa región de los factores. Con los diseños descritos en la sección anterior, se pueden realizar experimentos y obtener datos para estimar una aproximación cuadrática de la superficie de respuesta.

La ecuación de respuesta estimada permitirá al investigador localizar un punto de respuesta estacionario que quizás sea un máximo, un mínimo o un punto de deflexión en la superficie. Un examen de las curvas de nivel indicará qué tan sensible es la variable de respuesta a cada factor y el grado en que los factores afectan a las variables de respuesta.

Ejemplo 13.3 Respuesta de la duración de herramientas a la velocidad del torno y la profundidad de corte

Una compañía usaría una nueva herramienta de corte que ofrece un proveedor, éste asegura que la nueva herramienta reducirá los costos de producción porque durará más que el modelo anterior y el costo de reemplazo de la herramienta se reducirá. La vida de una herramienta de corte de metales depende de varias condiciones de operación como la velocidad del torno y la profundidad de corte.

El ingeniero de planta había determinado mediante estudios anteriores que la vida máxima de la herramienta se lograba, para la herramienta actual, con una velocidad de 400 y una profundidad de corte de 0.075. El ingeniero, que deseaba determinar la situación óptima para la nueva herramienta, usó un diseño central compuesto en un experimento para determinar la vida de la nueva herramienta al variar las velocidades del torno y las profundidades de corte dentro de la región de condiciones de operación óptimas urgentes para la vida máxima de la herramienta. Los datos del experimento se muestran en la tabla 13.5.

Tabla 13.5 Duración de una herramienta, a partir de un experimento factorial con la velocidad del torno y la profundidad de corte como factores de tratamiento, en un diseño central compuesto

<i>Factores originales</i>		<i>Factores codificados</i>		<i>Vida de la herramienta</i>
<i>Velocidad del torno</i>	<i>Profundidad del corte</i>	x_1	x_2	
600	0.100	+1	+1	154
600	0.050	+1	-1	132
200	0.100	21	11	166
200	0.050	21	21	83
683	0.075	$\sqrt{2}$	0	156
117	0.075	$-\sqrt{2}$	0	144
400	0.110	0	$\sqrt{2}$	166
400	0.040	0	$-\sqrt{2}$	91
400	0.075	0	0	167
400	0.075	0	0	175
400	0.075	0	0	170
400	0.075	0	0	176
400	0.075	0	0	156
400	0.075	0	0	170

Ecuación de superficie de respuesta estimada

El modelo de la superficie de respuesta de segundo orden de la ecuación (13.2) se ajusta a los datos mediante los procedimientos de regresión de mínimos cuadrados. La ecuación se puede estimar con un programa de computadora para análisis de regresión y el apéndice 13A.1 contiene un breve resumen de la estimación de mínimos cuadrados para los modelos de regresión. Una presentación detallada del análisis de regresión se encuentra en Rawlings (1988). La ecuación de la superficie de respuesta de segundo orden estimada para los datos de la vida de la herramienta de la tabla 13.5 es:

$$\hat{y} = 169 + 6.747x_1 + 26.385x_2 - 10.875x_1^2 - 21.625x_2^2 - 15.250x_1x_2$$

Particiones de sumas de cuadrados para el análisis de regresión

Las particiones de sumas de cuadrados en el análisis de varianza para el modelo de regresión se muestran en la tabla 13.6. Las sumas de cuadrados para el modelo de segundo orden completo es:

$$SCR(x_1, x_2, x_1^2, x_2^2, x_1x_2) = 10,946$$

Se hace una partición de la suma de cuadrados de regresión en reducciones para el modelo lineal y las componentes cuadráticas del modelo, con el principio de particiones de sumas de cuadrados del modelo reducido y el modelo completo.

La partición para las componentes lineales del modelo, x_1 y x_2 , o:

$$SCR(x_1, x_2) = 5,933$$

Tabla 13.6 Análisis de varianza para el modelo de superficie de respuesta cuadrático

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Total	13	11,317	
Regresión	5	10,946	2,189.2
Lineal (x_1, x_2)	2	5,933	2,966.5
Cuadrática (x_1^2, x_2^2, x_1x_2)	3	5,013	1,671.0
Error	8	371	46.4
Falta de ajuste	3	111	37.0
Error puro	5	260	52.0

es la suma de cuadrados de la regresión para el modelo reducido de primer orden $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + e$. La partición para las componentes cuadráticas es la diferencia entre la suma de cuadrados de regresión para el modelo completo y el modelo reducido, es decir:

$$SCR(x_1^2, x_2^2, x_1x_2 | x_1, x_2) = 10,946 - 5,933 = 5,013$$

Se hace una partición de la suma de cuadrados para el error, $SCE = 371$ en dos partes. La suma de cuadrados para el error experimental puro, $SCE(\text{error puro}) = 260$, con 5 grados de libertad se calcula a partir de las seis réplicas observadas en el centro del diseño con coordenadas de factor (V, D) = (400, 0.075). Las suma de cuadrados para el error con los 3 grados de libertad restantes, $SCE(\text{falta de ajuste}) = 111$, se pueden atribuir al error en la especificación del modelo de superficie de respuesta. Como los seis puntos centrales del diseño proporcionan una estimación del error experimental puro, la suma de cuadrados designada como falta de ajuste se puede usar para probar la significancia de la falta de ajuste en el modelo cuadrático.

Pruebas de hipótesis sobre el modelo de segundo orden

Las hipótesis de interés en el análisis son:

- Significancia del modelo completo de segundo orden:

$$H_0: \beta_1 = \beta_2 = \beta_{11} = \beta_{22} = \beta_{12} = 0$$

$$F_0 = \frac{2,189.2}{52} = 42.1 \quad H_0 \text{ se rechaza, ya que } F_0 > F_{0.05,5,5} = 5.05$$

- Significancia de las componentes lineales para el modelo:

$$H_0: \beta_1 = \beta_2 = 0$$

$$F_0 = \frac{2,189.2}{52} = 57.0 \quad H_0 \text{ se rechaza, ya que } F_0 > F_{0.05,2,5} = 5.79$$

- Significancia de las desviaciones cuadráticas del modelo lineal:

$$H_0: \beta_{11} = \beta_{22} = \beta_{12} = 0$$

$$F_0 = \frac{1,671}{52} = 32.1 \quad H_0 \text{ se rechaza, ya que } F_0 > F_{0.05,3,5} = 5.41$$

- Significancia de la falta de ajuste al modelo cuadrático:

$$F_0 = \frac{37}{52} = 0.71 \quad H_0 \text{ se rechaza, ya que } F_0 > F_{0.05,3,5} = 5.41$$

El modelo de regresión cuadrática completo es significativo y la falta de ajuste al modelo cuadrático no lo es; entonces se puede concluir que el modelo de segundo orden es una aproximación adecuada a la superficie de respuesta real. Una gráfica de curvas de nivel del modelo de superficie de respuesta cuadrático descrito en la figura 13.6 muestra una superficie máxima con la máxima duración de la herramienta en el centro de las curvas.

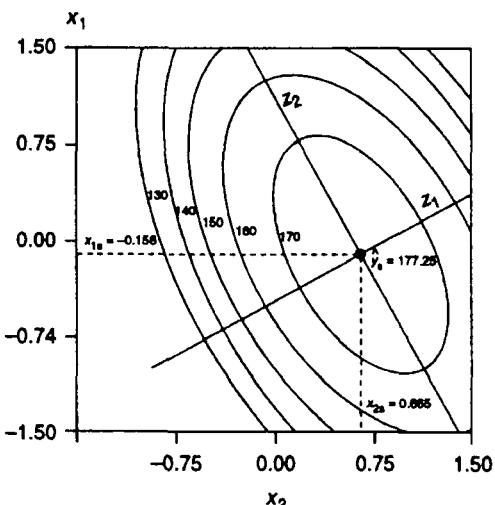


Figura 13.6 Gráfica de curvas de nivel de la superficie de respuesta para la ecuación de respuesta en el experimento de la vida de la herramienta, $\hat{y} = 169 + 6.747x_1 + 26.385x_2 - 10.875x_1^2 - 21.625x_2^2 + 15.250x_1x_2$

Las coordenadas de la gráfica de las curvas de nivel se despliegan para los valores codificados de los dos factores. La orientación de los contornos indica cierta interacción entre la velocidad del torno x_1 y la profundidad de corte x_2 ; por ejemplo, la vida de una herramienta de corte de 150 se puede mantener para velocidades mayores, si se incrementa x_1 y se disminuye la profundidad de corte, x_2 .

Las curvas también indican la sensibilidad relativa de la vida de la herramienta a los niveles de los factores codificados x_1 y x_2 . Las curvas de la vida de la herramienta aumentan con mayor rapidez cerca del máximo sobre el eje de profundidades codificadas x_2 que sobre el eje de las velocidades codificadas x_1 .

13.5 Exploración de superficies de respuesta

La ecuación cuadrática significativa y la gráfica de curvas de nivel de la ecuación proporcionan un panorama general de la relación entre la vida de la herramienta y los dos factores del diseño, velocidad del torno y profundidad de corte.

Las estimaciones de las coordenadas del punto estacionario en la superficie y una estimación de la respuesta en ese punto proporcionan una definición más específica de la superficie de respuesta. En ocasiones, es útil conocer la dirección y cantidad de cambio hecho en uno o varios niveles de los factores para lograr el cambio máximo en la respuesta.

Es posible determinar de manera más específica la sensibilidad de la respuesta a los factores del diseño con la forma canónica de la ecuación de respuesta. Localizar las coordenadas del punto estacionario y derivar la forma canónica de la ecuación de respuesta requieren cierto conocimiento de cálculo y álgebra matricial. Sin embargo, los resultados de los cálculos se entienden cuando se presentan con la forma gráfica de la figura 13.6.

Localización de las coordenadas del punto estacionario de la superficie de respuesta

Las coordenadas x_1 y x_2 del punto estacionario se obtienen de las derivadas parciales de la función de respuesta estimada respecto a x_1 y x_2 . La respuesta estimada para la duración de la herramienta es:

$$\hat{y} = 169 + 6.747x_1 + 26.385x_2 - 10.875x_1^2 - 21.625x_2^2 - 15.250x_1x_2 \quad (13.8)$$

Las derivadas parciales se igualan a 0:

$$\frac{\partial \hat{y}}{\partial x_1} = 0 \quad \frac{\partial \hat{y}}{\partial x_2} = 0$$

para producir las ecuaciones:

$$2(-10.875)x_1 + (-15.250)x_2 = -6.747$$

$$(-15.250)x_1 + 2(-21.625)x_2 = -26.385$$

Las soluciones de las ecuaciones para x_1 y x_2 son $\hat{x}_{1s} = -0.156$ y $\hat{x}_{2s} = -0.665$. Estos valores son las coordenadas de la respuesta máxima sobre la superficie en el punto estacionario indicado en la figura 13.6.

La respuesta estimada en el punto estacionario se encuentra al sustituir $\hat{x}_{1s} = -0.156$ y $\hat{x}_{2s} = 0.665$ en la ecuación (13.8); la respuesta estimada en el punto estacionario es:

$$\begin{aligned}\hat{y}_s &= 169 + 6.747(-0.156) + 26.385(0.665) - 10.875(-0.156)^2 \\ &\quad - 21.625(0.665)^2 - 15.250(-0.156)(0.665) = 177.25\end{aligned}$$

Dados $x_1 = (V - 400)/200$ y $x_2 = (D - 0.075)/0.025$, los valores de la velocidad del torno (V) y la profundidad de corte (D) en el punto estacionario son:

$$V = -0.156(200) + 400 = 368.8$$

y

$$D = 0.665(0.025) + 0.075 = 0.092$$

La solución general para un punto estacionario con cualquier número de variables x_i en la ecuación de respuesta se encuentra en el apéndice 13A.2.

Análisis canónico para simplificar la ecuación cuadrática

La forma canónica de una ecuación cuadrática es eficaz para visualizar la superficie y determinar la sensibilidad relativa de las variables de respuesta a cada factor. Es difícil visualizar la superficie mediante el examen de los coeficientes estimados para la forma normal de la ecuación de respuesta cuadrática. De la misma manera, es difícil determinar los cambios necesarios en los niveles de los factores para producir un cambio específico en la respuesta.

El análisis canónico gira los ejes de las variables x , a un nuevo sistema de coordenadas y el centro de este nuevo sistema se coloca en el punto de respuesta estacionario de la superficie. La forma canónica de la ecuación con dos variables es:

$$\hat{y} = \hat{y}_s + \lambda_1 Z_1^2 + \lambda_2 Z_2^2 \quad (13.9)$$

donde Z_1 y Z_2 son las variables de los ejes rotados. Observe que sólo se incluyen los términos cuadráticos de las variables canónicas Z_1 y Z_2 en la forma *canónica* de la ecuación de respuesta. El apéndice 13A.3 presenta una descripción de los cálculos necesarios para obtener la forma *canónica* de la ecuación de respuesta para la duración de la herramienta.

La forma *canónica* para la ecuación de respuesta de la vida de la herramienta es:

$$\hat{y} = 177.25 - 25.58Z_1^2 - 6.92Z_2^2 \quad (13.10)$$

donde el centro del nuevo sistema de coordenadas se localiza en $x_1 = -0.156$ y $x_2 = 0.665$ en el sistema de coordenadas original mostrado en la figura 13.6. Se determinó que la relación entre los dos sistemas de coordenadas (apéndice 13A.3) es:

$$Z_1 = 0.4603x_1 + 0.8877x_2 - 0.5185$$

$$Z_2 = 0.8877x_1 - 0.4603x_2 + 0.4446 \quad (13.11)$$

Observe que los ejes *canónicos* Z_1 y Z_2 están orientados junto con las curvas de nivel de la superficie. Los tamaños y signos de las λ_i indican el tipo de superficie de respuesta cuadrática que se estimó.

Los coeficientes λ_i para la superficie de la vida de la herramienta son $\lambda_1 = -25.58$ y $\lambda_2 = -6.92$, un examen de la superficie en la figura 13.6 revela que cualquier movimiento que se aleja del centro del sistema de coordenadas Z_1, Z_2 tiene como resultado una disminución en la respuesta. Así, cuando todos los coeficientes λ_i son negativos la superficie es máxima, como en el caso de la superficie de la vida de la herramienta en la figura 13.6.

Si los coeficientes λ_i son positivos, entonces el resultado de cualquier movimiento que se aleja del centro del sistema de coordenadas Z_1, Z_2 es un incremento en la respuesta y la superficie es mínima como se muestra en la figura 13.2a. Si un coeficiente es positivo y los demás negativos, digamos $\lambda_1 > 0$ y $\lambda_2 < 0$, entonces cualquier movimiento que se aleja de $(0, 0)$ a lo largo del eje Z_1 aumenta la respuesta y si se aleja por el eje Z_2 la disminuye. Así, la superficie es minimax o con forma de silla en el punto estacionario, como se muestra en la figura 13.2d. Si una de las $\lambda_i = 0$, la superficie es una cresta estacionaria (figura 13.2c) porque la respuesta no cambia en los ejes Z_i .

Las longitudes de los ejes principales de las elipses formadas por las curvas de nivel son proporcionales a $|\lambda_i|^{-1/2}$. Para la superficie de la vida de la herramienta $|-25.58|^{-1/2} = 0.20$ y $|-6.92|^{-1/2} = 0.38$, y la superficie ajustada se atenua a lo largo del eje Z_2 como se ve en la figura 13.6.

Para explicarlo, supongamos que la velocidad del torno y la profundidad de corte para una vida máxima en las coordenadas $x_1 = -0.156$ y $x_2 = 0.665$ no eran prácticas. El menor cambio en la duración de la herramienta cuando cambian la velocidad del torno y la profundidad de corte se exhibe en la superficie en la dirección del eje Z_2 cuando $Z_1 = 0$. Las coordenadas x_1 y x_2 en el eje Z_2 cuando $Z_1 = 0$ se pueden obtener de la primera de las ecuaciones (13.11). La pérdida mínima en la vida de la herramienta se encuentra en los valores de x_1 y x_2 que satisfacen $0.4603x_1 + 0.8877x_2 - 0.5185 = 0$.

Los coeficientes de x_i en las ecuaciones (13.11) proporcionan información acerca de las relaciones de la velocidad del torno y la profundidad de corte con la vida de la herramienta. Considerando que los coeficientes para la segunda ecuación que relacionan Z_2 con x_1 y x_2 , $Z_2 = 0.8877x_1 - 0.4603x_2 + 0.4446$. El par de coeficientes $(0.8877, -0.4603)$ indican una compensación entre la velocidad del torno y la profundidad de corte en la vida útil, porque en cierto grado, un incremento en la velocidad del torno se puede compensar con una disminución en la profundidad de corte sobre el eje Z_2 .

La ecuación de respuesta estimada en forma original o en forma *canónica* sólo es válida para la zona de los niveles de los factores incluida en el experimento. Cualquier intento para estimar la vida de la herramienta fuera de los límites

acotados por las velocidades de 117 a 683 y las profundidades de 0.04 a 0.11 será engañoso, por lo que es necesario un modelo de respuesta por completo diferente para determinar la duración de la herramienta fuera de la región de este estudio.

13.6 Diseños para mezclas de ingredientes

Algunos diseños de tratamientos tienen dos o más factores que son ingredientes de una mezcla en la que los porcentajes de cada uno deben sumar el 100% de la mezcla. Por lo tanto, los niveles de un factor no son independientes de los niveles de otros.

Muchos productos alimenticios, materiales para construcción y otros productos comerciales se forman con mezclas de dos o más ingredientes de una receta. Algunos ejemplos son:

- telas con mezcla de fibras de algodón y poliéster
- jugos de frutas de naranja, piña y manzana con agua
- concreto formado con agua, arena y cemento
- fórmulas de fertilizantes con nitrógeno, fósforo y potasio

Esta sección incluye una breve introducción para seleccionar diseños y estimar ecuaciones de superficies de respuesta para experimentos con mezclas. Cornell (1990) cubre con detalle el diseño y análisis para esta clase de experimentos.

Los niveles de los factores son las proporciones de los ingredientes

La variación en las proporciones de los ingredientes de las mezclas puede afectar las propiedades del producto final. Más que en las cantidades totales de los factores, las investigaciones de experimentos con mezclas se concentran en la relación de la variable de respuesta medida con las proporciones relativas de cada ingrediente presente en el producto.

Si x_1, x_2, \dots, x_k son las variables que representan las proporciones de los k ingredientes o componentes de la mezcla, los valores de las x_i están restringidos de manera que:

$$0 \leq x_i \leq 1 \quad i = 1, 2, \dots, k \quad (13.12)$$

y la suma de las proporciones de los k ingredientes en la mezcla es la unidad, es decir:

$$\sum_{i=1}^k x_i = x_1 + x_2 + \cdots + x_k = 1 \quad (13.13)$$

Si la proporción de un ingrediente es $x_i = 1$, entonces no hay otros ingredientes presentes en la mezcla y el producto es una mezcla pura o de un componente. Por ejemplo, un experimento con una mezcla de dos componentes para tela de algodón y poliéster, representado por la proporción de x_1 y x_2 , puede tener sólo algodón, en cuyo caso $x_1 = 1$ y $x_2 = 0$, o sólo poliéster, donde $x_1 = 0$ y $x_2 = 1$.

Los valores de x_1 y x_2 admisibles para un diseño de mezcla de dos componentes son valores coordinados sobre la recta $x_1 + x_2 = 1$ (mostrada en la figura 13.7).

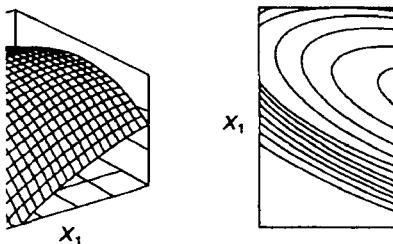


Figura 13.7 Espacio de los factores para una mezcla de dos componentes, $x_1 + x_2 = 1$

Descripción del espacio de los factores con un sistema de coordenadas Símplex

Los valores coordinados para un diseño de mezcla con tres componentes son los valores encontrados en el plano definido por $x_1 + x_2 + x_3 = 1$ en la figura 13.8a. La descripción geométrica del espacio de los factores para k componentes es la de un Símplex en $(k - 1)$ dimensiones. En la figura 13.8b se muestra el sistema Símplex de dos dimensiones para una mezcla de tres componentes como un triángulo equilátero.

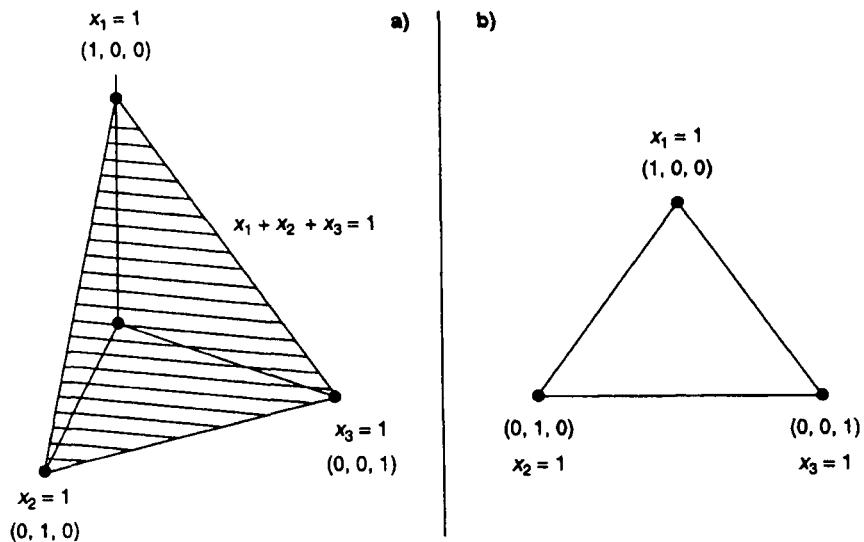


Figura 13.8 Espacio de los factores para a) una mezcla de tres componentes, $x_1 + x_2 + x_3 = 1$ y b) un sistema Símplex de tres componentes

Los vértices del triángulo representan mezclas de componentes sencillas con $x_i = 1$ y las demás igual a 0. Los lados del triángulo representan las coordenadas del diseño para mezclas de dos componentes con una $x_i = 0$. Las coordenadas del diseño en el interior del triángulo representan mezclas de tres componentes con $x_1 > 0, x_2 > 0$ y $x_3 > 0$. Cualquier combinación de proporciones de componentes para un experimento de mezcla debe estar en la frontera o dentro del triángulo de coordenadas de la figura 13.8b.

Los ejes de las variables x_i en el sistema Simplex de tres variables se muestran en la figura 13.9. El eje para la componente i es la recta que va del punto base $x_i = 0$ y $x_j = 1/(k - 1)$ para las componentes, $j \neq i$ al vértice, donde $x_i = 1$ y $x_j = 0$ para $j \neq i$. Por ejemplo, en el diseño con tres componentes de la figura 13.9, el eje x_1 va de la base de coordenadas $(0, \frac{1}{2}, \frac{1}{2})$ al vértice $(1, 0, 0)$.

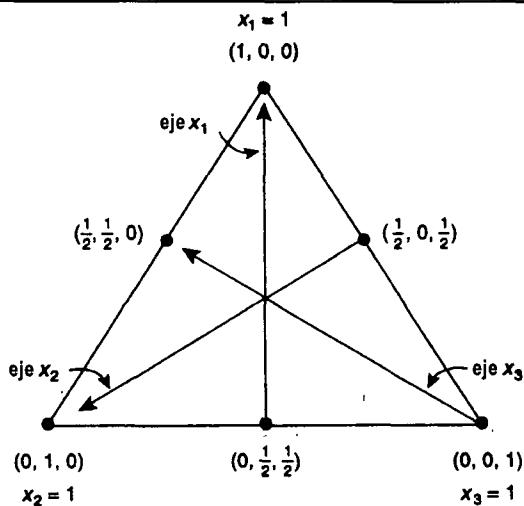


Figura 13.9 Ejes de coordenadas Simplex para una mezcla de tres componentes con coordenadas del diseño para un malla de retícula Simplex $\{3, 2\}$

Diseños de tratamientos para mezclas

Diseños de retícula Simplex

El arreglo formado por una distribución uniforme de coordenadas de diseño en una sistema simplex se conoce como *retícula*; vea la figura 13.9. Los diseño de retícula simplex se componen de una malla de coordenadas de diseño construidas de manera que permiten estimar las ecuaciones polinomiales de la superficie de respuesta.

La designación $\{k, m\}$ se usa en un diseño de retícula Simplex con k componentes para estimar una ecuación polinomial de superficie de respuesta de grado

m . Por ejemplo, el diseño de retícula Símplex $\{3, 2\}$ tiene tres componentes en la mezcla para estimar una ecuación de superficie de respuesta cuadrática.

Las proporciones de cada componente incluida en los diseños de retícula Símplex $\{k, m\}$ son:

$$x_i = 0, \frac{1}{m}, \frac{2}{m}, \dots, 1 \quad (13.14)$$

Los diseños consisten de todas las combinaciones posibles de los niveles de x_i , donde $\sum x_i = 1$ para cualquier combinación de proporciones.

Las combinaciones de proporciones de la mezcla mostradas en cada de la figura 13.9 son los valores de las coordenadas para una retícula Símplex de $\{3, 2\}$.

Las proporciones de cada x_i con $m = 2$ son $x_i = 0, \frac{1}{2}$, y 1. En la figura 13.9 se observa que $\sum x_i = 1$ para cada punto del diseño.

El diseño de retícula Símplex $\{k, 2\}$ para estimar ecuaciones de superficie de respuesta cuadráticas tiene sólo mezclas en las fronteras del sistema de coordenadas con una o más componentes ausentes en la mezcla. La retícula Símplex general $\{k, m\}$ consiste de mezclas de una componentes, mezclas de dos componentes, etcétera, hasta mezclas que consisten hasta de m componentes. Si $m = k$, habrá una sola mezcla en el centro del sistema de coordenadas en el experimento que contiene todas las componentes de las mezclas; por ejemplo, la retícula Símplex $\{3, 3\}$ incluirá la mezcla con proporciones de componentes $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ al igual que las mezclas de una y dos componentes con proporciones $\frac{1}{3}$ y $\frac{2}{3}$ para las ambas.

Diseños Símplex con centroides

El diseño Símplex con centroide es un diseño en el sistema de coordenadas símplex que se compone de mezclas que contienen 1, 2, 3, ... o k componentes en proporciones iguales. En consecuencia, existen k mezclas de una componente, todas las mezclas posibles de dos componentes con proporción $\frac{1}{2}$ para cada una, todas las mezclas posibles de tres componentes con proporción $\frac{1}{3}$ para cada una, etcétera, hasta una mezcla de k componentes con proporción de $\frac{1}{k}$. En la tabla 13.7 se comparan las mezclas para los diseños Símplex con centroides con las mezclas para los diseños de retícula símplex de $\{3, 2\}$ y $\{3, 3\}$.

Diseño Símplex con centroides ampliado

Las combinaciones de la mezcla para los diseños de retícula Símplex y Símplex con centroides se encuentran en las orillas del espacio de factores Símplex con la excepción de un punto centroide que contiene la mezcla de todas las componentes. Es posible obtener mezclas más completas si se aumentan los diseños centroides Símplex con mezclas sobre los ejes del espacio de factores.

Los puntos del diseño se colocan en cada eje, equidistantes al centro o centroide, hacia los vértices. Un diseño de k componentes tendrá k puntos de diseño adicionales con coordenadas:

x_1	x_2	...	x_k
$\frac{(k+1)}{2k}$	$\frac{1}{2k}$...	$\frac{1}{2k}$
$\frac{1}{2k}$	$\frac{(k+1)}{2k}$...	$\frac{1}{2k}$
\vdots	\vdots	\vdots	\vdots
$\frac{1}{2k}$	$\frac{1}{2k}$...	$\frac{(k+1)}{2k}$

Los puntos axiales adicionales proporcionarán una mejor distribución de la información a lo largo de la zona del experimento. Los tres puntos de diseño adicionales necesarios al aumentar el diseño Símplex con centroides para tres componentes son $(\frac{4}{6}, \frac{1}{6}, \frac{1}{6})$, $(\frac{1}{6}, \frac{4}{6}, \frac{1}{6})$ y $(\frac{1}{6}, \frac{1}{6}, \frac{4}{6})$. El diseño completo se describe en la figura 13.10.

Tabla 13.7 Diseños de retícula Símplex y un diseño Símplex con centroides para una mezcla con tres componentes

{3, 2} Retícula			{3, 3} Retícula			Centroide		
x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
1	0	0	1	0	0	1	0	0
0	1	0	0	1	0	0	1	0
0	0	1	0	0	1	0	0	1
$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{2}{3}$	$\frac{1}{3}$	0	$\frac{1}{2}$	$\frac{1}{2}$	0
$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{2}{3}$	0	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{2}$
0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{2}{3}$	$\frac{1}{3}$	0	$\frac{1}{2}$	$\frac{1}{2}$
			$\frac{1}{3}$	$\frac{2}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
			$\frac{1}{3}$	0	$\frac{2}{3}$			
			0	$\frac{1}{3}$	$\frac{2}{3}$			
			$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$			

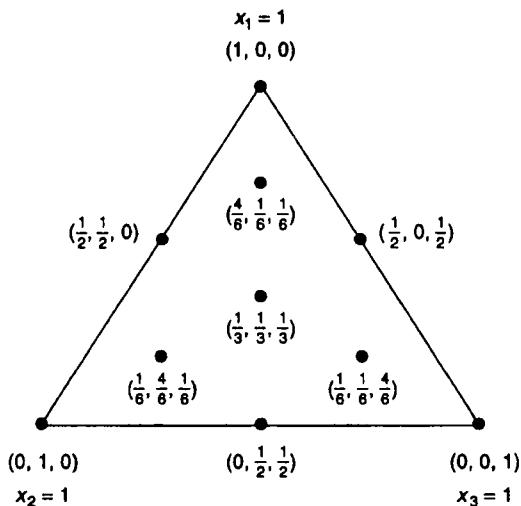


Figura 13.10 Diseño Simplex con centroide ampliado para un experimento de mezcla con tres componentes

Pseudocomponentes para ingredientes con cotas inferiores

Muchas mezclas requieren que todos los componentes estén presentes al menos en una proporción mínima. Es claro que el concreto requiere proporciones mínimas de agua, cemento y arena. Las cotas mínimas, L_i , para las proporciones de las componentes imponen la siguiente restricción:

$$0 \leq L_i \leq x_i \leq 1 \quad (13.15)$$

sobre las proporciones. Suponga que las cotas inferiores para cemento (x_1), agua (x_2) y arena (x_3) son:

$$0.10 \leq x_1 \quad 0.20 \leq x_2 \quad 0.30 \leq x_3$$

y se usará un diseño de retícula Simplex {3, 2} en el experimento. Las cotas inferiores sobre las proporciones de los componentes limitan el diseño a una subregión del espacio de factores original mostrado en la figura 13.9 y la tabla 13.7.

Para simplificar la construcción de las coordenadas del diseño, se construye un conjunto de pseudocomponentes codificando las variables de las componentes originales en un sistema de coordenadas Simplex para las variables de las pseudocomponentes \tilde{x}_i , con restricciones $0 \leq \tilde{x}_i \leq 1$. Si la cota inferior de la componente i es L_i y $L = \sum L_i$, entonces la pseudocomponente se calcula como:

$$\tilde{x}_i = \frac{x_i - L_i}{1 - L}$$

Es posible construir un diseño con las componentes originales basado en las coordenadas de las pseudocomponentes que se establece en un Simplex normal con $\sum \tilde{x}_i = 1$. Las proporciones de los componentes originales necesarios para las mezclas en el experimento se pueden derivar mediante la transformación inversa:

$$x_i = L_i + \tilde{x}_i(1 - L)$$

Para el ejemplo del concreto, las cotas inferiores fueron $L_1 = 0.10$, $L_2 = 0.20$ y $L_3 = 0.30$, con suma $L = 0.10 + 0.20 + 0.30 = 0.60$ y las pseudocomponentes son:

$$\tilde{x}_1 = \frac{x_1 - 0.10}{0.40} \quad \tilde{x}_2 = \frac{x_2 - 0.20}{0.40} \quad \tilde{x}_3 = \frac{x_3 - 0.30}{0.40}$$

y las transformaciones para regresar a las proporciones de los componentes originales a partir de las pseudocomponentes son:

$$x_1 = 0.10 + 0.40\tilde{x}_1 \quad x_2 = 0.20 + 0.40\tilde{x}_2 \quad x_3 = 0.30 + 0.40\tilde{x}_3$$

El diseño completo para el experimento de la mezcla de concreto se muestra en la tabla 13.8, con los valores de las coordenadas de las pseudocomponentes y de los componentes originales en la subregión del Simplex original.

Tabla 13.8 Coordenadas de pseudocomponentes y componentes originales en un diseño de retícula Simplex {3, 2} para el experimento de la mezcla de concreto

Pseudocomponentes			Componentes originales		
\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	Cemento	Aqua	Arena
1	0	0	0.50	0.20	0.30
0	1	0	0.10	0.60	0.30
0	0	1	0.10	0.20	0.70
$\frac{1}{2}$	$\frac{1}{2}$	0	0.30	0.40	0.30
$\frac{1}{2}$	0	$\frac{1}{2}$	0.30	0.20	0.50
0	$\frac{1}{2}$	$\frac{1}{2}$	0.10	0.40	0.50

13.7 Análisis de experimentos con mezclas

Polinomios canónicos para aproximar superficies

La forma general de la función polinomial usada para aproximar la superficie de respuesta lineal es:

$$\mu_y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \quad (13.16)$$

La restricción sobre las componentes de la mezcla, $x_1 + x_2 + \dots + x_k = 1$, crea una dependencia entre las x_i en la función lineal. Multiplicar β_0 por $(x_1 + x_2 + \dots + x_k)$ proporciona una nueva expresión del modelo como:

$$\mu_y = \beta_0 \left(\sum_{i=1}^k x_i \right) + \sum_{i=1}^k \beta_i x_i = \beta_1^* x_1 + \beta_2^* x_2 + \cdots + \beta_k^* x_k \quad (13.17)$$

donde $\beta_i^* = \beta_0 + \beta_i$, $i = 1, 2, \dots, k$. La nueva expresión de la ecuación con parámetros β_i^* se conoce como **polinomio canónico**. El polinomio canónico y el polinomio original son equivalentes porque uno se deriva del otro y el grado del polinomio y el número de componentes no cambian con la nueva expresión.

La función polinomial cuadrática usada para aproximar la superficie de respuesta es:

$$\mu_y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_i x_i^2 + \sum_{i < j} b_{ij} x_i x_j \quad (13.18)$$

El polinomio canónico cuadrático producido al incluir la restricción $\sum x_i = 1$ es:

$$\mu_y = \sum_{i=1}^k \beta_i^* x_i + \sum_{i < j} b_{ij}^* x_i x_j \quad (13.19)$$

donde $\beta_i^* = \beta_0 + \beta_i + \beta_{ii}$, y $\beta_{ij}^* = \beta_{ij} - \beta_{ii} - \beta_{jj}$. Los nuevos parámetros del polinomio canónico cuadrático para los tres componentes de la mezcla expresados en términos de los parámetros del polinomio original son:

$$\beta_1^* = \beta_0 + \beta_1 + \beta_{11} \quad \beta_2^* = \beta_0 + \beta_2 + \beta_{22} \quad \beta_3^* = \beta_0 + \beta_3 + \beta_{33}$$

$$\beta_{12}^* = \beta_{12} - \beta_{11} - \beta_{22} \quad \beta_{13}^* = \beta_{13} - \beta_{11} - \beta_{33} \quad \beta_{23}^* = \beta_{23} - \beta_{22} - \beta_{33}$$

La interpretación de los polinomios canónicos se ilustra con un experimento de mezcla de componentes de gasolina.

Ejemplo 13.4 Experimento de mezcla de gasolina

El octanaje de la gasolina depende de la proporción de varios componentes de petróleo mezclados para producir combustible. El objetivo de la mayor parte de los estudios de los estudios de mezclas de gasolina es desarrollar un modelo de mezclas lineal para determinar la mezcla de componentes de gasolina más rentable. Los coeficientes en el modelo lineal de mezcla, conocidos como *valores de fusión*, describen el comportamiento de mezclado de un componente de combustible dado. Sin embargo, la composición de la mezcla depende, entre otros de factores como la calidad de los componentes. Así, la linealidad de los componentes de la mezcla se pierde y deben considerarse modelos de interacción entre componentes o modelos cuadráticos más complejos.

El análisis de un experimento de mezcla se explica con una mezcla para obtener gasolina. Para evaluar la necesidad de un modelo de interacción, se diseñó el experimento para estimar un polinomio cuadrático canónico.

Se estableció un experimento de mezclas para evaluar el efecto de los tres componentes sobre la tasa de octano de la gasolina. Los componentes, alquilante (A), corrida ligera (B) y reformado (C), se usaron en un diseño Simplex con centroides con siete mezclas y se determinaron las tasas de octano para dos réplicas de cada una. Las tasas de octano para cada mezcla se muestran en la tabla 13.9.

Tabla 13.9 Tasas de octano en un experimento de mezclas de gasolina

Componentes*				
x_1	x_2	x_3	y_{ij}	\bar{y}_i
1	0	0	106.6, 105.0	105.80
0	1	0	83.3, 81.4	82.35
0	0	1	99.4, 91.4	95.40
$\frac{1}{2}$	$\frac{1}{2}$	0	94.1, 91.4	92.75
$\frac{1}{2}$	0	$\frac{1}{2}$	101.9, 98.0	99.95
0	$\frac{1}{2}$	$\frac{1}{2}$	92.3, 86.5	89.40
$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	96.3, 91.7	94.00

* x_1 = alquilante, x_2 = corrida ligera, x_3 = reformado

Fuente: R. D. Snee (1981), Developing blending models for gasoline and other mixtures. *Technometrics* 23, 119-130.

Estimación del modelo de superficie de respuesta polinomial canónico cuadrático

Por conveniencia se eliminará el asterisco de los coeficientes en las ecuaciones polinomiales canónicas. El modelo completo polinomial canónico para el experimento de mezcla de gasolina es:

$$y_{ij} = \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \beta_{12} x_{1j} x_{2j} + \beta_{13} x_{1j} x_{3j} + \beta_{23} x_{2j} x_{3j} + e_{ij} \quad i = 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad (13.20)$$

donde se supone que los errores experimentales e_{ij} son independientes, con distribución normal, con media 0 y varianza σ^2 . Además, $t = 7$ tratamientos de mezcla y $r = 2$ réplicas de cada una para un total de $N = rt = 14$ observaciones.

En principio, la hipótesis significativa es si la respuesta depende de los componentes de la mezcla, de acuerdo con el modelo cuadrático. Cuando la hipótesis nula es verdadera, el modelo reducido $y_{ij} = \beta_0 + e_{ij}$, donde $\beta_1 = \beta_2 = \beta_3 = \beta_0$ y $\beta_{12} = \beta_{13} = \beta_{23} = 0$ describe en forma adecuada la respuesta media.

Un análisis de varianza de los datos en la tabla 13.9 para los tratamientos de la mezcla en un diseño totalmente aleatorizado proporcionará una estimación del error experimental puro. La suma de cuadrados entre las mezclas es:

$$SCT = r \sum_{i=1}^t (\bar{y}_i - \bar{y})^2 = 669.32$$

con $(t - 1) = 6$ grados de libertad. La suma de cuadrados del error experimental es:

$$SCE = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2 = 73.74$$

con $(N - t) = 7$ grados de libertad. Así, la media cuadrática para el error experimental es $CME = 10.53$ con 7 grados de libertad.

Las estimaciones de mínimos cuadrados de los parámetros para los polinomios canónicos requieren un ajuste a un modelo de regresión sin el término usual de la ordenada β_0 . Muchos programas de regresión para computadora tienen la capacidad de ajustar el modelo de regresión sin el término de la ordenada, y proporcionarán las estimaciones correctas de mínimos cuadrados de los parámetros para los polinomios canónicos.

La ecuación cuadrática completa estimada es:

$$\hat{y} = 105.8x_1 + 82.3x_2 + 95.4x_3 - 5.1x_1x_2 - 2.4x_1x_3 + 2.3x_2x_3 \quad (13.21)$$

con una suma de cuadrados del error experimental $SCE_f = 73.76$ con $14 - 6 = 8$ grados de libertad.

Pruebas de hipótesis sobre el modelo

Una prueba para el modelo completo

Si la respuesta no depende de los componentes de la mezcla, el modelo reducido es $y_{ij} = \beta_0 + e_{ij}$ y la superficie tiene una altura constante. La suma de cuadrados del error experimental para este modelo reducido es:

$$SCE_r = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y})^2 = 743.05$$

con $N - 1 = 13$ grados de libertad.

La reducción de la suma de cuadrados para el modelo completo de superficie de respuesta cuadrático completo es:

$$SCR = SCE_r - SCE_f = 743.05 - 73.76 = 669.29$$

con $13 - 8 = 5$ grados de libertad. La suma de cuadrados para el modelo cuadrático proporciona los 5 o 6 grados de libertad para los tratamientos con 1 otro grado de libertad para la falta de ajuste al modelo cuadrático. El análisis de varianza se resume en la tabla 13.10.

La hipótesis nula para la ecuación de respuesta cuadrática es $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_0$ y $\beta_{12} = \beta_{13} = \beta_{23} = 0$. El estadístico de prueba $F_0 = CMR/CME = 133.86/10.53 = 12.71$ excede el valor crítico $F_{0.05,5,7} = 3.97$, y la hipótesis nula se rechaza.

Tabla 13.10 Análisis de varianza para el experimento de mezcla de gasolina

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Tratamientos	6	669.32	111.55
Regresión	5	669.29	133.86
Falta de ajuste	1	0.03	0.03
Error	7	73.74	10.53

Una prueba para los términos cuadráticos

El investigador deseará saber si es necesario el modelo cuadrático completo para aproximar la superficie de respuesta o si la superficie lineal es adecuada para explicar la relación entre la tasa de octano y la mezcla de componentes. Una prueba de la hipótesis nula $H_0: \beta_{12} = \beta_{13} = \beta_{23} = 0$ determinará si las componentes cuadráticas del modelo son necesarias. El principio del modelo completo y el modelo reducido se puede usar para determinar la significancia de las componentes cuadráticas, pero se requieren las sumas de cuadrados para el error experimental del modelo cuadrático completo y el modelo lineal reducido $y_{ij} = \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + e_{ij}$. La ecuación lineal del modelo reducido estimada es:

$$\hat{y} = 105.1x_1 + 82.1x_2 + 95.5x_3 \quad (13.22)$$

con una suma de cuadrados del error experimental $SCE_{lin} = 77.37$ con $14 - 3 = 11$ grados de libertad. La reducción de suma de cuadrados debida a los términos cuadráticos después del ajuste de los términos lineales es:

$$SCR(\text{cuadrática}) = SCE_{lin} - SCE_f = 77.37 - 73.76 = 3.61$$

con $11 - 8 = 3$ grados de libertad con una $CMR(\text{cuadrática}) = 3.61/3 = 1.20$. El estadístico de prueba $F_0 = CMR(\text{cuadrática})/CME = 1.20/10.53 = 0.11$ con región crítica $F_0 > F_{05,3,7} = 4.35$ no es significativo, y los términos cuadráticos no mejoran la aproximación del modelo a la superficie de respuesta. Así, un modelo lineal de mezclas es adecuado para este conjunto de componentes de una mezcla.

Una prueba para los términos lineales

Como las componentes cuadráticas del modelo casi no contribuyen a la suma de cuadrados de la regresión, es bastante obvio que los términos lineales proporcionan la mayor parte de la suma de cuadrados para el modelo de regresión. Una prueba formal de la hipótesis nula $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_0 = 0$ se deriva de la reducción de la suma de cuadrados para el modelo lineal:

$$SCR(\text{lineal}) = SCE_r - SCE_{lin} = 743.05 - 77.37 = 665.68$$

con $13 - 11 = 2$ grados de libertad con una $CMR(\text{lineal}) = 665.68/2 = 332.84$. El estadístico de prueba es $F_0 = CMR(\text{lineal})/CME = 332.84/10.53 = 31.61$ con una región crítica $F_0 > F_{05,2,7} = 4.74$, y la hipótesis nula se rechaza.

Interpretaciones para la ecuación de respuesta estimada

El polinomio canónico lineal estimado, $\hat{y} = 105.1x_1 + 82.1x_2 + 95.5x_3$ en la ecuación (13.22) proporciona un ajuste significativo y adecuado para la superficie de respuesta de la mezcla. El error estándar estimado para cada β_i , determinado a partir del programa de regresión, es $s_{\hat{\beta}_i} = 1.89$ y las estimaciones de los coeficientes individuales son significativas según la prueba t de Student, $t_0 = \hat{\beta}_i/s_{\hat{\beta}_i}$, con 7 grados de libertad.

El coeficiente $\hat{\beta}_i$ es la respuesta estimada en el vértice del diseño Símplex que representa la mezcla con 100% de ese componente o la mezcla de un componente; de otra manera, representa la respuesta estimada en el valor máximo para ese componente. Las variables del ejemplo 13.4 fueron las proporciones de x_1 = alquilate, x_2 = corrida ligera y x_3 = reformada. Por ejemplo, con 100% de alquilate y 0% de corrida ligera y reformada, la tasa de octano estimada es $\hat{y} = 105.1$. De la misma manera, con 100% de corrida ligera el octano estimado es $\hat{y} = 82.1$, y con 100% de reformada el octano estimado es $\hat{y} = 95.5$. La superficie lineal estimada se describe en la figura 13.11.

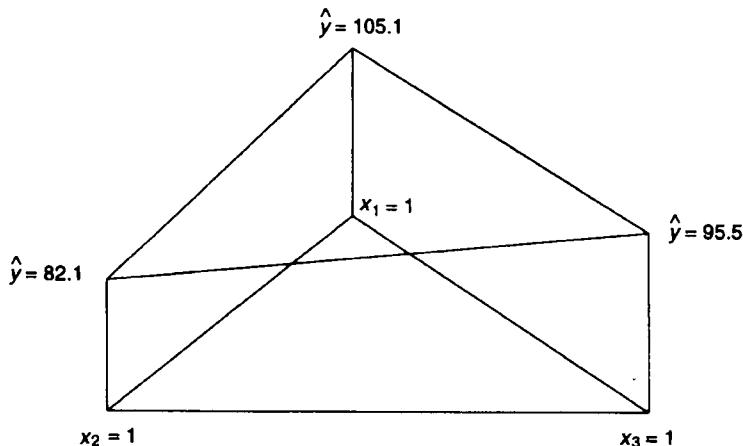


Figura 13.11 Superficie lineal estimada para el experimento de mezcla de gasolina

Las β_i describen el alejamiento de la superficie de respuesta. Una ilustración de la respuesta cuadrática para un sistema de dos componentes se describe en la figura 13.12. Si las dos componentes son aditivas, con una mezcla lineal de ellas la respuesta media es $\mu_i = \beta_1x_1 + \beta_2x_2$ (mostrada como una línea recta de la mezcla en la figura 13.12). Una mezcla cuadrática no lineal de las dos componentes

con $\beta_{12} > 0$ se muestra con la curva de respuesta $\mu_q = \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2$, donde el coeficiente β_1 representa la altura de la curva cuando $x_1 = 1$ y $x_2 = 0$, y β_2 representa la altura de la curva cuando $x_1 = 0$ y $x_2 = 1$ y el término $\beta_{12}x_1x_2$ contribuye a la respuesta siempre que $x_1 > 0$ y $x_2 > 0$. El mayor alejamiento de la línea de mezcla ocurre en $x_1 = x_2 = \frac{1}{2}$ cuando $\beta_{12}x_1x_2 = \frac{1}{4}\beta_{12}$. Se dice que la mezcla de dos componentes en la figura 13.12 es sinérgica porque la respuesta para la mezcla 1:1 en $x_1 = x_2 = \frac{1}{2}$ excede el promedio simple de las mezclas puras descritas por la recta de mezcla lineal. Si el coeficiente β_{12} fuera negativo, la mezcla no lineal estaría por debajo de esa recta y las componentes serían antagónicas entre sí.

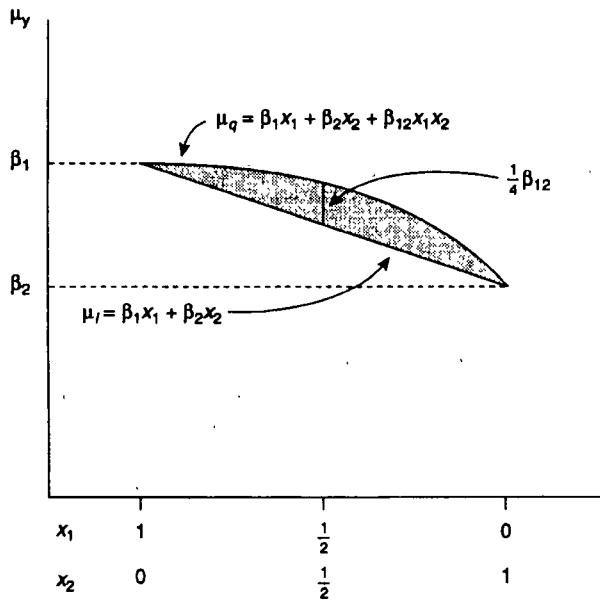


Figura 13.12 Comparación de la mezcla lineal y no lineal para un sistema de mezcla de dos componentes

EJERCICIOS PARA EL CAPÍTULO 13

- Se realizó un experimento factorial 2^2 para determinar si el volumen de dos reactivos afectaba la capacidad de un método de ensayo para medir los niveles de una droga específica en suero. Se probaron dos réplicas de los tratamientos en un diseño totalmente aleatorizado con dos réplicas adicionales en el centro del diseño y se hizo un muestreo de suero para cada prueba de una reserva tratada con una sola dosis de droga. A continuación se muestran los datos para los volúmenes del reactivo en μl .

<i>Reactivo</i>		
<i>A</i>	<i>B</i>	% de droga recuperada
10	20	32, 35
40	20	44, 47
10	50	51, 53
40	50	68, 72
25	35	48, 53

- a. Estime la varianza del error experimental.
- b. Estime la ecuación de respuesta lineal y los errores estándar. ¿Son significativos los efectos lineales de los reagentes?
- c. Estime β_{12} para la interacción AB y el error estándar de la estimación. ¿Existe una interacción significativa?
- d. Estime el alejamiento de la superficie lineal, $\beta_{11} + \beta_{22}$, y el error estándar de la estimación. ¿Es significativo el alejamiento?
- e. Determine los primeros cinco pasos en la trayectoria de mayor pendiente a partir del centro del diseño en etapas de una unidad en x_1 para el reactivo A y muestre los niveles de ambos factores en cada una.

2. Considere el experimento de vinilación de metil glucósidos del ejemplo 13.1. Suponga que el químico ha observado una respuesta máxima en $T = 175^\circ\text{C}$ y $P = 474.5$ psi en la trayectoria de mayor pendiente de la tabla 13.2. Utilice esos niveles como niveles promedio del factor en la ecuación (13.4) y proponga un diseño central compuesto rotatorio para estimar la ecuación de la superficie de respuesta cuadrática y muestre los niveles reales de T y P para cada combinación de tratamiento en él.

3. Muestre las coordenadas del diseño codificadas para un diseño central compuesto rotatorio con precisión uniforme de cuatro factores.

4. Describa la superficie de respuesta para las siguientes formas canónicas dadas para superficies de respuesta cuadráticas:
 - a. $\hat{y} = 100 - Z_1^2 - 2Z_2^2$
 - b. $\hat{y} = 50 + 2Z_1^2$
 - c. $\hat{y} = 75 + Z_1^2 - 2Z_2^2$

5. Un científico de animales estudió la relación entre el metabolismo de metionina, un aminoácido azufroso y de caroteno, vitamina A, conforme afectan el crecimiento de los pollos. Se pensó que los niveles óptimos de metionina y caroteno eran 0.9% de metionina en la dieta y 50 microgramos de caroteno al día. Se usó un diseño central compuesto rotatorio para el experimento. Se asignaron al azar ocho pollos a cada tratamiento dietético y se registraron sus aumentos de peso después de 38 días. El aumento promedio de peso para los tratamientos es el siguiente:

<i>Factores originales</i>		<i>Factores codificados</i>		
<i>Metionina</i>	<i>Caroteno</i>	x_1	x_2	<i>Aumento de peso</i>
1.183	85.36	+ 1	+ 1	445
1.183	14.64	+ 1	- 1	331
0.617	85.36	- 1	+ 1	443
0.617	14.64	- 1	- 1	336
1.183	50.00	$\sqrt{2}$	0	414
0.500	50.00	- $\sqrt{2}$	0	389
0.900	100.00	0	$\sqrt{2}$	435
0.900	0.00	0	- $\sqrt{2}$	225
0.900	50.00	0	0	442
0.900	50.00	0	0	412
0.900	50.00	0	0	418
0.900	50.00	0	0	440
0.900	50.00	0	0	441

- Estime la ecuación de la superficie de respuesta cuadrática para el aumento de peso y resuma las particiones de sumas de cuadrados en una tabla de análisis de varianza.
 - Pruebe la significancia del modelo cuadrático completo, las desviaciones cuadráticas del modelo lineal, la significancia de las componentes lineales del modelo y la falta de ajuste al modelo cuadrático. ¿Cuáles son sus conclusiones?
 - La superficie de respuesta tiene un máximo dentro de las coordenadas del diseño. Determine los niveles de metionina y caroteno que producen la respuesta máxima y estime esta última.
 - Calcule la ecuación canónica (vea el apéndice 13A.3) y describa la superficie de respuesta. Con base en la ecuación canónica, ¿Cuál es la relación entre metionina y caroteno?, ¿puede una de las dos usarse para compensar la otra en la dieta de los animales?
6. El experimento de vinilación de metil glucósidos usado en el ejemplo 13.1 (Marvel *et al.*, 1969) incluyó cuatro factores en un diseño rotatorio compuesto central colocado en un diseño de bloques incompleto. El porcentaje de conversión de metil glucósidos en un producto de vinilación era la variable de respuesta de interés. Los niveles reales y codificados de los cuatro factores usados en el experimento fueron:

	<i>Nivel codificado</i>	-2	-1	0	1	2
x_1	Tiempo, horas	1	3	5	7	9
x_2	Temperatura, °C	115	130	145	160	175
x_3	Presión, psi	250	325	400	475	550
x_4	Razón del solvente (agua/dioxano)	95	80	65	50	35

El porcentaje de conversión de metil glucósidos en un producto de vinilación para cada combinación de tratamiento en el experimento se muestra en la tabla que sigue:

<i>Bloque 1</i>					<i>Bloque 2</i>					<i>Bloque 3</i>				
<i>y</i>	<i>x</i> ₁	<i>x</i> ₂	<i>x</i> ₃	<i>x</i> ₄	<i>y</i>	<i>x</i> ₁	<i>x</i> ₂	<i>x</i> ₃	<i>x</i> ₄	<i>y</i>	<i>x</i> ₁	<i>x</i> ₂	<i>x</i> ₃	<i>x</i> ₄
10	-1	-1	-1	-1	56	-1	1	1	1	13	-2	0	0	0
21	1	1	-1	-1	7	-1	-1	-1	1	43	2	0	0	0
5	-1	-1	1	1	35	1	1	1	-1	2	0	-2	0	0
46	1	1	1	1	27	1	-1	1	1	52	0	2	0	0
6	1	-1	-1	1	19	1	-1	-1	-1	18	0	0	-2	0
16	1	-1	1	-1	52	1	1	-1	1	52	0	0	2	0
22	-1	1	1	-1	17	-1	-1	1	-1	15	0	0	0	-2
18	-1	1	-1	1	24	-1	1	-1	-1	58	0	0	0	2
21	0	0	0	0	33	0	0	0	0	34	0	0	0	0
31	0	0	0	0	30	0	0	0	0	39	0	0	0	0

- a. Estime la ecuación de la superficie de respuesta cuadrática para el porcentaje de conversión y resuma las particiones de sumas de cuadrados en una tabla de análisis de varianza.
- b. Pruebe la significancia del modelo cuadrático completo, las desviaciones cuadráticas del modelo lineal, la significancia de las componentes lineales del modelo y la falta de ajuste al modelo cuadrático. ¿Cuáles son sus conclusiones?
- c. Determine los niveles de los factores que producen la respuesta máxima con el modelo cuadrático. ¿Está el óptimo dentro de los niveles de los factores del diseño actual?
- d. Calcule la ecuación canónica a partir de las estimaciones del modelo cuadrático (vea el apéndice 13A.3) y describa la superficie de respuesta. Con base en la ecuación canónica, ¿Qué tipo de superficie se estimó?
7. Construya un diseño de bloques incompleto para un diseño central compuesto con tres factores de tratamiento tal que se satisfagan las condiciones de la ecuación (13.5). ¿Qué valor de α se requiere para que la relación de la ecuación (13.7) se cumpla?
8. Se planea un experimento para evaluar la calidad del sabor de un jugo de frutas que contiene jugo de naranja, piña, limón y papaya. Las proporciones mínimas permitidas de los cuatro jugos en la mezcla son naranja > .15, piña > .10, limón > .10 y papaya > .20. Diseñe un experimento de mezcla con un diseño de retícula Símplex. Enumere las pseudocomponentes y las componentes originales del diseño.
9. Una mezcla de pavimento se forma dispersando azufre líquido en asfalto líquido para producir un aglutinado de azufre-asfalto, que luego se mezcla con arena para producir la mezcla de pavimento. Debe realizarse un experimento de mezcla con mínimos de 5% de azufre y 10% de asfalto.
- Diseñe el experimento de mezcla con un diseño Símplex con centroides. Muestre las pseudocomponentes y las componentes originales del diseño.
 - Construya el diseño como un Símplex con centroides ampliado. Muestre las pseudocomponentes y las componentes originales.
10. Se realizó un experimento con el aglutinado de azufre-asfalto descrito en el ejercicio 13.9. El experimento se realizó con mínimos de 10% de azufre, 20% de asfalto y 50% de arena en las mezclas. Se prepararon dos réplicas de cada mezcla y un espécimen de cada réplica se sometió a una prueba de resistencia,

utilizando un diseño de retícula Símplex. Los siguientes son los datos de la resistencia con las proporciones de mezcla para las pseudocomponentes, donde x_1 es la pseudocomponente para el sulfuro, x_2 para el azufre y x_3 para la arena.

<i>Pseudocomponentes</i>			<i>Resistencia</i>
x_1	x_2	x_3	y_y
1	0	0	12.0, 13.7
0	1	0	2.4, 3.6
0	0	1	2.6, 4.3
0.5	0.5	0	18.9, 16.8
0.5	0	0.5	19.4, 17.1
0	0.5	0.5	4.6, 7.3

- Dados los valores mínimos para azufre, asfalto y arena, determine las proporciones reales de los tres componentes para cada mezcla de tratamiento en el experimento.
- Estime la varianza del error experimental.
- Estime el polinomio de la superficie de respuesta lineal y cuadrático. Determine la significancia de los términos lineales y la significancia de la adición cuadrática al modelo.
- Interprete los coeficientes del modelo.

Apéndice 13A.1 Estimación de mínimos cuadrados de los modelos de regresión

La estimación de mínimos cuadrados de los parámetros para el modelo de regresión sigue los procedimientos ilustrados en los capítulos anteriores para varios diseños de experimentos. La diferencia notoria para el modelo de regresión es la inclusión de las variables x_i de valores continuos que no estaban en los modelos de diseño de experimentos anteriores.

La estimación se explica con un modelo con dos variables x_i . El modelo de regresión lineal múltiple es:

$$y_j = \beta_0 x_{0j} + \beta_1 x_{1j} + \beta_2 x_{2j} + e_j \quad j = 1, 2, \dots, n$$

La ordenada o término constante para el modelo escrito en la forma general como $\beta_0 x_{0j}$ suele identificarse en el modelo sólo como β_0 , ya que la variable x_{0j} toma un valor constante $x_{0j} = 1$ para todas las observaciones.

Los estimadores de mínimos cuadrados para los coeficientes de regresión se encuentran diferenciando:

$$\sum_{j=1}^n e_j^2 = \sum_{j=1}^n (y_j - \beta_0 x_{0j} - \beta_1 x_{1j} - \beta_2 x_{2j})^2$$

con respecto a cada β_i e igualando el resultado a 0. Las derivadas parciales igualadas a 0 son:

$$\frac{\partial}{\partial \beta_0} = -2 \sum_{j=1}^n x_{0j}(y_j - \beta_0 x_{0j} - \beta_1 x_{1j} - \beta_2 x_{2j}) = 0$$

$$\frac{\partial}{\partial \beta_1} = -2 \sum_{j=1}^n x_{1j}(y_j - \beta_0 x_{0j} - \beta_1 x_{1j} - \beta_2 x_{2j}) = 0$$

$$\frac{\partial}{\partial \beta_2} = -2 \sum_{j=1}^n x_{2j}(y_j - \beta_0 x_{0j} - \beta_1 x_{1j} - \beta_2 x_{2j}) = 0$$

Las ecuaciones normales que se obtienen son:

$$\hat{\beta}_0 \sum x_{0j}^2 + \hat{\beta}_1 \sum x_{0j}x_{1j} + \hat{\beta}_2 \sum x_{0j}x_{2j} = \sum x_{0j}y_j$$

$$\hat{\beta}_0 \sum x_{0j}x_{1j} + \hat{\beta}_1 \sum x_{1j}^2 + \hat{\beta}_2 \sum x_{1j}x_{2j} = \sum x_{1j}y_j$$

$$\hat{\beta}_0 \sum x_{0j}x_{2j} + \hat{\beta}_1 \sum x_{1j}x_{2j} + \hat{\beta}_2 \sum x_{2j}^2 = \sum x_{2j}y_j$$

Las soluciones simultáneas de las ecuaciones para las $\hat{\beta}_i$ tienen como resultado los estimadores de mínimos cuadrados para las β_i . Las estimaciones para un problema dado se pueden obtener con cualquier paquete estadístico de computadora que incluya un programa para regresión lineal múltiple.

El lector interesado puede encontrar detalles de la metodología de análisis de regresión en Rawlings (1988). Aquí se presenta una breve descripción de la formulación del modelo, la construcción de las ecuaciones normales y sus soluciones, con la suposición de que se tienen conocimientos de la notación matricial. El modelo de regresión lineal múltiple se puede escribir en forma de matriz como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

donde:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{kn} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Si \mathbf{X}' denota la transpuesta de la matriz \mathbf{X} , las ecuaciones normales en forma matricial son:

$$\mathbf{X}' \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}' \mathbf{y}$$

donde:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \Sigma x_1 & \Sigma x_2 & \cdots & \Sigma x_k \\ \Sigma x_1 & \Sigma x_1^2 & \Sigma x_1 x_2 & \cdots & \Sigma x_1 x_k \\ \Sigma x_2 & \Sigma x_1 x_2 & \Sigma x_2^2 & \cdots & \Sigma x_2 x_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma x_k & \Sigma x_1 x_k & \Sigma x_2 x_k & \cdots & \Sigma x_k^2 \end{bmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} \Sigma y \\ \Sigma x_1 y \\ \Sigma x_2 y \\ \vdots \\ \Sigma x_k y \end{bmatrix}$$

Las soluciones de las ecuaciones normales para $\hat{\beta}$ se encuentran al multiplicar ambos lados de las ecuaciones por la matriz inversa $\mathbf{X}'\mathbf{X}$ denotada por $(\mathbf{X}'\mathbf{X})^{-1}$. La solución se escribe como:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Los cálculos se ilustran con los datos de dos variables independientes, x_1 y x_2 , para estimar los coeficientes del modelo de primer orden $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$. Las matrices de datos son:

$$\mathbf{y}' = [41, 52, 54, 73, 66, 67, 79]$$

$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 1 & 3 & 2 & 2 & 3 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 7 & 28 & 14 \\ 28 & 140 & 63 \\ 14 & 63 & 32 \end{bmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 432 \\ 1,884 \\ 921 \end{bmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 1.16 & -0.03 & -0.44 \\ -0.03 & 0.06 & -0.11 \\ -0.44 & -0.11 & 0.44 \end{bmatrix}$$

La solución a las ecuaciones normales es:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1.16 & -0.03 & -0.44 \\ -0.03 & 0.06 & -0.11 \\ -0.44 & -0.11 & 0.44 \end{bmatrix} \begin{bmatrix} 432 \\ 1,884 \\ 921 \end{bmatrix} = \begin{bmatrix} 31.43 \\ 3.57 \\ 8.00 \end{bmatrix}$$

La ecuación estimada es $\hat{y}_j = \hat{\beta}_0 x_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k$ y la suma de cuadrados del error experimental para el modelo completo de regresión se calcula con:

$$\begin{aligned} SCE_f &= \sum_{j=1}^n (y_j - \hat{y}_j)^2 = (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) = \mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y} \\ &= \sum_{j=1}^n y_j^2 - \hat{\beta}_0 \sum y_j - \hat{\beta}_1 \sum x_1 y_j - \dots - \hat{\beta}_k \sum x_k y_j \end{aligned}$$

con $n - k - 1$ grados de libertad.

El modelo reducido sin las variables x_i , $i = 1, 2, \dots, k$, es $y_j = \beta_0 x_{0j} + e_j = \mu + e_j$; la ecuación de regresión estimada con el modelo reducido es $\hat{y}_j = \hat{\beta}_0 x_{0j}$. La suma de cuadrados del error experimental para el modelo reducido es:

$$SCE_r = \sum_{j=1}^n (y_j - \hat{y}_j)^2 = \sum_{j=1}^n y_j^2 - \hat{\beta}_0 \sum_{j=1}^n y_j$$

con $n - 1$ grados de libertad.

Para el modelo reducido, $\hat{\beta}_0 = \bar{y}$ y $SCE_r = \sum_{j=1}^n (y_j - \bar{y})^2$. La suma de cuadrados para la regresión debido a la inclusión de las variables independientes x_i , $i = 1, 2, \dots, k$, en el modelo es:

$$SCR = SCE_f - SCE_r$$

con k grados de libertad.

Dado $\sum y_j^2 = 27,716$, la suma de cuadrados del error experimental para el modelo completo del ejemplo es:

$$SCE_f = 27,716 - 31.43(432) - 3.57(1,884) - 8.0(921) = 44.4$$

con $n - k - 1 = 7 - 2 - 1 = 4$ grados de libertad. La estimación de β_0 para el modelo reducido es $\bar{y} = 61.7$, y la suma de cuadrados del error experimental es $SCE_r = 1055.4$. La suma de cuadrados para la regresión es:

$$SCR = 1055.4 - 44.4 = 1011.0$$

con $k = 2$ grados de libertad.

Apéndice 13A.2 Localización de coordenadas para el punto estacionario

El modelo cuadrático estimado expresado en forma matricial es:

$$\hat{y} = \hat{\beta}_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} \quad (13A.1)$$

donde:

$$\mathbf{x}' = [x_1 \ x_2 \ \dots \ x_k] \quad \mathbf{b}' = [\hat{\beta}_1 \ \hat{\beta}_2 \ \dots \ \hat{\beta}_k]$$

$$\mathbf{B} = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \hat{\beta}_{13}/2 & \cdots & \hat{\beta}_{1k}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} & \hat{\beta}_{23}/2 & \cdots & \hat{\beta}_{2k}/2 \\ \hat{\beta}_{13}/2 & \hat{\beta}_{23}/2 & \hat{\beta}_{33} & \cdots & \hat{\beta}_{3k}/2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\beta}_{1k}/2 & \hat{\beta}_{2k}/2 & \hat{\beta}_{3k}/2 & \cdots & \hat{\beta}_{kk} \end{bmatrix}$$

El punto estacionario se encuentra igualando a 0 la derivada de \hat{y} respecto a x ,

$$\frac{\partial \hat{y}}{\partial x} = \mathbf{b} + 2\mathbf{Bx} = 0 \quad (13A.2)$$

El vector de las coordenadas del diseño para el punto estacionario es la solución de la ecuación (13A.2), es decir:

$$\mathbf{x}_s = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b} \quad (13A.3)$$

Al sustituir la solución en la ecuación (13A.1), la respuesta estimada en el punto estacionario es:

$$\hat{y}_s = \hat{\beta}_0 + \frac{1}{2}\mathbf{x}'_s\mathbf{b} \quad (13A.4)$$

Apéndice 13A.3 Forma canónica de la ecuación cuadrática

La forma canónica de la ecuación de respuesta cuadrática es:

$$\hat{y} = \hat{y}_s + \lambda_1 Z_1^2 + \lambda_2 Z_2^2 + \cdots + \lambda_k Z_k^2 \quad (13A.5)$$

donde los λ_i son los valores específicos de la matriz \mathbf{B} en la ecuación (13A.1) y Z_i son las variables asociadas con los ejes rotados que corresponden a los ejes de las curvas de nivel de la superficie de respuesta. El origen del sistema de coordenadas rotadas es el punto estacionario con todas las $Z_i = 0$ y respuesta \hat{y}_s .

Los valores específicos de \mathbf{B} son las raíces de la ecuación del determinante:

$$|\mathbf{B} - \lambda \mathbf{I}| = 0 \quad (13A.6)$$

donde \mathbf{I} es la matriz identidad. La relación entre la matriz \mathbf{B} y las λ_i es

$$\mathbf{Bm}_i = m_i \lambda_i, \quad i = 1, 2, \dots, k \quad (13A.7)$$

donde m_i son los vectores específicos correspondientes a las λ_i . Las m_i están normalizadas de manera que $m'_i m_i = 1$.

La relación entre las variables que representan los niveles codificados de los factores x y las variables de la ecuación canónica Z es:

$$\mathbf{Z} = \mathbf{M}'(\mathbf{x} - \mathbf{x}_s) \quad (13A.8)$$

donde las columnas de M son los vectores propios normalizados m_i .

Los distintos métodos necesarios para los cálculos se pueden encontrar en los libros de álgebra matricial como Graybill (1983), aunque los cálculos matriciales se pueden realizar con muchos programas de computadora. Los cálculos se ilustran con la ecuación de respuesta para el ejemplo 13.3:

$$\hat{y} = 169 + 6.747x_1 + 26.385x_2 - 10.875x_1^2 - 21.625x_2^2 - 15.25x_1x_2$$

La ecuación del determinante es:

$$\begin{vmatrix} -10.875 - \lambda & -7.625 \\ -7.625 & -21.625 - \lambda \end{vmatrix} = 0$$

con $\lambda^2 + 32.5\lambda + 177.03 = 0$. Las raíces de la ecuación cuadrática son $\lambda_1 = -25.58$ y $\lambda^2 = -6.92$. Entonces, la ecuación canónica con $\hat{y}_s = 177.25$ es:

$$\hat{y} = 177.25 - 25.58Z_1^2 - 6.92Z_2^2$$

La matriz de valores propios normalizados es:

$$\mathbf{M} = \begin{bmatrix} 0.4603 & 0.8877 \\ 0.8877 & -0.4603 \end{bmatrix}$$

Las coordenadas del punto estacionario son $x_{1s} = -0.156$ y $x_{2s} = 0.665$, y la relación entre las variables canónicas y las variables codificadas de los factores es:

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 0.4603 & 0.8877 \\ 0.8877 & -0.4603 \end{bmatrix} \begin{bmatrix} (x_1 + 0.156) \\ (x_2 - 0.665) \end{bmatrix}$$

o:

$$Z_1 = 0.4603(x_1 + 0.156) + 0.8877(x_2 - 0.665)$$

$$Z_2 = 0.8877(x_1 + 0.156) - 0.4603(x_2 - 0.665)$$

14 Diseños de parcelas divididas

En este capítulo se presenta el diseño de parcelas divididas “split plot” para experimentos factoriales y describe algunas características relativas a su estructura, composición de los errores experimentales, análisis y eficiencia relativa. Las extensiones y variaciones del diseño incluyen los diseños de parcelas divididas y de bloques divididos.

14.1 Parcelas de distintos tamaños en el mismo experimento

En los experimentos factoriales, a veces, un factor requiere más material experimental para su evaluación que otro factor. En las pruebas de campo en agronomía y horticultura, un factor como los métodos de cultivo puede requerir el uso de algún tipo de equipo que es más adecuado en parcelas grandes; mientras que otro factor, como nivel de fertilidad, se puede aplicar con facilidad a una parcela mucho más pequeña. La parcela de tratamiento de métodos de cultivo más grande, la *parcela completa*, se divide en *subparcelas* más pequeñas a las que se les aplican distintos tratamientos de cultivo o fertilidad. Esto se conoce como diseño de *parcelas divididas* y este ejemplo particular tiene dos unidades experimentales de diferente tamaño.

El experimento usado para el siguiente ejemplo explica la creación de un diseño de parcela dividida cuando se introduce un segundo factor en las subdivisiones de las unidades experimentales existentes para un experimento que está en progreso.

Ejemplo 14.1 Fertilizante de nitrógeno y acumulación de maleza en el pasto Pencross

En la mayoría de los campos de golf, el suelo es arena casi pura y se requiere irrigación frecuente y fertilizantes para mantener el pasto. El suelo arenoso tiene poca capacidad para retener el nitrógeno la fertilización se filtra con rapidez de la zona de la raíz después de irrigar. Administrar grandes dosis iniciales de nitrógeno

para retenerlo por períodos más largos es dañino para el pasto y los microbios del suelo, por lo que resultaría útil aplicar tasas moderadas de nitrógeno y retenerlo en la zona de la raíz.

Los fertilizantes de nitrógeno se fabrican en varias configuraciones químicas; se sabe que dos fertilizantes de uso común, sulfato de amonio y urea, liberan el nitrógeno con rapidez y se espera que se filtre en el suelo muy pronto. Otros, como el isobutildieno diurea (IBDU) y la urea cubierta de sulfuro, urea(SC), aunque más costosos, liberan las formas de nitrógeno más despacio y se espera retenerlo en el suelo más tiempo.

Un segundo factor que puede afectar la retención de nitrógeno es la hierba de pasto o pasto muerto que se junta. La hierba se recoge con frecuencia porque se piensa que acumula enfermedades, pero en realidad coadyuva a la retención del fertilizante de nitrógeno y, de manera parcial, podría aliviar algunas dificultades de la pérdida de nitrógeno en la arena.

Objetivo de investigación: un científico de suelos desea investigar los efectos del nitrógeno suministrado en distintas formas químicas y luego evaluar aquellos efectos combinados con la acumulación de hierba sobre la calidad del pasto sembrado.

Diseño del tratamiento: las cuatro formas de fertilizante de nitrógeno usadas en el estudio fueron 1) urea, 2) sulfato de amonio, 3) isobutildieno diurea (IBDU) y 4) urea con cubierta de sulfuro, urea(SC). Cada fertilizante debía suministrarse a una tasa de 1 libra de nitrógeno por 1000 pies cuadrados de pasto. Cualquier diferencia en las respuestas de los fertilizantes se podría atribuir a la forma de liberación del nitrógeno, ya que se suministró una cantidad equivalente de nitrógeno en cada forma.

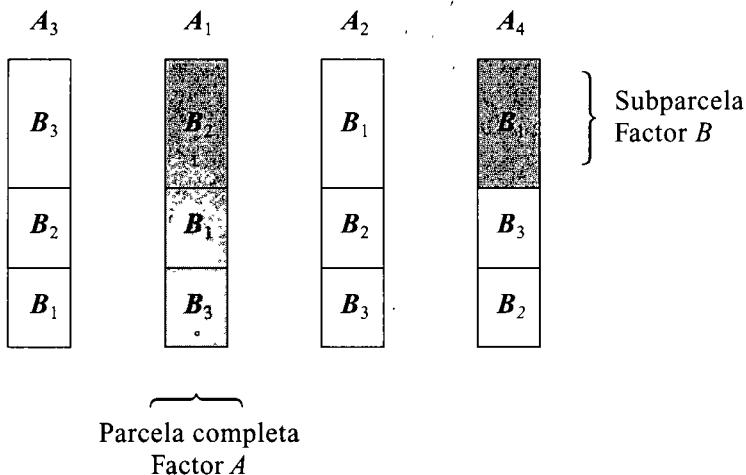
Diseño del experimento: se construyó un campo de golf y se sembró con pasto Pencross en las parcelas experimentales. Las parcelas de tratamiento se arreglaron en un diseño de bloques totalmente aleatorizado con dos réplicas.

Dos años más tarde se agregó el segundo factor de tratamiento: años de acumulación de hierba. Cada una de las ocho parcelas se dividió en tres subparcelas a las que se asignaron al azar los niveles del segundo factor. Los períodos que se permitió que se acumulara la hierba en las subparcelas de pasto fueron dos, cinco y ocho años.

El diseño obtenido de parcelas divididas tenía un factor de tratamiento de nitrógeno para parcelas completas en un diseño de bloques totalmente aleatorizado, con los años de acumulación de hierba como factor de tratamiento de las subparcelas.

En el cuadro 14.1 se muestra el plan para una réplica del diseño de parcelas divididas para este experimento, luego de la aleatorización. El tratamiento completo de nitrógeno, factor *A*, tiene cuatro niveles, y el factor de subparcelas, años de acumulación de hierba, factor *B*, tiene tres niveles.

Cuadro 14.1 Distribución experimental de parcelas divididas para el experimento de siembra de pasto



Fuente: Dr. I Pepper, Department of Soil and Water Science, University of Arizona.

