

# **Diseño Experimental**

Msc. Julio Hurtado Márquez

Esp. Roberto Gómez Fernández

*E-mail address:* `jhurtado@unitecnologica.edu.co`, `rgomez@unitecnologica.edu.co`



# Contents

Prefacio	vii
<b>Part 1. Diseños de experimentos</b>	<b>1</b>
Chapter 1. Introducción al Diseño de Experimentos	3
1. Objetivos de un diseño de experimento	4
2. Principios básicos del diseño de experimentos	4
3. Conceptos básicos	4
4. Etapas de un diseño de experimentos.	5
Chapter 2. <b>Experimentos con un solo factor</b>	7
1. Competencias	7
2. <b>Experimentos con un solo factor</b>	7
3. Familia de diseños para comparar tratamientos	8
4. Supuestos acerca del modelo estadístico	9
5. <b>Diseño completamente al azar (DCA) y ANOVA</b>	9
6. ANOVA para el diseño completamente al azar (DCA)	11
7. Comparaciones de rangos múltiples	14
8. Verificación de los supuestos del modelo	19
9. <b>Elección del tamaño de la muestra</b>	23
10. Ejercicios: experimentos con un sólo factor	23
Chapter 3. Diseños en bloques	33
1. Competencias	33
2. Diseños en bloques completos al azar	33
3. Diseño en cuadro latino (DCL)	37
4. <b>Diseño en cuadro grecolatino (DCGL)</b>	40
5. Ejemplo de un Diseño en cuadro latino (DCL) y en cuadro grecolatino (DCGL)	42
6. Diseño de bloques incompletos balanceados (BIBD)	47
7. Ejercicios: Diseños en bloques completos al azar	52
Chapter 4. Diseños Factoriales	61
1. Competencias	61
2. Introducción	61
3. Conceptos básicos en diseños factoriales	61
4. Diseños factoriales con dos factores	63
5. Diseños factoriales con tres factores	70
6. Ejercicios: Diseños Factoriales	73
Chapter 5. Diseños factoriales $2^k$	83

1. Competencias	83
2. Diseños factoriales $2^2$	83
3. Diseño factorial $2^3$	94
4. Diseño Factorial General $2^k$	104
5. Diseño factorial $2^k$ no replicado	109
6. Factoriales $2^k$ con punto al centro	126
7. Factoriales $2^k$ en bloques	128
8. Ejercicios: Diseños factoriales $2^k$	131
Chapter 6. Diseños factoriales fraccionados $2^{k-p}$	145
1. Competencias	145
2. Introducción	145
3. Diseños factoriales fraccionados $2^{k-1}$	146
4. El concepto de resolución	149
5. Construcción de fracciones $2^{k-1}$	150
6. Diseños factoriales fraccionados $2^{k-2}$	154
7. Diseño factorial fraccionado $2^{k-p}$	158
8. Ejercicios: Diseños factoriales fraccionados $2^{k-p}$	164
Chapter 7. Metodología de superficie de respuesta	169
1. Competencias	169
2. Introducción a la Metodología de Superficie de Respuesta (MSR)	169
3. Método del ascenso más pronunciado	171
4. Esquema de los elementos de la MSR en su contexto amplio	172
5. Análisis de una superficie de respuesta de segundo orden	180
6. Sistemas de cordilleras	187
7. Respuestas múltiples	188
8. Diseños de superficies de respuesta	188
9. Ejercicios: Metodología de superficie de respuesta	207
Chapter 8. Diseño Robusto	215
1. Introducción	215
2. Enfoques de <i>Gen'ichi Taguchi</i> en el diseño de experimentos	215
3. El concepto de robustez	220
4. Factores de control, de ruido y de señal	221
5. Arreglos ortogonales	225
6. Diseño con arreglo interno y externo (diseño de parámetros)	228
7. El cociente señal/ruido	229
8. Un ejemplo ilustrativo de un experimento robusto	231
9. Otro ejemplo ilustrativo de un diseño robusto	235
10. Ejercicios	240
Chapter 9. Diseños anidados y diseños en parcelas divididas	245
1. Competencias	245
2. Introducción	245
3. Diseños anidados	245
4. Modelo y análisis estadístico de diseños anidados	246
5. Modelo y análisis estadístico de diseños cruzado-anidados	253
6. Diseño en parcelas divididas ( <i>split-plot</i> )	265
7. Modelo y análisis estadístico de los diseños en parcelas divididas	271

CONTENTS

v

8. Ejercicios: Diseños anidados y diseños en parcelas divididas	277
Bibliography	283



## Prefacio

Estimados estudiantes de Maestría en Ingenierías y Especialización en Estadística Aplicada, bienvenidos al curso de Diseño de Experimentos. El interés principal de estas notas de clases es sentar las bases para la preparación de un documento fruto de investigación y trabajo en común, para el cual están cordialmente invitados, puesto que este curso los preparará para iniciar investigaciones en cualquiera de los campos de la Ingeniería, las Ciencias de la Salud, las Humanidades y las Ciencias Básicas, donde se vean involucrados los diseños experimentales, y es deseo del autor, incorporar a éstas notas esas futuras investigaciones. En este documento encontramos extensas partes tratadas textualmente en los textos: Análisis y Diseños de experimentos de los autores Humberto Gutiérrez Pulido y Román de la Vara Salazar; y Diseño y análisis de experimentos de Douglas Montgomery, que son los textos más importantes y reconocidos en este tema y de antemano, los invito a que los adquieran, pues, serán de gran ayuda para el éxito del curso. Muchas cosas se quedan por fuera de estas notas, pero esperamos que siembre el deseo de fomentar una cultura en este campo.





## **Part 1**

# **Diseños de experimentos**



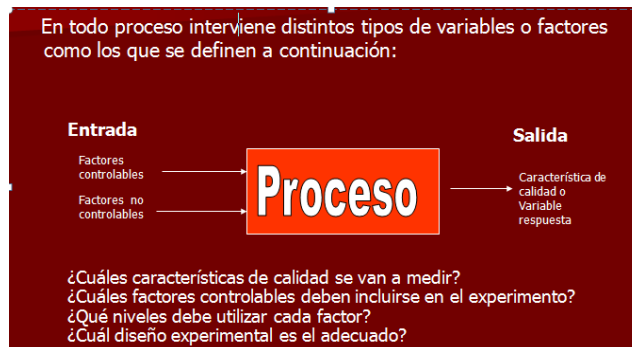
## CHAPTER 1

# Introducción al Diseño de Experimentos

El diseño experimental (DE) tiene inicio teórico a partir de 1935 por *Sir Ronald A. Fisher* (1890 – 1962), desarrollados en la Estación Agrícola Experimental de Rothamsted, en el Reino Unido, donde introdujo el concepto de *aleatorización* y *el análisis de varianza*. A lo largo de varias décadas, la teoría del diseño de experimentos y sus aplicaciones se consolidaron y expandieron, y en años recientes, recibieron un fuerte impulso por las contribuciones de *Genichi Taguchi*, un estadístico japonés ampliamente conocido en Occidente.

- *Se puede afirmar que el diseño experimental estudia procesos. Un proceso puede considerarse como una caja negra a la cual ingresan diversas variables que interactúan para producir un resultado. Las variables que ingresan al proceso se denominan variables de entrada (factores) y el resultado, variable de salida (respuesta). El nivel de la variable de salida depende de los niveles que adopten las variables de entrada. La búsqueda de combinaciones óptimas de las variables de entrada da lugar al diseño experimental, el cual es una prueba (o un conjunto de pruebas) durante la cual se realizan cambios sistemáticos y controlados a las variables de entrada para medir el efecto sobre la variable de salida óptima.*

El diseño experimental utiliza técnicas como la regresión múltiple, la superficie de respuesta, el diseño de parámetros y varias extensiones del análisis de varianza, así como los análisis gráficos y las comparaciones múltiples. Diseñar un experimento significa planear un experimento de modo que reúna la información pertinente al problema bajo investigación. El DE es la secuencia completa de pasos tomados con antelación, para asegurar que los datos se obtendrán apropiadamente de modo que permitan un análisis objetivo que conduzca a deducciones válidas con respecto al problema bajo estudio.



La necesidad de un diseño de experimento surge de la inquietud de responder a preguntas como:

- (1) ¿Cuáles son los factores que se estudiarán en esta investigación?
- (2) ¿Qué factores afectan las características que se van a analizar?
- (3) ¿Cómo se va a medir el efecto de los factores estudiados?
- (4) ¿Cuántas veces deberá ejecutarse el experimento?
- (5) ¿Cuál será la forma de análisis o tipo de diseño experimental?
- (6) ¿A partir de que valores se considera importante el efecto de un factor?

### 1. Objetivos de un diseño de experimento

- (1) Proporcionar la máxima cantidad de información pertinente al problema bajo investigación.
- (2) El diseño, plan o programa debe ser tan simple como sea posible.
- (3) Proporcionar métodos que permitan obtener la mayor cantidad de información válida acerca de una investigación, teniendo en cuenta el factor costo y el uso adecuado del material disponible mediante métodos que permitan disminuir el error experimental.

### 2. Principios básicos del diseño de experimentos

- (1) **Repetición.** Viene a ser la reproducción o réplica del experimento básico (asignación de un tratamiento a una unidad experimental). Las principales razones por las cuales es deseable la repetición son: primero por que proporciona una estimación del error experimental, siendo tal estimación confiable a medida que aumenta el número de repeticiones, y segundo permite estimaciones más precisas del tratamiento en estudio.
- (2) **Aleatorización.** Asignación al azar de tratamiento a las unidades experimentales. Una suposición frecuente en los modelos estadísticos de un diseño de experimentos es que: *las observaciones o los errores están distribuidos independientemente*, la aleatorización hace válida esta suposición.
- (3) **Control Local.** Cantidad de balanceo, bloqueo y agrupamiento de las unidades experimentales que se emplean en el diseño estadístico de un experimento.

### 3. Conceptos básicos

- (1) **Experimento.** Prueba o series de pruebas en las que se hacen cambios deliberados en las variables de entrada de un proceso (los factores que se estudian) o sistema para observar e identificar las razones de los cambios que pudieran observarse en la respuesta de salida.
- (2) **Unidad experimental.** Unidad a la cual se le aplica un sólo tratamiento (que puede ser una combinación de muchos factores) en una reproducción del experimento.
- (3) **Error experimental.** Describe la situación de no llegar a resultados idénticos con dos unidades experimentales tratadas de igual forma y refleja: (a) errores de experimentación, (b) errores de observación, (c) errores de medición, (d) variación del material experimental (esto es, entre unidades experimentales), (e) efectos combinados de factores extraños que

podrían influir las características en estudio, pero respecto a los cuales no se ha llamado la atención en la investigación.

- (4) **Agrupamiento.** Colocación de un conjunto de unidades experimentales homogéneas en grupos, de modo que los diferentes grupos puedan sujetarse a distintos tratamientos.
- (5) **Bloqueo.** Distribución de las unidades experimentales en bloques, de manera que las unidades dentro de un bloque sean relativamente homogéneas, de esta manera, la mayor parte de la variación predecible entre las unidades queda confundida con el efecto de los bloques.
- (6) **Balanceo.** Obtención de las unidades experimentales, el agrupamiento, el bloqueo y la asignación de los tratamientos a las unidades experimentales de manera que resulte una configuración balanceada.
- (7) **Tratamiento o combinación de tratamientos.** Conjunto particular de condiciones experimentales que deben imponerse a una unidad experimental dentro de los confines del diseño seleccionado.
- (8) **Factor.** Una variable independiente que tiene influencia sobre la respuesta de salida. Generalmente, se trabaja con más de una variable independiente y con los cambios que ocurren en la variable dependiente, cuando ocurren variaciones en una o más variables independientes.
- (9) **Confusión.** Cuando los efectos de dos o mas factores no se pueden separar.

#### 4. Etapas de un diseño de experimentos.

- (1) Enunciado o planteamiento del problema.
- (2) Formulación de hipótesis.
- (3) Proposición de la técnica experimental y el diseño.
- (4) Examen de sucesos posibles y referencias en que se basan las razones para la indagación que asegure que el experimento proporcionará la información requerida y en la extensión adecuada.
- (5) Consideración de los posibles resultados desde el punto de vista de los procedimientos estadísticos que se aplicarán, para asegurar que se satisfagan las condiciones necesarias y sean válidos estos procedimientos.
- (6) Ejecución del experimento.
- (7) Aplicación de las técnicas estadísticas a los resultados experimentales.
- (8) Extracción de conclusiones con medidas de la confiabilidad de las estimaciones generadas.
- (9) Valoración de la investigación completa y contrastación con otras investigaciones del mismo problema o similares.



## CHAPTER 2

# Experimentos con un solo factor

### 1. Competencias

- (1) Identificar dentro de la familia de los diseños experimentales, aquellos utilizados en la comparación de tratamientos.
- (2) Diferenciar los distintos modelos estadísticos y los análisis de varianzas en experimentos con un sólo factor.
- (3) Realizar las diversas pruebas de rangos múltiples y la comparación por contrastes.
- (4) Verificar los supuestos del modelo estadístico en diseños con un solo factor.

### 2. Experimentos con un solo factor

En este tipo de diseño de experimento se considera un *sólo factor de interés* y el objetivo es comparar más de dos tratamientos, con el fin de elegir la mejor alternativa entre las varias que existen, o por lo menos para tener una mejor comprensión del comportamiento de la variable de interés en cada uno de los distintos tratamientos.

*El interés del experimentador se centra en comparar los tratamientos en cuanto a sus medias poblacionales*, sin olvidar que también es importante compararlos en relación a sus varianzas y su capacidad actual y futura para cumplir con los requerimientos de calidad y productividad. La hipótesis fundamental a probar cuando se comparan varios tratamientos es:

$$\begin{aligned} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu \\ H_1 : \mu_i \neq \mu_j \text{ para algún } i \neq j \end{aligned}$$

con la cual se quiere decidir si los tratamientos son iguales estadísticamente en cuanto a sus medias, contra la alternativa de que al menos dos de ellos son diferentes.

La estrategia natural para resolver este problema es obtener una muestra representativa de mediciones en cada uno de los tratamientos con base en las medias y varianzas muestrales, construir un estadístico de prueba para decidir el resultado de dicha comparación.

En caso de aceptar  $H_0$ , se concluye que los tratamientos son iguales en cuanto a la media de cierta variable de salida de interés. En caso de rechazar  $H_0$  se acepta como verdadera la hipótesis alternativa  $H_1$ , que indica que al menos dos de los tratamientos son diferentes entre sí, y faltaría investigar precisamente cuales de ellos, son los causantes de las diferencias detectadas.

Es importante que las posibles diferencias entre los datos se deban principalmente al *factor de interés*, y no a los factores que no se consideran. Al existir otros factores no controlables se puede afectar sensiblemente las conclusiones.

**EJEMPLO 1.** *En el caso de comparar varias máquinas, si cada máquina es manejada por un operador diferente y se sabe que éste tiene una influencia en el resultado, entonces, es claro que el factor operador debe tomarse en cuenta si se quiere comparar a las máquinas de manera justa.*

Un operador más hábil puede ver a su máquina (aunque ésta sea la peor) como la que tiene el mejor desempeño, lo que impide una comparación adecuada de los equipos. Para evitar este sesgo habría dos maneras de anular el posible efecto del factor operador:

- (1) Utilizando el mismo operador en las cuatro máquinas. Esta estrategia no es aconsejable, ya que al utilizar el mismo operador, se elimina el efecto del factor operador, pero restringe la validez de la comparación a dicho operador, y es posible que el resultado no se mantenga al utilizar otros operadores.
- (2) Cada operador trabaje durante el experimento con cada una de las máquinas, esta estrategia es más recomendable, ya que al utilizar todos los operadores con todas las máquinas permite tener resultados de la comparación que son válidos para todos los operadores. Esta última de manera nulificar el efecto de operadores, recibe el nombre de *Bloqueo*.

**2.1. Factores de bloqueo.** Son factores adicionales al factor de interés que se incorporan de manera explícita en un experimento comparativo, para estudiar de manera más adecuada y eficaz al factor de interés.

**OBSERVACIÓN 1.** *Cuando se comparan varias máquinas, manejadas por operadores diferentes, es pertinente incluir explícitamente al factor operadores (bloques) para lograr el propósito del estudio. También se podrían controlar el tipo de material, lotes, tipo de producto, día, turno, etc. Se controlan factores que por conocimiento del proceso o experiencia previa, se sabe que pueden afectar en forma sensible el resultado de la comparación.*

### 3. Familia de diseños para comparar tratamientos

Los diseños experimentales más utilizados para comparar tratamientos son:

- (1) Diseño completamente al azar (DCA)
- (2) Diseño en bloque completamente al azar (DBCA)
- (3) Diseño en cuadro latino (DCL)
- (4) Diseño en cuadro grecolatino (DCGL)

La diferencia fundamental entre estos diseños es el número de factores de bloque que incorporan o controlan de forma explícita durante el experimento. La comparación de los tratamientos en cuanto a la respuesta media que logran, en cualquiera de estos diseños, se hace mediante la hipótesis

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu$$

$$H_1 : \mu_i \neq \mu_j \text{ para algún } i \neq j$$

que se prueba con la técnica estadística llamada *Análisis de Varianza* (ANOVA) con uno, dos, tres o cuatro criterios de clasificación, dependiendo del número de factores de bloques incorporados al diseño.



Tabla 2			
Diseño	Factores de bloqueo	ANOVA con	Modelo estadístico
DCA	0	un criterio	$Y_i = \mu + \tau_i + \varepsilon_i$
DBCA	1	dos criterios	$Y_{ij} = \mu + \tau_i + \gamma_j + \varepsilon_{ij}$
DCL	2	tres criterios	$Y_{ijk} = \mu + \tau_i + \gamma_j + \delta_k + \varepsilon_{ijk}$
DCGL	3	cuatro criterios	$Y_{ijkl} = \mu + \tau_i + \gamma_j + \delta_k + \varphi_l + \varepsilon_{ijkl}$

NOTA 1.  $Y$  es la variable de salida,  $\mu$  la media global,  $\tau_i$  el efecto del  $i$ -ésimo tratamiento,  $\varepsilon :=$  error aleatorio, y  $\gamma_j, \delta_k, \varphi_l :=$  son los efectos de tres factores de bloqueo.

El modelo estadístico que describe el comportamiento de la variable observada  $Y$  en cada diseño, incorpora un término adicional por cada factor de bloqueo controlado.

De acuerdo con los modelos dados en la tabla, para cada diseño comparativo se tienen al menos dos fuentes de variabilidad: *los tratamientos* o niveles del factor de interés y *el error aleatorio*. Se agrega una nueva fuente de variabilidad por cada factor de bloque que se controla directamente. Se observa que los diseños suponen que no hay efectos de interacción entre los factores, lo cual sería lo deseable que ocurra; de no ocurrir así, tal efecto se recarga al error y el problema de comparación no se resuelve con éxito.