Aleatorización alterada para diseños de parcelas divididas

La asignación aleatoria habitual para las combinaciones de tratamientos factorial a las unidades experimentales se alteró para cumplir con las necesidades particulares del experimento; por ejemplo, el diseño de parcelas divididas que se muestra en el cuadro 14.1 tiene una restricción adicional en la aleatorización respecto a la usual. En el diseño aleatorizado normal, las 12 combinaciones de tratamientos se asignan al azar a las 12 subparcelas; pero en el diseño de parcelas divididas, los tres niveles del factor *B* combinados con una sola forma de nitrógeno, factor *A*, están restringidos a la misma parcela completa. Entonces, el diseño de parcelas divididas es el producto de cambios específicos en la asignación aleatoria de las combinaciones de tratamientos factorial a las unidades experimentales.

Más posibilidades para dos tamaños de unidades

Si un factor, como temperatura, humedad o fotoperiodo, requiere cámaras de control ambiental, entonces cada nivel de ese factor requiere una cámara distinta. Se puede introducir un segundo factor de tratamiento, como el medio de cultivo para las plantas, e incluso un tercer factor de tratamiento dentro de cada cámara, con lo que se logra un uso económico de las cámaras para estudiar más de un factor a la vez. Las cámaras representan las parcelas completas y las subparcelas son las unidades a las que se aplica el segundo factor dentro de cada cámara.

En la investigación educativa se aplica un método de enseñanza a todos los alumnos de un salón como tratamiento de parcela completa; pero, se pueden usar subgrupos para estudiar un factor adicional, como el uso de ciertos materiales o microcomputadoras.

La investigación industrial puede necesitar réplicas de lotes obtenidos con mezclas de materias primas. El tratamiento de parcela completa, como mezclas de productos, puede ser un lote de materia prima, y el tratamiento de subparcelas, como tiempo de curado, se puede aplicar a sublotes de los lotes de mezcla de productos.

14.2 Dos errores experimentales para dos tamaños de parcelas

Como el análisis estadístico debe tomar en cuenta la presencia de dos tipos o tamaños diferentes de unidades experimentales en el experimento, los efectos del factor A se estiman a partir de las parcelas completas, mientras que los efectos del factor B y de la interacción AB se estiman a partir de las subparcelas para el factor B . Como las parcelas completas y subparcelas son unidades experimentales de distintos tamaños o tipos, tienen diferente precisión, lo que debe considerarse para hacer comparaciones entre las medias de tratamiento.

La consideración de dos errores separados es consecuencia del hecho de que las observaciones de distintas subparcelas en la misma parcela pueden tener una correlación positiva. La correlación refleja la naturaleza de las unidades experimentales que responden de manera similar cuando son adyacentes, como las subparcelas vecinas, los estudiantes en un salón de clase, los cultivos en una cámara de cultivo o las unidades del mismo lote de materia prima en un experimento industrial.

Se supone que existe una correlación ρ entre las observaciones en cualesquier dos subparcelas en la misma parcela completa y que las observaciones de dos parcelas completas diferentes no están correlacionadas. Dadas estas suposiciones, se puede demostrar que la varianza del error para los efectos principales de A por cada subparcela es $\sigma^2[1 + (b - 1)\rho]$ si hay b subparcelas en cada parcela completa. De la misma manera, para las comparaciones entre los efectos principales de B y las de la interacción AB la varianza del error por cada subparcela es $\sigma^2(1 - \rho)$.

Como consecuencia de estas diferencias en los errores asociados con las comparaciones de parcelas completas y subparcelas de tratamientos, la partición de la suma de cuadrados en el análisis de varianza se altera un poco con respecto a la partición para el diseño factorial de dos factores. Las particiones para los efectos del factor y los factores de bloque permanecen igual que los diseños factoriales usuales, pero el error experimental se divide en dos componentes. Una componente del error experimental se asocia con el factor de tratamiento de la parcela completa y la otra con el factor de tratamiento de la subparcela y la interacción.

14.3 Análisis para diseños de parcelas divididas

Modelo de parcelas divididas

Para diseñar parcelas divididas se usa una formulación de modelo mixto que refleje las distintas varianzas del error experimental para las subparcelas y la parcela completa, lo que incluye los efectos del error aleatorio por separado para ambas. Si el factor de tratamiento de la parcela completa se coloca en un diseño de bloques totalmente aleatorizado el modelo lineal es:

$$y_{ijk} = \mu + \alpha_i + \rho_k + d_{ik} + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad (14.1)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

donde μ es la media general, α_i es el efecto del i -ésimo nivel del factor A , ρ_k es el efecto del k -ésimo bloque, d_{ik} es el error aleatorio de la parcela completa, β_j es el efecto del j -ésimo nivel del factor B , $(\alpha\beta)_{ij}$ es el efecto de interacción entre ambos factores y e_{ijk} es el error aleatorio de la subparcela.

Se supone que los errores de la parcela completa y la subparcela son errores aleatorios independientes con distribución normal con media 0 y varianzas respectivas de σ_d^2 y σ_e^2 . La asignación aleatoria de los tratamientos a las unidades experimentales justifica la suposición de independencia para los dos errores aleatorios y la correlación igual entre los errores de las unidades en la subparcelas dentro de una misma parcela completa.

Análisis de varianza para parcelas divididas

En la tabla 14.1 se muestran los cuadrados medios esperados para el análisis de varianza mediante el método mixto de la ecuación (14.1), con los efectos fijos para los factores A y B .

Tabla 14.1 Cuadrados medios esperados para el análisis de varianza de parcelas divididas

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
Bloques	$r - 1$	CM bloques	
A	$a - 1$	CMA	$\sigma_e^2 + b\sigma_d^2 + rb\theta_a^2$
Error(1)	$(a - 1)(r - 1)$	$CME(1)$	$\sigma_e^2 + b\sigma_d^2$
B	$b - 1$	CMB	$\sigma_e^2 + ra\theta_b^2$
AB	$(a - 1)(b - 1)$	$CM(AB)$	$\sigma_e^2 + r\theta_{ab}^2$
Error(2)	$a(r - 1)(b - 1)$	$CME(2)$	σ_e^2

Los cuadrados medios esperados para el error(1) y el error(2) según el modelo mixto reflejan las diferencias en la variabilidad para los dos tipos de unidades

experimentales. Las varianzas del error esperadas para las parcelas completas son mayores que las de las subparcelas.

Pruebas de hipótesis sobre los efectos de factores

Los estadísticos F para probar las hipótesis nulas de la interacción y los efectos principales son:

- (interacción) $H_0: (\alpha\beta)_{ij} = 0$ contra $H_\alpha: (\alpha\beta)_{ij} \neq 0$ para alguna i, j

$$F_0 = \frac{CM(AB)}{CME(2)} \text{ con } (a-1)(b-1) \text{ y } a(r-1)(b-1) \text{ g.l.}$$

- (efectos principales del factor B) $H_0: \bar{\mu}_{1.} = \dots = \bar{\mu}_{.b}$ contra $H_a: \bar{\mu}_{1.} \neq \bar{\mu}_{.j}$ para alguna i, j

$$F_0 = \frac{CMB}{CME(2)} \text{ con } b-1 \text{ y } a(r-1)(b-1) \text{ g.l.}$$

- (efectos principales del factor A) $H_0: \bar{\mu}_1 = \dots = \bar{\mu}_{a.}$ contra $H_a: \bar{\mu}_i \neq \bar{\mu}_j$ para alguna i, j

$$F_0 = \frac{CMA}{CME(1)} \text{ con } a-1 \text{ y } (a-1)(r-1) \text{ g.l.}$$

Ejemplo 14.2 Datos y análisis

Una de las mediciones realizadas en las parcelas con pasto del ejemplo 14.1 fue el contenido de clorofila en la hierba cortada (mg/g) para las muestras de cada parcela. Los datos se muestran en la tabla 14.2 y el análisis de varianza en la 14.3.

Tabla 14.2 Contenido de clorofila (mg/g) en la hierba cortada

<i>Fuente del nitrógeno</i>	<i>Bloque</i>	<i>Años de acumulación de hierba</i>		
		2	5	8
Urea	1	3.8	5.3	5.9
	2	3.9	5.4	4.3
Sulfato de amonio	1	5.2	5.6	5.4
	2	6.0	6.1	6.2
IBDU	1	6.0	5.6	7.8
	2	7.0	6.4	7.8
Urea(SC)	1	6.8	8.6	8.5
	2	7.9	8.6	8.4

Tabla 14.3 Análisis de varianza del contenido de clorofila en recortes de pasto Pencross

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	23	48.78			
Bloque	1	0.51	0.51		
Nitrógeno (<i>N</i>)	3	37.32	12.44	29.62	0.010
Error (1)	3	1.26	0.42		
Hierba (<i>T</i>)	2	3.82	1.91	9.10	0.009
<i>N</i> × <i>T</i>	6	4.15	0.69	3.29	0.065
Error(2)	8	1.72	0.21		

Notas sobre los cálculos

El análisis de varianza para el diseño de parcelas divididas se puede calcular con muchos programas estadísticos disponibles. Las instrucciones para el análisis son equivalentes a las del diseño de tratamientos factoriales con dos factores, con excepción de la necesidad de calcular dos términos de error separados.

La suma de cuadrados para el error(1) es numéricamente equivalente a la suma de cuadrados del error en el diseño de experimento que se utilizó en parcelas completas, y con diseños totalmente aleatorizados es equivalente a la suma de cuadrados para parcelas completas dentro de los tratamientos del factor *A*. Para diseños de bloques totalmente aleatorizados, es equivalente a los cálculos de la suma de cuadrados para la interacción bloque × *A*. La suma de cuadrados del error(2) es, por lo común, la suma de cuadrados residual que el programa calcula en forma automática.

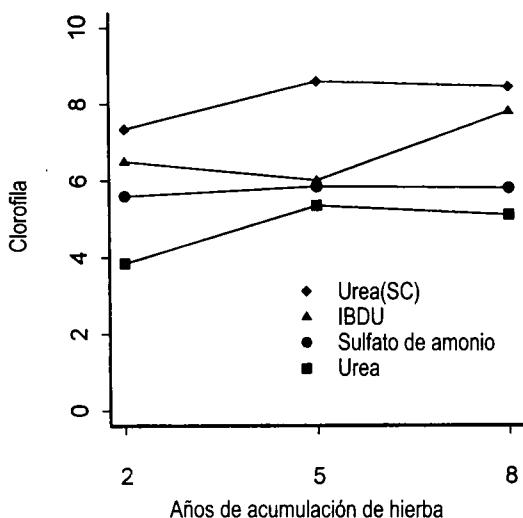
De manera alternativa, el análisis se puede realizar con los programas estadísticos escritos de manera específica para modelos mixtos que se han incorporado a muchos paquetes de cómputo, particularmente útiles para los diseños de parcelas divididas en diseños de bloques con bloques aleatorios. La estimación estadística en estos programas se basa en métodos de máxima verosimilitud que están fuera del objetivo de este libro.

La prueba para la interacción entre el nitrógeno y la hierba es $F_0 = 0.69/0.21 = 3.29$, y no es significativa con $Pr > F = .065$ (tabla 14.3). La prueba para las diferencias entre las medias de la hierba es $F_0 = 1.91/0.21 = 9.10$, y es significativa con $Pr > F = .009$. La prueba para las diferencias entre las medias de tratamiento de nitrógeno en parcelas completas es $F_0 = 12.44/0.42 = 29.62$, y es significativa con $Pr > F = .01$.

Las medias de celda y las medias marginales para la fuente de nitrógeno y los años de acumulación de hierba se muestran en la tabla 14.4. El pasto que recibió la urea(SC) tenía el contenido más alto de clorofila, seguido del IBDU y el sulfato de amonio, mientras que la urea rindió el menor contenido de clorofila. Esta jerarquización relativa del contenido de clorofila según la fuente de nitrógeno fue la misma para cada año de acumulación de hierba.

Tabla 14.4 Contenido medio de clorofila (mg/g) de los recortes de pasto Pencross

Fuente de nitrógeno	Años de acumulación de hierba			Medias de nitrógeno
	2	5	8	
Urea	3.85	5.35	5.10	4.77
Sulfato de amonio	5.60	5.85	5.80	5.75
IBDU	6.50	6.00	7.80	6.77
Urea(SC)	7.35	8.60	8.45	8.13
Medias de hierba	5.83	6.45	6.79	

**Figura 14.1** Gráfica del contenido medio de clorofila en recortes de pasto Pencross contra la fuente de nitrógeno para dos, cinco y ocho años de acumulación de hierba

La figura 14.1 muestra la gráfica del contenido medio de clorofila contra los años de acumulación de hierba por separado para cada *fuente* de nitrógeno. Las medias marginales en la tabla 14.4 indican un incremento en el contenido de clorofila cuando aumentan los años de acumulación de hierba. Sin embargo, la observación de las medias de celdas y sus valores graficados en la figura 14.1 muestran que esta tendencia general no se manifestó para cada fuente de nitrógeno por separado.

Esta diferencia en las tendencias entre las fuentes de nitrógeno indica la posibilidad de la presencia de alguna forma de interacción aun cuando la prueba global para la interacción no sea significativa. Algunas veces las particiones de la suma de cuadrados de la interacción revelan la presencia de ciertas formas de interacción entre los factores.

Interpretación con contrastes de regresión

El diseño factorial se compone de un factor cuantitativo, años de acumulación de hierba (T), y un factor cualitativo, fuente de nitrógeno (N). Este diseño conduce al uso de las particiones de regresión polinomial en el análisis de varianza para un factor cuantitativo con un factor cualitativo (presentado en el capítulo 6). Las particiones de suma de cuadrados de regresión lineal y cuadrática se pueden calcular para los años de acumulación (T) con las particiones correspondientes de la interacción nitrógeno \times hierba. Estas particiones se muestran en el análisis de varianza de la tabla 14.5.

Tabla 14.5 Análisis de varianza para el contenido de clorofila en recortes de pasto Pencross con particiones de regresión polinomial ortogonal para el factor hierba

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	23	48.78			
Bloques	1	0.51	0.51		
Nitrógeno (N)	3	37.32	12.44	29.62	0.010
Error(1)	3	1.26	0.42		
Hierba (T)	2	3.82	1.91	9.10	0.009
T lineal	1	3.71	3.71	17.67	0.003
T cuadrática	1	0.11	0.11	0.52	0.494
$N \times T$	6	4.15	0.69	3.29	0.065
$N \times T$ lineal	3	0.80	0.27	1.29	0.358
$N \times T$ cuadrática	3	3.36	1.12	5.33	0.028
Error(2)	8	1.72	0.21		

La interacción para las desviaciones cuadráticas de nitrógeno \times hierba es significativa con $F_0 = 1.12/0.21 = 5.33$ y $Pr > F = .028$. La partición de regresión lineal para hierba (T) es significativa con $F_0 = 3.71/0.21 = 17.67$ y $Pr > F = .003$. Así, existen desviaciones cuadráticas significativas de la respuesta lineal para los años de acumulación de hierba que difieren entre las fuentes de nitrógeno. Los distintos patrones se observan en la figura 14.1.

En este punto, pueden atenderse varios tipos de comparaciones para aclarar más la interpretación. Dado que hubo una componente de interacción significativa, será conveniente comparar las medias de nitrógeno para cada año de acumulación de hierba y determinar si la urea con cubierta de sulfuro siempre rindió el mayor contenido de clorofila en el pasto. Además, se puede calcular una regresión cuadrática del contenido de clorofila en los años de acumulación de hierba para cada fuente de nitrógeno y establecer las diferencias entre las fuentes de nitrógeno con respecto a los años de acumulación de hierba.

14.4 Errores estándar para las medias de los factores de tratamiento

En las pruebas de hipótesis, los errores estándar mostrados en la tabla 14.6 se usan para comparaciones entre las medias de tratamiento estimadas. Los grados de libertad asociados con cada error estándar son los del cuadrado medio usado en el error estándar, con la excepción única de comparación final mostrada en la tabla, donde el error estándar es una combinación ponderada de los dos cuadrados medios del error. En consecuencia, se pueden aproximar los grados de libertad adecuados mediante el procedimiento propuesto por Satterthwaite (1946) (Capítulo 5). La aproximación es:

$$\text{g.l.} = \frac{\frac{[(b-1)CME(2) + CME(1)]^2}{[(b-1)CME(2)]^2} + \frac{[CME(1)]^2}{f_1}}{f_2} \quad (14.2)$$

donde f_1 y f_2 son los grados respectivos para $CME(1)$ y $CME(2)$, a partir del análisis de varianza.

Tabla 14.6 Estimadores del error estándar para el diseño de parcelas divididas (A = factor de parcela completa, B = factor de subparcela)

Comparación de tratamientos	Estimador del error estándar
Diferencia entre dos medias A	$\sqrt{\frac{2CME(1)}{rb}}$
Diferencia entre dos medias B	$\sqrt{\frac{2CME(2)}{ra}}$
Diferencia entre dos medias B para el mismo nivel de A	$\sqrt{\frac{2CME(2)}{r}}$
Diferencia entre dos medias A para un nivel igual o diferente de que B	$\sqrt{\frac{2[(b-1)CME(2) + CME(1)]}{rb}}$
$\bar{y}_{uk} - \bar{y}_{vk}$ o bien $\bar{y}_{uk} - \bar{y}_{vm}$	

Debe observarse que las comparaciones que requieren errores estándar basados en la combinación lineal de medias cuadráticas $(b-1)CME(2) + CME(1)$, sólo tienen niveles de probabilidad aproximados para las pruebas de hipótesis y los intervalos de confianza. La combinación lineal de cuadrados medios ya no comparte

las mismas propiedades de la distribución de probabilidad de los cuadrados medios individuales.

Errores estándar para el experimento de siembra de pasto

Consultando la tabla 14.6, los errores estándar necesarios para el análisis del ejemplo 14.1 son las diferencias entre:

- dos medias de nitrógeno

$$\sqrt{\frac{2CME(1)}{rb}} = \sqrt{\frac{2(0.42)}{6}} = 0.37 \text{ con } 3 \text{ grados de libertad}$$

- dos medias de hierba

$$\sqrt{\frac{2CME(2)}{ra}} = \sqrt{\frac{2(0.21)}{8}} = 0.23 \text{ con } 8 \text{ grados de libertad}$$

- dos medias de hierba para la misma cantidad de nitrógeno

$$\sqrt{\frac{2CME(2)}{r}} = \sqrt{\frac{2(0.21)}{2}} = 0.46 \text{ con } 8 \text{ grados de libertad}$$

- dos medias de nitrógeno a igual o diferente nivel que la hierba

$$\sqrt{\frac{2[(b-1)CME(2) + CME(1)]}{rb}} = \sqrt{\frac{2[2(0.21) + 0.42]}{6}} = 0.53$$

donde los grados de libertad para el último error estándar obtenido a partir de la aproximación de Satterthwaite son:

$$g.l. = \frac{\frac{[2(0.21) + 0.42]^2}{[2(0.21)]^2} + \frac{[0.42]^2}{3}}{\frac{8}{3}} = 8.73 \text{ o bien } 9 \quad (14.3)$$

La varianza del error experimental para las subparcelas, $CME(2) = 0.21$, es la mitad de la varianza del error experimental para las parcelas completas, $CME(1) = 0.42$. Así, las comparaciones entre los años de acumulación de hierba y las medias de interacción serán más precisas que las comparaciones entre las medias de las fuentes de nitrógeno. La interrogante de si se hizo un trueque equitativo entre las ventajas y desventajas de dividir parcelas se responderá en la sección 14.6 con los cálculos de eficiencia relativa.

14.5 Características del diseño de parcelas divididas

La aleatorización para los diseños de parcelas divididas requiere que los niveles del factor A se asignen al azar a las unidades completas, de acuerdo con el protocolo para el diseño de experimentos en el que se arreglan las unidades completas, es decir, totalmente aleatorizado, bloque completo aleatorizado, etc. Los niveles del factor B se asignan al azar a las subunidades *dentro* de cada unidad completa, por separado para cada unidad completa.

El diseño también se puede describir como un diseño factorial *confundido* (presentado en el capítulo 11), donde las subunidades se consideran unidades experimentales, de manera que los niveles del factor A se aplican a los grupos o bloques de las subunidades. En consecuencia, las comparaciones entre los niveles de A se *confunden* con los bloques de las subunidades. El diseño de parcelas divididas suele recibir el nombre de *diseño factorial confundido*, en él los *efectos principales* se confunden con los bloques, mientras que en los diseños del capítulo 11, la confusión se limita a las *interacciones*.

Análisis de varianza para diseños experimentales comunes

Las fuentes de variación y los grados de libertad de las particiones de la suma de cuadrados se muestran en la tabla 14.7 para aquel diseño de parcelas divididas en el que las parcelas completas se arreglan en tres diseños experimentales comunes (totalmente aleatorizado, bloque completo aleatorizado y cuadrado latino). Cada diseño tiene r réplicas, a niveles de A y b niveles de B .

Tabla 14.7 Fuentes de variación y grados de libertad para el análisis de varianza de los diseños de parcelas divididas

<i>Totalmente aleatorizado</i>		<i>Bloque completo aleatorizado</i>		<i>Cuadrado latino</i>	
		<i>Parcelas completas</i>			
<i>Fuente</i>	<i>g.l.</i>	<i>Fuente</i>	<i>g.l.</i>	<i>Fuente</i>	<i>g.l.</i>
A	$a - 1$	Bloques	$r - 1$	Renglones	$a - 1$
Error(1)	$a(r - 1)$	A	$a - 1$	Columnas	$a - 1$
		Error(1)	$(r - 1)(a - 1)$	A	$a - 1$
				Error(1)	$(a - 1)(a - 2)$
<i>Subparcelas</i>					
B	$b - 1$	B	$b - 1$	B	$b - 1$
AB	$(a - 1)(b - 1)$	AB	$(a - 1)(b - 1)$	AB	$(a - 1)(b - 1)$
Error(2)	$a(r - 1)(b - 1)$	Error(2)	$a(r - 1)(b - 1)$	Error(2)	$a(a - 1)(b - 1)$

Diferente precisión para las medias de factores de parcelas completas y subparcelas

Los mayores grados de libertad asociados con las estimaciones del error experimental para las medias del factor B , de la interacción AB , y las comparaciones

entre ellas, indican que los efectos de B y AB se estiman con más precisión que los efectos de A medidos en las parcelas completas. Sin embargo, esto puede ser engañoso, ya que $CME(1)$ y $CME(2)$ tienen menos grados de libertad para el diseño de parcelas divididas que el cuadrado medio del error experimental en el diseño sin parcelas divididas para los efectos de dos factores. Por ejemplo, el mismo diseño factorial, pero en un diseño de bloques totalmente aleatorizado sin restricciones de parcelas divididas, tendría $(r - 1)(ab - 1)$ grados de libertad para el error experimental, que excede el $CME(1)$ o el $CME(2)$ por $(r - 1)(a - 1)$ y $a(a - 1)(b - 1)$ grados de libertad, respectivamente.

La experiencia práctica con los diseños de parcelas divididas ha demostrado que $CME(2)$ con frecuencia es menor que $CME(1)$, como se presume a partir del cuadrado medio esperado. En consecuencia, habrá un incremento en la precisión de las estimaciones de los efectos de B y AB que compensan la pérdida de los grados de libertad. Sin embargo, el error experimental promedio de todos los efectos de tratamiento es el mismo con o sin la característica de parcelas divididas, por tanto, no hay una ganancia neta con el diseño de parcelas divididas. Dicho de otra manera, si hay ganancia en la precisión de los efectos de B y AB , esta se compensa con la pérdida en la precisión de los efectos de A ; por lo tanto, efectos importantes e interesantes de A se pueden juzgar insignificantes. La evaluación de la eficiencia relativa del diseño de parcelas divididas en este contexto se presenta en la sección 14.6.

Ventajas del experimento de parcelas divididas

Ya se ha mencionado la ventaja principal de las parcelas divididas: cuando un factor requiere una mayor cantidad de material experimental que otro factor, como en los estudios de agronomía, o cuando existe la oportunidad de estudiar respuestas de un segundo factor, mientras que se da un uso eficiente de los recursos, como en los estudios en cámaras de cultivo. El experimento del ejemplo 14.1 explica la creación de un diseño de parcelas divididas con la introducción de un segundo factor a las subdivisiones de las unidades experimentales existentes de un experimento en proceso.

Desventajas del experimento de parcelas divididas

Las principales desventajas del diseño de parcelas divididas incluyen la posible pérdida en la precisión de las comparaciones de tratamientos y un incremento en la complejidad del análisis estadístico. El análisis de varianza y la estimación de los errores estándar para los distintos tipos de comparaciones de tratamientos requiere de más cálculos.

14.6 Eficiencia relativa de las comparaciones de subparcelas y parcelas completas

Es común atribuir, la eficiencia relativa de un diseño de experimento a la eficiencia como resultado del uso de bloques mediante algún factor y la de ignorar el factor de bloque. Con los diseños de parcelas divididas, es revelador considerar la

eficiencia relativa al usarlos en lugar de diseñar el mismo experimento sin la característica de las parcelas divididas. Por ejemplo, cuando el factor de tratamiento de la parcela completa se arregla en un diseño de bloques totalmente aleatorizado, es posible determinar cuál de los diseños es más eficiente en las comparaciones de tratamientos de parcelas completas y de interacción. Como se explica en la sección 14.5, existe una relación entre el incremento en la precisión de las medias de tratamiento en las subparcelas y la disminución en la precisión de las medias de tratamiento de las parcelas completas.

Eficiencia de las comparaciones en la subparcelas

Federer (1955) muestra que la eficiencia del diseño de parcelas divididas con respecto al diseño de bloques totalmente aleatorizado para las comparaciones en las subparcelas es:

$$RE = K \frac{a(b-1)CME(2) + (a-1)CME(1)}{(ab-1)CME(2)} \quad (14.4)$$

donde

$$K = \frac{(f_1 + 1)(f_2 + 3)}{(f_1 + 3)(f_2 + 1)}$$

$f_1 = a(b-1)(r-1)$, los grados de libertad para el error en la subparcela $CME(2)$, y $f_2 = (ab-1)(r-1)$, los grados de libertad para el error experimental del bloque totalmente aleatorizado.

Eficiencia de las comparaciones de parcelas completas

La eficiencia relativa de los diseños de parcelas divididas con respecto al diseño de bloques totalmente aleatorizado para las parcelas completas es:

$$RE = K \frac{a(b-1)CME(2) + (a-1)CME(1)}{(ab-1)CME(1)} \quad (14.5)$$

con $f_1 = (a-1)(r-1)$ para el error de la parcela completa $CME(1)$.

La eficiencia del diseño de parcelas divididas para las comparaciones de la hierba y la subparcela nitrógeno \times hierba con respecto al de 12 combinaciones distintas, asignadas al azar a las 12 parcelas en cada bloque, con $a = 4$, $b = 3$, $r = 2$, $f_1 = 8$ y $f_2 = 11$ es:

$$RE = \frac{(9)(14)}{(11)(12)} \times \frac{8(0.21) + 3(0.42)}{11(0.21)} = 0.95(1.27) = 1.21$$

o una ganancia del 21%. Para las comparaciones de nitrógeno en parcelas completas con $f_1 = 3$ y $f_2 = 11$, la eficiencia relativa es:

$$RE = \frac{(4)(14)}{(6)(12)} \times \frac{8(0.21) + 3(0.42)}{11(0.42)} = 0.78(0.64) = 0.49$$

o sólo 49%. Por lo tanto, con el diseño de parcelas divididas se obtuvo una ganancia de sólo un 21% en la eficiencia para las comparaciones de subparcelas y una pérdida del 51% para las comparaciones de las parcelas completas con respecto al diseño de bloques totalmente aleatorizado.

14.7 Diseño con doble subdivisión de parcelas para tres factores de tratamiento

La conveniencia de introducir un tercer factor en el diseño de tratamientos requiere una subdivisión de las subparcelas de manera que todos los niveles del tercer factor se apliquen a estas nuevas subdivisiones, que se conocen como subsubparcelas. El diseño, llamado de **subsubparcelas**, tiene tres tamaños o tipos diferentes de unidades experimentales y su análisis requiere el cálculo de una suma de cuadrados adicional para el error experimental asociado con las subsubparcelas. La partición de la suma de cuadrados para los factores A y B es la mostrada en la tabla 14.7 y si el tercer factor, C , tiene c niveles, las fuentes de variación adicionales y los grados de libertad asociados en el análisis de varianza son:

<i>Fuente</i>	<i>Grados de libertad</i>
C	$c - 1$
AC	$(a - 1)(c - 1)$
BC	$(b - 1)(c - 1)$
ABC	$(a - 1)(b - 1)(c - 1)$
Error(3)	$ab(r - 1)(c - 1)$

La suma de cuadrados para el error(3) es igual a la suma de cuadrados residual que acostumbra producir un programa de computadora. Deben darse instrucciones especiales para los cálculos separados de la suma de cuadrados para el error(2). Para los diseños de bloque totalmente aleatorizados, la suma de cuadrados para el error(2) se calcula como la suma de cuadrados agregada para la interacción bloques $\times B$ y bloques $\times AB$.

Todos los errores estándar aplicados a los efectos A y B mostrados en la tabla 14.6 son válidos y sólo necesitan que se incluya el valor de c como parte del divisor. Por ejemplo, el error estándar para la diferencia entre dos medias del factor A $\bar{y}_u - \bar{y}_v$ es $\sqrt{2CME(1)/rbc}$. Los errores estándar necesarios para las comparaciones que involucran el factor C se muestran en la tabla 14.8.

14.8 Diseño de bloques divididos

Los investigadores han usado numerosas variantes del diseño básico de parcelas divididas más allá de las subdivisiones de las unidades experimentales. Una variante muy común es el diseño de bloques divididos, donde se aplican los tratamientos en subunidades en una banda que cruza todas las unidades de parcelas

completas. En ocasiones este diseño recibe el nombre de diseño de *parcelas en banda*.

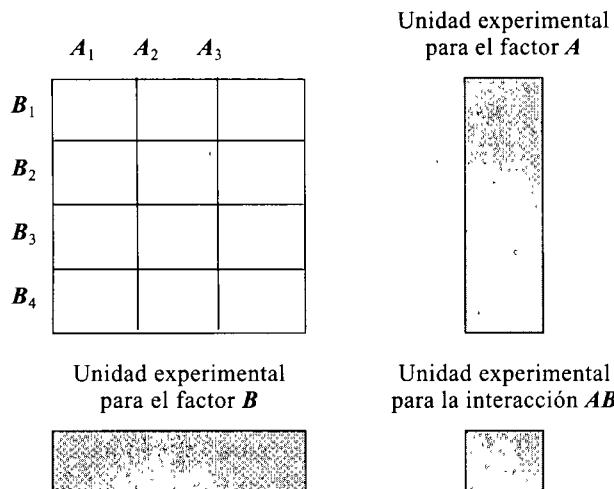
El diseño de bloques divididos puede ser útil en estudios agrícolas cuando dos factores de tratamiento deben usar parcelas grandes. Los niveles de un factor de tratamiento ***A*** se asignan al azar a las parcelas del diseño de bloques totalmente aleatorizado.

Tabla 14.8 Estimadores del error estándar para el diseño de subparcelas

Comparación de tratamientos	Estimador del error estándar
Dos medias de <i>C</i> ($\bar{y}_{u.} - \bar{y}_{v.}$)	$\sqrt{\frac{2CME(3)}{rab}}$
Dos medias de <i>C</i> en el mismo nivel de <i>A</i> ($\bar{y}_{i.u.} - \bar{y}_{i.v.}$)	$\sqrt{\frac{2CME(3)}{rb}}$
Dos medias de <i>C</i> en el mismo nivel de <i>B</i> ($\bar{y}_{iu.} - \bar{y}_{iv.}$)	$\sqrt{\frac{2CME(3)}{ra}}$
Dos medias de <i>C</i> en el mismo nivel de <i>A</i> y <i>B</i> ($\bar{y}_{yu.} - \bar{y}_{yv.}$)	$\sqrt{\frac{2CME(3)}{r}}$
Dos medias de <i>B</i> en un nivel igual o diferente de <i>C</i> ($\bar{y}_{ui.} - \bar{y}_{vi.}$ o bien $\bar{y}_{ui.} - \bar{y}_{vj.}$)	$\sqrt{\frac{2[(c-1)CME(3) + CME(2)]}{rac}}$
Dos medias de <i>B</i> en el mismo nivel de <i>A</i> y <i>C</i> ($\bar{y}_{uuj} - \bar{y}_{vij}$)	$\sqrt{\frac{2[(c-1)CME(3) + CME(2)]}{rc}}$
Dos medias de <i>A</i> en un nivel igual o diferente de <i>C</i> ($\bar{y}_{ui.} - \bar{y}_{vi.}$ o bien $\bar{y}_{ui.} - \bar{y}_{vj.}$)	$\sqrt{\frac{2[(c-1)CME(3) + CME(1)]}{rbc}}$
Dos medias de <i>A</i> en un nivel igual o diferente de <i>B</i> y <i>C</i> por ejemplo $\bar{y}_{uuj} - \bar{y}_{vij}$ o $\bar{y}_{uuj} - \bar{y}_{vlk}$)	$\sqrt{\frac{2[b(c-1)CME(3) + (b-1)CME(2) + CME(1)]}{rbc}}$

Las parcelas para el segundo factor ***B*** se construyen de la misma manera, pero se colocan perpendiculares a las parcelas del factor ***A***; los niveles del segundo factor ***B***

Cuadro 14.2 Unidades experimentales para un bloque de un diseño de bloques divididos



se asignan al azar a este segundo arreglo de parcelas a través del mismo bloque. El cuadro 14.2 presenta un bosquejo de un bloque para el diseño de bloques divididos, en donde hay $a = 3$ niveles del factor A y $b = 4$ niveles del factor B .

Tres unidades con diferente tamaño y tres errores experimentales

El diseño de bloques divididos tiene tres tamaños de unidades experimentales donde las unidades para los efectos principales de A y B son equivalentes a las parcelas completas, cada una con diferente orientación. La unidad para el efecto de interacción AB es la subparcela que se encuentra en la intersección de las dos parcelas completas para los respectivos niveles de A y B . En consecuencia, debe estimarse un error experimental separado a partir del análisis de varianza para cada uno de los tres efectos de tratamiento.

El modelo estadístico lineal para los diseños de bloques divididos difiere del diseño de parcelas divididas normal en que tiene términos de error aleatorio separados para cada uno de los tres efectos de tratamiento. También se incluye un efecto de bloque ρ_k en el modelo:

$$y_{ijk} = \mu + \rho_k + \alpha_i + d_{ik} + \beta_j + g_{jk} + (\alpha\beta)_{ij} + e_{ijk} \quad (14.6)$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, r$$

Los efectos del error aleatorio, d_{ik} , g_{jk} y e_{ijk} son los errores experimentales para las respectivas unidades asociadas con los efectos de A , B y AB , con varianzas σ_d^2 , σ_g^2 y σ_e^2 . El análisis de varianza para este modelo junto con las medias cuadráticas esperadas para los efectos de tratamiento fijos se muestran en la tabla 14.9.

Tabla 14.9 Descripción del análisis de varianza para el diseño de bloques divididos con dos factores de tratamiento

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Cuadrados medios esperados</i>
Bloques	$r - 1$	SC bloques	
<i>A</i>	$a - 1$	SCA	$\sigma_e^2 + b\sigma_d^2 + rb\theta_a^2$
Error(1)	$(r - 1)(a - 1)$	$SCE(1)$	$\sigma_e^2 + b\sigma_d^2$
<i>B</i>	$b - 1$	SCB	$\sigma_e^2 + a\sigma_g^2 + ra\theta_b^2$
Error(2)	$(r - 1)(b - 1)$	$SCE(2)$	$\sigma_e^2 + a\sigma_g^2$
<i>AB</i>	$(a - 1)(b - 1)$	$SC(AB)$	$\sigma_e^2 + r\theta_{ab}^2$
Error(3)	$(r - 1)(a - 1)(b - 1)$	$SCE(3)$	σ_e^2

Errores estándar para las medias del factor de tratamiento

Los errores estándar para los bloques divididos son más complejos que los del diseño de parcelas divididas debido a la alteración en el patrón de aleatorización de los niveles del factor *B*. Los errores estándar básicos de las medias de tratamiento y las diferencias de tratamiento se muestran en la tabla 14.10.

Los grados de libertad para los errores estándar con un solo cuadrado medio serán los asociados con el mismo. Para los errores estándar con dos cuadrados medios o más es necesario obtener los valores de los grados de libertad con la aproximación de Satterthwaite.

14.9 Información adicional sobre diseños de parcelas divididas

En este capítulo se estudió el diseño básico de parcelas divididas y varias extensiones del mismo, presentando un gran número de modificaciones y aplicaciones del diseño en otros casos como tratamientos en subunidades en diseño de cuadrado latino, con comparaciones *confundidas* entre los tratamientos de subunidades, y arreglos sistemáticos de los tratamientos de unidades completas (Cochran y Cox, 1957; Petersen, 1985). Robinson (1967) presentó un diseño de parcelas divididas y su análisis para subparcelas en un arreglo de bloque incompleto balanceado, y Coons *et al.* (1989) usaron un diseño de parcelas divididas con parcelas completas en un diseño de bloque incompleto balanceado. Mejza y Mejza (1996) proporcionaron las propiedades estadísticas de los diseños de parcelas divididas con bloques incompletos y Hering y Mejza (1997) demostraron las propiedades de los diseños de bloques divididos con bloques incompletos. Federer (1975) incluyó un extenso arreglo de modificaciones con la conexión a series de experimentos repetidos en distintos lugares y tiempos. Little y Rubin (1987) y Jarret (1978) proporcionaron detalles de discusiones relevantes sobre el análisis de los diseños de parcelas divididas con datos faltantes y las formulaciones para obtener los errores estándar para los contrastes

generales de tratamientos. Steel y Torrie (1980) realizaron estudios sobre el análisis de datos de cosechas repetidas para cultivos perennes como diseños.

Tabla 14.10 Estimadores de los errores estándar para el diseño de bloques divididos

<i>Comparación de tratamientos</i>	<i>Estimador del error estándar</i>
Dos medias A ($\bar{y}_u - \bar{y}_v$)	$\sqrt{\frac{2CME(1)}{rb}}$
Dos medias B ($\bar{y}_u - \bar{y}_v$)	$\sqrt{\frac{2CME(2)}{ra}}$
Dos medias A , mismo nivel de B ($\bar{y}_{uj} - \bar{y}_{vj}$)	$\sqrt{\frac{2[(b-1)CME(3) + CME(1)]}{rb}}$
Dos medias B , mismo nivel de A ($\bar{y}_{ju} - \bar{y}_{jv}$)	$\sqrt{\frac{2[(a-1)CME(3) + CME(2)]}{ra}}$
Cualesquiera dos medias, diferentes niveles de A y B ($\bar{y}_{uj} - \bar{y}_{vk}$)	$\sqrt{\frac{2[aCME(1) + bCME(2) + (ab-a-b)CME(3)]}{rab}}$

de parcelas divididas. Como las cosechas repetidas de la misma parcela caen dentro de la categoría general de mediciones repetidas y estudios longitudinales, en el siguiente capítulo se encuentra un análisis del tema.

EJERCICIOS PARA EL CAPÍTULO 14

1. Se realizó un experimento con sorgo, utilizando un diseño de parcelas divididas con dos factores de tratamiento: densidad de población de plantas e híbrido. Las parcelas completas se usaron para los cuatro niveles de densidad de población: 10, 15, 25 y 40 plantas por metro de fila. Había tres híbridos asignados al azar a las subparcelas de cada parcela. Los datos señalados en la tabla son los pesos de la semilla por planta, en gramos.
 - a. Escriba un modelo lineal para este experimento, explique los términos y calcule el análisis de varianza para los datos.
 - b. Construya un tabla de medias de celdas y medias marginales para este experimento y calcule los errores estándar estimados para la tabla de medias.

<i>Híbrido</i>	<i>Bloque</i>	<i>Peso de semillas (g) de una prueba con sorgo</i>			
		10	15	25	40
TAM 680	1	40.7	24.2	16.1	11.2
	2	37.8	44.4	17.6	12.7
	3	32.9	27.8	19.9	14.5
	4	43.1	34.1	20.1	15.4
RS 671	1	39.4	31.3	17.9	14.8
	2	47.8	34.5	30.5	17.3
	3	44.4	25.6	22.5	17.7
	4	49.0	50.4	25.2	18.7
Tx 399	1	68.7	26.2	20.5	18.9
Tx × 2536	2	56.2	48.1	28.2	26.2
	3	44.8	41.1	30.0	19.2
	4	59.3	46.0	24.7	22.0

Fuente: Dr. R. Voigt, Department of Plant Sciences, University of Arizona.

- c. Calcule los errores estándar estimados para las diferencias de dos medias observadas:
 - i) para híbridos.
 - ii) para poblaciones de plantas.
 - iii) para dos híbridos de la misma población de plantas.
- d. Pruebe las hipótesis para los efectos de interacción y principales, suponga efectos fijos para los híbridos y las poblaciones de plantas.
- e. Calcule la eficiencia relativa de este diseño de parcelas divididas para los tratamientos de subparcelas y parcelas con respecto al diseño de bloques completo aleatorizado normal e interprete.
- f. Obtenga una partición de suma de cuadrados para la población de plantas y la suma de cuadrados de la interacción con las particiones de regresión polinomial adecuadas e interprete los resultados. Los coeficientes para las particiones lineal, cuadrática y cúbica para los cuatro niveles de plantas por metro se muestran a continuación.

<i>Plantas</i>	10	15	25	40
<i>Lineal</i>	-0.546	-0.327	0.109	0.764
<i>Cuadrática</i>	0.513	-0.171	-0.741	0.399
<i>Cúbica</i>	-0.435	0.783	-0.435	0.087

- 2. Se realizó un experimento de parcelas divididas en un diseño totalmente aleatorizado con tratamientos de parcelas completas como un factorial de 2×2 (factores *A* y *B*) y los tratamientos de las subparcelas como tres niveles del factor *C*, había cuatro réplicas de las unidades experimentales. Suponga que todos los tratamientos eran fijos.
 - a. Escriba un modelo lineal para el experimento. Identifique cada componente del modelo y muestre los intervalos numéricos en el subíndice.

- b. Describa la tabla de análisis de varianza, muestre las fuentes de variación, los grados de libertad y los cuadrados medios esperados.
3. Suponga que los tratamientos de parcelas completas descritos en el ejercicio 14.2 se arreglaron en un diseño de cuadrado latino. Repita los incisos a) y b) para este caso.
4. Suponga que los tratamientos de subparcelas del ejercicio 14.2 eran un arreglo factorial de 3×3 de los factores C y D con el resto de las condiciones iguales. Repita los incisos a) y b) para este caso.
5. Un investigador de ciencias alimentarias desea realizar un experimento para evaluar el efecto de las condiciones de almacenamiento en frío sobre la calidad del alimento. Los dos factores de tratamiento que usará son la temperatura de almacenamiento y el material de los contenedores. El producto alimenticio se colocará en uno de los contenedores y se almacenará en una cámara de control de temperatura durante un periodo fijo luego del cual se medirán varias condiciones físicas y de calidad del producto en cada contenedor. El investigador tiene tres pequeñas cámaras de control de temperatura disponibles para el experimento, las temperaturas de almacenamiento son 2° , 4° y 8°C . Existen cuatro tipos de contenedores para el estudio: plástico sellado, cartón encerado sellado, cartón sellado y un contenedor sin sellar como control. Cuenta con cuatro posiciones en el centro de la cámara donde puede colocar los contenedores. Dibuje un diagrama de un plan para el experimento de manera que el investigador pueda tener tres réplicas del mismo. Utilice la siguiente guía para construir su diagrama.
- Use etiquetas I, II y III para identificar las tres cámaras de temperatura y etiquetas a , b , c y d para identificar las cuatro posiciones dentro de cada cámara.
 - Muestre el tipo de contenedor (C1, C2, C3 o C4) asignado a cada posición en cada cámara, lo mismo que la temperatura (2° , 4° , 8°) asignada a la cámara, para las tres réplicas en su diagrama.
 - Muestre el esquema de aleatorización que usó en la asignación del inciso b).
6. Un investigador especialista en una compañía de productos del mar investigó el crecimiento de bacterias en ostiones y mejillones sujeto a tres temperaturas de almacenamiento. Disponía de nueve unidades de almacenamiento en frío, de las que se seleccionaron tres unidades al azar para usar con cada nivel de temperatura. Los ostiones y los mejillones se almacenaron durante dos semanas en cada unidad, al cabo de dos semanas se hizo una cuenta de bacterias de una muestra de ostiones y mejillones. Al final de ejercicio se muestra el logaritmo de la cuenta de bacterias para cada muestra.
- El investigador pudo tener tres réplicas para el estudio con sólo tomar tres muestras aleatorias de cada alimento en una unidad de almacenamiento con una temperatura. De esta manera, sólo habría necesitado tres unidades de almacenamiento para el estudio, una para cada temperatura. Explique las posibles dificultades del estudio si lo hubiera realizado de esta manera.
 - ¿Hay un incremento significativo en el crecimiento bacterial cuando aumenta la temperatura? Justifique su respuesta.
 - ¿Existe una interacción significativa entre el tipo de alimento y el aumento (si lo hay) en el crecimiento de bacterias con la temperatura? Justifique su respuesta.
 - Escriba un modelo lineal para su análisis, establezca las suposiciones y explique los términos.
 - Determine si sus suposiciones sobre el modelo lineal son correctas para los datos.

<i>Unidad de enfriamiento</i>	<i>Temperatura (°C)</i>	<i>Marisco*</i>	<i>log(cuenta)</i>
1	0	1	3.6882
1	0	2	0.3565
2	0	1	1.8275
2	0	2	1.7023
3	0	1	5.2327
3	0	2	4.5780
4	5	1	7.1950
4	5	2	5.0169
5	5	1	9.3224
5	5	2	7.9519
6	5	1	7.4195
6	5	2	6.3861
7	10	1	9.7842
7	10	2	10.1352
8	10	1	6.4703
8	10	2	5.0482
9	10	1	9.4442
9	10	2	11.0329

*1 = ostiones, 2 = millones

7. Un experimento de parcelas divididas en un diseño de bloques completo aleatorizado evaluó los efectos de las tasas de nitrógeno, agua y fósforo sobre la eficiencia del uso del agua en un cultivo de maíz dulce con irrigación por goteo. Dos tasas de fósforo ($P_1 = 0$ y $P_2 = 245$ lb P₂O₅/acre) se asignaron al azar a parcelas completas en un diseño de bloques completo aleatorizado. Los tratamientos del factorial de 3 × 3 de nitrógeno (0, 130, y 260 lb N/acre) y agua (16, 22 y 28 pulgadas) se asignaron en forma aleatoria a las parcelas principales. Los datos mostrados en la tabla se refieren a la eficiencia en el uso del agua para cada parcela.
- Escriba un modelo lineal para el experimento, explique los términos y realice el análisis de varianza para los datos.
 - Construya una tabla de medias de celdas y medias marginales para este experimento y calcule los errores estándar estimados para la tabla de medias.
 - Calcule los errores estándar estimados para las diferencias entre dos medias observadas para:
 - las tasas de fósforo.
 - los niveles de agua.
 - las tasas de nitrógeno.
 - el agua por medias de celdas de nitrógeno.
 - Pruebe las hipótesis para todas las interacciones y efectos principales e interprete los resultados.
 - Realice las particiones de sumas de cuadrados para agua y nitrógeno en particiones de regresión polinomial lineal y cuadrática que incluyan la interacción. Interprete los resultados; use una gráfica de las medias observadas para ayudarle en la interpretación.

<i>Agua</i>	<i>Nitrógeno</i>	<i>Bloque 1</i>		<i>Bloque 2</i>	
		<i>P</i> ₁	<i>P</i> ₂	<i>P</i> ₁	<i>P</i> ₂
16	0	8.1	9.7	8.6	15.5
	130	36.0	34.2	34.5	33.1
	260	34.6	34.0	40.7	39.3
22	0	10.0	6.2	5.1	10.9
	130	21.5	19.7	19.9	21.9
	260	30.7	28.9	26.4	25.7
28	0	10.6	6.3	4.5	10.4
	130	19.4	19.7	21.7	19.9
	260	23.2	23.0	19.4	23.2

Fuente: Dr. T. Doerge, Department of Soil and Water Science, University of Arizona.

15 Diseños con mediciones repetidas

El tema central de este capítulo se refiere a mediciones repetidas de la variable de respuesta en cada unidad experimental, haciendo hincapié en experimentos con observaciones realizadas en ocasiones sucesivas en el tiempo. Se estudian las propiedades estadísticas de las observaciones y se presentan los métodos adecuados para analizar los datos.

15.1 Estudios de tendencias en el tiempo

La tendencia en el tiempo de las respuestas individuales al tratamiento es un aspecto importante para muchos experimentos, como aquellos en los que se pesan animales cada semana para supervisar su crecimiento bajo diferentes condiciones de nutrición o se cosechan parcelas de cultivos perennes como alfalfa varias veces. Las **mediciones repetidas** existen con frecuencia en pruebas clínicas, cuando la respuesta de los pacientes a un tratamiento se mide regularmente para supervisarla.

Las mediciones repetidas en cada unidad experimental proporcionan información sobre la tendencia en tiempo de la variable de respuesta bajo diferentes condiciones de tratamiento. Las tendencias en el tiempo pueden revelar qué tan rápido responden las unidades al tratamiento o durante cuánto tiempo se manifiestan los efectos del tratamiento en las unidades del estudio. También es posible evaluar las diferencias entre las tendencias de los tratamientos.

El resultado de observaciones repetidas es una mayor precisión

Las observaciones repetidas de la misma unidad experimental a través del tiempo, con frecuencia son más eficientes que el uso de una unidad experimental diferente para cada observación en el tiempo. No sólo se requieren menos unidades, lo que reduce costos, también la estimación de las tendencias en el tiempo serán más precisas. El aumento en la precisión se debe a que las mediciones en la misma unidad tienden a ser menos variables que las mediciones en unidades distintas; por tanto el efecto de las mediciones repetidas es similar al efecto del uso de bloques.

Ejemplo 15.1 Detección temprana de flebitis en terapia con amiodarone

Diseño del tratamiento: se diseñó el experimento descrito con detalle en el ejemplo 2.2, para explorar mecanismos de detección temprana de la flebitis durante la terapia con amiodarone. La flebitis es una inflamación de una vena que se presenta después de la administración intravenosa de drogas. Se administraron tres tratamientos intravenosos a los animales de prueba: 1) amiodarone con una solución como excipiente, 2) la solución excipiente sola y 3) una solución salina.

Diseño del experimento: los conejos usados como animales de prueba se asignaron al azar a los tres tratamientos en un diseño totalmente aleatorizado. Una aguja insertada en la vena de la oreja del conejo fue el medio de administración del tratamiento, por lo que se supervisó la temperatura de ambas orejas durante varias horas y se consideró que un incremento en la temperatura de la oreja tratada era un indicio temprano de flebitis. La diferencia en las temperaturas de las dos orejas (tratada menos no tratada) se usó como la variable de respuesta.

Mediciones repetidas: durante el estudio se observaron las temperaturas de cada conejo cada 30 minutos. En la tabla 15.1 se muestran las observaciones a los 0, 30, 60 y 90 minutos.

Una gráfica de perfiles revela la tendencia

En las figuras 15.1 se muestran las tendencias observadas con el tiempo para los tres tratamientos en el estudio, las gráficas de perfiles muestran un incremento en las diferencias observadas de temperatura para los conejos. Se observa un incremento menos definitivo con el tratamiento del excipiente solo y el perfil de la solución salina indica sólo una respuesta fluctuante sin tendencia definida.

El objetivo del análisis para este estudio será determinar si existe un tendencia significativa ascendente en la temperatura para cualquier tratamiento. De ser así, será importante determinar si el amiodarone es responsable del incremento y no la solución excipiente, o el procedimiento intravenoso en sí mediante el tratamiento salino. Así, serán importantes en el análisis los contrastes entre la tendencia del tratamiento de amiodarone y la de los dos tratamientos de control.

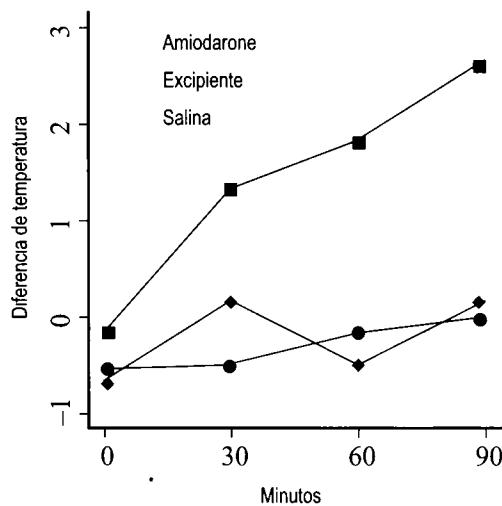
Diseños entre y dentro de unidades experimentales

Los diseños de mediciones repetidas se pueden describir en términos del diseño *entre sujetos* y el diseño *dentro sujetos*. Los diseños entre sujetos se refieren a los diseños en que una unidad experimental se asigna a un tratamiento. Los diseños dentro sujetos se refieren a las diferentes mediciones en cada unidad experimental. El diseño explicado en el ejemplo 15.1 tiene tres tratamientos en un diseño totalmente aleatorizado para el diseño entre sujetos. El diseño dentro sujetos consiste en las mediciones repetidas en cada conejo.

Tabla 15.1 Diferencias en las temperaturas ($^{\circ}\text{C}$), tratada menos no tratada, de las orejas de los conejos luego de 0, 30, 60 o 90 minutos después del tratamiento

Tratamiento	Conejo	Tiempo de observación (minutos)			
		0	30	60	90
Amiodarone	1	-0.3	-0.2	1.2	3.1
	2	-0.5	2.2	3.3	3.7
	3	-1.1	2.4	2.2	2.7
	4	1.0	1.7	2.1	2.5
	5	-0.3	0.8	0.6	0.9
	Media	-0.24	1.38	1.88	2.58
Excipiente	6	-1.1	-2.2	0.2	0.3
	7	-1.4	-0.2	-0.5	-0.1
	8	-0.1	-0.1	-0.5	-0.3
	9	-0.2	0.1	-0.2	0.4
	10	-0.1	-0.2	0.7	-0.3
	Media	-0.58	-0.52	-0.06	0.00
Salina	11	-1.8	0.2	0.1	0.6
	12	-0.5	0.0	1.0	0.5
	13	-1.0	-0.3	-2.1	0.6
	14	0.4	0.4	-0.7	-0.3
	15	-0.5	0.9	-0.4	-0.3
	Media	-0.68	0.24	-0.42	0.22

Fuente: G. Ward, Department of Pharmaceutical Sciences, University of Arizona.

**Figura 15.1** Gráfica de perfiles de las medias de cada tratamiento en cada periodo para el estudio de amiodarone

De manera alternativa, se pueden administrar dos tratamientos o más a cada sujeto. Suponiendo que se usan atletas como sujetos para un estudio de fisiología del ejercicio y después de completar el régimen de entrenamiento se les somete a pruebas en bandas caminadoras colocadas en dos posiciones distintas, horizontal e inclinada. El objetivo es determinar si existen diferencias entre los resultados de las pruebas con la caminadora horizontal y la inclinada. Si las pruebas en ellas se realizan con los atletas en orden aleatorio, cada atleta es un bloque en un diseño de bloques aleatorizado.

15.2 Relaciones entre las mediciones repetidas

Las relaciones entre las observaciones gobiernan los métodos estadísticos necesarios para el diseño de investigación específico usado en un estudio. En este capítulo se explora la correspondencia de las relaciones con el método de análisis para diseños de medidas repetidas junto con algunas estrategias útiles para el análisis.

Observaciones correlacionadas entre las mediciones repetidas

El orden de las mediciones a los 0, 30, 60 y 90 minutos en cada conejo en el ejemplo 15.1 no se puede aleatorizar con respecto al tiempo; así, es posible que los pares de mediciones repetidas en el mismo conejo tengan una correlación. En general, se supone que los pares de observaciones adyacentes en el tiempo tienen mayor correlación que los pares de observaciones más separadas. Y se supone que las observaciones a los 0 y 30 minutos en cualquiera de los conejos del estudio no tienen una correlación mayor que las observaciones a los 0 y 90 minutos.

La correlación entre dos variables, digamos y_1 y y_2 , se define como:

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \quad (15.1)$$

donde σ_1 y σ_2 son las desviaciones estándar de y_1 y y_2 , y σ_{12} es la covarianza entre y_1 y y_2 . Si el valor esperado o media de la variable y es $E(y) = \mu$, entonces su varianza es $\sigma^2 = E(y - \mu)^2$. La covarianza de dos variables, y_1 y y_2 , es $\sigma_{12} = E(y_1 - \mu_1)(y_2 - \mu_2)$. La covarianza es una medida de la manera en que dos variables varían juntas, si aumenta el valor de una variable mientras el de la otra aumenta, la covarianza es positiva y la correlación entre las variables es positiva. Las varianza y covarianzas teóricas para las mediciones repetidas sucesivas como y_1, y_2, y_3 y y_4 se explican en el cuadro 15.1 como la matriz Σ de 4×4 .

Suposiciones del análisis de varianza

Varianzas iguales para los grupos de tratamiento y observaciones independientes con distribución normal son las suposiciones usuales necesarias para un análisis de varianza válido de los datos. La independencia en las observaciones tiene como resultado un valor de cero para las covarianzas mostradas en el cuadro 15.1. Con

Cuadro 15.1 Matriz Σ de varianzas y covarianzas para cuatro mediciones repetidas

	y_1	y_2	y_3	y_4
y_1	σ_1^2	σ_{12}	σ_{13}	σ_{14}
y_2	σ_{21}	σ_2^2	σ_{23}	σ_{24}
y_3	σ_{31}	σ_{32}	σ_3^2	σ_{34}
y_4	σ_{41}	σ_{42}	σ_{43}	σ_4^2

estas suposiciones, σ^2 tiene el mismo valor para todos los grupos de tratamiento y tiempos de medición, y $\rho = 0$ o bien $\sigma_y = \sigma_j = 0$.

Las medias compuestas de simetría son iguales a la correlación entre las mediciones repetidas

Un experimento particular con asignación aleatoria de los tratamientos a las unidades experimentales es sólo una muestra aleatoria de todos los experimentos posibles que pudieran usarse. Hacer una asignación aleatoria no elimina la correlación entre las observaciones; sin embargo, la correlación esperada entre las unidades experimentales es constante con todas las aleatorizaciones posibles. Si varianzas y correlaciones son constantes, las covarianzas tendrán el valor constante $\sigma_y = \rho\sigma^2$ de la ecuación (15.1), esta condición se conoce como *simetría compuesta*. La matriz de varianzas y covarianzas con simetría compuesta se muestra en el cuadro 15.2.

Cuadro 15.2 Matriz Σ de varianzas y covarianzas para cuatro mediciones repetidas con simetría compuesta

	y_1	y_2	y_3	y_4
y_1	σ^2	$\rho\sigma^2$	$\rho\sigma^2$	$\rho\sigma^2$
y_2	$\rho\sigma^2$	σ^2	$\rho\sigma^2$	$\rho\sigma^2$
y_3	$\rho\sigma^2$	$\rho\sigma^2$	σ^2	$\rho\sigma^2$
y_4	$\rho\sigma^2$	$\rho\sigma^2$	$\rho\sigma^2$	σ^2

Los tratamientos de parcelas divididas se aleatorizan, las mediciones repetidas no

La suposición de simetría compuesta se usa para los errores de las observaciones en el experimento de parcelas divididas del capítulo 14 porque se hizo una asignación aleatoria de los tratamientos a las subparcelas. El sujeto y el factor de tratamiento entre sujetos del diseño de mediciones repetidas equivalen, respecti-

vamente, a la parcela completa y al factor de tratamiento para la parcela completa de un diseño de parcelas divididas, y las mediciones repetidas en un sujeto son análogas a la subparcela. La diferencia entre las observaciones de la subparcela y las mediciones repetidas es que en el diseño de parcelas divididas se hace una asignación aleatoria de los tratamientos a las subparcelas, mientras que no hay aleatorización en las mediciones repetidas. Si todas las mediciones repetidas de un sujeto tienen igual correlación, existe simetría compuesta y el diseño de mediciones repetidas se puede analizar como un diseño de parcelas divididas, empleando el tiempo de medición como factor de tratamiento de la subparcela. El análisis de varianza de parcelas divididas se explica en el capítulo 14 para varios diseños de experimentos entre sujetos, o parcelas completas.

La condición de Huynh-Feldt es menos estricta que la simetría completa

Huynh y Feldt (1979) mostraron que las condiciones necesarias para el análisis de varianza usual en los diseños de mediciones repetidas eran menos estrictas que la condición de simetría completa. Demostraron que la condición necesaria es tener la misma varianza de la diferencia para todos los pares posibles de observaciones tomadas en diferentes períodos, digamos, y_i y y_j , es decir:

$$\sigma_{(y_i - y_j)}^2 = 2\lambda \text{ para } i \neq j \quad (15.2)$$

para alguna $\lambda > 0$. Esta condición también se puede establecer como:

$$\sigma_y = \frac{1}{2}(\sigma_i^2 + \sigma_j^2) - \lambda \text{ para } i \neq j \quad (15.3)$$

La matriz de varianzas y covarianzas que satisface esta condición se conoce como **matriz tipo H**. Los cuadrados medios del análisis de varianza se pueden usar para probar las hipótesis sobre los tratamientos dentro sujetos si se cumple la condición de Huynh-Feldt.

Si cada sujeto recibe todos los tratamientos

La realidad de muchos estudios de investigación desde el punto de vista económico y de control del error experimental exige la obtención de más de una observación de cada unidad experimental. Por ejemplo, el considerable costo de mantener animales grandes hace necesario obtener toda la información posible sobre los tratamientos con cada animal. Además, la variabilidad de las observaciones entre los animales tiende a ser mucho mayor que entre observaciones múltiples en el mismo animal, por lo que el bloquizar los animales con los tratamientos administrados a cada animal aumenta la precisión de las comparaciones de tratamientos.

Cuando se administra cada tratamiento en orden aleatorio a cada sujeto, por ejemplo, B→D→A→C, los sujetos son bloques aleatorios en un diseño de bloques completo aleatorizado. En la tabla 15.2 se muestran los cuadrados medios esperados para el diseño de bloques completo aleatorizado, que es un análisis de modelo mixto con efectos de bloques aleatorios y de tratamiento fijos. El estadístico $F_0 = CMT/CME$, prueba la hipótesis nula de que no hay diferencias entre las medias de tratamiento.

Tabla 15.2 Cuadrados medios esperados para el análisis de varianza mixto de bloque completo aleatorizado

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrados medios esperados
Bloques	$r - 1$	MSB	$\sigma^2 + t\sigma_p^2$
Tratamientos	$t - 1$	MST	$\sigma^2 + r\theta_t^2$
Error	$(t - 1)(r - 1)$	MSE	σ^2

Transferencia de efectos

Los efectos de ciertos tipos de tratamiento se transfieren al siguiente periodo de tratamiento cuando se administran en secuencia a los sujetos. Esta trasferencias pueden sesgar severamente las estimaciones de las medias de tratamiento debido a que los tratamientos suministrados en periodos anteriores influyen en el efecto del tratamiento aplicado en periodos sucesivos.

Los efectos de transferencia son particularmente complicados en sujetos humanos y animales con un suministro sucesivo de tratamientos dietéticos o médicos que afectan su fisiología. Antes de suministrar un segundo tratamiento suele usarse un periodo de descanso, o “de lavado”, entre dos tratamientos sucesivos para eliminar los efectos del tratamiento más reciente. El tema del capítulo 16 es los diseños cruzados desarrollados para estos estudios.

15.3 Una prueba para la suposición Huynh-Feldt

Se puede usar el análisis de varianza univariado con cualquiera de los tres conjuntos de suposiciones sobre las mediciones repetidas presentadas en la sección 15.2, éstos eran independencia, simetría compuesta o condición Huynh-Feldt. La condición de Huynh-Feldt con la matriz tipo H para varianzas y covarianzas de las mediciones repetidas es la menos restrictiva de las tres suposiciones. Se pueden usar sencillos métodos univaridos para el análisis si se puede suponer que la condición Huynh-Feldt se cumple para las mediciones repetidas. La suposición de una matriz tipo H se puede evaluar con la prueba atribuida a Mauchly (1940), que se demuestra en esta sección.

La prueba Mauchly de la condición Huynh-Feldt para mediciones repetidas se explica con el estudio de amiodarone del ejemplo 15.1. Recuerde que el experimento tenía tres tratamientos asignados a las unidades experimentales en un diseño totalmente aleatorizado y las unidades experimentales se midieron en cuatro ocasiones sucesivas:

La prueba W de Mauchly (Mauchly, 1940), usada para la prueba de hipótesis de la forma tipo H para la Σ , se calcula en muchos paquetes de computadora que tienen programas para el análisis de diseño de mediciones repetidas. En el apéndice 15A.1 se encuentra una breve descripción de los detalles del estadístico de prueba.

El resultado de la prueba W de Mauchly para el estudio con amiodarone calculado por un programa computacional se muestra en la tabla 15.3. El estadístico de prueba Mauchly, $W = 0.852$, tiene una distribución aproximada a la variable chi-cuadrada con $v = p(p - 1)/2 - 1$ grados de libertad, donde p es el número de mediciones repetidas, $v = 4(3)/2 - 1 = 5$ para el estudio con amiodarone. El estadístico W no es significativo, con $P(\chi^2_5 > 1.72) = .886$, la probabilidad de la variable chi-cuadrada con 5 grados de libertad que excede el valor de chi-cuadrada de 1.72. Si el resultado de la prueba Mauchly es aceptable, entonces las pruebas F en el análisis de varianza univariado son válidas.

Tabla 15.3 Resultados de la prueba Mauchly para una matriz Σ tipo H

Prueba de esfericidad de Mauchly
$W = .852$
Aprox. chi-cuadrada = 1.72 con 5 g.l.
$P(\chi^2_5 > 1.72) = .886$

La prueba Mauchly para la matriz Σ de varianza-covarianza tiende a tener un potencial bajo a menos que los tamaños de muestras sean muy grandes y su capacidad para detectar alejamientos de la hipótesis nula no es muy buena a menos que los experimentos tengan un número grande de réplicas, Boik (1981) indica un potencial menor que .20 para algunos casos específicos de la prueba de Mauchly con tres grupos de tratamiento y hasta 12 sujetos por grupo. En consecuencia, no se recomienda confiar por completo en la prueba. Dada la incertidumbre asociada con la capacidad de la prueba de Mauchly para detectar alejamientos de las suposiciones del análisis de varianza, la decisión para usar el análisis de varianza univariado tendrá que basarse en la experiencia con las características específicas del material de investigación.

15.4 Análisis de varianza univariado para mediciones repetidas

Si se puede tener una seguridad razonable de que las suposiciones del análisis de varianza son válidas para las mediciones repetidas en cada conejo de los grupos de tratamiento, se pueden usar los cuadrados medios del análisis de varianza de parcelas divididas para probar las hipótesis sobre las medias de tratamiento y sus interacciones con el tiempo, pues los conejos en el estudio con amiodarone equivalen a las parcelas completas para los tres tratamientos intravenosos, y las mediciones repetidas en los conejos lo son a los tratamientos en las subparcelas.

Uso del modelo de parcelas divididas para el análisis

El modelo lineal para el experimento de parcelas divididas es:

$$y_{ijk} = \mu + \alpha_i + d_{ik} + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad (15.4)$$

$i = 1, 2, \dots, t \quad j = 1, 2, \dots, p \quad k = 1, 2, \dots, r$

donde μ es la media general, α_i es el efecto del i -ésimo tratamiento, d_{ik} es el error experimental aleatorio para los conejos dentro de tratamientos con varianza σ_d^2 , β_j es el efecto del j -ésimo tiempo, $(\alpha\beta)_{ij}$ es la interacción entre tratamientos y tiempo, y e_{ijk} es el error experimental aleatorio con distribución normal en las mediciones repetidas con varianza σ_e^2 . El análisis de varianza de las parcelas divididas para los datos del estudio con amiodarone se muestra en la tabla 15.4, (Este análisis se exhibió en el capítulo 14).

Tabla 15.4 Análisis de varianza de parcelas divididas para mediciones repetidas en el estudio con amiodarone, en un diseño totalmente aleatorizado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	59	93.28			
Tratamiento (A)	2	35.38	17.69	19.44	0.000
Error(1)	12	10.94	0.91		
Tiempo	3	16.08	5.36	9.24	0.000
$A \times$ Tiempo	6	10.06	1.68	2.90	0.021
Error(2)	36	20.82	0.58		

La prueba para la interacción entre tratamientos y tiempo, $F_0 = CM(A \times \text{tiempo})/CME(2) = 1.68/0.58 = 2.90$, es significativa con $Pr > F = .021$. La interacción significativa entre el tiempo y los tratamientos intravenosos indica que las respuestas de la temperatura de la oreja son diferentes entre los tres tratamientos.

Uso de los contrastes de regresión en las mediciones repetidas para estudiar las tendencias en el tiempo

La prueba global para la significancia de la interacción entre los tratamientos y el tiempo señala muy poco sobre las características específicas de la interacción, si ésta existe. La respuesta a los tratamientos individuales en el tiempo es una componente importante del análisis de mediciones repetidas. Debe investigarse la interacción como una diferencia entre tratamientos en la tendencia a través del tiempo.

Las tendencias observadas en el tiempo para los tres tratamientos en el estudio con amiodarone se encuentran en la figura 15.1; en ella, las gráficas de perfiles muestran un incremento en las diferencias de temperatura observadas para los conejos y se aprecian tendencias menos definitivas con los tratamientos del vehículo y salino.

Los tratamientos entre sujetos y el tiempo constituyen un diseño factorial con el tiempo como factor cuantitativo y los tratamientos de entre sujetos como factor cualitativo. Las particiones de regresión polinomial para un factor cualitativo en el

análisis de varianza se explica varias veces en los capítulos anteriores (vea el ejemplo 6.3).

Las particiones de sumas de cuadrados de regresión lineal, cuadrática y cúbica se pueden calcular con las particiones correspondientes a la interacción tratamiento \times tiempo. En la tabla 15.5 se muestran las particiones de sumas de cuadrados para el estudio con amiodarone y en el ejemplo 14.1 están las particiones de sumas de cuadrados de regresión polinomial.

Tabla 15.5 Análisis de varianza de parcelas divididas con contrastes polinomiales para mediciones repetidas del estudio con amiodarone

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Total	59	93.28			
Tratamiento (A)	2	35.38	17.69	19.44	0.000
Error(1)	12	10.94	0.91		
Tiempo (T)	3	16.08	5.36	9.24	0.000
T lineal	1	14.52	14.52	25.03	0.000
T cuadrática	1	0.60	0.60	1.03	0.315
T cúbica	1	0.96	0.96	1.66	0.205
$A \times T$	6	10.06	1.68	2.90	0.021
$A \times T$ lineal	2	7.80	3.90	6.72	0.003
$A \times T$ cuadrática	2	0.56	0.28	0.48	0.622
$A \times T$ cúbica	2	1.71	0.85	1.47	0.242
Error(2)	36	20.82	0.58		

Previamente se determinó que la interacción entre los tratamientos y el tiempo era significativa, pero la hipótesis probada con los cuadrados medios en la tabla 15.5 proporciona información más específica sobre la forma de interacción, que para tratamientos y la regresión lineal en el tiempo, $A \times T$ lineal, es significativa con $F_0 = 3.90/0.58 = 6.72$ y $Pr > F = .003$. Ni la regresión cúbica ni la cuadrática en el tiempo tienen una interacción significativa con el tiempo y entonces las tendencias lineales en el tiempo difieren entre los tratamientos.

Los contrastes lineales estimados para cada grupo de tratamiento y sus errores estándar indicarán de manera más específica la diferencia de las tendencias lineales entre los grupos de tratamiento. Los contrastes lineales, $a_{i(1)}$, se calculan para cada tratamiento, donde $a_{i(1)}$ es el contraste lineal para el i -ésimo tratamiento; por ejemplo, las medias para los grupos de tratamiento con amiodarone a los 0, 30, 60 y 90 minutos fueron -0.24, 1.38, 1.88 y 2.58 en la tabla 15.1 y el contraste lineal para el grupo de tratamiento con amiodarone es:

$$a_{1(1)} = \frac{\sum P_{1j} y_{1j}}{\sum P_{1j}^2} = \frac{(-3)(-0.24) + (-1)(1.38) + (1)(1.88) + (3)(2.58)}{(-3)^2 + (-1)^2 + 1^2 + 3^2} \\ = \frac{8.96}{20} = 0.45$$

con error estándar

$$s_{c_1} = \sqrt{\frac{CME(2)}{r(\sum P_{1j}^2)}} = \sqrt{\frac{0.58}{5(20)}} = 0.08$$

Los contrastes lineales para los tratamientos del excipiente y salino son $a_{2(1)} = 0.11$ y $a_{3(1)} = 0.10$, respectivamente, cada uno con error estándar $s_c = 0.08$.

El ICS del 95% para los contrastes lineales requiere una t de Bonferroni, $t_{0.05,3,36} = 2.51$. La t de Bonferroni tiene 36 grados de libertad, ya que los cuadrados medios del error, $CME(2) = 0.58$ de la tabla 15.5, tienen 36 grados de libertad. Los ICS del 95% se calculan como $a_{i(1)} \pm 2.51(0.08)$.

El ICS del 95% para el tratamiento con amiodarone es (0.25, 0.65), lo que indica un incremento lineal significativo de la temperatura en el tiempo con el tratamiento. Los intervalos respectivos para los tratamientos vehículo y salino son (-0.09, 0.31) y (-0.10, 0.30), que señalan que no hay cambio lineal significativo en la temperatura para ninguno de los dos.

Como ninguno de los dos tratamientos de control, excipiente y salino, tuvo un cambio significativo en las diferencias de temperatura, se puede concluir que el incremento significativo en la temperatura para el grupo de tratamiento con amiodarone estaba en función de la droga y no del vehículo o de la manipulación por la inyección intravenosa.

15.5 Análisis cuando no se cumplen las suposiciones del análisis univariado

Cuando las condiciones Huynh-Feldt no se satisfacen para las mediciones repetidas, los resultados del análisis univariado descritos en la sección 15.4 no son válidos, por lo que se sugieren varios análisis alternativos cuando el análisis de varianza usual no se puede usar.

Tres opciones para el análisis

El análisis multivariado es el método común más ventajoso; pero, en términos generales los métodos multivariados están fuera del objetivo de este libro y su uso directo no se considera en el texto. Una segunda alternativa hace ajustes conservadores al estadístico F_0 usual del análisis de varianza para aproximar mejor los niveles de significancia de las pruebas, alternativa que se explica en esta sección. Una tercera alternativa con características atractivas, también considerada en esta sección, usa los contrastes entre mediciones repetidas. El análisis de contrastes

usa las características del análisis multivariado aplicables de manera directa a las mediciones repetidas. Los resultados del estudio descrito en el ejemplo 15.2 son observaciones repetidas en las unidades experimentales para las cuales las suposiciones Huynh-Feldt no eran válidas.

Ejemplo 15.2 Humedad y actividad de microbios en el suelo

Un suelo agrícola productivo requiere cierto nivel de ventilación para mantener activo el crecimiento de la raíz de la planta y la actividad de los microbios de suelo. Un científico encontró que los niveles de oxigenación del suelo se habían afectado en los suelos fertilizados con lodos, ricos en nutrientes, de una planta de tratamiento de aguas negras. El nivel de ventilación del suelo se puede reducir por el alto contenido de agua del lodo agregado, a través de la compactación con maquinaria pesada usada para agregar los lodos e, irónicamente, por el aumento en la actividad de los microbios que ocurre al agregar los lodos con un alto contenido de materia orgánica.

Objetivo de la investigación: un objetivo de este estudio específico fue determinar los niveles de humedad en los cuales la ventilación se convierte en limitante para la actividad de microbios en el suelo.

Diseño del tratamiento: el estudio incluyó un tratamiento de control sin lodos fertilizantes y un contenido de agua de 0.24 kg agua/kg suelo; se usaron tres tratamientos con diferente contenido de humedad para los fertilizantes con lodos, cuyos niveles de humedad para el suelo fertilizado fueron 0.24, 0.26 y 0.28 kg agua/kg suelo.

Diseño del experimento: las muestras de suelo se asignaron al azar a los cuatro tratamientos en un diseño totalmente aleatorizado. Las muestras de suelo tratado se colocaron en contenedores sellados y se incubaron en condiciones favorables para la actividad de microbios. El suelo se compactó en los contenedores al grado experimentado en el campo.

La actividad microbiana, medida como el aumento de CO_2 , se usó como medida del nivel de oxigenación del suelo. La evolución $\text{CO}_2/\text{kg suelo/día}$ en cada contenedor se midió a los 2, 4, 6 y 8 días de iniciado el periodo de incubación. La actividad microbiana en cada muestra de suelo se registró como el porcentaje de incremento en el CO_2 producido por encima del nivel atmosférico. Los datos se muestran en la tabla 15.6.

La suposición del análisis multivariado no es válida

La prueba de Mauchly para determinar si se cumple la condición Huynh-Feldt para mediciones repetidas de CO_2 se muestra en la tabla 15.7. El estadístico de Mauchly es $W = 0.180$, y la aproximación de la distribución chi-cuadrada a W tiene $v = p(p - 1)/2 - 1 = 4(3)/2 - 1 = 5$ grados de libertad. El nivel de significancia para la prueba es $P(\chi^2_5 > 11.52) = .042$, y la hipótesis nula para la matriz tipo H se rechaza con un nivel de significancia de .05.

Tabla 15.6 Mediciones repetidas del aumento de CO₂ por la actividad de microbios en el suelo bajo distintas condiciones de humedad

(kg agua/kg suelo)	Humedad	Contenedor	% evolución CO ₂ /kg suelo/día			
			Día 2	Día 4	Día 6	Día 8
Control	1		0.22	0.56	0.66	0.89
	2		0.68	0.91	1.06	0.80
	3		0.68	0.45	0.72	0.89
	Media		0.53	0.64	0.81	0.86
0.24	4		2.53	2.70	2.10	1.50
	5		2.59	1.43	1.35	0.74
	6		0.56	1.37	1.87	1.21
	Media		1.89	1.83	1.77	1.15
0.26	7		0.22	0.22	0.20	0.11
	8		0.45	0.28	1.24	0.86
	9		0.22	0.33	0.34	0.20
	Media		0.30	0.28	0.59	0.39
0.28	10		0.22	0.80	0.80	0.37
	11		0.22	0.62	0.89	0.95
	12		0.22	0.56	0.69	0.63
	Media		0.22	0.66	0.79	0.65

Fuente: Dr. I. Pepper y J. Neilson, Department of Soil and Water Science, University of Arizona.

Tabla 15.7 Resultados de la prueba de Mauchly para la condición Huynh-Feldt en la medición de CO₂ en muestras de suelo

<i>Prueba de esfericidad de Mauchly</i>
$W = 0.180$
Aprox. chi-cuadrada = 11.52 con 5 g.l.
$P(\chi^2_5 > 11.52) = .042$

Possibles errores de inferencia arreglados

Dada una prueba significativa de Mauchly, se puede suponer que los resultados del análisis univariado de varianza no son válidos. Boik (1981) demostró que desviaciones muy pequeñas de la condición Huynh-Feldt afectan seriamente los errores tipo I y el poder de las pruebas *F* univariadas para el factor de mediciones repetidas en el caso del estudio, el día que se mide el CO₂. Una de las primeras modificaciones al análisis univariado de varianza era ajustar los valores de los estadísticos F_0 calculados en el análisis.

Ajustes para los estadísticos de prueba univariados

Si la condición Huynh-Feldt no se cumple para las mediciones repetidas, entonces el estadístico F_0 tiene sólo una distribución aproximada a la distribución F , con menos grados de libertad (Box, 1954a, 1954b). Greenhouse y Geisser (1959) sugirieron un ajuste $\bar{\epsilon}$, basado en el trabajo de Box, donde el numerador y el denominador de los grados de libertad para el estadístico F_0 se multiplican por $\bar{\epsilon}^2$. Se obtienen pruebas más conservadoras cuando se usa el ajuste $\bar{\epsilon}^2 \leq 1$ y la prueba requiere un valor mayor de F_0 para ser significativa. El cálculo del ajuste $\bar{\epsilon}^2$ se muestra en el apéndice 15A.2 y se calcula con la mayoría de los programas estadísticos para análisis de mediciones repetidas.

Huynh y Feldt (1976) sugirieron un ajuste menos conservador que el ajuste $\bar{\epsilon}^2$ de Greenhouse y Geisser. Huynh (1978) informó que el ajuste $\tilde{\epsilon}$ de Huynh-Feldt producía pruebas con probabilidades de errores tipo I más cercanas al valor seleccionado de α que las obtenidas con el ajuste de Greenhouse-Geisser. El ajuste Huynh-Feldt se calcula con la mayoría de los programas para análisis de mediciones repetidas y su cálculo se muestra en el apéndice 15A.2.

Los ajustes se aplican al estadístico F_0 del análisis de varianza de parcelas divididas usual para el factor de mediciones repetidas. No es necesario ajustar las pruebas sobre factores de tratamiento de parcelas completas, o entre sujetos, ya que los tratamientos se asignan al azar a las unidades experimentales. El análisis de varianza de parcelas divididas para mediciones repetidas se muestra en la tabla 15.8 con los tratamientos de humedad como factor de parcela completa y los días como factor de la subparcela.

Tabla 15.8 Análisis de varianza para las mediciones de CO₂ en muestras de suelo (programa SAS-GLM)

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F	G-G*	H-F†
Humedad	3	11.56	3.85	11.15	0.003		
Error(1)	8	2.77	0.35				
Día	3	0.49	0.16	1.22	0.324	0.317	0.324
Día 3 humedad	9	1.55	0.17	1.28	0.296	0.330	0.304
Error(2)	24	3.21	0.13				

*Nivel de significancia después del ajuste con la epsilon de Greenhouse-Geisser = 0.5245

†Nivel de significancia después del ajuste con la epsilon de Huynh-Feldt = 0.8755

El ajuste con $\bar{\epsilon}^2$ de Greenhouse-Geisser, mostrado al pie de la tabla 15.8, es $\bar{\epsilon}^2 = .5245$; y el ajuste con $\tilde{\epsilon}$ de Huynh-Feldt es $\tilde{\epsilon} = .8755$. El ajuste de Greenhouse-Geisser reduce los grados de libertad alrededor del 48%, mientras que el ajuste menos conservador de Huynh-Feldt reduce los grados de libertad cerca del 12%.

Los ajustes \overline{F} de Greenhouse-Geisser al numerador y denominador de los grados de libertad para la prueba de interacción entre los niveles de humedad y los días, día \times humedad, son $9(.5245) = 4.7$ y $24(.5245) = 12.6$, respectivamente. Con los grados de libertad ajustados, la probabilidad de exceder $F_0 = 1.28$ es $Pr > F = .330$, y el nivel de significancia de .296 sin el ajuste. El efecto neto del ajuste aumenta el valor P para el estadístico F_0 a una prueba más conservadora que tiene una probabilidad menor de rechazar la hipótesis nula. Con los ajustes menos conservadores de Huynh-Feldt se obtienen $9(.8755) = 7.9$ y $24(.8755) = 21$ grados de libertad para la prueba con $Pr > F = .304$.

Como la interacción entre los niveles de humedad y los días no fue significativa, se hacen ajustes similares a los grados de libertad de F_0 para una prueba de los efectos principales para los días. Los grados de libertad ajustados tienen sólo un ligero efecto en $Pr > F$, pero en cualquier caso los efectos principales para los días no son significativos.

El resultado de la prueba F para los efectos principales de los tratamientos de humedad es el F_0 usual para el análisis de varianza de los efectos principales de parcelas completas. El valor calculado para los tratamientos de humedad del suelo es $F_0 = 11.15$ con $Pr > F = .003$ y la hipótesis nula se rechaza. El aumento de CO₂ promedio difiere entre los tratamientos de humedad del suelo y la interacción no significativa indica que los niveles no cambian durante los ocho días de medición.

La pruebas entre las medias de tratamiento de humedad del suelo necesitan los errores estándar basados en los cuadrados medios del error equivalente de parcelas completas, $CME(1) = 0.35$, de la tabla 15.8. Por ejemplo, el error estándar de la diferencia entre dos medias es $\sqrt{2(0.35)/(3)(4)} = 0.24$, dados $r = 3$ réplicas y $p = 4$ mediciones repetidas para cada media de efecto principal.

Los contrastes en las mediciones repetidas proporcionan inferencias específicas

Las pruebas F basadas en los ajustes de Greenhouse-Geisser y de Huynh-Feldt están limitados a conclusiones globales sobre la igualdad de las medias de tratamiento, pero es común que las preguntas sobre consecuencias mayores involucren interesantes contrastes entre las medias de los tratamientos. Un análisis alternativo para mediciones repetidas se basa en los contrastes importantes entre las mediciones repetidas, este análisis proporciona el estadístico de prueba adecuado a partir de los métodos multivariados al mismo tiempo que usa el análisis univariado de varianza.

El análisis requiere el cálculo de un contraste entre las mediciones repetidas para cada unidad experimental. Los valores del contraste se usan como si fueran las observaciones originales en las unidades experimentales, y se utiliza para calcular un análisis univariado de varianza.

Contrastes de regresión polinomial para estudiar las tendencias en el tiempo

Los contrastes de regresión polinomial ortogonal son los estadísticos más útiles para investigar la tendencia en el tiempo. Por ejemplo, el contraste lineal para la evolución a través del tiempo de CO₂ en la humedad del suelo para cada muestra se calcula como:

$$z_{ij(1)} = (-3)(y_{ij1}) + (-1)(y_{ij2}) + (1)(y_{ij3}) + (3)(y_{ij4})$$

El modelo lineal para el contraste c es:

$$z_{ij(c)} = \mu_c + \tau_{i(c)} + e_{ij(c)} \quad (15.5)$$

donde $z_{ij(1)}$ es el valor del contraste para la j -ésima unidad experimental en el i -ésimo tratamiento, μ_c es la media general para el contraste, $\tau_{i(c)}$ es el efecto del tratamiento para el contraste y $e_{ij(c)}$ es el error experimental con distribución normal para el contraste con varianza σ_{ec}^2 .

Los contrastes de interés para el estudio actual, mostrados en la tabla 15.6, son los contrastes polinomiales lineal, cuadrático y cúbico de las observaciones repetidas en cada contenedor de suelo. Los valores de los contrastes para cada contenedor se muestran en la tabla 15.9. Por ejemplo, el contraste lineal para el primer contenedor en el tratamiento de control es $z_{11(1)} = (-3)(0.22) + (-1)(0.56) + (1)(0.66) + (3)(0.89) = 2.11$.

Tabla 15.9 Contrastess lineal, cuadrático y cúbico para las tres muestras de suelo de cada uno de los cuatro grupos de tratamiento en el estudio de aumento de CO₂

Tratamiento	Unidad	Contrastes		
		Lineal	Cuadrático	Cúbico
Control	1	2.11	-0.11	0.37
	2	0.51	-0.49	-0.33
	3	0.90	0.40	-0.60
	Media	1.17	-0.07	-0.19
0.24	4	-3.69	-0.77	0.77
	5	-5.63	0.55	-1.61
	6	2.45	-1.47	-0.85
	Media	-2.29	-0.56	-0.56
0.26	7	-0.35	-0.09	-0.05
	8	2.19	-0.21	-2.47
	9	-0.05	-0.25	-0.05
	Media	0.60	-0.18	-0.86
0.28	10	0.45	-1.01	0.15
	11	2.46	-0.34	-0.08
	12	1.36	-0.40	0.02
	Media	1.42	-0.58	0.03

Lineal(P_{1i}) (-3, -1, 1, 3); cuadrático(P_{2i}), (1, -1, -1, 1); cúbico(P_{3i}), (-1, 3, -3, 1)

Un análisis separado para cada contraste

Los análisis de varianza calculados para los contrastes lineal, cuadrático y cúbico se muestran en la tabla 15.10, siguiendo la forma usada en el ejemplo 15.1. Los contrastes polinomiales se calcularon para el factor día junto con sus contrastes de interacción con el factor humedad.

Tabla 15.10 Análisis de varianza para contrastes polinomiales el día de la medición de CO₂ en las muestras de suelo

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Día lineal	1	0.031	0.031	0.11	0.744
Día lineal × humedad	3	1.320	0.440	1.64	0.255
Error	8	2.143	0.268		
Día cuadrático	1	0.366	0.366	4.19	0.075
Día cuadrático × humedad	3	0.156	0.052	0.60	0.635
Error	8	0.698	0.087		
Día cúbico	1	0.093	0.093	2.02	0.193
Día cúbico × humedad	3	0.070	0.023	0.50	0.690
Error	8	0.369	0.046		

La diferencia principal del análisis en el ejemplo 15.1 es el cálculo de sumas de cuadrados del error separadas para cada contraste y su interacción con la humedad. Las estimaciones respectivas de la varianza del error experimental a partir de la fuente de variación «error» son $s_{e1}^2 = 0.268$, $s_{e2}^2 = 0.087$ y $s_{e3}^2 = 0.046$ para los contrastes lineal, cúbico y cuadrático. Muchos paquetes estadísticos de computadora pueden realizar este análisis para diseños de muestras repetidas.

Prueba para la interacción entre tratamientos y tendencias

La suma de cuadrados para una interacción entre un contraste y los tratamientos de humedad mide la variabilidad en el contraste de regresión entre los tratamientos de humedad; por lo tanto, mide la interacción entre el contraste y el tratamiento de humedad. El estadístico F_0 prueba la hipótesis $H_0: \tau_{t(c)} = 0$ para todos los tratamientos de humedad.

El estadístico de prueba para la no interacción entre el contraste lineal de tiempo y los tratamientos de humedad en el análisis de varianza para día lineal, $F_0 = CM(\text{día lineal} \times \text{humedad})/s_{e1}^2 = 0.440/0.268 = 1.64$ y con $Pr > F = .255$ se concluye que no hay interacción. De la misma manera, los estadísticos F_0 para las interacciones de día cuadrático y día cúbico con la humedad prueban las hipótesis nulas de no interacción entre los tratamientos de humedad y los contrastes cuadrático y cúbico.

Como ninguna interacción fue significativa, se pueden probar las hipótesis referentes a la existencia de cualquier tendencia en el aumento de CO₂ con el tiempo. Los cuadrados medios para los contrastes se usan para probar estas hipótesis sobre el efecto principal del tiempo.

El estadístico, $F_0 = CM(\text{día lineal}/s_{e1}^2) = 0.031/0.268 = 0.11$ prueba la hipótesis de que el contraste lineal es cero para todos los tratamientos, y la hipótesis no se rechaza con $Pr > F = .744$. De manera similar, los estadísticos F_0 para día cuadrático y día cúbico prueban las hipótesis de que los contrastes cuadrático y cúbico son iguales a cero para todos los tratamientos y ningún contraste es significativo.

De estas pruebas se puede concluir que no había una tendencia significativa en el aumento del CO₂ debido a la actividad microbiana durante los primeros ocho días del periodo de incubación, pero, el análisis de varianza (Tabla 15.8) indicó que los niveles promedio de CO₂ difieren entre todos los tratamientos.

El análisis de contrastes individuales es más conservador

El análisis de contrastes individuales no está libre de inconvenientes. Las pruebas basadas en el análisis mostrado en la tabla 15.10 son más conservadoras y menos poderosas que las basadas en el análisis de parcelas individuales normal de la tabla 15.8. Las pruebas F de la tabla 15.10 para los contrastes se basan en varianzas del error con 8 grados de libertad, mientras que las pruebas para los mismos contrastes a partir del análisis de parcelas divididas en la tabla 15.8 estaban basadas en una varianza del error con 24 grados de libertad. Incluso con los ajustes de Greenhouse-Geisser o de Huynh-Feldt en el análisis multivariado, las pruebas eran menos conservadoras que las del análisis de contrastes individuales.

Los contrastes pueden tener distintas varianzas de error experimental

El análisis basado en contrastes individuales señala un hecho común en el análisis de datos, debido al cual las varianzas del error asociadas con distintos contrastes puede ser muy diferentes. La disparidad en las varianzas del error para los contrastes polinomiales en el estudio de microbios en el suelo se puede ver en la tabla 15.10. Los cuadrados medios del error experimental para los contrastes lineal, cuadrático y cúbico fueron $s_{e1}^2 = 0.268$, $s_{e2}^2 = 0.087$ y $s_{e3}^2 = 0.046$. Entonces, existe una diferencia triple entre las varianzas del error de los contrastes lineal y cuadrático y casi de seis veces entre los contrastes lineal y cúbico. La varianza de error usada en el análisis univariado de la tabla 15.8, $CME = 0.13$, es la varianza promedio del error para los contrastes individuales y su valor global en el análisis de varianza de parcelas divididas puede conducir a conclusiones contradictorias. Por lo tanto, se recomienda evaluar la posible disparidad entre las varianzas del error para un grupo de contrastes en cualquier estudio.

15.6 Otros experimentos con propiedades de mediciones repetidas

En este capítulo se ha hecho una analogía entre las mediciones repetidas en una unidad experimental y las subparcelas de un experimento de parcelas divididas,

aunque las diferencian varias características. En general, las subparcelas son unidades experimentales distintas de menor tamaño que la parcela completa y los tratamientos se les asignan al azar a las distintas unidades experimentales. Estas dos propiedades de las subparcelas permiten hacer una suposición razonable de que la estructura de la varianza y covarianza de las observaciones es compatible con los requerimientos del análisis de varianza tradicional.

En un diseño de parcelas divididas tradicional, las subparcelas casi siempre representan variabilidad en el espacio, que es lo opuesto a la variabilidad en el tiempo asociada con las mediciones repetidas. Las relaciones de covarianza entre las observaciones en los tratamientos de la subparcelas serían similares a las de las mediciones repetidas en el tiempo si los tratamientos en la subunidad no se asignan al azar a las sup parcelas. La aleatorización no es posible en cierto tipo de experimentos de parcelas divididas que tienen una distribución en el espacio de las subunidades.

Mediciones repetidas en el espacio

Examinemos un estudio en el que un fisiólogo de animales ha planteado la hipótesis de que el tamaño del segmento de tejido afecta los resultados de un ensayo sobre concentración de tirosina. Se obtuvieron cuatro segmentos de tamaños diferentes del mismo músculo del diafragma de un solo animal y se realizó el ensayo de tirósina en cada segmento.

Los cuatro segmentos de diferente tamaño de un solo músculo de diafragma representan tratamientos dentro sujeto, donde los músculos del diafragma individuales son los sujetos. El diseño del experimento es un diseño de bloque completo con los músculos del diafragma como bloques completos y los tamaños del segmento de músculo como tratamientos.

Los cuatro segmentos se pueden tomar de lugares aleatorios del músculo del diafragma, pero no se conocen las relaciones entre los segmentos. Si existen correlaciones diferentes entre los segmentos de distintos tamaños, entonces las suposiciones usuales para el diseño de bloques completo aleatorizado no son apropiadas y debe considerarse un análisis de mediciones repetidas para los datos.

Diseños de tratamiento con gradiente

Los experimentos con diseños de tratamiento con gradiente para las subparcelas describen un segundo tipo de experimento de parcelas divididas sin asignación aleatoria de los tratamientos a las subparcelas. Los ejemplos incluyen experimentos de irrigación por aspersión para cultivos de agronomía usados para crear un diseño de tratamientos con gradiente. Por lo común, el objetivo de estos experimentos es asegurar las propiedades de tolerancia a la sequía de varios cultivos. Los cultivos se asignan al azar a las parcelas en el campo en un diseño de bloques completo aleatorizado. El diseño se ilustra en la figura 15.2 para un experimento con cinco cultivos en cada uno de los dos bloques completos.

Se coloca un sistema de irrigación por aspersión entre los dos bloques de parcelas de cultivo, los aspersores del sistema se pueden ajustar para emitir una cantidad grande de agua cerca de la línea del aspersor y menos cantidad en gradiente al

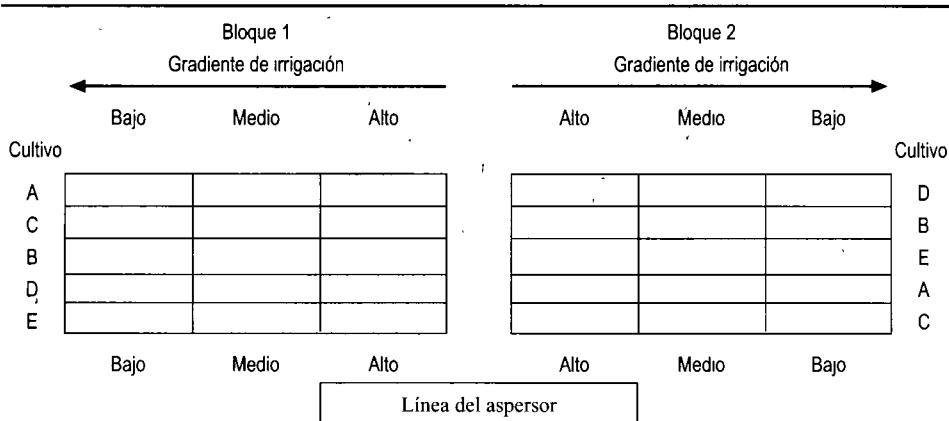


Figura 15.2 Ilustración de un diseño de tratamiento de gradiente para cultivos

alejarse de la línea. En consecuencia, se establece un tratamiento de gradiente de agua a lo largo de cada cultivo y se registran las cosechas y otras mediciones de las subparcelas establecidas en el gradiente de irrigación para cada parcela de cultivo.

En la figura 15.2 se muestran la subparcelas para cada cultivo de parcela completa. Los tratamientos de las subparcelas denominados “alto”, “medio” y “bajo” son las cantidades de agua aplicadas a las suparcelas, que no tiene una asignación aleatoria a las subparcelas.

Las suposiciones usuales del diseño de parcelas divididas quizá no sean adecuadas para las observaciones en las subparcelas al no haber aleatorización. Es más apropiado considerar las observaciones como mediciones repetidas y continuar con un análisis de datos mediante los métodos descritos en las secciones anteriores de este capítulo.

La prueba F para los efectos principales de los cultivos serán válidas en el análisis de varianza siempre que los cultivos se asignen al azar a las parcelas completas en cada bloque. La pregunta de investigación de interés inicial en estos estudios se relaciona con el desempeño diferencial de los cultivos a través del tratamiento con gradiente de agua. La prueba estadística de interacción cultivo \times nivel de agua en el análisis de subparcelas dentro sujetos tomará en cuenta esa hipótesis.

15.7 Otros modelos para la correlación entre mediciones repetidas

La estructura más general para la correlación entre mediciones repetidas se muestra en el cuadro 15.1, donde todas las varianzas y covarianzas, con $\sigma_{ij} = \sigma_{ji}$, tienen el potencial de ser únicas. Dos versiones basadas en las suposiciones que simplifican la estructura son la de simetría compuesta (cuadro 15.1) o las condiciones Huynh-Feldt para igualdad de varianzas entre todas las diferencias de las mediciones repetidas (ecuación 15.2). Bajo estas condiciones, una análisis de varianza de parcelas divididas directo proporcionará una inferencia válida.

Análisis de variables derivadas cuando las suposiciones del análisis de varianza no se cumplen

Los experimentos de laboratorio o de campo con una estructura estricta como el estudio con amiodarone del ejemplo 15.1 o el estudio de microbios en el suelo del ejemplo 15.2 suelen analizarse de manera eficaz mediante el análisis usado para el estudio de microbios en el suelo. Este tipo de análisis puede considerarse de **variables derivadas**, que en el ejemplo 15.2 eran los contrastes polinomiales ortogonales derivados de las mediciones repetidas en el tiempo.

Otras variables se pueden derivar de las observaciones repetidas en cada unidad, según la naturaleza del estudio. Por ejemplo, si las mediciones repetidas reflejan una respuesta de crecimiento o desgaste, puede ser razonable derivar variables basadas en un modelo de crecimiento o desgaste no lineal. En esos casos, las variables derivadas serían parámetros como pendiente, asymptota, punto de inflexión, entre otros.

La ventaja de este enfoque es su sencillez durante inferencia del análisis univariado de varianza normal. Una de sus desventajas es que las variables derivadas no necesariamente son independientes; por lo tanto, las inferencias para las variables derivadas no son independientes entre sí. El enfoque de variables derivadas se desglosa en estudios con datos incompletos debido a qué faltan valores en algunos puntos en el tiempo, o en estudios menos estructurados en los que los sujetos se pueden medir en distintos puntos del tiempo. En estas circunstancias, la suposición de una varianza común para los métodos de análisis de varianza normal ya no es válida.

Modelos con un número reducido de parámetros de correlación

Los modelos multivariados se pueden usar para la inferencia con la estructura de correlación general mostrada en el cuadro 15.1. Sin embargo, el modelo puede incluir un número innecesariamente grande de varianzas y covarianzas por estimar. Se han propuesto estructuras más sencillas con menos parámetros de covarianza que estimar, que tampoco son tan restrictivas como la simetría compuesta o la condición Huynh-Feldt y pueden ser más representativas del comportamiento de correlación.

Dos de estos modelos son el de **correlación serial** y el de **coeficientes aleatorios**. El modelo de correlación serial tiene errores correlacionados dentro de los sujetos o unidades definidas por la relación:

$$\rho_{(t_i - t_j)} = \frac{\sigma_{e_i e_j}}{\sigma^2}$$

donde $\rho_{(t_i - t_j)}$ es la correlación entre los errores en los tiempos i y j y $\sigma_{e_i e_j}$ es la covarianza entre los errores en los tiempos i y j en el mismo sujeto.

En el modelo de coeficientes aleatorios la correlación surge entre las mediciones repetidas como consecuencia de la suposición de que los efectos del tratamiento o los coeficientes de regresión varían a través de los sujetos o de las unidades experimentales. El ejemplo más sencillo es cuando la ordenada del perfil de respuesta en el tiempo varía entre las unidades, porque algunas de ellas tienen una respuesta baja intrínseca mientras la respuesta de otras es alta.

El modelo tiene efectos de tratamiento específicos del sujeto o de la unidad como resultado de esta suposición. Los modelos de coeficientes aleatorios son muy útiles cuando se desea hacer inferencias sobre individuos más que sobre poblaciones, ya que se pueden estimar los coeficientes del modelo para cada individuo con propósitos de predicción.

Los métodos de estimación estadística para las estructuras de correlación alternativas se basan principalmente en los métodos de la máxima verosimilitud y están fuera del alcance de este libro. Diggle, Liang y Zeger (1994) presentan con detalle los modelos y métodos de estimación disponibles para estudios con mediciones repetidas.

EJERCICIOS PARA EL CAPÍTULO 15

- Se realizó un estudio con sujetos humanos para medir los efectos de tres alimentos en los niveles de glucosa en el suero. Cada alimento se asignó al azar a los cuatro sujetos y se midió la glucosa en el suero de cada sujeto a los 15, 30 y 45 minutos después de ingerir el alimento. Los datos se muestran en la siguiente tabla:

Dieta	Sujeto	Tiempo (minutos)		
		15	30	45
1	1	28	34	32
	2	15	29	27
	3	12	33	28
	4	21	44	39
2	5	22	18	12
	6	23	22	10
	7	18	16	9
	8	25	24	15
3	9	31	30	39
	10	28	27	36
	11	24	26	36
	12	21	26	32

- Describa el estudio en términos de diseños entre sujetos y dentro sujetos.
- Calcule la media de las observaciones para cada dieta en cada tiempo de medición y elabore una gráfica de perfiles para cada tratamiento.
- Escriba el modelo lineal para un análisis de varianza de parcelas divididas, identifique los términos e indique las suposiciones necesarias para el análisis de los datos.
- Realice el análisis de varianza de parcelas divididas para los datos, pruebe las hipótesis necesarias y calcule las medias de tratamiento y sus errores estándar. ¿Cuáles son sus conclusiones?
- Obtenga las gráficas residuales del análisis de parcelas divididas e interprételas.
- Calcule las particiones de la suma de cuadrados para los contrastes lineal y cuadrático en el tiempo y sus interacciones con la dieta, pruebe las hipótesis nulas e interprete los resultados.

- g. Si dispone de un programa de computadora para el análisis de mediciones repetidas, pruebe la hipótesis de que se puede suponer la condición Huynh-Feldt para la matriz Σ de los errores experimentales para las varianzas y covarianzas dentro sujetos. ¿Cuál es su conclusión?
- h. Suponga que es necesario realizar un análisis de varianza separado para los contrastes de regresión lineal y cuadrática. Escriba el modelo lineal para uno de los contrastes, identifique los términos e indique las suposiciones necesarias para el modelo. Calcule el análisis de los contrastes e interprete los resultados. ¿Son comparables las varianzas del error experimental para los contrastes? ¿Difieren los resultados de los del inciso f)? Explique:
2. Un científico de suelos realizó un experimento para evaluar los efectos de compactación y humedad del suelo sobre la actividad microbiana. Se presentarán niveles reducidos de actividad microbiana en los suelos mal ventilados. Los niveles de ventilación pueden estar restringidos en suelos muy saturados o compactos. Las muestras de suelo tratadas se colocaron en contenedores sellados y se incubaron en condiciones propicias para la actividad microbiana. Ésta se midió en cada muestra de suelo como el porcentaje de incremento en el CO_2 producido por encima de los niveles atmosféricos.

El diseño fue un factorial de 3×3 con tres niveles de compactación del suelo (densidad del volumen = mg suelo/ m^3) y tres niveles de humedad del suelo (kg agua/kg suelo). Se prepararon dos réplicas de las unidades de contenedores de suelo para cada tratamiento.

Se registró la evolución $\text{CO}_2/\text{kg suelo/día}$, tres días sucesivos. Los datos para cada unidad son:

<i>Densidad</i>	<i>Humedad</i>	<i>Unidad</i>	<i>Día</i>		
			<i>1</i>	<i>2</i>	<i>3</i>
1.1	0.10	1	2.70	0.34	0.11
		2	2.90	1.57	1.25
		3	5.20	5.04	3.70
	0.20	4	3.60	3.92	2.69
		5	4.00	3.47	3.47
		6	4.10	3.47	2.46
	0.24	7	2.60	1.12	0.90
		8	2.20	0.78	0.34
		9	4.30	3.36	3.02
	1.4	10	3.90	2.91	2.35
		11	1.90	3.02	2.58
		12	3.00	3.81	2.69
1.6	0.10	13	2.00	0.67	0.22
		14	3.00	0.78	0.22
	0.20	15	3.80	2.80	2.02
		16	2.60	3.14	2.46
	0.24	17	1.30	2.69	2.46
		18	0.50	0.34	0.00

Fuente: Dr. I. Pepper y J. Neilson, Department of Soil and Water Science, University of Arizona.

- a. Describa el estudio en términos de diseños entre sujetos y dentro sujetos.

- b. Calcule la media de las observaciones para cada nivel de densidad de volumen y humedad del suelo en cada tiempo de medición y elabore una gráfica de perfiles de los resultados de cada tratamiento.
 - c. Escriba un modelo lineal para un análisis de varianza de parcelas divididas e indique las suposiciones necesarias para el análisis con este modelo.
 - d. Realice un análisis de parcelas divididas para los datos, pruebe las hipótesis necesarias y calcule las medias de tratamiento y sus errores estándar. ¿Cuáles son sus conclusiones?
 - e. Obtenga las gráficas residuales a partir del análisis de parcelas divididas e interpretelas.
 - f. Calcule las particiones de suma de cuadrados para los contrastes lineal, cuadrático y cúbico en el tiempo y sus interacciones con los tratamientos de densidad y humedad; pruebe las hipótesis nulas e interprete los resultados.
 - g. Si se dispone de un programa de computadora para análisis de mediciones repetidas, pruebe las hipótesis de que se puede suponer la condición Huynh-Feldt para la matriz Σ de los errores experimentales para las varianzas y covarianzas dentro sujetos. ¿Cuál es su conclusión?
 - h. Aplique el ajuste de Greenhouse-Geisser a las pruebas F del inciso d), ¿Difieren las conclusiones? Explique.
- i. Realice un análisis separado para los contrastes de regresión lineal y cuadrático. Escriba el modelo lineal para uno de los contrastes, identifique los términos e indique las suposiciones necesarias para el modelo. Calcule el análisis para los contrastes e interprete los resultados. ¿Son comparables las varianzas del error experimental para los contrastes?, ¿Difieren los resultados de los del inciso f)? Explique.
3. La tela de la ropa deportiva puede cambiar el estado de hidratación de la piel porque sirve como barrera a la disipación del agua generada por el cuerpo. Un científico de textiles realizó un estudio para evaluar el efecto de los tipos de fibra y el contenido de humedad de la tela sobre la pérdida de agua de la piel por evaporación.

Se usaron cinco hombres para el estudio, cada sujeto se usó para evaluar cinco tratamientos de tela; las telas de prueba eran telas de algodón y poliéster de uso común en la ropa deportiva y las pruebas se realizaron colocando una pieza de tela directamente sobre la piel del antebrazo del sujeto, se usó un instrumento para medir la cantidad de agua evaporada de la superficie de la piel del sujeto y las pruebas se realizaron en un entorno controlado a 70°F y 65% de humedad relativa.

Los tratamientos de telas fueron 1) algodón en equilibrio, 2) algodón saturado, 3) poliéster duro en equilibrio, 4) poliéster duro saturado y 5) poliéster suave saturado.

Dos de los tratamientos de la lista indican que la tela está en equilibrio, por lo que el contenido de humedad de las telas, se había dejado en equilibrio con el nivel de humedad del entorno antes de la prueba.

Se midió la pérdida de agua evaporada después de 60 minutos de aplicación de la tela a la piel. Los datos para todos los sujetos en cada tratamiento se muestran en la siguiente tabla:

Sujeto	Tratamiento				
	1	2	3	4	5
1	4.04	6.50	4.01	10.71	10.66
2	2.25	18.23	1.94	8.39	7.42
3	3.55	15.01	1.58	8.63	13.86
4	3.02	15.15	4.15	4.09	5.15
5	1.94	9.59	12.14	6.30	12.79

Fuente: Dr. K. Hatch, Family and Consumer Resources, University of Arizona.

- a. Describa el diseño.
- b. Escriba el modelo mixto del análisis de varianza para este estudio, identifique los términos y establezca las suposiciones necesarias para un análisis de los datos.
- c. El científico de textiles estaba interesado en cuatro contrastes específicos entre los tratamientos. Los contrastes de interés eran i) 1 contra 2, ii) 2 contra 4, iii) 2 contra 5 y iv) 3 contra 4. Escriba la tabla de contrastes, ¿son ortogonales? Explique.
- d. Realice un análisis de varianza de modelo mixto para los datos. Incluya un análisis de los contrastes enumerados en el inciso c). Interprete los resultados.
- e. Realice un análisis para cada contraste de la lista del inciso c). Escriba el modelo lineal para uno de los contrastes, identifique los términos e indique las suposiciones necesarias para el modelo. Calcule los análisis para los contrastes e interprete los resultados. ¿Son comparables las varianzas del error experimental?, ¿difieren las conclusiones de las del inciso d)? Explique.
4. Un agrónomo realizó una prueba de rendimiento con cinco cultivos de alfalfa en un diseño de bloques completo aleatorizado con tres réplicas. Cada parcela se cosechó cuatro veces en cada uno de dos años. En la tabla se muestra el rendimiento (lb/parcela) de las dos cosechas de cada año en ambas parcelas:

Cultivo	Bloque	1986		1987	
		Abril	Mayo	Abril	Mayo
1	1	20.4	23.2	14.8	22.9
	2	21.5	23.7	18.8	22.6
	3	21.1	23.4	14.3	22.1
2	1	19.1	22.4	14.5	19.2
	2	20.8	22.1	10.1	22.0
	3	20.5	23.5	12.0	21.5
3	1	19.3	22.1	14.5	19.5
	2	19.8	25.4	16.9	23.1
	3	20.5	24.8	16.7	20.1
4	1	23.2	25.6	14.9	19.5
	2	21.8	24.4	16.0	18.1
	3	22.2	26.8	16.7	21.0
5	1	21.4	24.5	14.1	21.3
	2	20.7	22.9	12.6	20.0
	3	18.7	21.8	14.3	21.4

Fuente: Dr. M. Ottoman, Department of Plant Science, University of Arizona.

- a. Describa el estudio en términos de diseño entre sujetos y dentro sujetos.
- b. Calcule la media de las observaciones para cada cultivo en cada cosecha de los dos años, y haga una gráfica de perfiles de los resultados de cada cultivo.
- c. Escriba un modelo lineal para un análisis de varianza de parcelas divididas, identifique los términos e indique las suposiciones necesarias para el análisis de los datos.

- d. Si dispone de un programa de computadora para análisis de mediciones repetidas, obtenga la matriz Σ de los errores experimentales para las varianzas y covarianzas dentro sujetos a partir de las cuatro mediciones en los meses y años. Pruebe la hipótesis de que se puede suponer la condición Huynh-Feldt para la matriz Σ . ¿Cuál es su conclusión?
- e. El agrónomo desea comparar las cosechas de los cultivos. La significancia de la interacción entre cultivos, años y meses de cosecha deben determinarse antes de que el agrónomo pueda comparar los cultivos promedio a través de los meses y años. Debe probarse la suposición de Huynh-Feldt para la matriz Σ de los errores experimentales para el análisis de mes \times cultivo y para el análisis de año \times mes \times cultivo si su programa tiene la capacidad. Realice las pruebas, calcule el análisis de parcelas divididas para los datos y pruebe las hipótesis para las interacciones de los cultivos con los meses y los años. Use los ajustes épsilon de G-G y H-F para las pruebas F si es necesario. ¿Cuáles son sus conclusiones de este análisis?
- f. Obtenga las gráficas residuales del análisis de parcelas divididas e interprételas.
5. Un agrónomo realizó un experimento para evaluar la tolerancia a la sequía de cultivos de cebada. Usó un sistema de riego por aspersión para crear un diseño de tratamiento con gradiente de agua en cada parcela. En la sección 15.6 se encuentra la descripción del diseño. Los cuatro cultivos se asignaron al azar a las parcelas en un diseño de bloques completo aleatorizado. La cantidad de agua aplicada a cada parcela disminuía al alejarse del aspersor. Se midió la cosecha de grano en subparcelas de 12 pies cuadrados en cada parcela de cultivo, a cuatro distancias del aspersor. Los datos (gramos de cebada por 12 pies cuadrados) para las subparcelas se muestran a continuación:

<i>Variedad</i>	<i>Bloque</i>	<i>Distancia del aspersor*</i>			
		<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
1	1	416.7	376.1	328.9	178.1
	2	490.2	513.7	438.4	348.1
	3	341.2	452.0	541.5	458.8
	2	644.7	555.4	587.8	413.7
	2	526.8	481.4	490.3	468.1
	3	540.6	504.3	495.9	523.3
3	1	388.9	491.8	355.0	222.2
	2	298.8	407.3	500.0	320.3
	3	386.7	388.4	492.4	438.2
4	1	512.0	598.9	442.1	186.0
	2	484.8	542.5	463.1	383.2
	3	368.5	547.8	702.9	445.3

*1 = más cercano al aspersor, 4 = más lejano al aspersor

Fuente: Dr. M. Ottoman, Department of Plant Sciences,
University of Arizona.

- a. Describa el estudio en términos de diseños entre sujetos y dentro sujetos.
- b. Calcule la media de las observaciones para cada cultivo en cada distancia y presente una gráfica de perfiles de los resultados.
- c. Escriba el modelo lineal para un análisis de varianza de parcelas divididas para los datos, identifique los términos e indique las suposiciones necesarias para el análisis de los datos.

- d. Realice un análisis de parcelas divididas para los datos, pruebe las hipótesis necesarias y calcule las medias de tratamiento y sus errores estándar. ¿Cuáles son sus conclusiones?
- e. Calcule las particiones de suma de cuadrados para los contrastes lineal, cuadrático y cúbico de la distancia del aspersor; pruebe la hipótesis nula e interprete los resultados.
- f. Si dispone de un programa de computadora para el análisis de mediciones repetidas, pruebe la hipótesis de que puede suponerse la condición de Huynh-Feldt para la matriz Σ de los errores experimentales para las varianzas y covarianzas de sujetos-dentro. ¿Cuál es su conclusión?
- g. Aplique los ajustes de Greenhouse-Geisser para las pruebas F del inciso d), ¿difieren las conclusiones? Explique.
- h. Realice un análisis separado de los contrastes de regresión lineal, cuadrático y cúbico. Escriba el modelo lineal para uno de los contrastes, identifique los términos e indique las suposiciones necesarias para el modelo. Calcule el análisis de los contrastes e interprete los resultados. ¿Son comparables las varianzas del error experimental para los contrastes? ¿Difieren los resultados de los del inciso e)? Explique.

Apéndice 15A.1 Prueba de esfericidad de Mauchly

La condición Huynh-Feldt para la matriz de varianzas y covarianzas de las p mediciones repetidas de los sujetos necesita $(p - 1)$ contrastes ortogonales normalizados para que las mediciones repetidas no se correlacionen y tengan varianzas iguales. Sean Σ la matriz de covarianzas de las mediciones repetidas y C una matriz de $(p - 1) \times p$, donde los renglones son contrastes ortogonales normalizados de las p mediciones repetidas. La condición Huynh-Feldt necesaria para la covarianza de los contrastes es $C\Sigma C' = \lambda I$, donde I es la matriz identidad y C' es la transpuesta de C . Si se satisface la condición, se dice que la matriz de covarianzas λI es esférica.

Sea s_{ij} el elemento en el i -ésimo renglón y la j -ésima columna de la matriz de covarianzas de $p \times p$ para los errores experimentales dentro sujetos S , con v grados de libertad. Se eligen $(p - 1)$ contrastes ortogonales normalizados en las p mediciones repetidas. Sea C la matriz de $(p - 1) \times p$, donde los renglones son contrastes ortogonales normalizados en las p mediciones repetidas. Calcule la matriz de $(p - 1) \times (p - 1)$, CSC' . El estadístico de prueba (Mauchly, 1940) para la hipótesis nula $H_0: C\Sigma C' = \lambda I$ es:

$$W = \frac{(p - 1)^{p - 1} | CSC' |}{(\text{tr} CSC')^{p - 1}} \quad (15A.1)$$

donde CSC' es la traza de la matriz. La traza de una matriz es la suma de sus elementos diagonales. El estadístico de prueba se pone en la escala adecuada para mejorar la exactitud de sus aproximaciones por la distribución chi-cuadrada. El factor de escala para la aproximación chi-cuadrada con $f = \frac{1}{2}p(p - 1) - 1$ grados de libertad es:

$$\gamma = v - \frac{2p^2 - 3p + 3}{6(p - 1)} \quad (15A.2)$$

La hipótesis nula se rechaza al nivel de significancia γ si $W > \chi^2_{\alpha, f}$.

Apéndice 15A.2 Ajuste de grados de libertad para el análisis de varianza de mediciones repetidas

La $\widehat{\epsilon}$ de Greenhouse-Geisser (Greenhouse y Geisser, 1959), y la $\tilde{\epsilon}$ de Huynh-Feldt (Huynh y Feldt, 1976) fueron propuestas como ajustes de los grados de libertad para las pruebas F en el análisis de factores de tratamiento dentro-sujetos. Los ajustes, basados en el trabajo de Box (1954a, 1954b) se desarrollaron para diseños con un sólo factor de tratamiento dentro sujetos. La $\tilde{\epsilon}$ de Huynh-Feldt es una función simple de la $\widehat{\epsilon}$ de Greenhouse-Geisser. Los cálculos se describen en los siguientes párrafos comenzando con los cálculos para $\widehat{\epsilon}$.

Sea s_{ij} el elemento en el i -ésimo renglón y la j -ésima columna de la matriz unida de covarianzas $p \times p$ para los errores experimentales dentro sujetos, S . Se eligen $q = (p - 1)$ contrastes ortogonales normalizados en las p mediciones repetidas. Sea C una matriz de $q \times p$, donde los renglones son contrastes ortogonales normalizados en las p mediciones repetidas. Calcule la matriz de $q \times q$, $A = CSC'$ con elementos $\{a_{ij}\}$. El ajuste $\widehat{\epsilon}$ de Greenhouse-Geisser se calcula con:

$$\widehat{\epsilon} = \frac{\left(\sum_{i=1}^q a_{ii} \right)^2}{q \sum_{i=1}^q \sum_{j=1}^q a_{ij}^2} \quad (15A.3)$$

El ajuste $\tilde{\epsilon}$ de Huynh-Feldt se calcula con:

$$\tilde{\epsilon} = \frac{(Nq\widehat{\epsilon} - 2)}{q(v - q\widehat{\epsilon})} \quad (15A.4)$$

donde N es el número de sujetos y v son los grados de libertad del error para el error experimental del análisis de varianza entre sujetos. Si el experimento se compone de N sujetos, cada uno con p tratamientos, entonces $v = N - 1$. Con r réplicas de los t tratamientos entre sujetos en un diseño de bloques completo aleatorizado, entonces $v = (t - 1)(r - 1)$.

Dado un estadístico F_0 dentro sujetos basado en cuadrados medios con v_1 y v_2 grados de libertad, los grados de libertad ajustados para la prueba son ϵv_1 y ϵv_2 , donde $\widehat{\epsilon}$ se usa para el ajuste Greenhouse-Geisser y $\tilde{\epsilon}$ se usa para el ajuste Huynh-Feldt. La $\tilde{\epsilon}$ de Huynh-Feldt no se usa si $\tilde{\epsilon} \geq 1$.

16 Diseños cruzados

El tema principal de este capítulo es el diseño y análisis de experimentos con distintos tratamientos, suministrados en períodos sucesivos a las unidades experimentales. El análisis de experimentos incluye una evaluación del efecto de un tratamiento que puede transferirse y afectar la respuesta de otro tratamiento en el siguiente período. En este capítulo se presentan diseños específicamente construidos para estimar con eficiencia los efectos directos y transferidos de los tratamientos.

16.1 Suministro de todos los tratamientos a cada unidad experimental

El **estudio cruzado** (“crossover”) describe experimentos con tratamientos suministrados en secuencia a cada unidad experimental: un tratamiento se administra a una unidad experimental en un período específico, después del cual se administra otro tratamiento a la misma unidad. Los tratamientos se aplican en forma sucesiva a la unidad hasta que ha recibido todos los tratamientos.

Los diseños cruzados pueden aumentar la precisión y reducir los costos

Cuando se comparan tratamientos en la misma unidad experimental, la variación entre-unidades se elimina del error experimental. Así, las unidades experimentales individuales se usan como bloques para disminuir el error experimental y aumentar la eficiencia del experimento. En general, las comparaciones de tratamientos en los diseños por bloques son más precisas que las de los diseños no bloquizados, porque la variación de la unidad experimental o bloque se elimina de las comparaciones entre tratamientos. Así, la ventaja principal de los estudios cruzados es el aumento en la precisión de las comparaciones de tratamientos.

Cuando se dispone de un número limitado de unidades, los diseños cruzados, que casi siempre se usan con sujetos animales o humanos ayudan a economizar recursos. El gasto de mantener animales grandes y la dificultad para reclutar sujetos humanos y lograr suficientes réplicas, hacen que los diseños cruzados sean más atractivos, ya que requieren menos unidades para una cantidad igual de réplicas de tratamiento. A continuación se presenta un experimento con diseño cruzado, a manera de ejemplo.

Ejemplo 16.1 Digestión de alimentos en ganado vacuno

Los efectos asociativos se presentan en las dietas de animales cuando se combinan alimentos y el resultado en el animal es diferente del pronosticado para la suma de los ingredientes individuales. Se ha demostrado que la adición de fibra a las dietas de los rumiantes influye en varios factores de aprovechamiento, como el tiempo de retención del rumiado. Sin embargo, la información sobre los efectos asociativos relativos de distintas fibras era escasa, en especial sobre dietas de alimentos mixtos.

Hipótesis de investigación: un científico de animales planteó la hipótesis de que la fuente de la fibra podía influir en el aprovechamiento de dietas mixtas en novillos, alterando la digestión rumiante de otros ingredientes de la dieta.

Diseño de tratamientos: la dieta mixta básica para los novillos era un 65% de concentrado basado en alimento hojuelas de sorgo y un 35% de fibra. Se usaron tres tratamientos de fibra con (A) 35% de alfalfa como tratamiento de control, (B) 17.5% de paja de trigo y 17.5 de alfalfa, y (C) 17.5% de cáscara de semilla de algodón y 17.5% de alfalfa.

Diseño del experimento: se disponía de 12 novillos para el estudio y se les suministró cada una de las tres dietas de fibra en una de las seis secuencias posibles, cada dieta en cada secuencia se dio a dos novillos durante 30 días. Antes de recopilar datos se estableció un periodo de 21 días para que los novillos se adaptaran al cambio de dieta.

En la tabla 16.1 se muestra el coeficiente de digestión de fibra detergente neutral (FDN) calculado para cada novillo con cada dieta. Este coeficiente indica el porcentaje de fibra digerida por el novillo.

Tabla 16.1 Coeficientes de digestión de FDN para dos novillos en cada secuencia de tres dietas de fibra en un diseño cruzado

Novillo:	Secuencia											
	1		2		3		4		5		6	
	1	2	3	4	5	6	7	8	9	10	11	12
Periodo I	(A)	50	55	(B)	44	51	(C)	35	41	(A)	54	58
Periodo II	(B)	61	63	(C)	42	45	(A)	55	56	(C)	48	51
Periodo III	(C)	53	57	(A)	57	59	(B)	47	50	(B)	51	54

Fuente: J. Moore, Department of Animal Science, University of Arizona.

Otros ejemplos: se prueba una droga nueva y una común, en un diseño cruzado con pacientes que sufren de asma bronquial agudo, para determinar si la nueva droga mejora su respiración más que la común. La nueva droga se administra a un paciente durante la primera semana y la común se da al mismo paciente durante la segunda semana.

Al utilizar un diseño cruzado con técnicos como operadores, se prueban tres diseños del tablero de control para un instrumento de laboratorio; el técnico opera el instrumento con cada diseño de tablero. Cada técnico prueba los tres diseños en días sucesivos, un tablero distinto cada día.

Diseño para evitar confundir los efectos del tiempo con el tratamiento

Una comparación entre dos tratamientos en la misma unidad experimental también es una comparación entre dos períodos, pues tanto tratamientos como períodos pueden contribuir a las diferencias observadas. Las pruebas cruzadas se diseñan para evitar confundir los efectos de períodos y tratamientos. Por ejemplo, un grupo de unidades experimentales recibirá la secuencia A→B, y un segundo grupo la secuencia B→A; de esta manera ambos tratamientos se administran en cada período, y las comparaciones entre tratamientos son independientes de las comparaciones entre períodos.

El efecto de acarreo puede persistir después del período de tratamiento

Una desventaja del diseño cruzado es la posibilidad de que un tratamiento suministrado en un período influya en la respuesta del siguiente período de tratamiento. Los efectos de un tratamiento que continúan durante el siguiente período de tratamiento son *efectos acarreados*.

En general, los tratamientos se administran en diseños cruzados durante el tiempo suficiente para permitir que su efecto se manifieste en los sujetos. Se retira el sujeto del tratamiento durante un período de descanso o de limpieza, entre dos tratamientos, para que recupere su estado fisiológico o psicológico original.

Por ejemplo, después de administrar una droga a un paciente, se trata de que el período de descanso permita eliminar o “limpiar” los residuos en su sistema, para que no estén presentes en los períodos sucesivos de tratamiento. Aunque se dé suficiente tiempo de descanso para que desaparezca la droga inicial del sistema, el estado psicológico puede haberse alterado lo suficiente para tener algún efecto en las respuestas de los posteriores períodos de tratamiento. En los estudios cruzados no se puede ignorar el potencial de un efecto acarreado.

Como se supone que los efectos acarreados están presentes en los estudios, los diseños construidos específicamente para medirlos se presentarán en este capítulo. Se han desarrollado varios diseños para estudios cruzados que cumplen con las necesidades específicas de distintos problemas de investigación. El material del capítulo está restringido a unos cuantos diseños cruzados elementales para explicar los principios básicos del diseño y análisis de estos estudios, a fin de obtener las metas específicas de la investigación.

Se usará el diseño cruzado básico del ejemplo 16.1 para explicar la relación entre los efectos acarreados de los tratamientos y los otros efectos presentes en el

experimento, además de proporcionar los fundamentos para el modelo estadístico básico y el análisis de estudios cruzados.

Diseño balanceado renglón-columna para el estudio de digestión

El diseño cruzado para las dietas de fibra es un diseño balanceado renglón-columna donde los períodos y los novillos son los renglones y las columnas del diseño. Cada dieta de fibra se suministra una vez a cada novillo y cuatro veces en cada periodo del diseño. En el capítulo 8, este diseño recibió el nombre de diseño de rectángulo latino.

Las seis secuencias de tratamientos de dieta se muestran en el cuadro 16.1. Las seis secuencias de dietas con fibra se asignaron al azar para los 12 novillos, pero no se aleatorizó el orden de suministro de las dietas de los mismos. Cada una de las seis secuencias de tres tratamientos debe presentarse con la misma frecuencia para evitar confundir los efectos del periodo con los del tratamiento y obtener un diseño *balanceado* para los efectos acarreados.

Cuadro 16.1 Diseño cruzado con seis secuencias de tratamiento para tres tratamientos en tres períodos

Periodo	Secuencia					
	1	2	3	4	5	6
I	A	B	C	A	B	C
	↓	↓	↓	↓	↓	↓
II	B	C	A	C	A	B
	↓	↓	↓	↓	↓	↓
III	C	A	B	B	C	A

Diseños para balancear los efectos acarreados

Un diseño cruzado está balanceado para los efectos acarreados cuando cada tratamiento sigue a cada uno de los otros tratamientos el mismo número de veces. Cada tratamiento se aplica con la misma frecuencia en cada periodo, y una vez en cada sujeto.

El diseño cruzado para el estudio de dietas con fibra del cuadro 16.1 es un diseño balanceado. La dieta A sigue a la dieta B dos veces, una vez en cada una de las secuencias 5 y 6; y sigue a la dieta C dos veces, una vez en cada secuencia 2 y 3. De la misma manera, las dietas B y C siguen a cada uno de los otros tratamientos dos veces. El balance se aplica sólo a los efectos acarreados de primer orden que alteran la respuesta en el primer periodo que sigue al suministro del tratamiento. Un efecto acarreado de segundo orden altera la respuesta en el segundo periodo que sigue al suministro del tratamiento.

El periodo de adaptación de 21 días, en el estudio de la dieta con fibra, es equivalente al periodo de descanso, o de limpieza usado en otros estudios para evitar transferencias. En ausencia de efectos acarreados, el efecto de cualquier dieta ante-

rior no se manifiesta en la fisiología digestiva del animal en el periodo actual, y las medidas de la digestión reflejan sólo los efectos directos de la dieta actual.

Como el tratamiento de dieta de control, con 35% de alfalfa quebrada era la dieta normal para los novillos, los tratamientos se diseñaron según la hipótesis de que fuentes de fibra alternas influyen en la utilización de las dietas mixtas durante la digestión del rumiante. Las dietas B y C proporcionan fibra (de la paja del trigo, o las cáscaras de semilla de algodón mezcladas con alfalfa). Las demás partes de la dieta, que incluyen grano y minerales, eran iguales en todas las dietas. Si la hipótesis de investigación es cierta, la digestión de la fibra dietética diferirá entre los tres tratamientos.

Las tres medias de tratamiento calculadas con los datos de la tabla 16.1 son $\bar{y}_A = 56.6$, $\bar{y}_B = 53.3$ y $\bar{y}_C = 47.2$. La media observada mide el efecto directo de un tratamiento, en los períodos en que estaba activo para los sujetos, más los efectos acarreados de los otros dos tratamientos. Los efectos directos de la dieta A no se confunden con los períodos o los novillos porque se midieron en cada período para cada novillo. Al no haber efectos acarreados $\bar{y}_A = 56.6$ es una estimación no sesgada de μ_A en el diseño balanceado. Como la dieta A sigue a la dieta B en dos secuencias y a la dieta C en otras dos secuencias, la respuesta a la dieta A puede aumentar o disminuir por los efectos acarreados de las dietas B o C administradas a los novillos en períodos anteriores.

16.2 Análisis de diseños cruzados

Modelo lineal para diseños cruzados

Un diseño cruzado tiene n grupos de secuencias de tratamientos y r_i sujetos en el i -ésimo grupo. Existen t tratamientos y cada grupo de sujetos recibe los tratamientos en orden diferente durante p períodos. Sea y_{ijk} la observación del j -ésimo sujeto de la i -ésima secuencia de tratamiento en el período k .

El modelo lineal para un diseño cruzado es:

$$y_{ijk} = \mu + \alpha_i + b_{ij} + \gamma_k + \tau_{d(i,k)} + \lambda_{c(i,k-1)} + e_{ijk} \quad (16.1)$$

$$i = 1, 2, \dots, n \quad j = 1, 2, \dots, r_i \quad k = 1, 2, \dots, p \quad d, c = 1, 2, \dots, t$$

donde μ es la media general, α_i es el efecto de la secuencia de tratamiento i , b_{ij} es el efecto aleatorio con varianza σ_b^2 para el j -ésimo sujeto en la i -ésima secuencia de tratamiento, γ_k es el efecto del período y e_{ijk} es el error aleatorio con varianza σ^2 para el sujeto en el período k . El efecto directo del tratamiento administrado en el período k del grupo de secuencias i es $\tau_{d(i,k)}$, y $\lambda_{c(i,k-1)}$ es el efecto acarreado del tratamiento administrado en el período $k - 1$ del grupo de secuencias i . El valor del efecto acarreado para la respuesta observada en el primer período es $\lambda_{c(i,0)} = 0$, ya que no existe acarreo.

Para simplificar, los respectivos efectos directos y acarreados de los tratamientos se identifican como $\tau_1, \tau_2, \dots, \tau_t$ y $\lambda_1, \lambda_2, \dots, \lambda_t$, respectivamente. Los valores esperados para las observaciones en la primera y segunda secuencias del estudio de dieta con fibra están en y_{ik} , que representa la secuencia i en el período k :

Secuencia 1 (A→B→C)	$E(y_{11}) = \mu + \alpha_1 + \gamma_1 + \tau_1$
	$E(y_{12}) = \mu + \alpha_1 + \gamma_2 + \tau_2 + \lambda_1$
	$E(y_{13}) = \mu + \alpha_1 + \gamma_3 + \tau_3 + \lambda_2$
Secuencia 2 (B→C→A)	$E(y_{21}) = \mu + \alpha_2 + \gamma_1 + \tau_2$
	$E(y_{22}) = \mu + \alpha_2 + \gamma_2 + \tau_3 + \lambda_2$
	$E(y_{23}) = \mu + \alpha_2 + \gamma_3 + \tau_1 + \lambda_3$

Los efectos acarreados sólo se presentan en las observaciones del segundo y tercer periodo. Por ejemplo, $\lambda_{c(1,0)} = 0$ en la secuencia 1 porque ningún tratamiento precede a la dieta A, pero las observaciones en los periodos 2 y 3 contienen los efectos acarreados de las dietas A y B, λ_1 y λ_2 , respectivamente. De la misma manera, en la secuencia 2 los efectos acarreados de las dietas B y C, λ_2 y λ_3 , se encuentran en las observaciones respectivas del segundo y tercer periodos.

La suposición del modelo univariado es satisfactoria

Las observaciones en cada unidad experimental son mediciones repetidas con el tiempo en condiciones diferentes y representan una observación multivariada en la unidad experimental. Las suposiciones necesarias para el análisis de varianza univariado se presentan en el capítulo 15, como “diseños con mediciones repetidas”. Kenward y Jones (1989) y Diggle *et al.* (1994) estudian los detalles de los modelos de mediciones repetidas para diseños cruzados que están más allá del alcance de este libro.

El análisis de varianza se puede usar para los diseños cruzados dependiendo de las relaciones entre las varianzas y covarianzas de los errores experimentales para las medidas repetidas. El análisis univariado se puede usar si ninguna de las suposiciones de independencia, simetría compuesta, o la condición Huynh-Feldt es adecuada para los errores experimentales.

La condición Huynh-Feldt de varianzas iguales para todas las diferencias posibles entre las medidas repetidas es la suposición menos restrictiva para los errores experimentales. Se supondrá que la condición Huynh-Feldt se satisface para los datos.

Análisis de varianza para diseños cruzados

Análisis sin efectos acarreados

Si el diseño cruzado es un diseño renglón-columna balanceado, se puede usar el análisis de varianza *en ausencia de efectos acarreados* descrito en el capítulo 8. Los sujetos y los períodos son los renglones y las columnas del diseño, y los efectos directos de tratamiento son ortogonales a renglones y columnas. Las sumas de cuadrados para renglones, columnas, tratamientos y error experimental, se pueden calcular con el análisis de varianza que se encuentra en la tabla 8.10.

Análisis con efectos acarreados

El análisis de varianza para el modelo con efectos de tratamiento acarreados de la ecuación (16.1) se describe en la tabla 16.2. La separación en las particiones de la suma

Tabla 16.2 Tabla de análisis de varianza para un diseño cruzado con n secuencias, p periodos, t tratamientos y r_i sujetos, en la i -ésima secuencia; $N = \sum_i^n r_i$

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Total	$Np - 1$	SC Total	
Entre-sujetos:			
Secuencia	$n - 1$	SCS	CMS
Secuencia sujetos-dentro	$(N - n)$	SCW	CMW
Dentro-sujeto:			
Periodo	$p - 1$	SCP	CMP
Tratamientos (directos)	$t - 1$	SCT	CMT
Tratamientos (transferidos)	$t - 1$	SCC	CMC
Error	$(N - 1)(p - 1) - 2(t - 1)$	SCE	CME

de cuadrados en grupos dentro-sujeto y entre-sujetos indica la correspondencia con las mediciones repetidas del análisis univariado de parcelas divididas, donde los sujetos son las parcelas completas y las mediciones repetidas durante p períodos son las subparcelas.

Los efectos de tratamiento y transferidos no ortogonales

Cuando en un diseño balanceado renglón-columna se presentan los efectos acarreados, éstos no son ortogonales a los directos ni a los bloques. En el cuadro 16.2 se observa la relación entre ambos efectos. Para los efectos acarreados cada periodo y acarreo son diseños de bloques completos, mientras que las secuencias forman un diseño de bloque incompleto balanceado, y cada par de ellos se presenta con dos de las secuencias. Los efectos de tratamiento directos y acarreados no son ortogonales debido a que no se presentan en todas sus combinaciones posibles, pues nunca están juntos para el mismo tratamiento en la misma observación.

Cuadro 16.2 Efectos τ_i y λ_i del modelo para el estudio de una dieta con alfalfa

Periodo	Secuencia					
	1	2	3	4	5	6
I	τ_1	τ_2	τ_3	τ_1	τ_2	τ_3
II	$\tau_2 + \lambda_1$	$\tau_3 + \lambda_2$	$\tau_1 + \lambda_3$	$\tau_3 + \lambda_1$	$\tau_1 + \lambda_2$	$\tau_2 + \lambda_3$
III	$\tau_3 + \lambda_2$	$\tau_1 + \lambda_3$	$\tau_2 + \lambda_1$	$\tau_2 + \lambda_3$	$\tau_3 + \lambda_1$	$\tau_1 + \lambda_2$

Antes de hacer inferencias sobre la diferencia entre los efectos directos del tratamiento, se debe determinar la significancia de los efectos acarreados; si aparecen en un estudio, es necesario ajustar la estimación de las diferencias entre las medias de tratamiento.

Sumas de cuadrados ajustadas para efectos no ortogonales

En el análisis de varianza, las sumas de cuadrados respectivas de los efectos directos y acarreados deben ajustarse con la otra.

Las sumas de cuadrados ajustadas se calculan a partir de las diferencias entre las sumas de cuadrados del error experimental de los modelos completo y reducido. La suma de cuadrados de los efectos acarreados ajustada para los efectos directos requiere de las sumas de cuadrados del error de los modelos:

$$\text{Modelo reducido} \quad y_{ijk} = \mu + \alpha_i + b_{ij} + \gamma_k + \tau_d + e_{ijk} \quad (16.2)$$

$$\text{Modelo completo} \quad y_{ijk} = \mu + \alpha_i + b_{ij} + \gamma_k + \tau_d + \lambda_c + e_{ijk}$$

y es $SCC = SCE_{r(\lambda)} - SCE_f$, donde $SCE_{r(\lambda)}$ es la suma de cuadrados del error calculada a partir del modelo reducido sin λ_c , y SCE_f la calculada del modelo completo.

La adición de cuadrados de los efectos directos ajustada para los efectos acarreados requiere por su parte de las sumas de cuadrados del error en los modelos:

$$\text{Modelo reducido} \quad y_{ijk} = \mu + \alpha_i + b_{ij} + \gamma_k + \lambda_c + e_{ijk} \quad (16.3)$$

$$\text{Modelo completo} \quad y_{ijk} = \mu + \alpha_i + b_{ij} + \gamma_k + \lambda_c + \tau_d + e_{ijk}$$

y es $SCT = SCE_{r(\tau)} - SCE_f$, donde $SCE_{r(\tau)}$ es la suma de cuadrados del error, calculada a partir del modelo reducido sin τ_d , y SCE_f calculada del modelo completo.

La codificación de datos para programas de cómputo exige una atención especial

Muchos programas de cómputo para estadística calculan las particiones de sumas de cuadrados necesarias para los diseños balanceado y no balanceados. Debe prestarse atención especial al archivo de codificación de datos para incluir los efectos acarreados en el modelo. Algunas estrategias alternativas para los efectos acarreados se pueden encontrar en Ratkowsky, Alldredge y Cotton (1990) y en Milliken y Johnson (1984). La codificación detallada para el archivo de datos del ejemplo 16.1 se muestra en el apéndice 16A.1, y las fórmulas para los cálculos manuales están disponibles para algunos diseños balanceados en Cochran y Cox (1975), Petersen (1985) y Gill (1978). Una breve descripción de las fórmulas críticas para las sumas de cuadrados de tratamientos ajustados para diseños balanceados aparece en el apéndice 16A.2.

Análisis de varianza para las dietas con fibra

El análisis de varianza para el estudio de dietas con fibra se muestra en la tabla 16.3. Los dos análisis describen el ajuste de los modelos completo y reducido para obtener las sumas de cuadrados ajustadas para los efectos de tratamiento directos y acarreados.

Tabla 16.3 Análisis de varianza para la digestión de FDN para el estudio de la dieta con fibra en un diseño cruzado

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Secuencia	5	331.67	66.33	3.48	.080
Novillos en secuencia	6	114.33	19.06		
Periodo	2	288.17	144.08	16.01	.000
Dieta	2	559.50	279.75*	31.09	.000
Acarreo	2	18.37	9.19†	1.02	.380
Error	18	161.96	9.00		

*Cuadrado medio de los efectos de tratamiento directos no ajustados para los efectos acarreados, *CMT*(no ajustada).

†Cuadrado medio de los efectos de tratamiento acarreados ajustados para los efectos directos, *CMC*(ajustada).

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr < F
Secuencia	5	331.67	66.33	3.48	.080
Novillos en secuencia	6	114.33	19.06		
Periodo	2	288.17	144.08	16.01	.000
Acarreo	2	129.60	64.80*	7.20	.005
Dieta	2	448.28	224.14†	24.91	.000
Error	18	161.96	9.00		

*Cuadrado medio de los efectos de tratamiento transferidos no ajustados para los efectos directos, *CMC*(no ajustada).

†Cuadrados medios de los efectos de tratamiento directos ajustados para los efectos acarreados, *CMT*(ajustada).

El primer análisis de la tabla 16.3 proporciona el cuadrado medio de los efectos acarreados ajustados para los efectos directos, *CM*(transf., aj.) = 9.19, y el segundo proporciona el cuadrado medio de los efectos directos ajustado para los efectos acarreados, *CM*(dieta, aj.) = 224.14. Las diferencias entre los cuadrados medios ajustados y no ajustados reflejan la relación no ortogonal entre los efectos de tratamientos acarreados y directos. Los cuadrados medios no ajustados para los efectos directos de tratamiento en el primer análisis es *CM*(dieta, no aj.) = 279.75, y los cuadrados medios no ajustados para los efectos acarreados en el segundo análisis es *CM*(acarreo, no aj) = 64.80.

Pruebas de hipótesis para los efectos de tratamiento acarreados

La hipótesis nula para los efectos acarreados es $H_0: \lambda_1 = \lambda_2 = \lambda_3 = 0$. Ésta es la hipótesis de interés inicial del estudio porque las inferencias sobre los efectos directos de tratamiento dependen de la presencia de los efectos acarreados. El estadístico de prueba para la hipótesis de que no hay efectos acarreados es

$$F_0 = \frac{CM(\text{acarreo, aj.)}}{CME} = \frac{9.19}{9.00} = 1.02$$

con 2 y 18 grados de libertad. El nivel de significancia para la prueba es $Pr > F = .380$, y la hipótesis nula no se rechaza. Los efectos acarreados no son significativos en el estudio de la dieta con fibra.

El estadístico $F_0 = CM(\text{dieta, aj.)}/CME = 224.14/9.00 = 24.90$ con 2 y 18 grados de libertad, prueba la hipótesis nula para los efectos directos de tratamiento. La hipótesis nula se rechaza con un nivel de significancia de $Pr > F = .000$. Existen diferencias significativas en la digestión de FDN entre las tres dietas con fibra.

La significancia de los efectos acarreados no ajustados para los efectos directos de tratamiento del segundo análisis de varianza en la tabla 16.3, de nuevo refleja la relación no ortogonal entre los efectos directos y acarreados. La correlación entre sus estimaciones conduce a dificultades para la distinción entre sus contribuciones separadas. La suma de cuadrados para los efectos acarreados no ajustados según los efectos directos:

$$Sc(\text{acarreo, no aj.)} = 129.60$$

es mucho más grande que la suma de cuadrados ajustada:

$$SC(\text{acarreo, aj.)} = 18.37$$

porque en la suma de cuadrados no ajustada los efectos directos están mezclados con los acarreados. En consecuencia, la significancia de los cuadrados medios no ajustados se debe a una combinación de la variación en los efectos directos y acarreados, y no es posible determinar qué conjunto de efectos contribuye a la significancia.

Si los efectos acarreados no son significativos, es común basar todas las inferencias en las medias de tratamiento no ajustadas según los efectos acarreados. En este caso, se usará el análisis para los diseños de cuadrado latino del capítulo 8. No obstante, un estudio detallado de Abeyasekera y Curnow (1984) sobre las consecuencias de ignorar los ajustes para los efectos acarreados, los llevó a recomendar ajustar siempre las medias de tratamiento para los efectos acarreados, sin importar la significancia de la prueba para efectos acarreados.

Interpretación de tratamientos con contrastes múltiples

De acuerdo con la hipótesis de investigación, la eficiencia de la digestión rumiante depende de la fuente de la fibra. La dieta normal o de control A incluía 35% de alfalfa quebrada. Las dietas alteradas tenían la mitad de la alfalfa y 17.5% de paja de trigo (dieta B) o 17.5% de cáscara de semilla de algodón (dieta C). La hipótesis nula de que no hay cambio en la digestión de FDN debido a las dietas alteradas se puede probar

mediante comparaciones de la dieta de control con las dietas alteradas usando el método de Dunnett. Las estimaciones de los contrastes entre las medias ajustadas de las dietas alteradas y la dieta A, $\hat{\mu}_A - \hat{\mu}_B$ y $\hat{\mu}_A - \hat{\mu}_C$, se muestran en la tabla 16.4, junto con los intervalos de confianza simultáneos del 95% para las comparaciones de control con los otros tratamientos, de acuerdo con el método de Dunnett.

Tabla 16.4 Estimaciones de contrastes entre la dieta de control y las alteradas para el estudio de dietas con fibra en un diseño cruzado

Contraste	Estimación	Error estándar	95% ICS (L, U)
Dieta A – Dieta B	4.06	1.37	(0.77, 7.35)
Dieta A – Dieta C	9.63	1.37	(6.34, 12.92)

El error estándar estimado de la diferencia entre dos medias ajustadas es 1.37 en la tercera columna de la tabla 16.4. El estadístico de dos lados de Dunnett para dos comparaciones es $d_{05,2,18} = 2.40$. El ICS del 95% indica que el reemplazo de la mitad de la alfalfa por cáscaras de semilla de algodón (C) en la dieta redujo la digestión de FDN en mayor grado que el reemplazo por paja de trigo (dieta B), ya que el límite inferior del intervalo para la dieta C está más lejos de 0 que el de la dieta B.

El valor crítico para una prueba de confianza de desigualdades con el método de Dunnett es $|\hat{\mu}_A - \hat{\mu}_i| > D(2, .05) = 1.37$ (2.40) = 3.28. La diferencia para la dieta A contra la dieta B, 4.06, y la diferencia de la dieta A contra la dieta C, 9.63, son significativas porque exceden el valor crítico. El reemplazo de la mitad de la alfalfa por cáscaras de semilla de algodón (dieta C) en la dieta reduce la digestión de FDN en un estimado de 9.63%, y el reemplazo por paja de trigo (dieta B) la reduce en un estimado de 4.06%.

16.3 Diseños balanceados para estudios cruzados

Las consideraciones prácticas determinan si un estudio cruzado es adecuado para el problema de investigación. Los diseños son más eficaces cuando los tratamientos manifiestan un efecto en el sujeto dentro de un periodo razonablemente corto para proporcionar un lapso de estudio de duración manejable. Una extensa publicación sobre los estudios cruzados se puede encontrar en Jones y Kenward (1989) y en Ratkowsky, Evans y Alldredge (1993).

Cuándo pueden tener éxito los diseños cruzados

Los estudios que miden las respuestas a los cambios en las dietas de animales se han realizado con éxito empleando diseños cruzados. Además, estos diseños reducen los costos de mantener las mayores cantidades de animales necesarias para proporcionar réplicas de tratamientos equivalentes con otros diseños. En especial, el ahorro en costos es pertinente para eliminar la variabilidad de sujeto a sujeto en las respuestas de estrés o ansiedad en seres humanos. En experimentos diseñados

para comparar sistemas diferentes de ensamble de componentes u operación instrumental con cada sujeto; la diferente tensión fisiológica o psicológica se mide con facilidad en cada sujeto.

Cuándo los diseños cruzados no tienen éxito

Los diseños no deben usarse en pruebas clínicas médicas para condiciones agudas como el dolor postoperatorio. Si un tratamiento en la secuencia cura la condición aguda, no queda nada que tratar en los períodos sucesivos. Los diseños se pueden usar de manera eficaz para estudiar el tratamiento de condiciones persistentes, como la artritis, cuando los administrados alivian los síntomas, en lugar de curarlos.

Las consideraciones prácticas influyen en la elección del diseño

Existen otras consideraciones prácticas que influyen en la selección de si un diseño de estudio cruzado es factible. La posibilidad de perder sujetos en el curso de un estudio aumenta con el número de períodos. Los diseños más deseables, los balanceados para los efectos acarreados, aumentan la complejidad cuando se añaden tratamientos al estudio. En consecuencia, el orden correcto de suministro de tratamientos a los sujetos se hace más difícil de manejar en estudios grandes. La capacidad de interpretación de resultados puede disminuir con la complejidad del diseño.

Se han desarrollado muchos diseños diferentes para estudios cruzados, a fin de cumplir con los requerimientos específicos del problema de investigación práctica o para contar con buenas propiedades estadísticas. Tres categorías de diseños, introducidas en esta sección, son útiles para varios tipos de estudios de investigación que pueden usar de manera adecuada los estudios cruzados, pero están restringidas a diseños balanceados para tres tratamientos o más y tres períodos o más. Como los diseños de dos períodos no son eficientes para estimar las medias de tratamiento y los diseños de dos tratamientos requieren una organización especial de las secuencias de tratamientos para estimar todos los efectos de interés en el estudio, los diseños de dos tratamientos se estudian en otra sección.

Las categorías de diseños presentadas en el resto de esta sección se basan en la relación entre el número de tratamientos t y el número de períodos p ; son diseños con $p = t$, $p = t + 1$ y $p < t$.

Comparación entre la eficiencia de los diseños cruzados y los cuadrados latinos

La eficiencia de un diseño para la estimación estadística de los efectos directos y acarreados es una consideración importante en la elección del diseño. La varianza de la diferencia entre dos efectos de tratamiento, σ_d^2 , forma la base de las medidas de eficiencia. El diseño de cuadrado latino, en ausencia de efectos acarreados, proporciona la varianza mínima, $2\sigma^2/r$, y es la norma de comparación. La razón de la varianza del cuadrado latino entre el diseño cruzado mide la eficiencia relativa del diseño cruzado como:

$$RE = \frac{2\sigma^2/r}{\sigma_d^2} \times 100 \quad (16.4)$$

La medida de eficiencia relativa se aplica a efectos cruzados o directos. La varianza de la diferencia, σ_d^2 , pertenece a cualquier efecto de tratamiento ajustado para el otro.

Los diseños más convenientes tienen un balance de varianza donde la varianza de la diferencia entre dos efectos de tratamiento, directo o acarreado, es la misma para todos los pares de tratamientos. El balance definido por la relación entre los efectos directos y acarreados es la otra característica conveniente en un diseño. El diseño está balanceado si el efecto directo de cada tratamiento se asocia con una frecuencia igual al efecto acarreado de primer orden de cada uno de los otros tratamientos. Patterson y Lucas (1962) presentan algunas fórmulas generales para calcular las eficiencias de diseños balanceados.

Diseños para igual número de periodos y tratamientos, $p = t$

Diseños de Williams construidos a partir de cuadrados latinos

Williams (1949) presentó métodos para construir diseños cruzados balanceados a partir de arreglos de cuadrados latinos. Si el número de tratamientos es par, $t = 4, 6, \dots$, se puede construir un diseño balanceado en un cuadrado latino. Si el número de tratamientos es impar, $t = 3, 5, 7, \dots$ se requieren dos cuadrados latinos particulares.

Se usaron dos cuadrados latinos de 3×3 , que producen seis secuencias de tratamiento, para el estudio de dietas con fibra del ejemplo 16.1. El diseño se reproduce en el cuadro 16.3 y muestra dos cuadrados latinos de 3×3 . Se utilizó un cuadrado latino para las tres primeras secuencias de tratamiento en el experimento, mientras que el segundo cuadrado se usó para las tres últimas.