Un efecto de interacción entre dos factores hace referencia a que el efecto de cada factor depende del nivel en que se encuentra el otro.

#### 4. Supuestos acerca del modelo estadístico

- (1) *Aditividad*: Los factores o componentes del modelo estadístico son aditivos, es decir la variable respuesta es la suma de los efectos del modelo estadístico.
- (2) *Linealidad*: La relación existente entre los factores o componentes del modelo estadístico es del tipo lineal.
- (3) *Normalidad*: Los valores resultado del experimento provienen de una distribución de probabilidad *Normal* con media  $\mu$  y variancia  $\sigma^2$ .
- (4) *Independencia*: Los resultados observados de un experimento son independientes entre sí.
- (5) *Variancias Homogéneas (Homocedasticidad)*: Las diversas poblaciones generadas por la aplicación de dos o más tratamientos tienen variancias homogéneas (variancia común).

#### 5. Diseño completamente al azar (DCA) y ANOVA

El diseño completamente al azar (DCA) es el más simple de todos los diseños, sólo considera dos fuentes de variabilidad: *los tratamientos* y *el error aleatorio*. Se llama completamente al azar por que todas las corridas experimentales se realizan en orden aleatorio completo; en apariencia no existe ninguna restricción a la aleatorización, es decir, si durante el estudio se hacen en total  $N$  pruebas, estas se corren al azar, de manera que los posibles efectos ambientales y temporales se repartan equitativamente entre los tratamientos.

DEFINICIÓN 1. En un DCA se tienen  $k$  poblaciones o tratamientos, independientes y con medias desconocidas  $\mu_1, \mu_2, \dots, \mu_k$ , y varianzas también desconocidas pero que se suponen iguales  $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$  ( $= \sigma^2$ ). Las poblaciones pueden ser  $k$  métodos de producción,  $k$  tratamientos,  $k$  grupos, etc. y sus medias se refieren o son medidas en términos de la variable de respuesta. Se comparan las poblaciones, en principio mediante la hipótesis de igualdad de medias.

Los datos generados por un DCA para comparar dichas poblaciones, se pueden escribir como en la tabla 2. El elemento  $Y_{ij}$  en esta tabla es la  $j$ -ésima observación que se hizo en el tratamiento  $i$ ;  $n_i$  son las repeticiones observadas en el tratamiento  $i$ . Se recomienda utilizar el mismo número de repeticiones ( $n_i = n$ ) en cada tratamiento, a menos que hubiera alguna razón para no hacerlo. Cuando  $n_i = n$  para toda  $i$ , se dice que el diseño es *balanceado*.

OBSERVACIÓN 2. El número de tratamientos  $k$  es determinado por el investigador y depende del problema particular de que se trata. El número de observaciones por tratamiento ( $n$ ) debe escogerse con base en la variabilidad que se espera observar en los datos, a la diferencia mínima que el experimentador considera que es importante detectar y al nivel de confianza que se desea tener en las conclusiones. Con este tipo de consideraciones, generalmente se recomiendan entre 10 y 30 mediciones en cada tratamiento. Se usa  $n = 10$  cuando las mediciones dentro de cada tratamiento tienen un comportamiento consistente (con poca dispersión). En el otro extremo, se recomienda  $n = 30$  cuando las mediciones muestran bastante dispersión. Con  $n = 10$  se pueden detectar **diferencias de medias** mayores o iguales a  $1.5\sigma$  con probabilidad alta y con  $n = 30$  se pueden detectar diferencias iguales o mayores a  $0.7\sigma$ . Cuando es costoso o tardado realizar las pruebas para cada tratamiento, se puede seleccionar un número menor de repeticiones, con lo cual sólo se podrán detectar diferencias grandes entre los tratamientos.

(5.1)

Tabla 3. Tratamientos				
$T_1$	$T_2$	$T_3$	$\dots$	$T_k$
$Y_{11}$	$Y_{21}$	$Y_{31}$	$\dots$	$Y_{k1}$
$Y_{12}$	$Y_{22}$	$Y_{32}$	$\dots$	$Y_{k2}$
$Y_{13}$	$Y_{23}$	$Y_{33}$	$\dots$	$Y_{k3}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Y_{1n_1}$	$Y_{2n_2}$	$Y_{3n_3}$	$\dots$	$Y_{kn_k}$

NOTA 2. En el caso de que los tratamientos tengan efecto, las observaciones  $Y_{ij}$  de la tabla 3 se podrán describir con el modelo lineal dado por

(5.2) 
$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

donde  $\mu$  es el parámetro de escala común a todos los tratamientos, llamado *media global*,  $\tau_i$  es un parámetro que mide el efecto del tratamiento  $i$  y  $\varepsilon_{ij}$  el error aleatorio (más el error experimental) atribuible a la medición. Este tipo de modelo se llama **Modelo de Efectos Fijos**.

NOTA 3. Cuando son muchos los tratamientos que se desea comparar es conveniente comparar sólo una muestra de la población de tratamientos, de modo que  $\tau_i$  pasa a ser una variable aleatoria con su propia varianza  $\sigma_{\tau_i}^2$ , que deberá estimarse a partir de los datos. Este tipo de modelo se llama **Modelo de Efectos Aleatorios**.

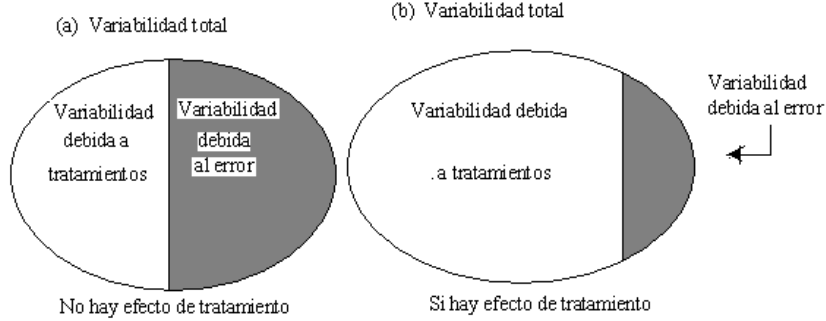


Figura 1. Variación total en un DCA

## 6. ANOVA para el diseño completamente al azar (DCA)

El nombre de análisis de varianza (ANOVA) viene del hecho de que se utilizan cocientes de varianzas para probar la hipótesis de igualdad de medias. La idea general de esta técnica es separar la variación total en dos partes: la variabilidad debida a los *tratamientos* y la debida al *error*. Cuando la primera predomina “claramente” sobre la segunda es cuando se concluye que los tratamientos tienen efecto, es decir, las medias son diferentes. Cuando los tratamientos contribuyen igual o menos que el error, se concluye que las medias son iguales (vea figura 1).

**6.1. Notación de puntos.** Sirve para representar de manera abreviada cantidades numéricas que se pueden calcular a partir de los datos en un DCA, como los de la tabla 3, en la cual  $Y_{ij}$  representa la  $j$ -ésima observación en el  $i$ -ésimo tratamiento, con  $i = 1, 2, \dots, k$  y  $j = 1, 2, \dots, n_i$ . Las cantidades de interés son las siguientes:

$Y_{i.}$  = Suma de las observaciones del  $i$ -ésimo tratamiento.

$\bar{Y}_{i.}$  = Media de las observaciones del  $i$ -ésimo tratamiento.

$Y_{..}$  = Suma o total de las  $N = n_1 + n_2 + \dots + n_k$  observaciones.

$\bar{Y}_{..}$  = Media global o promedio de todas las observaciones.

Note que el punto indica la suma sobre el correspondiente subíndice.

$$(6.1) \quad Y_{i.} = \sum_{j=1}^{n_i} Y_{ij}; \quad \bar{Y}_{i.} = \frac{\sum_{j=1}^{n_i} Y_{ij}}{n_i}; \quad Y_{..} = \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}; \quad \bar{Y}_{..} = \frac{Y_{..}}{N}$$

**6.2. ANOVA en el DCA.** El objetivo es probar las hipótesis de igualdad de respuestas medias dadas por

$$(6.2) \quad \begin{aligned} H_0 &: \mu_1 = \mu_2 = \dots = \mu_k = \mu \\ H_1 &: \mu_i \neq \mu_j \text{ para algún } i \neq j \end{aligned}$$

la cual puede escribirse en forma equivalente como

$$(6.3) \quad \begin{aligned} H_0 &: \tau_1 = \tau_2 = \dots = \tau_k = 0 \\ H_1 &: \tau_i \neq 0, \quad \text{para algún } i \end{aligned}$$

donde  $\tau_i$  es el efecto del  $i$ -ésimo tratamiento sobre la variable respuesta. Si se acepta  $H_0$  se confirma que los efectos sobre la respuesta de los  $k$  tratamientos son estadísticamente nulos y en caso de rechazar se estaría concluyendo que al menos un efecto es diferente de cero.

Para probar la hipótesis dada por las relaciones (6.2) o (6.3) mediante la técnica de ANOVA con un criterio de clasificación, lo primero es decomponer la variabilidad total de los datos en sus dos componentes: la variabilidad debida a *tratamientos* y la debida al *error aleatorio*, como se hace a continuación. Una medida de la variabilidad total presente en las observaciones de la tabla 3 es la suma total de cuadrados dada por

$$SCT = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}^2 - \frac{Y_{..}^2}{N}$$

donde  $Y_{..}$  es la suma de los  $N = \sum_{i=1}^k n_i$  datos del experimento. Sumando y restando adentro del paréntesis la media del  $i$ -ésimo tratamiento ( $\bar{Y}_{i.}$ ) y desarrollando el cuadrado se tiene:

$$\begin{aligned} SCT &= \sum_{i=1}^k \sum_{j=1}^{n_i} [(Y_{ij} - \bar{Y}_{i.}) + (\bar{Y}_{i.} - \bar{Y}_{..})]^2 \\ (6.4) \quad SCT &= \sum_{i=1}^k n_i (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2 \end{aligned}$$

donde el primer componente de 6.4 es la suma de los cuadrados de los tratamientos,

$SCTr$  ( $SCTr = \sum_{i=1}^k \frac{Y_{i.}^2}{n_i} - \frac{Y_{..}^2}{N}$ ); y el segundo es la suma de los cuadrados del error,

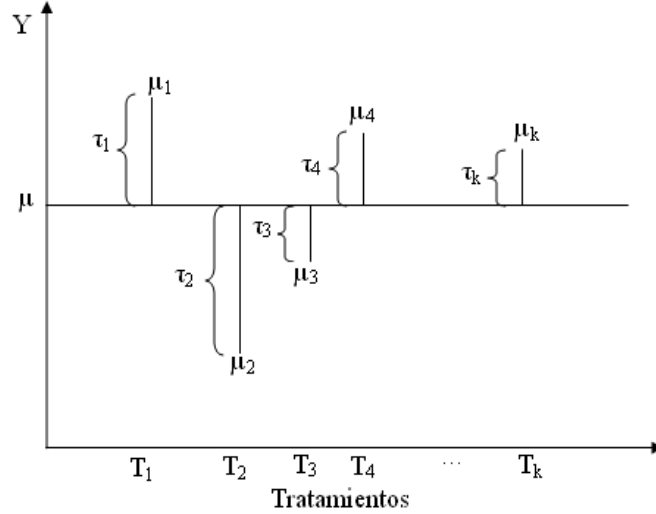
$SCE$  ( $SCE = SCT - SCTr$ ). Se puede escribir:  $SCT = SCTr + SCE$ . Como hay un total de  $N$  observaciones, la  $SCT$  tiene  $N - 1$  grados de libertad. Hay  $k$  tratamientos o niveles del factor de interés, así  $SCTr$  tiene  $k - 1$  grados de libertad, mientras que la  $SCE$  tiene  $N - k$ . Los grados de libertad respectivos a los términos de la igualdad cumplen la relación similar dada por:  $N - 1 = (k - 1) + (N - k)$ . Las sumas de cuadrados divididas entre sus respectivos grados de libertad se llaman *cuadrados medios*. El cuadrado medio de tratamientos y el cuadrado medio del error se denotan por:

$$(6.5) \quad CMT_r = \frac{SCTr}{k - 1} \quad \text{y} \quad CME = \frac{SCE}{N - k}$$

Cuando la hipótesis nula  $H_0$  es verdadera, ambos cuadrados medios son estimadores insesgados de la varianza común  $\sigma^2$  de los tratamientos que se comparan. Los valores esperados de los cuadrados medios están dados por:

$$(6.6) \quad E(CME) = \sigma^2 \quad \text{y} \quad E(CMT_r) = \sigma^2 + \frac{\sum_{i=1}^k n_i \tau_i^2}{N - k}$$

Figura 2 Representación de los efectos de tratamientos en el DCA



Se construye el estadístico de prueba como sigue: se sabe que  $SCE$  y  $SCTr$  son independientes, por lo que  $SCE/\sigma^2$  y  $SCTr/\sigma^2$  son dos variables aleatorias independientes con distribución *ji-cuadrada* con  $N - k$  y  $k - 1$  grados de libertad, respectivamente. Entonces, bajo el supuesto de que la hipótesis  $H_0$  es verdadera, el estadístico

$$(6.7) \quad F_0 = \frac{CMT_r}{CME}$$

debe seguir una distribución  $F$  con  $(k - 1)$  grados de libertad en el numerador y  $(N - k)$  grados de libertad en el denominador. De las ecuaciones (6.6) y (6.7) se deduce que si  $F_0$  es grande se rechaza la hipótesis de que no hay efectos de tratamientos, en cambio si  $F_0$  es pequeño se confirma la validez de  $H_0$ . Así, para un nivel de significancia  $\alpha$  prefijado, se rechaza  $H_0$  si  $F_0 > F_{\alpha, k-1, N-k}$ , donde  $F_{\alpha, k-1, N-k}$  es el percentíl  $(1 - \alpha) \times 100$  de la distribución  $F$ . También se rechaza  $H_0$  si  $\text{valor} - p < \alpha$ , donde el *valor - p* es el área bajo la distribución  $F_{k-1, N-k}$  a la derecha del estadístico  $F_0$ , es decir,  $\text{valor} - p = P(F > F_0)$ .

Toda la información necesaria para calcular el estadístico  $F_0$  hasta llegar al *valor - p* se escribe en la llamada tabla de análisis de varianza (ANOVA).

Tabla 4. ANOVA para el DCA					
Variabilidad	SC	g.l	CM	F valor	valor P
Entre tratamientos	$SCTr$	$k-1$	$CMT_r = \frac{SCT_r}{k-1}$	$F_0 = \frac{CMT_r}{CME}$	$P(F_{k-1, N-k} > F_0)$
Dentro tratamientos	$SCE$	$N-k$	$CME = \frac{SCE}{N-k}$		
Total	$SCT$	$N-1$			

6.2.1. *Ejemplo ilustrativo.* Un fabricante de calzado desea mejorar la calidad de las suelas, las cuales se pueden hacer con uno de los cuatro tipos de cuero  $A$ ,  $B$ ,  $C$  y  $D$  disponibles en el mercado. Para ello, prueba los cueros con una máquina que hace pasar los zapatos por una superficie abrasiva; la suela de los zapatos se desgasta

al pasarla por dicha superficie. Como criterio de desgaste se usa la pérdida de peso después de un número fijo de ciclos. Se prueban en orden aleatorio 24 zapatos, seis de cada tipo de cuero. Al hacer las pruebas en orden completamente al azar se evitan sesgos y las mediciones en un tipo de cuero resultan independiente de las demás. Los datos (en miligramos) sobre el desgaste de cada tipo de cuero se muestran en la tabla

Tipo de cuero	Desgaste						Promedio
A	264	260	258	241	262	255	256.7
B	208	220	216	200	213	206	210,5
C	220	263	219	225	230	228	230.8
D	217	226	215	224	220	222	220.7

¿Existen diferencias entre el desgaste promedio de los diferentes tipos de cuero?

La respuesta a esta pregunta es el resultado de contrastar las hipótesis

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D = \mu$$

$$H_A : \mu_i \neq \mu_j \text{ para algún } i \neq j$$

El resultado se analiza con Statgraphics

Tabla 5. ANOVA para el DCA					
Variabilidad	$SC$	g.l	$CM$	F valor	$valor - p$
Tipo de cuero	7072.33	3	2357.44	23.24	0.0000
Error	2029.0	20	101.45		
Total	9101.33	23			

La conclusión de la hipótesis se lee en la última columna de la tabla de ANOVA. Como el  $valor - p = 0.0000$  es menor que la significancia prefijada  $\alpha = 0.05$ , se rechaza  $H_0$  y se acepta que al menos un par de cueros tienen un desgaste promedio diferente.

Si al menos un cuero se desgasta diferente de otro, entonces ¿cuáles tipos de cuero son diferentes entre sí? Se realizan todas las comparaciones posibles dos a dos entre las medias de tratamiento, usando los métodos de comparaciones múltiples. También se recomienda el uso del gráfico de medias (means plot) y el diagrama de cajas (box-and-whisker plot) como se presentan a continuación en las figuras 3a y 3b

## 7. Comparaciones de rangos múltiples

Cuando se rechaza  $H_0 : \mu_1 = \mu_2 = \dots = \mu_k = \mu$  y se acepta  $H_1 : \mu_i \neq \mu_j$ , para algún  $i \neq j$ ; es necesario investigar cuales tratamientos resultaron diferentes o cuales tratamientos provocan la diferencia. La respuesta consiste en hacer la siguiente prueba

$$(7.1) \quad \begin{aligned} H_0 : \mu_i &= \mu_j \\ H_1 : \mu_i &\neq \mu_j \text{ para toda } i \neq j \end{aligned}$$

**7.1. Método LSD (diferencia mínima significativa).** Una vez que se rechazó  $H_0$  en el ANOVA, el problema es probar la igualdad de todos los posibles pares

Figura 3a Diagrama de Cajas y bigotes

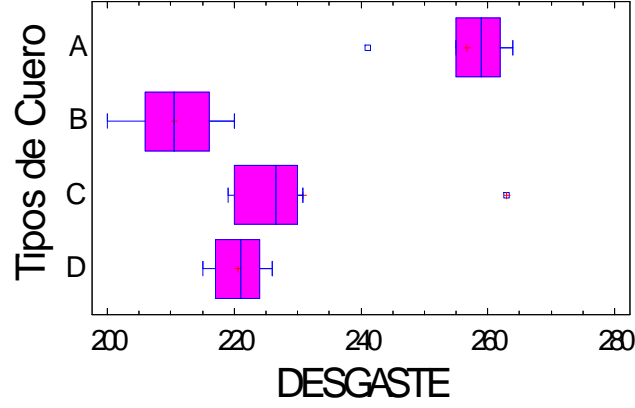
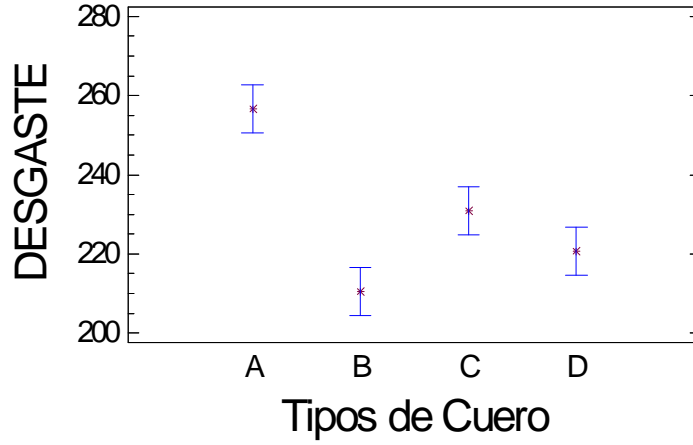


Figura 3b Diagrama de medias



de medias con la hipótesis 7.1. Para  $k$  tratamientos se tienen en total  $k(k-1)/2$  pares de medias. Se rechaza  $H_0 : \mu_i = \mu_j$  si ocurre  $|\bar{Y}_{i\cdot} - \bar{Y}_{j\cdot}| > LSD$ , donde

$$LSD = t_{\alpha/2, N-k} \cdot \sqrt{CME \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}$$

$LSD$  se llama diferencia mínima significativa (least significant difference), ya que es la diferencia mínima que debe haber entre dos medias muestrales para poder considerar que los tratamientos son significativamente diferentes. Note que si el diseño es *balanceado*, es decir, si  $n_i = n_j$  la diferencia mínima significativa se reduce

a

$$LSD = t_{\alpha/2, N-k} \cdot \sqrt{\frac{2}{n} CME}$$

Es una prueba para comparar dos medias y su uso en comparaciones simultáneas se justifica sólo en las siguientes condiciones: (a) La prueba  $F$  resulta significativa. (b) Las comparaciones fueron planeadas antes de ejecutar el experimento.

**7.2. Método de Tukey (HSD).** En este procedimiento se usa la distribución de probabilidad de rango estudentizado, que representamos con  $Q_{\alpha; m, n}$ , donde  $m$  son los grados de libertad del numerador y  $n$  los grados de libertad del denominador, el cual representamos con  $.$  Se rechaza  $H_0 : \mu_i = \mu_j$ , si ocurre  $|\bar{Y}_i - \bar{Y}_j| > Tukey$ , donde

$$Tukey = Q_{\alpha, k, N-k} \cdot \sqrt{\frac{CME}{2} \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}.$$

Cuando el diseño es *balanceado*, es decir, si  $n_i = n_j = n$

$$Tukey = Q_{\alpha, k, N-k} \cdot \sqrt{\frac{1}{n} CME}.$$

Este procedimiento es llamado también *Diferencia Significativa Honesta*, se utiliza para realizar comparaciones múltiples de medias cuando a posteriori el diseño evidencia diferencia entre los tratamientos. Esta prueba es similar a la prueba de Duncan en cuanto a su procedimiento y además es más exigente.

**7.3. Método de Duncan.** Si las  $k$  muestras son de igual tamaño, se acomodan los  $k$  promedios en orden ascendente y el error estándar de cada promedio se estima con  $\sqrt{\frac{1}{n} CME}$ . Si alguna o todas las muestras tienen tamaños diferentes se reemplaza  $n$  por la media armónica de las  $n_i$  al calcular el error estándar de los promedios. La media armónica de las  $n_i$  está dada por

$$n_{AR} = k \left( \sum_{i=1}^k \frac{1}{n_i} \right)^{-1}$$

De la tabla de rangos significantes de Duncan se obtiene el valor crítico  $r_{\alpha}(p, l)$ , donde  $p = 2, 3, \dots, k$  y  $l$  = grado de libertad para el error

$$DUNCAN = r_{\alpha}(p, l) \cdot \sqrt{\frac{1}{n} CME}$$

Se rechaza  $H_0 : \mu_i = \mu_j$  si ocurre  $|\bar{Y}_i - \bar{Y}_j| > DUNCAN$ . Este procedimiento es utilizado para realizar comparaciones múltiples de medias; para realizar esta prueba no es necesario realizar previamente la prueba  $F$  y que ésta resulte significativa; sin embargo, es recomendable efectuar esta prueba después que la prueba  $F$  haya resultado significativa, a fin de evitar contradicciones entre ambas pruebas

**OBSERVACIÓN 3.** Con referencia al ejemplo 6.2.1, el test  $LSD$  sugiere que el cuero tipo  $A$  se diferencia significativamente de los otros tipos de cuero, pero entre los cueros tipo  $B$  y  $D$  y los cueros tipo  $D$  y  $C$  no hay diferencia entre ellos si se les compara en los niveles medios. Se observa una diferencia significativa entre los cueros tipo  $B$  y  $C$ . Los contrastes  $Tukey$  y  $Duncan$  muestran las mismas conclusiones obtenidas en  $LSD$ .



Tests LSD $\alpha = 0.05$			
METODO	$n_i$	LS Mean	Grupos
CUERO B	6	210,500	X
CUERO D	6	220,667	X X
CUERO C	6	230,833	X
CUERO A	6	256,667	X

**7.4. Método de Dunnet.** En ocasiones uno de los  $k$  tratamientos a comparar es el llamado *tratamiento control* y el interés fundamental es comparar los  $k - 1$  tratamientos restantes con dicho control. En muchos casos el *tratamiento control* se refiere a la *ausencia de tratamiento*, es decir, a un grupo de especímenes de prueba a los que no se les aplica la sustancia o aditivo de los que se están comparando. Por ejemplo, al comparar varios medicamentos para el *resfriado* es conveniente que uno de los tratamientos sea el que los pacientes *no utilicen ningún medicamento*, esto sirve como referencia para decidir la posible utilidad de los medicamentos.

Por facilidad denotemos como tratamiento control al  $k$ -ésimo tratamiento. Comparar respecto al control implica probar las  $k - 1$  hipótesis dadas por

$$\begin{aligned} H_0 : \mu_i &= \mu_k \\ H_1 : \mu_i &\neq \mu_k \text{ para toda } i = 1, 2, \dots, k - 1 \end{aligned}$$

De la tabla de rangos significantes de Dunnet se obtiene el valor crítico  $D_\alpha(k - 1, l)$ , donde  $p = 2, 3, \dots, k$  y  $l = \text{g.l para el error } CME$

$$DUNNET = D_\alpha(k - 1, l) \cdot \sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right) CME}$$

La hipótesis  $H_0$  se rechaza si

$$|\bar{Y}_i - \bar{Y}_j| > DUNNET.$$

**7.5. Comparación por contrastes.** En la práctica no siempre interesa probar sólo las  $k(k - 1)/2$  hipótesis dos a dos dadas por 7.1 y no siempre estas hipótesis dos a dos interesan todas por igual. En ocasiones el objetivo del estudio lleva a contrastar hipótesis que involucran a más de dos medias.

**7.5.1. Contraste.** Una expresión de la forma  $C = \sum_{i=1}^k c_i \mu_i$  es una combinación de las medias poblacionales de interés, donde los coeficientes  $c_i$  son números reales.

La combinación lineal  $C$  se llama contraste si  $\sum_{i=1}^k c_i = 0$ .

Muchas hipótesis estadísticas de interés son *contrastos* y en particular las hipótesis de comparación de medias lo son. Por ejemplo, ya antes hemos visto que la hipótesis  $H_0 : \mu_i = \mu_j$  para toda  $i \neq j$ , se puede escribir equivalentemente como  $H_0 : \mu_i - \mu_j = 0$ , donde se observa que el contraste correspondiente es la combinación lineal  $c_i \mu_i + c_j \mu_j$  con  $c_i = 1$  y  $c_j = -1$ , e interesa verificar si es estadísticamente igual a cero.

Supongamos que interesa probar si el contraste por  $C = \sum_{i=1}^k c_i \mu_i$  es igual a cero. Si las  $k$  poblaciones objeto de estudio son normales con media  $\mu_i$  y varianza

$\sigma_i^2$ , ( $i = 1, 2, \dots, k$ ) el contraste  $C \sim N(\mu_C, \sigma_C^2)$  donde

$$\mu_C = \sum_{i=1}^k c_i \mu_i \text{ y } \sigma_C^2 = \sum_{i=1}^k \frac{c_i^2}{n_i} \sigma_i^2.$$

Cuando las varianzas de los tratamientos son iguales y el diseño experimental es balanceado ( $n_i = n$  para cada  $i$ ), la varianza del contraste se reduce

$$\sigma_C^2 = \frac{\sigma^2}{n} \sum_{i=1}^k c_i^2.$$

Usando el *CME*, para estimar a  $\sigma^2$  y  $\bar{Y}_i$  para estimar la media  $\mu_i$  se puede ver que un intervalo al  $100(1 - \alpha)\%$  de confianza para el contraste  $C$  está dado por

$$IC(C) = \sum_{i=1}^k c_i \bar{Y}_i \pm t_{\alpha/2, N-k} \cdot \sqrt{\frac{CME}{n} \cdot \sum_{i=1}^k c_i^2},$$

donde  $t_{\alpha/2, N-k}$ , es un punto porcentual de la distribución *t de Student* con  $N - k$  grados de libertad. En caso de que el intervalo contenga al cero se concluye que el contraste  $C$  es estadísticamente igual a cero.

**7.5.2. Contrastes ortogonales.** En el caso de un diseño *balanceado* dos contrastes  $C_1 = \sum_{i=1}^k c_{1i} \mu_i$  y  $C_2 = \sum_{i=1}^k c_{2i} \mu_i$  son ortogonales, si la suma del producto de los coeficientes es igual a cero, esto es, si  $\sum_{i=1}^k c_{1i} c_{2i} = 0$ ; para el diseño *desbalanceado* los contrastes son ortogonales si:

$$\sum_{i=1}^k n_i c_{1i} c_{2i} = 0.$$

**EJEMPLO 2.** En el problema de los  $k = 4$  métodos de ensamble (3) se pueden construir grupos de contrastes ortogonales de tamaño tres. Una posibilidad de elección se muestra en la siguiente tabla 6:

Tabla 6				
$c_1$	$c_2$	$c_3$	$c_4$	Contraste ortogonales
2	-1	-1	0	$2\mu_A - \mu_B - \mu_C$
0	1	-1	0	$\mu_B - \mu_C$
1	1	1	-3	$\mu_A + \mu_B + \mu_C - 3\mu_D$

Es fácil ver que los tres contrastes definidos en esta tabla son ortogonales entre sí. Por ejemplo, el primero y el segundo son ortogonales porque  $(2 \times 0) + (-1 \times 1) + (-1 \times -1) + (0 \times 0) = 0$ , y lo mismo pasa con los otros dos posibles productos. Observe también que con cada contraste se puede definir una hipótesis estadística, como se hace en el siguiente método de *Sheffé*.

**7.5.3. Método de Sheffé.** Este método está diseñado para probar todos los contrastes que pudieran interesar al experimentador, sin el inconveniente de inflar por ello el *error tipo I* (detección de diferencias que no existen). Supongamos que interesa contrastar las hipótesis

$$\begin{aligned} H_0 : 2\mu_A &= \mu_B + \mu_C \\ H_1 : 2\mu_A &\neq \mu_B + \mu_C \end{aligned}$$

donde la hipótesis nula se puede escribir alternativamente como  $H_0 : 2\mu_A - \mu_B - \mu_C = 0$ , lo que implica que la hipótesis está definida por el contraste  $C_0 = 2\mu_A - \mu_B - \mu_C$ . De manera que el contraste estimado está dado por

$$\hat{C}_0 = 2\bar{Y}_A - \bar{Y}_B - \bar{Y}_C$$

y su varianza estimada es

$$V(\hat{C}_0) = CME \sum_{i=1}^k c_i^2 / n_i$$

donde  $n_i$  es el número de mediciones en el tratamiento  $i = A, B, C$ . Los intervalos de confianza al  $100(1 - \alpha)\%$  de confianza para todos los contrastes tienen la forma

$$\hat{C} \pm \sqrt{(k-1) V(\hat{C}) F_{\alpha, k-1, N-k}}$$

$\hat{C}$  representa a cualquier posible contraste y  $F_{\alpha, k-1, N-k}$  es el percentil  $100(1 - \alpha)\%$  de una distribución  $F$  con  $k-1$  grados de libertad en el numerador y  $N-k$  grados de libertad en el denominador. Si el intervalo resultante para un contraste particular, digamos  $C_0$ , no contiene al cero, se concluye que el contraste y su hipótesis nula asociada son estadísticamente significativos. De manera equivalente, el método de *Sheffé* rechaza la hipótesis nula si el contraste asociado es

$$|\hat{C}| > \sqrt{(k-1) V(\hat{C}) F_{\alpha, k-1, N-k}}$$

## 8. Verificación de los supuestos del modelo

La validez de los resultados obtenidos en cualquier ANOVA queda supeditado a que *los supuestos del modelo se cumplan*, los cuales son: *normalidad*, *varianza constante* (igual varianza de los tratamientos) e *independencia* para la variable respuesta  $Y$ ; los cuales se traducen en supuestos sobre el *término error* ( $\varepsilon$ ) en los diferentes modelos. En la práctica, lo que se hace es tomar la muestra de *residuos* para comprobar los supuestos del modelo, ya que esta se puede ver como una muestra aleatoria de una distribución normal con *media cero* y *varianza constante*  $\sigma^2$ . Los residuos  $e_{ij} = Y_{ij} - \hat{Y}_{ij}$  se definen como la diferencia entre la respuesta observada ( $Y_{ij}$ ) y la respuesta predicha por el modelo ( $\hat{Y}_{ij}$ ), lo que permite hacer un diagnóstico más directo de la calidad del modelo, ya que su magnitud señala qué tan bien describe los datos el modelo.

**8.1. Verificación de los supuestos del modelo en el DCA.** Consideremos el modelo en el DCA

$$(8.1) \quad Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

El modelo (8.1) es un modelo teórico que se supone describirá lo que ocurra en el experimento. Cuando se realiza el ANOVA y sólo si éste resulta significativo, entonces se procede estimar el modelo ajustado o modelo de trabajo dado por

$$\hat{Y}_{ij} = \hat{\mu} + \hat{\tau}_i$$

donde  $\hat{Y}_{ij}$  es la respuesta predicha,  $\hat{\mu}$  es la media global estimada y  $\hat{\tau}_i$  es el efecto estimado del tratamiento  $i$ ; los gorros indican que son *estimadores*, es decir, valores calculados a partir de los datos del experimento. El término del error desaparece del modelo estimado por el hecho de que su valor esperado, es igual a cero ( $E(\varepsilon_{ij}) = 0$ ). Como la media global  $\mu$  se estima con  $\bar{Y}_{..}$  y el efecto del tratamiento  $\tau_i$  con  $\bar{Y}_{i.} - \bar{Y}_{..}$ , el modelo ajustado del DCA se puede escribir como

$$(8.2) \quad \hat{Y}_{ij} = \bar{Y}_{..} + (\bar{Y}_{i.} - \bar{Y}_{..}) = \bar{Y}_{i.}$$

La *respuesta predicha* para cada observación es la *media muestral del tratamiento* correspondiente. El *residual* o *residuo* asociado a la observación  $Y_{ij}$  se define como la diferencia entre  $Y_{ij}$  y el valor predicho  $\hat{Y}_{ij}$  por el modelo, es decir

$$e_{ij} = \hat{\varepsilon}_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - \bar{Y}_{i.}$$

NOTA 4. En el DCA los residuos se obtienen restando a cada valor observado la media muestral del tratamiento a que pertenece. Los  $N$  residuos  $e_{ij}$  representan una muestra aleatoria de la variable  $\varepsilon_{ij}$ . Los supuestos del modelo lineal en términos de los residuos, suponen que:

1. Los  $e_{ij}$ . siguen una distribución normal con media cero.
2. Los  $e_{ij}$  son independientes entre sí.
3. Los tratamientos tienen una varianza constante  $\sigma^2$ .

**8.2. Pruebas gráficas.** Para comprobar cada supuesto existen pruebas analíticas y gráficas. Las *pruebas gráficas* se pueden aplicar razonablemente con pocos datos, cosa que no sucede con las pruebas analíticas; estas últimas pierden drásticamente su potencia con pocos datos. El inconveniente que tienen las gráficas es que no son “exactas”, aun así, proporcionan la evidencia suficiente en contra o a favor de los supuestos. Su uso requiere fuerte evidencia visual para concluir que el supuesto en cuestión no se cumple, requiere que la evidencia en contra de un supuesto esté soportada por más de dos puntos. Cuando es uno o dos los puntos que se salen del comportamiento esperado de las gráficas se puede tratar de un problema de *puntos aberrantes*, no de violación del supuesto en cuestión. En este caso debe investigarse la obtención de dichas mediciones atípicas, ya que ese tipo de puntos pueden afectar sensiblemente los resultados del análisis.

Es mejor prevenir en lo posible que los supuestos no se violen, aplicando los tres principios básicos del diseño de experimentos: *repetición*, *aleatorización* y *bloqueo*. Es fácil encontrar situaciones en las que por no aplicar alguno de estos principios, no se cumple alguno de los supuestos del modelo.

OBSERVACIÓN 4. No aleatorizar el orden en que se corren las pruebas colleva a que no se cumpla el supuesto de independencia.

**8.3. Normalidad.** Un procedimiento para verificar el cumplimiento del supuesto de normalidad de los residuos, consiste en graficar los residuos en papel o *gráfica de probabilidad normal* que se incluye casi en todos los paquetes estadísticos. Esta gráfica tiene las escalas de tal manera que si los residuos siguen una distribución normal, al graficarlos deben tender a estar *alineados* en una *línea recta*; si claramente no se alinean se concluye que el supuesto de normalidad no se cumple.

OBSERVACIÓN 5. El ajuste de los puntos a una recta no tiene que ser perfecto, dado que el análisis de varianza resiste pequeñas y moderadas desviaciones al supuesto de normalidad.

**8.3.1. Prueba de Shapiro-Wilks para normalidad.** Consideremos una muestra aleatoria de datos  $X_1, X_2, \dots, X_n$  que proceden de cierta distribución desconocida. Se quiere verificar si dichos datos fueron generados por un proceso normal mediante las hipótesis estadísticas

$H_0$ : Los datos proceden de una distribución normal

$H_1$ : Los datos no proceden de una distribución normal

Los pasos para la prueba de Shapiro-Wilks son:

- (1) Se ordenan los datos de menor a mayor. Denotemos los datos ordenados por  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ .
- (2) De la tabla Shapiro-Wilks para este procedimiento, se obtienen los coeficientes  $a_1, a_2, \dots, a_k$ , donde  $k$  es aproximadamente  $n/2$ .
- (3) Se calcula el estadístico  $W$  definido como

$$W = \frac{1}{(n-1)S^2} \left[ \sum_{i=1}^k a_i (X_{(n-i+1)} - X_{(i)}) \right]^2$$

donde  $S^2$  es la varianza muestral.