Cuadro 16.3 Diseño cruzado balanceado con dos cuadrados latinos de 3×3

Periodo	<i>Cuadro 1</i>			<i>Cuadro 2</i>		
	1	2	3	4	5	6
I	A	B	C	A	B	C
	↓	↓	↓	↓	↓	↓
II	B	C	A	C	A	B
	↓	↓	↓	↓	↓	↓
III	C	A	B	B	C	A
				$E_d = 80.0$	$E_c = 44.44$	

En el cuadro 16.4 se muestra el diseño de Williams para $t = 4, 5$ y 6 tratamientos y balanceado para los efectos acarreados. En los diseños con cuadrado latino para un número par de tratamientos, cada tratamiento sigue una vez a cada uno de los demás y los efectos directos se asocian una vez con los efectos acarreados de primer orden, para todos los tratamientos. En los diseños con un número impar de

tratamientos, cada tratamiento sigue a los otros dos veces y los efectos directos de cada tratamiento se asocian dos veces con el efecto acarreado de primer orden de todos los demás tratamientos. La eficiencia relativa de los diseños de Williams para los efectos directos E_d y la eficiencia relativa E_c para los efectos acarreados se muestra abajo de cada diseño en los cuadros 16.3 y 16.4.

Cuadro 16.4 Diseño cruzado de cuadrado latino balanceado para cuatro, cinco y seis tratamientos

Grupos de secuencias con cuatro tratamientos					Grupos de secuencias con seis tratamientos						
Periodo	1	2	3	4	Periodo	1	2	3	4	5	6
I	A	B	C	D	I	A	B	C	D	E	F
II	D	A	B	C	II	C	D	E	F	A	B
III	B	C	D	A	III	B	C	D	E	F	A
IV	C	D	A	B	IV	E	F	A	B	C	D
$E_d = 90.91$		$E_c = 62.50$			V	F	A	B	C	D	E
					VI	D	E	F	A	B	C
						$E_d = 96.55$		$E_c = 77.78$			

Cinco tratamientos

Periodo	Grupo de secuencias 1					Grupo de secuencias 2				
	1	2	3	4	5	6	7	8	9	10
I	A	B	C	D	E	A	B	C	D	E
II	B	C	D	E	A	C	D	E	A	B
III	D	E	A	B	C	B	C	D	E	A
IV	E	A	B	C	D	E	A	B	C	D
V	C	D	E	A	B	D	E	A	B	C
	$E_d = 94.74$					$E_c = 72.00$				

Diseños a partir de conjuntos ortogonales de cuadrados latinos

Un conjunto completo de cuadrados latinos ortogonales proporciona un diseño cruzado balanceado para todos los órdenes de efectos acarreados. Un conjunto ortogonal de cuadrado latino para t tratamientos, requiere $t - 1$ cuadrados. Dos cuadrados latinos son ortogonales cuando ambos se superponen y cada tratamiento de un cuadrado aparece una vez con cada tratamiento del otro cuadrado.

Se usaron dos cuadrados latinos ortogonales de 3×3 para el estudio de dietas con fibra y forman el diseño de Williams mostrado en el cuadro 16.3.

Observe que cuando el cuadrado 1 se superpone al cuadrado 2, la etiqueta del tratamiento A en el cuadrado 1 está una vez con cada una de las etiquetas A, B y C.

del cuadrado 2. El mismo patrón es cierto para las etiquetas B y C del cuadrado 1, cuando se superpone en el cuadrado 2.

El conjunto completo de cuadrados ortogonales requiere $t(t - 1)$ sujetos para cualquier número de tratamientos. Por ejemplo, el conjunto ortogonal de tres cuadrados latinos para cuatro tratamientos requiere 12 grupos de sujetos.

La eficiencia del diseño para estimar los efectos directos y acarreados a partir de un conjunto completo de cuadrados ortogonales es la misma que para los diseños de Williams. El diseño de Williams requiere sólo $t - 1$ sujetos para t par y $2(t - 1)$ sujetos para t impar, por lo que la única ventaja real de un conjunto completo ortogonal de diseños de cuadrados latinos sobre los diseños de Williams es que el conjunto completo es balanceado para los efectos acarreados de todos los órdenes hasta $(p - 1)$. Por ejemplo, el diseño con cuatro tratamientos está balanceado para efectos acarreados de primero, segundo y tercer orden.

Cuando sólo están presentes en el estudio efectos acarreados de primer orden, los diseños de Williams tienen ahorros evidentes en el número de secuencias de tratamiento para más de tres tratamientos.

Diseños extendidos para efectos ortogonales directos y acarreados, $p = t + 1$

Los efectos directos y acarreados no son ortogonales en diseños en los que el número de períodos es igual al número de tratamientos. Lucas (1957) señaló que la no ortogonalidad de los diseños balanceados tiene como resultado una eficiencia mucho menor para estimar los efectos acarreados, que para estimar los efectos directos de tratamientos. Las diferencias en la eficiencia de estimación para ambos efectos se observan en el cuadro 16.4. La extensión de un período extra puede eliminar la no ortogonalidad y proporcionar diseños con estimaciones independientes de los efectos acarreados y directos que tienen una precisión más parecida en las estimaciones. Lucas (1957) proporcionó la primera descripción formal de los diseños extendidos.

El diseño extendido de período extra más sencillo se deriva de un diseño cruzado balanceado repitiendo en el período $(p + 1)$ el tratamiento administrado al sujeto en el período p . En el cuadro 16.5 se muestra un diseño de período extra para $t = 4$ tratamientos en $p = 5$ períodos, derivado del diseño de Williams para cuatro tratamientos.

Cada tratamiento administrado a un sujeto en el período IV se repite en el mismo sujeto en el período V, y está precedido por cada uno de los otros tratamientos en el diseño, incluyendo él mismo. Por ejemplo, el tratamiento A está precedido por sí mismo en la secuencia 3, por B en la secuencia 2, por C en la secuencia 4 y por D en la secuencia 3.

El diseño de período extra de Williams está totalmente balanceado para los efectos acarreados, ya que cada uno se presenta igual número de veces en cada secuencia y con cada efecto directo. Así, las estimaciones de los efectos directos son las mismas, haya o no efectos acarreados en el modelo y viceversa. La adición de un período extra desbalancea al diseño con respecto a las secuencias y los efectos directos debido a que los tratamientos aparecen un número desigual de veces dentro de cada secuencia. Observe en el cuadro 16.5 que uno de los tratamientos aparece dos veces en una secuencia, mientras que los otros aparecen sólo una vez en la misma.

Las eficiencias de los conjuntos ortogonales de cuadrados latinos con un período extra son las mismas que para los diseños de Williams con un período extra.

Cuadro 16.5 Diseño cruzado de periodo extra para cuatro tratamientos

Periodo	Grupo de secuencias			
	1	2	3	4
I	A ↓	B ↓	C ↓	D ↓
II	D ↓	A ↓	B ↓	C ↓
III	B ↓	C ↓	D ↓	A ↓
IV	C ↓	D ↓	A ↓	B ↓
V	C	D	A	B

$E_d = 96.00$ $E_c = 80.00$

La ventaja de aumentar la precisión al estimar los efectos acarreados debe ponderarse contra la desventaja de extender la duración y costo del experimento. Los diseños sólo deben usarse si existe una sospecha fuerte de que se pueden presentar efectos acarreados en el estudio. Pueden obtenerse algunos sesgos en las respuestas, si los sujetos o los investigadores, están conscientes de que los dos últimos períodos de medición involucran el mismo tratamiento.

Diseños para un ciclo de tratamiento incompleto, $p < t$

En los diseños balanceados presentados en las secciones anteriores, cada sujeto recibe todos los tratamientos, pero consideraciones prácticas o éticas pueden evitar el suministro del ciclo completo de tratamientos a cada sujeto. Puede ser más práctico administrar sólo parte de los tratamientos a cada sujeto y reducir las posibilidades de que se retire del estudio; también algunas consideraciones éticas pueden evitar que el investigador administre demasiados tratamientos a un paciente en una prueba clínica.

Patterson (1951, 1952) describe métodos para construir diseños balanceados para estudios que requieren que el número de períodos sea menor que el número de tratamientos, $p < t$. Los sujetos son bloques incompletos en el diseño porque no reciben todos los tratamientos. Los diseños de bloques incompletos se pueden obtener si se elimina uno o más períodos del conjunto completo de cuadrados latinos ortogonales. Los diseños incompletos también se pueden derivar de cuadrados latinos incompletos o de los cuadrados de Youden presentados en el capítulo 9 para los diseños renglón-columna. A manera de explicación, en el cuadro 16.6 se muestra un diseño para siete tratamientos y cuatro períodos.

Cuadro 16.6 Diseño incompleto para siete tratamientos y cuatro períodos

Periodo	Secuencia						
	1	2	3	4	5	6	7
1	A	B	C	D	E	F	G
2	B	C	D	E	F	G	A
3	D	E	F	G	A	B	C
4	G	A	B	C	D	E	F

Periodo	Secuencia						
	8	9	10	11	12	13	14
1	A	B	C	D	E	F	G
2	G	A	B	C	D	E	F
3	E	F	G	A	B	C	D
4	B	C	D	E	F	G	A

Petersen (1985) y Patterson y Lucas (1985) presentan tablas de diseños de cuadrados latinos incompletos para estudios cruzados. Jones y Kenward (1989) proporcionan los diseños de cuadrados latinos incompletos más eficientes para $t = 3, 4, 5, 6$ y 7 tratamientos.

16.4 Diseños cruzados para dos tratamientos

Quizá el diseño que más se usa en los estudios cruzados es el de dos tratamientos, dos secuencias y dos períodos. En ellos, un grupo recibe la secuencia de tratamiento A→B y el otro la secuencia B→A, como se muestra en el cuadro 16.7. Cada uno de los tratamientos aparece en cada secuencia y en cada periodo del arreglo de cuadrado latino de 2×2 . Así, los efectos de los tratamientos no se confunden con los efectos de las secuencias o los períodos, que son los renglones y columnas equivalentes al diseño de cuadrado latino.

Cuadro 16.7 Diseño cruzado de 2×2

Secuencia	Periodo	
	I	II
1	A	→ B
2	B	→ A

Modelo lineal para dos tratamientos

El modelo lineal para el diseño con dos tratamientos es:

$$y_{ijk} = \mu + \alpha_i + b_j + \gamma_k + \tau_d + \lambda_c + e_{ijk} \quad (16.5)$$

$$i, k, d, c = 1, 2 \quad j = 1, 2, \dots, r_i$$

donde μ es la media general, α_i es el efecto de la secuencia, b_j es efecto del sujeto aleatorio con media 0 y varianza σ_b^2 , γ_k es el efecto del periodo, τ_d es el efecto directo del tratamiento, λ_c es el efecto acarreado y e_{ijk} es el error aleatorio independiente con media 0 y varianza σ^2 .

Efectos acarreadados confundidos en el diseño de 2×2

El diseño cruzado de 2×2 tiene una deficiencia significativa. Los efectos acarreados se confunden con otros efectos del estudio y la significancia de los efectos acarreados no se puede probar a partir del análisis de la ecuación (16.5). Las medias observadas para los sujetos en cada periodo de cada secuencia del diseño, $\bar{y}_{11}, \bar{y}_{12}, \bar{y}_{21}$ y \bar{y}_{22} , contienen toda la información disponible para estimar los efectos del factor de diseño para secuencias y periodos, y los efectos de tratamiento acarreados. Los valores esperados de las medias de celdas, $E(\bar{y}_{ik}) = \mu_{ik} = \mu + \alpha_i + \gamma_k + \tau_d + \lambda_c$, se muestran en el cuadro 16.8.

Cuadro 16.8 Medias de celdas esperadas para el diseño cruzado de 2×2

Secuencia	Medias de celdas esperadas, μ_{ij}	
	Periodo I	Periodo II
A → B	$\mu + \alpha_1 + \gamma_1 + \tau_1$	$\mu + \alpha_1 + \gamma_2 + \tau_2 + \lambda_1$
B → A	$\mu + \alpha_2 + \gamma_1 + \tau_2$	$\mu + \alpha_2 + \gamma_2 + \tau_1 + \lambda_2$

Los valores esperados simplificados se obtienen una vez aplicada la relación $\tau_1 + \tau_2 = 0$, basada en la definición de los efectos de tratamiento. La suma de dos efectos de tratamiento es 0, de manera que $\tau_2 = -\tau_1$ y $-\tau_1$ se puede sustituir por τ_2 en los valores esperados. Es posible hacer sustituciones similares para otros efectos de factores. Dadas las sustituciones $\alpha_2 = -\alpha_1$, $\gamma_2 = -\gamma_1$, y $\lambda_2 = -\lambda_1$, los valores esperados son:

$$E(\bar{y}_{11}) = \mu_{11} = \mu + \alpha_1 + \gamma_1 + \tau_1$$

$$E(\bar{y}_{12}) = \mu_{12} = \mu + \alpha_1 - \gamma_1 - \tau_1 + \lambda_1$$

$$E(\bar{y}_{21}) = \mu_{21} = \mu - \alpha_1 + \gamma_1 - \tau_1$$

$$E(\bar{y}_{22}) = \mu_{22} = \mu - \alpha_1 - \gamma_1 + \tau_1 - \lambda_1$$

Los cinco parámetros a estimar son μ , α_1 , γ_1 , τ_1 y λ_1 , y sólo se dispone de cuatro medias observadas para la estimación. No es posible obtener estimaciones de parámetros, para algunos efectos, sin incluir los parámetros de otros. En particular, no es posible separar en el análisis el efecto de la secuencia α_1 o del tratamiento directo τ_1 del efecto acarreado λ_1 . Si uno de los parámetros se puede eliminar del modelo, el resto se podrá estimar a partir de las medias de celdas.

Eliminar algunos parámetros del modelo para estimar los efectos

Se han propuesto varios modelos alternativos y métodos de análisis para el diseño cruzado de 2×2 (Grizzle, 1965, 1974; Hills y Armitage, 1979; Milliken y Johnson, 1984). Sin importar el modelo, debe eliminarse uno de los parámetros para estimar los otros y probar las hipótesis sobre ellos. En ciertas circunstancias, es posible dar argumentos para eliminar del modelo los efectos de secuencia o de periodo. Eliminar cualquiera de estos dos parámetros pone en riesgo el análisis porque las inferencias se hacen bajo las suposiciones amplias de que no hay periodo o secuencia. Los efectos se dejan en el modelo porque es necesario probar su significancia en el estudio, no porque se piense que en realidad están presentes.

Es posible hacer una prueba de significancia para los efectos acarreados del diseño cruzado de 2×2 con la suposición cuestionable de que no existe algún otro efecto; en estas circunstancias, el diseño no se debe usar. Jones y Kenward (1989) consideran el diseño adecuado si los efectos de secuencia, periodo y transferibles son despreciables en cuanto al tamaño relativo de los efectos del tratamiento. Por otro lado, una asamblea del US Food and Drug Administration recomendó que el diseño no se usara para la evaluación de medicamentos.

Un poco de ayuda de las observaciones base y de limpieza

Los anexos al diseño de 2×2 revisado con detalle por Jones y Kenward (1989) incluyen observaciones base tomadas antes del suministro del tratamiento que permiten estimar los efectos acarreados y probar su significancia. Sin embargo, los efectos directos y acarreados todavía no son ortogonales en el diseño y sus estimaciones tienen una alta correlación, $\rho = 0.87(\sigma^2/n)$. Aunque los dos efectos ya no se confunden, la alta correlación entre sus estimaciones causa dificultades para distinguir sus contribuciones separadas. Cuando las estimaciones de los efectos tienen una alta correlación, pueden obtenerse pruebas de significancia contradictorias en el análisis. Un efecto puede ser no significativo cuando un segundo efecto está en el modelo y puede ser significativo cuando se elimina el segundo efecto.

La inclusión de observaciones de limpieza tomadas entre dos periodos de tratamiento reduce la correlación entre las estimaciones de los dos efectos a $\rho = 0.50(\sigma^2/n)$. Aunque la correlación se reduce, todavía existe una alta correlación entre los efectos directos y acarreados. Las estimaciones no se correlacionan si se incluye un segundo periodo de limpieza, después de terminar el segundo tratamiento.

Así, solamente es posible obtener un diseño totalmente ortogonal de dos tratamientos con la adición de los datos base y de limpieza, que requieren un total de cinco períodos de medición.

Adición de un periodo extra para estimar los efectos acarreados

Las otras adiciones al diseño cruzado de 2×2 revisadas por Jones y Kenward (1989) incluyen diseños para los dos tratamientos administrados en varias combinaciones de secuencias y períodos. Al agregar secuencias o períodos de tratamiento aumenta la capacidad de estimar los efectos importantes en el estudio cruzado. En esta sección se presentan dos de los diseños para dos secuencias en tres períodos.

El diseño de *reanudación* para dos tratamientos, mostrado en el cuadro 16.9, se usa con frecuencia como un anexo al diseño de 2×2 para estimar los efectos directos y acarreados. Sin embargo, el diseño del *periodo extra* para dos tratamientos, mostrado en el mismo cuadro, es mejor para la estimación de los efectos de tratamiento.

Cuadro 16.9 Diseño de reanudación y de periodo extra para dos tratamientos

Secuencia	<i>Periodo de reanudación</i>			Secuencia	<i>Periodo extra</i>						
	I	II	III		I	II	III				
1	A	→	B	→	A	1	A	→	B	→	B
2	B	→	A	→	B	2	B	→	A	→	A

Los dos primeros períodos de ambos diseños proporcionan la base del diseño cruzado de 2×2 , pero los diseños difieren en la administración del tratamiento durante el tercer período. El tercer período de tratamiento regresa al del primero en el diseño de reanudación, mientras que el del segundo período se repite al tercero en el diseño de período extra.

Ambos diseños tienen seis medias de celdas para la estimación de los efectos del diseño y de tratamiento, y se puede emplear el análisis de varianza de la tabla 16.2, basado en el modelo de la ecuación (16.5) para las particiones de sumas de cuadrados. Un análisis de ambos diseños proporciona pruebas válidas para todos los efectos del modelo.

El diseño de período extra supera al diseño de reanudación

En las secciones anteriores sobre diseños balanceados se afirma que los efectos de tratamiento directos y acarreados son ortogonales en el diseño de período extra, de manera que las estimaciones de estos efectos no se correlacionan. Los efectos son no ortogonales en el diseño de reanudación, y la correlación entre los efectos directos y acarreados es $\rho = 0.87(\sigma^2/n)$.

La superioridad adicional de los diseños de periodo extra reside en su mayor eficiencia para la estimación de los efectos de tratamiento directos y acarreados. Las varianzas de las estimaciones para estos efectos en el diseño de periodo extra son:

$$\sigma_{\hat{\tau}}^2 = 0.19 \left(\frac{\sigma^2}{n} \right) \quad \text{y} \quad \sigma_{\hat{\lambda}}^2 = 0.25 \left(\frac{\sigma^2}{n} \right)$$

y las varianzas para los efectos en el diseño de reanudación son:

$$\sigma_{\hat{\tau}}^2 = 0.75 \left(\frac{\sigma^2}{n} \right) \quad \text{y} \quad \sigma_{\hat{\lambda}}^2 = 1.0 \left(\frac{\sigma^2}{n} \right)$$

Las propiedades estadísticas de los diseños de periodo extra son superiores a las del diseño de reanudación, ya que las estimaciones de los efectos de tratamiento directos y acarreados no se correlacionan y las varianzas de las estimaciones son menores.

Hasta este punto, la elección del diseño se ha enfocado en la estimación y las pruebas de significancia para el diseño principal y los factores de tratamiento en el diseño cruzado: secuencia, periodo y efectos de tratamiento directos y acarreados. Se pueden construir otros diseños para facilitar la estimación de los efectos adicionales, que son potencialmente importantes en algunos estudios. Estos efectos incluyen las interacciones entre los efectos directos y acarreados, y las interacciones entre los efectos de tratamiento y los factores de diseño, como períodos y secuencias. Los diseños para la estimación y pruebas de estos modelos de efectos adicionales se pueden consultar en Jones y Kenward (1989).

EJERCICIOS PARA EL CAPÍTULO 16

- Un fabricante farmacéutico realizó un estudio cruzado con un medicamento contra las convulsiones (DPH) empleado para manejar lesiones psicomotrices. Se administró una sola dosis de DPH a un sujeto y se midió el nivel de droga en el plasma 12 horas después de administrarlo. Los cuatro tratamientos fueron (A) 100 mg de DPH genérico en solución, (B) 100 mg de DPH procesado en cápsula, (C) 100 mg de DPH genérico en cápsula y (D) 300 mg de DPH procesado en cápsula.

Se usó un diseño de cuadrado latino cruzado de 4×4 para el estudio. Se asignaron dos sujetos a cada una de las cuatro secuencias de tratamiento y se suministró una sola dosis de cada tratamiento a cada uno, con un intervalo de dos semanas entre los tratamientos. Los niveles de DPH en el plasma se muestran en la tabla al final del ejercicio.

- ¿Está balanceado el diseño para los efectos acarreados? Explique.
- Escriba un modelo lineal para el estudio, explique los términos y establezca las suposiciones necesarias para el análisis.
- Calcule el análisis de varianza para los datos y pruebe la significancia de los efectos directos y acarreados.
- El fabricante deseaba saber si la biodisponibilidad del DPH en forma de cápsula era tan alta como la de forma líquida, si la biodisponibilidad del producto genérico era equivalente a la de su producto, y si

- Se repara para la prueba de alta calidad dietaria rante
- A asi ser per a. b. c. d. e.

- había un efecto de dosis (300 mg contra 100 mg) en la biodisponibilidad. Establezca un conjunto razonable de contrastes entre los cuatro tratamientos, pruebe su significancia e interprete los resultados.
- e. Obtenga las gráficas residuales para el análisis e interprételas.

Secuencia	Sujeto	Periodo			
		I	II	III	IV
ABDC	1	2.3	1.7	1.8	4.5
	2	1.0	1.2	1.1	3.3
BCAD	3	1.9	4.4	1.4	1.5
	4	0.9	2.1	1.0	0.8
CDBA	5	3.2	0.8	0.8	0.9
	6	2.2	0.8	0.9	0.8
DACB	7	1.4	1.2	3.6	1.1
	8	0.9	0.8	2.9	0.8

Fuente: K. S. Albert *et al.* (1974), Bioavailability of diphenylhydantoin. *Clinical Pharmacology and Therapeutics*, 16, 727-735.

2. Se usó un diseño cruzado para comparar medicinas que controlan la hipertensión. Se usaron dos drogas, A y B, solas y combinadas. La combinación de las dos drogas se etiquetó como droga C. Los sujetos se asignaron al azar a una de las seis secuencias de tratamiento y cada periodo de tratamiento duró cuatro semanas, sin período de descanso entre tratamientos. La presión sistólica de los sujetos al final de cada periodo se muestra en la tabla al principio de la siguiente página.
- ¿Está balanceado el diseño para los efectos acarreados? Explique.
 - Escriba un modelo lineal para el estudio, explique los términos y establezca las suposiciones necesarias para el análisis.
 - Calcule el análisis de varianza para los datos y pruebe la significancia de los efectos directos y acarreados.
 - Calcule por separado los contrastes entre la droga C (la combinación de A y B) y las drogas A y B, junto con su error estándar. ¿Existe una diferencia significativa en la presión sistólica para la combinación de las dos drogas y la de las dos drogas solas? ¿Hay una diferencia significativa en la presión con las drogas A y B?
 - Obtenga las gráficas residuales para el análisis e interprételas.
3. Se realizó un estudio de la digestión de novillos en un diseño de cuadrado latino cruzado con periodo extra para evaluar los efectos de la fibra de baja calidad en la digestión. Las fibras de baja calidad usadas en la prueba eran (A) cáscara de semilla de algodón, (B) paja de pasto bermuda, y (C) paja de trigo, y la fibra de alta calidad usada como control era (D) alfalfa. Se asignó al azar un novillo a cada secuencia de cuatro dietas. Los novillos recibieron cada dieta durante 30 días y se midió la digestión en la materia seca durante la última semana de cada prueba, lo que proporcionó un ajuste de 21 días de adaptación a cada dieta.

Secuencia	Sujeto	Periodo		
		I	II	III
ABC	1	174	146	164
	2	145	125	130
	3	230	174	200
	4	240	130	195
ACB	5	192	150	160
	6	194	208	160
	7	175	152	175
	8	202	160	180
BAC	9	184	192	176
	10	140	150	150
	11	155	230	226
	12	180	185	190
BCA	13	136	132	138
	14	145	154	166
	15	194	210	190
	16	180	180	190
CAB	17	206	220	210
	18	160	180	145
	19	188	200	190
	20	185	197	182
CBA	21	180	180	208
	22	210	160	226
	23	185	180	200
	24	190	145	160

Fuente: B. J. Jones y M. G. Kenward (1989), *Design and Analysis of Cross-Over Trials*, London, Chapman Hall.

La dieta de fibra suministrada en el cuarto periodo se repitió durante el quinto. Los datos de la materia seca para cada novillo, en cada secuencia, se muestran en la siguiente tabla:

Novillo	Periodo				
	I	II	III	IV	V
1	75(A)	76(B)	79(C)	81(D)	79(D)
2	79(C)	73(A)	79(D)	75(B)	77(B)
3	81(D)	79(C)	75(B)	72(A))	73(A)
4	76(B)	79(D)	72(A)	76(C)	73(C)

- ¿Está balanceado el diseño para los efectos acarreados? Explique.
- ¿Tiene alguna ventaja este diseño respecto al diseño de cuatro períodos? Explique.
- El modelo lineal para este diseño cruzado de periodo extra es:

$$y_{ij} = \mu + \alpha_i + \gamma_j + \tau_d + \lambda_c + e_{ij}$$

$$i = 1, 2, \dots, 4 \quad j = 1, 2, \dots, 5 \quad d, c = 1, 2, \dots, 4$$

donde μ es la media general, α_i es el efecto de la secuencia de novillos, γ_j es el efecto del período, τ_d es el efecto directo del tratamiento, λ_c es el efecto transferido y e_{ij} es el error aleatorio independiente con distribución normal con varianza σ^2 .

- Calcule el análisis de varianza para los datos. Realice los cálculos en una computadora usando un ajuste secuencial para los efectos directos y acarreados. Observe que las sumas de cuadrados para estos efectos son las mismas, sin importar su orden en el ajuste secuencial de los términos. ¿Por qué?
 - Pruebe la significancia de los efectos directos y acarreados.
 - Utilice el método de Dunnett para comparar la dieta de control, alfalfa, con cada dieta e interprete los resultados.
 - Obtenga las gráficas residuales para el análisis e interprételas.
- Se usó un diseño cruzado de período extra para comparar dos medicamentos para el control de la hipertensión. Los sujetos se asignaron al azar a una de las secuencias de tratamiento de drogas. Cada tratamiento duró seis semanas con una semana de descanso o limpieza entre tratamientos. Las presiones diastólicas de los sujetos, medidas al final de cada período, fueron:

Secuencia	Sujeto	Período		
		I	II	III
ABB	1	73	92	75
	2	90	90	80
	3	95	75	75
	4	80	80	90
	5	90	90	70
	6	45	60	45
	7	70	60	80
	8	122	101	90
BAA	9	60	70	100
	10	100	85	80
	11	50	45	70
	12	65	70	70
	13	88	88	88
	14	70	70	80
	15	60	90	70

- ¿Está balanceado el diseño para los efectos acarreados? Explique.

- b. ¿Tiene alguna ventaja este diseño con respecto al de dos períodos simple, sin usar el tercer período, como se muestra en los datos de la tabla? Explique.
- c. Escriba un modelo lineal para el estudio, explique los términos y establezca las suposiciones necesarias para el análisis.
- d. Calcule el análisis de varianza para los datos y pruebe la significancia de los efectos de tratamiento directos y acarreados.
- e. Calcule el contraste entre las drogas A y B junto con el error estándar. ¿Difieren A y B de manera significativa respecto de la presión diastólica?
- f. Obtenga las gráficas residuales para el análisis e interprételas.
5. Se realizó un estudio cruzado para evaluar cuatro distribuciones de teclados. Se usaron cuatro voluntarios familiarizados con la distribución común de teclados, cada sujeto usó las cuatro distribuciones en secuencia y se le asignó al azar una secuencia de distribuciones. Cada distribución se usó durante cuatro días para las actividades normales de captura de texto y datos. El quinto día realizaron todos la misma tarea con el teclado respectivo y se registró el número de errores. Ninguno de los sujetos sabía que hacían una prueba el último día. El número de errores registrado para cada sujeto en cada distribución fue:

<i>Sujeto</i>	<i>Periodo</i>			
	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
1	7(D)	2(B)	1(A)	5(C)
2	1(A)	4(C)	6(D)	3(B)
3	6(C)	1(A)	3(B)	7(D)
4	3(B)	6(D)	3(C)	1(A)
5	4(C)	5(D)	1(A)	2(B)
6	6(D)	4(C)	2(B)	0(A)
7	1(A)	3(B)	4(C)	5(D)
8	2(B)	2(A)	7(D)	4(C)
9	5(D)	0(A)	3(C)	3(B)
10	0(A)	4(D)	2(B)	3(C)
11	3(C)	2(B)	7(D)	0(A)
12	2(B)	4(C)	0(A)	6(D)

- a. ¿Está balanceado el diseño para los efectos acarreados? Explique.
- b. Escriba un modelo lineal para el estudio, explique los términos y establezca las suposiciones necesarias para el análisis. ¿Piensa que una suposición de distribución normal para las observaciones es válida en este estudio? Explique.
- c. Calcule el análisis de varianza para los datos y pruebe la significancia de los efectos directos y acarreados.
- d. Obtenga las gráficas residuales para el análisis e interprételas. ¿Cree que una transformación de los datos mejorará el análisis? Si es así, ¿qué transformación sería adecuada?

Apéndice 16A.1 Codificación de archivos de datos para estudios cruzados

Efectos acarreados codificados como niveles del factor

Este esquema de codificación fue sugerido por Ratkowsky, Alldredge y Cotton (1990) para el programa GLM en SAS. Con este esquema, las particiones de sumas de cuadrados correctas se producen también mediante el programa MANOVA en SPSS. Las estimaciones de las medias de tratamiento ajustadas pueden diferir en alguna cantidad constante entre programas debido a los distintos algoritmos de cálculo usados, pero las estimaciones de contraste entre medias de tratamiento y errores estándar serán congruentes. Con este esquema de codificación algunos programas no terminan el análisis.

Los códigos básicos para los niveles de un factor en el archivo de datos son una secuencia de enteros 1, 2, 3, Se requiere un código para el nivel de cada factor para cada observación en el archivo de datos, pero los efectos acarreados no se presentan durante las observaciones del primer periodo de un estudio cruzado y no hay “nivel” que codificar para ellos. Algunos programas permiten que haya un código para el factor, aunque el factor no se encuentre en la observación. Para los ejemplos de este capítulo, los niveles del factor acarreado se identificaron con los niveles 0, 1, 2, ..., t , donde “0” se usó para las observaciones del primer periodo, se asignó el código “1” a la observación si aparecía el tratamiento 1 del periodo anterior, un código “2” si el tratamiento 2 estaba en el periodo anterior y así sucesivamente.

El código completo para el archivo de datos usado en el ejemplo 16.1 incluye las columnas de códigos para el periodo (P), la secuencia (S), al animal dentro de la secuencia (A), la dieta (D) y los factores de acarrreo (C) al igual que las columnas de los datos para la variable de respuesta DFN.

P	S	A	D	C	FDN	P	S	A	D	C	FDN	P	S	A	D	C	FDN
1	1	1	1	0	50	2	1	1	2	1	61	3	1	1	3	2	53
1	1	2	1	0	55	2	1	2	2	1	63	3	1	2	3	2	57
1	2	1	2	0	44	2	2	1	3	2	42	3	2	1	1	3	57
1	2	2	2	0	51	2	2	2	3	2	45	3	2	2	1	3	59
1	3	1	3	0	35	2	3	1	1	3	55	3	3	1	2	1	47
1	3	2	3	0	41	2	3	2	1	3	56	3	3	2	2	1	50
1	4	1	1	0	54	2	4	1	3	1	48	3	4	1	2	3	51
1	4	2	1	0	58	2	4	2	3	1	51	3	4	2	2	3	54
1	5	1	2	0	50	2	5	1	1	2	57	3	5	1	3	1	51
1	5	2	2	0	55	2	5	2	1	2	59	3	5	2	3	1	55
1	6	1	3	0	41	2	6	1	2	3	56	3	6	1	1	2	58
1	6	2	3	0	46	2	6	2	2	3	58	3	6	2	1	2	61

Observe que la codificación para el factor acarreado C es “0” en el primer periodo ($P = 1$). El código del factor acarreado en el segundo periodo ($P = 2$) es

idéntico al de la dieta (D) para el mismo novillo en el periodo anterior ($P = 1$). De la misma manera, los códigos de las dietas para el segundo periodo son los usados para el código de acarreo en el tercer periodo.

Efectos acarreados codificados como covariadas

Milliken y Johnson (1984) introdujeron los efectos acarreados en el modelo como covariadas. El método para codificar suele llamarse de códigos de variables *ficticias*, “dummy” o *indicadoras* y se basa en el conocimiento de que una restricción debe colocarse en las soluciones para los efectos del factor en las ecuaciones normales. Una restricción en la relación entre las soluciones que se puede usar para los efectos acarreados es $\hat{\lambda}_1 + \hat{\lambda}_2 + \dots + \hat{\lambda}_t = 0$, y que el conocimiento de $t - 1$ efectos proporciona en forma automática el valor para el efecto restante; la relación es $\hat{\lambda}_1 + \hat{\lambda}_2 + \hat{\lambda}_3 = 0$ con tres efectos tratamiento acarreados. El conocimiento de los valores para dos efectos acarreados proporciona de manera automática el valor para el tercero, por lo tanto, $\hat{\lambda}_3 = -\hat{\lambda}_1 - \hat{\lambda}_2$ y los efectos acarreados tienen 2 grados de libertad.

Se necesitan sólo $t - 1$ covariadas para que el modelo lineal incluya los parámetros para los efectos acarreados debido a las restricciones sobre ellos. El modelo para el estudio de dietas de fibra se usa como ejemplo. El modelo expresado con dos covariadas para los efectos acarreados es:

$$y_{ijk} = \mu + \alpha_i + b_j + \gamma_k + \tau_d + \lambda_1 X_1 + \lambda_2 X_2 + e_{ijk} \quad (16A.1)$$

$$i = 1, 2, \dots, 6 \quad j = 1, 2 \quad k = 1, 2, 3 \quad d = 1, 2, 3$$

Un método para codificar las covariadas de los efectos acarreados en el periodo k es:

$$X_1 = \begin{cases} 1 & \text{si } d = 1 \text{ en periodo } (k-1) \\ -1 & \text{si } d = 3 \text{ en periodo } (k-1) \\ 0 & \text{de otra manera} \end{cases}$$

Así, $X_1 = 1$ cuando la dieta A se encuentra en el periodo anterior y $X_1 = -1$ cuando la dieta C aparece en el periodo anterior.

También:

$$X_2 = \begin{cases} 1 & \text{si } d = 2 \text{ en periodo } (k-1) \\ -1 & \text{si } d = 3 \text{ en periodo } (k-1) \\ 0 & \text{de otra manera} \end{cases}$$

con lo que $X_2 = 1$ cuando la dieta B aparece en el periodo anterior y $X_2 = -1$ cuando lo hace la dieta C en el periodo anterior.

La codificación completa para el archivo de datos del ejemplo 16.1 incluye las columnas de códigos para periodo (P), secuencia (S), animal dentro de la secuencia (A), dieta (D) y las covariadas para los efectos acarreados X_1 y X_2 , al igual que las columnas de datos para la variable de respuesta FDN.

No existe efecto acarreado en el primer periodo ($P = 1$) y las covariadas se codifican como $X_1 = X_2 = 0$. La observación de la dieta B en el segundo periodo ($P = 2$) de la

P	S	A	D	X_1	X_2	FDN	P	S	A	D	X_1	X_2	FDN	P	S	A	D	X_1	X_2	FDN
1	1	1	1	0	0	50	2	1	1	2	1	0	61	3	1	1	3	0	1	53
1	1	2	1	0	0	55	2	1	2	2	1	0	63	3	1	2	3	0	1	57
1	2	1	2	0	0	44	2	2	1	3	0	1	42	3	2	1	1	-1	-1	57
1	2	2	2	0	0	51	2	2	2	3	0	1	45	3	2	2	1	-1	-1	59
1	3	1	3	0	0	35	2	3	1	1	-1	-1	55	3	3	1	2	1	0	47
1	3	2	3	0	0	41	2	3	2	1	-1	-1	56	3	3	2	2	1	0	50
1	4	1	1	0	0	54	2	4	1	3	1	0	48	3	4	1	2	-1	-1	51
1	4	2	1	0	0	58	2	4	2	3	1	0	51	3	4	2	2	-1	-1	54
1	5	1	2	0	0	50	2	5	1	1	0	1	57	3	5	1	3	1	0	51
1	5	2	2	0	0	55	2	5	2	1	0	1	59	3	5	2	3	1	0	55
1	6	1	3	0	0	41	2	6	1	2	-1	-1	56	3	6	1	1	0	1	58
1	6	2	3	0	0	46	2	6	2	2	-1	-1	58	3	6	2	1	0	1	61

primera secuencia ($S = 1$) sigue a la dieta A del primer periodo; entonces, se requiere un efecto acarreado para la dieta A y los códigos de las covariadas son $X_1 = 1$ y $X_2 = 0$. La observación de la dieta A en el tercer periodo ($P = 3$) de la segunda secuencia ($S = 2$) sigue a la dieta C del segundo periodo; así, se requiere un efecto acarreado para la dieta C con códigos de covariadas $X_1 = -1$ y $X_2 = -1$.

La suma de cuadrados para las dos covariadas, X_1 y X_2 , será la suma de cuadrados para los efectos acarreados ajustados por los efectos directos. Este método de análisis se puede usar con cualquier programa estadístico que permita incluir covariadas en el modelo. Cada programa tendrá una sintaxis única para la representación de las covariadas agregadas al modelo.

Apéndice 16A.2 Suma de cuadrados del tratamiento para los diseños balanceados

Las cantidades mostradas al lado de los cálculos manuales de las sumas de cuadrados ajustadas y los efectos acarreados en los diseños cruzados balanceados y los de periodo extra balanceados son válidas para un número igual de sujetos por secuencia de tratamiento. Con la notación de Cochran y Cox (1957), las fórmulas necesarias son:

T_i = totales de tratamiento, $i = 1, 2, \dots, t$

R_i = total de observaciones en los períodos inmediatos siguientes al tratamiento i

F_i = total de secuencias en las que el tratamiento i es el último

P_1 = total del periodo 1

G = gran total de todas las observaciones

Además, es t = número de tratamientos, r = número de sujetos por secuencia y m = número de cuadrados latinos usados en el diseño.

Diseños de Williams: para los diseños de Williams se calcula:

$$\widehat{T}_i = (t^2 - t - 1)T_i + tR_i + F_i + P_1 - tG$$

para la suma de cuadrados ajustada por efectos directos de tratamiento:

$$SCT(\text{ajustada}) = \frac{\sum \widehat{T}_i^2}{rmt(t^2 - t - 1)(t + 1)(t - 2)}$$

Se calcula:

$$\widehat{R}_i = tT_i + t^2R_i + tF_i + tP_1 - (t + 2)G$$

para la suma de cuadrados ajustada por efectos acarreados de tratamiento:

$$SSC(\text{ajustada}) = \frac{\sum \widehat{R}_i^2}{rmt^3(t + 1)(t - 2)}$$

Se calculan las sumas de cuadrados no ajustadas usuales para períodos, secuencias, sujetos dentro de las secuencias y tratamientos, según el diseño del rectángulo latino del capítulo 8. Las sumas de cuadrados del error experimental para completar la partición de suma de cuadrados para el análisis de varianza, como se muestra en la tabla 16.2, es:

$$SCE = SC \text{ Total} - SCS - SCW - SCP - SCT(\text{no ajustada}) - SCC$$

Las estimaciones de las diferencias entre las medias de tratamiento ajustadas se pueden encontrar con:

$$\widehat{\mu}_i - \widehat{\mu}_j = \frac{\widehat{T}_i - \widehat{T}_j}{rmt(t + 1)(t - 2)}$$

La estimación del error estándar para el contraste con medias ajustadas a partir de un diseño cruzado balanceado es:

$$s_{(\widehat{\mu}_i - \widehat{\mu}_j)} = \sqrt{\frac{2CME}{N} \cdot \frac{(t^2 - t - 1)}{(t + 1)(t - 2)}}$$

donde N es el número de réplicas por tratamiento.

Diseño de periodo extra de Williams: los ajustes necesarios para los cálculos en los diseños de periodo extra son:

$$\widehat{T}_i = (t + 1)T_i - F_i - G$$

con la suma de cuadrados ajustada de los tratamientos como:

$$SCT(\text{ajustado}) = \frac{\sum \widehat{T}_i^2}{mt(t + 1)(t + 2)}$$

Se calcula:

$$\widehat{R}_t = tR_t + P_1 + G$$

para la suma de cuadrados ajustada por los efectos acarreados como:

$$SCC(\text{ajustado}) = \frac{\sum \widehat{R}_t^2}{mt^3}$$

La diferencia estimada entre dos medias de tratamiento ajustadas es:

$$\widehat{\mu}_i - \widehat{\mu}_j = \frac{\widehat{T}_i - \widehat{T}_j}{mt(t+2)}$$

con una estimación del error estándar:

$$S(\widehat{\mu}_i - \widehat{\mu}_j) = \sqrt{\frac{2CME}{mt} \cdot \frac{(t+1)}{(t+2)}}$$

17 Análisis de covarianza

El tema principal de este capítulo es el uso de información adicional en las unidades experimentales como una práctica de control local para reducir la estimación del error experimental. Los valores de las medias de tratamiento en un estudio de investigación dependen de covariadas que cambian entre las unidades experimentales y tienen una relación significativa con la variable de respuesta principal. En este capítulo se usa el análisis de covarianza para eliminar la influencia de las covariadas sobre las comparaciones de los tratamientos en diseños totalmente aleatorizados y diseños de bloques completos.

17.1 Control local con una covariada medida

En los experimentos se usan varias técnicas de control local para gobernar la varianza del error experimental. Estas prácticas de control reducen el error experimental y aumentan la precisión de las estimaciones para las medias de tratamiento y las pruebas de hipótesis. Con frecuencia se usan variables concomitantes o relacionadas para seleccionar y agrupar unidades para controlar la variación del error experimental.

Muchas variables concomitantes, o *covariadas*, pueden medirse en cualquier momento durante el experimento y al analizar los resultados se puede evaluar su influencia sobre la variable de respuesta. El análisis de covarianza, que combina la metodología de regresión con el análisis de varianza, evalúa la influencia de la covariada sobre la variable de respuesta y permite comparar tratamientos con una base común relativa a los valores de las covariadas.

A menudo, muchos factores ajenos a los factores de tratamiento influyen en la variable de respuesta. El uso de bloques, según estos factores de influencia, es uno de los principales medios que usan los investigadores para controlar el error experimental. Cuando se construyen bloques de unidades con valores similares para los factores, los tratamientos del experimento se pueden comparar entre sí en un entorno más homogéneo.

Muchas veces, las condiciones experimentales impiden la formación de bloques de unidades similares por muchas razones, ya que puede tenerse un conocimiento incompleto del material experimental, o los efectos de factores externos quizás no aparezcan hasta después de iniciado el experimento y es posible que se tengan muy pocas unidades o valores similares para completar bloques. Aunque el investigador conozca los factores que influyen, el número de factores adicionales tal vez haga imposible usarlos todos como criterios de bloque.

Los estudios típicos con variables adicionales, que influyen en las comparaciones de tratamientos, abarcan:

- Pruebas clínicas donde la edad, peso, género, historia clínica u ocupación de los pacientes pueden influir en su respuesta a los tratamientos. Aunque los investigadores pueden bloquizar dos o tres de estos factores, no es posible ignorar su influencia en el resto de las variables.
- Pruebas con árboles frutales, que se bloquizan según el suelo o los gradientes de irrigación, pero donde no se puede ignorar la influencia de tratamientos anteriores y los registros históricos de cosechas.
- Pruebas de alimentación, donde pueden construirse bloques de animales de acuerdo con la camada y los pesos iniciales, pero la cantidad de alimento consumida durante el estudio también tendrá influencia en la medida del aumento de peso.

El estudio descrito en el ejemplo 17.1 explica un experimento en el que la variable de respuesta no sólo es afectada por el tratamiento aplicado a los sujetos del estudio, sino también por una covariada que se midió en cada sujeto antes del estudio.

Ejemplo 17.1 Efectos del ejercicio en la ventilación de oxígeno

Un método clínico común para evaluar la capacidad cardiovascular de los individuos es mediante pruebas de ejercicio en una caminadora de banda. Una de las medidas obtenidas durante las pruebas, el consumo máximo de oxígeno, se considera el mejor índice de capacidad de trabajo y la función cardiovascular máxima.

La medida del consumo máximo de oxígeno de un individuo depende de varios factores, como el modo y el protocolo de prueba, además de la condición física y la edad del sujeto. Un protocolo de prueba común en la caminadora es el protocolo de inclinación, donde el grado y velocidad se aumentan hasta que el sujeto queda exhausto.

Hipótesis de investigación: los investigadores de la fisiología del ejercicio tenían la opinión de que las condiciones de la prueba en caminadora debería simular lo más posible el tipo de entrenamiento cardiovascular del sujeto, para lograr su máxima oxigenación durante la prueba. Se planteó la hipótesis de que el entrenamiento aeróbico con escalón se simulaba mejor con el protocolo de inclinación que con el régimen de carrera plana.

Diseño del tratamiento: los dos tratamientos seleccionados para el estudio fueron un programa de 12 semanas de entrenamiento aeróbico con escalón y 12 semanas de carrera plana al aire libre. Se probó a los sujetos en la caminadora antes y después del periodo de 12 semanas con el protocolo de inclinación. Si la

hipótesis fuera cierta, los sujetos en el entrenamiento aeróbico con escalón mostrarían un mayor incremento en su oxigenación máxima que los del entrenamiento en terreno plano y éstos quizás estarían limitados por la fatiga de ciertos músculos en la prueba con protocolo de inclinación y no lograrían la oxigenación máxima antes de quedar exhaustos.

Diseño del experimento: los sujetos fueron 12 hombres sanos que no participaban en programas de ejercicio regulares. Se asignaron seis al azar a cada grupo en un diseño totalmente aleatorizado y se les tomaron varias medidas respiratorias en la caminadora antes del periodo de 12 semanas. No había diferencias en las medidas respiratorias de los dos grupos de sujetos antes del tratamiento.

La medida de interés para este ejemplo es el cambio en la ventilación máxima (litros/minuto) de oxígeno para el periodo de 12 semanas. Las observaciones de los 12 sujetos y sus edades se muestran en la tabla 17.1, junto con el grupo de medias y errores estándar.

Tabla 17.1 Cambio máximo de ventilación (litros/minuto) luego de un programa de ejercicio de 12 semanas

Grupo	Edad	Cambio	Grupo	Edad	Cambio
Aeróbicos	31	17.05	Carrera	23	- 0.87
	23	4.96		22	- 10.74
	27	10.40		22	- 3.27
	28	11.05		25	- 1.97
	22	0.26		27	7.50
	24	2.51		20	- 7.25
Media	25.83	7.71		23.17	- 2.77
Err. Est.	1.40	2.55		1.01	2.54

Fuente: D. Allen, *Exercise Physiology*, University of Arizona.

¿Se relaciona el desempeño en la caminadora con la edad?

En la figura 17.1 se muestra una gráfica del cambio en la ventilación máxima (*y*) para cada sujeto con su edad (*x*) en el eje horizontal. La gráfica sugiere una relación positiva fuerte entre la edad del sujeto y su cambio en la ventilación máxima en la caminadora, sin importar el grupo de tratamiento al que pertenecen. Así, parece haber una considerable variación del error experimental dentro de cada grupo, asociada con las diferencias de edad.

El protocolo de estudio requería que los sujetos elegibles fueran hombres sanos entre los 20 y 35 años con un estilo de vida sedentario. Aunque los dos grupos de sujetos no difieren en medidas de ventilación máxima de la prueba, un análisis de varianza de un lado revela una diferencia significativa entre los dos grupos en el cambio después del entrenamiento. El grupo de ejercicio aeróbico tenía un cambio mayor en la tasa de ventilación que el grupo de carrera, pero el grupo aeróbico se compone de

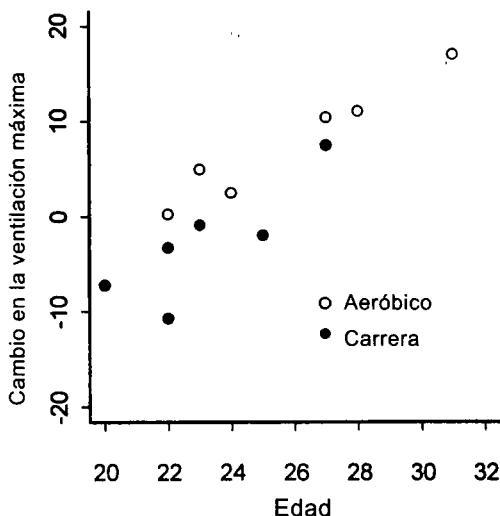


Figura 17.1 La relación entre la ventilación máxima y la edad en el estudio de ejercicio físico

un grupo de hombres mayores. Debe determinarse si la diferencia es resultado del ejercicio o de las diferencias de edad en los dos grupos. Es decir, suponiendo que los cambios en la media de ventilación máxima se comparan dentro de la misma edad para ambos grupos, ¿será la media del grupo de aeróbicos significativamente mayor que la media del grupo de carrera? Se puede usar el análisis de covarianza para ayudar a responder la pregunta y para determinar si la relación entre el cambio de ventilación máxima y la edad contribuye de manera significativa a la variación del error experimental.

17.2 Análisis de covarianza para diseños totalmente aleatorizados

Modelo lineal y análisis de covarianza

El experimento sobre fisiología del ejercicio se realizó con un diseño totalmente aleatorizado, en dos grupos de tratamiento. Suponiendo una relación lineal entre la variable de respuesta y y una covariada x , el modelo lineal para el diseño totalmente aleatorizado es:

$$y_{ij} = \mu_i + \beta(x_{ij} - \bar{x}_i) + e_{ij} \quad (17.1)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde μ_i es la media del tratamiento, β es el coeficiente para la regresión lineal de y_{ij} sobre x_{ij} , y las e_{ij} son errores experimentales aleatorios con distribución normal,

con media 0 y varianza σ^2 . Dos suposiciones clave adicionales para este modelo son que el coeficiente de regresión β es el mismo para todos los grupos de tratamiento y que los tratamientos no influyen la covariada x .

El primer objetivo del análisis de covarianza es determinar si la adición de la covariada ha reducido la estimación de la varianza de error experimental. Si la reducción es significativa, entonces se obtienen las estimaciones de las medias de grupos de tratamiento \bar{y}_{ia} ajustadas al mismo valor que la covariada x para cada grupo de tratamiento y se determina la significancia de las diferencias de tratamiento con base en las medias de tratamiento ajustadas.

Modelos alternativos para evaluar la contribución de la covariada

El análisis requerirá estimaciones de mínimos cuadrados de los parámetros para los modelos completo y reducido, que son:

- el modelo completo $y_{ij} = \mu_i + \beta(x_{ij} - \bar{x}_{..}) + e_{ij}$
- un modelo reducido sin covariada $y_{ij} = \mu_i + e_{ij}$
- un modelo reducido sin efectos de tratamiento $y_{ij} = \mu + \beta(x_{ij} - \bar{x}_{..}) + e_{ij}$

Se requiere el modelo reducido sin la covariada para evaluar la influencia de la covariada, y el modelo reducido sin efectos de tratamiento para la evaluación de la significancia de los efectos de tratamiento en presencia de la covariada.

Las estimaciones de mínimos cuadrados de los parámetros se derivan del modelo completo para obtener:

$$SCE_f = \sum [y_{ij} - \hat{\mu}_i - \hat{\beta}(x_{ij} - \bar{x}_{..})]^2$$

con $N - t - 1$ grados de libertad; el modelo reducido sin la covariada para:

$$SCE_r = \sum (y_{ij} - \hat{\mu}_i)^2$$

con $N - t$ grados de libertad; y el modelo reducido sin efectos de tratamiento para:

$$SCE_r^* = \sum [y_{ij} - \hat{\mu} - \hat{\beta}(x_{ij} - \bar{x}_{..})]^2$$

con $N - 2$ grados de libertad.

La reducción de la suma de cuadrados debida a la adición de la covariada x al modelo se obtiene como la diferencia:

$$SC(\text{Covariada}) = SCE_r - SCE_f$$

con un grado de libertad. La suma de cuadrados del tratamiento ajustada, después de ajustar la covariada es:

$$SCT(\text{ajustada}) = SCE_r^* - SCE_f$$

con $t - 1$ grados de libertad.

Las SCE para cada modelo ajustado a los datos de fisiología del ejercicio (tabla 17.1) son $SCE_f = 70.39$, $SCE_r = 389.30$ y $SCE_r^* = 142.18$.

Particiones de suma de cuadrados para el análisis de covarianza

Es posible calcular las reducciones de sumas de cuadrados con la mayoría de los programas de cómputo para análisis de covarianza y producirán la información necesaria que se encuentra en la siguiente presentación. Las particiones de suma de cuadrados para el cambio en la tasa de ventilación máxima, con la covariada edad en el estudio de fisiología del ejercicio, se muestra en el análisis de varianza de la tabla 17.2. Observe que sin la covariada edad, la varianza del error experimental estimada es $CME_r = SCE_r/(N - t) = 389.3/10 = 38.03$. La adición de la covariada ha reducido la estimación de $CME = 7.82$ en la tabla 17.2. Entonces, usar la covariada edad como control local para reducir σ^2 a través de covariadas parece ser eficaz. La ganancia en eficiencia debida a la covariada se explica más adelante en esta sección.

Tabla 17.2 Análisis de covarianza para el cambio de ventilación máxima con la edad como covariada, en un estudio de fisiología del ejercicio

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Regresión	1	318.91	318.91*	40.78	.000
Grupo	1	71.79	71.79†	9.18	.014
Error	9	70.39	7.82‡		

*CM(covariada) †CMT(ajustada) ‡CME_f

Pruebas de hipótesis sobre covariadas y tratamientos

Se determina la significancia de la reducción debida a la covariada con una prueba de la hipótesis nula $H_0: \beta = 0$. El estadístico de prueba es:

$$F_0 = \frac{CM(\text{Covariada})}{CME} = \frac{318.91}{7.82} = 40.78 \quad (17.2)$$

con valor crítico $F_{05,1,9} = 5.12$. La hipótesis nula se rechaza con $Pr > F = .000$ en la tabla 17.2, pues la adición de la covariada ha reducido de manera significativa la varianza del error experimental.

La relación significativa entre el cambio en la tasa de ventilación máxima y la edad de los sujetos indica la necesidad de evaluar la significancia de los efectos del tratamiento, luego de ajustar la covarianza. La hipótesis nula para las medias de tratamiento ajustadas es $H_0: \mu_1 = \mu_2$, y el estadístico de prueba es:

$$F_0 = \frac{CMT(\text{ajustada})}{CME} = \frac{71.79}{7.82} = 9.18 \quad (17.3)$$

con valor crítico $F_{0.05,1,9} = 5.12$. La hipótesis nula se rechaza con $Pr > F = 0.014$ en la tabla 17.2, y se concluye que las medias de tratamiento ajustadas por la edad son diferentes.

Medias de tratamiento ajustadas a un valor común de la covariada

Las estimaciones de las medias de tratamiento se ajustan a un valor común de la covariada si la inclusión de ésta en el modelo reduce de manera significativa la varianza del error experimental. Las medias de tratamiento se pueden ajustar a cualquier valor de la covariada, pero, en general, se ajustan a la media global \bar{x} como:

$$\bar{y}_{ia} = \bar{y}_i - \hat{\beta} (\bar{x}_i - \bar{x}) \quad (17.4)$$

La estimación del coeficiente de regresión, que la mayoría de los programas calculan en forma automática es:

$$\hat{\beta} = \frac{\sum_{i=1}^r \sum_{j=1}^r (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)}{\sum_{i=1}^r \sum_{j=1}^r (x_{ij} - \bar{x}_i)^2} = \frac{169.10}{89.67} = 1.886 \quad (17.5)$$

y la ecuación de regresión estimada para el i -ésimo grupo de tratamiento será:

$$\hat{y}_i = \bar{y}_i + \hat{\beta} (x_i - \bar{x}_i) \quad (17.6)$$

y son:

$$\hat{y}_{1j} = 7.70 + 1.886(x_{1j} - 25.83)$$

para el grupo de aeróbicos y:

$$\hat{y}_{2j} = -2.77 + 1.886(x_{2j} - 23.17)$$

para el grupo de carrera. La línea de regresión para cada grupo de tratamiento se muestra en la figura 17.2.

Las medias de tratamiento ajustadas a la edad promedio $\bar{x} = 24.5$ son:

$$\bar{y}_{1a} = 7.70 - 1.886(25.83 - 24.5) = 5.19$$

$$\bar{y}_{2a} = -2.77 - 1.886(23.17 - 24.5) = 0.26$$

En la figura 17.2 se muestran las medias de tratamiento no ajustadas y ajustadas con la línea de regresión calculada para cada grupo de tratamiento.

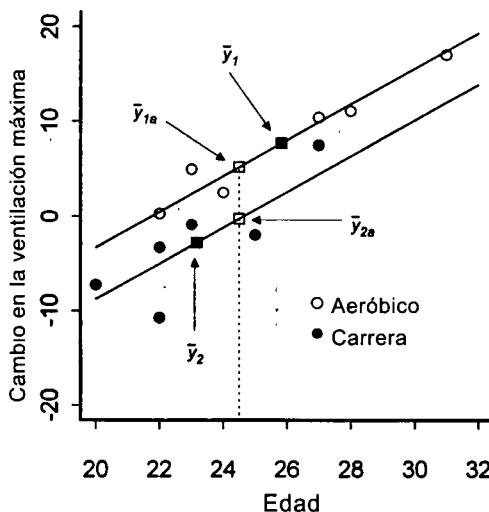


Figura 17.2 Regresión entre la ventilación máxima y la edad en el estudio de fisiología del ejercicio con medias de tratamiento ajustadas.

La diferencia entre las medias de tratamiento no ajustadas $\bar{y}_1 - \bar{y}_2 = 7.71 - (-2.77) = 10.48$ era mucho mayor que la diferencia entre las medias de tratamiento ajustadas a una edad común, $\bar{y}_{1A} - \bar{y}_{2A} = 5.19 - (-0.26) = 5.45$. Parte de la diferencia entre las medias no ajustadas se debe a la diferencia de más de dos años en los promedios de edad de los sujetos, en los dos grupos de tratamiento, $\bar{x}_1 = 25.83$ y $\bar{x}_2 = 23.17$. Las medias ajustadas son las estimaciones del cambio medio de ventilación máxima a una edad común; de esta manera, la diferencia entre las medias ajustadas refleja sólo el efecto del ejercicio en el cambio de la tasa de ventilación máxima.

Errores estándar para las medias de tratamiento ajustadas

Dos cantidades útiles para el cálculo de los errores estándar y la eficiencia relativa en el análisis de covarianza son las sumas de cuadrados para los tratamientos, T_{xx} , y el error experimental, E_{xx} , del análisis de varianza para la covariada edad, x . Estas sumas de cuadrados y sus valores para este estudio son:

$$T_{xx} = r \sum_{i=1}^t (\bar{x}_{i..} - \bar{x}..)^2 = 21.31$$

y

$$E_{xx} = \sum_{i=1}^r \sum_{j=1}^{t_i} (x_{ij} - \bar{x}_{i..})^2 = 89.67$$

El estimador del error estándar para una media de tratamiento ajustada es:

$$s_{\bar{y}_{ia}} = \sqrt{CME \left[\frac{1}{r_i} + \frac{(\bar{x}_i - \bar{x})^2}{E_{xx}} \right]} \quad (17.7)$$

El error estándar estimado para la media ajustada del grupo de ejercicio aeróbico es:

$$s_{\bar{y}_{1a}} = \sqrt{7.82 \left[\frac{1}{6} + \frac{(25.83 - 24.50)^2}{89.67} \right]} = 1.21$$

La estimación del error estándar para la media ajustada del grupo de carrera será la misma que para la media del grupo de ejercicio, ya que la cantidad $(\bar{x}_i - \bar{x})^2$ en la ecuación (17.7) es la misma para ambos.

El estimador del error estándar para la diferencia entre dos medias de tratamiento ajustadas no siempre está disponible en los programas de cómputo, sin algo de programación por parte del usuario. Se calcula como:

$$s_{(\bar{y}_{1a} - \bar{y}_{2a})} = \sqrt{CME \left[\frac{1}{r_i} + \frac{1}{r_j} + \frac{(\bar{x}_i - \bar{x}_j)^2}{E_{xx}} \right]} \quad (17.8)$$

La diferencia entre las medias de tratamiento ajustadas para los grupos de ejercicio aeróbico y carrera es $\bar{y}_{1a} - \bar{y}_{2a} = 5.19 - (-0.26) = 5.45$, con una estimación del error estándar de:

$$s_{(\bar{y}_{1a} - \bar{y}_{2a})} = \sqrt{7.82 \left[\frac{1}{6} + \frac{1}{6} + \frac{(25.83 - 23.17)^2}{89.67} \right]} = 1.80$$

Aun cuando todas las r_i son iguales, el error estándar de la diferencia varía entre los pares de tratamientos con más de dos grupos de tratamiento debido al término $(\bar{x}_i - \bar{x}_j)$ de la ecuación (17.8); en la práctica, la variación de la estimación es ligera. Un solo error estándar promedio de las diferencias sugerido por Finney (1946) para simplificar el análisis con números de réplicas iguales es:

$$s_{(\bar{y}_{1a} - \bar{y}_{2a})} = \sqrt{\frac{2CME}{r} \left[1 + \frac{T_{xx}}{(t-1)E_{xx}} \right]} \quad (17.9)$$

La sustitución de T_{xx} en lugar de $(\bar{x}_i - \bar{x}_j)^2$ en la ecuación (17.7) proporciona el error estándar promedio para las medias ajustadas.

¿Se obtuvo mayor eficiencia con la covariada?

Que el esfuerzo necesario para el ajuste de la covariada valga la pena depende del aumento en la eficiencia de las medias estimadas. La eficiencia del ajuste de la covarianza con respecto al análisis, sin el ajuste, se basa en la razón de las respectivas varianzas para las estimaciones de las medias de tratamiento. La estimación de la varianza del error experimental a partir del modelo reducido sin la covarianza es $CMC_r = SCE_r/(N - t) = 389.3/10 = 38.93$. La varianza promedio sugerida por

Finney (1946) se puede usar para la estimación con el ajuste de covarianza. La eficiencia se calcula como:

$$E = \frac{CME_r}{CME \left[1 + \frac{T_{xx}}{(t-1)E_{xx}} \right]} \quad (17.10)$$

Dados $CMM_r = 38.93$, $T_{xx} = 21.33$, $E_{xx} = 89.67$, $CME = 7.82$ y $t = 2$, la eficiencia del ajuste de covarianza es $E = 38.93/9.68 = 4.0$. Así, sin este ajuste para la edad, se requeriría del cuádruple de sujetos en el estudio de ejercicio para lograr la misma precisión en las medias de tratamiento estimadas.

Suposiciones críticas para un análisis de covarianza válido

La validez de las inferencias a partir del análisis de varianza requiere una suposición de errores independientes y homogéneos con distribución normal. Una evaluación de las suposiciones con respecto a los errores experimentales homogéneos y con distribución normal se puede lograr mediante una estimación de los residuales para cada observación como $\hat{\epsilon}_y = y_{ij} - \bar{y}_{\cdot i} - \hat{\beta}(x_{ij} - \bar{x}_{\cdot i})$. Muchos programas de cómputo proporcionan las estimaciones de los residuales que pueden emplearse en gráficas normales y de residuales contra los valores estimados para evaluar las suposiciones, como se describe en el capítulo 4.

Otras suposiciones críticas para la validez de las inferencias, a partir del análisis de covarianza, son: 1) los tratamientos no afectan a la covariada x , 2) existe una relación lineal entre la variable de respuesta y y la covariada x y 3) el coeficiente de regresión β es el mismo para todos los grupos de tratamiento.

¿Afectan los tratamientos a la covariada?

Cuando los tratamientos afectan a la covariada x al igual que a la variable de respuesta principal y , la respuesta resultante es multivariada y el ajuste de covarianza para las medias de tratamiento no es adecuado. En estos casos, debe realizarse un análisis de la respuesta bivariada (x, y) mediante los métodos de análisis multivariado. El ajuste de covarianza es apropiado, si se mide antes de la administración del tratamiento, ya que así los tratamientos todavía no han tenido oportunidad de afectar su valor. Si la covariada se mide junto con la variable de respuesta, entonces debe decidirse si pudo quedar afectada por los tratamientos, antes de considerar el ajuste.

¿Todos los tratamientos tienen el mismo coeficiente de regresión?

Las comparaciones entre las medias de tratamientos ajustadas son independientes del valor de la covariada si las líneas de regresión para los grupos de tratamiento son paralelas. Si la relación entre y y x difiere entre los grupos de tratamiento como se muestra en la figura 17.3, entonces las diferencias entre las medias de tratamiento ajustadas depende de manera crucial del nivel de x que se elige para el ajuste.

La heterogeneidad de los coeficientes de regresión entre los grupos de tratamiento se parece a la interacción entre los factores de un diseño de tratamiento factorial estándar.

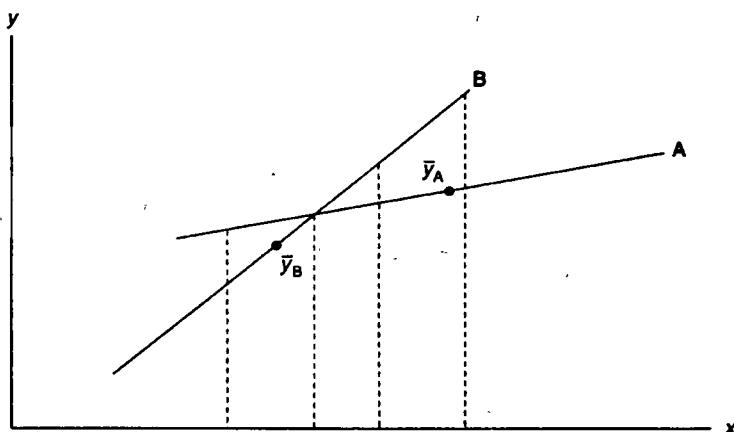


Figura 17.3 Relaciones de regresión diferentes entre la variable de respuesta y y la covariada x .

En este caso, se hacen comparaciones con los efectos simples de un factor en cada nivel de los otros factores. Debe usarse una estrategia similar, en el caso de covariadas continuas con coeficientes de regresión distintos, para los grupos de tratamiento.

Con las relaciones no lineales existen dificultades similares; en estas circunstancias, las inferencias respecto a las respuestas deben incluir una descripción completa que incluya los efectos de los tratamientos y de la covariada.

Evaluación del modelo de regresión separado

Modelo lineal con regresiones separadas para cada tratamiento

El modelo lineal con coeficientes de regresión diferentes para cada grupo de tratamiento es:

$$y_{ij} = \mu_i + \beta_i(x_{ij} - \bar{x}_i) + e_{ij} \quad (17.11)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde β_i es el coeficiente de regresión para el i -ésimo grupo de tratamiento. Se puede derivar una prueba para la igualdad de los coeficientes de regresión, $H_0: \beta_1 = \beta_2 = \dots = \beta_t$, a partir del análisis de los dos modelos alternativos, el modelo con una regresión común para todos los tratamientos y el modelo con regresiones separadas, que implican la presencia de interacción entre los tratamientos y la covariada; entonces, el modelo alterno se puede escribir de manera que incluya un término para la interacción entre la covariada y los tratamientos.

Ejemplo 17.2 Definición auditiva y entorno cultural

Hendrix, Carter y Scott (1982) describieron un estudio realizado para determinar la diferencia en la capacidad para definir, de manera auditiva los sonidos del entorno, respecto a varios factores. El estudio se diseñó para probar los efectos de un tratamiento sobre la definición auditiva. Los sujetos, pertenecientes a dos grupos culturales diferentes, participaron en pruebas previas y posteriores al tratamiento.

Sólo se usa una porción de los datos de este ejemplo para explicar los efectos de coeficientes de regresión heterogéneos para la covariada. Se realizará el análisis para determinar si el aumento en la calificación auditiva entre las pruebas previas y posteriores a la administración del tratamiento difiere entre los sujetos, con la calificación de la prueba previa como covariada. Los datos del aumento en calificaciones y las calificaciones de la prueba previa para sujetos mujeres en el grupo de tratamiento auditivo se muestran en la tabla 17.3.

Tabla 17.3 Calificaciones previas e incrementadas para la definición auditiva en dos grupos culturales

Cultura	Prueba previa	Ganancia	Cultura	Prueba previa	Ganancia
1	64	7	2	52	2
	39	32		50	12
	69	2		59	6
	56	20		42	10
	67	4		62	1
	39	26		35	9
	32	34		41	6
	62	8		36	8
	64	4		37	8
	66	2		64	6
Media	55.8	13.6		47.8	6.8

Partición de suma de cuadrados para la homogeneidad de las regresiones

Dado el modelo $y_{ij} = \mu_i + \beta(x_{ij} - \bar{x}_i) + e_{ij}$ en la ecuación (17.1), con la suposición de una pendiente común, como el modelo 1 con los $t+1$ parámetros usuales, μ_i y β . El modelo $y_{ij} = \mu_i + \beta_i(x_{ij} - \bar{x}_i) + e_{ij}$ se define en la ecuación (17.11), con la suposición de pendientes diferentes para cada tratamiento, como el modelo 2 con $2t$ parámetros, μ_i y β_i . La suma de cuadrados para el error experimental del modelo 1, digamos SCE_1 , tendrá $(N - t - 1)$ grados de libertad, y la suma de cuadrados para el error experimental del modelo 2, digamos SCE_2 , tendrá $(N - 2t)$ grados de libertad.

El análisis de varianza para los datos de definición auditiva con el modelo 1 en la tabla 17.4 contiene $SCE_1 = 519.38$, con 17 grados de libertad. El análisis de varianza para el modelo 2 se muestra en la tabla 17.5 con 2 grados de libertad para la reducción de la suma de cuadrados de regresión separada, calculada dentro de

cada grupo cultural. La suma de cuadrados del error se redujo a $SCE_2 = 147.22$, con 16 grados de libertad.

Tabla 17.4 Análisis de covarianza del aumento en calificaciones con covariada con calificación previa a la prueba del estudio de definición auditiva

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Regresión	1	1 061.12	1 061.12	34.73	.000
Grupo	1	647.47	647.47	21.19	.000
Error	17	519.38*	30.55		

* SCE_1

Tabla 17.5 Análisis de covarianza del aumento en calificación suponiendo regresiones separadas del aumento en las calificaciones previas a la prueba para cada grupo cultural

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Regresión	1	610.24	610.24	66.33	.000
Regresión dentro del grupo	2	1 433.28	716.64	77.89	.000
Error	16	147.22*	9.20		

* SCE_2

La suma de cuadrados para probar la homogeneidad de los coeficientes de regresión para los grupos de tratamiento es la diferencia en la suma de cuadrados del error experimental para los dos modelos, es decir:

$$\begin{aligned} SC(\text{Homogeneidad}) &= SCE_1 - SCE_2 \\ &= 519.38 - 147.22 \\ &= 372.16 \end{aligned} \quad (17.12)$$

con $(N - t - 1) - (N - 2t) = (t - 1)$ o $17 - 16 = 1$ grado de libertad. El estadístico F_0 para probar la hipótesis nula de coeficientes de regresión iguales, $H_0: \beta_1 = \beta_2 = \dots = \beta_t$ es:

$$F_0 = \frac{CM(\text{Homogeneidad})}{ECM_2} \quad (17.13)$$

con un valor crítico $F_{\alpha,(t-1),(N-2t)}$. La prueba para el estudio de ejercicio es $F_0 = (372.16/1)/9.20 = 40.45$, que excede a $F_{.05,1,16} = 4.49$, y se rechaza la hipótesis nula de coeficientes de regresión iguales para los dos grupos culturales. La regresión del aumento en las calificaciones previas es diferente para los dos grupos culturales y no es apropiado el ajuste de las medias de los grupos culturales con la misma calificación previa de una regresión común.

Muchos programas estadísticos tienen la capacidad de calcular directamente las sumas de cuadrados necesarias para la homogeneidad de los coeficientes de regresión en la ecuación (17.12), cuando se especifica como un efecto de interacción de *tratamiento por covariada* en el programa. En la tabla 17.6 se muestra el análisis de varianza para el modelo de covarianza con un tratamiento mediante un término de interacción covariada. Observe que la suma de cuadrados para la interacción grupos \times regresión equivale a la suma de cuadrados para la homogeneidad, ecuación (17.12), derivada de los modelos alternativos con regresiones comunes y diferentes para los grupos de tratamiento.

Tabla 17.6 Análisis de covarianza para el aumento en calificaciones, que incluye la interacción entre la covariada de calificación previa y los grupos culturales

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr > F
Grupo	1	610.24	610.24	66.32	.000
Regresión	1	770.63	770.63	83.75	.000
Grupos \times regresión	1	372.16	372.16	40.45	.000
Error	16	147.22	9.20		

Coeficiente de regresión estimado para cada tratamiento

Si las regresiones son diferentes de manera significativa entre los grupos de tratamiento, la estimación de mínimos cuadrados del coeficiente de regresión para el i -ésimo grupo de tratamiento se calcula a partir de los valores x_{ij} y y_{ij} dentro de los grupos de tratamiento como:

$$\hat{\beta}_i = \frac{\sum_{j=1}^r (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)}{\sum_{j=1}^r (x_{ij} - \bar{x}_i)^2} \quad (17.14)$$

La ecuación de regresión estimada para el i -ésimo grupo de tratamiento es $\hat{y}_i = \bar{y}_i + \hat{\beta}_i(x_{ij} - \bar{x}_i)$. Las estimaciones de los coeficientes de regresión para los dos grupos culturales y sus errores estándar (entre paréntesis) son $\hat{\beta}_1 = -0.91(0.07)$ y $\hat{\beta}_2 = -0.16(0.09)$. Ambas estimaciones están disponibles en los programas de cómputo para modelos lineales. La estimación del error estándar para el coeficiente de regresión estimado, $\hat{\beta}_i$, es la raíz cuadrada de $CME/\sum_{j=1}^r (x_{ij} - \bar{x}_i)^2$.

Las ecuaciones de regresión estimadas para los dos grupos culturales son:

$$\text{Grupo 1: } \hat{y}_1 = 55.8 - 0.91(x - 13.6)$$

$$\text{Grupo 2: } \hat{y}_2 = 47.8 - 0.16(x - 6.8)$$

Las ecuaciones de regresión se graficaron en la figura 17.4 junto con las observaciones para cada grupo cultural. Está claro que las líneas de regresión son

diferentes y no tendría sentido una comparación de las medias de los grupos culturales con la media global previa a la prueba de $\bar{x} = 51.8$.

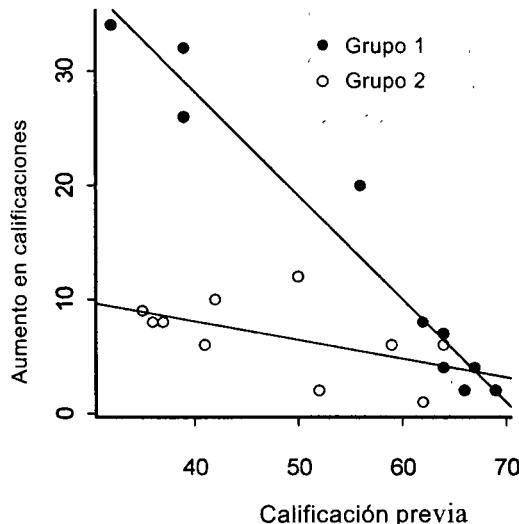


Figura 17.4 Regresión entre el aumento en calificaciones y las calificaciones previas con regresiones separadas para los grupos culturales.

En su lugar, se consideran las líneas como estimaciones del efecto simple de las calificaciones previas en cada nivel de grupo cultural y se hacen inferencias basadas en ello. En ambos grupos culturales, el aumento en las calificaciones fue menor con calificaciones previas altas, pero la tasa de disminución era mucho mayor en el grupo 1. De hecho, una prueba t de Student simple para $H_0: \beta_2 = 0$ es $t_0 = -0.16/0.09 = -1.78$, que lleva a que no se rechace la hipótesis nula; entonces, no existe evidencia convincente que indique cambios en el aumento con las calificaciones previas en el grupo 2.

Es obvio que el aumento en las calificaciones del tratamiento en el grupo 1 es mucho mayor cuando los sujetos tienen baja definición auditiva y parecen beneficiarse más con la capacitación, que los del grupo 2 con definición auditiva inicial baja. Cuando los sujetos de cualquier grupo tienen una elevada definición auditiva inicial previa, la ganancia por la capacitación fue más o menos baja o despreciable.