- (4) Si el valor  $W$  del estadístico es mayor que su valor crítico al nivel  $\alpha$  seleccionado en la tabla, se rechaza la normalidad de los datos.

#### 8.4. Varianza Constante.

**8.4.1. Predichos contra los residuos.** Se puede verificar el supuesto de que los tratamientos tienen la misma varianza es graficando los *predichos* contra los *residuos* ( $\hat{Y}_{ij}$  vs.  $e_{ij}$ ), generalmente  $\hat{Y}_{ij}$  va en el eje  $X$  (horizontal) y los residuos  $e_{ij}$  en el eje vertical. Si los puntos en la gráfica de residuos contra los predichos se *distribuyen aleatoriamente* en una banda horizontal (sin ningún patrón claro y contundente), entonces es señal de que se cumple el supuesto de que los tratamientos tienen igual varianza.

**8.4.2. Factor contra residuos.** Al graficar de niveles del *factor contra residuos*, si se cumple el supuesto de varianza constante, se espera que la amplitud de la dispersión de los puntos en cada nivel de factor tenderá a ser similar; y no se cumplirá el supuesto si hay diferencias fuertes en esta amplitud.

Otra interpretación de la gráfica de *factor contra residuos* es que cuando los tratamientos o niveles muestran una diferente dispersión de sus residuales correspondientes, es que el factor o los tratamientos tienen efecto significativo sobre la variabilidad de la respuesta. Con base en esta información se podría buscar proponer un nivel de operación para dicho factor que minimice la dispersión y optimice la media.

**OBSERVACIÓN 6.** Cuando hay una evidencia contundente de que no se cumple el supuesto de varianza constante, se debe ver en qué sentido se ven afectadas las conclusiones que se obtienen con el ANOVA y las pruebas de rangos múltiples. Por ejemplo, si se aprecia que el mejor tratamiento también es el que tiene menor dispersión, entonces se debe mantener tal tratamiento como la elección correcta, y ver si es de interés investigar por qué la diferencia en variabilidad con algunos de los otros tratamientos. Pero si al que se había considerado el mejor tratamiento, es el que tiene la varianza más grande, entonces es difícil mantener a tal tratamiento como la elección correcta. En este caso se debe replantear la decisión y el análisis. Una forma de volver a hacer el análisis y reconsiderar la situación es transformar

los datos u observaciones  $Y_{ij}$  de manera que se disminuyan las diferencias en dispersión, y se pueda ver más claramente lo que ha pasado en el experimento. Existe gran cantidad de transformaciones propuestas que logran lo anterior, entre las más frecuentes se encuentra la logarítmica y la raíz cuadrada. Así se hace la transformación, se saca logaritmo a los datos u observaciones por ejemplo, y con los datos transformados se vuelve a hacer el análisis completo.

**8.4.3. Prueba de Bartlett de homogeneidad de varianzas.** Cuando se tienen  $k$  poblaciones o tratamientos independientes, cada una con distribución normal  $(N(\mu_i, \sigma_i^2), i = 1, 2, \dots, k)$ , donde las varianzas son desconocidas. Se quiere probar la hipótesis de igualdad de varianzas dada por

$$(8.3) \quad \begin{aligned} H_0 : \sigma_1^2 &= \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2 \\ H_1 : \sigma_i^2 &\neq \sigma_j^2 \text{ para algún } i \neq j \end{aligned}$$

Mediante un DCA se obtienen  $k$  muestras aleatorias de tamaños  $n_i$  ( $i = 1, 2, \dots, k$ ) de dichas poblaciones, de modo que el total de mediciones es  $N = n_1 + n_2 + \dots + n_k$ . El estadístico de prueba para la hipótesis (8.3) está dado por

$$\begin{aligned} \chi_0^2 &= 2.3026 \frac{q}{c}, \text{ donde} \\ q &= (N - k) \log_{10} S_p^2 - \sum_{i=1}^k (n_i - 1) \log_{10} S_i^2 \\ c &= 1 + \frac{1}{3(k-1)} \left( \sum_{i=1}^k (n_i - 1)^{-1} - (N - k)^{-1} \right) \\ S_p^2 &= \frac{1}{N - k} \sum_{i=1}^k (n_i - 1) S_i^2 \end{aligned}$$

donde  $S_i^2$  es la varianza muestral del tratamiento  $i$ . Bajo la hipótesis nula de igualdad de varianza, el estadístico  $\chi_0^2$  sigue una distribución ji-cuadrada con  $k - 1$  grados de libertad, por lo que se rechaza  $H_0$  cuando  $\chi_0^2 > \chi_{\alpha, k-1}^2$ . Observe que el estadístico  $q$ , en el numerador del estadístico  $\chi_0^2$ , es grande en la medida de que las varianzas muestrales  $S_i^2$  son diferentes; y es igual a cero cuando estas son iguales.

**OBSERVACIÓN 7.** *La prueba de Bartlett es sensible a la falta de normalidad de las poblaciones de interés, por lo que debe comprobarse el cumplimiento de este supuesto.*

**8.5. Independencia.** La suposición de independencia en los residuos puede verificarse si se grafica el *orden* en que se colectó un dato contra el *residuo* correspondiente, si se detecta una tendencia o patrón no aleatorio claramente definido, entonces es evidencia de que existe una correlación entre los errores y por lo tanto el supuesto de independencia no se cumple. Si el comportamiento de los puntos es aleatorio dentro de una banda horizontal, el supuesto se está cumpliendo.

La violación de este supuesto generalmente indica deficiencias en la planeación y ejecución del experimento, puede ser una indicación de que no se aplicó en forma correcta el principio de aleatorización, o simplemente que, en la medida en que el experimento se fue ejecutando fueron apareciendo factores que afectaron la respuesta observada. Por ello, en caso de tener problemas con este supuesto, las conclusiones que se pueden obtener del análisis son endebles, y por tanto es mejor revisar lo hecho y tratar de investigar por qué no se cumplió con ese supuesto de independencia, para *reconsiderar* la situación.

### 9. Elección del tamaño de la muestra

Para decidir el *número de réplicas* que se hará por cada tratamiento (tamaño de muestra), si se esperan diferencias pequeñas entre tratamientos será necesario un tamaño grande de muestra. *Estimar el tamaño muestral muchas veces requiere cierto conocimiento previo sobre la varianza del error experimental.* Si se recurre a la experiencia en los diseños se involucran *un solo factor* varía entre *cinco y diez*. Incliniéndose por un extremo de este rango e incluso salirse de ese rango, tomando en cuenta las siguientes consideraciones:

- (1) Cuando se espera poca variabilidad entre los tratamientos, mayor será la cantidad de réplicas si se quiere detectar diferencias significativas y viceversa. Si se esperan grandes diferencias, quizá con pocas réplicas sea suficiente.
- (2) Si se espera mucha variación dentro de cada tratamiento, debido a variación de fuentes no controladas como *métodos de medición, medio ambiente, materia prima, etc.*, entonces será necesario más réplicas.
- (3) Si son varios tratamientos (*cuatro o más*), entonces éste es un punto favorable para reducir el número de réplicas.

Además de lo anterior, considerar los costos y tiempo del experimento global. De aquí que si se toman en cuenta las consideraciones de arriba se puede establecer el tamaño de muestra que permita responder en una primera fase las preguntas más importantes que se quieren responder con el experimento.

**9.1. Elección del tamaño de muestra por intervalo de confianza.** Si se tiene el número de tratamientos que desea probar,  $k$ ; y que tomando las consideraciones dadas antes tiene una propuesta inicial del número de réplicas que va a utilizar,  $n_0$ . También tiene una idea aproximada del valor de  $\sigma$  (la desviación estándar del error aleatorio), y tiene una idea de la magnitud de las diferencias,  $d_T$  entre tratamientos que le interesa detectar.

La diferencia mínima significativa ( $LSD$ ) entre tratamientos está dada por la expresión

$$LSD = d_T = t_{\alpha/2, N-k} \cdot \sqrt{\frac{2}{n} CME}$$

donde

$$n = (t_{\alpha/2, N-k})^2 \cdot \frac{2}{d_T^2} CME$$

Si la significancia es  $\alpha = 0.05$ , entonces en esta fórmula se van a hacer las siguientes sustituciones:  $N = k \cdot n_0$ ,  $CME = \sigma^2$ ,  $LSD = d_T$ ; de esta forma el tamaño de muestra que tentativamente se debe usar está dado

$$n = (t_{0.025, k(n_0-1)})^2 \cdot \frac{2}{d_T^2} \sigma^2$$

El valor de  $n$  arrojado por esta fórmula dará una idea del número de réplicas por tratamiento de acuerdo a las consideraciones iniciales.

### 10. Ejercicios: experimentos con un sólo factor

**EJERCICIO 1.** *Explique en qué consiste y cuándo se deben aplicar cada uno de los siguientes diseños comparativos: diseño completamente al azar, diseño en bloques completos al azar, diseño en cuadro latino y diseño en cuadro grecolatino.*

EJERCICIO 2. Supongamos que se desea probar la igualdad entre sí de cinco medias, una alternativa para hacer esto sería comparar de dos en dos utilizando la prueba *T* de Student, por ejemplo, y al final tomar una decisión. Explique por qué esto aumenta el error tipo I.

EJERCICIO 3. ¿Qué mide el cuadrado medio del error en un experimento? ¿Qué son los grados de libertad para una suma de cuadrados en un análisis de varianza?

EJERCICIO 4. A continuación se muestra parte del ANOVA para comparar cinco tratamientos con cuatro réplicas cada uno

Variabilidad	SC	g.l	CM	F valor	valor P
Tratamiento	800				
Error	400				
Total					

Agregar en esta tabla los grados de libertad, el cuadrado medio y la razón *F* para cada una de las fuentes de variación. Explique de manera esquemática cómo calcularía el valor *P* o significancia observada para ver si hay diferencia entre tratamientos. ¿Con la información disponible se pueden hacer conjeturas sobre si hay diferencias significativas entre tratamientos? Argumente. Anote el modelo estadístico y formule la hipótesis pertinente.

EJERCICIO 5. Se desea investigar el efecto del pH en el crecimiento de cierto microorganismo en un medio específico. Para ello se realiza un experimento, teniendo como punto de partida la misma cantidad de microorganismos. Se hacen cuatro repeticiones y se obtienen los siguientes resultados. ¿Estos datos son evidencia suficiente para afirmar que los niveles de pH donde se logra menor y mayor crecimiento son el 3 y e12, respectivamente? Explique

Nivel de pH	Crecimiento promedio (en %)
1	80
2	105
3	75

EJERCICIO 6. Se desea investigar la influencia de la temperatura en el rendimiento de un proceso químico, en particular interesa investigar un rango de temperatura entre 60 a 120° C. Se tienen recursos para realizar 20 corridas experimentales. a) Los niveles de temperatura con los que experimentan son 60, 65, 70 y 120; haciendo cinco repeticiones con cada nivel. ¿Considera que es adecuado el diseño experimental usado? Argumente, y de ser necesario proponga alternativas. b) El orden en que decidieron hacer las corridas experimentales para facilitar el trabajo experimental fue primero las cinco del nivel bajo de temperatura, luego las cinco del siguiente, y así hasta finalizar. ¿Es correcto lo que hicieron? Argumente. c) Para hacer el análisis estadístico, comparan mediante una prueba *T* de Student de dos en dos niveles de temperatura, y con base en esto obtuvieron conclusiones. ¿Es adecuado tal análisis?, argumente, y en su caso proponga alternativas.

EJERCICIO 7. Describa en qué consiste cada uno de los supuestos del modelo en un análisis de varianza, y explique la forma típica en que estos supuestos se verifican.

EJERCICIO 8. ¿Qué son y cuándo se aplican las pruebas para comparar medias?



EJERCICIO 9. En una industria química se prueban diferentes mezclas para ver si difieren en cuanto al peso molecular final. Se prueban cuatro diferentes mezclas, con cinco repeticiones cada una. A continuación se muestra una parte de la tabla del análisis de varianza y los promedios obtenidos para cada mezcla

Variabilidad	valor $P$	Mezcla	Peso medio
Tratamiento	0.01	A	10000
Error		B	7000
Total		C	8000
		D	7500

- (a) ¿Las mezclas difieren de manera significativa en cuanto a su peso molecular?  
 (b) Con el análisis de varianza y de acuerdo al promedio, ¿se puede garantizar que la mezcla B es con la que se logra un menor peso molecular? Argumente.  
 (c) Si al verificar los supuestos de varianza constante (igual varianza entre las mezclas), éste no se cumple, ¿qué significa esto? ¿Se puede seguir apoyando la conclusión del inciso (a)?

EJERCICIO 10. Se hace un estudio sobre la efectividad de tres marcas de spray para matar moscas. Para ello, cada spray se aplica a un grupo de 100 moscas, y se cuenta el número de moscas muertas, expresado en porcentajes. Se hacen seis réplicas y los resultados obtenidos se muestran en seguida

	Número de réplicas					
Marca de Spray	1	2	3	4	5	6
1	72	65	67	75	62	73
2	55	59	68	70	53	50
3	64	74	61	58	51	69

- a) Formule la hipótesis adecuada y el modelo estadístico. b) ¿Existe diferencia entre la efectividad promedio de los sprays? c) ¿Hay algún spray mejor? Argumente. d) Dé un intervalo al 95%, de confianza para la efectividad promedio (porcentaje) de cada una de las marcas. e) Dibuje las gráficas de medias y los diagramas de caja simultáneos. Interpretelos. Verifique los supuestos de normalidad y de igual varianza entre las marcas.

EJERCICIO 11. En un centro de investigación se realiza un estudio para comparar varios tratamientos que al aplicarse previamente a los frijoles crudos reducen su tiempo de cocción. Estos tratamientos son a base de bicarbonato de sodio ( $\text{NaHCO}_3$ ) y cloruro de sodio o sal común ( $\text{NaCl}$ ). El primer tratamiento es el tratamiento control, que consiste en no aplicar ningún tratamiento. El tratamiento  $T_2$ , es el remojo en agua con bicarbonato de sodio, el  $T_3$ ; es en agua con sal común y el  $T_4$  es en agua con una combinación de ambos ingredientes en proporciones iguales. La variable de respuesta es el tiempo de cocción en minutos. Los datos se

muestran en la siguiente tabla:

Control	$T_2$	$T_3$	$T_4$
213	76	57	84
214	85	67	82
204	74	55	85
208	78	64	92
212	82	61	87
200	75	63	79
207	82	63	90

a) ¿De qué manera el experimentador debe aleatorizar los experimentos y el material experimental? b) Dé ejemplos de factores que deben estar fijos durante las pruebas experimentales, para que éstos no afecten los resultados y las conclusiones. c) Formule y pruebe la hipótesis de que las medias de los tratamientos son iguales. d) Obtenga el diagrama de caja y el gráfico de medias. Interprete. e) ¿Hay algún tratamiento mejor? ¿Cuál es el tiempo de cocción esperado para el mejor tratamiento? Algo importante a cuidar en un experimento es que no haya efectos colaterales no deseados que son causados por el tratamiento ganador, piense en este caso qué posibles efectos colaterales podría causar el mejor tratamiento. g) ¿Se cumplen los supuestos del modelo? Verifique gráficamente. h) Pruebe la hipótesis de igualdad de varianzas entre tratamientos (que corresponde a un supuesto).

EJERCICIO 12. Para estudiar la confiabilidad de ciertos tableros electrónicos para carros, se someten a un envejecimiento acelerado durante 100 horas a determinada temperatura, midiéndose como variable de interés la intensidad de corriente que circula entre dos puntos, cuyos valores aumentan con el deterioro. Se probaron 20 módulos repartidos equitativamente en cinco temperaturas y los resultados obtenidos son:

20°C	40°C	60°C	80°C	100°C
15	17	23	28	45
18	21	19	32	51
13	11	25	34	57
12	16	22	31	48

a) Realice el análisis de varianza para estos datos, para estudiar si la temperatura afecta la intensidad de corriente promedio. b) ¿La temperatura tiene efecto sobre la variabilidad de las intensidades? Es decir, verifique si hay igual varianza entre los diferentes tratamientos.

EJERCICIO 13. En una empresa de manufactura se propone un tratamiento para reducir el porcentaje de productos defectuosos. Para validar esta propuesta se diseñó un experimento en el que se producía con o sin la propuesta de mejora. Cada corrida experimental consistió en producir un lote y la variable de respuesta es el porcentaje de producto defectuoso. Se hicieron 25 réplicas para cada tratamiento. Los datos obtenidos se muestran a continuación:

	Porcentaje de producto defectuoso												
Con tratam	5.3	4.0	4.0	4.0	2.6	2.1	5.1	4.1	4.1	3.2	5.1	2.2	4.1
	2.2	1.1	2.0	3.0	3.1	2.1	1.2	3.3	2.1	4.0	2.0	3.0	
Sin tratam	8.0	13.2	7.2	8.2	9.1	6.7	12.2	16.3	9.2	6.4	7.2	17.2	
	8.7	11.3	4.5	6.6	9.2	10.2	10.6	13.3	5.2	6.2	8.0	4.8	

(a) ¿Las diferencias son significativas estadísticamente? (b) ¿Cuál es el porcentaje de defectos que se esperan con el nuevo tratamiento? (c) Cuantifique el nivel de reducción logrado con el tratamiento propuesto.

EJERCICIO 14. Una compañía farmacéutica desea evaluar el efecto que tiene la cantidad de almidón en la dureza de las tabletas. Se decidió producir lotes con una cantidad determinada de almidón. Se decidió que las cantidades de almidón a aprobar fueran 2%, 5% y 10%. La variable de respuesta sería el promedio de la dureza de 20 tabletas de cada lote, obteniéndose los siguientes resultados

% de almidon	Dureza			
2	4.3	5.2	4.8	4.5
5	6.5	7.3	6.9	6.1
10	9.0	7.8	8.5	8.1

a) ¿Con los datos experimentales, hay evidencia suficiente de que el porcentaje de almidón influye en la dureza en las tabletas. (b) Realice los análisis complementarios necesarios. (c) ¿Qué recomendaría al fabricante? Se desea maximizar la dureza de las tabletas. (d) Verifique los supuestos.

EJERCICIO 15. Los datos que se presentan en seguida son rendimientos en toneladas por hectárea de un pasto con tres niveles de fertilización nitrogenada. El diseño fue completamente aleatorizado, con cinc. repeticiones por tratamiento

Niveles de Nitrogeno		
1	2	3
14.8230	25.151	32.605
14.6760	25.401	32.460
14.7200	25.131	32.256
14.5141	25.031	32.669
15.0650	25.267	32.111

a) ¿Las diferencias muestrales hacen obvias la presencia de diferencias poblacionales? Comente. b) Obtenga el análisis de varianza e interprételo. c) Analice los residuos. ¿Hay algún problema?