Los intervalos de confianza o pruebas de hipótesis entre las medias estimadas de los grupos para cualquier valor dado de calificación previa, digamos $x = x_0$, $\hat{y}_1|x_0 - \hat{y}_2|x_0$ se puede construir con la varianza del contraste, como sigue:

$$s_{\hat{y}_1|x_0 - \hat{y}_2|x_0}^2 = CME \left[\frac{1}{r_1} + \frac{1}{r_2} + \frac{(\bar{x}_0 - \bar{x}_1)^2}{\sum (x_{1j} - \bar{x}_{1.})^2} + \frac{(\bar{x}_0 - \bar{x}_2)^2}{\sum (x_{2j} - \bar{x}_{2.})^2} \right] \quad (17.15)$$

17.3 Análisis de covarianza para diseños de experimentos bloqueados

El análisis de covarianza se puede aplicar a cualquier diseño de experimento con una extensión directa de los principios aplicados en la sección anterior al diseño totalmente aleatorizado.

Ejemplo 17.3 Prueba de disponibilidad de nutrientes con cebada en un invernadero

Los métodos de control en bosques y praderas con mantos acuíferos afectan el estado de los nutrientes en cualquier tipo de vegetación y suelo complejos. El conocimiento de las relaciones entre suelo; planta y nutrientes es esencial para su manejo.

La disponibilidad de ciertos nutrientes en estos suelos se evalúa mediante una técnica de cultivo en macetas dentro de un invernadero con plantas de cebada. Las plantas de prueba crecen en suelos fertilizados a un nivel óptimo y sin los nutrientes en cuestión; si hay deficiencia en los nutrientes del suelo, las plantas cultivadas con fertilizante mostrarán un mayor crecimiento que las cultivadas con los nutrientes omitidos.

Objetivo de investigación: en un estudio de este tipo, un investigador quería determinar la disponibilidad de nitrógeno y fósforo en un campo dominado por vegetación de chaparral. Recolectó muestras de suelo bajo los matorrales de la vegetación dominante en la pradera, en la falda del monte y muestras compuestas para una evaluación mediante cultivo en macetas, con disponibilidad de nitrógeno y fósforo.

Diseño de tratamientos: se usaron cuatro nutrientes para el estudio: 1) como está, sin agregar fertilizante, 2) completo, con fertilizante completo, 3) N_0 , nitrógeno omitido y 4) P_0 , fósforo omitido. Los tratamientos de nutrientes se agregaron como soluciones al suelo, mezclado y colocado en macetas de plástico en el invernadero.

Diseño del experimento: las macetas de tratamiento se colocaron en una banca del invernadero en un diseño de bloques completo aleatorizado para controlar la variación del error experimental, causada por los gradientes de luz y temperatura dentro del invernadero.

Se cultivaron las plantas de cebada en las macetas durante siete semanas, momento en que las plantas se cosecharon, secaron y pesaron. Una plaga de hoja afectó el crecimiento de las plantas a medio experimento. Se supuso que la plaga afectaría el crecimiento y al final del experimento se midió el porcentaje del área de la hoja afectada, en cada contenedor, antes de la cosecha. El peso total en seco



de las plantas de cebada y el porcentaje del área de la hoja afectada con la plaga se muestran en la tabla 17.7, para cada maceta.

Tabla 17.7 Materia seca total y en gramos y porcentaje del área de la hoja con plaga x en plantas de cebada

Bloque	Tratamiento							
	Como está		Completo		N_0		P_0	
	y	x	y	x	y	x	y	x
1	23.1	13	30.1	7	26.4	10	26.2	8
2	20.9	12	31.8	5	27.2	9	25.3	9
3	28.3	7	32.4	6	28.6	6	29.7	7
4	25.0	9	30.6	7	28.5	6	26.0	7
5	25.1	8	27.5	9	30.8	5	24.9	9
Media	24.48	9.8	30.48	6.8	28.3	7.2	26.42	8.0

Fuente: Dr. J. Kemmedson, *Renewable Natural Resources*, University of Arizona.

Modelo lineal para un diseño de bloque completo aleatorizado

El modelo de los efectos para el diseño de bloques completo aleatorizado se expresa como:

$$y_{ij} = \mu + \tau_i + \rho_j + \beta(x_{ij} - \bar{x}_.) + e_{ij} \quad (17.15)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

donde μ es la media general, τ_i es el efecto de tratamiento, ρ_j es el efecto de bloque, β es la regresión de y sobre x y e_{ij} son errores aleatorios independientes con distribución normal, media 0 y varianza σ^2 . Además, se supone que los tratamientos o bloques no afectan a la covariada, y que la regresión es la misma para todos los tratamientos.

Modelos alternos para evaluar la influencia de la covariada

Si es necesario se usan otros modelos completos y reducidos para evaluar la influencia de la covariada y también la significancia de los efectos de tratamiento después del ajuste de covariada. Los modelos necesarios y sus sumas de cuadrados del error experimental son:

- el modelo completo, $y_{ij} = \mu + \tau_i + \rho_j + \beta(x_{ij} - \bar{x}_.) + e_{ij}$, y SCE_f con $(r - 1)(t - 1) - 1$ grados de libertad.
- el modelo usual de bloque completo aleatorizado sin covariada, $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$, y SCE_r con $(r - 1)(t - 1)$ grados de libertad.

- sin efectos de tratamiento, $y_{ij} = \mu + \rho_j + \beta(x_{ij} - \bar{x}_{..}) + e_{ij}$, y SCE_r^* con $(N - r - 1)$ grados de libertad.

Por último, aunque no es totalmente necesario para el análisis, la suma de cuadrados ajustada por bloques se puede calcular usando el modelo con:

- sin efectos de bloques, $y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + e_{ij}$, y SCE_r^{**} con $(N - t - 1)$ grados de libertad.

La SCE para cada modelo, ajustada a los datos de cebada (tabla 17.7) es $SCE_f = 12.577$, $SCE_r = 39.167$, $SCE_r^* = 37.437$ y $SCE_r^{**} = 20.363$.

Particiones de la suma de cuadrados para el análisis de covarianza

La reducción de la suma de cuadrados, después de agregar la covariada x al modelo de bloques completo aleatorizado usual, se obtiene como la diferencia:

$$\begin{aligned} SC(\text{covariada}) &= SCE_r - SCE_f \\ &= 39.167 - 12.577 = 26.590 \end{aligned}$$

con un grado de libertad. La suma de cuadrados del tratamiento ajustada, después de incluir la covariada y los efectos de bloque es:

$$\begin{aligned} SCT(\text{ajustada}) &= SCE_r^* - SCE_f \\ &= 37.437 - 12.577 = 24.860 \end{aligned}$$

con $(t - 1) = 3$ grados de libertad. La suma de cuadrados del bloque ajustada, después de incluir la covariada y los efectos de tratamiento es:

$$\begin{aligned} SCB(\text{ajustada}) &= SCE_r^{**} - SCE_f \\ &= 20.363 - 12.577 = 7.786 \end{aligned}$$

El análisis de covarianza para la prueba de disponibilidad de nutrientes con cebada se muestra en la tabla 17.8.

La significancia de la covariada requiere una prueba de la hipótesis nula $H_0: \beta = 0$ con el estadístico:

Tabla 17.8 Análisis de covarianza para la producción de materia seca de plantas de cebada con el porcentaje del área de la hoja dañada por plaga como covariada

Fuente de variación	Grados de libertad	Sumas de cuadrados	Cuadrados medios	F	Pr > F
Regresión	1	26.590	26.590	23.263	0.001
Bloque	4	7.786	1.947	1.703	0.219
Tratamiento	3	24.860	8.287	7.250	0.006
Error	11	12.577	1.143		

$$F_0 = \frac{CM(\text{Covariada})}{CME} = \frac{26.590}{1.143} = 23.263$$

que es significativa con $Pr > F = .001$ en la tabla 17.8. Entonces, la relación entre el porcentaje del área de hoja dañada por la plaga y la producción de materia seca de plantas de cebada es significativa.

La estimación del coeficiente de regresión, que proporcionan la mayoría de los programas de cómputo, es $\hat{\beta} = -0.863$. El coeficiente negativo indica que la producción de materia seca disminuye cuando aumenta la incidencia de la plaga en la planta.

La hipótesis nula de que no hay diferencia entre las medias de tratamiento se prueba con el estadístico:

$$F_0 = \frac{CMT(\text{ajustada})}{CME} = \frac{8.287}{1.143} = 7.25$$

y la hipótesis nula se rechaza con $Pr > F = .006$ en la tabla 17.8.

Medias de tratamiento ajustadas y sus errores estándar

Las medias de tratamiento ajustadas se calculan de la misma manera que para el diseño totalmente aleatorizado con la ecuación (17.4) y se muestra en la tabla 17.9 junto con sus errores estándar. Los errores estándar mostrados en las ecuaciones (17.7) y (17.9) se pueden usar para las medias de tratamiento ajustadas con las sumas de cuadrados para los tratamientos, $T_{xx} = 26.55$, y el error, $E_{xx} = 35.70$, del análisis de varianza para la covariada, la infección por plaga, usando los datos de la tabla 17.7. El error estándar promedio estimado de la diferencia entre dos medias de tratamiento ajustadas es:

$$s_{(\bar{y}_{ia} - \bar{y}_{Ja})} = \sqrt{\frac{2(1.143)}{5} \left[1 + \frac{26.55}{(3)35.7} \right]} = 0.76 \quad (17.16)$$

Interpretaciones con contrastes múltiples

Se puede usar el intervalo de confianza del 95% de Dunnett de un lado con el tratamiento “completo” como control para determinar si el suelo era deficiente en nutrientes.

Tabla 17.9 Medias ajustadas y sus errores estándar a partir del análisis de covarianza para la producción de materia seca de las plantas de cebada, con el porcentaje de daño en el área de la hoja como covariada

Tratamiento	Media ajustada	Error estándar
Como está	26.08	0.58
Completo	29.49	0.52
N_0	27.65	0.50
P_0	26.46	0.48

Si el suelo es deficiente, el crecimiento de la planta con el tratamiento completo o de control excede el de los tratamientos sin uno o más de los nutrientes; por lo tanto, si el resultado de la cota superior del intervalo para el tratamiento *menos* la de control es negativo, el tratamiento es deficiente en nutrientes.

De la tabla VI del apéndice, el valor crítico del estadístico de Dunnett para un intervalo de un lado es $d_{05,3,11} = 2.31$. El criterio de Dunnett es $D(3,.05) = 2.31(0.76) = 1.76$. Las cotas superiores de un ICS del 95% para las diferencias entre los tratamientos y el tratamiento de control completo, se muestran en la tabla 17.10.

Tabla 17.10 Cotaz superiores de intervalos de confianza simultáneos del 95% con el método de Dunnett para comparar el tratamiento de control (completo) con otros tratamientos

Tratamiento	Media ajustada (\bar{y}_i)	($\bar{y}_i - \bar{y}_c$)	Cotas superiores de un ICS del 95%
Como está	26.08	- 3.41	- 1.65
Completo	$\bar{y}_c = 29.49$	-	-
N_0	27.65	- 1.84	- 0.08
P_0	26.46	- 3.03	- 1.27

La diferencia entre el tratamiento sin añadir nutrientes y el tratamiento de control es $26.08 - 29.29 = - 3.41$, con una cota superior, -1.65, en el intervalo de confianza de un lado. Así, el suelo es deficiente en algún nutriente no especificado. Las pruebas concretas para las deficiencias en nitrógeno y fósforo requieren comparaciones entre el tratamiento completo y los tratamientos N_0 y P_0 . La diferencia entre el tratamiento deficiente en nitrógeno, N_0 , y el de control es $27.65 - 29.49 = - 1.84$, con cota superior de -0.08. La diferencia entre el tratamiento deficiente en fósforo, P_0 , y el de control es $26.46 - 29.49 = - 3.03$ con cota superior de -1.27. Aunque el suelo era deficiente tanto en nitrógeno como fósforo, la cota superior para la comparación del fósforo está más lejos de 0 que la de la comparación del nitrógeno, lo que indica una mayor deficiencia de fósforo que de nitrógeno.

17.4 Consecuencias prácticas del análisis de covarianza

La aplicación práctica del análisis de covarianza se explica nada más con diseños de bloques totalmente aleatorizados y completos aleatorizados, pero su uso puede extenderse a cualquier diseño de experimentos y tratamientos, al igual que a estudios por observación de estructuras complejas y estudios que requieren el uso de covariadas múltiples para el ajuste. El objetivo de este capítulo fue introducir las ideas básicas que fundamentan el uso de información adicional en las unidades básicas del estudio. Las fórmulas manuales específicas para el análisis dependerán del diseño particular del experimento y los tratamientos empleados para el estudio. No obstante, en todos los casos, el uso de los modelos completo y reducido, como se explicó con los dos diseños de este capítulo, permiten evaluar la influencia de la covariada en la reducción del error experimental y la significancia de las medias de tratamiento ajustadas.

Se proporcionaron presentaciones amplias de los usos y malos usos de las covariadas en estudios de investigación en dos ediciones especiales de *Biometrics* (1957), volumen 13, número 3 y (1982), volumen 38, número 3. De interés particular son los artículos de Cochran (1957), Smith (1957) y Cox y McCullagh (1982). Surgieron varios aspectos relativos al uso de covariadas, entre ellos la posibilidad de aplicarlas en ciertas situaciones y la relación entre los bloques y las covariadas.

El análisis de covarianza combina las características de dos modelos

El análisis de covarianza combina las características de varios modelos para el análisis de varianza y regresión para hacer la partición de la variación total en componentes atribuibles a: 1) efectos de tratamiento, 2) cualquiera de las covariadas y 3) el error experimental aleatorio y la variación asociada con cualquier diseño de factores de bloque. La intención básica es comparar tratamientos con un valor común para la covariada.

Casos en que las covariadas son superiores a los bloques para controlar el error

En apariencia, el análisis de covarianza ofrece una alternativa al uso de bloques para reducir el error experimental. Los diseños de bloque restringen el número de criterios que es razonable usar para el control local. Por otro lado, la covarianza permite usar cualquier número de factores que sean necesarios; también utiliza mejor los valores exactos de los factores cuantitativos, mientras que el bloqueo agrupa los mismos factores en clases de valores. La ventaja de la covariada parece grande cuando existen más factores como candidatos potenciales para variables de bloque.

Casos en que los bloques son superiores a las covariadas para controlar el error

La covarianza puede quedar en desventaja sin el uso de bloques, porque la asignación aleatoria de los tratamientos a las unidades experimentales pueden tener como resultado una distribución desigual de tratamientos entre los niveles de la covariada.

Cualquier asociación entre la covariada y los tratamientos confunde los efectos de ambos en la variable de respuesta. El uso de bloques sobre los valores de las covariadas hace una distribución pareja de los tratamientos entre los niveles de la covariada y evita confundir los efectos de tratamiento y las covariadas.

El bloque es más eficaz con los factores cualitativos que producen una variación que se reconoce entre las unidades experimentales. Estos factores incluyen el estudio de prácticas administrativas cuando varios técnicos participan en las tareas a realizar, o se tienen que llevar a cabo en días diferentes. Los lotes de materiales también proporcionan un criterio de bloque cualitativo natural y eficaz.

Siempre que haya un alto reconocimiento de las diferencias sistemáticas, el bloque es el medio más eficiente para reducir la variación del error experimental, que mantiene la ortogonalidad necesaria para evitar confundir los efectos de las covariadas con los tratamientos. En resumen, el bloque se recomienda como la primera opción para reducir la varianza del error experimental, con ajustes derivados de la información adicional y, si es necesario, con el análisis de covarianza, para mejorar más la precisión.

Los estudios comparativos por observación están en desventaja

El análisis de los estudios comparativos por observación se pueden beneficiar en forma parecida con el análisis que incluye covariadas para reducir la varianza del error y ajustar las medias de grupos para las diferencias en sus valores covariados. Los estudios por observación tienen la desventaja de que las unidades no se pueden aleatorizar para los grupos de tratamiento definidos. Existe la posibilidad de una influencia en la respuesta por parte de covariadas no observadas adicionales que se asocian con los grupos de tratamiento, con esto se introduce una incógnita sesgada en las comparaciones de los grupos. Los estudios experimentales tienen la desventaja de que los efectos de estas variables se distribuyen entre las unidades con la asignación aleatoria y es mucho menos probable que su influencia se confunda con los efectos de los tratamientos.

Peligros al extrapolalar los datos

Por último, se debe ser cauteloso al ajustar las medias de tratamiento a un valor común de la covariada. Aun cuando las regresiones sean paralelas y no haya posibilidad de que los efectos de tratamiento afecten a la covariada, los valores de esta pueden ser muy diferentes para los grupos de tratamiento. Si esos valores se separan para los grupos de tratamiento, entonces los ajustes tendrán que extrapolarse a un valor de la covariada que no es común a ningún grupo. Un ejemplo extremo para explicarlo es la situación en la que se usa el ingreso como covariada para el ajuste al comparar un grupo de ejecutivos corporativos con los trabajadores del nivel bajo. El ajuste se aplica a la región extrapolada de ingreso promedio, no incluida en ninguno de los grupos y la comparación se hace entre dos grupos en circunstancias sin sentido. Incluso si los ajustes extrapolados fueran válidos, los errores estándar de los valores extrapolados serían muy grandes.

EJERCICIOS PARA EL CAPÍTULO 17

1. Se realizó un experimento sobre la resistencia de puntos de soldadura para tres tipos de aleación de acero. Se hicieron seis soldaduras en cada aleación y se midió la fuerza necesaria para separarlas. Se midió el diámetro de la soldadura porque se pensó que afectaba su resistencia. Los datos se muestran en la siguiente tabla, donde y = resistencia de soldadura y x = diámetro de la soldadura.

Aleación	y	x	Aleación	y	x	Aleación	y	x
1	37.5	12.5	2	57.5	16.5	3	38.0	15.5
1	40.5	14.0	2	69.5	17.5	3	44.5	16.0
1	49.0	16.0	2	87.0	19.0	3	53.0	19.0
1	51.0	15.0	2	92.0	19.5	3	55.0	18.0
1	61.5	18.0	2	107.0	24.0	3	58.5	19.0
1	63.0	19.5	2	119.5	22.5	3	60.0	20.5

- a. Use el diámetro de la soldadura como covariada de la resistencia y escriba un modelo lineal para el experimento, identifique cada término y establezca las suposiciones necesarias para el modelo.
- b. Realice un análisis de covarianza y pruebe la significancia de la covariada; ajuste las medias de tratamiento.
- c. Calcule las medias ajustadas de tratamiento, sus errores estándar y un error estándar promedio de la diferencia entre las dos medias ajustadas.
- d. Grafique la recta de regresión para cada aleación y muestre las medias observadas, ajuste las medias por cada aleación.
- e. Calcule la eficiencia del ajuste covariado.
- f. Pruebe la hipótesis de regresiones homogéneas para cada aleación.
- g. Analice los resultados del experimento y la eficacia del ajuste covariado.
- h. El nivel de significancia de la prueba de regresiones homogéneas era $Pr > F = .068$. Cuando la fuente de variación para grupo \times regresión no es significativa, se obtiene una suma de cuadrados para el error en el análisis de covarianza normal si se supone un coeficiente de regresión homogéneo. En realidad, lo que se hace es agrupar la partición de la suma de cuadrados de grupo \times regresión con la suma de cuadrados del error del análisis de covarianza para el modelo con coeficientes de regresión separados, para cada grupo de tratamiento. Hendrix *et al.* (1982) sugirió que, en ese caso, debía usarse un nivel de significancia de $\alpha = .20$ a $.25$, ya que es similar al problema de probar modelos con especificaciones incompletas, como los presentados por Bozivich, Bancroft y Hartley (1956). ¿Qué piensa sobre esta estrategia?
- i. Suponga que usted se adhiere a la filosofía de un nivel de significancia de $\alpha = .20$ a $.25$ al probar la hipótesis nula de coeficientes de regresión iguales para todos los grupos de tratamiento. Entonces se rechaza la hipótesis nula del inciso f). Realice un análisis de covarianza con las estimaciones de regresión separadas para cada grupo de tratamiento y calcule los coeficientes de regresión estimados y sus errores estándar para cada aleación. ¿Cuál es su inferencia a partir del estudio en este punto?
2. Un nutriólogo realizó un experimento para evaluar los efectos de cuatro complementos vitamínicos en el aumento de peso de animales de laboratorio. El experimento se realizó con un diseño totalmente

aleatorizado, con cinco animales, en jaulas separadas, para cada tratamiento. Como el consumo de calorías difiere entre los animales e influye en el aumento de peso, el investigador lo midió en cada animal. Los datos del aumento de peso (y = gramos) y el consumo de calorías (x = calorías/10) son:

<i>Dieta</i>	<i>y</i>	<i>x</i>									
1	48	35	2	65	40	3	79	51	4	59	53
1	67	44	2	49	45	3	52	41	4	50	52
1	78	44	2	37	37	3	63	47	4	59	52
1	69	51	2	73	53	3	65	47	4	42	51
1	53	47	2	63	42	3	67	48	4	34	43

- Determine si la dieta influye el consumo de calorías al grado de que este último quede invalidado como covariada.
 - Use el consumo de calorías como covariada del aumento de peso y escriba un modelo lineal para el experimento. Identifique cada término y establezca las suposiciones del modelo.
 - Realice un análisis de covarianza, pruebe la significancia de la covariada y ajuste las medias de tratamiento.
 - Calcule las medias de tratamiento ajustadas, sus errores estándar y el error estándar promedio de la diferencia entre dos medias ajustadas.
 - Grafique la recta de regresión para cada dieta, muestre las medias observadas y ajuste las medias de cada dieta.
 - Calcule la eficiencia del ajuste de covarianza.
 - Pruebe la hipótesis de regresiones homogéneas para cada dieta.
 - Analice los resultados del experimento y la eficacia del ajuste de covarianza.
3. Un científico de plantas realizó un experimento para estudiar los efectos del nivel de irrigación por goteo en el crecimiento, cosecha y calidad del maíz dulce. Se usaron tres niveles de irrigación (15.8, 24.0 y 28.5 pulgadas de agua aplicadas), y se arregló el experimento con un diseño de bloques completo aleatorizado para controlar la variabilidad del campo. Una de las variables de respuesta medidas fue el peso del desperdicio por parcelas, o la cantidad de maíz dulce en la parcela, que era inadecuado para la venta. El número de plantas por parcela varió y esto afectaría la cosecha. Como se optimizó la humedad del suelo para establecer la mejor cosecha, los niveles de irrigación impuestos, después de establecer la cosecha, no afectaban el número de plantas por parcela. La cosecha de desperdicio observada (y = toneladas métricas/hectárea) y x = número de plantas en una sección de 40 pies de filas se muestran a continuación:

<i>Bloque</i>	<i>Nivel de irrigación</i>					
	1		2		3	
	<i>y</i>	<i>x</i>	<i>y</i>	<i>x</i>	<i>y</i>	<i>x</i>
1	1.5	45	1.9	54	1.1	43
2	3.1	58	1.8	57	1.8	60
3	3.8	61	2.9	55	3.7	71
4	3.3	59	2.3	56	1.8	48

- a. Use el número de plantas en 40 pies de fila como covariada para la cosecha de desperdicio, escriba un modelo lineal para el experimento, identifique cada término y establezca las suposiciones necesarias para el modelo.
 - b. Realice el análisis de covarianza, pruebe la significancia de la covariada y ajuste las medias de tratamiento.
 - c. Muestre el valor del coeficiente de regresión. Calcule las medias de tratamiento ajustadas, sus errores estándar y el error estándar promedio de la diferencia entre dos medias ajustadas.
 - d. Calcule la eficiencia del ajuste de covarianza.
 - e. Analice los resultados del experimento y la eficacia del ajuste covariado.
4. Se realizó un experimento con un diseño de bloque completo aleatorizado para estudiar el efecto del control natural, *bacillus*, y un insecticida químico estándar para el control de una plaga de gusanos en un plantío. Los tratamientos incluyeron cuatro fuentes de *bacillus* (tratamientos 1-4), un tratamiento químico estándar (tratamiento 5) y un control sin tratamiento (tratamiento 6). Los tratamientos se aplicaron a plantas crecidas en campos de cultivo. Se contó el número de gusanos (*cuenta*) en cada planta antes del tratamiento. Se contó el número de gusanos vivos (*vivos*) 20 horas después de aplicar los tratamientos. A continuación se encuentran los datos de cada parcela:

Tratamiento	Bloque							
	1		2		3		4	
	Cuenta	Vivos	Cuenta	Vivos	Cuenta	Vivos	Cuenta	Vivos
<i>Bacillus 1</i>	15	17	25	26	18	21	23	26
<i>Bacillus 2</i>	19	18	21	22	20	19	9	20
<i>Bacillus 3</i>	19	19	19	21	21	23	25	22
<i>Bacillus 4</i>	22	14	31	26	17	17	19	19
<i>Químico</i>	17	5	22	6	26	13	18	10
<i>Control</i>	22	25	14	19	22	26	23	27

- a. Use la cuenta de gusanos antes del tratamiento como covariada del número de gusanos vivos 20 horas después del tratamiento, escriba un modelo lineal para el experimento, identifique cada término y establezca las suposiciones necesarias para el modelo.
 - b. Realice el análisis de covarianza, pruebe la significancia de la covariada y ajuste las medias de tratamiento.
 - c. Muestre el valor del coeficiente de regresión. Calcule las medias de tratamiento ajustadas, sus errores estándar y el error estándar promedio de la diferencia entre las dos medias ajustadas.
 - d. Calcule la eficiencia del ajuste covariado.
 - e. Analice los resultados del experimento y la eficacia del ajuste covariado
 - f. La variable de respuesta es una medida de la cuenta y es posible que no tenga una distribución normal. ¿Verificó las suposiciones de varianza homogénea y distribuciones normales de los errores experimentales en el modelo? Si no fue así, hágalo ahora; si es necesario, tome las acciones correctivas (según la presentación del capítulo 4) y répita los incisos b) y e).
5. El análisis de covarianza se puede usar para estimar valores faltantes en un diseño de bloques y proporcionar una estimación no sesgada de la suma de cuadrados del tratamiento para probar la hipótesis de que no hay diferencia entre las medias de tratamiento con un valor faltante (Coons, 1957).

Se introduce una covariada x para el valor faltante. Se asigna el valor $x = -1$ a la covariada del valor y faltante y $x = 0$ a los otros valores de y . Se asigna el valor $y = 0$ al dato que falta. Calcule un análisis de covarianza normal con los valores asignados a y y a la covariada x . La estimación del valor faltante es la del coeficiente de regresión $\beta = E_{xy}/E_{xx}$, y la suma de cuadrados del tratamiento ajustada en el análisis de covarianza es la suma de cuadrados correcta para probar la hipótesis nula de que no hay diferencia entre los tratamientos.

Como demostración de la técnica, utilice los datos del diseño de bloques completo aleatorizado del ejercicio 17.3 e ignore la covariada, x = número de plantas, para este ejercicio. Suponga que la observación $y = 2.9$ en el nivel de irrigación 2 del bloque 3 falta y asignele el valor $y = 0$. Construya la nueva covariada con $x = -1$ para el valor faltante y $x = 0$ para el resto de los valores de y , y realice el análisis de covarianza. Estime el valor faltante y use los cuadrados medios ajustados de los tratamientos para probar la hipótesis de que no hay diferencia entre las medias de tratamiento.

6. Describa un estudio en su campo donde se use la covariada, además o en lugar de un criterio de bloque para reducir la varianza del error y ajustar las medias de tratamiento. Proporcione una justificación para el uso de los valores covariados en el modelo estadístico, en lugar de bloques de acuerdo con la covariada, justifique su uso basado en que los tratamientos no la afectan. Si es posible, utilice un ejemplo de un artículo del periódico o su propia experiencia de investigación.

Referencias

- Abeyasekera, S., and Curnow, R. N. (1984). "The desirability of adjusting for residual effects in a crossover design." *Biometrics* 40, 1071–1078.
- Addelman, S. (1970). "Variability of treatments and experimental units in the design and analysis of experiments." *Journal of the American Statistical Association* 65, 1095–1108.
- Anderson, R. L. (1960). "Uses of variance component analysis in the interpretation of biological experiments." *Bulletin of the International Statistical Institute* 37, 71–90.
- Anderson, V. L., and McLean, R. A. (1974). *Design of experiments: A realistic approach*. New York: Marcel Dekker.
- Anscombe, F. J. (1948). "The transformation of Poisson, binomial, and negative-binomial data." *Biometrika* 35, 246–254.
- Bailey, R. A. (1986). "Randomization, constrained." In *Encyclopedia of Statistical Sciences*, vol. 7, 524–530. S. Kotz and N. L. Johnson (eds.). New York: Wiley.
- Bailey, R. A. (1987). "Restricted randomization. A practical example." *Journal of the American Statistical Association* 82, 712–719.
- Bainbridge, T. R. (1963). "Staggered, nested designs for estimating variance components." *ASQC Annual Convention Transactions*, 93–103.
- Bartlett, M. S. (1947). "The use of transformations." *Biometrics* 3, 39–52.
- Bennett, C. A., and Franklin, N. L. (1954). *Statistical analysis in chemistry and the chemical industry*. New York: Wiley.
- Berk, K. N., and Picard, R. R. (1991). "Significance tests for saturated orthogonal arrays." *Journal of Quality Technology* 23, 79–89.
- Bicking, C. A. (1954). "Some uses of statistics in the planning of experiments." *Industrial Quality Control* 10, 20–24.
- Biometrics (1957). "Special Issue on the Analysis of Covariance." *Biometrics* 13, 261–405.
- Biometrics (1982). "Special Issue: Analysis of Covariance." *Biometrics* 38, 540–753.

- Boik, R. J. (1981). "A priori tests in repeated measures designs: Effects of non-sphericity." *Psychometrika* 46, 241-255.
- Bose, R. C., Clatworthy, W. H., and Shrikhande, S. S. (1954). "Tables of partially balanced designs with two associate classes." North Carolina Agricultural Experiment Station. *Technical Bulletin*, no. 107.
- Bose, R. C., and Nair, K. R. (1939). "Partially balanced incomplete block designs." *Sankhya* 4, 337-372.
- Box, G. E. P. (1954a). "Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification." *Annals of Mathematical Statistics* 25, 290-302.
- Box, G. E. P. (1954b). "Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effect of inequality of variances and correlation between errors in the two-way classification." *Annals of Mathematical Statistics* 25, 484-498.
- Box, G. E. P., and Behnken, D. W. (1960). "Some new three level designs for the study of quantitative variables." *Technometrics* 2, 455-476.
- Box, G. E. P., and Cox, D. R. (1964). "An analysis of transformations." *Journal of the Royal Statistical Society, B* 26, 211-243.
- Box, G. E. P., and Draper, N. R. (1987). *Empirical model-building and response surfaces*. New York: Wiley.
- Box, G. E. P., and Hunter, J. S. (1957). "Multifactor experimental designs for exploring response surfaces." *Annals of Mathematical Statistics* 28, 195-241.
- Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978). *Statistics for experimenters. An introduction to design, data analysis, and model building*. New York: Wiley.
- Box, G. E. P., and Meyer, R. D. (1986). "An analysis for unreplicated fractional factorials." *Technometrics* 28, 11-18.
- Box, G. E. P., and Wilson, K. G. (1951). "On the experimental attainment of optimum conditions." *Journal of the Royal Statistical Society, B* 13, 1-45.
- Box, J. F. (1978). *R. A. Fisher, the life of a scientist*. New York: Wiley.
- Box, M. J., and Draper, N. R. (1974). "On minimum-point second-order designs." *Technometrics* 16, 613-616.
- Bozovich, H., Bancroft, T. A., and Hartley, H. O. (1956). Power of analysis of variance test procedures for incompletely specified models. *Annals of Mathematical Statistics* 27, 1017-1043.
- Bradu, D., and Gabriel, K. R. (1978). "The biplot as a diagnostic tool for models of two-way tables." *Technometrics* 20, 47-68.
- Brown, M. B., and Forsythe, A. B. (1974). "Robust tests for the equality of variances." *Journal of the American Statistical Association* 69, 364-367.
- Carmer, S. G., Nyquist, W. E., and Walker, W. M. (1989). "Least significant differences for combined analyses of experiments with two- or three-factor treatment designs." *Agronomy Journal* 81, 665-672.
- Carmer, S. G., and Swanson, M. R. (1973). "An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods." *Journal of the American Statistical Association* 68, 66-74.

- Carmer, S. G., and Walker, W. M. (1985). "Pairwise multiple comparisons of treatment means in agronomic research." *Journal of Agronomic Education* 14, 19–26.
- Clatworthy, W. H. (1973). "Tables of two-associate-class partially balanced designs." National Bureau of Standards, *Applied Mathematics Series*, no. 63. Washington, D.C.
- Cleveland, W. S. (1993). *Visualizing data*. Summit, N.J.: Hobart Press.
- Cochran, W. G. (1947). "Some consequences when the assumptions for the analysis of variance are not satisfied." *Biometrics* 3, 22–38.
- Cochran, W. G. (1957). "Analysis of covariance: Its nature and uses." *Biometrics* 13, 261–281.
- Cochran, W. G. (1965). *Sampling techniques*, 2d ed. New York: Wiley.
- Cochran, W. G. (1983). *Planning and analysis of observational studies*, L. E. Moses and F. Mosteller (eds.). New York: Wiley.
- Cochran, W. G., and Cox, G. M. (1957). *Experimental designs*, 2d ed. New York: Wiley.
- Collett, D. (1991). *Modelling binary data*. London: Chapman and Hall.
- Conner, W. S., and Zelen, M. (1959). "Fractional factorial experimental designs for factors at three levels." National Bureau of Standards, *Applied Mathematics Series*, no. 54. Washington, D.C.
- Conover, W. J., Johnson, M. E., and Johnson, M. M. (1981). "A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data." *Technometrics* 23, 351–361.
- Cook, R. D., and Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.
- Coons, I. (1957). "The analysis of covariance as a missing plot technique." *Biometrics* 13, 387–405.
- Coons, J. M., Kuehl, R. O., Oebker, N. F., and Simons, N. R. (1989). "Germination of eleven tomato phenotypes at constant or alternating high temperatures." *HortScience* 24(6), 927–930.
- Cornell, J. A. (1990). *Experiments with mixtures: Designs, models, and the analysis of mixture data*, 2d ed. New York: Wiley.
- Cox, D. R. (1958). *Planning of experiments*. New York: Wiley.
- Cox, D. R. (1970). *Analysis of binary data*. London: Chapman and Hall.
- Cox, D. R., and McCullagh, P. (1982). "Some aspects of analysis of covariance." *Biometrics* 38, 541–561.
- Daniel, C. (1959). "Use of half-normal plots in interpreting factorial two-level experiments." *Technometrics* 1, 311–342.
- Daniel, C. (1976). *Applications of statistics to industrial experimentation*. New York: Wiley.
- Daniel, C. (1978). "Patterns in residuals in the two-way layout." *Technometrics* 20, 385–395.
- Diamond, W. J. (1989). *Practical experiment designs for engineers and scientists*, 2d ed. New York: Van Nostrand Reinhold.
- Diggle, P. J., Liang, K., and Zeger, S. L. (1994). *Analysis of longitudinal data*. Oxford: Oxford University Press.

- Dobson, A. J. (1990). *An introduction to generalized linear models*. London: Chapman and Hall.
- Draper, N. R. (1985). "Small composite designs." *Technometrics* 27, 173–180.
- Draper, N. R., and Lin, D. K. J. (1990). "Small response-surface designs." *Technometrics* 32, 187–194.
- Dunnett, C. W. (1955). "A multiple comparisons procedure for comparing several treatments with a control." *Journal of the American Statistical Association* 50, 1096–1121.
- Dunnett, C. W. (1964). "New tables for multiple comparisons with a control." *Biometrics* 20, 482–491.
- Edgington, E. S. (1987). *Randomization tests*, 2d ed. New York: Marcel Dekker.
- Einot, I., and Gabriel, K. R. (1975). "A study of the powers of several methods of multiple comparisons." *Journal of the American Statistical Association* 70, 574–583.
- Eisenhart, C. (1947). "The assumptions underlying the analysis of variance." *Biometrics* 3, 1–21.
- Federer, W. T. (1955). *Experimental design: Theory and application*. New York: Macmillan.
- Federer, W. T. (1975). "The misunderstood split-plot." In *Applied Statistics*, R. P. Gupta (ed.). Amsterdam: North Holland.
- Finner, H. (1990). "On the modified S-method and directional errors." *Communications in Statistics—Theory and Methods*, A19, 41–53.
- Finney, D. J. (1946). "Standard errors of yields adjusted for regression on an independent measurement." *Biometrics* 2, 53–55.
- Finney, D. J. (1978). *Statistical methods in biological assay*, 3d ed. London: Griffin.
- Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.
- Fisher, R. A. (1926). "The arrangement of field experiments." *Journal of the Ministry of Agriculture of Great Britain* 33, 503–513.
- Fisher, R. A. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.
- Fisher, R. A. (1960). *The design of experiments*, 7th ed. New York: Hafner.
- Fisher, R. A., and Yates, F. (1963). *Statistical tables for biological, agricultural and medical research*, 6th ed. Edinburgh: Oliver and Boyd.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*, 2d ed. New York: Wiley.
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York: Wiley.
- Gates, C. E. (1995). "What really is experimental error in block designs?" *The American Statistician* 49, 362–363.
- Gates, C. E., and Shiue, C. (1962). "The analysis of variance of the S-stage hierarchical classification." *Biometrics* 18, 529–536.
- Gaylor, D. W., and Hopper, F. N. (1969). "Estimating the degrees of freedom for linear combinations of mean squares by Satterthwaite's formula." *Technometrics* 11, 691–706.

- Gill, J. L. (1978). *Design and analysis of experiments in the animal and medical sciences*. Vol. 1. Ames, Iowa: Iowa State University Press.
- Goldsmith, C. H., and Gaylor, D. W. (1970). "Three stage nested designs for estimating variance components." *Technometrics* 12, 487–498.
- Grandage, A. (1958). "Orthogonal coefficients for unequal intervals. Query 130." *Biometrics* 14, 287–289.
- Graybill, F. A. (1961). *An introduction to linear statistical models*. Vol. 1. New York: McGraw-Hill.
- Graybill, F. A. (1983). *Matrices with applications in statistics*, 2d ed. Belmont, CA: Wadsworth.
- Greenberg, B. G. (1951). "Why randomize?" *Biometrics* 7, 309–322.
- Greenhouse, S. W., and Geisser, S. (1959). "On methods in the analysis of profile data." *Psychometrika*, 24, 95–112.
- Grizzle, J. E. (1965). "The two-period change-over design and its use in clinical trials." *Biometrics* 21, 467–480.
- Grizzle, J. E. (1974). "Corrigenda to Grizzle (1965)." *Biometrics* 30, 727.
- Haaland, P. D. (1989). *Experimental design in biotechnology*. New York: Marcel Dekker.
- Hader, R. J. (1973). "An improper method of randomization in experimental design." *The American Statistician* 27, 82–84.
- Harshbarger, B. (1949). "Triple rectangular lattices." *Biometrics* 5, 1–13.
- Harshbarger, B., and Davis, L. L. (1952). "Latinized rectangular lattices." *Biometrics* 8, 73–84.
- Hartley, H. O. (1950). "The maximum *F*-ratio as a shortcut test for heterogeneity of variance." *Biometrika* 37, 308–312.
- Hartley, H. O., and Searle, S. R. (1969). "A discontinuity in mixed-model analysis." *Biometrics* 25, 573–576.
- Hartley, H. O., and Rao, J. N. K. (1967). "Maximum likelihood estimation for the mixed analysis of variance model." *Biometrika* 54, 93–108.
- Hayter, A. J. (1984). "A proof of the conjecture that the Tukey–Kramer multiple comparisons procedure is conservative." *Annals of Statistics* 12, 61–75.
- Hayter, A. J. (1986). "The maximum familywise error rate of Fisher's least significant difference test." *Journal of the American Statistical Association* 81, 1000–1004.
- Hayter, A. J. (1990). "A one-sided Studentized range test for testing against a simple order alternative." *Journal of the American Statistical Association* 85, 778–785.
- Hayter, A. J., and Liu, W. (1996). "Exact calculations for the one-sided Studentized range test for testing against a simple ordered alternative." *Computational Statistics & Data Analysis* 22, 17–25.
- Henderson, C. R. (1953). "Estimation of variance and covariance components." *Biometrics* 9, 226–252.
- Hendrix, L. J., Carter, M. W., and Scott, D. T. (1982). Covariance analysis with heterogeneity of slopes in fixed models. *Biometrics* 38, 641–650.
- Hering, F. and Mejza, S. (1997). "Incomplete split-block designs." *Biometrical Journal* 39, 227–238.

- Hicks, C. R. (1973). *Fundamental concepts in the design of experiments*, 2d ed. New York: Holt, Rinehart & Winston.
- Hills, M., and Armitage, P. (1979). "The two-period cross-over clinical trial." *British Journal of Clinical Pharmacology* 8, 7-20.
- Hinkelmann, K., and Kempthorne, O. (1994). *Design and Analysis of Experiments*. New York: Wiley.
- Hochberg, Y., and Tamhane, A. C. (1987). *Multiple comparison procedures*. New York: Wiley.
- Hocking, R. R. (1973). "A discussion of the two-way mixed model." *American Statistician* 27, 148-152.
- Hocking, R. R. (1985). *The analysis of linear models*. Monterey, Calif.: Brooks/Cole.
- Hocking, R. R., and Speed, F. M. (1975). "A full rank analysis of some linear model problems." *Journal of the American Statistical Association* 70, 706-712.
- Hsu, J. C. (1984). "Constrained simultaneous confidence intervals for multiple comparisons with the best." *Annals of Statistics* 12, 1136-1144.
- Hsu, J. C. (1996). *Multiple comparisons. Theory and methods*. London: Chapman and Hall.
- Hung, H. M., O'Neill, R. T., Bauer, P., and Köhne, K. (1977). "The behavior of the *P*-value when the alternative hypothesis is true." *Biometrics* 53, 11-22.
- Hunter, J. S. (1985). "Statistical design applied to product design." *Journal of Quality Technology* 17, 210-221.
- Hunter, J. S. (1989). "Let's all beware the Latin Square." *Quality Engineering* 1(4), 453-465.
- Hurlbert, S. H. (1984). "Pseudoreplication and the design of ecological field experiments." *Ecological Monographs* 54, 187-211.
- Huynh, H. (1978). "Some approximate tests for repeated measurement designs." *Psychometrika* 43, 161-175.
- Huynh, H., and Feldt, L. S. (1970). "Conditions under which mean square ratios in repeated measurements designs have exact *F*-distributions." *Journal of the American Statistical Association* 65, 1582-1589.
- Huynh, H., and Feldt, L. S. (1976). "Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split-plot designs." *Journal of Educational Statistics* 1, 69-82.
- Ito, P. K. (1980). "Robustness of ANOVA and MANOVA test procedures." In *Handbook of Statistics*, Vol. 1. P. R. Krishnaiah (ed.), 199-236. Amsterdam: North-Holland.
- Jarrett, R. G. (1978). "The analysis of designed experiments with missing observations." *Applied Statistics* 27, 38-46.
- John, J. A. (1966). "Cyclic incomplete block designs." *Journal of the Royal Statistical Society, B* 28, 345-360.
- John, J. A. (1987). *Cyclic designs*. London: Chapman and Hall.
- John, J. A., and Williams, E. R. (1995). *Cyclic and computer-generated designs*, 2d ed., London: Chapman and Hall.
- John, P. W. M. (1971). *Statistical design and analysis of experiments*. New York: Macmillan.

- Johnson, D., and Graybill, F. (1972). "An analysis of a two-way model with interaction and no replication." *Journal of the American Statistical Association* 67, 862–868.
- Johnson, N. R., and Leone, F. C. (1977). *Statistics and experimental design in engineering and the physical sciences*, Vol. II. New York: Wiley.
- Jones, B., and Kenward, M. G. (1989). *Design and analysis of cross-over trials*. London: Chapman and Hall.
- Jones, D. (1984). "Use, misuse, and role of multiple-comparison procedures in ecological and agricultural entomology." *Environmental Entomology* 13, 635–649.
- Kempthorne, O. (1952). *The design and analysis of experiments*. New York: Wiley.
- Kempthorne, O. (1966). "Some aspects of experimental inference." *Journal of the American Statistical Association* 61, 11–34.
- Kempthorne, O. (1975). "Inference from Experiments and Randomization." In *A Survey of Statistical Design and Linear Models*, 303–331. J. N. Srivastava (ed.). Amsterdam: North Holland.
- Keuls, M. (1952). "The use of the Studentized range in connection with an analysis of variance." *Euphytica* 1, 112–122.
- Kish, L. (1987). *Statistical design for research*. New York: Wiley.
- Koch, G. G. (1983). "Intraclass Correlation Coefficient." In *Encyclopedia of Statistical Sciences*, Vol. 4, 212–217. S. Kotz and N. L. Johnson (eds.). New York: Wiley.
- Kramer, C. Y. (1956). "Extension of multiple range tests to group means with unequal numbers of replications." *Biometrics* 12, 309–310.
- Lenth, R. V. (1989). "Quick and easy analysis of unreplicated factorials." *Technometrics* 31, 469–473.
- Lentner, M., Arnold, J. C., and Hinkelmann, K. (1989). "The efficiency of blocking: How to use MS(Blocks)/MS(Error) correctly." *The American Statistician* 43, 106–108.
- Leone, F. C., Nelson, L. S., Johnson, N. L., and Eisenstat, S. (1968). "Sampling distributions of variance components. II. Empirical studies of unbalanced nested designs." *Technometrics* 10, 719–737.
- Levene, H. (1960). "Robust tests for equality of variances." In *Contributions to Probability and Statistics*, 278–292. I. Olkin (ed.). Stanford, Calif.: Stanford University Press.
- Little, R. J. A., and Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Lucas, H. L. (1957). "Extra-period Latin-square change-over designs." *Journal of Dairy Science* 40, 225–239.
- Mandel, J. (1971). "A new analysis of variance model for non-additive data." *Technometrics* 13, 1–18.
- Manly, B. F. J. (1991). *Randomization and Monte Carlo methods in biology*. London: Chapman and Hall.
- Marvel, J. T., Berry, J. W., Kuehl, R. O., and Deutschmann, A. J. (1969). "Vinylation of Methyl α -D-Glucopyranoside." *Carbohydrate Research* 9, 295–303.

- Mason, R. L., Gunst, R. F., and Hess, J. L. (1989). *Statistical design and analysis of experiments with applications to engineering and science*. New York: Wiley.
- Mauchly, J. W. (1940). "Significance test for sphericity of a normal n -variate distribution." *Annals of Mathematical Statistics* 11, 204–209.
- McCullagh, P., and Nelder, J. A. (1989). *Generalized linear models*, 2d ed. London: Chapman and Hall.
- McIntosh, M. S. (1983). "Analysis of combined experiments." *Agronomy Journal* 75, 153–155.
- Mead, R. (1988). *The design of experiments. Statistical principles for practical application*. Cambridge, UK: Cambridge University Press.
- Mejza, I., and Mejza, S. (1996). "Incomplete split-plot designs generated by GDBIBD(2)." *Calcutta Statistical Association Bulletin* 46, 117–127.
- Milliken, G. A., and Johnson, D. E. (1984). *Analysis of messy data. Vol. I: Designed experiments*. Belmont, Calif.: Lifetime Learning Publications.
- Montgomery, D. C. (1997). *Design and analysis of experiments*, 4th ed. New York: Wiley.
- Moon, T. E., Levine, N., Cartmel, B., Bangert, J., Rodney, S., Schreiber, M., Peng, Y., Ritenbaugh, C., Meyskens, F., Alberts, D., and the Southwest Skin Cancer Prevention Study Group. (1995). "Design and recruitment for retinoid skin cancer prevention (SKICAP) trials." *Cancer Epidemiology, Biomarkers, and Prevention* 4, 661–669.
- Mosteller, F., and Tukey, J. W. (1977). *Data analysis and regression. A second course in statistics*. Reading, Mass.: Addison-Wesley.
- Myers, R. H., and Montgomery, D. C. (1995). *Response surface methodology. Process and product optimization using designed experiments*. New York: Wiley.
- Myers, R. H., Vining, G. G., Giovannitti-Jensen, A., and Myers, S. L. (1992). "Variance dispersion properties of second-order response surface designs." *Journal of Quality Technology* 24, 1–11.
- Nelder, J. A., and Wedderburn, R. W. M. (1972). "Generalized linear models." *Journal of the Royal Statistical Society A* 135, 370–384.
- Nelson, L. A., and Rawlings, J. O. (1983). "Ten common misuses of statistics in agronomic research and reporting." *Journal of Agronomic Education* 12, 100–105.
- Newman, D. (1939). "The distribution of the range in samples from a normal population, expressed in terms of an independent estimate of standard deviation." *Biometrika* 31, 20–30.
- Notz, W. (1982). "Minimal-point second-order designs." *Journal of Statistical Planning and Inference* 6, 47–58.
- Ostle, B., and Mensing, R. W. (1975). *Statistics in research*, 3d ed. Ames, Iowa: The Iowa State University Press.
- Patterson, H. D. (1951). "Change-over trials." *Journal of the Royal Statistical Society, B* 13, 256–271.
- Patterson, H. D. (1952). "The construction of balanced designs for experiments involving sequences of treatments." *Biometrika* 39, 32–48.

- Patterson, H. D., and Lucas, H. L. (1962). "Change-over designs." North Carolina Agricultural Experiment Station *Technical Bulletin*, no. 147.
- Patterson, H. D., and Thompson, R. (1971). Recovery of interblock information when block sizes are unequal. *Biometrika* 58, 545–554.
- Patterson, H. D., and Williams, E. R. (1976). "A new class of resolvable incomplete block designs." *Biometrika* 63, 83–92.
- Patterson, H. D., Williams, E. R., and Hunter, E. A. (1978). "Block designs for variety trials." *Journal of Agricultural Science* 90, 395–400.
- Petersen, R. G. (1985). *Design and analysis of experiments*. New York: Marcel Dekker.
- Pignatiello, J. J., and Ramberg, J. S. (1985). *Journal of Quality Technology* 17, 198–206. Discussion of article by R. N. Kackar (1985). "Off-line quality control, parameter design, and the Taguchi method." *Journal of Quality Technology* 17, 176–188.
- Pignatiello, J. J., and Ramberg, J. S. (1991). "Top ten triumphs and tragedies of Genichi Taguchi." *Quality Engineering* 4(2), 211–225.
- Plackett, R. L., and Burman, J. P. (1946). "The design of optimum multifactorial experiments." *Biometrika* 33, 305–325.
- Preece, D. A. (1983). "Latin squares, Latin cubes, Latin rectangles, etc." In *Encyclopedia of Statistical Sciences*, Vol. 4, 504–510. S. Kotz and N. L. Johnson (eds.). New York: Wiley.
- Ratkowsky, D. A., Alldredge, J. R., and Cotton, J. W. (1990). "Analyzing balanced or unbalanced Latin squares and other repeated measures designs for carryover effects using the GLM procedure." *SAS Users Group International, 15th Annual Conference Proceedings*, 1353–1358.
- Ratkowsky, D. A., Evans, M. A., and Alldredge, J. R. (1993). *Cross-over experiments: Design, analysis, and application*. New York: Marcel Dekker.
- Rawlings, J. O. (1988). *Applied regression analysis: A research tool*. Pacific Grove, Calif.: Brooks/Cole.
- Robinson, J. (1967). "Incomplete split-plot designs." *Biometrics* 23, 793–802.
- Roquemore, K. G. (1976). "Hybrid designs for quadratic response surfaces." *Technometrics* 18, 419–423.
- Satterthwaite, F. E. (1946). "An approximate distribution of estimates of variance components." *Biometrics* 2, 110–114.
- Saville, D. J. (1990). "Multiple comparison procedures: The practical solution." *The American Statistician* 44, 174–180.
- Scheffé, H. (1953). "A method for judging all contrasts in the analysis of variance." *Biometrika* 40, 87–104.
- Scheffé, H. (1959). *The analysis of variance*. New York: Wiley.
- Searle, S. R. (1971). *Linear models*. New York: Wiley.
- Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance components*. New York: Wiley.
- Searle, S. R. (1987). *Linear models for unbalanced data*. New York: Wiley.
- Searle, S. R., Speed, F. M., and Henderson, H. V. (1981). "Some computational and model equivalencies in analysis of variance of unequal-subclass-numbers data." *The American Statistician* 35, 16–33.

- Smith, H. F. (1957). "Interpretation of adjusted treatment means and regressions in analysis of covariance." *Biometrics* 13, 282–308.
- Smith, J. R., and Beverly, J. M. (1981). "The use and analysis of staggered nested factorial designs." *Journal of Quality Technology* 13, 166–173.
- Speed, F. M., and Hocking, R. R. (1976). "The use of the R()-notation with unbalanced data." *The American Statistician* 30, 30–33.
- Speed, F. M., Hocking, R. R., and Hackney, O. P. (1978). "Methods of analysis of linear models with unbalanced data." *Journal of the American Statistical Association* 73, 105–112.
- Steel, R. G. D., and Torrie, J. H. (1980). *Principles and procedures statistics: A biometrical approach*, 2d ed. New York: McGraw Hill.
- Taguchi, G. (1986). *Introduction to quality engineering*. Tokyo: Asian Productivity Organization.
- Tanur, J. F., Mosteller, F., Kruskal, W., Link, R., Pietero, R., Rising, G., and Lehman, E. (1978). *Statistics: A guide to the unknown*. San Francisco: Holden-Day.
- Thompson, W. A., Jr. (1962). The problem with negative estimates of variance components. *Annals of Mathematical Statistics* 33, 273–289.
- Tukey, J. W. (1949a). "Comparing individual means in the analysis of variance." *Biometrics* 5, 99–114.
- Tukey, J. W. (1949b). "One degree of freedom for non-additivity." *Biometrics* 5, 232–242.
- Tukey, J. W. (1955). "Answer to a query on non-additivity in Latin squares." *Biometrics* 11, 111–113.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass.: Addison-Wesley.
- Tukey, J. W. (1994). "The problem of multiple comparisons." In H. L. Braun (ed.), *The collected works of John W. Tukey*. Vol. VIII, Chapter 1, pp. 1–300. New York: Chapman and Hall.
- Urquhart, N. S., and Weeks, D. L. (1978). "Linear models in messy data: Some problems and alternatives." *Biometrics* 34, 696–705.
- Urquhart, N. S., Weeks, D. L., and Henderson, C. R. (1973). "Estimation associated with linear models." *Communications in Statistics* 1, 303–330.
- Velleman, P. F., and Hoaglin, D. C. (1981). *Applications, basics, and computing of exploratory data analysis*. Boston, Mass.: Duxbury Press.
- Vining, G. G. (1998). *Statistical methods for engineers*. Pacific Grove, Calif.: Duxbury Press.
- Voss, D. T. (1988). "Generalized modulus-ratio tests for analysis of factorial designs with zero degrees of freedom for error." *Communications in Statistics—Theory and Methods* 17, 3345–3359.
- Whitaker, E. R., Williams, E. R., and John, J. A. (1998). CycDesigN, CSIRO Forestry and Forest Products, Canberra, Australia, and The University of Waikato, Hamilton, New Zealand.
- Williams, E. J. (1949). "Experimental designs balanced for the estimation of residual effects of treatments." *Australian Journal of Scientific Research* 2, 149–168.

- Williams, E. R. (1986). Row and column designs with contiguous replicates. *Australian Journal of Statistics* 28, 154–163.
- Williams, E. R., and Talbot, M. (1993). ALPHA+: *Experimental designs for variety trials. Design user manual*. CSIRO, Canberra and SASS, Edinburgh.
- Williams, J. S. (1962). “A confidence interval for variance components.” *Biometrika* 49, 278–281.
- Yates, F. (1934). “The analysis of multiple classifications with unequal numbers in the different classes.” *Journal of the American Statistical Association* 29, 51–66.
- Yates, F. (1936a). “Incomplete randomized blocks.” *Annals of Eugenics* 7, 121–140.
- Yates, F. (1936b). “A new method of arranging variety trials involving a large number of varieties.” *Journal of Agricultural Science* 26, 424–455.
- Yates, F. (1940a). “The recovery of interblock information in balanced incomplete block designs.” *Annals of Eugenics* 10, 317–325.
- Yates, F. (1940b). “Lattice squares.” *Journal of Agricultural Science* 30, 672–687.
- Yates, F. (1948). “Contribution to the discussion of ‘The Validity of Comparative Experiments’ by F. J. Anscombe.” *Journal of the Royal Statistical Society, A* 111, 204–205.
- Youden, W. J. (1937). “Use of incomplete block replications in estimating tobacco-mosaic virus.” *Contributions from Boyce Thompson Institute* 9, 41–48.
- Youden, W. J. (1940). “Experimental designs to increase accuracy of greenhouse studies.” *Contributions from Boyce Thompson Institute* 11, 219–228.
- Youden, W. J. (1956). Special invited paper delivered to the Institute of Mathematical Statistics, December 8, 1956.
- Youden, W. J. (1972). “Randomization and experimentation.” *Technometrics* 14, 13–22.
- Zahn, D. A. (1975a). “Modifications of and revised critical values for the half-normal plot.” *Technometrics* 17, 189–200.
- Zahn, D. A. (1975b). “Empirical study of the half-normal plot.” *Technometrics* 17, 201–211.

Apéndice

TABLAS

Tabla I Distribución normal estándar; los elementos son $P(Z \geq z) = \alpha$

Tabla II Distribución t de Student; $t_{\alpha,v}$ para $P(t \geq t_{\alpha,v}) = \alpha$

Tabla III Distribución Chi-cuadrada; $X^2_{\alpha,v}$ para $P(X^2 \geq X^2_{\alpha,v}) = \alpha$

Tabla IV Distribución F ; F_{α,v_1,v_2} para $P(F \geq F_{\alpha,v_1,v_2}) = \alpha$

Tabla V Estadístico t de Bonferroni; $t_{\alpha/2,k,v}$ para $P(|t| \geq t_{\alpha/2,k,v}) = \alpha$

Tabla VI Comparaciones múltiples contra el mejor y pruebas de Dunnett;
 $d_{\alpha,k,v}$ para $P(d \geq d_{\alpha,k,v}) = \alpha$

Tabla VII Intervalos de Student; $q_{\alpha,k,v}$ para $P(q \geq q_{\alpha,k,v}) = \alpha$

Tabla VIII Prueba $F Max$; $F_{\alpha}Max$ para $P(FMax \geq F_{\alpha}Max) = \alpha$ para t
grupos de tratamiento y v grados de libertad

Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos

Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios

Tabla XI Polinomios ortogonales

Tabla XII 10 000 dígitos aleatorios

Tabla XIII 400 permutaciones aleatorias de los primeros 10 enteros

Tabla I Distribución normal estándar; los elementos $P(Z \geq z) = \alpha$.

<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.00	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.10	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.20	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.30	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.40	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.50	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.60	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.70	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.80	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.90	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.00	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.10	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.20	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.30	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.40	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.50	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.60	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.70	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.80	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.90	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.00	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.10	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.20	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.30	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.40	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.50	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.60	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.70	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.80	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.90	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.00	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

Adaptado con permiso de *CRC Standard Probability and Statistics Tables and Formulae* (1991), editado por W. H. Beyer, CRC Press, Boca Ratón, Florida.

Tabla II Distribución t de Student; $t_{\alpha,v}$ para $P(t \geq t_{\alpha,v}) = \alpha$.

$v \setminus \alpha$	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	636.309
2	.289	.816	1.886	2.920	4.303	6.965	9.925	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.768
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	3.291

Calculado con MINITAB por R. O. Kuehl.

Tabla III Distribución Chi-cuadrada; $X_{\alpha,v}^2$ para $P(X^2 \geq X_{\alpha,v}^2) = \alpha$.

$v \setminus \alpha$	0.995	0.99	0.975	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.025	0.01	0.005
1	.0000393	.000157	.000982	.00393	.0158	.102	.455	1.32	2.71	3.84	5.02	6.63	7.88
2	.0100	.0201	.0506	.103	.211	.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	.0717	.115	.216	.352	.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	.207	.297	.484	.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	.412	.554	.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	.676	.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

Adaptado de *Biometrika Tables for Statisticians*, vol. I, 1966, editado por E. S. Pearson y H. O. Hartley, con permiso de Biometrika Trustees.

Tabla IV Distribución F ; F_{α,v_1,v_2} para $P(F \geq F_{\alpha,v_1,v_2}) = \alpha$.

Tabla IV Distribución $F; F_{\alpha, v_1, v_2}$ para $P(F \geq F_{\alpha, v_1, v_2}) = \alpha$.

		$\alpha = 0.10$																	
$v_2 \setminus v_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

Adaptado de *Biometrika Tables for Statisticians*, vol. I, 1966, editado por E. S. Pearson y H. O. Hartley, con permiso de Biometrika Trustees.

Tabla IV Distribución $F; F_{\alpha, v_1, v_2}$ para $P(F \geq F_{\alpha, v_1, v_2}) = \alpha$.

		$\alpha = 0.05$																		
$v_2 \setminus v_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40	
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76	
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73	
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71	
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69	
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67	
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65	
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64	
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62	
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51	
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39	
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25	
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00	

Tabla IV Distribución $F; F_{\alpha, v_1, v_2}$ para $P(F \geq F_{\alpha, v_1, v_2}) = \alpha$.

$v_2 \setminus v_1$	$\alpha = 0.05$																			
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40	
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76	
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73	
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71	
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69	
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67	
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65	
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64	
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62	
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51	
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39	
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25	
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00	

Tabla IV Distribución $F; F_{\alpha, v_1, v_2}$ para $P(F \geq F_{\alpha, v_1, v_2}) = \alpha$.

		$\alpha = 0.01$																	
$v_2 \backslash v_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Tabla V Estadística de la Prueba.

Tabla V Estadístico t de Bonferroni; $t_{\alpha/2,k,v}$ para $P(|t| \geq t_{\alpha/2,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.01$														
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20
5	4.77	5.25	5.60	5.89	6.14	6.35	6.54	6.71	6.87	7.01	7.15	7.27	7.39	7.50	7.98
6	4.32	4.70	4.98	5.21	5.40	5.56	5.71	5.84	5.96	6.07	6.17	6.26	6.35	6.43	6.79
7	4.03	4.36	4.59	4.79	4.94	5.08	5.20	5.31	5.41	5.50	5.58	5.66	5.73	5.80	6.08
8	3.83	4.12	4.33	4.50	4.64	4.76	4.86	4.96	5.04	5.12	5.19	5.25	5.32	5.37	5.62
9	3.69	3.95	4.15	4.30	4.42	4.53	4.62	4.71	4.78	4.85	4.91	4.97	5.02	5.08	5.29
10	3.58	3.83	4.00	4.14	4.26	4.36	4.44	4.52	4.59	4.65	4.71	4.76	4.81	4.85	5.05
11	3.50	3.73	3.89	4.02	4.13	4.22	4.30	4.37	4.44	4.49	4.55	4.60	4.64	4.68	4.86
12	3.43	3.65	3.81	3.93	4.03	4.12	4.19	4.26	4.32	4.37	4.42	4.47	4.51	4.55	4.72
13	3.37	3.58	3.73	3.85	3.95	4.03	4.10	4.16	4.22	4.27	4.32	4.36	4.40	4.44	4.60
14	3.33	3.53	3.67	3.79	3.88	3.96	4.03	4.09	4.14	4.19	4.23	4.28	4.31	4.35	4.50
15	3.29	3.48	3.62	3.73	3.82	3.90	3.96	4.02	4.07	4.12	4.16	4.20	4.24	4.27	4.42
16	3.25	3.44	3.58	3.69	3.77	3.85	3.91	3.96	4.02	4.06	4.10	4.14	4.18	4.21	4.35
17	3.22	3.41	3.54	3.65	3.73	3.80	3.86	3.92	3.97	4.01	4.05	4.09	4.12	4.15	4.29
18	3.20	3.38	3.51	3.61	3.69	3.76	3.82	3.87	3.92	3.96	4.00	4.04	4.07	4.10	4.23
19	3.17	3.35	3.48	3.58	3.66	3.73	3.79	3.84	3.88	3.93	3.96	4.00	4.03	4.06	4.19
20	3.15	3.33	3.46	3.55	3.63	3.70	3.75	3.80	3.85	3.89	3.93	3.96	3.99	4.02	4.15
25	3.08	3.24	3.36	3.45	3.52	3.58	3.64	3.68	3.73	3.76	3.80	3.83	3.86	3.88	4.00
30	3.03	3.19	3.30	3.39	3.45	3.51	3.56	3.61	3.65	3.68	3.71	3.74	3.77	3.80	3.90
40	2.97	3.12	3.23	3.31	3.37	3.43	3.47	3.51	3.55	3.58	3.61	3.64	3.67	3.69	3.79
60	2.91	3.06	3.16	3.23	3.29	3.34	3.39	3.43	3.46	3.49	3.52	3.54	3.57	3.59	3.68
120	2.86	3.00	3.09	3.16	3.22	3.26	3.31	3.34	3.37	3.40	3.43	3.45	3.47	3.49	3.58
∞	2.81	2.94	3.02	3.09	3.14	3.19	3.23	3.26	3.29	3.32	3.34	3.36	3.38	3.40	3.48

k = número de comparaciones

Tomado de B. J. R. Bailey (1977) "Tables of the Bonferroni t Statistic", *Journal of the American Statistical Association* 72, 469-478. Reimpreso con permiso de *Journal of the American Statistical Association*. Copyright 1977 por American Statistical Association. Todos los derechos reservados.

Tabla V Estadístico t de Bonferroni; $t_{\alpha/2,k,v}$ para $P(|t| \geq t_{\alpha/2,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.05$														
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20
5	3.16	3.53	3.81	4.03	4.22	4.38	4.53	4.66	4.77	4.88	4.98	5.08	5.16	5.25	5.60
6	2.97	3.29	3.52	3.71	3.86	4.00	4.12	4.22	4.32	4.40	4.49	4.56	4.63	4.70	4.98
7	2.84	3.13	3.34	3.50	3.64	3.75	3.86	3.95	4.03	4.10	4.17	4.24	4.30	4.36	4.59
8	2.75	3.02	3.21	3.36	3.48	3.58	3.68	3.76	3.83	3.90	3.96	4.02	4.07	4.12	4.33
9	2.69	2.93	3.11	3.25	3.36	3.46	3.55	3.62	3.69	3.75	3.81	3.86	3.91	3.95	4.15
10	2.63	2.87	3.04	3.17	3.28	3.37	3.45	3.52	3.58	3.64	3.69	3.74	3.79	3.83	4.00
11	2.59	2.82	2.98	3.11	3.21	3.29	3.37	3.44	3.50	3.55	3.60	3.65	3.69	3.73	3.89
12	2.56	2.78	2.93	3.05	3.15	3.24	3.31	3.37	3.43	3.48	3.53	3.57	3.61	3.65	3.81
13	2.53	2.75	2.90	3.01	3.11	3.19	3.26	3.32	3.37	3.42	3.47	3.51	3.55	3.58	3.73
14	2.51	2.72	2.86	2.98	3.07	3.15	3.21	3.27	3.33	3.37	3.42	3.46	3.50	3.53	3.67
15	2.49	2.69	2.84	2.95	3.04	3.11	3.18	3.23	3.29	3.33	3.37	3.41	3.45	3.48	3.62
16	2.47	2.67	2.81	2.92	3.01	3.08	3.15	3.20	3.25	3.30	3.34	3.38	3.41	3.44	3.58
17	2.46	2.66	2.79	2.90	2.98	3.06	3.12	3.17	3.22	3.27	3.31	3.34	3.38	3.41	3.54
18	2.45	2.64	2.77	2.88	2.96	3.03	3.09	3.15	3.20	3.24	3.28	3.32	3.35	3.38	3.51
19	2.43	2.63	2.76	2.86	2.94	3.01	3.07	3.13	3.17	3.22	3.26	3.29	3.32	3.35	3.48
20	2.42	2.61	2.74	2.85	2.93	3.00	3.06	3.11	3.15	3.20	3.23	3.27	3.30	3.33	3.46
25	2.38	2.57	2.69	2.79	2.86	2.93	2.99	3.03	3.08	3.12	3.15	3.19	3.22	3.24	3.36
30	2.36	2.54	2.66	2.75	2.82	2.89	2.94	2.99	3.03	3.07	3.10	3.13	3.16	3.19	3.30
40	2.33	2.50	2.62	2.70	2.78	2.84	2.89	2.93	2.97	3.01	3.04	3.07	3.10	3.12	3.23
60	2.30	2.46	2.58	2.66	2.73	2.79	2.83	2.88	2.91	2.95	2.98	3.01	3.03	3.06	3.16
120	2.27	2.43	2.54	2.62	2.68	2.74	2.78	2.82	2.86	2.89	2.92	2.95	2.97	3.00	3.09
∞	2.24	2.39	2.50	2.58	2.64	2.69	2.73	2.77	2.81	2.84	2.87	2.89	2.91	2.94	3.02

k = número de comparaciones

Tabla VI Comparaciones múltiples contra el mejor y pruebas de Dunnett;
 $d_{\alpha,k,v}$ para $P(d \geq d_{\alpha,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.05$ (de un lado)												
	2	3	4	5	6	7	8	9	10	11	12	15	20
5	2.44	2.68	2.85	2.98	3.08	3.16	3.24	3.30	3.36	3.41	3.45	3.57	3.72
6	2.34	2.56	2.71	2.83	2.92	3.00	3.07	3.12	3.17	3.22	3.26	3.37	3.50
7	2.27	2.48	2.62	2.73	2.82	2.89	2.95	3.01	3.05	3.10	3.13	3.23	3.36
8	2.22	2.42	2.55	2.66	2.74	2.81	2.87	2.92	2.96	3.01	3.04	3.14	3.25
9	2.18	2.37	2.50	2.60	2.68	2.75	2.81	2.86	2.90	2.94	2.97	3.06	3.18
10	2.15	2.34	2.47	2.56	2.64	2.70	2.76	2.81	2.85	2.89	2.92	3.01	3.12
11	2.13	2.31	2.44	2.53	2.60	2.67	2.72	2.77	2.81	2.85	2.88	2.96	3.07
12	2.11	2.29	2.41	2.50	2.58	2.64	2.69	2.74	2.78	2.81	2.84	2.93	3.03
13	2.09	2.27	2.39	2.48	2.55	2.61	2.66	2.71	2.75	2.78	2.82	2.90	3.00
14	2.08	2.25	2.37	2.46	2.53	2.59	2.64	2.69	2.72	2.76	2.79	2.87	2.97
15	2.07	2.24	2.36	2.44	2.51	2.57	2.62	2.67	2.70	2.74	2.77	2.85	2.95
16	2.06	2.23	2.34	2.43	2.50	2.56	2.61	2.65	2.69	2.72	2.75	2.83	2.93
17	2.05	2.22	2.33	2.42	2.49	2.54	2.59	2.64	2.67	2.71	2.74	2.81	2.91
18	2.04	2.21	2.32	2.41	2.48	2.53	2.58	2.62	2.66	2.69	2.72	2.80	2.89
19	2.03	2.20	2.31	2.40	2.47	2.52	2.57	2.61	2.65	2.68	2.71	2.79	2.88
20	2.03	2.19	2.30	2.39	2.46	2.51	2.56	2.60	2.64	2.67	2.70	2.77	2.87
24	2.01	2.17	2.28	2.36	2.43	2.48	2.53	2.57	2.60	2.64	2.66	2.74	2.83
30	1.99	2.15	2.25	2.33	2.40	2.45	2.50	2.54	2.57	2.60	2.63	2.70	2.79
40	1.97	2.13	2.23	2.31	2.37	2.42	2.47	2.51	2.54	2.57	2.60	2.67	2.75
60	1.95	2.10	2.21	2.28	2.35	2.39	2.44	2.48	2.51	2.54	2.56	2.63	2.72
120	1.93	2.08	2.18	2.26	2.32	2.37	2.41	2.45	2.48	2.51	2.53	2.60	2.68
∞	1.92	2.06	2.16	2.23	2.29	2.34	2.38	2.42	2.45	2.48	2.50	2.56	2.64

Tomado de C. W. Dunnett (1955) "A Multiple Comparison Procedure for Comparing Several Treatments with a Control", *Journal of the American Statistical Association* 50, 1112-1118. Reimpreso con permiso de *Journal of the American Statistical Association*. Copyright 1955 por The American Statistical Association. Todos los derechos reservados. C. W. Dunnett (1964) "New Tables for Multiple Comparison with a Control", *Biometrics* 20, 482-491; y tablas adicionales producidas por C. W. Dunnett en 1980.