EJERCICIO 16. Un químico del departamento de desarrollo de un laboratorio farmacéutico desea conocer cómo influye el tipo de aglutinante utilizado en tabletas de ampicilina de 500 mg en el porcentaje de friabilidad para la cual se eligen los siguientes aglutinantes: polivinilpirrolidona (PVP), carboximetilcelulosa sódica (CMC), grenetina. Los resultados del diseño experimental son los siguientes.

Aglutinante	% de friabilidad				
PVP	0.485	0.250	0.073	0.205	0.161
CMC	9.64	9.37	9.53	9.86	9.79
GNT	0.289	0.275	0.612	0.152	0.137

(a) Especifica el nombre del diseño experimental. (b) ¿Sospecha que hay efecto significativo del tipo de aglutinante sobre la variable de respuesta? (c) Escriba las hipótesis a probar y el modelo estadístico. (d) Realice el análisis adecuado para probar las hipótesis e interprete los resultados. (e) Revise los supuestos. ¿Hay algún problema?

EJERCICIO 17. En el Anova de un solo factor con  $k$  tratamientos y  $n$  observaciones por tratamientos, sea  $\mu = (1/k) \sum \mu_i$ . (a) Exprese  $E(\bar{Y}_{..})$  [sugerencia:  $\bar{X}_{..} = (1/k) \sum \bar{Y}_{i.}$ ] en términos de  $\mu$ . (b) Calcule  $E(\bar{Y}_{i.}^2)$  [Sugerencia para la variable aleatoria  $Y$ ,  $E(Y^2) = V(Y) + [E(Y)]^2$ ] (c) Calcule  $E(\bar{Y}_{..}^2)$  (d) Calcule  $E(SCTr)$  y luego muestre que  $E(MCTr) = \sigma^2 + \frac{n}{k-1} \sum (\mu_i - \mu)^2$  (e) Con el resultado del inciso (d) ¿cuál es  $E(MCTr)$  cuando  $H_0$  es verdadera? Cuando  $H_0$  es falsa, ¿cómo se compara  $E(MCTr)$  con  $\sigma^2$ ?

EJERCICIO 18. En un estudio sobre el efecto conjunto del ozono y el dióxido de azufre como agentes contaminantes del aire se dispuso cierto tipo de semillas de habichuelas en cámaras que se mantuvieron a lo largo de toda la experiencia abiertas siempre por su parte superior. Utilizando un diseño completamente aleatorizado se asignaron seis de estas cámaras a un primer tratamiento que consistió en fumigarlas repetidamente con dióxido de azufre, mientras que en las otras seis restantes no se efectuó ningún tratamiento. Posteriormente, se consiguió un ambiente de ozono a través de un proceso de filtrado del carbono en tres de las cámaras que habían sido fumigadas con dióxido de azufre y también en tres de las cámaras del grupo que no había sido fumigado. Las tres cámaras en ambos grupos fueron seleccionadas también mediante un diseño completamente aleatorizado. Transcurrido un mes, se registró la cosecha total (en kg.) de habichuelas en cada cámara, obteniéndose los siguientes resultados:

Ozono ausente		Ozono presente	
Dióxido ausente	Dióxido presente	Dióxido ausente	Dióxido presente
1.52	1.49	1.15	0.65
1.85	1.55	1.30	0.76
1.39	1.21	1.57	0.69

a) A la vista de estos resultados contrastar si existe un efecto conjunto del ozono y el dióxido de azufre. Utilizar  $\alpha = 0.05$ . b) A la vista del resultado obtenido, contrasta las posibles diferencias entre todos los grupos utilizando el método de Bonferroni y de Tukey. Construye los subgrupos homogéneos y comenta los resultados obtenidos. c) Construir y resolver tres contrastes de medias (ortogonales entre si) que aporten la misma información que aporta el contraste global de anova. Comprobar la descomposición de la suma de cuadrados entre grupos global en la suma de las sumas de cuadrados de cada uno de los contrastes. d) Como la validez del test F exige el cumplimiento de ciertas hipótesis, valora si en este problema se dan dichas condiciones. Caso de que no se cumplieran propon una transformación adecuada y valora entonces si existe el efecto que antes has estudiado.

EJERCICIO 19. Un equipo de fisiólogos investiga las consecuencias de una inundación en el metabolismo radical de dos especies de arboles: el abedul de los ríos que es tolerante a las inundaciones y el abedul europeo que no es tolerante a las mismas. Se seleccionaron ocho plántulas de cada especie, cuatro de ellas fueron anegadas durante todo un día y las otras cuatro fueron utilizadas como controles. Posteriormente se midió la concentración de adenosina trifosfato (ATP) en las raíces de cada planta. Los datos (nmol de ATP por mg. de tejido) se presentan en

la siguiente tabla:

Abedul de los ríos		Abedul Europeo	
Anegada	Control	Anegada	Control
1.45	1.70	0.21	1.34
1.19	2.04	0.58	0.99
1.05	1.49	0.11	1.17
1.07	1.91	0.27	1.30

Si para estos datos,  $SC(\text{entre}) = 4.5530$  y  $SC(\text{dentro}) = 0.47438$ : a) Completar la tabla ANOVA y efectuar el test de la  $F$  con  $\alpha = 0.05$ . b) Suponiendo que cada una de las cuatro poblaciones consideradas tiene la misma desviación típica, utilizar los datos para calcular un estimador de dicha desviación típica.

EJERCICIO 20. Los siguientes datos provienen de un experimento realizado en la estación experimental de Rothamsted. El objetivo era medir la eficacia de tres insecticidas, el clorodinitrobenzono (CN), el carbón disulfido (CD) y un preparado propio denominado cymag (CM). Cada insecticida se aplicó a dosis normal (1) y doble (2). Por último se contó con un grupo control al que no se aplicó ningún insecticida. Los pesticidas se aplicaron antes de la siembra del trigo, y los datos recogidos muestran el incremento del número de gusanos encontrados en cada parcela después de la recolección del trigo.

Control	Insecticida					
	1CN	1CD	1CM	2CN	2CD	2CM
466	222	194	306	92	166	28
421	219	221	176	114	172	179
561	332	308	215	80	111	165
433	298	256	199	128	80	82

Para estos datos se obtiene la siguiente tabla ANOVA:

Fuentes de variación	SC	gl	CM	F	p-valor
Insecticida	392447	6	65408	22.473	4.007e-08
Residual	61121	21	2911		

a) Contrasta la posible igualdad de medias en el número de gusanos que proliferaron con los diferentes tratamientos, planteando el contraste de hipótesis adecuado. b) A la vista del resultado obtenido, contrasta las posibles diferencias entre todos los grupos utilizando el método de Bonferroni y de Tukey. Construye los subgrupos homogéneos y comenta los resultados obtenidos. c) Completa tus conclusiones considerando las siguientes comparaciones: - Una comparación entre el grupo control y los insecticidas. - Una comparación entre los insecticidas a dosis 1 y los insecticidas a dosis 2. - Una comparación entre los insecticidas a dosis 1. - Una comparación entre los insecticidas a dosis 2. Para cada una de estas comparaciones: Indica qué grados de libertad corresponden a cada comparación. Indica cuáles de ellas son contrastes. Descompón en contrastes ortogonales las comparaciones con más de 1 grado de libertad. ¿Qué parejas de contrastes resultan ser ortogonales entre sí? Resuelve los contrastes y comenta los resultados. d) Como la validez del test  $F$  exige el cumplimiento de ciertas hipótesis, valora si en este problema se dan dichas condiciones. Caso de que no se cumplieran propón una transformación adecuada y valora entonces si existe el efecto que antes has estudiado.

EJERCICIO 21. Para controlar el posible impacto medioambiental que supondrá el incendio de varias fábricas de tejidos próximas a un bosque, se determinó el tiempo (en segundos) que tardaban en arder 5 vestidos, elegidos al azar, realizados en cada una de ellas. Del análisis de los datos se obtuvieron los siguientes resultados:

Fábrica	1	2	3	4	5
Media	16.78	11.76	10.24	11.98	15.26
Desv. Típica	1.167	2.3298	1.1437	1.862	0.9182

a) Si medimos la peligrosidad de una fábrica por el tiempo que tardan en arder sus vestidos, ¿hay evidencia para pensar que los tejidos de las fábricas influyen en su peligrosidad? b) Calcula los subgrupos homogéneos resultantes de aplicar la corrección de Bonferroni y el método de Tukey. ¿Qué conclusiones puedes extraer de estos grupos homogéneos? ¿Cuáles son tus conclusiones globales sobre el análisis realizado?

EJERCICIO 22. Los siguientes resultados proceden de un estudio sobre la posible reacción de los áfidos sanos de una colonia ante la presencia de un áfido con un proceso infeccioso causado por hongos. Se han considerado los siguientes estadios de la infección: una primera etapa en la que el áfido no ha sido infectado aún, la etapa siguiente en la que comienza la infección, la etapa de mayor desarrollo de la infección que acaba con la muerte del áfido y, finalmente una etapa post-mortem. La siguiente tabla recoge la media, desviación típica y tamaño muestral del porcentaje de áfidos sanos de la colonia que mostraron algún tipo de respuesta al áfido infectado.

	Preinfección	Inicio infección	Infección desarrollada	Post-mortem
Media	78.38	9.82	5.22	2.42
Desv. Típica	10.68	9.66	7.35	3.11
Tamaño muestral	6	6	6	6

a) Propón de forma justificada un diseño que valore la respuesta de los áfidos sanos al individuo enfermo según los diferentes estadios del proceso infeccioso. ¿Se puede afirmar que hay diferencias en el comportamiento de los áfidos sanos en los diferentes estadios de la infección? b) ¿Crees que los áfidos sanos se comportan de forma diferente antes y después de la infección? ¿Encuentras diferencias significativas en el comportamiento de los áfidos sanos en las tres etapas de la etapa infecciosa?

EJERCICIO 23. Se sabe que los cucos utilizan los nidos de otras especies para depositar sus huevos. Cada hembra vigila un vasto territorio en busca de nidos en construcción y observa el comienzo de la puesta ajena. Aprovecha el momento en que los propietarios están ausentes para poner en el nido elegido uno de sus huevos, hurtando del mismo uno de los huevos originales para restablecer su número. De acuerdo con una teoría los cucos imitan en tamaño los huevos de las especies elegidas como hospederas lo que ha dado lugar a distintas subespecies de cucos dependiendo de la especie elegida para ser parasitada (y por tanto del huevo del cuco). Con el objetivo de valorar si de verdad se puede considerar la existencia de diferentes subespecies, se ha realizado un estudio de las longitudes de los huevos de cucos

encontrados en los nidos de tres especies. Los resultados son los siguientes:

<i>Nidos</i>	<i>Carriceros</i>	<i>Petirrojos</i>	<i>Lavanderas</i>
<i>Longitud Media en mm</i>	<i>23.09</i>	<i>23.12</i>	<i>24.14</i>
<i>Desviación Típica en mm</i>	<i>0.90</i>	<i>1.07</i>	<i>0.57</i>
<i>Tamaño Muestral</i>	<i>15</i>	<i>15</i>	<i>15</i>

a) Describe con todo detalle el diseño del experimento y valora la existencia de subespecies de cucos distintas, de acuerdo con la teoría anterior. Utiliza 0.01 como nivel de significatividad. b) Realiza un análisis que te permita clasificar los cucos en una, dos o tres subespecies diferentes. Utiliza también 0.01 como nivel de significatividad.





## CHAPTER 3

# Diseños en bloques

### 1. Competencias

Identificar las características generales y los usos que se le dan a los diseños en bloques.

Explicar la definición del diseño en bloques completos al azar; así como su hipótesis, modelo estadístico y análisis de varianza.

Describir la selección y la aleatorización del diseño en cuadro latino y su diferencia con el diseño en cuadro grecolatino.

### 2. Diseños en bloques completos al azar

En un diseño en bloques completos al azar (DBCA) se consideran tres fuentes de variabilidad: *el factor de tratamientos, el factor de bloques y el error aleatorio*, es decir, se tienen tres posibles "culpables" de la variabilidad presente en los datos. La palabra completo en el nombre del diseño se debe a que en cada bloque se prueban todos los tratamientos, es decir, que los bloques están completos.

La aleatorización se hace dentro de cada bloque; no se realiza de manera total como en el diseño completamente al azar. El hecho de que existan bloques hace que no sea práctico a que incluso sea imposible aleatorizar en su totalidad.

Los factores de bloqueo que aparecen en la práctica son: *turno, lote, día, tipo de material, línea de producción, operador, máquina, método, etc.* La imposibilidad de aleatorizar de bloque a bloque se aprecia claramente cuando se bloquean factores como día o turno, ya que no tiene sentido pensar en seleccionar al azar el orden de los días o los turnos, porque es imposible regresar el tiempo.

Supongamos una situación experimental con  $k$  tratamientos y  $b$  bloques. El aspecto de los datos para este caso se muestra en la siguiente tabla. Considerando una repetición en cada combinación de tratamiento y bloque.

Bloques	Tratamientos				
	1	2	3	...	$k$
1	$Y_{11}$	$Y_{21}$	$Y_{31}$	...	$Y_{k1}$
2	$Y_{12}$	$Y_{22}$	$Y_{32}$	...	$Y_{k2}$
3	$Y_{13}$	$Y_{23}$	$Y_{33}$	...	$Y_{k3}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$b$	$Y_{1b}$	$Y_{2b}$	$Y_{3b}$	...	$Y_{kb}$

**2.1. Modelo estadístico.** Cuando se decide utilizar un DBCA, el experimentador piensa que cada medición será el resultado del efecto del tratamiento donde se encuentre, del efecto del bloque a que pertenece y de cierto error que se espera sea aleatorio. Si actuara otro tipo de efecto o factor adicional a estos tres, dicho efecto se carga en el error, el cual deja de ser aleatorio, y como consecuencia puede

enmascarar el efecto del factor de interés al realizar el ANOVA. De aquí la importancia de aplicar el principio de bloqueo, evitando que otro factor no contemplado en el experimento afecte los datos en el experimento. El modelo estadístico para este diseño está dado por

$$Y_{ij} = \mu + \tau_i + \gamma_j + \varepsilon_{ij}; \begin{cases} i = 1, 2, \dots, k \\ j = 1, 2, \dots, b \end{cases}$$

donde  $Y_{ij}$  es la medición que corresponde al tratamiento  $i$  y al bloque  $j$ ;  $\mu$  es la media global poblacional;  $\tau_i$  es el efecto debido al tratamiento  $i$ , y  $\gamma_j$  es el efecto debido al bloque  $j$ , y  $\varepsilon_{ij}$  es el error aleatorio atribuible a la medición  $Y_{ij}$ . Se supone que los errores se distribuyen de manera normal con media cero y varianza constante  $\sigma^2$ , y que son independientes.

**2.2. Hipótesis a probar.** Como ya se ha mencionado, la hipótesis de interés es la misma para todos los diseños comparativos, y está dada por

$$(2.1) \quad \begin{aligned} H_0 : \mu_1 = \mu_2 = \dots = \mu_k = \mu \\ H_1 : \mu_i \neq \mu_j \text{ para algún } i \neq j \end{aligned}$$

que tambien puede expresarse

$$(2.2) \quad \begin{aligned} H_0 : \tau_1 = \tau_2 = \dots = \tau_k = 0 \\ H_1 : \tau_i \neq 0 \text{ para algún } i \end{aligned}$$

En cualquiera de estas hipótesis la afirmación a probar es que la respuesta media poblacional lograda con cada tratamiento es la misma para los  $k$  tratamientos, y que por tanto cada respuesta media  $\mu_i$  es igual a la respuesta media global poblacional,  $\mu$ . De manera alternativa, se puede afirmar que los efectos de tratamientos sobre la variable de respuesta son nulos, porque cuando el efecto  $\tau_i = \mu_i - \mu = 0$ , entonces necesariamente la respuesta media del tratamiento es igual a la media global ( $\mu_i = \mu$ ).

**2.3. Análisis de varianza para DBCA.** La hipótesis dada por (2.1 o 2.2) se prueba con un análisis de varianza con dos criterios de clasificación: el factor de tratamientos y el factor de bloque

Tabla 1 Tabla ANOVA para el DBCA					
Fuentes de variabilidad	SC	GL	CM	F <sub>0</sub>	valor P
Tratamientos	SCTr	k-1	CMTTr	$\frac{CMTTr}{CME}$	$P(F > F_0)$
Bloques	SCB	b-1	CMB	$\frac{CMB}{CME}$	$P(F > F_0)$
Error	SCE	(k-1)(b-1)	CME		
Total	SCT	N-1			

Los cálculos necesarios pueden ser manuales, pero siempre práctico hacerlos con un software estadístico, porque además proporciona muchas otras opciones gráficas y

tabulares útiles, no sólo el ANOVA. Aquí

$$SCT = \sum_{j=1}^b \sum_{i=1}^k Y_{ij}^2 - \frac{Y_{..}^2}{N}$$

$$SCTr = \sum_{i=1}^k \frac{Y_{i.}^2}{b} - \frac{Y_{..}^2}{N}$$

$$SCB = \sum_{j=1}^b \frac{Y_{.j}^2}{k} - \frac{Y_{..}^2}{N}$$

y la del error se obtiene por sustracción como  $SCE = SCT - SCTr - SCB$

**EJEMPLO 3. (Comparación de cuatro métodos de ensamble)** Un equipo de mejora investiga el efecto de cuatro métodos de ensamble  $A$ ,  $B$ ,  $C$  y  $D$ , sobre el tiempo de ensamble en minutos. En primera instancia la estrategia experimental es aplicar cuatro veces los cuatro métodos de ensamble en orden completamente aleatorio (las 16 pruebas en orden aleatoria). Los tiempos de ensamble obtenidos se muestran en la tabla 2 siguiente. Si se usa este diseño llamado diseño completamente al azar (DCA), se supone que de acuerdo a como se ha hecho el estudio, además del método de ensamble, no existe ningún otro factor que pueda influir de manera significativa sobre la variable de respuesta (tiempo de ensamble). Si los experimentadores se dan cuenta que hay cuatro operadores y consideran que esto puede afectar de manera significativa los tiempos de ensamble, y por ende la comparación de los métodos, entonces debe utilizar el diseño en bloques completos al azar (DBCA). Los datos para este diseño se pueden ver en la tabla 3 que aparece a continuación

Tabla 2 DCA					Tabla 3 DBCA				
Metodo de ensamble				,	Metodo				Operador
A	B	C	D		A	B	C	D	
6	7	11	10		1	6	7	10	10
8	9	16	12		2	9	10	16	13
7	10	11	11		3	7	11	11	11
8	8	13	9		4	8	8	14	9

Para comparar los cuatro métodos se plantea la hipótesis

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D = \mu$$

$$H_1 : \mu_i \neq \mu_j \text{ para algún } i \neq j = A, B, C, D$$

la cual se prueba mediante el análisis de varianza dado en la tabla 5. De esta tabla se observa que para los métodos se obtuvo un *valor*  $P = 0.0029 < \alpha = 0.05$ , por lo que se rechaza la hipótesis  $H_0$ , de que el tiempo medio poblacional de los métodos de ensamble son iguales, y se acepta que al menos dos de los métodos son diferentes en cuanto al tiempo promedio que requieren. De la misma manera para operadores, como *valor*  $P = 0.0298 < \alpha = 0.05$  el factor de bloques (operadores) también afecta, es decir, existen diferencias entre los operadores en cuanto al tiempo promedio. Sin embargo, recordemos que no es objetivo del experimento comparar a los operadores, y su control en el estudio se utiliza para lograr una comparación

más justa y precisa de los métodos de ensamble.