Tabla VI Comparaciones múltiples contra el mejor y pruebas de Dunnett;
 $d_{\alpha,k,v}$ para $P(|d| \geq d_{\alpha,k,v}) = \alpha$.

v^k	$\alpha = 0.05$ (de dos lados)												
	2	3	4	5	6	7	8	9	10	11	12	15	20
5	3.03	3.29	3.48	3.62	3.73	3.82	3.90	3.97	4.03	4.09	4.14	4.26	4.42
6	2.86	3.10	3.26	3.39	3.49	3.57	3.64	3.71	3.76	3.81	3.86	3.97	4.11
7	2.75	2.97	3.12	3.24	3.33	3.41	3.47	3.53	3.58	3.63	3.67	3.78	3.91
8	2.67	2.88	3.02	3.13	3.22	3.29	3.35	3.41	3.46	3.50	3.54	3.64	3.76
9	2.61	2.81	2.95	3.05	3.14	3.20	3.26	3.32	3.36	3.40	3.44	3.53	3.65
10	2.57	2.76	2.89	2.99	3.07	3.14	3.19	3.24	3.29	3.33	3.36	3.45	3.57
11	2.53	2.72	2.84	2.94	3.02	3.08	3.14	3.19	3.23	3.27	3.30	3.39	3.50
12	2.50	2.68	2.81	2.90	2.98	3.04	3.09	3.14	3.18	3.22	3.25	3.34	3.45
13	2.48	2.65	2.78	2.87	2.94	3.00	3.06	3.10	3.14	3.18	3.21	3.29	3.40
14	2.46	2.63	2.75	2.84	2.91	2.97	3.02	3.07	3.11	3.14	3.18	3.26	3.36
15	2.44	2.61	2.73	2.82	2.89	2.95	3.00	3.04	3.08	3.12	3.15	3.23	3.33
16	2.42	2.59	2.71	2.80	2.87	2.92	2.97	3.02	3.06	3.09	3.12	3.20	3.30
17	2.41	2.58	2.69	2.78	2.85	2.90	2.95	3.00	3.03	3.07	3.10	3.18	3.27
18	2.40	2.56	2.68	2.76	2.83	2.89	2.94	2.98	3.01	3.05	3.08	3.16	3.25
19	2.39	2.55	2.66	2.75	2.81	2.87	2.92	2.96	3.00	3.03	3.06	3.14	3.23
20	2.38	2.54	2.65	2.73	2.80	2.86	2.90	2.95	2.98	3.02	3.05	3.12	3.22
24	2.35	2.51	2.61	2.70	2.76	2.81	2.86	2.90	2.94	2.97	3.00	3.07	3.16
30	2.32	2.47	2.58	2.66	2.72	2.77	2.82	2.86	2.89	2.92	2.95	3.02	3.11
40	2.29	2.44	2.54	2.62	2.68	2.73	2.77	2.81	2.85	2.87	2.90	2.97	3.06
60	2.27	2.41	2.51	2.58	2.64	2.69	2.73	2.77	2.80	2.83	2.86	2.92	3.00
120	2.24	2.38	2.47	2.55	2.60	2.65	2.69	2.73	2.76	2.79	2.81	2.87	2.95
∞	2.21	2.35	2.44	2.51	2.57	2.61	2.65	2.69	2.72	2.74	2.77	2.83	2.91

Tabla VI Comparaciones múltiples contra el mejor y pruebas de Dunnett;
 $d_{\alpha,k,v}$ para $P(d \geq d_{\alpha,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.01$ (de un lado)												
	2	3	4	5	6	7	8	9	10	11	12	15	20
5	3.90	4.21	4.43	4.60	4.73	4.85	4.94	5.03	5.11	5.17	5.24	5.39	5.59
6	3.61	3.88	4.07	4.21	4.33	4.43	4.51	4.59	4.64	4.70	4.76	4.89	5.06
7	3.42	3.66	3.83	3.96	4.07	4.15	4.23	4.30	4.35	4.40	4.45	4.57	4.72
8	3.29	3.51	3.67	3.79	3.88	3.96	4.03	4.09	4.14	4.19	4.23	4.34	4.48
9	3.19	3.40	3.55	3.66	3.75	3.82	3.89	3.94	3.99	4.04	4.08	4.18	4.31
10	3.11	3.31	3.45	3.56	3.64	3.71	3.78	3.83	3.88	3.92	3.96	4.06	4.18
11	3.06	3.25	3.38	3.48	3.56	3.63	3.69	3.74	3.79	3.83	3.86	3.96	4.08
12	3.01	3.19	3.32	3.42	3.50	3.56	3.62	3.67	3.71	3.75	3.79	3.88	3.99
13	2.97	3.15	3.27	3.37	3.44	3.51	3.56	3.61	3.65	3.69	3.73	3.81	3.92
14	2.94	3.11	3.23	3.32	3.40	3.46	3.51	3.56	3.60	3.64	3.67	3.76	3.87
15	2.91	3.08	3.20	3.29	3.36	3.42	3.47	3.52	3.56	3.60	3.63	3.71	3.82
16	2.88	3.05	3.17	3.26	3.33	3.39	3.44	3.48	3.52	3.56	3.59	3.67	3.78
17	2.86	3.03	3.14	3.23	3.30	3.36	3.41	3.45	3.49	3.53	3.56	3.64	3.74
18	2.84	3.01	3.12	3.21	3.27	3.33	3.38	3.42	3.46	3.50	3.53	3.61	3.71
19	2.83	2.99	3.10	3.18	3.25	3.31	3.36	3.40	3.44	3.47	3.50	3.58	3.68
20	2.81	2.97	3.08	3.17	3.23	3.29	3.34	3.38	3.42	3.45	3.48	3.56	3.65
24	2.77	2.92	3.03	3.11	3.17	3.22	3.27	3.31	3.35	3.38	3.41	3.48	3.57
30	2.72	2.87	2.97	3.05	3.11	3.16	3.21	3.24	3.28	3.31	3.34	3.41	3.50
40	2.68	2.82	2.92	2.99	3.05	3.10	3.14	3.18	3.21	3.24	3.27	3.34	3.42
60	2.64	2.78	2.87	2.94	3.00	3.04	3.08	3.12	3.15	3.18	3.20	3.27	3.35
120	2.60	2.73	2.82	2.89	2.94	2.99	3.03	3.06	3.09	3.12	3.14	3.20	3.28
∞	2.56	2.68	2.77	2.84	2.89	2.93	2.97	3.00	3.03	3.06	3.08	3.14	3.21

Tabla VI Comparaciones múltiples contra el mejor y pruebas de Dunnett;
 $d_{\alpha,k,v}$ para $P(|d| \geq d_{\alpha,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.01$ (de dos lados)													
	2	3	4	5	6	7	8	9	10	11	12	15	20	
5	4.63	4.98	5.22	5.41	5.56	5.69	5.80	5.89	5.98	6.05	6.12	6.30	6.52	
6	4.21	4.51	4.71	4.87	5.00	5.10	5.20	5.28	5.35	5.41	5.47	5.62	5.81	
7	3.95	4.21	4.39	4.53	4.64	4.74	4.82	4.89	4.95	5.01	5.06	5.19	5.36	
8	3.77	4.00	4.17	4.29	4.40	4.48	4.56	4.62	4.68	4.73	4.78	4.90	5.05	
9	3.63	3.85	4.01	4.12	4.22	4.30	4.37	4.43	4.48	4.53	4.57	4.68	4.82	
10	3.53	3.74	3.88	3.99	4.08	4.16	4.22	4.28	4.33	4.37	4.42	4.52	4.65	
11	3.45	3.65	3.79	3.89	3.98	4.05	4.11	4.16	4.21	4.25	4.29	4.39	4.52	
12	3.39	3.58	3.71	3.81	3.89	3.96	4.02	4.07	4.12	4.16	4.19	4.29	4.41	
13	3.33	3.52	3.65	3.74	3.82	3.89	3.94	3.99	4.04	4.08	4.11	4.20	4.32	
14	3.29	3.47	3.59	3.69	3.76	3.83	3.88	3.93	3.97	4.01	4.05	4.13	4.24	
15	3.25	3.43	3.55	3.64	3.71	3.78	3.83	3.88	3.92	3.95	3.99	4.07	4.18	
16	3.22	3.39	3.51	3.60	3.67	3.73	3.78	3.83	3.87	3.91	3.94	4.02	4.13	
17	3.19	3.36	3.47	3.56	3.63	3.69	3.74	3.79	3.83	3.86	3.90	3.98	4.08	
18	3.17	3.33	3.44	3.53	3.60	3.66	3.71	3.75	3.79	3.83	3.86	3.94	4.04	
19	3.15	3.31	3.42	3.50	3.57	3.63	3.68	3.72	3.76	3.79	3.83	3.90	4.00	
20	3.13	3.29	3.40	3.48	3.55	3.60	3.65	3.69	3.73	3.77	3.80	3.87	3.97	
24	3.07	3.22	3.32	3.40	3.47	3.52	3.57	3.61	3.64	3.68	3.70	3.78	3.87	
30	3.01	3.15	3.25	3.33	3.39	3.44	3.49	3.52	3.56	3.59	3.62	3.69	3.78	
40	2.95	3.09	3.19	3.26	3.32	3.37	3.41	3.44	3.48	3.51	3.53	3.60	3.68	
60	2.90	3.03	3.12	3.19	3.25	3.29	3.33	3.37	3.40	3.42	3.45	3.51 ⁺	3.59	
120	2.85	2.97	3.06	3.12	3.18	3.22	3.26	3.29	3.32	3.35	3.37	3.43	3.51	
∞	2.79	2.92	3.00	3.06	3.11	3.15	3.19	3.22	3.25	3.27	3.29	3.35	3.42	

Tabla VII Intervalos de Student; $q_{\alpha,k,v}$ para $P(q \geq q_{\alpha,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.05$																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	18.1	26.7	32.8	37.2	40.5	43.1	45.4	47.3	49.1	50.6	51.9	53.2	54.3	55.4	56.3	57.2	58.0	58.8	59.6	
2	6.09	8.28	9.80	10.89	11.73	12.43	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.36	16.57	16.77	
3	4.50	5.88	6.83	7.51	8.04	8.47	8.85	9.18	9.46	9.72	9.95	10.16	10.35	10.52	10.69	10.84	10.98	11.12	11.24	
4	3.93	5.00	5.76	6.31	6.73	7.06	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.67	8.80	8.92	9.03	9.14	9.24	
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	
6	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32	6.49	6.65	6.79	6.92	7.04	7.14	7.24	7.34	7.43	7.51	7.59	
7	3.34	4.16	4.68	5.06	5.35	5.59	5.80	5.99	6.15	6.29	6.42	6.54	6.65	6.75	6.84	6.93	7.01	7.08	7.16	
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.65	
10	3.15	3.88	4.33	4.66	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.12	6.20	6.27	6.34	6.41	6.47	
11	3.1	3.82	4.26	4.58	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.98	6.06	6.14	6.20	6.27	6.33	
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40	5.51	5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	
13	3.06	3.73	4.15	4.46	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.93	6.00	6.06	6.11	
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.56	5.64	5.72	5.79	5.86	5.92	5.98	6.03	
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	5.65	5.72	5.79	5.85	5.91	5.96	
16	3.00	3.65	4.05	4.34	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	
17	2.98	3.62	4.02	4.31	4.52	4.70	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.55	5.61	5.68	5.74	5.79	5.84	
18	2.97	3.61	4.00	4.28	4.49	4.67	4.83	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	
19	2.96	3.59	3.98	4.26	4.47	4.64	4.79	4.92	5.04	5.14	5.23	5.32	5.39	5.46	5.53	5.59	5.65	5.70	5.75	
20	2.95	3.58	3.96	4.24	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43	5.50	5.56	5.61	5.66	5.71	
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.50	5.55	5.59	
30	2.89	3.48	3.84	4.11	4.30	4.46	4.60	4.72	4.83	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.48	
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74	4.82	4.90	4.98	5.05	5.11	5.17	5.22	5.27	5.32	5.36	
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	
∞	2.77	3.32	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80	4.84	4.89	4.93	4.97	5.01	

Adaptado de *Biometrika Tables for Statisticians*, vol 1, 1966, editado por E. S. Pearson y H. O. Hartley, con permiso de The Biometrika Trustees.

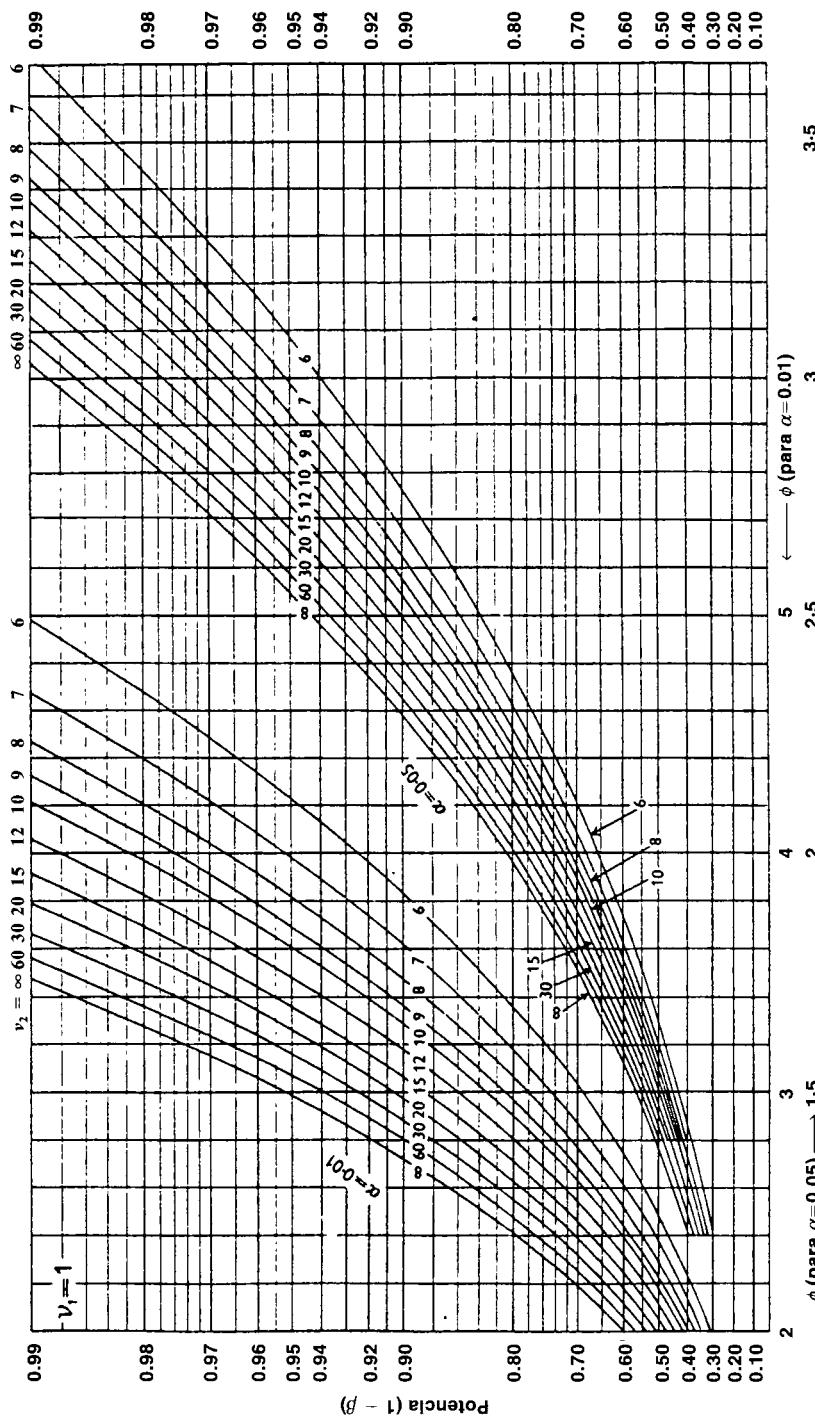
Tabla VII Intervalos de Student; $q_{\alpha,k,v}$ para $P(q \geq q_{\alpha,k,v}) = \alpha$.

$v \setminus k$	$\alpha = 0.01$																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	90.0	135	164	186	202	216	227	237	246	253	260	266	272	272	282	286	290	294	298	
2	14.0	19.0	22.3	24.7	26.6	28.2	29.5	30.7	31.7	32.6	33.4	34.1	34.8	35.4	36.0	36.5	37.0	37.5	37.9	
3	8.26	10.6	12.2	13.3	14.2	15.0	15.6	16.2	16.7	17.1	17.5	17.9	18.2	18.5	18.8	19.1	19.3	19.5	19.8	
4	6.51	8.12	9.17	9.96	10.6	11.1	11.5	11.9	12.3	12.6	12.8	13.1	13.3	13.5	13.7	13.9	14.1	14.2	14.4	
5	5.70	6.97	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93	
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.49	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	
8	4.74	5.63	6.20	6.63	6.96	7.24	7.47	7.68	7.87	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.32	7.49	7.65	7.78	7.91	8.03	8.13	8.23	8.32	8.41	8.49	8.57	
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.48	7.60	7.71	7.81	7.91	7.99	8.07	8.15	8.22	
11	4.39	5.14	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	
12	4.32	5.04	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.34	7.42	7.48	7.55	
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05	7.12	7.20	7.27	7.33	7.39	
15	4.17	4.83	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	
16	4.13	4.78	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.80	6.87	6.94	7.00	7.05	
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65	6.72	6.79	6.85	6.91	6.96	
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	6.52	6.59	6.65	6.71	6.76	6.82	
24	3.96	4.54	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	
40	3.82	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.60	5.69	5.77	5.84	5.90	5.96	6.02	6.07	6.12	6.17	6.21	
60	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.79	5.84	5.89	5.93	5.98	6.02	
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.51	5.56	5.61	5.66	5.71	5.75	5.79	5.83	
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65	

Tabla VIII Prueba F Max; $F_{\alpha}Max$ para $P(FMax \geq F_{\alpha}Max) = \alpha$ para t grupos de tratamiento y v grados de libertad.

$v \setminus t$	$\alpha = 0.05$											
	2	3	4	5	6	7	8	9	10	11	12	
2	39.0	87.5	142	202	266	333	403	475	550	626	704	
3	15.4	27.8	39.2	50.7	62.0	72.9	83.5	93.9	104	114	124	
4	9.60	15.5	20.6	25.2	29.5	33.6	37.5	41.1	44.6	48.0	51.4	
5	7.15	10.8	13.7	16.3	18.7	20.8	22.9	24.7	26.5	28.2	29.9	
6	5.82	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	19.7	20.7	
7	4.99	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.1	15.8	
8	4.43	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.2	12.7	
9	4.03	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.3	10.7	
10	3.72	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.01	9.34	
12	3.28	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.25	7.48	
15	2.86	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.77	5.93	
20	2.46	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.49	4.59	
30	2.07	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.36	3.39	
60	1.67	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.33	2.36	
∞	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
$v \setminus t$	$\alpha = 0.01$											
	2	3	4	5	6	7	8	9	10	11	12	
2	199	448	729	1036	1362	1705	2063	2342	2813	3204	3605	
3	47.5	85	120	151	184	21(6)	24(9)	28(1)	31(0)	33(7)	36(1)	
4	23.2	37	49	59	69	79	89	97	106	113	120	
5	14.9	22	28	33	38	42	46	50	54	57	60	
6	11.1	15.5	19.1	22	25	27	30	32	34	36	37	
7	8.89	12.1	14.5	16.5	18.4	20	22	23	24	26	27	
8	7.50	9.9	11.7	13.2	14.5	15.8	16.6	17.9	18.9	19.8	21	
9	6.54	8.5	9.9	11.1	12.1	13.1	13.9	14.7	15.3	16.0	16.6	
10	5.85	7.4	8.6	9.6	10.4	11.1	11.8	12.4	12.9	13.4	13.9	
12	4.91	6.1	6.9	7.6	8.2	8.7	9.1	9.5	9.9	10.2	10.6	
15	4.07	4.9	5.5	6.0	6.4	6.7	7.1	7.3	7.5	7.8	8.0	
20	3.32	3.8	4.3	4.6	4.9	5.1	5.3	5.5	5.6	5.8	5.9	
30	2.63	3.0	3.3	3.4	3.6	3.7	3.8	3.9	4.0	4.1	4.2	
60	1.96	2.2	2.3	2.4	2.4	2.5	2.5	2.6	2.6	2.7	2.7	
∞	1.00	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	

Adaptado de *Biometrika Tables for Statisticians*, vol 1, 1966, editado por E. S. Pearson y H. O. Hartley, con permiso de The Biometrika Trustees.

Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

Adaptado con permiso de *Biometrika Tables for Statisticians*, vol 1, 1966, editado por E. S. Pearson y H. O. Hartley.

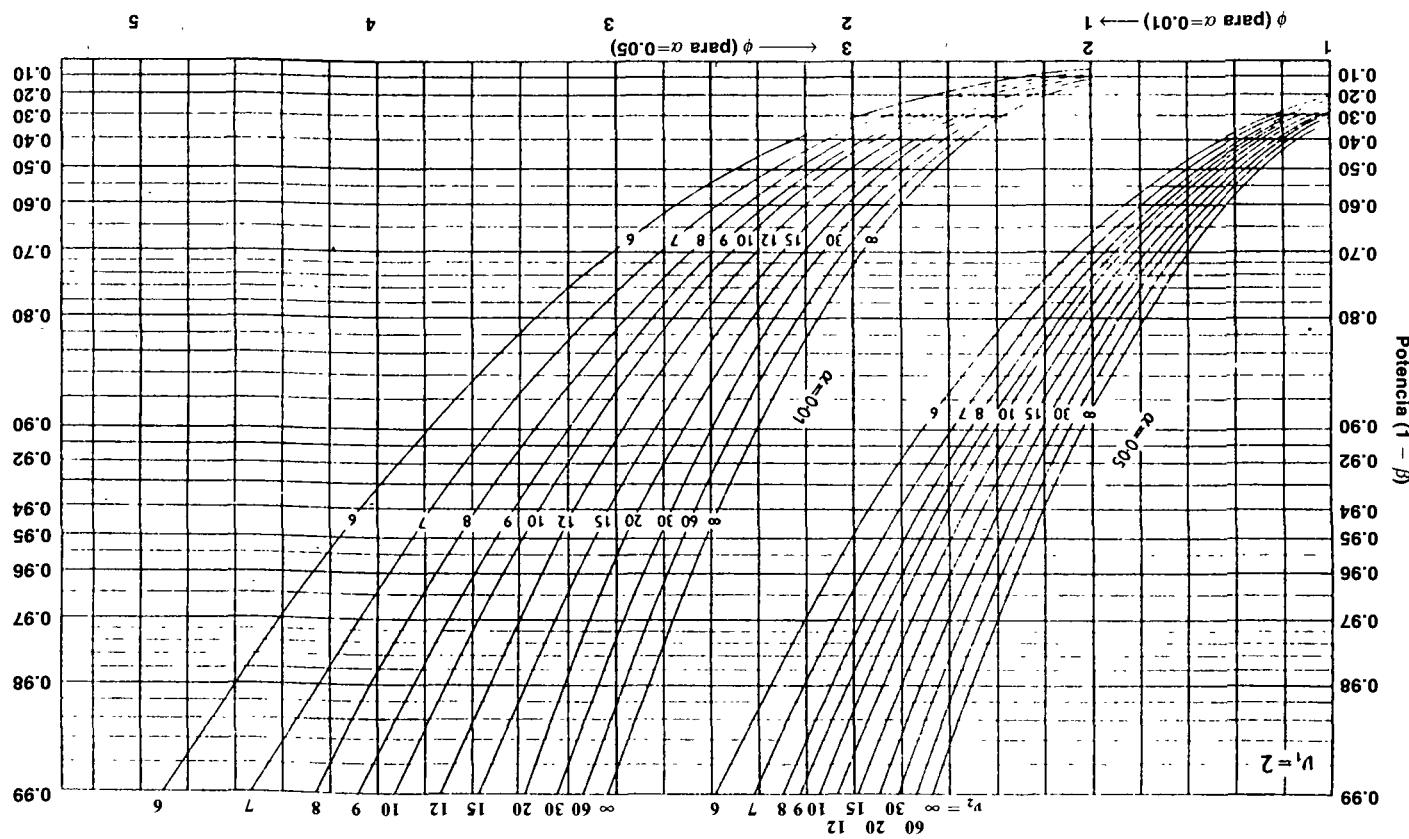
Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

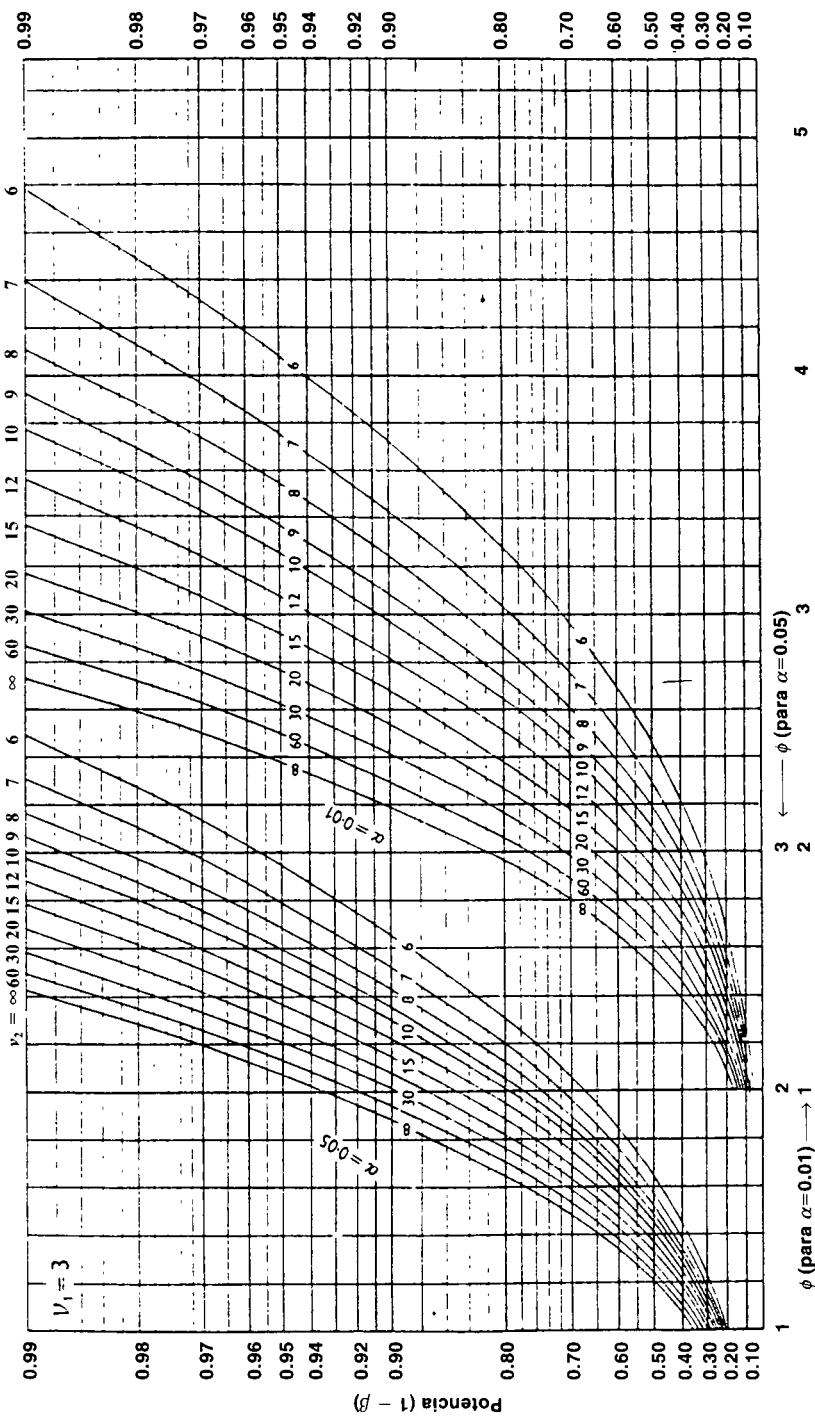
Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

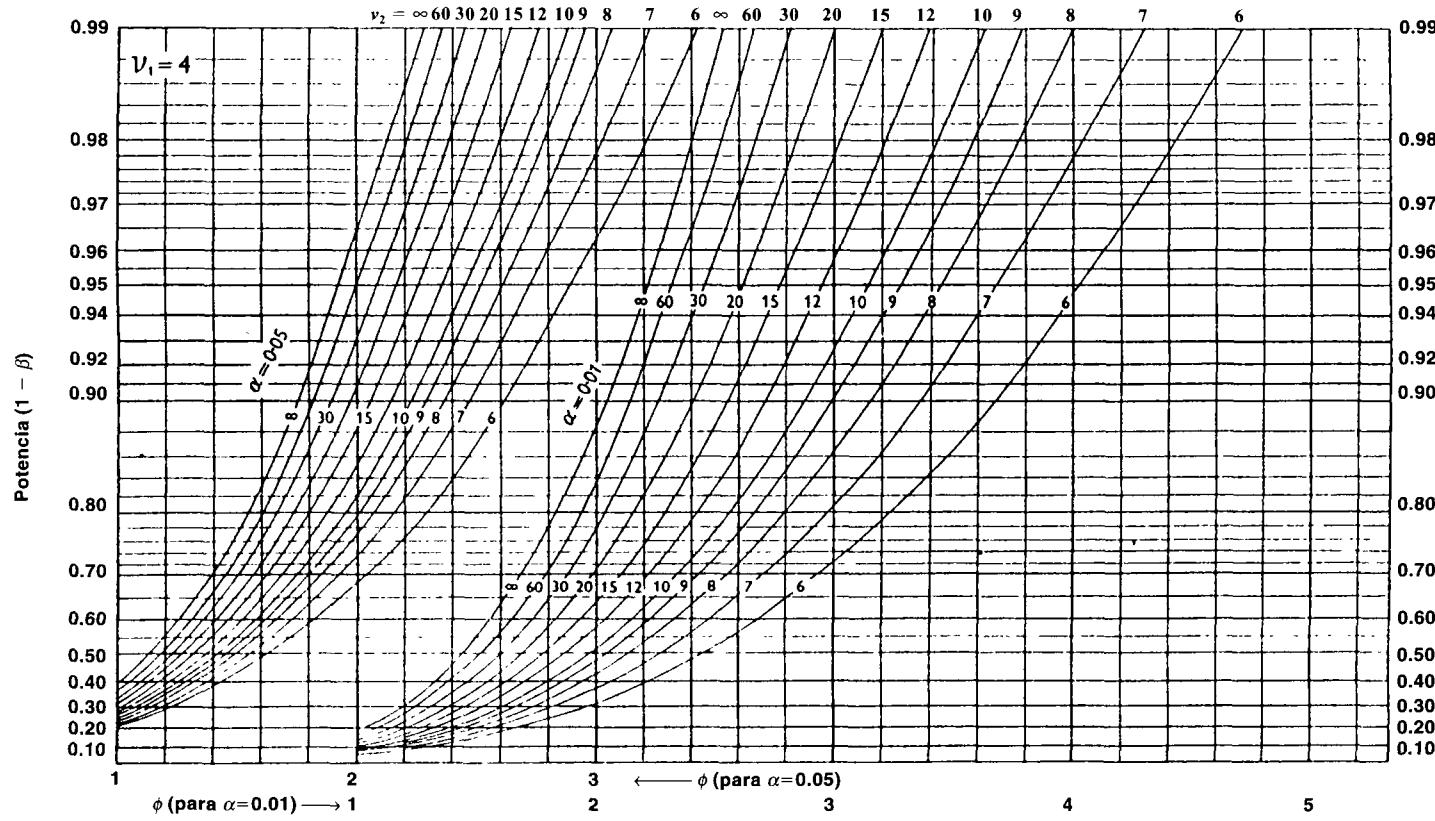
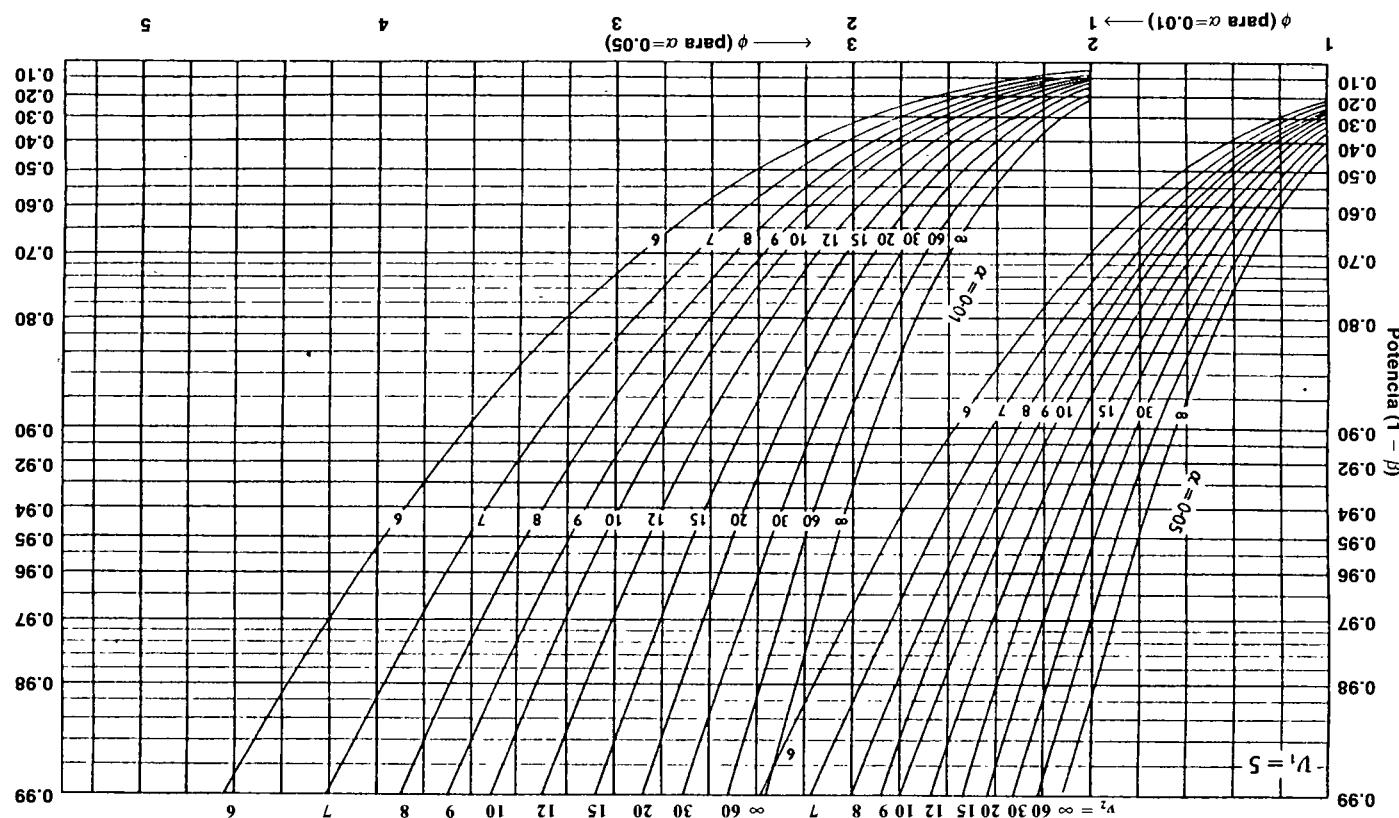


Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.



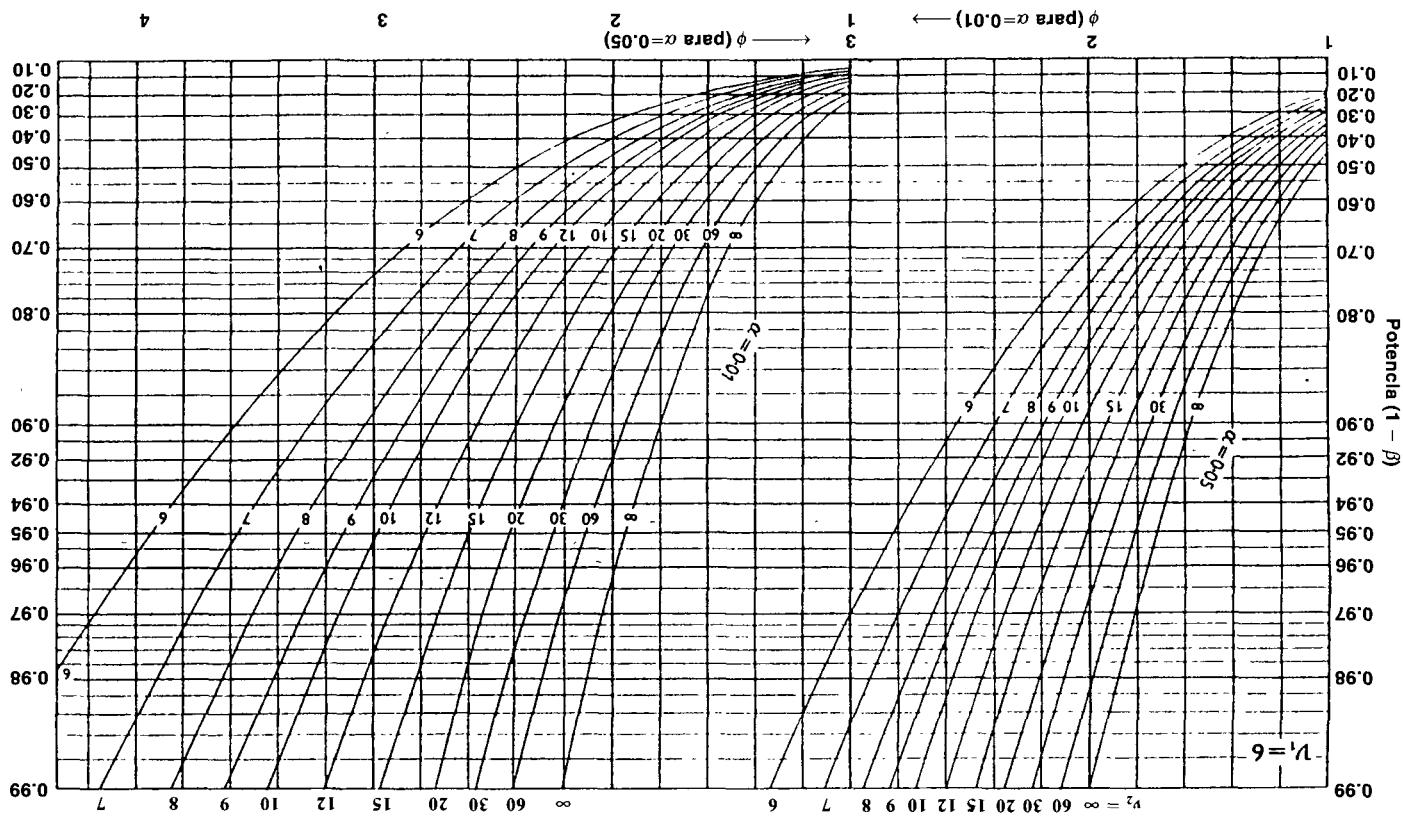
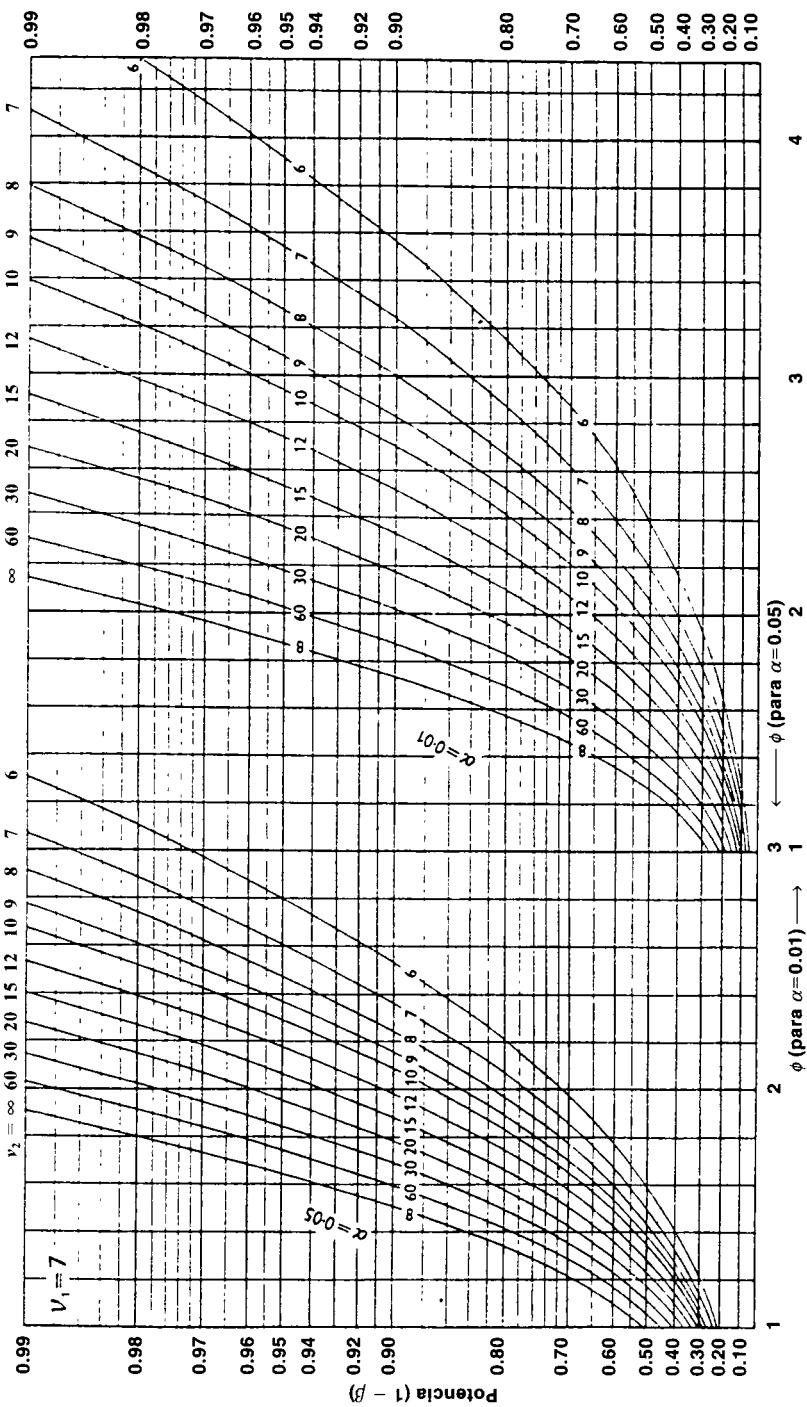


Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.**Tabla IX** Curvas de notación de la prueba F para análisis de varianza del modelo de efectos fijos.

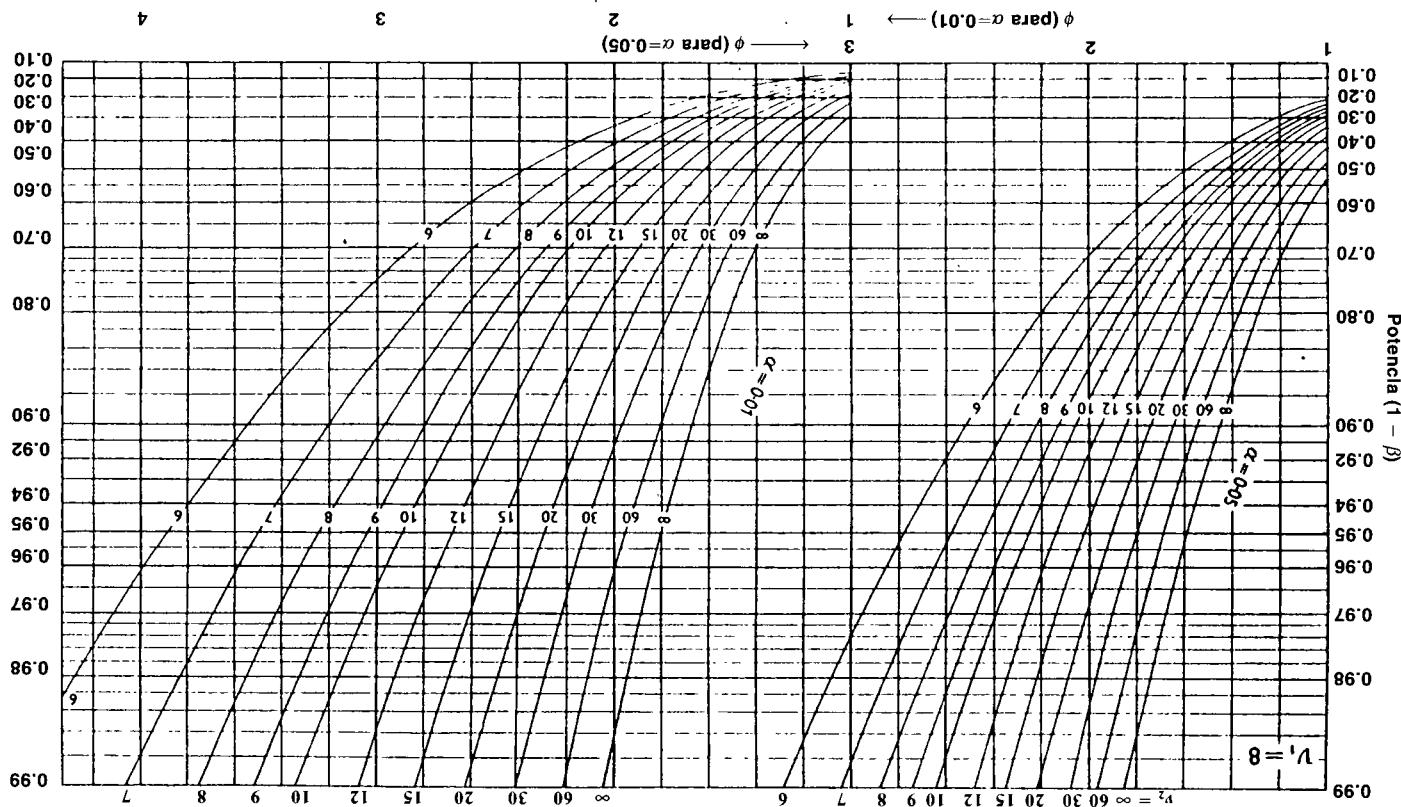


Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

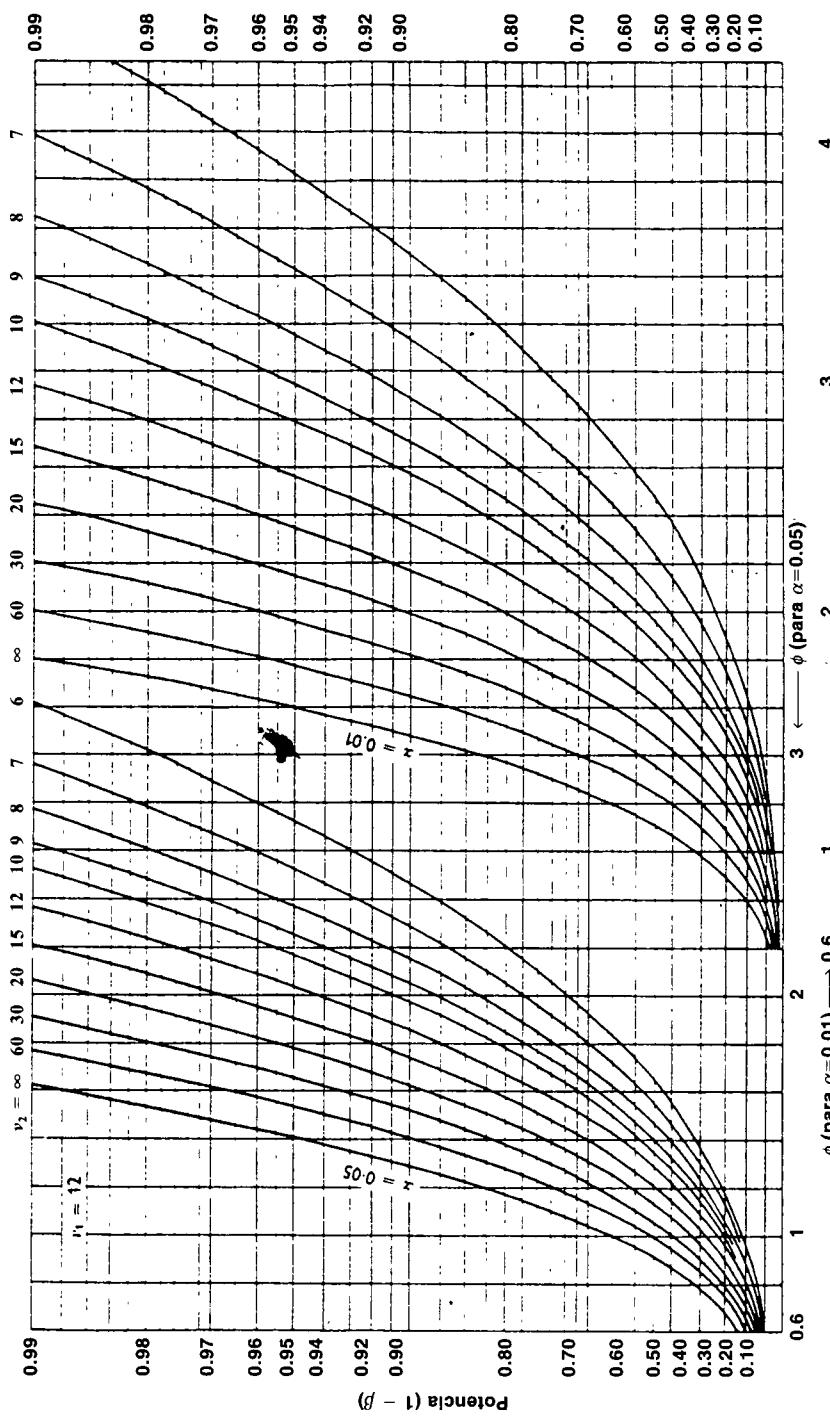


Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

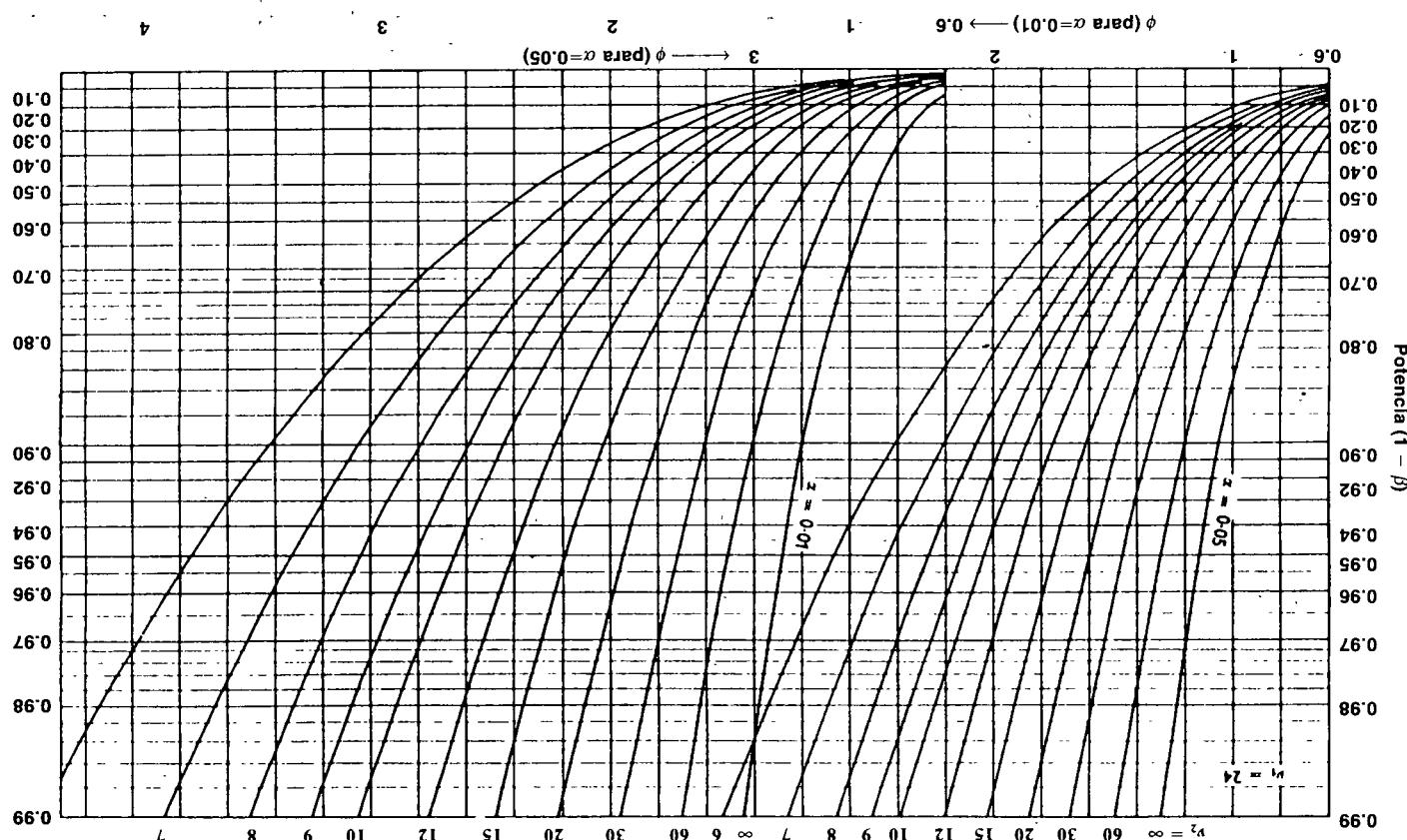
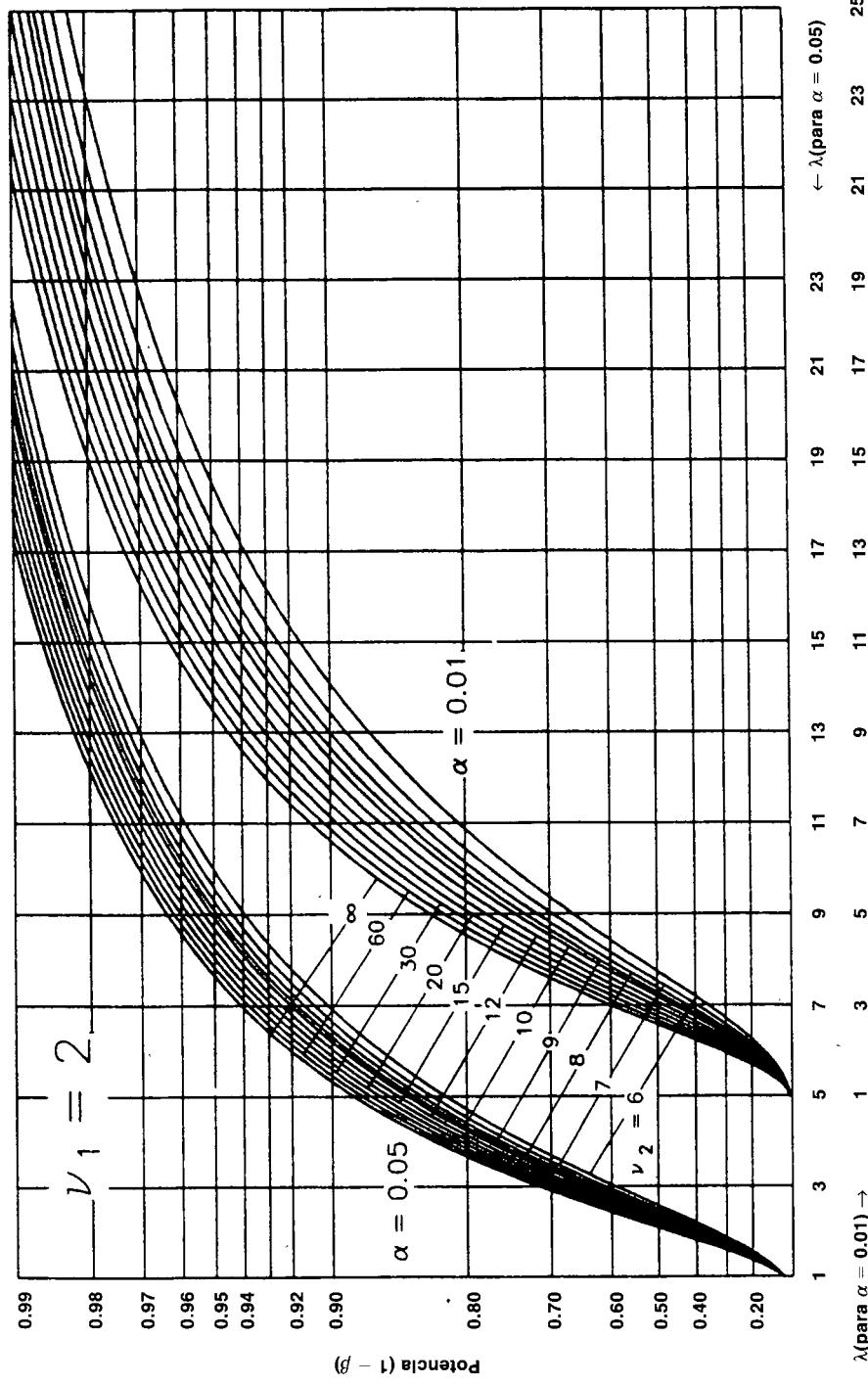


Tabla IX Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos fijos.

Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

Calculado por R. D. Axelson y R. O. Kuehl.

Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

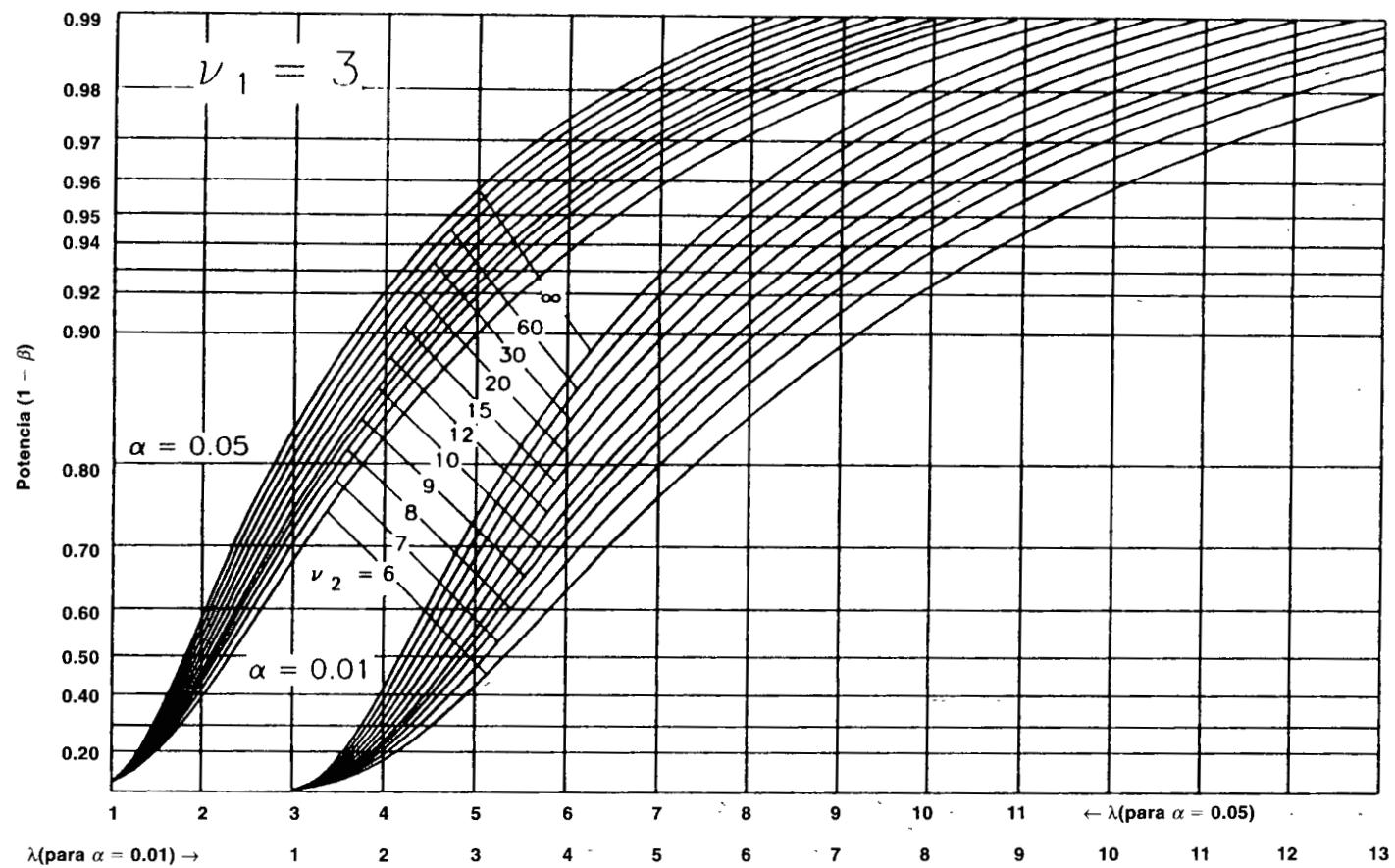


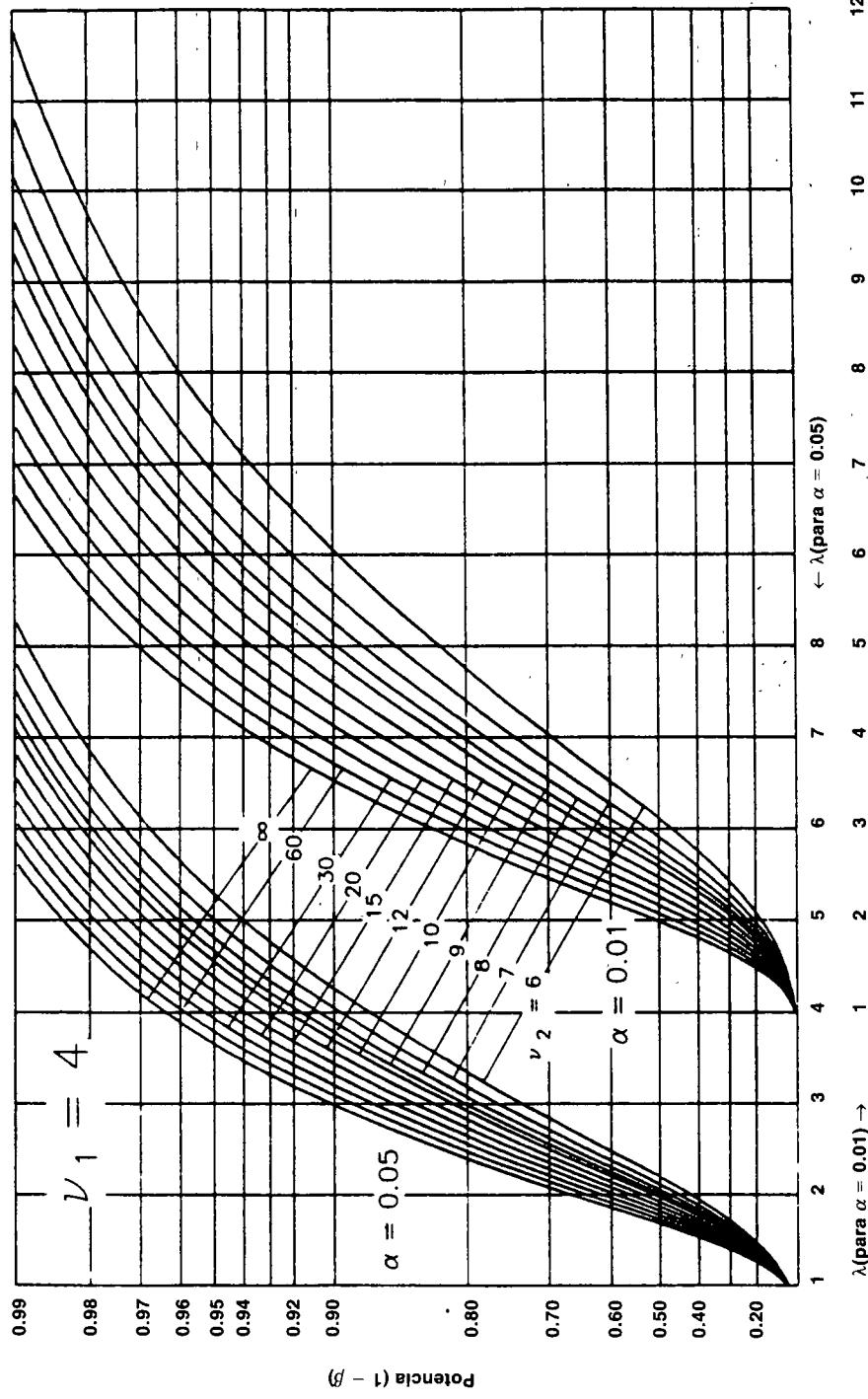
Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

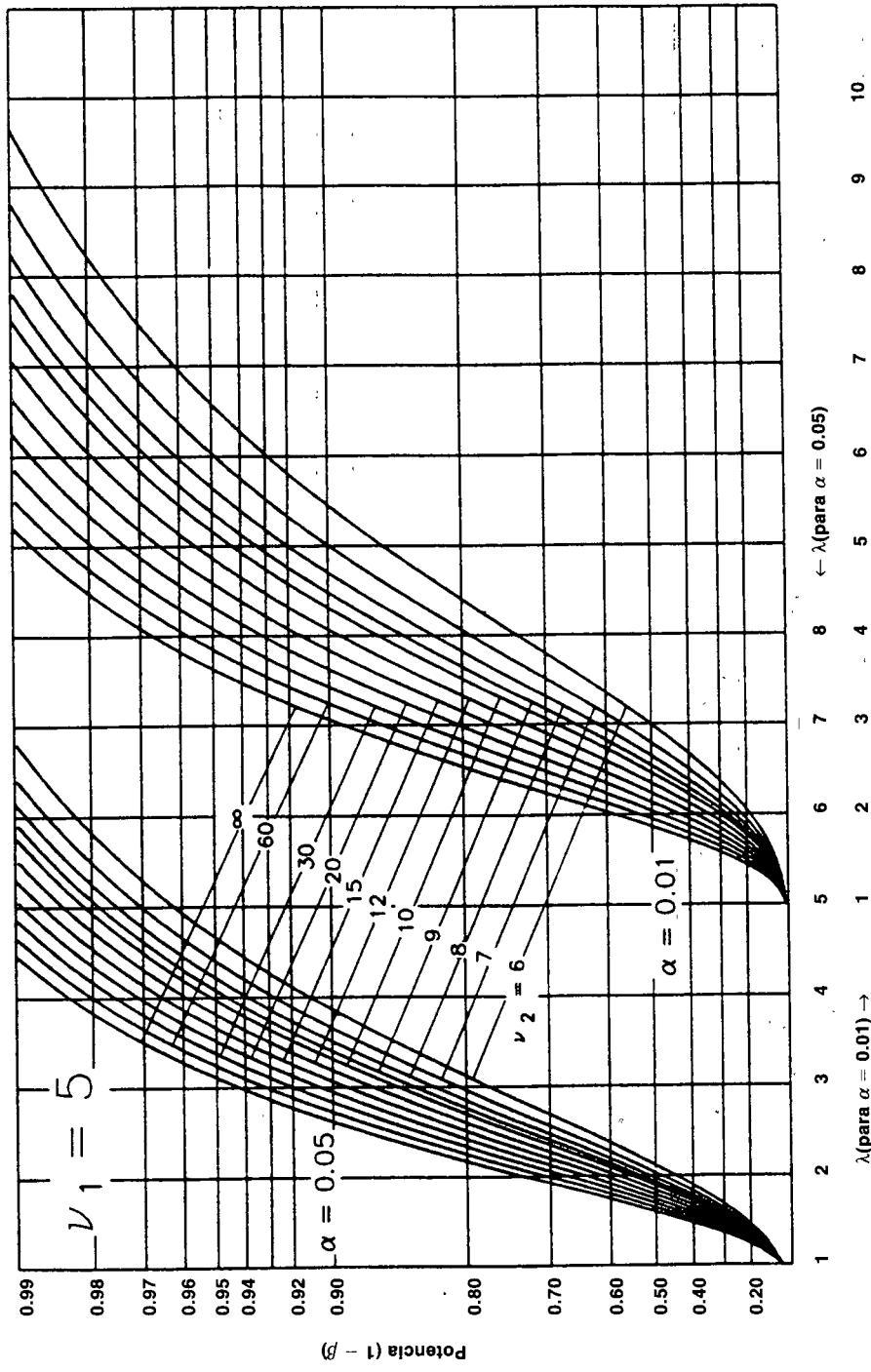


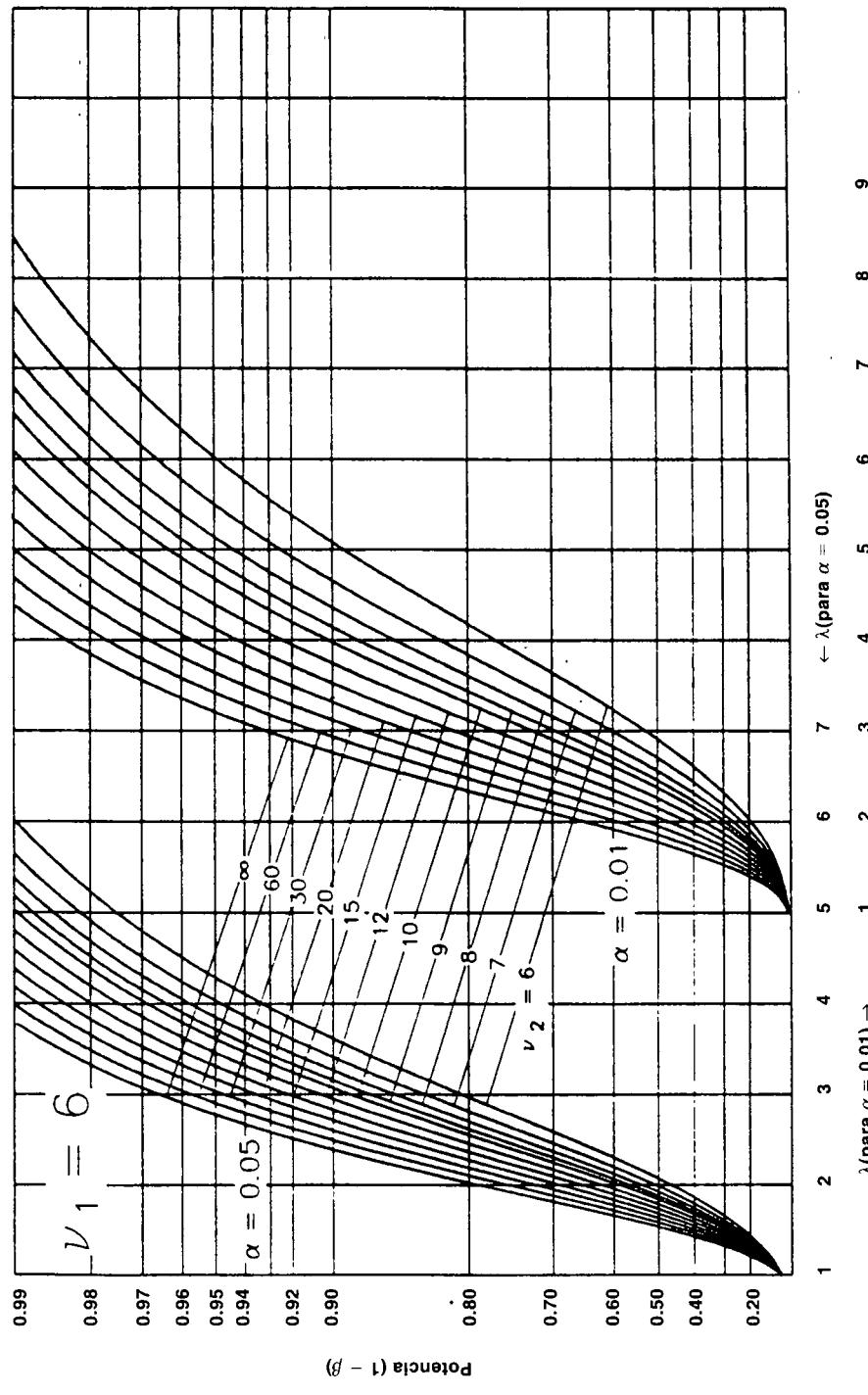
Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

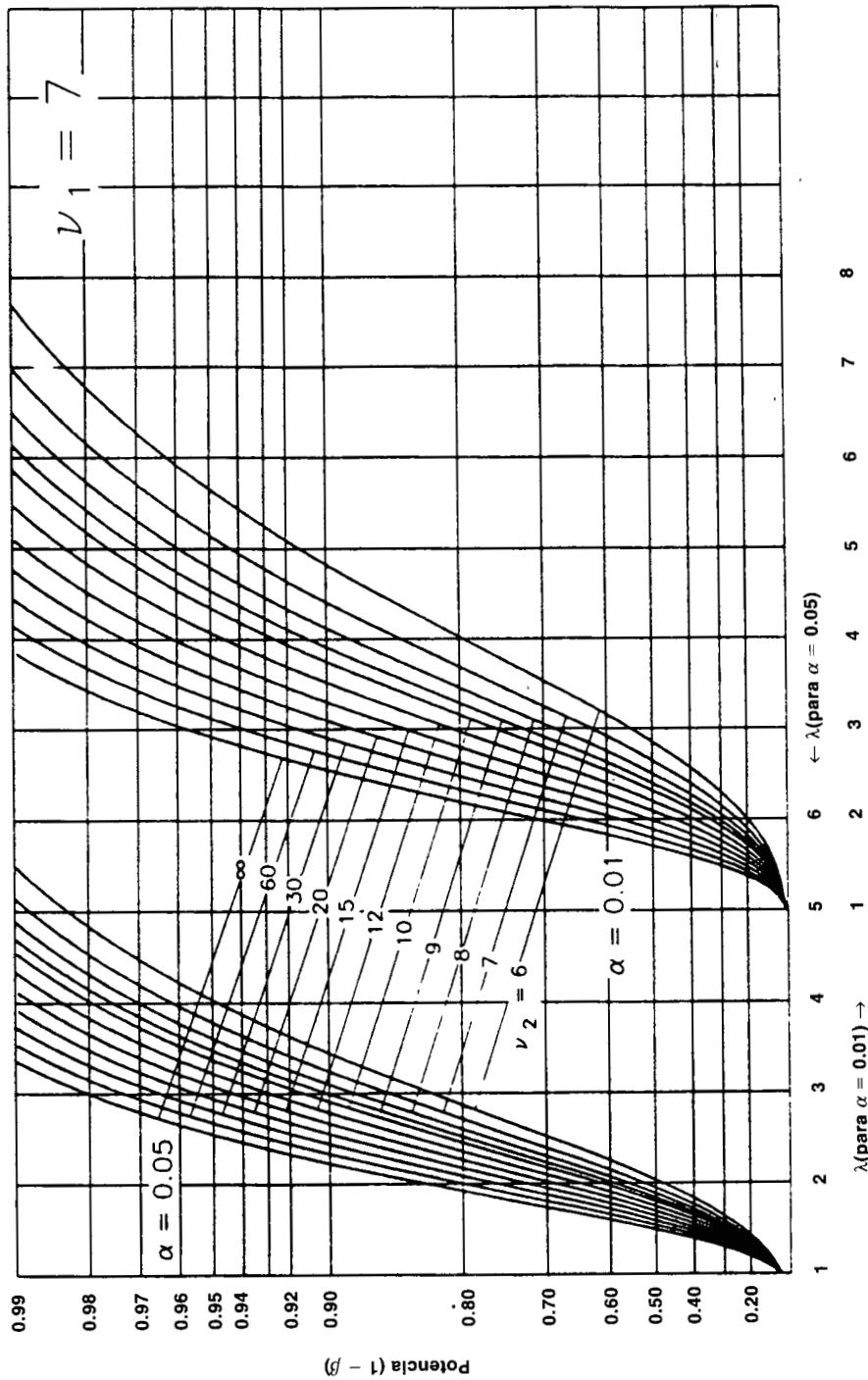


Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

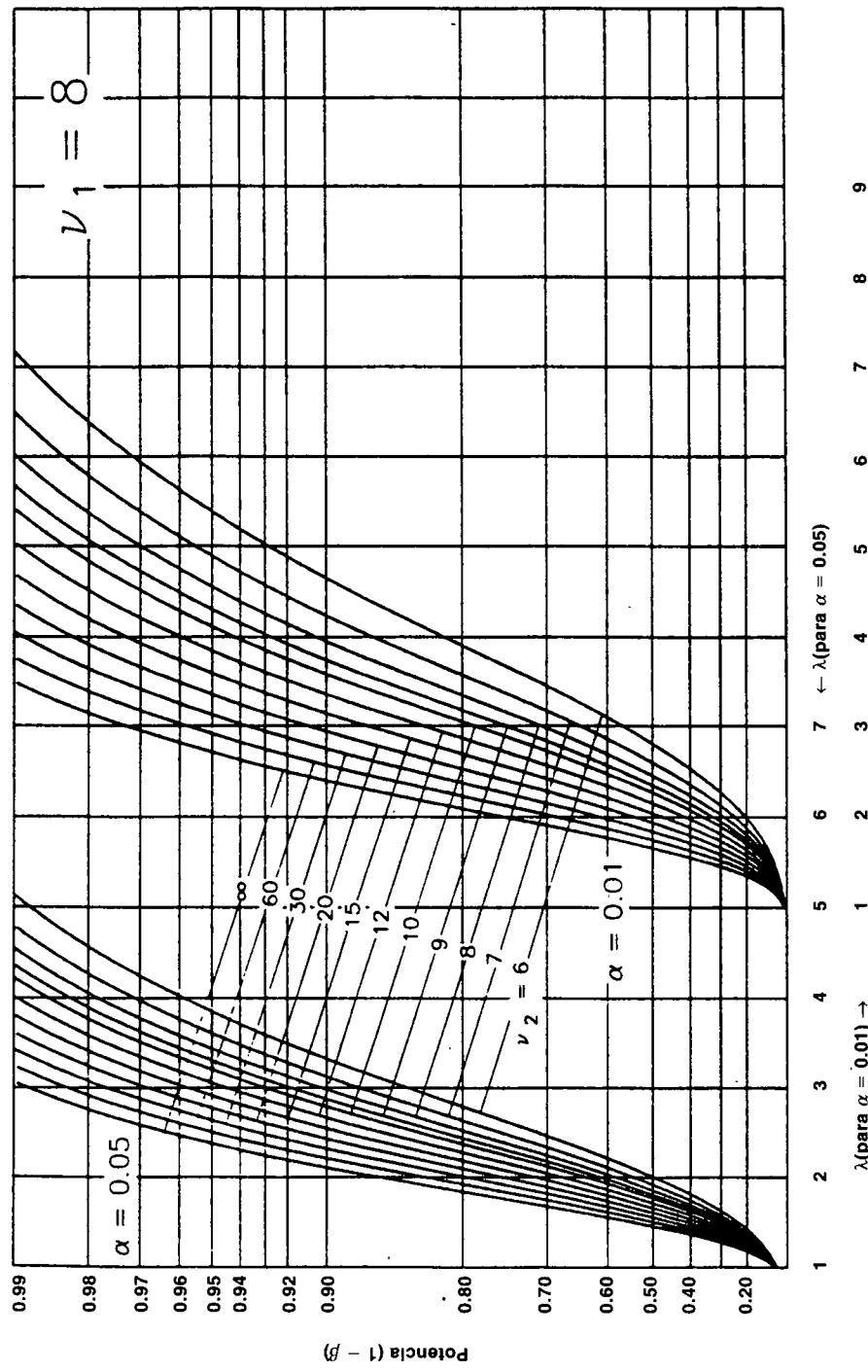


Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

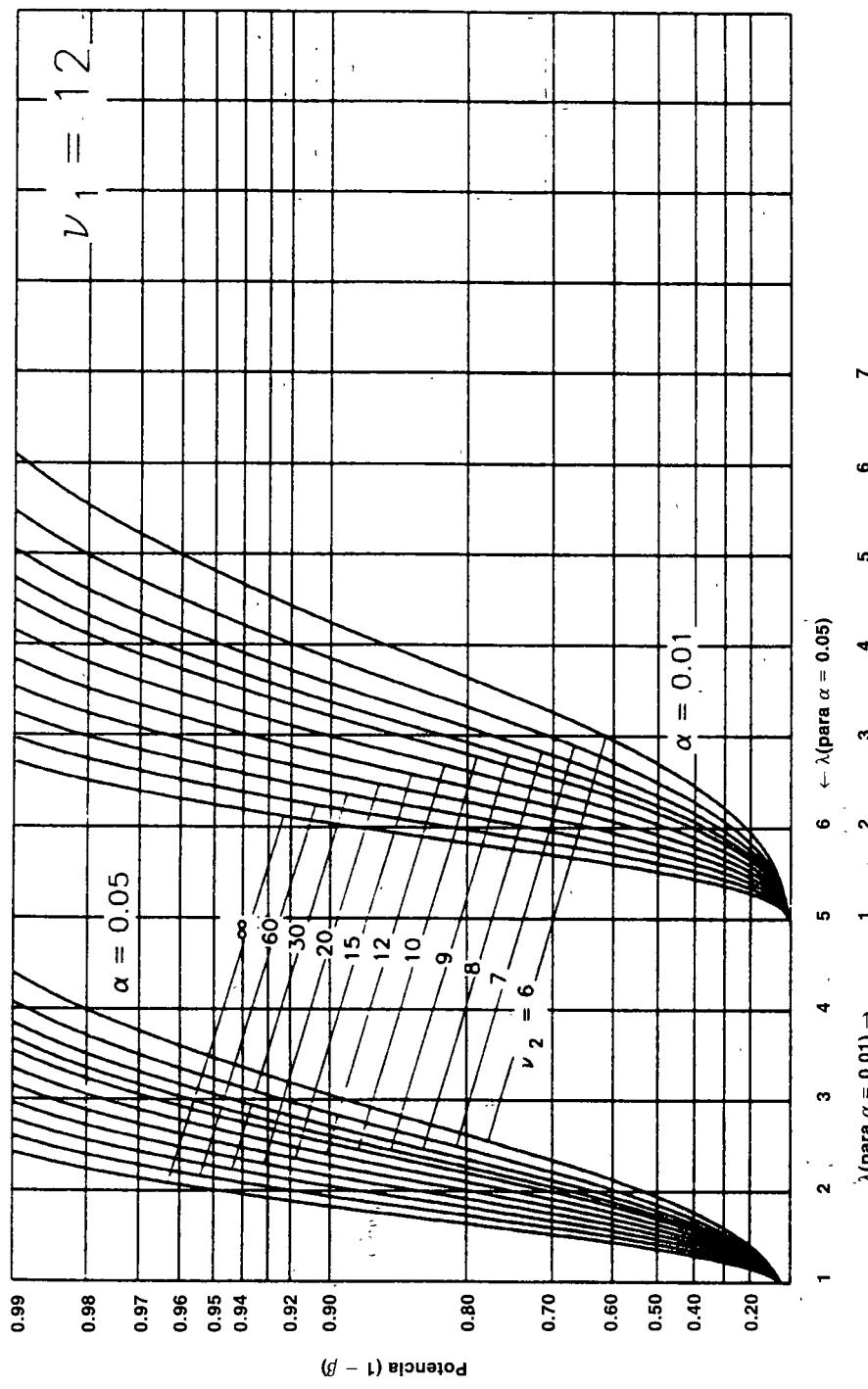


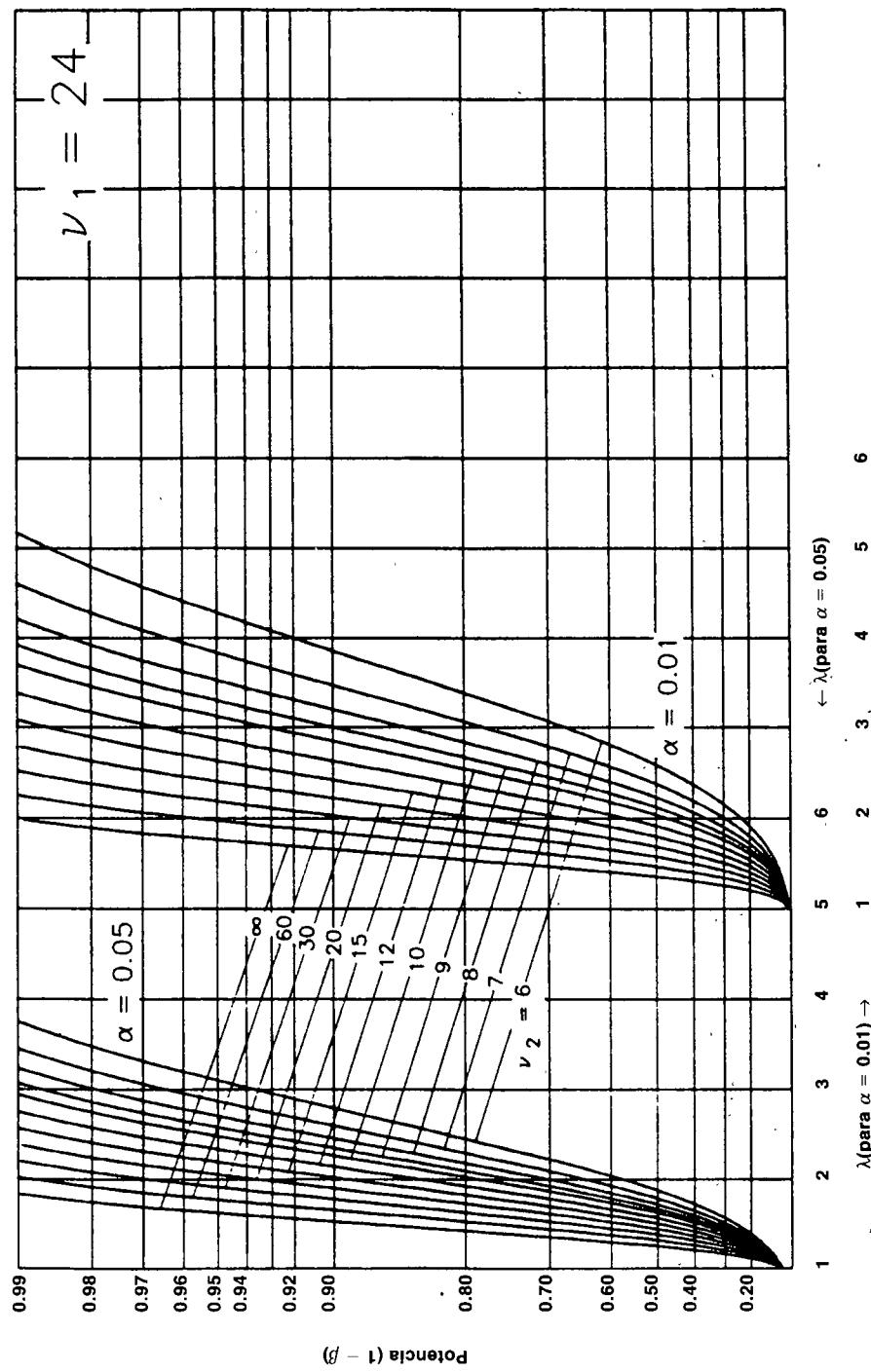
Tabla X Curvas de potencia de la prueba F para análisis de varianza del modelo de efectos aleatorios.