Tabla 4. ANOVA para Diseño completamente al azar (DCA)					
Fuentes de variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Métodos de ensamble	69,5	3	23,1667	9,42	0,0018
Error	29,5	12	2,45833		
Total	99,0	15			

Tabla 5 Tabla ANOVA para el DBCA					
Fuentes de variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Métodos de ensamble	61,5	3	20.5	10.25	0,0029
Bloques	28,5	3	9.5	4.75	0.0298
Error	18,0	9	2.0		
Total	108,0	15			

#### 2.4. Comparación de parejas de medias de tratamiento en el DBCA.

Como se rechaza la hipótesis de igualdad de los cuatro tratamientos o métodos de ensamble, es natural preguntarse cuáles de ellos son diferentes entre sí. Para averiguarlo se utiliza alguna de las pruebas que se vieron en la sección *Comparaciones o pruebas de rango múltiples*. Con estas fórmulas se sustituye el número de réplicas por el número de bloques, y también se debe cambiar los grados de libertad del error, que en el caso de bloques está dado por  $(k-1)(b-1)$ . Por ejemplo, recordemos que la diferencia mínima significativa (LSD) para dos tratamientos,  $i$  y  $l$ , en un DCA está dada por

$$LSD = t_{\alpha/2, N-k} \cdot \sqrt{\frac{2}{n} CME}$$

Entonces, en bloque esta expresión se transforma en

$$LSD = t_{\alpha/2, (k-1)(b-1)} \cdot \sqrt{\frac{2}{b} CME}$$

donde  $b$  es el número de bloques. De aquí que en el ejemplo 3  $t_{0.025;9} = 2.26$ , entonces  $LSD = 2.26 \cdot \sqrt{\frac{2}{4} * 2} = 2.26$

Tabla 4 Test LSD			
Método	$n_i$	LS Mean	Grupo
A	4	7,50	X
B	4	9,00	X X
C	4	10,75	X X
D	4	12,75	X

se concluye que el tratamiento  $A$  es diferente de  $C$  y  $D$ , y que el tratamiento  $B$  es diferente de  $D$ . Las otras tres comparaciones ( $A$  con  $B$ ,  $B$  con  $C$  con  $D$ ) aceptan la hipótesis de igualdad. De acuerdo con esto, y dadas las respuestas medias muestrales  $\bar{Y}_A = 7.5$ ,  $\bar{Y}_B = 9.0$ ,  $\bar{Y}_C = 10.75$   $\bar{Y}_D = 12.75$ , se concluye que el método  $A$  es mejor que los métodos  $C$  y  $D$ , pero el método  $A$  no es mejor que el  $B$ .

**2.5. Efecto de bloque .** La tabla de ANOVA también proporciona una prueba para el efecto de los bloques. En el segundo renglón de la tabla se verifica la hipótesis

$$H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_b = 0$$

$$H_1 : \gamma_j \neq 0 \text{ para algún } j$$

que en caso de resultar significativa implica que el factor de bloques tiene influencia sobre la variable de respuesta, y debe ser tomado en cuenta para mejorar la calidad de ésta; si no se rechaza, y se acepta que los bloques son iguales en respuesta media, entonces se tiene el argumento a favor de no controlar este factor en futuros experimentos sobre esta misma respuesta.

Si fuera de interés el estudio del factor de bloque al mismo nivel del factor de tratamientos, entonces se debería correr el experimento aleatorizando completamente el orden de todas las combinaciones posibles entre bloques y tratamientos. Si ése fuera el caso y suponiendo que sea posible aleatorizar totalmente, el resultado sería un diseño factorial  $k \times b$ , que se presentara más adelante. Otro supuesto del diseño de bloques al azar es que no existe efecto de interacción entre el factor de bloque y el factor de tratamientos.

### 3. Diseño en cuadro latino (DCL)

Diseño en el que se controlan dos factores de bloque y uno de tratamientos con la misma cantidad de niveles. Los tratamientos se representan por letras latinas y se distribuyen en forma adecuada en un cuadro, por tal razón se llama Diseño en cuadro latino (DCL). En este diseño se tienen cuatro fuentes de variabilidad que pueden afectar la respuesta observada: *los tratamientos, el factor de bloque I* (columnas), *el factor de bloque II* (renglones) y *el error aleatorio*. Sean  $A, B, C, \dots, K$ , los  $k$  tratamientos a comparar, por tanto ambos factores de bloques tienen también  $k$  bloques cada uno. El aspecto de los datos se muestra en la tabla 5

Tabla 5 Tabla ANOVA para el DCL					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Tratamientos	SCTr	k-1	CMTr	$\frac{CMT_r}{CME}$	$P(F > F_0)$
Bloques	SCB1	k-1	CMB1	$\frac{CMB_1}{CME}$	$P(F > F_0)$
Columnas	SCB2	k-1	CMB2	$\frac{CMB_2}{CME}$	$P(F > F_0)$
Error	SCE	(k-2)(k-1)	CME		
Total	SCT	k <sup>2</sup> -1			

**3.1. Análisis del diseño.** El modelo estadístico que se supone describe el comportamiento de las observaciones está dado por

$$Y_{ijl} = \mu + \tau_i + \gamma_j + \delta_l + \varepsilon_{ijl}$$

donde  $Y_{ijl}$  es la observación del tratamiento  $i$ , en el nivel  $j$  del factor columna y en el nivel  $l$  del factor renglón;  $\varepsilon_{ijl}$ , es el error atribuible a dicha observación. De acuerdo con este modelo, la variabilidad total presente en los datos se puede descomponer como

$$SCT = SCTr + SCB1 + SCB2 + SCE$$

y los grados de libertad correspondientes son

$$k^2 - 1 = (k - 1) + (k - 1) + (k - 1) + (k - 2)(k - 1)$$

El ANOVA para el diseño en cuadro latino se muestra en la tabla 5. En él se prueba la hipótesis sobre los efectos de tratamiento, del factor renglón y del factor columna. Otra vez, la hipótesis fundamental es la de los tratamientos; las otras dos proporcionan un adicional al objetivo inicial y permiten comprobar la relevancia de controlar los factores de bloque

3.1.1. *Ejemplo de un DCL: Comparación de cuatro marcas de llantas.*

*Una compañía de mensajería está interesada en determinar cuál marca de llantas tiene mayor duración, medida ésta en términos del desgaste. Para ello, planea un experimento en cuadro latino en el que se comparan las cuatro marcas de llantas sometiénolas a una prueba de 32000 kilómetros de recorrido, utilizando cuatro diferentes tipos de auto y las cuatro posiciones posibles de las llantas en el auto.*

El factor de interés: tipo de llanta o marca; los factores de bloques: el tipo de carro y la posición de la llanta en el carro, porque por experiencia se sabe que el tipo de carro y la posición de la llanta tienen efecto en el desgaste de la misma.

El diseño utilizado es el cuadro latino de la tabla 6 que permite estudiar los tres factores, cada uno en cuatro niveles. La elección del cuadro latino se hace antes de obtener los datos, para ello, a partir de un cuadro latino inicial, se aleatorizan las columnas y los renglones; y después las diferentes marcas de llanta se asignan de manera aleatoria a las letras latinas que denotan los niveles del factor de interés. Las pruebas se hacen al mismo tiempo con choferes a quienes se les instruye para que manejen de manera similar, sobre el mismo terreno para los cuatro automóviles. Al hacer las pruebas de los cuatro autos al mismo tiempo se evita el efecto del ambiente en el desgaste; asimismo, el conductor y el tipo de terreno podrían influir, pero se considera suficiente mantenerlos lo más homogéneo posible durante el experimento. Los datos observados se muestran en la tabla 7. Se mide la diferencia máxima entre el grosor de la llanta nueva y el grosor de la llanta después de haber recorrido los 32000 kilómetros. Obviamente a mayor diferencia en grosor, mayor desgaste. Las unidades de medición son milésimas de pulgada.

Tabla 6 Diseño en cuadro latino (DCL)						
		Bloques II (columnas)				
		1	2	3	...	$k$
Bloques I (renglones)	1	$A=Y_{111}$	$B=Y_{221}$	$C=Y_{331}$	...	$K=Y_{kk1}$
	2	$B=Y_{212}$	$C=Y_{322}$	$D=Y_{432}$	...	$A=Y_{k2}$
	3	$C=Y_{313}$	$D=Y_{423}$	$E=Y_{523}$	...	$B=Y_{k3}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$k$	$K=Y_{k1k}$	$A=Y_{12k}$	$B=Y_{23k}$	...	$J=Y_{jkk}$

Tabla 7. DCL en la comparación de llantas					
		Carro (columnas)			
		1	2	3	4
Posición (renglones)	1	$C = 12$	$D = 11$	$A = 13$	$B = 8$
	2	$B = 14$	$C = 12$	$D = 11$	$A = 3$
	3	$A = 17$	$B = 14$	$C = 10$	$D = 9$
	4	$D = 13$	$A = 14$	$B = 13$	$C = 9$

El ANOVA resultante se muestra en la tabla 8, se observa que existen diferencias entre las marcas de llanta y entre los tipos de carro, a un nivel de significancia de  $\alpha = 0.05$ ; no hay evidencia suficiente para concluir que la posición tiene un efecto

importante, puesto que su correspondiente *valor p* es mayor que 0.05.

Tabla 8. ANOVA para el DCL					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Marca	30.6875	3	10.2292	11.42	0.0068
Posicion	6.1875	3	2.0625	2.30	0.1769
Carro	38.6875	3	12.8958	14.40	0.0038
Error	5.375	6	0.895833		
Total	80.9375	15			

**Interpretación.** Para investigar cuáles marcas de llantas son diferentes entre sí, se aplica la prueba LSD y se obtienen los resultados de la siguiente tabla:

Tabla 9 Test LSD			
Método	$n_i$	LS Mean	Grupos
C	4	10.75	X
D	4	11,00	X
B	4	11,25	X
A	4	14,25	X

Las conclusiones sobre los seis pares de marcas se leen en la columna de grupos homogéneos, donde se observa que la marca *A* no tiene *X* en la primera columna y es la única con *X* en la segunda columna, lo que indica que es diferente al resto de las marcas; y, considerando que mientras la diferencia máxima en grosor sea mayor, la llanta se desgasta más, se concluye que la marca *A* sufre mayor desgaste que las otras tres, por lo que es la peor llanta. Entre las tres marcas restantes (*C*, *D* y *B*) no se encontró una diferencia significativa en cuanto al desgaste medio, se concluye que desde el punto de vista estadístico y a la luz de los resultados experimentales, estas tres marcas de llantas pueden considerarse iguales. Esto no quiere decir que sean idénticas, sino que sus diferencias son menores, y no se alcanzan a detectar en el análisis del experimento.

### ANOVA para el DCL

Tabla 10 ANOVA DCL			
variabilidad	SC	GL	F
Tratamientos	$SCTr = \sum_{i=1}^k \frac{Y_{i..}^2}{k} - \frac{Y_{...}^2}{N}$	$k - 1$	$\frac{CMT_r}{CME}$
Bloque I	$SCB1 = \sum_{j=1}^k \frac{Y_{.j.}^2}{k} - \frac{Y_{...}^2}{N}$	$k - 1$	$\frac{CMB1}{CME}$
Bloque II	$SCB2 = \sum_{l=1}^k \frac{Y_{..l}^2}{k} - \frac{Y_{...}^2}{N}$	$k - 1$	$\frac{CMB2}{CME}$
Error	$SCE = SCT - SCTr - SCB1 - SCB2$	$(k - 2)(k - 1)$	
Total	$SCT = \sum_{i=1}^k \sum_{j=1}^k \sum_{l=1}^k Y_{ijl}^2 - \frac{Y_{...}^2}{N}$	$k^2 - 1$	

**3.2. Selección y aleatorización de un cuadro latino.** No cualquier arreglo de letras latinas en forma de cuadro es cuadro latino, la regla fundamental es que cada letra debe aparecer sólo una vez en cada renglón y en cada columna. Un

cuadro latino estándar es aquel en el que en la primera columna y en el primer renglón aparecen las letras en orden alfabético, acotinuación se presenta un cuadro latino estándar de tamaño cuatro

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>B</i>	<i>C</i>	<i>D</i>	<i>A</i>
<i>C</i>	<i>D</i>	<i>A</i>	<i>B</i>
<i>D</i>	<i>A</i>	<i>B</i>	<i>C</i>

existen además los siguientes tres cuadros latinos de dimensión cuatro:

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>B</i>	<i>A</i>	<i>D</i>	<i>C</i>
<i>C</i>	<i>D</i>	<i>B</i>	<i>A</i>
<i>D</i>	<i>C</i>	<i>A</i>	<i>B</i>

, 

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>B</i>	<i>D</i>	<i>A</i>	<i>C</i>
<i>C</i>	<i>A</i>	<i>D</i>	<i>B</i>
<i>D</i>	<i>C</i>	<i>B</i>	<i>A</i>

, 

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>B</i>	<i>A</i>	<i>D</i>	<i>C</i>
<i>C</i>	<i>D</i>	<i>A</i>	<i>B</i>
<i>D</i>	<i>C</i>	<i>B</i>	<i>A</i>

Para cuatro tratamientos se pueden construir un total de 576 cuadros latinos de los cuales cuatro son estándar. La selección del diseño debería legir uno al azar de los 576 posibles, pero es prácticamente imposible construirlos a todos para seleccionar uno al azar. Sin embargo, ocurre que dado un cuadro latino, cualquier intercambio de columnas o de renglones es también cuadro latino, por eso la estrategia de selección y aleatorización recomendada en la práctica es la siguiente:

- (1) Se construye el cuadro latino estándar más sencillo.
- (2) Se aleatoriza el orden de los renglones (o columnas) y posteriormente se aleatoriza el orden de las columnas (o renglones).
- (3) Por último, los tratamientos a comparar se asignan en forma aleatoria a las letras latinas.

El cuadro latino tiene dos restricciones a la aleatorización debido a los dos factores de bloque, lo que implica que a la hora de correr el experimento no hay ningún margen de aleatorización. Es decir, se puede correr por columna o por renglón segun convenga. Lo que no es correcto es hacer todas las pruebas de un tratamiento, y luego todas las de otro, y así sucesivamente, puesto que se puede introducir ruido adicional debido a factores no controlables que cambian con el tiempo.

#### 4. Diseño en cuadro grecolatino (DCGL)

Con el diseño en cuadro grecolatino (DCGL) se controlan tres factores de bloque además del factor de tratamientos. Se llama cuadro grecolatino porque los cuatro factores involucrados se prueban en la misma cantidad de niveles, de aquí que se pueda escribir como un cuadro (véase tabla 10); y se utilizan letras latinas para denotar a los tratamientos y letras griegas para nombrar los niveles o bloques del tercer factor del bloque. Al igual que en el cuadro latino, cada letra (latinas y griegas) debe aparecer sólo una vez en cada renglón y en cada columna. Además, cada par de letras debe aparecer sólo una vez en todo el arreglo. En la tabla 10 se



presenta el aspecto de los datos del diseño en cuadro latino de dimensión  $k = 4$ .

Tabla 10 Diseño en cuadro grecolatino (DCGL)					
		Columnas			
		1	2	3	4
Renglones	1	$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
	2	$B\delta$	$A\gamma$	$D\beta$	$C\alpha$
	3	$C\beta$	$D\alpha$	$A\delta$	$B\gamma$
	4	$D\gamma$	$C\delta$	$B\alpha$	$A\beta$

El modelo estadístico que describe a las mediciones en un cuadro grecolatino está dado por

$$Y_{ijlm} = \mu + \tau_i + \gamma_j + \delta_l + \varphi_m + \varepsilon_{ijlm}$$

donde  $Y_{ijlm}$  es la observación o respuesta que se encuentra en el tratamiento  $i$  ( $i$ -ésima letra latina), en el renglón  $j$ , en la columna  $l$  y en la  $m$ -ésima letra griega;  $\tau_i$  es el efecto del tratamiento  $i$ ;  $\gamma_j$  es el efecto del renglón  $j$ ;  $\delta_l$  representa el efecto de la columna  $l$ ;  $\varphi_m$  representa el efecto de la  $m$ -ésima letra griega, que son los niveles del tercer factor de bloque;  $\varepsilon_{ijlm}$  representa el error aleatorio atribuible a la medición  $Y_{ijlm}$ . Es importante no confundir las letras griegas del modelo que representan efectos, con las letras griegas en el diseño que representan a los niveles del tercer factor de bloque. La variabilidad total presente en los datos se puede partir de la manera usual como

$$SCT = SCTr + SCB1 + SCB2 + SCB3 + SCE$$

donde las sumas  $SCB1$ ,  $SCB2$  y  $SCB3$  miden la variabilidad debida a los factores de bloque renglón, columna y de letras griegas, respectivamente. Para  $k$  tratamientos, los grados de libertad correspondientes a cada suma son

$$k^2 - 1 = (k - 1) + (k - 1) + (k - 1) + (k - 1) + (k - 3)(k - 1)$$

Un bosquejo del análisis de varianza se muestra en la tabla 11. En ella se prueban las hipótesis de igualdad de letras latinas (tratamientos), de renglones, de columnas y de letras griegas.