Tabla XI Polinomios ortogonales.

X_j	$t = 3$			$t = 4$			$t = 5$				$t = 6$					$t = 7$					
	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_5	P_1	P_2	P_3	P_4	P_5	P_6
1	-1	1	-3	1	-1		-2	2	-1	1	-5	5	-5	1	-1	-3	5	-1	3	-1	1
2	0	-2	-1	-1	3		-1	-1	2	-4	-3	-1	7	-3	5	-2	0	1	-7	4	-6
3	1	1	1	-1	-3		0	-2	0	6	-1	-4	4	2	-10	-1	-3	1	1	-5	15
4			3	1	1		1	-1	-2	-4	1	-4	-4	2	10	0	-4	0	6	0	-20
5							2	2	1	1	3	-1	-7	-3	-5	1	-3	-1	1	5	15
6											5	5	5	1	1	2	0	-1	-7	-4	-6
7																3	5	1	3	1	1
$\sum_{j=1}^t \{P_j(X_j)\}^2$	2	6	20	4	20		10	14	10	70	70	84	180	28	252	28	84	6	154	84	924
λ	1	3	2	1	$\frac{10}{3}$		1	1	$\frac{5}{6}$	$\frac{35}{12}$	2	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{7}{12}$	$\frac{21}{10}$	1	1	$\frac{1}{6}$	$\frac{7}{12}$	$\frac{7}{20}$	$\frac{77}{60}$
$t = 8$																					
X_j	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_1	P_2	P_3	P_4	P_5	P_6
1	-7	7	-7	7	-7	1	-4	28	-14	14	-4	4	-9			6	-42	18	-6	3	
2	-5	1	5	-13	23	-5	-3	7	7	-21	11	-17	-7			2	14	-22	14	-11	
3	-3	-3	7	-3	-17	9	-2	-8	13	-11	-4	22	-5			-1	35	-17	-1	10	
4	-1	-5	3	9	-15	-5	-1	-17	9	9	-9	1	-3			-3	31	3	-11	6	
5	1	-5	-3	9	15	-5	0	-20	0	18	0	-20	-1			-4	12	18	-6	-8	
6	3	-3	-7	-3	17	9	1	-17	-9	9	9	1	1			-4	-12	18	6	-8	
7	5	1	-5	-13	-23	-5	2	-8	-13	-11	4	22	3			-3	-31	3	11	6	
8	7	7	7	7	7	1	3	7	-7	-21	-11	-17	5			-1	-35	-17	1	10	
9							4	28	14	14	4	4	7			2	-14	-22	-14	-11	
10													9			6	42	18	6	3	
$\sum_{j=1}^t \{P_j(X_j)\}^2$	168	168	264	616	2184	264	60	2772	990	2002	468	1980	330	132	8580	2860	780	660			
λ	2	1	$\frac{2}{3}$	$\frac{7}{12}$	$\frac{7}{10}$	$\frac{11}{60}$	1	3	$\frac{5}{6}$	$\frac{7}{12}$	$\frac{3}{20}$	$\frac{11}{60}$	2	$\frac{1}{2}$	$\frac{5}{3}$	$\frac{5}{12}$	$\frac{1}{10}$	$\frac{11}{240}$			

Adaptado de *Biometrika Tables for Statisticians*, vol. 1, 1966, editado por E. S. Pearson y H. O. Hartley, con permiso de The Biometrika Trustees.

Tabla XII 10 000 dígitos aleatorios.

40092	81747	62038	32614	91062	95535	85635	46183	52525	86502
04040	77141	68126	39738	00241	90805	90476	31866	43819	72869
27744	59344	15764	27705	86115	10818	33029	47185	78385	47791
03012	35878	31415	02956	73710	73199	71014	68069	42556	25033
03756	14835	94850	14777	02637	98934	18678	52385	53701	97816
32815	29692	41490	78596	23967	53733	32292	33595	89471	51662
13702	22046	10820	74519	09094	42165	96372	43048	82936	34352
90988	43962	72505	00500	42042	73894	50219	71653	90329	61575
62355	05504	69481	95689	06051	27611	72996	91049	80893	71328
49076	28508	71392	13453	51243	18476	54021	05696	11779	15863
36273	09222	95599	32931	21871	98668	99815	94235	42693	51724
68993	03491	80984	48526	90843	29395	51200	97425	39987	36822
39920	31267	60848	05034	03234	64340	09434	53143	05462	34473
54411	06333	36913	08937	96649	36916	29084	26470	41677	97141
36589	59006	78749	47682	81371	87395	46397	88578	53407	45513
50992	10021	61655	46972	77136	72126	75735	54573	21754	34005
89336	50439	97392	55854	86657	02737	25120	67498	40731	63664
54553	95650	25878	12743	41984	83063	34507	74733	54853	04773
87225	55691	25090	31227	86467	65431	19301	21752	57547	99668
89701	46487	48959	69554	59786	04082	30302	18492	82721	90393
04307	47780	07824	26546	52978	22552	95034	57905	33641	63382
73569	68313	08308	32795	27608	20956	92853	60375	10364	56966
14979	67281	49874	26421	25443	48250	15627	85412	67329	71950
63914	98384	78123	83289	46311	38100	89622	74531	91274	67461
86390	62009	13480	13275	85547	67966	55299	24147	21527	38419
32007	57426	35585	85429	18445	49946	93822	81152	08889	94365
01836	33548	68505	64965	21008	99708	75050	86951	00673	73955
32452	64986	81436	50212	35425	13971	11767	48191	08463	83295
91424	82977	32735	80372	33184	56791	72463	94831	11268	26694
38244	80884	35719	07613	73294	39929	89473	92680	85910	06938
26010	21397	37765	63680	04611	37636	92158	23594	14873	09599
67054	42605	18159	00922	60210	80042	76776	48738	64415	99442
83513	91260	18937	89770	87417	66732	54208	99378	77020	84131
65512	49102	25594	51608	64913	61490	73883	29741	66253	68670
77810	94342	98901	37379	05225	97931	51528	10666	37760	89043
07968	43404	73783	69894	01357	34726	89661	89342	30408	82356
85370	05580	25323	86903	31950	82648	57589	58573	58332	55619
60923	13031	43571	62495	96389	46218	61748	42218	95879	97246
85495	42421	04882	67652	77291	12909	73833	45871	71388	82924
06328	30206	35409	78693	67795	64053	64309	51064	02516	29597
76447	35060	98444	14312	04058	85851	61351	22310	36618	90115
78866	29368	84660	13189	86648	88606	30739	67032	70300	45249
02448	59524	35486	74847	64272	26363	89711	29296	66690	35704
90016	85192	20040	97348	38337	39697	74895	21585	09245	63855
30806	23228	45362	65937	86967	98883	50230	70750	34638	60422
61224	84197	05618	34422	46881	35190	65108	83574	91763	74682
69058	80489	83177	79446	28867	06358	21001	13830	51949	33314
29731	88669	90396	86789	06896	21048	21903	65161	69464	18954
55831	36845	20465	09528	42037	57535	15875	83455	15612	99552
54654	11577	44676	07215	37288	90974	48352	07127	62359	36357

Calculado por R. D. Axelson y R. O. Kuehl.

Tabla XII 10 000 dígitos aleatorios.

26862	23753	16576	48185	26188	37728	25811	97801	49522	37980
88683	98306	07900	23707	99948	33525	74273	47346	28068	67382
11649	19469	97566	73268	84821	17286	78963	69433	27656	20853
11539	08341	75699	97873	28867	21957	62406	20655	43287	20888
56151	48566	02341	00975	07214	50747	88664	98610	27313	64162
69530	17535	27838	40610	50380	78140	07497	36527	50167	46282
15818	76764	92511	03502	03169	06702	59346	35361	77128	18245
50336	94817	84848	65960	76977	95100	63306	19380	46687	31515
28406	21331	85488	88274	48787	44075	86769	09610	02927	33135
72796	02583	10970	77788	83695	69263	20800	32501	91294	98923
19569	04841	41380	89265	01799	11059	93494	66138	13304	12325
75557	66470	34334	82621	74506	43534	26514	11624	42412	76386
11018	38364	01933	32767	56338	85843	84565	62142	29687	17325
91226	45783	50106	60482	03949	91614	06628	65204	65570	32278
14651	00799	86285	40491	42792	51441	91259	33280	21469	14226
45076	54205	08493	36125	46555	08566	12084	62827	00321	56568
26297	93649	18706	64045	25840	79460	46865	92165	93820	34790
92147	37854	71738	19132	03564	68174	93762	37616	05231	55992
78686	88798	72430	69712	59483	93455	12339	77474	55033	80771
39337	46482	68067	62770	22840	32935	56812	14135	19428	19452
22252	15585	55870	52572	14044	26189	89418	52296	38387	75269
57240	75441	73297	63779	59276	17017	85257	30527	12722	35094
04322	54208	11657	11593	70223	60145	99744	91446	70523	31993
27666	19664	78042	51445	47023	30784	69771	82770	55057	52078
53574	41181	17109	61114	83876	95597	41259	37891	84897	89452
57403	53175	09644	44756	43387	57407	48145	42504	63603	37123
47663	74298	42107	67427	31764	81212	60736	57428	14755	42171
23443	29258	09203	79058	55323	12976	32399	17237	85784	02217
51643	09345	84572	28560	87805	02876	45418	73275	39984	96349
80051	75581	23727	58813	75587	23116	83055	86758	00086	80536
43926	60090	60696	72918	51179	49157	07915	65483	78143	10451
10254	05670	92734	24229	38248	84834	02502	09713	34985	31628
59972	80565	94063	83430	54499	83544	62526	58198	51422	43211
60281	89538	02140	93302	13432	15221	69413	13619	12440	49242
16811	18472	29709	39684	72700	32277	03202	66143	91079	90161
01902	59532	16742	87759	74717	38255	98729	42024	47121	20304
41733	54639	64241	70303	57220	78894	00192	03386	38487	27908
43355	69142	53361	09024	11162	46052	93213	17886	91801	98031
70729	22517	60942	41250	37148	24255	03358	38681	54169	01229
99082	83541	65511	79195	35411	30115	34826	92891	16193	49832
06263	94134	73079	02018	93129	25999	62186	62716	39844	71598
25600	31706	45184	04039	12757	17543	57061	35372	14213	75793
38681	10679	56604	69704	00663	87405	30603	90223	22174	06601
80481	16829	91125	21262	03530	61672	46950	15458	10962	44808
47010	50624	57003	15278	31486	44583	73363	00197	45102	14852
77358	22444	92701	91049	78909	10623	57688	61120	83266	83874
38510	77101	96660	96889	78941	07529	22497	96254	04580	48105
81400	38106	00865	27113	91974	34324	39319	71900	27482	66416
99793	34087	52787	02700	29651	15165	64316	69420	08167	94154
55128	29344	80171	25092	80960	14721	59328	06345	91581	82399

Tabla XII 10 000 dígitos aleatorios.

27716	39430	74892	55062	77137	37327	49419	30022	98867	91330
88339	55349	44443	05790	77996	60410	41369	22078	31121	13179
37968	25059	37540	04687	53545	26673	40810	48176	06334	98994
15917	55975	42519	36245	01377	03940	29064	34000	65999	20130
70360	58659	09025	65477	56751	44808	61307	88623	89037	68398
00645	21621	36761	53781	73718	58716	81236	17964	70404	57782
89194	42646	63316	59445	18655	46159	16175	17398	72326	11634
55980	75632	55491	14356	56454	00347	62540	86968	68038	58485
13171	81146	28515	59825	88041	74881	65046	53098	59673	49321
86112	27709	22965	16439	21368	60597	39390	78060	05766	87653
43520	50388	33109	80118	52631	28818	77139	37368	90715	76626
38819	71158	07453	94947	52652	20698	05598	87513	45302	94056
33517	96172	08551	91814	11724	65092	65214	17879	79958	06629
83251	07374	39275	31071	41215	42198	15736	20401	62220	15274
41449	51749	75858	33478	82088	93677	32579	56769	13978	17131
84803	30272	85826	18935	24913	79617	72859	79433	94633	40520
84633	75447	06239	52081	94605	29414	90088	86603	01489	17407
01638	69018	93805	73036	07877	41875	73920	87707	71362	30956
28674	59616	07748	09703	49503	74360	14030	35940	10022	42009
83293	38075	26568	69884	37289	14617	18927	93178	30591	23286
68589	63091	88890	38333	71939	37194	21497	13995	35084	15634
45432	44958	68132	99054	63817	30182	73247	57265	53261	25571
11054	89885	77436	86004	68590	73162	92891	11720	71241	89156
42514	56335	12990	36130	46050	93096	63546	78406	62320	51346
07047	90954	35462	86360	24284	09102	25426	15567	25015	88667
17561	05233	52549	90437	92488	57833	83542	05540	91798	98934
05911	53885	56556	40850	14137	97683	84795	51261	66590	77263
17612	20461	62032	47587	17035	83371	94995	50315	41561	21500
06111	59316	99244	53028	21979	10783	44488	22348	98002	18472
74570	09252	44243	51742	74563	07363	56710	15735	73850	23924
44776	03402	35319	10026	27044	19091	51864	70008	56725	34529
83366	42305	34398	66549	99205	74699	57967	96962	17378	43224
01073	59291	56232	91521	98063	96809	25210	96350	46086	22857
42767	58685	31976	44953	26620	49569	85394	15279	74548	74456
13285	58416	71593	12007	65202	61687	48650	27987	23498	72343
76273	17196	84114	42240	14314	49976	68714	33538	06533	87191
24056	80981	20168	29047	20856	65202	16533	28520	11420	96331
49676	14882	65467	73855	03075	61191	64955	48555	91374	80971
64991	57009	77897	83153	95013	86277	70494	18046	56932	20434
28089	98413	31829	62562	48896	20935	94369	48925	97773	37099
96534	97921	13604	70506	29624	44005	92916	41939	35505	90644
50836	96073	16019	48998	21487	50844	90073	76007	87108	54137
28746	47219	30895	23263	60178	50722	15734	60946	89808	82700
80584	53948	44644	76912	98636	95840	78725	77378	30973	15414
32514	27970	73626	19182	31168	63792	07136	97354	59750	85256
15164	71201	39626	58961	33178	38641	85647	45973	19310	84457
39324	98893	22000	96405	15322	77189	92508	53154	47615	14436
83484	60576	62727	96233	01802	98190	76160	59963	84478	45070
42484	28641	47335	80096	84105	19111	74979	42318	84907	52870
65911	27235	53577	04477	95852	70219	82456	19107	99280	08322

Tabla XII 10 000 dígitos aleatorios.

94221	07150	80463	03658	07013	55333	56464	04052	26053	81893
84431	13952	08149	80802	09136	71238	55048	09919	40389	61875
05455	88879	13449	32774	53964	18829	21872	74294	85282	92231
70079	21122	27004	58731	21160	80238	68780	82384	20602	92898
11446	25521	48438	72828	21550	35582	32768	69773	38451	79113
45708	85141	04684	34025	71208	19992	51615	71231	88276	58194
79659	33506	87595	03537	59265	92924	83027	41286	01315	14180
29884	97311	35812	45807	77277	94148	67863	24496	87510	42172
43279	67351	27095	17012	94794	30437	68917	13450	03254	05915
76413	06486	82936	33337	78926	25558	33856	14237	17551	59814
22492	13665	36925	48848	50459	49600	52743	20066	76869	79832
58291	55383	67639	66799	62393	04016	17872	55180	62652	10594
27126	62422	62990	56929	84904	19431	98746	55471	68157	65921
00842	24512	32501	35772	28232	10170	72564	21429	20642	27757
73267	31540	46725	32423	36226	89466	48202	25074	75094	60558
00964	87753	51908	44409	93930	08466	98989	17080	34531	77828
19944	78118	79260	23252	34690	76692	19405	11712	51394	48362
70288	68077	60880	23921	11896	20005	08466	86560	31270	69837
79468	67331	29400	11097	23247	71989	00401	05865	67119	80106
46373	06322	61880	71922	24235	47410	66986	26995	12239	86031
61516	21667	78846	71110	41710	64814	57890	26749	11543	88532
41099	45697	50877	86382	62166	07811	91211	05631	26154	32851
11024	75195	61800	59089	42760	38788	66395	47948	35330	74353
70868	80366	38791	71706	26600	01965	70249	43459	66862	15575
87057	86347	29685	79945	79152	46353	73810	50543	60423	98381
22388	64789	68609	63841	44728	45553	38269	97970	56182	01730
46942	66348	10183	50169	19373	54011	74267	03039	40453	44193
29806	89043	20745	66576	16824	75842	76690	65355	04887	68928
14617	79295	99744	62723	68156	47838	17085	81836	43124	47199
25809	85025	49691	68786	43859	50193	72474	26608	21241	66151
34780	60969	42518	07858	17204	98628	46176	04768	96909	23402
94677	73718	64271	59301	31606	66479	97437	87625	69066	53509
12250	10164	28403	22306	32881	58027	08882	01821	67718	74001
15253	78715	01384	02794	91983	00765	88148	56337	31053	17782
74787	90110	69156	62945	69569	13906	55002	99567	49338	95426
57379	47097	40942	61028	56704	29036	32246	57688	76317	67112
25663	48264	05635	98825	24427	77318	19674	44417	51022	69841
88330	25095	40796	63642	62300	23016	32759	43912	14635	06138
27651	22887	84669	92353	24804	40367	78290	45034	17498	93457
44555	85772	01754	36006	91187	72110	23796	63114	12759	12726
43299	87460	52469	24806	79248	53238	39640	49745	48627	56065
08740	06533	80838	38622	15149	08813	83839	96731	26326	13098
01029	77011	70206	54340	62941	29887	87785	78345	90186	65065
77885	85449	53717	52475	89906	40846	13143	50758	29288	77939
09852	08312	18396	34775	60762	52712	97343	75834	22023	87492
00298	02548	91069	14424	10383	88365	27797	58311	75041	45545
22831	43108	29166	08517	07100	78997	90983	13085	19481	86602
43584	61723	29354	96769	77087	33578	93149	23121	12969	04072
69116	44955	37-85	63158	44994	41991	43607	82729	72299	06758
49035	19029	73264	77376	96988	72124	70367	68128	89698	06584

Tabla XIII 400 permutaciones aleatorias de los primeros 10 enteros.

9	5	6	1	6	5	3	1	6	6	1	3	1	7	2	7	6	10	10	3
7	8	9	5	1	1	8	8	5	5	9	10	2	4	5	5	4	2	4	2
10	7	1	8	5	6	1	3	3	9	3	1	3	6	9	9	9	4	3	5
6	6	7	2	3	10	4	10	9	8	8	2	5	1	7	1	2	1	9	1
3	4	4	4	9	3	10	5	1	7	7	8	7	5	6	8	1	3	8	10
2	10	2	9	7	9	5	6	7	2	5	5	4	9	1	6	8	8	2	8
8	9	3	3	2	4	2	7	4	4	4	6	10	2	4	10	5	9	1	7
1	2	8	7	10	8	7	9	8	3	6	9	9	3	3	4	10	7	6	9
4	3	10	6	8	2	6	4	10	1	2	4	6	10	10	3	3	5	7	6
5	1	5	10	4	7	9	2	2	10	10	7	8	8	8	2	7	6	5	4
5	5	10	3	9	3	3	2	9	9	3	10	2	6	9	2	2	7	9	5
4	9	8	4	7	5	10	5	6	3	10	7	7	3	7	10	4	3	2	2
7	8	7	2	2	6	8	1	10	10	9	4	8	1	3	6	7	6	6	1
3	7	1	9	10	9	9	6	4	4	4	1	3	5	4	3	8	5	1	9
10	3	3	10	1	8	1	10	1	8	5	5	6	4	8	5	1	1	3	8
2	1	2	6	5	4	4	9	2	5	6	9	10	10	2	1	9	10	8	10
8	10	5	5	3	7	2	3	3	7	7	8	5	7	1	9	6	4	7	7
1	2	6	7	6	10	5	4	5	6	8	6	9	9	6	8	5	2	5	3
6	6	4	1	4	2	7	8	7	1	2	3	4	8	10	7	10	9	4	4
9	4	9	8	8	1	6	7	8	2	1	2	1	2	5	4	3	8	10	6
1	9	4	9	2	3	10	10	7	6	9	5	3	7	8	3	1	6	1	8
4	6	10	10	5	6	7	5	4	3	7	7	5	9	7	4	7	1	8	10
9	8	9	4	8	1	2	9	8	2	10	4	10	6	2	5	5	5	3	4
2	3	7	3	7	10	5	1	1	4	1	6	1	1	5	9	2	7	2	3
10	4	8	5	1	7	9	4	5	7	6	9	4	10	1	8	8	10	9	1
8	10	6	1	3	5	1	8	3	1	8	10	6	2	9	10	4	2	7	5
5	7	3	2	10	2	4	2	10	10	2	3	2	3	10	6	6	4	6	6
7	5	2	7	6	4	3	6	6	9	5	1	7	5	3	7	10	9	4	9
3	2	1	6	4	9	6	3	9	8	3	8	8	8	4	1	9	8	10	2
6	1	5	8	9	8	8	7	2	5	4	2	9	4	6	2	3	3	5	7
3	8	4	4	8	7	8	10	8	2	1	2	10	4	7	7	2	2	7	9
9	7	8	9	6	6	6	7	1	4	8	4	7	6	4	8	5	1	4	4
2	1	6	1	3	3	10	6	10	5	2	5	3	2	9	4	3	10	1	3
1	10	2	3	2	4	7	2	2	8	9	7	4	10	6	6	8	7	6	2
5	6	9	7	9	1	2	9	4	10	3	9	9	8	3	2	4	8	9	8
8	2	10	8	10	2	1	3	9	1	7	3	1	3	8	9	1	3	10	7
10	9	3	6	1	9	4	8	3	6	6	1	8	5	2	1	9	5	5	6
6	5	7	2	4	8	5	1	6	9	4	10	6	1	1	3	6	6	3	1
4	4	1	5	7	10	9	4	5	3	5	6	2	9	5	10	7	4	8	5
7	3	5	10	5	5	3	5	7	7	10	8	5	7	10	5	10	9	2	10
1	5	7	5	4	8	8	2	3	4	3	2	6	5	1	7	8	5	1	1
5	3	4	9	9	7	3	6	3	2	1	1	8	4	3	9	4	7	8	9
4	1	9	8	2	9	7	8	7	7	6	9	9	3	8	2	9	1	9	5
10	7	1	1	6	1	9	2	1	1	10	8	4	6	7	6	5	9	4	10
2	2	3	3	10	2	4	9	8	9	8	3	7	10	6	3	7	6	7	6
3	10	10	10	5	4	2	7	3	3	4	4	4	5	1	5	5	3	4	10
8	6	5	6	8	5	5	10	5	5	2	5	2	7	2	8	1	10	2	8
6	9	8	2	3	10	6	5	6	8	7	10	10	2	10	4	10	8	5	4
7	4	2	7	1	3	1	1	10	10	5	7	1	8	4	1	2	2	6	7
9	8	6	4	7	6	10	4	4	6	9	6	3	9	9	10	6	3	3	3

Calculado por R. D. Axelson y R. O. Kuehl.

Tabla XIII 400 permutaciones aleatorias de los primeros 10 enteros.

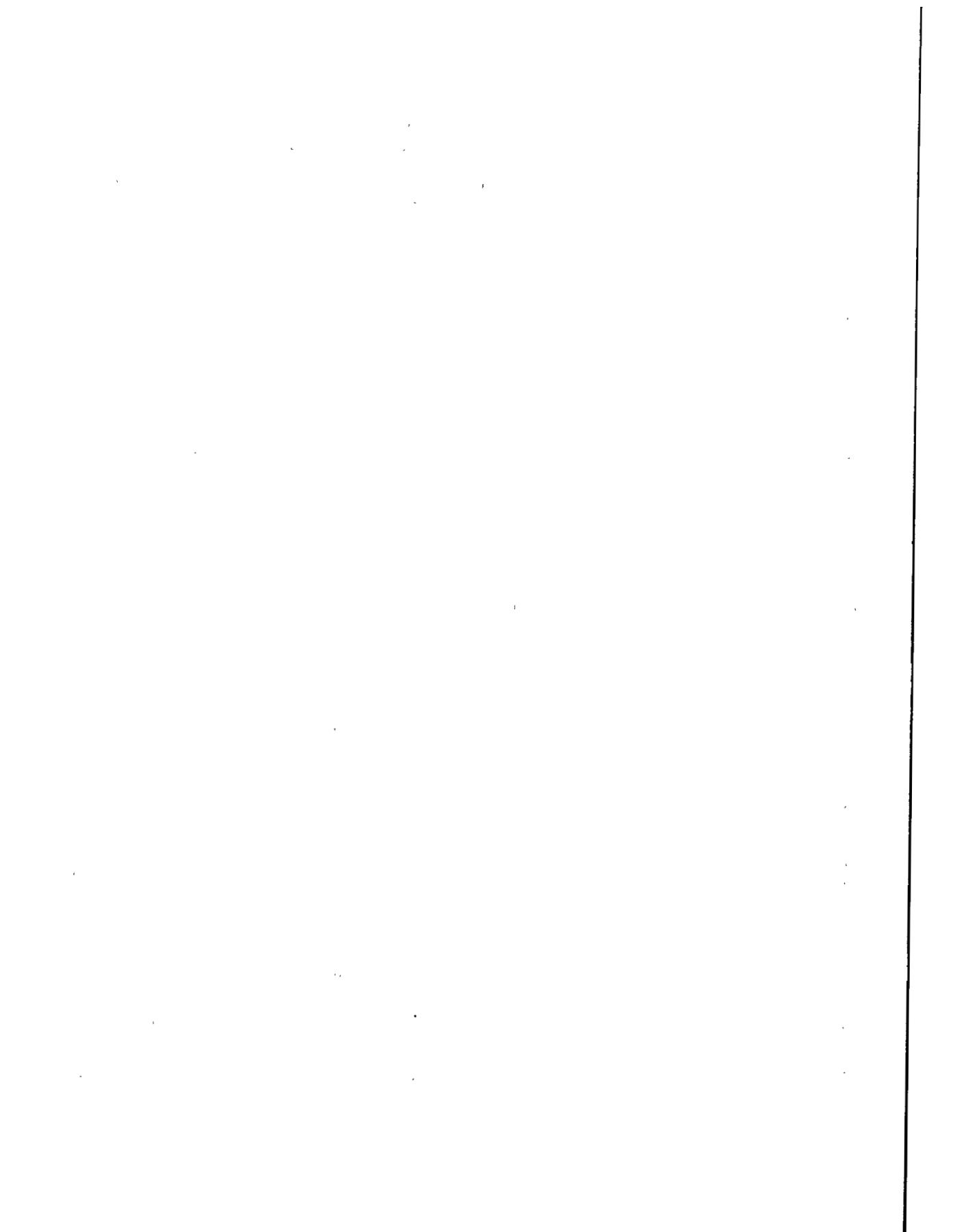
5	2	4	1	5	4	9	3	5	3	9	2	1	8	9	4	4	7	5	4
8	7	1	8	8	3	7	10	9	9	10	3	2	10	7	3	7	9	10	6
7	1	7	9	7	7	10	1	8	5	7	1	7	3	3	8	2	4	7	10
1	3	2	2	9	8	1	6	4	1	3	9	9	8	6	2	6	5	4	2
9	4	5	6	1	1	6	8	6	2	5	6	6	4	8	10	10	2	8	5
2	9	6	4	2	9	4	2	7	7	1	4	5	9	4	9	9	6	6	7
10	8	9	10	4	10	2	9	3	6	4	7	8	5	1	1	8	8	3	8
4	5	8	5	10	2	8	5	1	8	2	8	3	2	2	7	3	1	2	9
6	10	3	3	3	5	3	4	10	4	8	5	10	1	5	5	5	3	1	3
3	6	10	7	6	6	5	7	2	10	6	10	4	7	10	6	1	10	9	1
8	5	2	4	5	8	2	7	10	7	5	1	10	5	2	4	6	3	2	3
3	7	8	3	8	5	9	2	1	2	7	10	3	9	6	9	8	2	7	10
4	8	10	9	4	3	8	3	6	3	8	5	4	7	7	7	9	6	4	2
6	9	1	10	3	10	4	1	4	10	3	8	1	3	3	2	2	1	9	1
5	4	7	7	10	4	6	6	3	8	6	9	6	6	1	10	7	8	1	5
7	1	4	2	7	2	7	5	8	9	10	6	5	8	4	6	4	4	5	9
1	6	5	1	9	9	10	10	5	1	4	2	9	10	5	8	3	5	10	6
2	2	3	8	6	6	1	8	9	6	2	7	7	2	8	5	10	7	8	4
10	3	9	5	2	7	3	9	7	4	9	3	2	4	9	3	5	9	3	7
9	10	6	6	1	1	5	4	2	5	1	4	8	1	10	1	1	10	6	8
9	5	3	4	2	9	9	4	5	9	3	2	8	10	3	6	7	7	10	4
6	2	1	9	4	6	3	7	9	3	9	4	3	2	4	7	6	9	4	1
1	8	7	1	1	5	8	1	1	6	5	5	2	6	9	2	8	3	6	5
7	4	5	2	10	3	7	10	4	7	10	10	5	1	2	8	10	10	2	8
8	3	8	10	5	10	1	3	7	4	7	9	7	4	1	4	9	5	5	3
4	1	10	6	3	7	4	6	3	1	6	8	4	8	5	3	5	2	7	2
10	10	6	7	8	4	5	5	6	5	8	1	10	9	7	9	3	8	8	9
2	7	9	3	9	2	2	9	2	2	1	6	6	3	10	5	1	4	3	7
5	9	2	5	6	8	10	2	10	8	4	3	9	5	8	1	4	6	1	6
3	6	4	8	7	1	6	8	8	10	2	7	1	7	6	10	2	1	9	10
1	9	1	7	3	3	2	9	9	3	4	7	5	8	7	6	6	2	1	4
4	6	9	2	2	5	1	5	7	8	10	9	3	4	5	1	10	1	6	2
10	4	6	4	5	1	6	7	8	9	2	10	8	2	2	5	9	3	7	9
2	2	5	6	8	4	8	4	5	6	5	3	9	10	4	3	4	7	9	1
5	7	7	3	6	10	4	1	6	1	1	4	4	7	9	9	2	6	10	8
9	3	3	5	1	2	5	6	1	5	9	8	1	6	10	10	3	8	4	10
8	1	4	1	10	8	10	10	10	2	7	2	2	1	1	4	7	4	2	6
7	5	10	9	9	6	9	3	3	7	8	5	6	3	6	2	8	5	8	5
3	10	8	10	4	7	3	2	4	10	6	1	10	5	3	7	1	10	3	7
6	8	2	8	7	9	7	8	2	4	3	6	7	9	8	8	5	9	5	3
6	8	9	9	6	7	9	2	5	3	8	6	8	9	6	10	3	2	7	1
8	6	7	10	4	8	7	8	8	1	10	9	9	2	4	2	1	10	6	3
9	10	10	8	5	2	5	1	6	9	7	7	1	5	2	8	9	1	10	5
10	2	2	2	2	1	3	5	10	7	5	8	4	4	10	1	4	5	9	10
1	5	5	6	7	9	8	9	3	2	4	3	6	3	8	6	5	9	3	9
7	7	4	4	1	6	1	4	7	8	3	2	7	7	1	4	10	6	4	6
3	9	1	3	9	3	2	6	2	6	1	10	10	10	7	9	6	8	5	4
4	1	3	7	8	4	6	7	4	5	9	4	2	8	3	7	7	3	2	2
5	3	6	5	10	5	4	3	9	10	2	1	5	1	5	3	2	4	8	7
2	4	8	1	3	10	10	10	1	4	6	5	3	6	9	5	8	7	1	8

Tabla XIII 400 permutaciones aleatorias de los primeros 10 enteros.

10	8	8	2	8	1	7	6	7	6	1	7	2	5	4	4	9	1	3	7
4	1	4	8	6	6	8	4	1	8	4	8	1	8	2	8	1	3	10	4
7	4	3	3	10	4	5	1	8	2	2	6	7	2	9	1	5	10	4	10
1	10	10	4	4	7	3	9	6	10	3	5	5	8	5	7	8	4	1	5
6	9	5	10	2	9	2	5	9	9	9	1	8	3	8	9	4	2	8	8
2	2	6	6	1	3	9	2	3	5	10	4	4	9	7	6	7	6	9	6
3	6	7	5	5	10	1	10	5	3	7	9	10	7	1	5	3	8	5	9
8	5	9	1	9	8	4	7	10	1	8	10	3	10	3	2	2	9	6	3
9	3	2	9	3	2	10	3	4	4	5	3	9	6	10	3	6	5	7	2
5	7	1	7	7	5	6	8	2	7	6	2	6	1	6	10	10	7	2	1
10	1	10	9	5	5	1	6	5	8	9	4	3	3	7	10	8	8	1	10
7	5	3	10	4	10	3	9	3	10	4	3	5	4	10	6	9	4	7	2
4	3	7	4	2	8	6	8	1	4	2	6	8	9	2	5	4	9	6	7
5	2	9	1	3	2	7	10	2	2	7	8	1	10	9	4	10	2	4	6
9	10	2	3	6	9	9	4	7	5	10	10	6	5	6	2	7	1	10	1
6	9	8	7	9	7	10	7	8	9	5	2	9	1	8	9	1	10	2	3
3	6	6	6	10	1	2	2	4	6	6	7	10	6	3	7	6	6	8	8
2	4	1	2	7	3	8	1	6	1	8	5	4	7	4	1	5	7	9	5
1	8	5	8	1	6	5	5	9	7	3	9	7	8	5	3	3	3	5	9
8	7	4	5	8	4	4	3	10	3	1	1	2	2	1	8	2	5	3	4
3	3	2	8	9	1	10	7	10	9	1	4	8	2	5	9	9	7	1	1
2	8	10	5	3	7	3	1	2	3	7	3	9	1	3	4	3	2	9	4
9	7	5	9	1	2	4	3	9	7	6	9	7	7	2	1	2	10	8	9
6	2	6	2	2	10	1	5	6	5	9	10	4	3	10	2	10	8	7	10
7	9	4	7	10	4	5	6	3	6	3	5	1	5	8	7	6	3	4	8
5	1	3	1	7	5	7	8	8	1	2	1	5	9	4	6	8	1	2	6
8	4	7	6	8	8	8	9	4	10	10	6	2	4	9	8	1	4	5	2
4	6	1	4	5	6	2	10	5	2	5	8	10	6	7	3	4	6	10	7
1	5	8	10	6	3	9	2	7	4	4	2	3	10	6	10	7	5	3	3
10	10	9	3	4	9	6	4	1	8	8	7	6	8	1	5	5	9	6	5
7	3	5	2	9	2	2	1	6	7	1	7	9	7	2	1	1	7	2	2
4	5	10	9	10	9	1	9	1	6	3	5	8	6	4	7	5	5	7	10
3	7	6	10	5	8	4	6	7	3	5	4	10	5	10	10	2	9	5	1
5	1	9	4	4	7	10	7	10	1	10	6	5	3	8	3	7	4	1	3
2	2	3	8	1	6	8	2	5	9	8	1	3	1	6	8	8	2	3	4
10	4	2	7	2	3	7	4	2	5	4	2	1	9	9	9	3	1	9	8
1	9	1	5	7	1	6	5	3	8	2	8	4	8	7	6	10	10	8	6
8	10	4	6	6	4	5	3	4	2	6	9	7	2	1	2	4	8	10	7
6	8	8	3	8	5	3	10	8	10	7	10	2	10	5	4	9	6	6	5
9	6	7	1	3	10	9	8	9	4	9	3	6	4	3	5	6	3	4	9
2	3	4	6	2	5	4	4	1	4	1	3	6	1	10	6	5	2	3	9
4	4	1	8	5	6	5	10	10	5	10	7	4	9	9	1	6	4	6	6
8	10	3	2	8	4	2	6	2	1	3	5	2	5	3	9	1	6	7	8
6	5	10	4	3	2	3	8	7	8	9	1	9	4	6	8	7	3	10	5
9	1	2	1	10	3	1	3	9	6	4	8	10	10	4	5	8	5	4	4
5	8	7	3	6	9	6	9	4	9	8	2	3	2	7	3	9	7	8	10
7	7	9	9	4	8	8	5	6	2	2	10	5	7	8	10	2	1	2	3
1	6	5	10	7	7	9	1	5	10	7	4	1	3	1	7	4	10	5	7
3	2	6	7	9	1	10	2	3	7	6	6	7	6	5	2	3	8	9	1
10	9	8	5	1	10	7	7	8	3	5	9	8	8	2	4	10	9	1	2

Tabla XIII 400 permutaciones aleatorias de los primeros 10 enteros.

3	9	7	9	1	2	10	5	3	5	5	5	3	5	4	2	5	3	7	4
5	7	9	5	5	7	5	4	6	10	10	2	5	10	2	3	3	7	6	7
9	4	6	10	6	4	8	9	5	8	3	10	10	2	7	10	4	1	3	6
10	3	3	2	10	8	2	6	1	9	4	4	9	8	5	5	8	4	8	8
7	2	5	6	9	1	9	1	10	3	9	3	1	4	6	9	10	5	4	10
2	8	4	3	8	6	6	3	4	6	6	7	4	1	1	7	6	6	9	5
4	10	2	4	3	9	7	7	8	2	1	6	2	6	3	6	1	10	2	1
1	5	10	7	4	5	3	8	7	7	7	8	7	9	9	8	7	8	10	2
8	1	8	1	7	10	1	10	9	4	8	9	6	3	10	1	9	9	1	3
6	6	1	8	2	3	4	2	2	1	2	1	8	7	8	4	2	2	5	9
8	4	7	3	10	1	8	4	9	4	4	8	7	9	3	1	10	10	9	3
4	6	9	1	6	10	4	1	6	7	8	5	10	6	6	7	5	7	10	9
6	2	8	5	7	2	2	6	4	8	9	10	4	4	10	10	7	8	5	5
2	1	5	7	8	8	1	5	5	5	6	7	6	3	8	8	1	9	1	1
10	9	6	9	9	3	5	9	2	2	3	6	8	10	7	5	4	3	2	2
3	7	1	6	5	9	6	3	1	6	7	9	3	2	9	3	2	4	6	10
7	3	4	8	3	6	3	7	8	3	5	4	1	5	2	2	8	5	3	4
9	8	10	2	2	7	10	10	10	9	2	1	2	8	4	6	3	6	7	7
1	10	3	10	4	4	7	2	3	10	1	2	5	7	1	9	6	2	4	6
5	5	2	4	1	5	9	8	7	1	10	3	9	1	5	4	9	1	8	8
6	5	5	1	1	6	9	8	7	2	2	2	1	6	5	7	3	6	7	8
2	7	2	10	8	5	5	7	3	9	6	3	9	5	8	1	6	2	10	9
4	6	10	3	10	9	3	2	8	7	7	8	3	7	10	3	5	8	6	5
3	9	7	5	7	10	7	6	9	3	3	5	6	4	7	10	8	1	2	7
10	8	1	6	4	7	2	4	4	6	10	1	10	1	4	9	10	3	1	2
9	1	4	9	2	8	6	3	5	5	5	9	2	10	6	5	1	5	4	1
5	2	8	8	5	3	8	5	1	10	1	6	5	9	1	8	9	7	5	10
1	10	6	4	3	4	4	10	2	1	8	7	8	2	3	2	4	9	8	3
8	3	3	7	6	1	1	1	6	4	4	4	4	3	9	6	2	4	9	6
7	4	9	2	9	2	10	9	10	8	9	10	7	8	2	4	7	10	3	4
4	1	5	1	4	5	2	3	10	4	2	4	4	5	8	5	6	2	10	2
10	9	1	7	2	10	4	6	1	10	4	8	1	9	5	4	1	6	6	4
7	4	9	8	1	9	6	1	9	2	6	2	7	7	10	1	3	4	7	7
6	2	10	3	7	7	1	4	3	7	10	10	5	8	7	10	10	8	1	9
3	7	4	2	10	3	5	8	5	1	7	6	8	2	4	6	5	5	9	3
9	10	6	10	9	6	3	7	7	6	5	9	9	1	3	8	8	9	8	6
5	8	2	6	5	2	7	5	2	9	1	7	10	10	6	3	4	3	2	8
8	6	3	4	8	1	9	9	8	8	8	5	3	3	1	9	9	10	4	1
1	5	8	5	6	4	10	2	4	3	3	3	2	6	2	7	7	7	5	5
2	3	7	9	3	8	8	10	6	5	9	1	6	4	9	2	2	1	3	10
9	2	2	2	3	1	5	7	2	2	10	1	4	1	6	2	2	2	9	10
7	5	10	6	4	5	7	6	6	1	8	5	2	5	3	3	8	10	9	1
4	3	1	7	10	9	1	2	8	3	3	10	9	7	10	8	7	7	8	2
6	10	9	9	6	3	8	1	1	6	6	2	7	9	4	10	6	5	6	6
10	1	7	10	2	2	6	3	9	10	4	9	10	4	9	9	4	6	5	8
2	7	8	3	1	6	4	9	10	5	2	6	3	2	6	6	3	8	2	7
3	6	6	5	8	8	9	8	7	7	7	8	6	3	8	4	1	1	7	9
8	9	5	4	9	10	10	10	4	9	1	7	8	10	2	5	10	2	1	3
5	8	3	1	5	7	3	5	3	8	5	3	1	8	5	1	5	4	3	4
1	4	4	8	7	4	2	4	5	4	9	4	5	6	7	7	9	3	4	10



Respuestas a ejercicios selectos

Capítulo 2

1. a. $y_{ij} = \mu_i + e_{ij}$ $i = 1, 2, 3$ $j = 1, 2, \dots, 5$
 μ_i = media del tipo de semáforo, e_{ij} = error experimental

b. e_{ij} aleatorio, independiente con media 0 y varianza σ^2 igual para todos los semáforos

c.

Fuente	gl	SC	CM
Semáforo	2	1 202.63	601.31
Error	12	137.83	11.49
Total	14	1 340.46	

d. 35.5, 20.9, 14.0 error estándar = 1.52

e. (32.2, 38.8), (17.6, 24.2), (10.7, 17.3)

f. $F_0 = 52.4$, $F(0.05, 2, 12) = 3.89$, se rechaza

g. $5\mu_1 = 177.4$, $5\mu_2 = 104.5$, $5\mu_3 = 70.0$

3. a. $y_{ij} = \mu_i + e_{ij}$ $i = 1, 2, \dots, 5$ $j = 1, 2, \dots, 5$

μ_i = media de tratamiento, e_{ij} = error experimental

b. e_{ij} aleatorio, independiente con media 0 y varianza σ^2 para todos los tratamientos

c.

Fuente	gl	SC	CM
Tratamiento	4	48 569	12 142
Error	20	3 110	156
Total	24	51 679	

d. 86.37, 112.85, 206.16, 110.69, 87.97 error estándar = 5.59

e. (74.71, 98.03), (101.19, 124.51), (194.50, 217.82), (99.03, 122.35), (76.31, 99.63)

f. $F_0 = 78.08$, $F(0.05, 4, 20) = 2.87$, se rechaza

g. $5\mu_1 = 431.87$, $5\mu_2 = 564.26$, $5\mu_3 = 1 030.78$, $5\mu_4 = 553.47$, $5\mu_5 = 439.86$

5. a. $y_{ij} = \mu_i + e_{ij}$ $i = 1, 2, \dots, 5$ $j = 1, 2, \dots, r_i$ $r_1 = 4$, $r_2 = r_3 = r_4 = r_5 = 2$

μ_i = media de tratamiento, e_{ij} = error experimental

b. e_{ij} aleatorio, independiente con media 0 y varianza σ^2 para todos los tratamientos

c.

Fuente	gl	SC	CM
Tratamiento	4	0.175969	0.043992
Error	7	0.000094	0.000013
Total	11	0.176062	

- d. media(err est), 0.049(0.0018), 0.086(0.0026), 0.116(0.0026), 0.187(0.0026), 0.397(0.0026)
e. (0.045, 0.053), (0.080, 0.092), (0.110, 0.122), (0.181, 0.193), (0.391, 0.403)
f. $F_0 = 3.28475$, $F(0.05, 4, 7) = 4.12$, se rechaza
g. $4\mu_1 = 0.197$, $2\mu_2 = 0.171$, $2\mu_3 = 0.231$, $2\mu_4 = 0.374$, $2\mu_5 = 0.794$

7. a. Use $D = 30$, $t = 5$, $\sigma^2 = 156$ en la ecuación (2.26). Si $r = 7$, $\Phi = 2$ y la potencia se aproxima a 0.93 en el apéndice IX para $v_1 = 4$ y $v_2 = 30$

Capítulo 3

1. a. estimación(err est), (i) 18.03(1.86), (ii) 6.90(2.14)
b. (i) 1803.603, (ii) 119.025, $SCT = 1202.628$
c. $t(0.025, 12) = 2.179$, (i) $t_0 = 9.71$, se rechaza, (ii) $t_0 = 3.22$, se rechaza
d. $ECM = 11.49$, $F(0.05, 1, 12) = 4.75$, (i) $F_0 = 94.34$, se rechaza, (ii) $F_0 = 10.36$, se rechaza
e. $t_0^2 = F_0$

3.

Fuente	gl	SC	CM	F ₀
Nitrógeno	4	4 994.8	1 248.7	5.61
Lineal	1	2 958.4	2 958.4	13.29
Cuadrático	1	1 783.1	1 783.1	8.01
Cúbico	1	67.6	67.6	0.30
Cuarto	1	185.7	185.7	0.83
Error	15	3 338.0	222.5	

F_0 para lineal y cuadrático excede $F(0.05, 1, 15) = 4.54$, entonces cuadrático es el mejor ajuste polinomial.
Coeficiente estimado (err est) $\bar{y} = 142.60(3.34)$, $a_1 = 8.60(2.36)$, $a_2 = -5.64(1.99)$, $a_3 = 1.30(2.36)$, $a_4 = -0.81(0.89)$

La ecuación de regresión cuadrática es $\hat{y} = 114.11 + 0.6234N - 0.0023N^2$

Nitrógeno	Media	Predicción	Err Est	IC 95%
0	112.0	114.1	7.0	(99.2, 129.0)
50	145.5	139.6	4.5	(130.0, 149.2)
100	149.0	153.9	5.2	(142.8, 165.0)
150	157.5	156.8	4.5	(147.2, 166.4)
200	149.0	148.5	7.0	(133.6, 163.4)

5. a. $d(0.05, 7, 8) = 2.81$ $M = 2.81 \sqrt{2(6.02)/2} = 6.89$

Paciente	\bar{y}_i	$\max_{j \neq i} \bar{y}_j$	D_i	$D_i - M$	$D_i + M$	ICS 95%
1	167.0	244.6	-77.6	-84.5	-70.7	(-84.5, 0)
2	185.5	244.6	-59.1	-66.0	-52.2	(-66.0, 0)
3	104.0	244.6	-140.6	-147.5	-133.7	(-147.5, 0)
4	214.9	244.6	-29.7	-36.6	-22.8	(-36.6, 0)
5	149.0	244.6	-95.6	-102.5	-88.7	(-102.5, 0)
6	169.4	244.6	-75.2	-82.1	-68.3	(-82.1, 0)
7	160.5	244.6	-84.1	-91.0	-77.2	(-91.0, 0)
8	244.6	214.9	29.7	22.8	36.6	(0, 36.6)

- b. Paciente 8 con $D_i + M > 0$, se elige como sujeto con colesterol más alto

7. $d(.05, 3, 16) = 2.59, D(.05, 3) = 2.59 \sqrt{2(6.875)/5} = 4.3$

Tratamiento	Media	$\bar{y}_i - \bar{y}_c$	ICS 95%	
			(L, U)	
Control	42.2	—	—	—
3% glucosa	29.0	-13.2	(-17.5, -8.9)	
3% fructuosa	27.6	-14.6	(-18.9, -10.3)	
3% sucrosa	34.0	-8.2	(-12.5, -3.9)	

Todos los tratamientos difieren significativamente del control ya que ninguno de los intervalos incluye el 0. 3% de fructuosa es el más lejano al control y 3% de sucrosa es menos lejano según los límites superiores.

9. $F(.05, 2, 12) = 3.89$

Contraste	Estimación	Err Est	S(.05)
Premedido vs otros	18.03	1.86	5.19
Mitad vs completo	6.90	2.14	5.97

Ambos contrastes son significativos ya que las estimaciones exceden $S(.05)$

11. $s^2 = 155.51 \text{ 20 g}$

DMS(.05) = 16.45, DHS(5, .05) = 23.60, SNK(2, .05) = 16.45,
SNK(3, .05) = 19.95, SNK(4, .05) = 22.08, , SNK(5, .05) = 23.60

Tratamiento	Medias	Pares significativos por cada criterio*		
		DMS(.05)	DHS(5, .05)	SNK(k, .05)
1. Premuda	86.37	2,3,4	2,3,4	2,3,4
2. Ayuno	112.85	1,3,5	1,3,5	1,3,5
3. 60 gr salvado	206.16	1,2,4,5	1,2,4,5	1,2,4,5
4. 80 gr salvado	110.69	1,3,5	1,3	1,3,5
5. Mezcla de malta	87.97	2,3,4	2,3	2,3,4

*Cada renglón de números identifica los tratamientos declarados significativos diferentes de los tratamientos mostrados en la primera columna de ese renglón.

- e. El DHS proporciona mayor fuerza de inferencia con las desigualdades de confianza que el DMS o SNK, éstos proporcionan pruebas de homogeneidad sólo en el sentido débil respecto a las tasas de error del experimento.

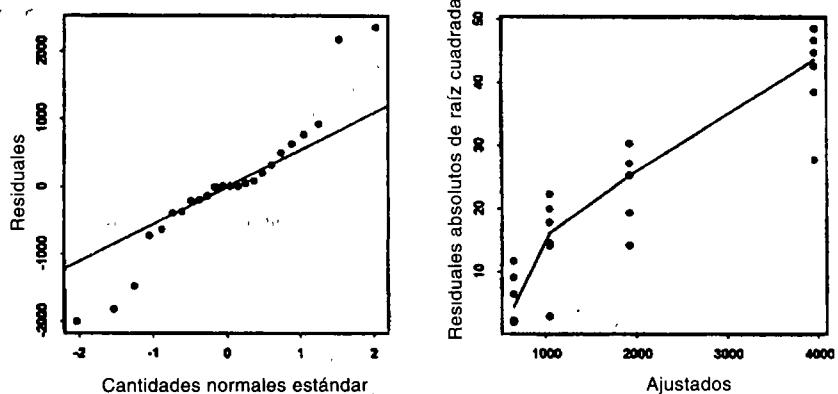
f.

Contraste	Estimación	ICS 95% de Tukey
Salv60–Salv80	95.5	(71.9, 119.0)
Salv60–Ayuno	93.3	(69.7, 117.0)
Salv60–Malta	118.0	(94.6, 142.0)
Salv60–Premuda	120.0	(96.2, 143.0)
Salv80–Ayuno	-2.2	(-25.8, 21.4)
Salv80–Malta	22.7	(-0.9, 46.3)
Salv80–Premuda	24.3	(0.7, 47.9)
Ayuno –Malta	24.9	(1.3, 48.5)
Ayuno –Premuda	26.5	(2.9, 50.1)
Malta –Premuda	1.6	(-22.0, 25.2)

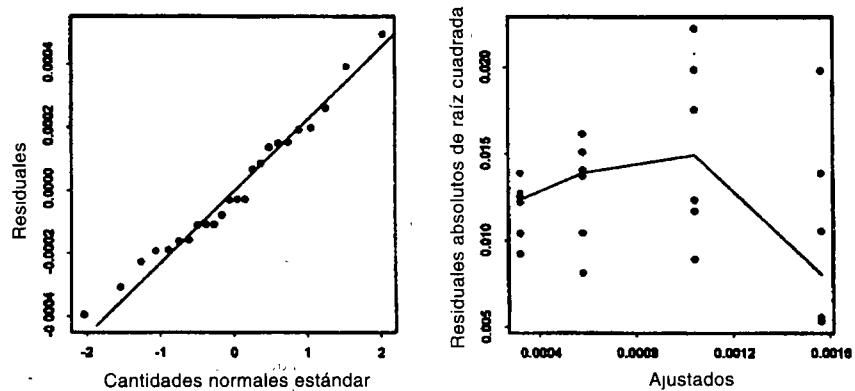
- g. Ancho de estimación del intervalo al igual que magnitud y dirección de separación entre dos medias. Es claro que la separación entre la etapa de Salv60 y las demás es mucho mayor que la separación entre cualquier otro par de etapas.

Capítulo 4

1. a.



- Prueba Levene(Med) para varianzas homogéneas, $F_0 = 23.42$, $F(.05, 3, 20) = 3.10$, se rechaza
 b. $\hat{\beta} = 1.73$, $\hat{p} = 1 - 1.73 = -0.73$, use la transformación recíproca de escala de potencias, $x = 1/y$
 c. Gráficas residuales después de la transformación recíproca



- Prueba de Levene(Med) para varianzas homogéneas, $F_0 = 1.56$, $F(.05, 3, 20) = 3.10$, no se rechaza
 d. Realice un análisis del valor codificado $x = 1000 \times (1/y)$

Fuente	gl	SC	CM	F_0
Lineal	1	4.73	4.73	87.02
Cuadrático	1	0.63	0.63	11.55
Cúbico	1	0.04	0.04	0.69
Error	20	1.09	0.05	

El polinomio cuadrático es significativo, $F(.05, 1, 20) = 4.35$

3. a. Binomial

b. sen^{-1} , donde $\sqrt{\pi} = (y + 3/8)(n + 3/4)$

c.

Fuente	gl	SC	CM	F_0
Clones	8	0.05	0.01	1.00
Error	9	0.05	0.01	

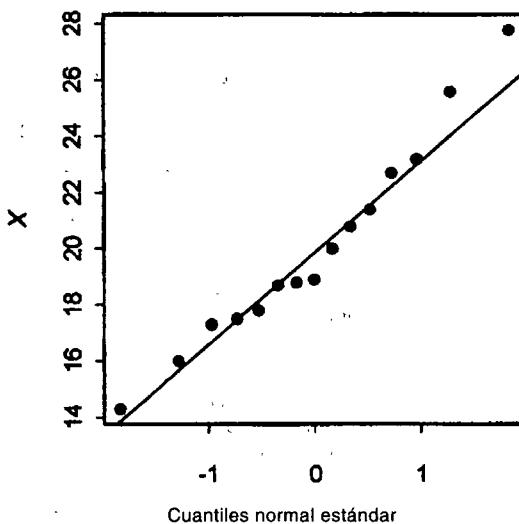
Medias: 0.47, 0.45, 0.40, 0.32, 0.49, 0.44, 0.46, 0.40, 0.48

d. $d(0.5, 8, 9) = 2.81 \quad M = 2.81 \sqrt{2(0.01)/2} = 0.281$

Todos los clones se eligen, lo que indica que no hay diferencias entre los clones más enraizados y los otros.

5. a. Los valores f son 0.0333, 0.1000, 0.1667, 0.233, 0.3000, 0.3667, 0.4333, 0.5000, 0.5667, 0.6333, 0.7000, 0.7667, 0.8333, 0.9000, 0.9667
 $g(f)$ son -1.83, -1.28, -0.97, -0.73, -0.52, -0.34, -0.17, 0.00, 0.17, 0.34, 0.52, 0.73, 0.97, 1.28, 1.83

b.



Capítulo 5

1. a. $y_{ij} = \mu + a_i + e_{ij} \quad i = 1, 2, \dots, 5 \quad j = 1, 2, \dots, 8$

a_i = efectos aleatorios de machos con media 0 y varianza σ_a^2

e_{ij} = efectos aleatorios de progenie con media 0 y varianza σ_e^2

Fuente	gl	SC	MC	E(CM)
Machos	4	5 591.1	1 397.8	$\sigma_e^2 + 8\sigma_a^2$
Progenie	35	16 232.7	463.8	σ_e^2

b. $\hat{\sigma}_e^2 = 463.8, \hat{\sigma}_a^2 = 116.7$

$\hat{\sigma}_e^2$ 90% IC:(326.722), $\hat{\sigma}_a^2$ IC 90%: (-4, 1 387)

c. $F_0 = 3.01, F(.05, 4, 35) = 2.65$, se rechaza

d. $\hat{\rho}_I = 0.20$, IC 90%: (.02, 0.67)

4. a. $y_{ij} = \mu + a_i + e_{ij} \quad i = 1, 2, \dots, 8 \quad j = 1, 2, 3, 4$

a_i = efectos aleatorios de lotes de semilla de algodón con media 0 y varianza σ_a^2

e_{ij} = efectos aleatorios de muestras de semilla dentro de los lotes con media 0 y varianza σ_e^2

Fuente	gl	SC	CM	E(CM)
Lotes	7	13 696.5	1 956.6	$\sigma_e^2 + 4\sigma_a^2$
Muestras	24	5 548.2	231.2	σ_e^2

- b. $\widehat{\sigma}_e^2 = 231.2$, $\widehat{\sigma}_d^2 = 431.4$
c. 662.6
d. 34.9% para muestras dentro de lotes y 65.1% para lotes
e. 25.7
f. Vea aplicaciones en las secciones 5.8 o 5.10. Cualquiera determina asignaciones óptimas de muestras con base en el costo o el número de muestras por lote para lograr una prueba de hipótesis fuerte.

6. a. $y_{ij} = \mu + \tau_i + e_{ij} + d_{ijk}$ $i = 1, 2, \dots, j = 1, 2, 3, k = 1, 2, 3$

τ_i = efecto fijo de fabricantes

e_{ij} = efectos aleatorios de filtros con media 0 y varianza σ_e^2

d_{ijk} = efectos aleatorios de pruebas de filtros con media 0 y varianza σ_d^2

Fuente	gl	SC	CM	E(CM)
Fabricante	1	1.8689	1.8689	$\sigma_d^2 + 3\sigma_e^2 + 9\theta_i^2$
Filtros	4	3.7413	0.9353	$\sigma_d^2 + 3\sigma_e^2$
Pruebas	12	0.3047	0.0254	σ_d^2

- b. $F_0 = 2.0$, $F(0.05, 1, 4) = 7.71$ no se rechaza

- c. Media(err est): 0.47(0.32), 1.12(0.32)

95% IC: (-0.42, 1.36), (0.23, 2.01)

- d. $c_1 = 200$, $c_2 = 1$, $\widehat{\sigma}_d^2 = 0.0254$, $\widehat{\sigma}_e^2 = 0.3033$

$n = 5$ pruebas, $r = 8$ filtros

8. a. $y_{ij} = \mu + a_i + e_{ij} + d_{ijk}$ $i = 1, 2, \dots, 15$ $j = 1, 2, \dots, r$, $k = 1, 2, \dots, n_j$

a_i = efectos aleatorios de campo con media 0 y varianza σ_a^2

e_{ij} = efectos aleatorios de sección con media 0 y varianza σ_e^2

d_{ijk} = efectos aleatorios de muestra de suelo con media 0 y varianza σ_d^2

Fuente	gl	SC	CM	E(CM)
Campos	14	14.43	1.03	$\sigma_d^2 + 1.19\sigma_e^2 + 2.38\sigma_a^2$
Secciones	15	11.53	0.77	$\sigma_d^2 + 1.2\sigma_e^2$
Muestras	6	8.80	1.47	σ_d^2

- b. $\widehat{\sigma}_d^2 = 1.47$, $\widehat{\sigma}_e^2 = -0.58$, $\widehat{\sigma}_a^2 = 0.11$

- c. $M = 0.78$, $v = 15$, $F_0 = 0.78/0.77 = 1.01$, $F(0.05, 14, 15) = 2.42$, no se rechaza

- d. $F = 0.77/1.47 = 0.52$, $F(0.05, 15, 6) = 3.94$, no se rechaza

Capítulo 6

1. a. $y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$ $i = 1, 2, 3$ $j = 1, 2$ $k = 1, 2, 3, 4$

α_i = efecto fijo de alcohol, β_j = efecto fijo de base

$(\alpha\beta)_{ij}$ = efecto de interacción entre alcohol y base

e_{ijk} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Alcohol	2	5.40	2.70
Base	1	6.51	6.51
Interacción	2	22.57	11.28
Error	18	36.76	2.04

- b. Medias marginales de alcohol. 89.98, 91.04, 90.10; error estándar = 0.50

Medias marginales base. 89.95, 90.89; error estándar = 0.41

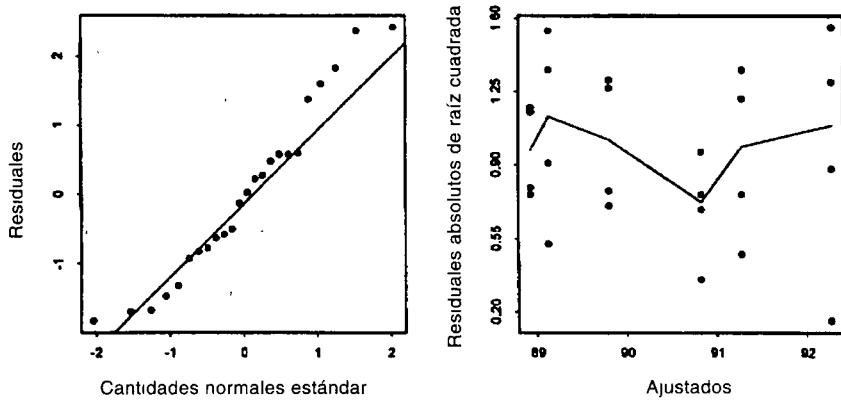
Base	Alcohol		
	1	2	13
1	90.83	89.80	88.93
2	89.13	92.28	91.28

Error estándar de medias de celdas = 0.71

- c. $F_0 = 5.53, F(.05, 2, 18) = 3.55$. Se rechaza
d. Contraste. (Base 1 - Base 2) para cada alcohol, $s_c = 1.01, t$ de Bonferroni (.025, 3, 18) = 2.64

Alcohol	c	t_0	
1	1.70	1.68	no se rechaza
2	-2.48	-2.46	no se rechaza
3	-2.35	-2.33	no se rechaza

- e. Prueba de Levene(Med), $F_0 = 0.51$, no se rechaza la hipótesis de varianzas de tratamiento iguales

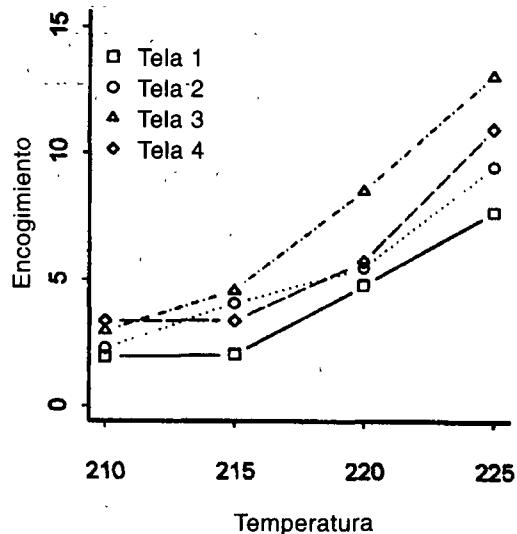


3. a. $y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad i = 1, 2, 3, 4 \quad j = 1, 2, 3, 4 \quad k = 1, 2$
 α_i = efecto fijo de tela, β_j = efecto fijo de temperatura
 $(\alpha\beta)_{ij}$ = efecto de interacción entre tela y temperatura
 e_{ijk} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Tela (F)	3	41.88	13.96
Temperatura (T)	3	283.94	94.65
T lineal	1	262.14	262.14
T cuadrática	1	21.78	21.78
Tela \times temperatura	9	15.86	1.76
F \times T lineal	3	11.31	3.77
F \times T cuadrática	3	2.69	0.90
Error	16	0.80	0.05

- b. $F_0 = 35.24, F(.05, 9, 16) = 2.54$, se rechaza
c. Temperatura lineal $F_0 = 5.242.9$, temperatura cuadrática, $F_0 = 435.6, F(0.05, 1, 16) = 4.49$, se rechazan ambas hipótesis nulas
d. Tela \times temperatura lineal $F_0 = 75.4$
Tela \times temperatura cuadrática $F_0 = 18.0$
 $F(0.05, 3, 16) = 3.24$, se rechaza para tela \times temperatura lineal y se rechaza para tela \times temperatura cuadrática

e.



5. a.

Fuente	gl	CM	F_0
Insecticida (I)	4	199.84	
Herbicida (H)	4	171.70	
I lineal	1	4.96	0.03
I cuadrática	1	553.01	3.18
I cúbica	1	162.90	0.94
H lineal	1	572.91	3.29
H cuadrática	1	108.75	0.63
H cúbica	1	4.65	0.03
I lin \times H lin	1	3 102.49	17.83
I lin \times H cuad	1	236.60	1.36
I lin \times H cúb	1	6.76	0.04
I cuad \times H lin	1	1 180.30	6.78
I cuad \times H cuad	1	195.00	1.12
I cuad \times H cúb	1	37.03	0.21
I cúb \times H lin	1	379.80	2.18
I cúb \times H cuad	1	23.41	0.13
I cúb \times H cúb	1	18.49	0.11
Error	75	174.00	

b. Se rechaza la hipótesis nula si $F_0 \geq F(.05, 1, 7.5) = 3.97$

c. $wgt = 105 - 1.17I - 40.7H + 0.00749I^2 + 1.72(I)(H) - 0.0415(I^2)(H)$

7. a.

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + e_{ijkl}$$

$$i, j, k, l = 1, 2, 3$$

α_i = efecto fijo de textura de suelo, β_j = efecto fijo de salinidad, γ_k = efecto fijo del agua, $(\alpha\beta)_{ij}$ = interacción entre textura y salinidad, $(\alpha\gamma)_{ik}$ = interacción entre textura y agua, $(\beta\gamma)_{jk}$ = interacción entre salinidad y agua, $(\alpha\beta\gamma)_{ijk}$ = interacción de tres factores, e_{ijkl} = error experimental con media 0 y varianza σ^2

Fuente	gl	SC	CM	F_0
Textura (T)	2	12.46	6.23	110.98
Salinidad (S)	2	58.13	29.07	517.77
Agua (W)	2	26.28	13.14	234.09
$T \times S$	4	1.28	0.32	5.72
$T \times W$	4	3.04	0.76	13.52
$S \times W$	4	20.62	5.15	91.82
$T \times S \times W$	8	1.65	0.21	3.68
Error	27	1.52	0.06	

b. Suelo \times salinidad \times medios de agua

Salinidad	2		8		10				
Agua	0	5	15	0	5	15	0	5	15
Arena arcillosa	0.54	1.85	3.39	0.09	1.10	0.13	0.07	0.11	0.20
Barro	0.96	2.35	5.40	0.21	0.29	0.58	0.14	0.29	0.39
Arcilla	1.44	3.08	5.56	0.62	0.95	1.64	0.49	0.80	2.41

Error estándar de celdas de medias = 0.17

Textura \times medias de salinidad

Textura	Salinidad			Agua		
	2	8	16	0	5	15
A. Arcillosa	1.93	0.10	0.12	0.23	0.69	1.24
Barro	2.90	0.36	0.27	0.43	0.98	2.12
Arcilla	3.36	1.07	1.23	0.85	1.61	3.20

Error estándar de medias de celdas = 0.10

Textura \times medias de agua

Textura	Agua		
	0	5	15
A. Arcillosa	1.93	0.10	0.12
Barro	2.90	0.36	0.27
Arcilla	3.36	1.07	1.23

Error estándar de medias de celdas = 0.10

Salinidad \times medias de agua

Salinidad	Agua		
	0	5	15
2	0.98	2.42	4.78
8	0.30	0.45	0.78
16	0.23	0.40	1.00

Error estándar de medias de celdas = 0.10

Medias de textura. 0.72, 1.18, 1.89, error estándar = 0.06

Medias de salinidad: 2.73, 0.51, 0.54, error estándar = 0.06

Medias de agua: 0.50, 1.09, 2.19, error estándar = 0.06

c. Se rechaza para textura, salinidad y agua, $F(.05, 2, 27) = 3.35$ Se rechaza para todas las interacciones de dos factores, $F(.05, 4, 27) = 2.73$ Se rechaza para la interacción de tres factores, $F(.05, 8, 27) = 2.31$

d.	Fuente	gl	SC	CM	F_0
	T	2	12.46	6.23	110.98*
	$S(\text{lin})$	1	38.55	38.55	686.70*
	$S(\text{cuad})$	1	19.58	19.58	348.84*
	$W(\text{lin})$	1	26.28	26.28	468.04*
	$W(\text{cuad})$	1	0.01	0.01	0.13
	$T \times S(1)$	2	0.96	0.48	8.55*
	$T \times S(q)$	2	0.32	0.16	2.89
	$T \times W(1)$	2	2.98	1.49	26.57*
	$T \times W(q)$	2	0.05	0.03	0.48
	$S(1) \times W(1)$	1	12.18	12.18	216.90*
	$S(1) \times W(q)$	1	0.13	0.13	2.29
	$S(q) \times W(1)$	1	8.30	8.30	147.79*
	$S(q) \times W(q)$	1	0.02	0.02	0.30
	$T \times S(1) \times W(1)$	2	1.27	0.63	11.27*
	$T \times S(1) \times W(q)$	2	0.19	0.09	1.67
	$T \times S(q) \times W(1)$	2	0.18	0.09	1.58
	$T \times S(q) \times W(q)$	2	0.02	0.01	0.18

*Se rechaza la hipótesis nula si $F_0 \geq F(05, 1, 27) = 4.21$

o $F_0 \geq F(.05, 2, 27) = 3.35$.

11. a. Las columnas **SC Yates** y **CM Yates** son los cálculos de las medias cuadráticas ponderadas, que se pueden obtener con la mayoría de los programas (MINITAB GLM con "Adj. SS", SAS GLM con "Type III SS", o Splus con la función "summary.aov"). Observe la columna **SC secuencial** (o **SS sequential**). Las sumas de cuadrados de interacción y error son idénticas a las de **Yates**. Las particiones de sumas de cuadrados de interacción y error se obtienen con un ajuste *secuencial* de términos al modelo en el orden presentado en la columna **Fuente** (o **Source**) y se pueden usar en forma indistinta para probar la interacción. Muchos programas también producen el ajuste secuencial (MINITAB GLM con "Seq SS", SAS GLM con "Type I SS" y Splus con la función "summary.aov").

Fuente	gl	SC secuencial	SC Yates	CM Yates
Base	1	8.897	14.242	14.242
Alcohol	2	4.105	7.675	3.838
Interacción	2	29.240	29.240	14.620
Error	10	19.873	19.873	1.987

$F_0 = CM(\text{interacción})/CME = 14.260/1.987 = 7.36$, se rechaza la hipótesis nula de no interacción ya que $F_0 > F(.05, 2, 27) = 3.35$. No es necesario probar los efectos principales.

- b. Media(error est)
 Base: 89.8(0.049), 91.8(0.61). Alcohol: 90.0(0.64), 92.0(0.81), 90.4(0.54)
 Interacción: $\hat{\mu}_{ij} = \text{media de celda observada}$ y $s_{\hat{\mu}_{ij}} = \sqrt{CME/r_{ij}}$
 91.1(1.00), 89.3(0.81), 88.9(0.70), 89.0(0.81), 94.7(1.41), 91.8(0.81)
 c. Nivel de alcohol, diferencia(error est), t_0
 Alcohol 1, 2.02(1.29), 1.57. Alcohol 2, -5.43(1.63), -3.33. Alcohol 3, -2.84(1.08), -2.64
 Usar t de Bonferroni para 3 comparaciones y 10 grados de libertad y rechazar H_0 si $|t_0| > 2.87$
 d. Las celdas de medias permanecen iguales. Las medias marginales observadas para base: 89.5, 91.0 y para alcohol: 89.8, 90.6, 91.8.

Capítulo 7

1. a. $y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$ $i = 1, 2, 3, 4$ $j = 1, 2, 3, 4, 5$ $k = 1, 2$
 a_i = efecto aleatorio de corrida, media 0, varianza σ_a^2

- b_j = efecto aleatorio de paciente, media 0, varianza σ_b^2
 $(ab)_{ij}$ = efecto aleatorio de corrida . paciente, media 0, varianza σ_{ab}^2
 e_{ijk} = error experimental, media 0, varianza σ^2

Fuente	gl	CM	E(CM)
Corrida (R)	3	364.46	$\sigma^2 + 2\sigma_{ab}^2 + 10\sigma_a^2$
Pacientes (P)	4	13.152.03	$\sigma^2 + 2\sigma_{ab}^2 + 8\sigma_b^2$
$R \times P$	12	9.12	$\sigma^2 + 2\sigma_{ab}^2$
Error	20	6.61	σ^2

- c. 35.53; 1 642.86; 1.26
d. Interacción: $H_0: \sigma_{ab}^2 = 0$, $H_a: \sigma_{ab}^2 \neq 0$, $F_0 = 1.38$, $F(.05, 12, 20) = 2.28$ no se rechaza
Pacientes: $H_0: \sigma_b^2 = 0$, $H_a: \sigma_b^2 \neq 0$, $F_0 = 1442.11$, $F(.05, 4, 12) = 3.26$, se rechaza
Corridas: $H_0: \sigma_a^2 = 0$, $H_a: \sigma_a^2 \neq 0$, $F_0 = 39.96$, $F(.05, 3, 12) = 3.49$, se rechaza

3. a. $y_{ijk} = \mu + \alpha_i + b_{j(i)} + C_{k(j)}$ $i = 1, 2, 3$ $j = 1, 2, 3, 4$ $K = 1, 2$
 α_i = efecto fijo de aleación
 $b_{j(i)}$ = efecto aleatorio de moldes dentro de aleación, media 0, varianza $\sigma_{b(a)}^2$
 $C_{k(j)}$ = efecto aleatorio de barras dentro de moldes, media 0, varianza $\sigma_{c(b)}^2$

Fuente	gl	CM	E(CM)
Aleación	2	14.69	$\sigma_{c(b)}^2 + 2\sigma_{b(a)}^2 + 8\theta_a^2$
Molde(aleación)	9	0.52	$\sigma_{c(b)}^2 + 2\sigma_{b(a)}^2$
Barra(molde)	12	0.36	$\sigma_{c(b)}^2$

- c. $F_0 = 28.25$, $F(.05, 2, 9) = 4.26$, se rechaza
d. Medias(err est): 14.7(0.25), 16.7(0.25), 14.1(0.25)
IC 95%: (14.1, 15.3), (16.1, 17.3), (13.5, 14.7)
e. Barras: 0.36, moldes: 0.08
5. a. $y_{ijk} = \mu + \alpha_i + b_j + (ab)_{ij} + c_{k(j)} + e_{ijkl}$
 α_i = efecto fijo de máquina
 b_j = efecto aleatorio de parcela, media 0, varianza σ_b^2
 $(ab)_{ij}$ = efecto aleatorio de máquina por interacción de parcela, media 0, varianza σ_{ab}^2
 $c_{k(j)}$ = efecto aleatorio de muestra dentro de parcela, media 0, varianza $\sigma_{c(b)}^2$
 $(ac)_{ik(j)}$ = efecto aleatorio de máquina por interacción de muestra añadida dentro de la parcela, media 0, varianza $\sigma_{ac(b)}^2$
 e_{ijkl} = error aleatorio, media 0, varianza σ^2

Fuente	gl	CM*	E(CM)
Máquina (M)	1	25.29	$\sigma^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{ab}^2 + 12\sigma_a^2$
Parcela (P)	2	24.82	$\sigma^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{c(h)}^2 + 4\sigma_{ab}^2 + 8\sigma_b^2$
$M \times P$	2	0.10	$\sigma^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{ab}^2$
Muestra (S) dentro de (P)	3	8.86	$\sigma^2 + 2\sigma_{ac(b)}^2 + 4\sigma_{c(h)}^2$
$M \times S$ dentro de P	3	0.54	$\sigma^2 + 2\sigma_{ac(b)}^2$
Error	12	0.117	σ^2

*Valor de media cuadrática $\times 10^3$

- c. $F_0 = MCM/CM(MP) = 252.9$, $F(.05, 1, 2) = 18.51$, se rechaza

Capítulo 8

1. a. $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$ $i = 1, 2, \dots, 6$ $j = 1, 2, \dots, 8$
 τ_i = efecto fijo de método de irrigación, ρ_j = efecto fijo de bloque
 e_{ij} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Bloques	7	457 507	63 358
Irrigación	5	47 842	9 568
Error	35	16 888	4 197

- b. Los errores experimentales son aleatorios e independientes con media 0 y varianza σ^2 . Los efectos de tratamiento y de bloque son aditivos. La aleatorización justifica la suposición de errores aleatorios independientes. La relación entre los efectos de tratamiento es consistente entre los bloques.
c. 22.9, 32.4
d. $d(.05, 5, 35) = 2.64$ $D(5, .05) = 85.5$

Irrigación	Media	$\bar{y}_i - \bar{y}_c$	IC 95%
			(L, U)
Anegado	229.6	control	
Goteo	299.6	70.0	(-15.5, 155.5)
En canal	290.4	60.8	(-24.7, 146.3)
Rocío	223.8	-5.8	(-91.3, 79.7)
Aspersor	292.0	62.4	(-23.1, 147.9)
Aspersor + rocío	291.0	61.4	(-24.1, 146.9)

Ningún tratamiento es significativamente diferente de anegado, ya que todos los intervalos incluyen 0.

- e. $ER = 0.99(3.17) = 3.14$. Más de tres veces el número de réplicas requerido por el diseño totalmente aleatorizado para tener la misma varianza de la media de tratamiento.

2. a. $y_{ijk} = \mu + \tau_i + \rho_j + e_{ijk}$ $i = 1, 2, \dots, 5$ $j = 1, 2, \dots, 5$
 τ_i = efecto fijo de fertilizante, ρ_j = efecto fijo de bloque
 e_{ijk} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Bloques	4	4.95	1.24
Fertilizante	4	72.12	18.03
Error	16	12.76	0.80

b.

Contraste	SC	F_0
Sin fert vs otros $(1, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4})$	3.13	3.93
Nitrógeno $(0, -1, +1, -1, +1)$	4.32	5.42*
P ₂ O ₅ $(0, -1, -1, +1, +1)$	64.44	80.79*
Interacción $(0, +1, -1, -1, +1)$	0.22	0.28

*Se rechaza la hipótesis nula si $F_0 \geq F(.05, 1, 16) = 4.49$

- c. 0.45, 0.80, 0.80, 0.80
e. $ER = 0.98(1.09) = 1.07$

5. a. $y_{ij} = \mu + \rho_i + \gamma_j + \tau_K + e_{ij}$ $i, j, k = 1, 2, 3, 4$
 ρ_i = efecto fijo de periodo (renglón), γ_j = efecto fijo de técnico (columna)
 τ_K = efecto fijo de método de construcción
 e_{ij} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Tiempo	.3	467.19	155.73
Técnico	3	17.19	5.73
Método	3	145.69	48.56
Error	6	22.88	3.81

- b. 0.98, 1.38
 c. $q(.05, 4, 6) = 4.90$, $HSD(4, .05) = 4.80$, 95% ICS son $\bar{y}_i - \bar{y}_j \pm 4.80$
 Los ICS 95% para diferencias que no incluyen 0 son (A, D), (A, C)
 d. $ER = .93(11) = 10.2$

7. a. $y_{ijk} = \mu + \tau_i + \rho_j + (\tau\rho)_{ij} + e_{ijk}$ $i = 1, 2, \dots, 5$ $j, k = 1, 2$

τ_i = efecto fijo de tasa de nitrógeno, ρ_j = efecto fijo de bloque

$(\tau\rho)_{ij}$ = efecto fijo de interacción tasa \times bloque

e_{ijk} = error experimental, media 0, varianza σ^2

El efecto de interacción se puede incluir en el modelo y su partición de suma de cuadrados, calculada en virtud de que se incluyeron dos parcelas de cada tratamiento en cada bloque. Así, la varianza del error experimental se puede estimar por separado como la varianza entre parcelas tratadas en forma parecida dentro de cada bloque.

Fuente	gl	SC	CM
Bloques	1	1 022.45	1 022.45
Tasas	4	4 813.00	1 203.25
Tasa lineal	1	2 856.10	2 856.10
Tasa cuadrática	1	1 716.07	1 716.07
Tasa cúbica	1	60.03	60.03
Bloque \times tasa	4	287.80	71.95
Error	10	422.50	42.25

- b. $F_0 = 1.70$, $F(.05, 4, 10) = 3.48$, no se rechaza.
 c. Vea partición SC en la tabla de análisis de varianza.
 Lineal: $F_0 = 67.6$, Cuadrática: $F_0 = 40.62$, $F(.05, 1, 10) = 4.96$
 Se rechaza la hipótesis nula para efectos lineales y cuadráticos
 d. $F_0 = 1.42$, no se rechaza la hipótesis nula; cúbica no significativa

9. a.

Fuente	gl	SC	CM
Total	45	614 708	
Bloques (no aj.)	7	432 384	61 769
Irrigación (aj.)	5	51 923	10 385
Error	33	130 402	3 952

- b. $F_0 = 2.63$, $F(.05, 5, 33) = 2.50$, se rechaza
 c. Media(err est): 290.1(24.1), 290.4(22.2), 223.8(22.2), 292.0(22.2), 291.0(22.2), 213.7(24.1)
 d. Mayor que los tratamientos sin parcelas perdidas
 e. (i) 34.1, (ii) 32.8, (iii) 32.8, (iv) 31.4
 f. Errores estándar aumentados si alguno o ambos tratamientos tienen parcelas perdidas; potencia disminuida

10. a.

Fuente	gl	Cuadrática de medios*			
		Año 1		Año 2	
		Primavera	Invierno	Primavera	Invierno
Tratamiento	3	32.55	41.05	193.15	48.65
Error	4	12.99	18.11	12.19	45.31

*Valor cuadrática de medios $\times 10^3$

b. $F_{máx} = 45.31/12.19 = 3.72$, $F_{0.05} \text{ máx} = 20.6$, $t = 4$, $v = 4$

No se rechaza la hipótesis de varianzas iguales

c.

Fuente	gl	CM	E(CM)
Año (Y)	1	13.61	$\sigma^2 + 2\sigma_{ys}^2 + 4\sigma_n^2 + 8\sigma_{ys}^2 + 16\sigma_y^2$
Estación (S)	1	1911.01	$\sigma^2 + 2\sigma_{ys}^2 + 8\sigma_{ys}^2 + 16\theta_s^2$
Tratamiento (T)	3	257.70	$\sigma^2 + 2\sigma_{ys}^2 + 4\sigma_n^2 + 8\theta_t^2$
$Y \times S$	1	94.61	$\sigma^2 + 2\sigma_{ys}^2 + 8\sigma_{ys}^2$
$T \times S$	3	15.94	$\sigma^2 + 2\sigma_{ys}^2 + 4\theta_t^2$
$T \times Y$	3	28.47	$\sigma^2 + 2\sigma_{ys}^2 + 4\sigma_n^2$
$T \times Y \times S$	3	13.27	$\sigma^2 + 2\sigma_{ys}^2$
Error	16	22.15	σ^2

*Valor cuadrático de medios $\times 10^3$

d. Estación: $F_0 = 1911.01/94.61 = 20.2$, $F(0.05, 1, 1) = 161.4$, no se rechaza

Tratamientos: $F_0 = 257.7/28.47 = 9.05$, $F(0.05, 3, 3) = 9.28$, no se rechaza

$Y \times S$: $F_0 = 94.61/13.27 = 7.13$, $F(0.05, 1, 3) = 10.13$, no se rechaza

$T \times S$: $F_0 = 15.94/13.27 = 1.20$, $F(0.05, 3, 3) = 9.28$, no se rechaza

$T \times Y$: $F_0 = 28.47/13.27 = 2.15$, $F(0.05, 3, 3) = 9.28$, no se rechaza

$T \times Y \times S$: $F_0 = 13.27/22.15 = 0.60$, $F(0.05, 3, 16) = 3.24$, no se rechaza

Capítulo 9

1. a. $\lambda = 1$

b. $E = \frac{2}{3}$

c. $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$ $i = 1, 2, 3, 4$ $j = 1, 2, 3, 4, 5, 6$

τ_i = efecto fijo de temperatura, ρ_j = efecto fijo de bloque

e_{ij} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Bloque(no aj.)	5	613.66	122.73
Temperatura(aj.)	3	718.29	239.43
Temp. lineal	1	677.08	677.08
Temp. cuadrática	1	7.33	7.33
Temp. cúbica	1	33.87	33.87
Error	3	41.00	13.67

Prueba de temperatura. $F_0 = 17.51$, $F(0.05, 3, 3) = 9.28$, se rechaza

d. Medias: 26.46, 23.82, 10.07, 3.61; error estándar = 2.50

e. 3.70

f. Vea la tabla de análisis de varianza para particiones de SC

Línea: $F_0 = 49.53$, $F(0.05, 1, 3) = 10.13$, se rechaza

Cuadrática: $F_0 = 0.54$, $F(0.05, 1, 3) = 10.13$, no se rechaza

g. No. F_0 para cúbica = 2.48, $F(0.05, 1, 3) = 10.13$, no se rechaza

3. a. $y_{im} = \mu + \tau_i + \rho_j + \gamma_m + e_{im}$ $i, m = 1, 2, \dots, 7$ $j = 1, 2, 3, 4$

τ_i = efecto fijo de tipo de intersección, ρ_j = efecto fijo de observador

γ_m = efecto fijo de ciudad

e_{im} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Observador	3	146.51	48.84
Ciudad(no aj.)	6	1 342.07	233.68
Intersección(aj.)	6	9 972.46	1 662.08
Error	12	368.85	30.74

- b. 4.19
 c. plan 9B.5, $E = 0.88$
 d. $d(.05, 6, 12) = 2.58, M = 2.58(4.19) = 10.8$

Interacción	\bar{y}_i	máx \bar{y}_i $j \neq i$	D_i	$D_i - M$	$D_i + M$	ICS 95%
						(L, U)
1	68.4	28.6	39.88	29.0	50.6	(0, 50.6)
2	28.6	28.7	-0.1	-10.9	10.77	(-10.9, 10.7)
3	28.7	28.6	0.1	-10.7	10.9	(-10.7, 10.9)
4	40.4	28.6	11.8	1.0	22.6	(0, 22.6)
5	44.8	28.6	16.2	5.4	27.0	(0, 27.0)
6	85.2	28.6	56.6	45.8	67.4	(0, 67.4)
7	32.6	28.6	4.0	-6.8	14.8	(-6.8, 14.8)

Selección de tipos de información 2, 3 y 7 con $P(SC) = 0.95$.

5. a. $t = 5, r = 4, k = 4, b = 5$
 b. $\lambda = 3$

Capítulo 10

1. a. $y_{ijk} = \mu + \tau_i + \gamma_j + \rho_{m(j)} + e_{ijm}$ $i = 1, 2, \dots, 8$ $j = 1, 2, \dots, 7$ $m = 1, 2$
 τ_i = efecto de variedad, γ_j = efecto de réplica, $\rho_{m(j)}$ = efecto de bloque dentro de la réplica
 e_{ijm} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Réplica	6	67.35	11.22
Bloque(no aj.)	7	336.90	48.13
Variedad(aj.)	7	1 245.99	178.00
Error	35	358.19	10.23

Prueba de hipótesis nula para variedad: $F_0 = 17.4, F(.05, 7, 35) = 2.29$, se rechaza
 Medias variedad ($\bar{\mu}$): 43.3, 54.4, 38.4, 51.8, 47.5, 49.4, 42.5, 51.0
 Error estándar = 1.29

- b. 1.85
 c. 0.86

3. a. $y_{ijlm} = \mu + \tau_i + \beta_m + \rho_{i(m)} + \gamma_{l(m)} + e_{ijlm}$
 τ_i = efecto de variedad, β_m = efecto de réplica, $\rho_{i(m)}$ = efecto de renglón dentro de la réplica
 $\gamma_{l(m)}$ = efecto de columna dentro de la réplica
 e_{ijlm} = error experimental, media 0, varianza σ^2

Fuente	gl	SC	CM
Réplica	3	5.12	1.71
Columna(no aj.)	8	29.30	3.66
Renglón(no aj.)	8	37.50	4.69
Variedad(aj.)	8	130.87	16.36
Error	8	6.04	0.75

Prueba de hipótesis nula para variedad: $F_0 = 21.81$, $F(.05, 8, 8) = 3.44$, se rechaza

Resultados abajo de los obtenido con GLM en SAS:

Medias de variedad (\bar{Y}_i): 59.1, 53.4, 55.1, 53.0, 55.2, 54.2, 54.4, 48.7, 56.9

Error estándar = 0.60

- b. 0.87 de GLM en SAS
- c. 0.50

Capítulo 11

1. Bloque I, unidades 1-8: (1), ab, ac, ad, bc, bd, cd, abcd
Bloque II, unidades 9-16: a, b, c, d, abc, abd, acd, bcd,
3. Confundir ABC en rép I, ABD en rép II, ACD en rép III y BCD en rép IV.
Rép I, Bloque I: (1), d, ab, ac, bc, abd, acd, bcd
Rép I, Bloque II: a, b, c, ad, bd, cd, abc, abcd
Rép II, Bloque I: (1), c, ab, ad, bd, abc, acd, bcd
Rép II, Bloque II: a, b, d, ac, bc, cd, abd, abcd
Rép III, Bloque I: (1), b, ac, ad, cd, abc, abd, bcd
Rép III, Bloque II: a, c, d, ab, bc, bd, acd, abcd
Rép IV, Bloque I: (1), a, bc, bd, cd, abc, abd, acd
Rép IV, Bloque II: b, c, d, ab, ac, ad, bcd, abcd
5. a. A, AC, ABC, CDE,
b. Sí, porque otra elección para definir contrastes no confunde efectos principales o de interacciones de dos factores.
c. Use el plan en la tabla 11A.1 con ACE, ABF y ABCD como contrastes definidos. No se confunden los efectos principales ni de interacción de dos factores.

7. a.

Fuente	gl	SC	CM	F ₀
Réplica	1	11.28	11.28	
Bloques	2	56.31	28.16	
A (velocidad)	1	639.03	639.03	80.28*
B (ángulo)	1	11.28	11.28	1.42
C (lubricación)	1	52.53	52.53	6.60*
D (aleación)	1	124.03	124.03	15.58*
AB	1	132.03	132.03	16.59*
AC	1	3.78	3.78	0.48
AD	1	34.03	34.03	4.28
BC	1	2.53	2.53	0.32
BD	1	0.03	0.03	0.00
CD	1	26.28	26.28	3.30
ABC	1	205.03	205.03	25.76*
ABD	1	175.78	175.78	22.08*
ACD	1	7.03	7.03	0.88
BCD	1	7.03	7.03	0.88
Error	14	111.44	7.96	

*Se rechaza la hipótesis nula si $F_0 \geq F(.05, 1, 14) = 4.60$

- b. Efecto(estimado):
 $A(-8.94), B(-1.19), C(-2.56), D(3.94), AB(4.06), AC(0.69), AD(-2.06),$
 $BC(-0.56), BD(-0.06), CD(1.81), ABC(-5.06), ABD(4.69), ACD(-0.94),$
 $BCD(-0.94)$, error estándar = 1.00
- c. Prepare un diseño de medias de celdas velocidad \times ángulo para cada nivel de lubricación y cada nivel de aleación para interpretar la interacción velocidad \times ángulo como cambio de niveles de lubricación o de aleación. Aunque los efectos principales de velocidad, lubricación y aleación son significativos, lo mismo que la interacción velocidad \times ángulo, se examinará la modificación de sus efectos mediante las interacciones significativas de tres factores
9. a. las interacciones generalizadas son AC y B^2
 b. las interacciones generalizadas son AB^2C y BC
 Elija (b) para evitar confundir un efecto principal.

Capítulo 12

1. 1. 1/12, 1, 0, 1, 16
 3. 1/4, 2, 1, 3, 16
 5. 1/4, 2, 1, 3, 32
 7. 1/16, 4, 11, 15, 8
 9. 1/32, 5, 26, 31, 16
 11. 1/64, 6, 57, 63, 32
3. a. $ABEF$
 b. IV
 c. $A = BCD, ACDEF, BEF; B = ACD, BCDEF, AEF; C = ABD, DEF, ABCEF$
 $D = ABC, CEF, ABDEF; E = ABF, CDF, ABCDE; F = ABE, CDE, ABCDF$
 $AB = CD, EF, ABCDEF; AC = BD, ADEF, BCEF; AD = BC, ACEF, BDEF$
 $AE = BF, ACDF, BCDE; AF = BE, ACDE, BCDF; BC = AD, BDEF, ACEF$
 $BD = AC, BCEF, ADEF; BE = AF, ACED, BCDF; BF = AE, ACDF, BCDE$
 $CD = AB, EF, ABCDEF; CE = DF, ABCF, ABED; CF = DE, ABDF, ABCE$
 $DE = CF, ABCE, ABDF, DF = CE, ABDE, ABCF; EF = AB, CD, ABCDEF$
5. Sí. Suponga que el diseño 2^{5-2}_{III} con factores A, B, C, D y E con 8 unidades con niveles de D y E generadas según $D = AB$, y $E = AC$. Se hace un doblez invertido en el diseño agregado un segundo diseño 2^{5-2}_{III} con todos los signos del primero invertidos, incluso la columna I de manera que $I = F$, un sexto factor. Entonces $D = ABF$ y $E = ACF$. Diseñe generadores con cuatro letras para obtener un diseño de resolución IV con 16 unidades para un diseño 2^{6-2}_{IV} .
7. a. D, E y BE ya que dejan de estar confundidos con otros efectos principales o interacciones de dos factores.
 b. Use $I = ABC = CDE$. Ya no están confundidos todos los efectos principales con interacciones de dos factores significativas u otros efectos principales. Sin embargo, $AD = BE$ y $AE = BD$, y no hay interacciones de dos factores potenciales agrupadas como pares.
9. 2^{4-1}_{IV}
11. a. V
 b. $E = -ABCD$
 c. Vea la tabla 12.3 y cambie los signos de los alias
 d. No hay interacciones de tres factores o mayores
 e. Efecto(estimado):
 $A(0.145), B(0.0875), C(0.0375), D(-0.0375), E(-0.47), AB(0.015), AC(0.095),$

$AD(0.03)$, $AE(-0.1525)$, $BC(-0.0675)$, $BD(0.1625)$, $BE(-0.405)$, $CD(0.0725)$,
 $CE(0.135)$, $DE(-0.315)$

Elija E , BE y DE como efectos significativos de la gráfica de calificaciones normales $CME = 0.03978$ del análisis de varianza. Error estándar del efecto = 0.0997

Fuente	gl	SC	CM
E	1	0.88360	0.88360
BE	1	0.65610	0.65610
DE	1	0.39690	0.39690
Error	12	0.47738	0.03978

La gráfica de probabilidad normal para evaluar este modelo revela que los residuales no son usuales con la distribución normal supuesta

Capítulo 13

1. a. $s^2 = 6.3$, 5 g.l.
b. $\hat{\beta}_0 = 50.25$, $\hat{\beta}_1 = 7.5$, $\hat{\beta}_2 = 10.75$, error estándar = 0.89
 $H_0: \beta_1 = 0$, $t_0 = 8.42$; $H_0: \beta_2 = 0$, $t_0 = 12.08$; $t(.025, 5) = 2.571$, se rechazan ambos; efectos lineales significativos
c. $\hat{\beta}_{12} = 1.5$ error estándar = .089, $t_0 = 1.69$, $t(.025, 5) = 2.571$, no se rechaza, interacciones no significativas
d. $\bar{y}_f - \bar{y}_c = -0.25$, error estándar = 1.98 $t_0 = -0.13$, $t(.025, 5) = 2.571$, no se rechaza, no hay alejamiento significativo de la superficie lineal
e. $(A, B), (25, 35), (40, 56.5), (55, 78), (70, 99.5), (85, 121), (100, 142.5)$

3.

Diseño 2 ^a				Axial			
x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
- 1	- 1	- 1	- 1	- 2	0	0	0
+ 1	- 1	- 1	- 1	+ 2	0	0	0
- 1	+ 1	- 1	- 1	0	- 2	0	0
+ 1	+ 1	- 1	- 1	0	+ 2	0	0
- 1	- 1	+ 1	- 1	0	0	- 2	0
+ 1	- 1	+ 1	- 1	0	0	+ 2	0
- 1	+ 1	+ 1	- 1	0	0	0	- 2
+ 1	+ 1	+ 1	- 1	0	0	0	+ 2
- 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				

Centro*			
x_1	x_2	x_3	x_4
0	0	0	0

*7 réplicas

5. a. $\hat{y} = 431 + 4.04x_1 + 64.75x_2 - 8.80x_1^2 - 44.55x_2^2 + 1.75x_1x_2$

Fuente	gl	SC	CM
Regresión	5	47 552	9 510
Lineal	2	33 669	16 835
Cuadrática	3	13 883	4 628
Error	7	2 795	399
Falta de ajuste	3	1 964	655
Error puro	4	831	208

- b. Modelo completo: $F_0 = 45.7$, $F(.05, 5, 4) = 6.266$, se rechaza
 Componentes cuadráticas: $F_0 = 22.3$, $F(.05, 3, 4) = 6.59$, se rechaza
 Componentes lineales. $F_0 = 80.9$, $F(.05, 2, 4) = 6.94$, se rechaza
 Falta de ajuste: $F_0 = 3.2$, $F(.05, 3, 4) = 6.59$, no se rechaza
 El modelo cuadrático es significativo sin faltas significativas de ajuste
- c. Metionine: $x_1 = 0.302$, $M = 0.986$
 Carotene: $x_2 = 0.733$, $C = 75.91$
- d. $\hat{y} = 455 - 44.572x_1^2 - 8.742x_2^2$
 $Z_1 = -0.725 - 0.024x_1 + 1.000x_2$
 $Z_2 = -0.320 - 1.000x_1 + 0.024x_2$

Del diseño especificado en la tabla 13.4 para diseños de bloque casi rotatorios y ortogonales

Bloque I			Bloque II		
-1	-1	-1	1.682	0	0
+1	-1	-1	-1.682	0	0
-1	+1	-1	0	1.682	0
+1	+1	-1	0	-1.682	0
-1	-1	+1	0	0	1.682
+1	-1	+1	0	0	-1.682
-1	+1	+1	0	0	0
+1	+1	+1	0	0	0
0	0	0			
0	0	0			

$$\alpha = 1.789 \text{ para la ecuación (13.7)}$$

10. a.	Sulfuro	Asfalto	Arena
	0.3	0.2	0.5
	0.1	0.4	0.5
	0.1	0.2	0.7
	0.2	0.3	0.5
	0.2	0.2	0.6
	0.1	0.3	0.6

- b. 2.02, 6 df
 c. Ecuación de seudocomponentes:
 $\hat{y} = 12.85\tilde{x}_1 + 3.00\tilde{x}_2 + 3.45\tilde{x}_3 + 39.70\tilde{x}_1\tilde{x}_2 + 40.40\tilde{x}_1\tilde{x}_3 + 10.90\tilde{x}_2\tilde{x}_3$
 Ecuación de componentes reales:
 $\hat{y} = -559.70x_1 - 140.95x_2 - 58.70x_3 + 992.50x_1x_2 + 1010.00x_1x_3 + 272.50x_2x_3$

Fuente	gl	SC	CM
Regresión	5	491.62	98.32
Lineal	2	262.42	131.21
Cuadrática	3	229.20	76.40
Error	6	12.10	2.02

Prueba para lineal: $F_0 = 64.96$, $F(.05, 2, 6) = 5.14$, se rechaza

Prueba para cuadrática. $F_0 = 37.82$, $F(.05, 3, 6) = 4.76$, se rechaza

Capítulo 14

1. a. $y_{ijk} = \mu + \alpha_i + \rho_k + d_{ik} + \beta_j + (\alpha\beta)ij + e_{ijk}$ $i, k = 1, 2, 3, 4$ $j = 1, 2, 3$
 α_i = efecto fijo de densidad de población de plantas, β_j = efecto fijo híbrido, $(\alpha\beta)_{ij}$ = efecto de interacción fijo población \times híbrido, ρ_k = efecto fijo de bloque, d_{ik} = efecto aleatorio de toda la parcela con media 0 y σ_d^2 , e_{ijk} = error aleatorio de subparcela con media 0 y varianza σ_e^2

Fuente	gl	SC	CM
Bloques	3	408.98	136.33
Población (P)	3	6 429.39	2 143.13
P lineal	1	5 658.23	5 658.23
P cuadrática	1	767.64	767.64
P cúbica	1	3.51	3.51
Error(1)	9	466.54	51.84
Híbrido (H)	2	881.41	440.70
$H \times P$	6	207.51	34.58
$H \times P$ lineal	2	72.74	36.37
$H \times P$ cuadrática	2	92.30	46.15
$H \times P$ cúbica	2	42.47	21.24
Error(2)	24	595.70	24.82

b.

Peso medio de semillas (g)

Híbrido	Plantas por metro de fila				Medias de híbridos
	10	15	25	40	
TAM 680	38.6	32.6	18.4	13.5	25.8
RS 671	45.2	35.5	24.0	17.1	30.4
Tx 399	57.3	40.4	25.9	21.6	36.3
\times Tx 2536					
Medias de población	47.0	36.1	22.8	17.4	

Factor(media de error estándar): Población(2.1), Híbrido(1.2), Celdas (2.5)

- c. (i) 1.8, (ii) 2.9, (iii) 4.11
d. Interacción: $F_0 = 1.39$, $F(.05, 6, 24) = 2.51$, no se rechaza
Población: $F_0 = 41.34$, $F(.05, 3, 9) = 3.86$, se rechaza
Híbrido: $F_0 = 17.76$, $F(.05, 2, 24) = 3.40$, se rechaza
e. Subparcela: $ER = 0.98(1.3) = 1.27$
Parcela completa: $ER = 0.88(0.62) = 0.55$
f. Vea en la tabla de análisis de varianza las particiones polinomiales de sumas de cuadrados Ninguna partición de interacción es significativa con $F(.05, 2, 24) = 3.40$ como valor crítico.
Población lineal: $F_0 = 109.15$, población cuadrática; $F_0 = 14.81$, Población cúbica: $F_0 = 0.07$, $F(.05, 1, 9) = 5.12$, se rechaza para lineal y cuadrática

3. a. $y_{ijklm} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \rho_k + \gamma_l + d_{kl} + \theta_m + (\alpha\theta)_{im} + (\beta\theta)_{jm}$
 $(\alpha\beta\theta)_{ilm} + e_{klm}$ $i, j = 1, 2$ $k, l = 1, 2, 3, 4$ $m = 1, 2, 3$
 α_i = efecto fijo A, β_j = efecto fijo B, $(\alpha\beta)_{ij}$ = efecto A \times B de interacción fijo, ρ_k = efecto de renglón fijo, γ_l = efecto de columna fijo, d_{kl} = error aleatorio de parcela completa con media 0 y varianza σ_d^2 , θ_m = efecto fijo C, $(\alpha\theta)_{im}$ = efecto de interacción fijo A \times C, $(\beta\theta)_{jm}$ = efecto de interacción fijo B \times C, $(\alpha\beta\theta)_{ilm}$ = efecto de interacción fijo A \times B \times C, e_{klm} = error aleatorio de subparcela con media 0 y varianza σ_e^2

b.

Fuente	gl	Cuadrática de medios esperada
Renglones	3	
Columnas	3	
A	1	$\sigma_e^2 + 3\sigma_d^2 + 24\theta_a^2$
B	1	$\sigma_e^2 + 3\sigma_d^2 + 24\theta_b^2$
AB	1	$\sigma_e^2 + 3\sigma_d^2 + 12\theta_{ab}^2$
Error(1)	6	$\sigma_e^2 + 3\sigma_d^2$
C	2	$\sigma_e^2 + 16\theta_c^2$
AC	2	$\sigma_e^2 + 8\theta_{ac}^2$
BC	2	$\sigma_e^2 + 8\theta_{bc}^2$
ABC	2	$\sigma_e^2 + 4\theta_{abc}^2$
Error(2)	24	σ_e^2

7. a. $y_{ijkl} = \mu + \alpha_i + \rho_j + d_{ij} + \beta_k + \gamma_l + (\beta\gamma)_{kl} + (\alpha\beta)_{ik} + (\alpha\gamma)_{il} + (\alpha\beta\gamma)_{ikl} + e_{ijkl}$ $i, j = 1, 2$ $k, l = 1, 2, 3$
 α_i = efecto fijo de fósforo, ρ_j = efecto fijo de bloque, d_{ij} = error aleatorio de parcela completa con media 0 y varianza σ_d^2 , β_k = efecto fijo de nitrógeno, γ_l = efecto fijo de agua, $(\beta\gamma)_{kl}$ = efecto de interacción fijo nitrógeno \times agua, $(\alpha\beta)_{ik}$ = efecto de interacción fijo fósforo \times nitrógeno, $(\alpha\gamma)_{il}$ = efecto de interacción fijo fósforo \times agua, $(\alpha\beta\gamma)_{ikl}$ = efecto de interacción fijo fósforo \times nitrógeno \times agua, e_{ijkl} = error aleatorio de subparcela con media 0 y varianza σ_e^2

Fuente	gl	SC	CM
Bloques	1	0.67	0.67
Fósforo (P)	1	1.25	1.25
Error(1)	1	27.56	27.56
Nitrógeno (N)	2	2 768.65	1 384.32
Agua (W)	2	751.84	375.92
NW	4	242.08	60.52
PN	2	12.71	6.35
PW	2	0.81	0.40
PNW	4	13.87	3.47
Error(2)	16	101.13	6.32

b.

Aqua	Nitrógeno	Fósforo		Medias de	
		0	245	<i>N</i> \times <i>W</i>	
16	0	8.4	12.6	10.5	
	130	35.3	33.7	34.55	
	260	37.7	36.7	37.2	
22	0	7.6	8.6	8.1	
	130	20.7	20.8	20.8	
	260	28.6	27.3	27.9	
28	0	7.6	8.4	8.0	
	130	20.6	19.8	20.2	
	260	21.3	23.1	22.2	

Agua	Fósforo		Fósforo		
	0	245	Nitrógeno	0	245
16	27.1	27.6	0	7.8	9.8
22	18.9	18.9	130	25.5	24.8
28	16.5	17.1	260	29.22	29.00

Medias(err est): $PNW(1.8)$, $NW(1.3)$, $PW(1.0)$, $PN(1.0)$

Medias P : 20.8, 21.2, error estándar = 1.2

Medias N : 8.8, 25.1, 29.1, error estándar = 0.7

Medias W : 27.4, 18.9, 16.8, error estándar = 0.7

- c. (i) 1.7, (ii) 1.0, (iii) 1.0, (iv) 1.8

- d. PNW : $F_0 = 0.55$, $F(.05, 4, 16) = 3.01$, no se rechaza

PW : $F_0 = 0.06$, $F(.05, 2, 16) = 3.63$, no se rechaza

PN : $F_0 = 1.01$, $F(.05, 2, 16) = 3.63$, no se rechaza

NW : $F_0 = 9.58$, $F(.05, 4, 16) = 3.01$, se rechaza

W : $F_0 = 59.48$, $F(.05, 2, 16) = 3.63$, se rechaza

N : $F_0 = 219.03$, $F(.05, 2, 16) = 3.63$, se rechaza

P : $F_0 = 0.05$, $F(.05, 1, 1) = 161.4$, no se rechaza

- e.

Fuente	gl	CM	F_0
N lineal (N_l)	1	2 464.43	389.92*
N cuadrática (N_q)	1	304.22	48.13*
W lineal (W_l)	1	672.04	106.33*
W cuadrática (W_q)	1	79.80	12.63*
$N_l W_l$	1	154.38	24.43*
$N_l W_q$	1	0.46	0.07
$N_q W_l$	1	40.89	6.47*
$N_q W_q$	1	46.35	7.33*
PN_l	1	7.04	1.11
PN_q	1	5.67	0.90
PW_l	1	0.01	0.00
PW_q	1	0.80	0.13
$PN_l W_l$	1	9.77	1.55
$PN_l W_q$	1	0.01	0.00
$PN_q W_l$	1	0.46	0.07
$PN_q W_q$	1	3.64	0.58

*Se rechaza la hipótesis nula ya que $F_0 \geq F(0.05, 1, 16) = 4.49$

Use un modelo de regresión cuadrática con N y W para la interpretación.

Capítulo 15

1. a. Diseño sujetos-entre: diseño totalmente aleatorizado con tres tratamientos de dieta y cuatro sujetos asignados al azar a cada dieta.

Diseño sujetos-dentro: tres mediciones repetidas de cada sujeto a 15, 20 y 45 minutos

- b.

Dieta	Minutos		
	15	30	45
1	19.0	35.0	31.5
2	22.0	20.0	11.5
3	26.0	27.3	35.8

c. $y_{ijk} = \mu + \alpha_i + d_{ik} + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$ $i, j = 1, 2, 3$ $k = 1, 2, 3, 4$

α_i = efecto fijo de dieta, β_j = efecto fijo de tiempo, $(\alpha\beta)_{ij}$ = efecto fijo dieta \times tiempo

d_{ik} = error experimental aleatorio independiente con distribución normal para dietas de sujetos-dentro con media 0 y varianza σ_d^2

e_{ijk} = error experimental aleatorio con distribución normal de mediciones repetidas con media 0 y varianza σ_e^2

Se supone que las e_{ijk} satisfacen la condición de Huynh-Feldt de varianzas iguales de la diferencia entre todos los períodos.

d.

Fuente	gl	SC	CM
Dieta (D)	2	1 020.67	510.33
Error(1)	9	413.33	45.93
Tiempo (T)	2	170.17	85.08
T lineal	1	92.04	92.04
T cuadrática	1	78.13	78.13
$D \times T$	4	869.67	217.42
$D \times T$ lineal	2	631.08	315.54
$D \times T$ cuadrática	2	238.58	119.29
Error(2)	18	128.17	7.12

Interacción dieta \times tiempo: $F_0 = 30.5$, $F(.05, 4, 18) = 2.93$, se rechaza
 Tiempo: $F_0 = 11.9$, $F(.05, 2, 18) = 3.55$, se rechaza

Dieta: $F_0 = 11.1$, $F(.05, 2, 9) = 4.26$, se rechaza

Medias de dieta: 28.5, 17.8, 29.7, error estándar = 1.96

Medias de tiempo: 22.3, 27.4, 26.3, error estándar = 0.77

Medias de celdas: vea medias en el inciso b., error estándar = 1.33

e. Vea particiones de sumas de cuadrados en la tabla de análisis de varianza

$D \times T$ lineal: $F_0 = 44.3$, $F(.05, 2, 18) = 3.55$, se rechaza

$D \times T$ cuadrática. $F_0 = 16.8$, $F(.05, 2, 18) = 3.55$, se rechaza

T lineal: $F_0 = 12.9$, $F(.05, 1, 18) = 4.41$, se rechaza

T cuadrática. $F_0 = 11.0$, $F(.05, 1, 18) = 4.41$, se rechaza

f. $W = 0.41$, aproximación chi-cuadrada = 7.12, 2 gl, $P = .0285$.

g.

Fuente	gl	SC	CM	F_0
Tiempo(lineal)	1	92.04	92.04	12.12*
Dieta $\times T$ (lineal)	2	631.08	315.54	41.53*
Error	9	68.40	7.60	
Tiempo (cuadrática)	1	78.13	78.13	11.76*
Dieta $\times T$ (cuadrática)	2	238.58	119.29	17.96*
Error	9	59.76	6.64	

*se rechaza la hipótesis nula

3. a. Los sujetos son bloques aleatorios en un diseño de bloques completo aleatorizado y cada sujeto recibe todos los tratamientos en orden aleatorio

b. $y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$ $i, j = 1, 2, \dots, 5$

τ_i = efecto fijo de tela, ρ_j = efecto aleatorio de sujeto, e_{ij} = error experimental aleatorio con media 0 y varianza σ^2 , se supone que es independiente y con distribución normal. Además suponga que no hay efectos transferidos de tratamiento

c.

Contraste	Tratamiento				
	1	2	3	4	5
(i) 1 vs. 2	1	-1	0	0	0
(ii) 2 vs. 4	0	1	0	-1	0
(iii) 2 vs. 5	0	1	0	0	-1
(iv) 3 vs. 4	0	0	1	-1	0

Los pares de contrastes ortogonales son (i) y (iv), y (iii) y (iv). La suma de coeficientes de productos cruzados de otros pares no es igual a cero

d.

Fuente	gl	SC	CM
Sujetos	4	17.94	4.49
Tratamientos	4	316.28	79.07
Error	16	227.08	14.19

Prueba para tratamientos: $F_0 = 5.57$, $F(.05, 4, 16) = 3.01$, se rechaza

Contraste	SC	F ₀
1 vs. 2	246.81	17.39*
2 vs. 4	69.48	4.90*
2 vs. 5	21.32	1.50
3 vs. 4	20.45	1.44

*Se rechaza la hipótesis nula ya que
 $F_0 \geq F(.05, 1, 16) = 4.49$.

5. a. Diseño sujetos-entre: diseño de bloque completo aleatorizado con cuatro variedades de parcelas asignadas al azar a cada uno de tres bloques
 Diseño sujetos-dentro: cuatro niveles de aplicación de agua en cuatro subparcelas espaciadas a distancias iguales a lo largo de la parcela de variedad.

b.

Variedad	Distancia			
	1	2	3	4
1	416.0	447.3	436.3	328.3
2	570.7	513.7	524.7	468.4
3	358.1	429.2	449.1	326.9
4	455.1	563.1	536.0	338.2

- c. $y_{ijk} = \mu + \alpha_i + \rho_k + d_{ik} + \beta_j + (\alpha\beta)_j + e_{ijk}$ $i, j = 1, 2, 3, 4$ $k = 1, 2, 3$
 α_i = efecto fijo de variedad, ρ_k = efecto fijo de bloque, β_j = efecto fijo de distancia, $(\alpha\beta)_j$ = efecto de interacción fijo variedad \times distancia, d_{ik} = error aleatorio de parcela competente con media 0 y varianza σ_d^2 , e_{ijk} = error aleatorio de subparcela con media 0 y varianza σ_e^2 . Se supone que los errores de la subparcela satisfacen la condición de Huynh-Feldt de varianzas iguales de las diferencias entre todas las distancias.

d.

Fuente	gl	SC	CM
Bloques	2	27 000	13 500
Variedades (V)	3	128 063	42 688
Error(1)	6	41 902	6 984
Distancia (D)	3	119 128	39 709
D lineal	1	39 145	39 145
D cuadrática	1	76 217	76 217
D cúbica	1	3 766	3 766
$V \times D$	9	44 837	4 982
$V \times D$ lineal	3	7 499	2 500
$V \times D$ cuadrática	3	36 473	12 158
$V \times D$ cúbica	3	864	288
Error(2)	24	177 735	7 406

Variedades: $F_0 = 6.11$, $F(.05, 3, 6) = 4.76$, se rechaza

Distancia. $F_0 = 5.36$, $F(.05, 3, 24) = 3.01$, se rechaza

Interacción $V \times D$ $F_0 = 0.67$, $F(.05, 9, 24) = 2.30$, no se rechaza

Medias de variedad: 407.0, 519.4, 390.8, 473.1 error estándar = 24.1

Medias de distancia: 450.0, 488.3, 486.5, 365.4, error estándar = 24.8

Medias de celdas: véa la tabla en el inciso b., error estándar = 49.7

- e. Vea particiones de sumas de cuadrados en la tabla de análisis de varianza.
 $V \times D$ lineal. $F_0 = 0.34$, $V \times D$ cuadrática: $F_0 = 1.64$, $V \times D$ cúbica: $F_0 = 0.04$, $F(.05, 3, 24) = 3.01$ No se rechazan las hipótesis nulas de interacción
 D lineal: $F_0 = 5.29$, D cuadrática: $F_0 = 10.29$, D cúbica. $F_0 = 0.51$, $F(.05, 1, 24) = 4.26$, se rechazan las particiones lineal y cuadrática
- f. $W = 0.29$, aproximación chi-cuadrada = 5.82, 5 gl, $P = .32$
- g. Épsilon G-G = .5634, gl ajustados: (distancia, 1.69), (variedad \times distancia, 5.07), (error, 13.57); las conclusiones no cambian.

h.	Fuente	gl	CM	F_0	$P > F$
	Bloques	2	55 834		
	D lineal	1	39 145	11.4	.0149
	$V \times D$ lineal	3	2 500	0.7	.5718
	Error	6	3 346		
	Bloques	2	1 044		
	D cuadrática	1	76 217	21.7	.0035
	$V \times D$ cuadrática	3	12 158	3.5	.0916
	Error	6	3 516		
	Bloques	2	2 138		
	D cúbica	1	3 766	1.3	.3052
	$V \times D$ cúbica	3	288	0.11	.9594
	Error	6	2 997		

Capítulo 16

1. a. Sí. Cada tratamiento sigue los demás tratamientos una vez, y cada tratamiento ocurre una vez en cada periodo y una vez en cada sujeto.

$$y_{ijk} = \mu + \alpha_i + b_j + \gamma_k + \tau_{d(i, k)} + \lambda_{c(i, k-1)} + e_{ijk}$$

i, k, d, c = 1, 2, 3, 4 j = 1, 2

α_i = efecto fijo de secuencia de tratamientos, b_j = efecto aleatorio de sujeto dentro de la secuencia con media 0 y varianza σ_b^2 , γ_k = efecto de periodo fijo, $\tau_{d(i, k)}$ = efecto directo fijo del tratamiento en el periodo k de la secuencia i, $\lambda_{c(i, k-1)}$ = efecto transferido fijo del tratamiento administrado en periodo k - 1 de la secuencia i, e_{ijk} = error aleatorio de la subparcela en el periodo k con media 0 y varianza σ_e^2 . Se supone que los errores experimentales tiene distribución normal y al menos satisfacen la condición de Huynh-Feldt de varianzas iguales para las diferencias entre todos los periodos. Además se supone que no hay efectos de transferencia más altos que de primer orden.

c.	Fuente	gl	SC	CM
	Secuencia	3	2.76	0.92
	Sujetos	4	4.71	1.18
	Periodos	3	0.05	0.02
	Droga(aj.)	3	24.25	8.08
	Transferencia(aj.)	3	0.06	0.02
	Error	15	1.99	0.13

Droga: $F_0 = 62.15$, $F(.05, 3, 15) = 3.29$, se rechaza

Transferencia: $F_0 = 0.15$, $F(.05, 3, 15) = 3.29$, no se rechaza

d.

Media de mínimos cuadrados para tratamiento

Contraste	A	B	C	D	c	s_c
A vs. B	1.17	1.16	3.28	1.17	0.01	0.19
B vs. C	0	1	-1	0	-2.12	0.19
B vs. D	0	1	0	-1	-0.01	0.19

A vs. B: $t_0 = 0.05$, B vs. C: $t_0 = -11.16$, B vs. D: $t_0 = -0.05$,

t de Bonferroni: $t(.025, 3, 15) = 2.69$, se rechaza para B vs. C, no se rechaza para A vs. B y B vs. D

3. a. Sí. Cada tratamiento sigue a los demás tratamientos, incluso a sí mismo una vez
 b. Sí. Los efectos de tratamiento directos y transferidos son ortogonales, mientras que no son ortogonales en el diseño de cuatro períodos.

c.	Fuente	gl	SC	CM
	Novillo	3	20.95	6.98
	Periodo	4	11.70	2.93
	Fibra	3	114.55	38.18
	Transferencia	3	3.25	1.08
	Error	6	12.50	2.08

d. Fibra: $F_0 = 18.36$, $F(.05, 3, 6) = 4.76$, se rechaza

Transferencia: $F_0 = 0.52$, $F(.05, 3, 6) = 4.76$, no se rechaza

e. $s(\widehat{\mu}_i - \widehat{\mu}_c) = 0.93$, $d(.05, 3, 6) = 3.10$, se rechaza si $|\widehat{\mu}_i - \widehat{\mu}_c| \geq D(3, .05) = 2.88$

Tratamiento	$\widehat{\mu}_i$	$\widehat{\mu}_i - \widehat{\mu}_{ci}$	ICS 95%
			(L, U)
Secilla de algodón	72.79	- 6.67	(- 9.55, - 3.79)
Bermuda	75.74	- 3.88	(- 6.76, - 1.00)
Trigo	77.49	- 2.113	(- 5.01, 0.75)
Alfalfa	79.62	Control	

La materia seca de semilla de algodón y bermuda es menor que la de alfalfa ya que los ICS no incluyen el 0.

5. a. Si. Cada tratamiento precede a los demás tratamientos tres veces.

b. $y_{ijk} = \mu + \alpha_i + \gamma_j + \tau_{d(i,j)} + \lambda_{c(i,j-1)} + e_{ij}$ $i = 1, 2, \dots, 12$ $j, d, c = 1, 2, 3, 4$

α_i = efecto fijo de secuencia, γ_j = efecto fijo de periodo, $\tau_{d(i,j)}$ = efecto fijo de tratamiento directo de teclado en periodo j y secuencia i , $\lambda_{c(i,j-1)}$ = efecto fijo de transferencia de tratamiento administrado en periodo $(j-1)$ y secuencia i ,

e_{ij} = error aleatorio con media 0 y varianza σ^2 . Se supone que los errores experimentales tienen distribución normal y que al menos satisfacen la condición de Huynh-Feldt de varianzas iguales entre todos los periodos. Suponga que no hay efectos de transferencia de orden más alto que el primero. La suposición de distribución normal puede no ser válida pues la variable de respuesta es el número discreto de errores cometidos por los sujetos.

Fuente	gl	SC	CM
Secuencia	11	12.23	1.11
Periodo	3	0.40	0.13
Distribución(aj.)	3	158.73	52.91
Transferencia(aj.)	3	1.21	0.40
Error	27	13.58	0.50

Distribución: $F_0 = 105.82$, $F(.05, 3, 27) = 2.96$, se rechaza

Transferencia: $F_0 = 0.80$, $F(.05, 3, 27) = 2.96$, no se rechaza

Capítulo 17

1. a. $y_{ij} = \mu_i + \beta(x_{ij} - \bar{x}) + e_{ij}$ $i = 1, 2, 3$ $j = 1, 2, \dots, 6$

μ_i = media de aleación, x_{ij} = diámetro de soldadura, β = coeficiente para regresión lineal de resistencia de soldadura, y , sobre el diámetro de soldadura, x , e_{ij} = error experimental aleatorio con media 0 y varianza σ^2 .

Suponga que los errores tienen distribución normal con varianza homogénea; β es la misma para todas las aleaciones, y los tratamientos de aleación no afectan el diámetro de soldadura

b.	Fuente	gl	SC	CM
	Regresión	1	2 838.92	2 838.92
	Aleación	2	2 005.97	1 002.99
	Error	14	716.16	51.15

Regresión: $F_0 = 55.50$, $F(.05, 1, 14) = 4.50$, se rechaza, $\hat{\beta} = 5.505$

Aleación: $F_0 = 19.61$, $F(.05, 2, 14) = 3.74$, se rechaza

c.	Aleación	\bar{y}_{ia}	Err est
	1	61.73	3.29
	2	78.04	3.25
	3	50.88	2.92

Error estándar promedio de la diferencia = 4.63

- d. Grafique las siguientes rectas de regresiones:

$$\text{Aleación 1: } \hat{y} = 50.42 + 5.505(x - 15.83)$$

$$\text{Aleación 2: } \hat{y} = 88.75 + 5.505(x - 19.83)$$

$$\text{Aleación 3: } \hat{y} = 51.50 + 5.505(x - 18.00)$$

e. $E = 3.69$

f. $CM(\text{homogeneidad}) = 129.13$, 2 gl, $CME = 38.16$, 12 gl

$$F_0 = 3.38$$
, $F(.05, 2, 12) = 3.89$, no se rechaza

i. coeficiente(err est): $\hat{\beta}_1 = 3.92(1.07)$, $\hat{\beta}_2 = 7.35(0.96)$, $\hat{\beta}_3 = 4.19(1.44)$

3. a. $y_{ij} = \mu + \tau_i + \rho_j + \beta(x_{ij} - \bar{x}_j) + e_{ij}$ $i = 1, 2, 3$ $j = 1, 2, 3, 4$

τ_i = efecto de irrigación, ρ_j = efecto de bloque, x_{ij} = número de plantas, β = coeficiente para regresión lineal de cosecha, y , sobre número de plantas, e_{ij} = error experimental aleatorio con media 0 y varianza σ^2

Suponga errores con distribución normal con varianzas homogéneas, β es la misma para todas las irrigaciones, y los tratamientos de irrigación no afectan el número de plantas.

b.	Fuente	gl	SC	CM
	Bloques	3	1.50	0.50
	Regresión	1	1.08	1.08
	Irrigación	2	1.52	0.76
	Error	5	0.44	0.09

Regresión. $F_0 = 12.0$, $F(.05, 1, 5) = 6.61$, se rechaza, $\hat{\beta} = 0.063$

Irrigación. $F_0 = 8.44$, $F(.05, 2, 5) = 5.79$, se rechaza

c.	Irrigación	\bar{y}_{ia}	Err est
	1	2.91	0.15
	2	2.23	0.15
	3	2.11	0.15

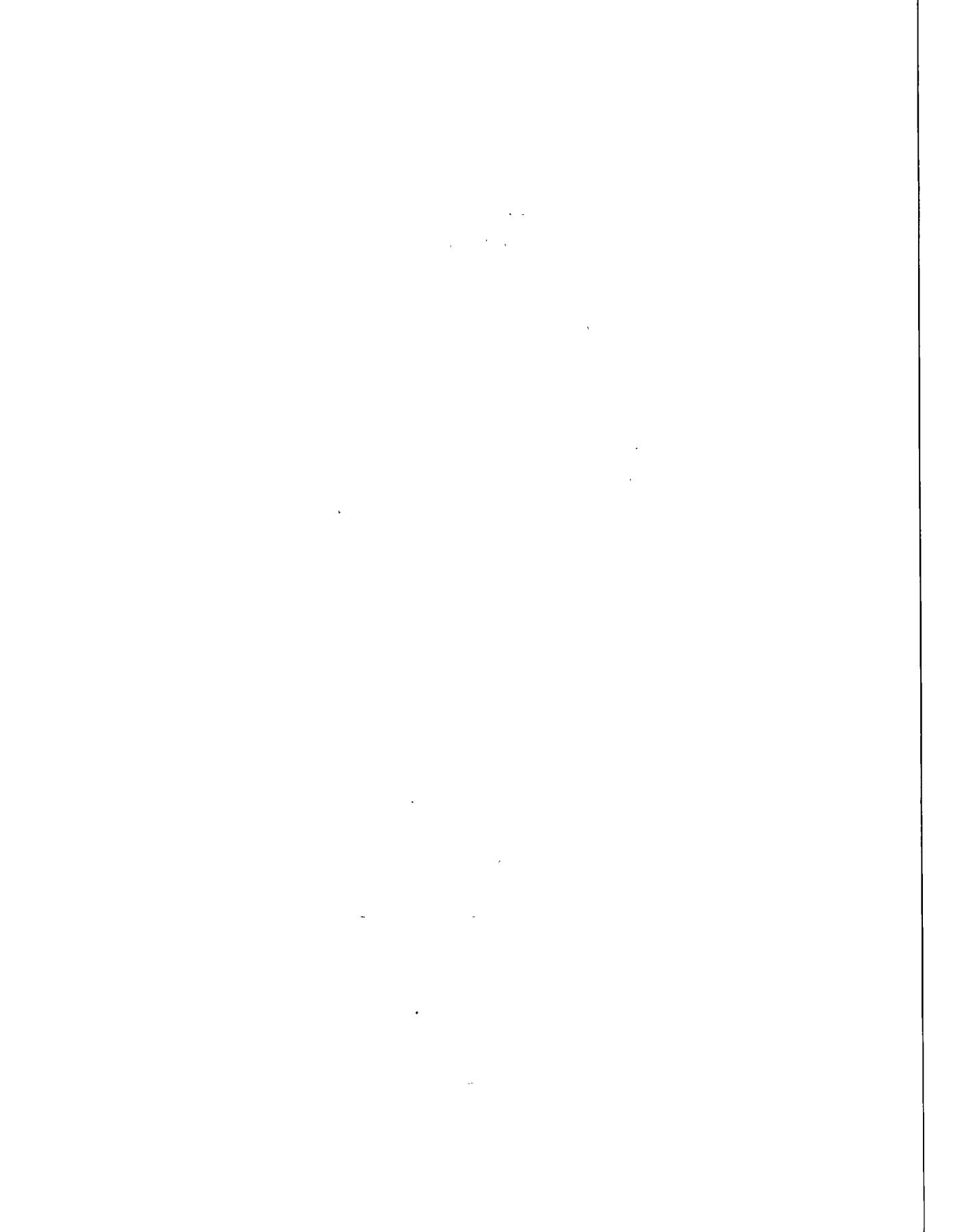
Error estándar promedio de la diferencia = 0.21

d. $E = 2.78$

5. a. $E_{xy} = 1.825$, $E_{xx} = 0.50$, $\hat{\beta} = 3.65$ = estimación de valor faltante

Fuente	gl	SC	CM
Bloques(no aj.)	3	1.95	
Regresión (x)	1	6.66	6.66
Irrigación (aj.)	2	1.38	0.69
Error	5	1.24	0.25

Irrigación: $F_0 = 2.76$, $F(.05, 2, 5) = 5.79$, no se rechaza



Índice

- Aditividad**
 - bloques completos aleatorizados, 266
 - cuadrado latino, 286
 - de los factores, 182, 205
- Ajustada**
 - medias de tratamiento, 556
 - suma de cuadrados, 316
 - total de tratamiento, 316, 344
- Aleatorización**, 20. Véase *Diseños individuales*
 - pruebas de aleatorización, 21
 - restringida, 24
- Aleatorización**, 39
- Análisis canónico**, 445, 467
- Análisis de covarianza**
 - eficiencia, 558
 - errores estándar, 557, 568
 - medias de tratamiento ajustadas, 556, 568
 - modelo lineal, 553, 560, 566
 - suposiciones, 559
- Análisis de varianza**
 - datos faltantes, 60, 208, 291, 486
 - hipótesis, 43, 47, 57, 123
 - números desiguales de submuestras, 164
 - series de experimentos, 296
 - submuestras, 159
 - véase también *diseños individuales*
- Análisis intrabloques**, 319
- Aproximación de Satterthwaite**, 168
- Archivos de datos**, 41
- Balance**, 312
 - bloque incompleto, 312
 - efectos acarreados, 529
 - parcial, 323
- Bloqueo por interacción de tratamiento**, 271
- Bloqueo**
 - aleatorio, 275
 - criterios, 10, 264, 310
 - eficiencia relativa de, 24, 273, 286
 - propósito, 8, 263, 310
 - renglón-columna, 276, 320, 342
- suma de cuadrados para**, 267, 316, 342
- tamaño**, 354
- Bonferroni**, 95
- Box-Béhnenk**, diseños de, 435
- Clases asociadas**, 324
- Coeficiente de variación**, 19
- Comparaciones en pares**, 107
 - DHS de Tukey, 108
 - Einot y Gabriel, 114
 - Hayter, 116
 - mínima diferencia significativa (MDS), 111
 - Student-Newman-Keuls, 112
- Comparaciones múltiples**, 91
 - con el mejor, 98
 - con un control, 104
- Comparaciones**. Véase *Contrastes*
- Componentes de la varianza**, 151, 234
 - análisis de estimaciones puntuales de la varianza, 152, 234
 - estimación, 258
 - estimaciones negativas, 155
 - modelos, 151, 234, 245, 252
- Componentes de varianza**. Véase *Varianza, componentes*
- Confundido**, 369
 - análisis de varianza, 371, 376
 - completo, 369
 - contrastos de definición, 379, 383
 - errores estándar, 376
 - factoriales 2n, 369
 - factoriales 3n, 383
 - método general, 378, 383
 - parcial, 372
- Contraste de definición**
 - factorial 2n, 379
 - factorial 3n, 383
- Contrastes**, 74
 - con diferente número de réplicas desiguales, 82, 103, 106, 109, 114
 - en pares, 107

- errores estándar, 75
- mediciones repetidas, 500, 506
- ortogonales, 81
- planeados, 74, 94, 115
- regresión. Véase Polinomios ortogonales
- suma de cuadrados, 76
- Control local, 8
- Correlación
 - en mediciones repetidas, 495, 511
 - intraclass, 155
 - serial, 512
 - subparcela, 472
- Covariada, 14
- Covarianza, 70, 495
- Cuadrados medios esperados
 - anidados y de factores cruzados, 253
 - bloque dividido, 486
 - de los factores, 186
 - diseño totalmente aleatorizado, 55, 62, 70
 - factores anidados, 245, 248, 251
 - modelos aleatorios, 152, 157, 234, 237
 - modelos mixtos, 239, 241, 242
 - parcelas divididas, 473
 - reglas, 255
 - series de experimentos, 295
 - submuestras, 162, 167
- Cuantil, 126
- Datos faltantes
 - diseño totalmente aleatorizado, 60
 - diseños de parcelas divididas, 486
 - diseños por bloque, 291
 - factoriales, 208
 - modelo de efectos aleatorios, 157, 166
- Diseño de bloques divididos, 483
- Diseño de bloques incompletos balanceados, 312
 - aleatorización, 313
 - análisis de varianza, 316
 - errores estándar, 318
 - medias de tratamiento, 318
 - modelo lineal, 315
- Diseño de bloques incompletos parcialmente balanceados, 322
- Diseño de bloques incompletos, 310
 - aleatorización, 313
 - balanceado, 312
 - cíclico, 345
 - diseños ∞ , 348
 - eficiencia, 325
 - elección, 352
 - estudios cruzados, 335
 - factoriales de 2^n , 369
 - factoriales de 3^n , 383
 - latinizado, 350
- parcialmente balanceado, 322
- renglón columna, 320, 342
- resoluble, 339
- retícula balanceada, 343
- retícula cuadrada, 343
- retícula rectangular, 341
- superficies de respuesta, 436
- Diseño de cuadrado latino, 275
 - aleatorización, 279
 - análisis de varianza, 282
 - eficiencia relativa, 286
 - errores estándar, 283-4
 - modelo estadístico lineal, 281
 - múltiple, 287
 - no aditividad, 286
- Diseño de los tratamientos, 5
- Diseño de parcelas divididas. Véase Diseños de bloques divididos
- Diseño de reanudación, 539
- Diseño de rectángulo latino, 288
- Diseño de tratamiento factorial, 7, 175
 - análisis de varianza, 186, 202
 - confundido, 369
 - de bloques completos, 289
 - de bloques incompletos, 369, 383
 - efectos, 177
 - errores estándar, 188, 204, 212, 217
 - modelo estadístico, 181, 201
 - réplica fraccionaria, 393
 - réplicas desiguales, 208
- Diseño dentro sujetos, 493
- Diseño entre sujetos, 493
- Diseño experimental, 13
- Diseño jerárquico. Véase Factor anidado
- Diseño totalmente aleatorizado, 13
 - aleatorización, 39
 - análisis de varianza, 55, 62
 - análisis, 47, 56, 59, 62
 - con diferente número de réplicas, 60
 - con igual número de réplicas, 38
 - modelo estadístico lineal, 43
- Diseños a, 348
 - características, 348
 - construcción, 349
 - eficiencia, 349
 - renglón-columna latinizado, 351
- Diseños anidados desequilibrados, 249
- Diseños centrales compuestos, 431
- Diseños cíclicos, 345
- Diseños con mediciones repetidas, 492
 - análisis de varianza, 500, 505, 508
 - contrastos de regresión, 500, 506
 - en el espacio, 510
 - error estándar, 502, 506

- Diseños cruzados de Williams, 532-533
- Diseños cruzados, 520
 - análisis de varianza, 526
 - balance para efectos acarreados, 523
 - codificación de archivos de datos, 545
 - de tratamiento incompleto, 535
 - dos tratamientos, 536
 - eficiencia, 531
 - errores estándar, 548-9
 - modelo lineal, 524
 - periodo extra, 535, 539
 - reanudación, 539
 - Williams, 532
- Diseños de bloques completos aleatorizados, 264
 - aleatorización, 265-266
 - análisis de varianza, 268
 - bloques aleatorios, 275
 - eficiencia relativa (ER), 273
 - errores estándar, 269
 - modelo estadístico, 266
 - no aditividad, 271
 - submuestras, 272
 - unidades experimentales múltiples por bloque, 271
- Diseños de factores anidados, 243
 - análisis de varianza, 246
 - cuadrados medios esperados, 246
 - desequilibrados, 249
 - errores estándar, 247
 - modelo lineal, 245
- Diseños de factores cruzados y anidados, 251
 - análisis de varianza, 254
 - cuadrados medios esperados, 253
 - modelo estadístico, 252
- Diseños de medidas realizados varias veces, 492
- Diseños de parcelas divididas, 469
 - aleatorización, 471
 - análisis de varianza, 473, 480
 - eficiencia relativa, 481
 - errores estándar, 478, 484
 - errores experimentales, 472
 - modelo mixto, 473
 - unidades experimentales, 471
- Diseños de retícula cuadrada balanceada, 343
- Diseños de retícula rectangulares, 341
 - latinizado, 351
- Diseños de retícula, 341
- Diseños de superficie de respuesta
 - Box-Behnken, 435
 - centrales compuestos, 431
 - de bloques incompletos, 436
 - factoriales 2n, 426
 - factoriales 3n, 431
 - rotatorio, 433
- Diseños de tratamiento con gradiente, 510
- Diseños para mezclas, 447
 - análisis, 453
 - de retícula simplex, 449
 - interpretaciones, 458
 - seudocomponentes, 452
 - simplex con centroídes ampliado, 450
 - simplex con centroídes, 450
 - sistema de coordenadas simplex, 448
- Diseños Plackett-Burman, 411
- Diseños renglón-columna anidados, 344
- Diseños renglón-columna, 320, 342, 523
- Diseños resolubles, 339
 - análisis de varianza, 342, 345
- Ecuaciones normales
 - en bloques incompletos balanceados, 337
 - en diseños factoriales, 225, 229
 - para modelo completo, 47
 - regresión, 464
- Efectos
 - acarreados, 522
 - aditividad, 182, 205
 - aleatorio, 148
 - de interacción, 178
 - de los factores, 177
 - del tratamiento, 53
 - directos, 524
 - fijos, 148, 182
 - principal, 178
 - simple, 177
- Eficiencia relativa. Véase *Diseños individuales*
- definición, 24
- verificación rápida, 274
- Eficiencia
 - análisis de covarianza, 558
 - de diseños de bloques incompletos, 325
 - relativa, 24
- Error con respecto a la comparación, 92
- Error estándar de efectos
 - factoriales de, 369
 - factoriales fraccionarios, 405
- Error estándar de la diferencia
 - análisis de covarianza, 558, 568-569
 - bloques completos aleatorizados, 269
 - bloques incompletos, 318
 - cuadrados latinos, 284
 - diseños cruzados, 548
 - modelos mixtos, 240
 - parcelas divididas, 478, 484
 - prueba de Dunnett, 104
 - prueba MDS, 110
- Errores estándar de la media
 - análisis de covarianza, 558
 - bloques completos aleatorizados, 269

- bloques incompletos, 318
- con submuestras, 162
- cuadrados latinos, 283
- diseño totalmente aleatorizado, 59, 62
- diseños anidados, 248
- factoriales, 188, 204
- modelos aleatorios, 155
- réplicas desiguales, 103, 109, 114, 213
- Error experimental, 4
 - varianza del, 4, 18, 43
- Error respecto al experimento, 92
 - control débil, 94
 - control fuerte, 94
- Errores estándar de los contrastes, 75
- Error
 - experimental, 4
 - tipo I y II, 19, 92
- Estimaciones de efectos de tratamiento
 - diseños renglón-columna, 345
 - factoriales, 226
 - intrabloques, 337
- Estudios de evaluación, i15, 392
- Estudios por observación comparativos, 4
- Experimento, 3
- F desplazada, 63
- Factor de eficiencia
 - diseño de bloques incompleto renglón-columna anidado, 345
 - diseños cruzados, 531
 - diseños ∞ , 349
 - en diseños de retícula, 341
 - para los diseños de bloque incompletos, 326
- Factor, 7, 176
 - cualitativo, 8, 177
 - cuantitativo, 8, 177
 - diversidad de, 392
 - notación, 176
- Factoriales 2n, 363
 - bloques incompletos, 369
 - columna identidad, 368
 - efectos, 364
 - etiquetas, 363
 - orden estándar, 367
- Factoriales 3n, 383
- Factoriales fraccionarios, 391
 - aleatorización, 398
 - análisis, 399
 - construcción, 396, 407, 409
 - diseño de Plackett-Burman, 411
 - diseño saturado, 411
 - error estándar, 405
 - fracción complementaria, 396
 - fracción principal, 396
- generador del diseño, 394
- relación de definición, 395
- resolución del diseño, 398
- técnica de doblez invertido, 413
- Fisher, Ronald A., 1, 156
- Grados de libertad, 54
 - aproximación de Satterthwaite, 168
- Gráfica cuantil-cuantil, 127
- Gráfica de perfiles, 493
- Gráfica de probabilidad normal, 127
- Gráfica de residuales, 128
 - de dispersión residual ajustada, 143
 - de probabilidad normal, 127
 - de separación-localización, 129
- Greenhouse-Geisser, ajuste, 505, 519
- Hipótesis de investigación, 5
- Homogeneidad de varianzas, 47
 - prueba F máx, 130
 - prueba Levene(Med), 128
- Huynh-Feldt
 - ajuste, 505, 519
 - condición de, 497-8
- Inferencia simultánea, 94
- Inferencia
 - de desigualdad confiable, 95, 108, 114
 - de direcciones confiables, 95
 - fuerza de, 94
 - intervalo de confianza, 80, 95
- Información interbloques, 320
- Interacción como error, 205
- Interacción
 - por tratamiento en cada bloqueo, 271, 286
 - entre factores, 175, 178
 - generalizada, 377, 381, 386
 - tratamiento por experimento, 296
- Intervalos de confianza
 - contrastos, 76
 - inferencia, 80, 94
 - media de tratamiento, 59
 - o valores P, 79
 - simultáneos, 96
 - varianza, 153
- Inusitados, 131
- Logit, 134
- Media de tratamiento, 42
 - estimación con mínimos cuadrados, 47, 181, 212, 318
- Método DHS de Tukey, 107
- Mínima diferencia significativa (MDS), 110
- Mínimos cuadrados, 47
 - bloque incompleto balanceado, 336
 - diseño totalmente aleatorizado, 47
 - factoriales, 225
 - regresión, 463

- Modelo completo**
 - análisis de covarianza, 555, 567
 - datos faltantes, 217, 291
 - diseño de bloques incompleto balanceado, 315
 - diseño mixto, 455
 - diseño totalmente aleatorizado, 43
 - diseños cruzados, 527
 - factoriales, 227
 - regresión, 463
- Modelo de medias de celdas**, 42, 181, 201, 211
- Modelo estadístico**, 151, 233
 - cuadrados medios esperados, 152, 234, 237
- Modelo lineal**, 45
 - generalizado, 140
- Modelo mixto**, 239
 - cuadrados medios esperados, 241
 - errores estándar, 240
 - mediciones repetidas, 499
- Modelo reducido**. Véase **Modelo completo**
- Modelo**. Véase **Modelo estadístico lineal**
- Muestreo**. Véase **Submuestas**
- No aditividad**. Véase **Aditividad**
- No ortogonal**, Véase **Ortogonalidad**
- Número de réplicas**
 - para factoriales, 208, 255
 - para la prueba F, 63
 - para la prueba z, 18
 - para varianzas, 158
- Observación base**, 538
- Ortogonalidad**
 - bloques y tratamientos, 291, 315, 344
 - contrastos, 81
 - diseños cruzados, 526, 534
 - factoriales, 209
- Parámetros no centrados**, 63, 208
- Parcela completa**, 469
- Particiones de suma de cuadrados**
 - análisis de covarianza, 555, 562, 567-568
 - análisis de superficie de respuesta, 442
 - bloques completos aleatorizados, 267
 - bloques incompletos, 315, 325
 - contrastos, 76
 - cuadrado latino, 281
 - diseño completo aleatorizado, 52
 - factoriales, 185, 202
 - polinomios ortogonales, 87
 - series de experimentos, 298
 - submuestas, 161
- Periodo de descanso (limpieza)**, 522
- Placebo**, 7, 27, 60
- Planes de diseño**
 - cíclicos, 360
 - diseños ∞ , 361
 - en bloques incompletos balanceados, 330
- en cuadrados latinos incompletos**, 332
- en cuadrados latinos**, 307
- referencias**, 313, 321, 324, 353, 416, 439
- Polinomios ortogonales**, 85, 190
 - errores estándar, 89
- Potencia**, 19, 63, 158
- Probits**, 134
- Prueba de Dunnett**, 105
- Prueba de intervalos múltiples**, 111
- Prueba de Scheffé**, 97
- Prueba de uniformidad**, 11
- Prueba del estadístico (w)de Mauchly**, 498
- Prueba doble ciega**, 28
- Prueba F Máx**, 114
- Prueba F**, 57
 - aproximada, 168, 236
- Prueba Levene (Med)**, 128
- Prueba Student-Newman-Keuls (SNK)**, 111-112
- Regla pares-impares**, 371, 397
- Regresión**, 441, 463
- Réplica**, 16
 - desiguales, Véase **Datos faltantes**
- Residuales**, 124
 - bloques completos aleatorizados, 272
 - de Student, 132
 - estandarizados, 131
 - factoriales, 190
 - valor absoluto, 128
- Residuo mod m**, 378
- Resolución del diseño**, 398
- Seudoréplicas**, 18
- Simetría compuesta**, 496
- Similaridad**, 11
- Submuestas**, 159
 - cuadrados medios esperados, 162, 167
 - diseños de bloques incompletos, 272
 - números desiguales, 164
 - solución óptima de la asignación, 163
- Subparcela**, 469
- Suma de cuadrados**
 - de los tratamientos ajustada, 316, 345, 527, 548-509, 555
 - error experimental, 49
 - no ajustada en bloques incompletos, 316
- Sumas de cuadrados medios ponderados**, 215
- Superficie de respuesta**
 - análisis canónico, 445, 467
 - análisis, 442
 - experimentos secuenciales, 425
 - exploración, 445
 - modelo de primer orden, 425
 - modelo de segundo orden, 425
 - punto estacionario, 444, 466

- Taguchi, G., 413
- Tamaño de muestra. Véase Número de réplicas
- Técnica, 8
- Transformaciones, 133
 - arco seno o angular, 134
 - escalera de exponentes, 135
 - exponentes, 135
 - log, 135
 - raíz cuadrada, 133
- Tratamiento de control, 6, 104
- Tratamientos, 4
- Trayectoria de la mayor pendiente, 429, 432
- Unidad de observación, 16
- Unidad experimental, 4
 - vs. unidad de observación, 16
- Valor f, 142
- Valores P, 58
 - o intervalos de confianza, 79
- Varianza
 - de una diferencia, 325, 531-532
 - de una media, 25, 162, 247, 275
 - error experimental, 4, 17, 43
 - intervalo de confianza, 153
 - pruebas de homogeneidad, 128

Esta obra se terminó de imprimir en enero del 2001
En los talleres de Litográfica Ingramex S.A. de C.V.
Centeno 162-1, Col. Granjas Esmeralda
C.P. 09810 México, D.F.

2a.

edición

Robert O. Kuehl

Principios
estadísticos
para
el diseño
y
análisis
de
investigaciones

Este libro enseña los principios del diseño de estadísticas y análisis de datos para efectuar estudios científicos comparativos. Profundiza en todos los pasos del diseño de experimentos y presenta al estudiante un punto de vista estructural. Por otro lado, a fin de proporcionar una visión amplia de las dimensiones de esta disciplina, los ejercicios y problemas presentan gran variedad de conjuntos de datos tomados de diferentes áreas tales como ingenierías, industria, cuestiones agrícolas e investigaciones químicas; además, la mayoría de los ejemplos se basan en casos reales. Finalmente, el autor introduce un enfoque muy moderno en este campo, que es la inferencia estadística simultánea.

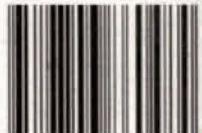
MÉXICO Y AMÉRICA CENTRAL
Tel. (525)281-2906
editor@thomsonlearning.com.mx
Mexico, D.F., MÉXICO

AMÉRICA DEL SUR
Tel./Fax (5411)4325-2236
thomson@thomsonlearning.com.ar
Buenos Aires, ARGENTINA

EL CARIBE
Tel. (787)758-7580
thomson@coqui.net
Hato Rey, PUERTO RICO

PACÍFICO ANDINO
Tel. (571)340-9470
clthomson@andinet.com
Bogotá, COLOMBIA

ISBN 970-686-048-7



9 789706 860484

Diseño de experimentos



THOMSON
LEARNING