Tabla 11 ANOVA DCGL			
variabilidad	SC	GL	F
Tratamientos	$SCTr = \sum_{i=1}^k \frac{Y_{i...}^2}{k} - \frac{Y_{....}^2}{N}$	$k - 1$	$\frac{CMT_r}{CME}$
Renglones	$SCB1 = \sum_{j=1}^k \frac{Y_{.j..}^2}{k} - \frac{Y_{....}^2}{N}$	$k - 1$	$\frac{CMB1}{CME}$
Columnas	$SCB2 = \sum_{l=1}^k \frac{Y_{..l.}^2}{k} - \frac{Y_{....}^2}{N}$	$k - 1$	$\frac{CMB2}{CME}$
Letras griegas	$SCB3 = \sum_{m=1}^k \frac{Y_{...m}^2}{k} - \frac{Y_{....}^2}{N}$	$k - 1$	$\frac{CMB3}{CME}$
Error	$SCE = SCT - SCTr - SCB1 - SCB2 - SCB3$	$(k - 3)(k - 1)$	
Total	$SCT = \sum_{i=1}^k \sum_{j=1}^k \sum_{l=1}^k \sum_{m=1}^k Y_{ijlm}^2 - \frac{Y_{....}^2}{N}$	$k^2 - 1$	

EJEMPLO 4. *Comparación de cuatro marcas de llantas*

En el ejemplo de las marcas de llantas, visto como un diseño en cuadro latino, se controlaron dos factores de bloque: tipo de carro y posición. Supongamos ahora que en algún momento se consideró controlar un factor de bloque adicional, como son los conductores. Con esto el diseño a utilizar habría sido el cuadro grecolatino dado por

Tabla 12 Diseño en cuadro grecolatino (DCGL)					
		Columnas			
		1	2	3	4
Renglones	1	$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
	2	$B\gamma$	$A\delta$	$D\alpha$	$C\beta$
	3	$C\delta$	$D\gamma$	$A\beta$	$B\alpha$
	4	$D\beta$	$C\alpha$	$B\delta$	$A\gamma$

donde las letras griegas  $\alpha$ ,  $\beta$ ,  $\gamma$  y  $\delta$  denotan a los cuatro conductores; las columnas son el tipo de carro y los renglones son la posición de la llanta. El problema con este diseño es que cada tipo de automovil tiene que ser probado con cada conductor, lo que aumenta el costo experimento al tener que utilizar 16 carros en total en lugar de cuatro que requiere el cuadro latino con el inconveniente adicional de que cada chofer tiene que hacer cuatro veces el recorrido de 32000 kilómetros. Entonces resulta impráctico considerar al factor conductor y es mejor estrategia cuidar sólo que sean conductores similares y que se ajusten a ciertas reglas durante el recorrido.

Si el factor de ruido con letras griegas fuera algo que tiene que ver con cada llanta de manera independiente, el costo experimental sería el mismo, al requerir sólo cuatro carros en el diseño grecolatino.

### 5. Ejemplo de un Diseño en cuadro latino (DCL) y en cuadro grecolatino (DCGL)

- (1) Suponga que un experimentador estudia los efectos que tienen *cinco formulaciones diferentes de la carga propulsora*, utilizada en los sistemas de expulsión de la tripulación de un avión basado en la *rápidez de combustión*. Cada formulación se hace con un lote de materia prima que solo alcanza para probar cinco formulaciones. Además, las formulaciones son preparadas por varios operadores, y puede haber diferencias sustanciales en las habilidades y experiencias de los operadores. Por lo tanto, hay dos factores perturbadores que serán “calculados en promedio” en el diseño: *los lotes de materia prima y los operadores*. El diseño apropiado para este problema consiste en probar cada formulación exactamente una vez con cada uno de los cinco operadores. Al diseño resultante se llama *diseño de cuadro latino (DCL)*, que usaremos para eliminar las dos fuentes perturbadoras.
- (2) *Factor de interés*: Formulaciones de la carga propulsora.
- (3) *Niveles del Factor*: cinco formulaciones  $A$ ,  $B$ ,  $C$ ,  $D$ , y  $E$  (cinco niveles  $k = 5$ )
- (4) *Variable de interés*  $Y =$  Rapidez de combustión
- (5) *Replicas por nivel*  $n = 1$ .

- (6) Suponga que la secuencia de la prueba es el cuadro latino estandar

A	B	C	D	E
B	C	D	E	A
C	D	E	A	B
D	E	A	B	C
E	A	B	C	D

- (7) Datos de la rapidez de combustión
- $Y$
- ()

Tabla 6. Diseño en cuadro latino (DCL)						
		columnas: $l$ (operadores)				
		1	2	3	4	5
renglones: $j$ (materia prima)	1	A=24	B=20	C=19	D=24	E=24
	2	B=17	C=24	D=30	E=27	A=36
	3	C=18	D=38	E=26	A=27	B=21
	4	D=26	E=31	A=26	B=23	C=22
	5	E=22	A=30	B=20	C=29	D=31

- (8) Modelo estadístico

$$Y_{ijl} = \mu + \tau_i + \gamma_j + \delta_l + \varepsilon_{ijl}$$

donde  $Y_{ijl}$  := rapidez de combustión de la  $i$ -ésima formulación, realizada por el  $j$ -ésimo operador (factor columna) con el  $l$ -ésimo lote de materia prima (factor renglón);  $\tau_i$  := es la medida del efecto de la  $i$ -ésima formulación a la rapidez de combustión,  $\varepsilon_{ijl}$  := es el error aleatorio y  $\mu$  := es la media global real de todas las formulaciones.

- (9) Hipótesis del problema

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D = \mu_E = \mu$$

$$H_1 : \mu_i \neq \mu_j \text{ para algunos } i, j$$

- (10) Significancia de la Prueba:
- $\alpha = 0.05$

- (11) Codificamos los datos restando a cada observación 25. Se obtiene el siguiente resumen de los datos

Tabla 6. Diseño en cuadro latino (DCL)									
		columnas: $l$ (operadores)					total		
		1	2	3	4	5	$Y_{.j}$	$Y_{.j}^2$	$\sum_{i,l=1}^5 Y_{ijl}^2$
renglones: $j$ (materia prima)	1	A=24	B=20	C=19	D=24	E=24	111	12321	2489
	2	B=17	C=24	D=30	E=27	A=36	134	17956	3790
	3	C=18	D=38	E=26	A=27	B=21	130	16900	3614
	4	D=26	E=31	A=26	B=23	C=22	128	16384	3326
	5	E=22	A=30	B=20	C=29	D=31	132	17424	3586
$Y_{..l}$		107	143	121	130	134			$Y_{.j}=635$
$Y_{..l}^2$		11449	20449	14641	16900	17956			81395
total							635	80985	16805

(12) Totales de tratamientos

Letra latina	total del tratamiento
<i>A</i>	$Y_{1..} = 143$
<i>B</i>	$Y_{2..} = 101$
<i>C</i>	$Y_{3..} = 112$
<i>D</i>	$Y_{4..} = 149$
<i>E</i>	$Y_{5..} = 130$
	$Y_{i..}^2 = 82295$

(13) Fuentes de variabilidad: Estadísticos de la ANOVA

$$SCT = \sum_{i,j,l=1}^5 Y_{ijl}^2 - \frac{Y_{...}^2}{N} = 16805 - \frac{635^2}{25} = 676$$

$$SC_{formulaciones} = SCTr = \sum_{i=1}^5 \frac{Y_{i..}^2}{5} - \frac{Y_{...}^2}{N} = \frac{82295}{5} - \frac{635^2}{25} = 330$$

$$SC_{lotes} = SCB1 = \sum_{j=1}^5 \frac{Y_{.j.}^2}{5} - \frac{Y_{...}^2}{N} = \frac{80985}{5} - \frac{635^2}{25} = 68$$

$$SC_{operadores} = SCB2 = \sum_{l=1}^5 \frac{Y_{...l}^2}{5} - \frac{Y_{...}^2}{N} = \frac{81395}{5} - \frac{635^2}{25} = 150$$

$$SCE = SCT - SCTr - SCB1 - SCB2 = 676 - 68 - 150 - 330 = 128$$

(14) Anova para rapidez de combustión

Tabla ANOVA para rapidez de combustión en un DCL				
Fuente	Sumas de cuad	Gl	Cuadrado Medio	Cociente-F
Formulaciones	330	4	82.50	7.73
Lotes de materia prima	68	4	17.00	valor p
Operadores	150	4	37.5	=0.0025
Error	128	12	10.67	
Total	676	24		

Como  $F_{0.05;4,12} = 3.26 < 7.73$  Se concluye que hay una diferencia significativa en la rapidez de combustión media generada por las diferentes formulaciones de la carga propulsora. También hay indicios de que hay diferencias entre los operadores, por lo que la formación en bloques de este factor fue una buena precaución. No hay evidencia sólida de una diferencia entre los lotes de materia prima, por lo que al parecer en este experimento particular hubo una preocupación innecesaria en esta fuente de variabilidad. Sin embargo, la formación de bloques de los lotes de materia prima es por lo general una buena idea.

(15) Haciendo uso de Statgraphics: Contrste LSD Grafica de medias

(16) **(Cuadro Grecolatino: DCGL)** Suponga que en el experimento de la carga propulsora un factor adicional: los montajes de prueba, podría ser importante. Sea que haya cinco montajes de prueba denotados con las

5. EJEMPLO DE UN DISEÑO EN CUADRO LATINO (DCL) Y EN CUADRO GRECOLATINO (DCGL)

Analysis of Variance for Rapidez de combustion - Type III Sums of Squares					
Source	Sun of Squares	Df	Mean Square	F-Ratio	P-Value
MAIN EFFECTS					
A:Carga propulsora	330,0	4	82,5	7,73	0,0025
B:Lotes	68,0	4	17,0	1,59	0,2391
C:Operadores	150,0	4	37,5	3,52	0,0404
RESIDUAL	128,0	12	10,6667		
TOTAL (CORRECTED)	676,0	24			

All F-ratios are based on the residual mean square error.

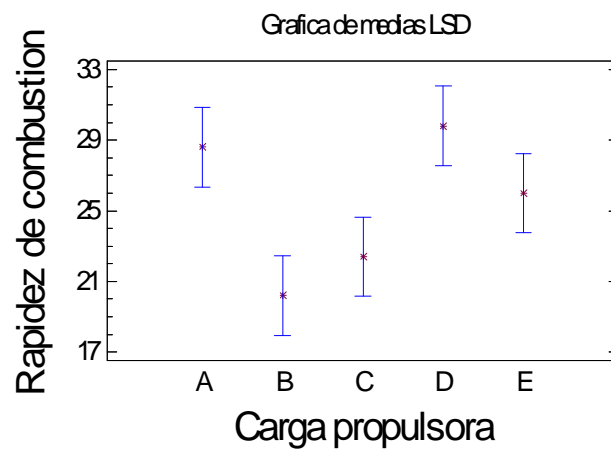
Multiple Range Tests for Rapidez de combustion by Carga propulsora

Method: 95,0 percent LSD				
Level	Count	LS Mean	LS Sigma	Homogeneous Groups
B	5	20,2	1,46059	X
C	5	22,4	1,46059	XX
E	5	26,0	1,46059	XX
A	5	28,6	1,46059	X
D	5	29,8	1,46059	X

Contrast	Difference	+/- Limits
A - B	*8,4	4,50055
A - C	*6,2	4,50055
A - D	-1,2	4,50055
A - E	2,6	4,50055
B - C	-2,2	4,50055
B - D	*-9,6	4,50055
B - E	*-5,8	4,50055
C - D	*-7,4	4,50055
C - E	-3,6	4,50055
D - E	3,8	4,50055

\* denotes a statistically significant difference.



letras griegas  $\alpha, \beta, \gamma, \delta$ , y  $\varepsilon$

		columnas: $l$ (operadores)					total			
		1	2	3	4	5	$Y_{.j..}$	$Y_{.j..}^2$	$\sum_{i,l,k=1}^5 Y_{ijlk}^2$	
renglones: $j$ (materia prima)	1	A $\alpha$ =24	B $\gamma$ =20	C $\varepsilon$ =19	D $\beta$ =24	E $\delta$ =24	111	12321	2489	
	2	B $\beta$ =17	C $\delta$ =24	D $\alpha$ =30	E $\gamma$ =27	A $\varepsilon$ =36	134	17956	3790	
	3	C $\gamma$ =18	D $\varepsilon$ =38	E $\beta$ =26	A $\delta$ =27	B $\alpha$ =21	130	16900	3614	
	4	D $\delta$ =26	E $\alpha$ =31	A $\gamma$ =26	B $\varepsilon$ =23	C $\beta$ =22	128	16384	3326	
	5	E $\varepsilon$ =22	A $\beta$ =30	B $\delta$ =20	C $\alpha$ =29	D $\gamma$ =31	132	17424	3586	
$Y_{..l.}$		107	143	121	130	134				
$Y_{..l.}^2$		11449	20449	14641	16900	17956				81395
total							635	80985	16805	

Letra latina	total del tratamiento	Letra griega	total del ensamble
$A$	$Y_{1...} = 143$	$\alpha$	$Y_{...1} = 135$
$B$	$Y_{2...} = 101$	$\beta$	$Y_{...2} = 119$
$C$	$Y_{3...} = 112$	$\gamma$	$Y_{...3} = 122$
$D$	$Y_{4...} = 149$	$\delta$	$Y_{...4} = 121$
$E$	$Y_{5...} = 130$	$\varepsilon$	$Y_{...5} = 138$
	$\sum_{i=1}^5 Y_{i...}^2 = 82295$		$\sum_{k=1}^5 Y_{...k}^2 = 80955$

Modelo estadístico

$$Y_{ijklk} = \mu + \tau_i + \theta_j + \omega_l + \psi_k + \varepsilon_{ijklk}$$

donde  $Y_{ijklk} :=$  rapidez de combustión de la  $i$ -ésima formulación, realizada por el  $j$ -ésimo operador (factor columna) con el  $l$ -ésimo lote de materia prima (factor renglón) con el  $k$ -ésimo montaje;  $\tau_i :=$  es la medida del efecto de la  $i$ -ésima formulación a la rapidez de combustión,  $\varepsilon_{ijl} :=$  es el error aleatorio y  $\mu :=$  es la media global real de todos las formulaciones.

Hipótesis del problema

$$\begin{aligned} H_0 : \mu_A &= \mu_B = \mu_C = \mu_D = \mu_E = \mu \\ H_1 : \mu_i &\neq \mu_j \text{ para algunos } i, j \end{aligned}$$

Estadísticos para el Anova

$$SCT = \sum_{i,j,l,k=1}^5 Y_{ijkl}^2 - \frac{Y_{\dots}^2}{N} = 16805 - \frac{635^2}{25} = 676$$

$$SC_{formulaciones} = SCTr = \sum_{i=1}^5 \frac{Y_{i\dots}^2}{5} - \frac{Y_{\dots}^2}{N} = \frac{82295}{5} - \frac{635^2}{25} = 330$$

$$SC_{lotes} = SCB1 = \sum_{j=1}^5 \frac{Y_{\cdot j \dots}^2}{5} - \frac{Y_{\dots}^2}{N} = \frac{80985}{5} - \frac{635^2}{25} = 68$$

$$SC_{operadores} = SCB2 = \sum_{l=1}^5 \frac{Y_{\cdot \cdot l \cdot}^2}{5} - \frac{Y_{\dots}^2}{N} = \frac{81395}{5} - \frac{635^2}{25} = 150$$

$$SC_{ensambles} = SCB3 = \sum_{k=1}^5 \frac{Y_{\cdot \cdot \cdot k}^2}{5} - \frac{Y_{\dots}^2}{N} = \frac{80955}{5} - \frac{635^2}{25} = 62$$

$$SCE = SCT - SCTr - SCB1 - SCB2 - SCB3 = 676 - 330 - 68 - 150 - 62 = 66$$

Tabla ANOVA para rapidez de combustión					
Fuente	Sumas de cuad	Gl	Cuadrado Medio	Cociente-F	valor p
Formulaciones	330	4	82.50	10	0,0033
Lotes de materia prima	68	4	17.00	2,06	0,1783
Operadores	150	4	37.5	4,55	0,0329
Montajes de la prueba	62	4	15.50	1,88	0,2076
Error	66	8	8.25		
Total	676	24			

Al comparar los dos diseños DCL y DCGL, se observa que al sacar la variabilidad debida a los montajes de prueba, el error experimental disminuye. Sin embargo, al disminuir el error experimental, se han reducido también los grados de libertad de 12 a 8. Por lo tanto, la estimación de error tiene menos grado de libertad y la prueba puede ser menos sensible.

## 6. Diseño de bloques incompletos balanceados (BIBD)

**6.1. Introducción.** Es posible que en algunos experimentos que usan diseños por bloques no puedan realizarse los ensayos de todas las combinaciones de tratamiento dentro de cada bloque. Situaciones como éstas ocurren debido a escasez en los recursos del experimento, o por el tamaño físico de los bloques. En estos casos es posible usar diseños aleatorizados por bloques en los que cada tratamiento no está presente en cada bloque. Estos diseños se conocen como Diseños Aleatorizados por Bloques Incompletos (BIBD)

**6.2. Diseño por BIBD.** Cuando las comparaciones entre todos los tratamientos tienen la misma importancia, éstas deben elegirse de manera que ocurran en forma balanceada dentro de cada bloque, lo que significa que cualquier par de tratamientos ocurren juntos el mismo número de veces que cualquier otro par, este tipo de diseño se llama un *diseño balanceado por bloques incompletos*.

6.2.1. *Análisis Estadístico.* Suponga que existen  $a$  tratamientos y  $b$  bloques y que se prueban  $k$  tratamientos en cada bloque, que cada tratamiento sucede  $r$  veces en el diseño (o se repite  $r$  veces) y que hay un total de  $N = ar = bk$  observaciones. Más aún, el número de veces que cada par de tratamientos ocurre en el mismo bloque es:

$$\lambda = r(k-1)/(a-1)$$

Se dice que el diseño es *simétrico* si  $a = b$ .

El parámetro  $\lambda$  debe ser un entero. Para deducir la relación de  $\lambda$ , considérese cualquier tratamiento, por ejemplo el 1. Como el tratamiento 1 ocurre en  $r$  bloques y hay otros  $k-1$  tratamientos en cada uno de esos bloques, existen  $r(k-1)$  observaciones en un bloque que contiene al tratamiento 1. Estas  $r(k-1)$  observaciones deben representar al resto de los  $a-1$  tratamientos  $\lambda$  veces. Por lo tanto,  $\lambda(a-1) = r(k-1)$ .

EJEMPLO 5. Para  $a = 4$  (nº de tratamientos),  $b = 4$ , (nº de bloques)  $k = 3$  (tratamientos en cada bloque), un diseño BIBD puede ser construido con:  $r = bk/a = 3$  (cada tratamiento sucede 3 veces en el diseño)  $\lambda = r(k-1)/(a-1) = 3(3-1)/(4-1) = 2$  (el número de veces que cada par de tratamientos ocurre en el mismo bloque)

bloques	Tratamientos		
1	1	2	3
2	1	2	4
3	1	3	4
4	2	3	4

EJEMPLO 6. Para  $a = 4$  (nº de tratamientos),  $b = 6$ , (nº de bloques)  $k = 2$  (tratamientos en cada bloque), un diseño BIBD puede ser construido con:  $r = bk/a = 3$  (cada tratamiento sucede 3 veces en el diseño)  $\lambda = r(k-1)/(a-1) = 3(2-1)/(4-1) = 1$  (el número de veces que cada par de tratamientos ocurre en el mismo bloque)

bloques	Tratamientos	
1	1	2
2	3	4
3	1	3
4	2	4
5	1	4
6	2	3

EJEMPLO 7. Supongamos  $b = 4$  bloques incompletos para investigar  $a = 6$  tratamientos.

bloques	Tratamientos		
1	1	2	3
2	1	3	6
3	2	4	5
4	4	5	6

Aunque  $r = 2$  (cada tratamiento sucede 2 veces en el diseño), pero no es un diseño BIBD porque  $\lambda$  no es igual para todas las parejas de tratamientos. Los pares de tratamientos (1, 5), (2, 6) y (3, 4) no ocurren en todos los bloques, el resto de parejas aparecen una sola vez en el mismo bloque.



6.2.2. *El modelo estadístico del BIBD es:*

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

En donde  $Y_{ij}$  es la  $i$ -ésima observación del  $j$ -ésimo bloque,  $\mu$  es la media general,  $\tau_i$  es el efecto del  $i$ -ésimo tratamiento,  $\beta_j$  es el efecto del  $j$ -ésimo bloque, y  $\varepsilon_{ij}$  es la componente del error aleatorio  $NID(0, \sigma^2)$ .

La variación total en los datos se expresa mediante la suma total de cuadrados corregidos (o ajustados).

$$SCT = \sum_i \sum_j Y_{ij}^2 - \frac{Y_{..}^2}{N}$$

La variabilidad total puede ser descompuesta

$$SCT = SCTr_{(aj)} + SCB + SCE$$

En donde corrige la suma de cuadrados de tratamiento para separar los efectos de tratamiento y de bloque. Esta corrección es necesaria porque cada tratamiento ocurre en un conjunto diferente de  $r$  bloques. Por esta razón las diferencias entre los totales de tratamientos no corregidos,  $Y_1, Y_2, \dots, Y_a$ , también son afectadas por las diferencias entre los bloques.

La suma de cuadrados de los bloques es:

$$SCB = \frac{1}{k} \sum_{j=1}^b Y_{.j}^2 - \frac{Y_{..}^2}{N}$$

en donde  $Y_{.j}$  es el total del  $j$ -ésimo bloque. La  $SCB$  tiene  $b - 1$  grados de libertad. La suma de cuadrados de tratamiento corregida (o ajustada) es:

$$SCTr_{(aj)} = \frac{k \sum_{i=1}^a Q_i^2}{\lambda a}$$

En donde  $Q_i$  es el total corregido del  $i$ -ésimo tratamiento, el cual se calcula mediante

$$Q_i = Y_{i.} - \frac{1}{k} \sum_{j=1}^b n_{ij} Y_{.j}, i = 1, 2, \dots, a$$

con  $n_{ij} = 1$  si el tratamiento  $i$  ocurre en el bloque  $j$ ,  $n_{ij} = 0$  en otro caso. Por lo tanto,  $\frac{1}{k} \sum_{j=1}^b n_{ij} Y_{.j}$  es el promedio de los totales de los bloques en los que se aplica el tratamiento  $i$ . La suma de los totales de tratamiento corregidos siempre será 0. La  $SCTr_{(aj)}$  tiene  $a - 1$  grados de libertad. La suma de cuadrados del error se calcula por diferencia  $SCE = SCT - SCTr_{(aj)} - SCB$  y tiene  $N - a - b + 1$  grados de libertad.

La estadística apropiada para probar la igualdad de los efectos de tratamiento es

$$F = CMT_{r_{(aj)}} / CME$$

## 6.2.3. Análisis de varianza para el diseño BIBD.

Tabla Tabla ANOVA para el BIBD				
Variabilidad	SC	GL	CM	F <sub>0</sub>
Tratamientos ajustados	$\frac{k \sum_{i=1}^a Q_i^2}{\lambda a}$	a-1	$CMT r_{(a,j)}$	$\frac{CMT r_{(a,j)}}{CME}$
Bloques	$\frac{1}{k} \sum_{j=1}^b Y_{\cdot j}^2 - \frac{Y_{\cdot\cdot}^2}{N}$	—	—	—
Error	SCE	N-a-b+1	CME	
Total	$\sum_i \sum_j Y_{ij}^2 - \frac{Y_{\cdot\cdot}^2}{N}$	N-1		

EJEMPLO 8. Supóngase que un ingeniero químico cree que el tiempo de reacción en un proceso químico es función del catalizador empleado. De hecho 4 catalizadores están siendo investigados. El procedimiento experimental consiste en seleccionar un lote de materia prima, cargar una planta piloto, aplicar cada catalizador a ensayos separados de dicha planta y observar el tiempo de reacción. Debido a que las variaciones en los lotes de materia prima pueden afectar el comportamiento del catalizador, el ingeniero decide controlar este factor por medio de bloques. Sin embargo, cada lote es lo suficientemente grande para permitir el ensayo de 3 catalizadores únicamente. Por lo tanto, es necesario utilizar un diseño aleatorizado por bloques incompletos. El diseño BIBD, junto con las observaciones recopiladas aparecen en la siguiente tabla:

	Bloque (Lote de Materia Prima)				
Tratamiento (Catalizador)	1	2	3	4	$Y_{i\cdot}$
1	73	74	—	71	218
2	—	75	67	72	214
3	73	75	68	—	216
4	75	—	72	75	222
$Y_{\cdot j}$	221	224	207	218	870 = $Y_{\cdot\cdot}$

Considérense los datos de la Tabla para el experimento de los catalizadores. Éste es un diseño BIBD con  $a = 4$ ,  $b = 4$ ,  $k = 3$ ,  $r = 3$ ,  $\lambda = 2$  y  $N = 12$ . A continuación vamos a realizar el análisis de estos datos.

La Suma Total de Cuadrados y de bloques se calculan como:

$$SCT = \sum_i \sum_j Y_{ij}^2 - \frac{Y_{\cdot\cdot}^2}{N} = 63156 - \frac{(870)^2}{12} = 81$$

$$SCB = \frac{1}{3} \sum_{j=1}^4 Y_{\cdot j}^2 - \frac{Y_{\cdot\cdot}^2}{12} = \frac{(221)^2 + (224)^2 + (207)^2 + (218)^2}{3} - \frac{(870)^2}{12} = 55$$

Para calcular la suma de cuadrados de tratamiento corregida que tome en cuenta los bloques, primero hay que determinar los totales de tratamientos corregidos:

$$\begin{aligned}
 Q_i &= Y_{i.} - \frac{1}{3} \sum_{j=1}^4 n_{ij} Y_{.j}, \quad i = 1, 2, 3, 4 \\
 Q_1 &= 218 - \frac{221 + 224 + 218}{3} = -\frac{9}{3} \\
 Q_2 &= 214 - \frac{207 + 224 + 218}{3} = -\frac{7}{3} \\
 Q_3 &= 216 - \frac{221 + 207 + 224}{3} = -\frac{4}{3} \\
 Q_4 &= 222 - \frac{221 + 207 + 218}{3} = \frac{20}{3}
 \end{aligned}$$

Se calcula ahora la suma de cuadrados de tratamiento corregida:

$$\begin{aligned}
 SCTR_{(aj)} &= \frac{k \sum_{i=1}^a Q_i^2}{\lambda a} \\
 &= \frac{3 \left[ \left(-\frac{9}{3}\right)^2 + \left(-\frac{7}{3}\right)^2 + \left(-\frac{4}{3}\right)^2 + \left(\frac{20}{3}\right)^2 \right]}{(2)(4)} = 22.75
 \end{aligned}$$

La suma de cuadrados del error se calcula por diferencia:

$$\begin{aligned}
 SCE &= SCT - SCTR_{(aj)} - SCB \\
 &= 81 - 22.75 - 55 = 3.25
 \end{aligned}$$

Realizando la Tabla para el Análisis de la Varianza:

Tabla Tabla ANOVA para el BIBD					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor p
Tratamientos ajustados	22.75	3	7.58	11.66	0.0107
Bloques	55.00	3	-	-	-
Error	3.25	5	0.65		
Total	81.00	11			

6.2.4. *Evaluación de los efectos del bloque:* En ocasiones, se desea evaluar los efectos de los bloques. Para lograrlo se requiere una descomposición alterna de  $SCT$ , en otras palabras,

$$\begin{aligned}
 SCT &= SCTR + SCB_{(aj)} + SCE \\
 Q'_j &= Y_{.j} - \frac{1}{r} \sum_{i=1}^a n_{ij} Y_{i.}, \quad j = 1, 2, \dots, b \\
 SCB_{(aj)} &= \frac{r \sum_{j=1}^b (Q'_j)^2}{\lambda b}
 \end{aligned}$$

**Análisis de Varianza para tiempo - Suma de Cuadrados Tipo III**

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
EFFECTOS PRINCIPALES					
A:catalizadores	22,75	3	7,58333	11,67	0,0107
B:BLOQUE	66,0833	3	22,0278	33,89	0,0010
RESIDUOS	3,25	5	0,65		
TOTAL (CORREGIDO)	81,0	11			

Todas las razones-F se basan en el cuadrado medio del error residual

Como  $a = b = 4$ , el diseño balanceado por bloques incompletos es simétrico. Por lo tanto,

$$\begin{aligned}
 Q'_1 &= 221 - \frac{218 + 216 + 222}{3} = \frac{7}{3} \\
 Q'_2 &= 224 - \frac{218 + 214 + 216}{3} = \frac{24}{3} \\
 Q'_3 &= 207 - \frac{214 + 216 + 222}{3} = \frac{-31}{3} \\
 Q'_4 &= 218 - \frac{218 + 214 + 222}{3} = 0 \\
 SCB_{(aj)} &= \frac{3[(\frac{7}{3})^2 + (\frac{24}{3})^2 + (\frac{-31}{3})^2 + 0^2]}{(2)(4)} = 66.08 \\
 SCTr &= \frac{(218)^2 + (214)^2 + (216)^2 + (222)^2}{3} - \frac{(870)^2}{12} = 11.67
 \end{aligned}$$

El resumen del Análisis de la Varianza para el diseño BIBD simétrico, se muestra en la siguiente tabla:

Tabla Tabla ANOVA para el BIBD					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor p
Tratamientos ajustados	22.75	3	7.58	11.66	0.0107
Tratamientos no ajustados	11.67	3	-	-	
Bloques	55.00	3	-	-	-
Bloques ajustados	66.08	3	22.03	33.90	0.0010
Error	3.25	5	0.65		
Total	81.00	11			

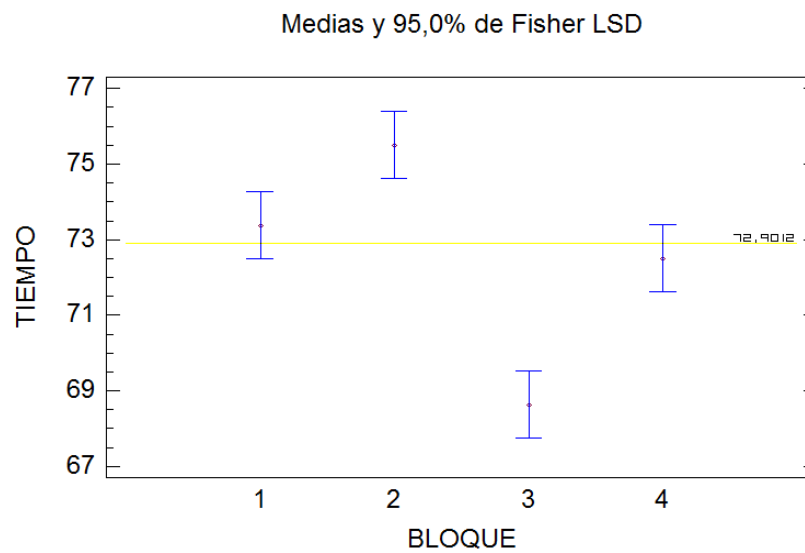
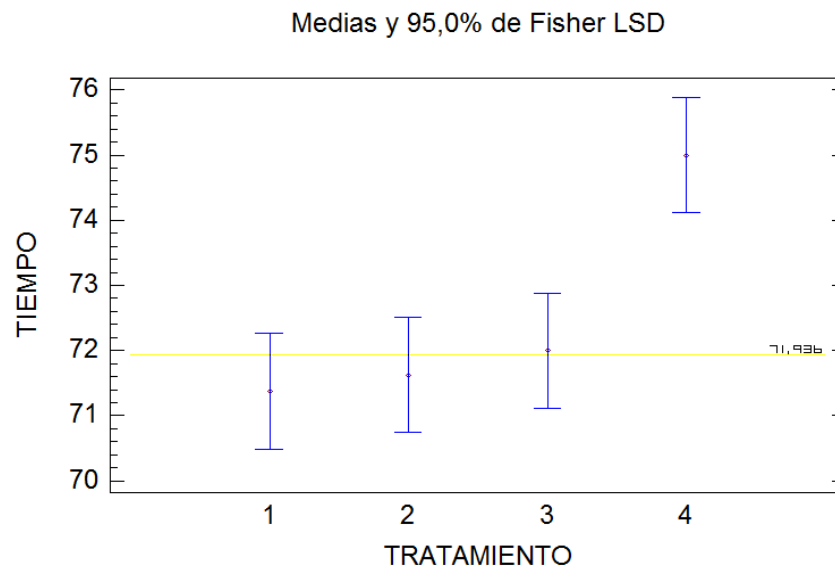
Hay que observar que la suma de cuadrados asociadas con cada media de cuadrados en la tabla anterior no es igual a la suma total de cuadrados, o sea que

$$SCT \neq SCTr_{(aj)} + SCB_{(aj)} + SCE$$

A continuación se presenta el análisis del problema usando statgraphics y las respectivas gráficas de medias

**7. Ejercicios: Diseños en bloques completos al azar**

EJERCICIO 24. 1. ¿En qué situaciones se aplica un diseño en bloques completos al azar? 2. ¿Qué diferencia hay entre un DBCA y los diseños en cuadro latino? 3.



Apoyándose en el modelo estadístico para un diseño en bloques, ¿por qué a través de este diseño se reduce el error aleatorio?

EJERCICIO 25. A continuación se muestra parte del ANOVA para un diseño en bloques, que tiene tres tratamientos y cinco bloques, con una sola repetición en

tratamiento-bloque.

Tabla 5					
Fuente	Sumas de cuad	Gl	Cuadrado Medio	Cociente-F	valor p
Tratamientos	600				
Bloque	850				
Error	500				
Total		14			

a) Agregar en esta tabla los grados de libertad, el cuadrado medio y la razón  $F$  para cada una de las fuentes de variación. b) Interprete en forma práctica para cada caso lo que está estimando el cuadrado medio. c) Escriba el modelo estadístico y las hipótesis pertinentes. d) Apoyándose en tablas de la distribución  $F$ , decida si se aceptan o se rechazan las hipótesis.

EJERCICIO 26. Realice el problema anterior, pero ahora suponga que no se haya bloqueado. ¿Se hubiesen obtenido las mismas conclusiones? Argumente.

EJERCICIO 27. Aunque en el análisis de varianza para un diseño en bloques completos al azar también se puede probar la hipótesis sobre si hay diferencia entre los bloques, se dice que esta hipótesis se debe ver con ciertas reservas. Explique por qué.

EJERCICIO 28. Explique por qué el adjetivo azar en el nombre del diseño en bloques completos al azar.

EJERCICIO 29. Una compañía farmacéutica realizó un experimento para comprobar los tiempos promedio (en días) necesarios para que una persona se recupere de los efectos y complicaciones que siguen a un resfriado común. En este experimento se compararon las personas que tomaron distintas dosis diarias de vitamina C. Para hacer el experimento se contactó a un número determinado de personas, que en cuanto les daba el resfriado empezaban a recibir algún tipo de dosis (las cuales se iban rotando). Si la edad de éstas es una posible fuente de variabilidad, explique con detalle cómo aplicaría la idea de bloqueo para controlar tal fuente de variabilidad.

EJERCICIO 30. A continuación se muestran los datos para un diseño en bloques al azar

		Tratamiento			Totales por bloque
		1	2	3	
Bloque	1	3	7	4	$Y_{.1} =$
	2	4	9	6	$Y_{.2} =$
	3	2	3	3	$Y_{.3} =$
	4	6	10	7	$Y_{.4} =$
Totales por tratamiento		$Y_{1.} =$	$Y_{2.} =$	$Y_{3.} =$	

(a) Complete los totales que se piden en la tabla anterior. (b) Calcule las sumas de cuadrados correspondientes:  $SCTr$ ,  $SCB$ ,  $SCT$  y  $SCE$ . (c) Obtenga la tabla de análisis de varianza y anote las principales conclusiones. (d) Obtenga la diferencia mínima significativa ( $LSD$ ) para comparar tratamientos en este diseño en bloques.

EJERCICIO 31. Se hace un estudio sobre la efectividad de tres marcas de atomizador para matar moscas. Para ello, cada atomizador se aplica a un grupo de 100

moscas, y se cuenta el número de moscas muertas, expresado en porcentajes. Se hicieron seis réplicas, pero éstas se hicieron en días diferentes, por ello se sospecha que puede haber algún efecto importante debido a esta fuente de variación. Los datos obtenidos se muestran a continuación.

		Tratamiento					
		1	2	3	4	5	6
Marca de atomizador	1	75	65	67	75	62	73
	2	55	59	68	70	53	50
	3	64	74	61	58	51	69

- (a) Suponiendo un DBCA, formule las hipótesis adecuadas y el modelo estadístico. (b) ¿Existe diferencia entre la efectividad promedio de los atomizadores? c) ¿Hay algún atomizador mejor? Argumente. d) ¿Hay diferencias significativas en los resultados de diferentes días en que se realizó el experimento? Argumente. e) Verifique los supuestos de normalidad y de igual varianza entre las marcas.

**EJERCICIO 32.** En una empresa lechera se tienen varios silos para almacenar leche (cisternas de 60 000 L). Un aspecto crítico para que se conserve la leche es la temperatura de almacenamiento. Se sospecha que en algunos silos hay problemas, por ello durante cinco días se decide registrar la temperatura a cierta hora crítica. Obviamente la temperatura de un día a otro es una fuente de variabilidad que podría impactar la variabilidad total.

		Día				
		lunes	martes	miércoles	jueves	viernes
Silo	A	4.0	4.0	5.0	0.5	3.0
	B	5.0	6.0	2.0	4.0	4.0
	C	4.5	4.0	3.5	2.0	3.0
	D	2.5	4.0	6.5	4.5	4.0
	E	4.0	4.0	3.5	2.0	4.0

- (a) En este problema, ¿cuál es el factor de tratamiento y cuál el factor de bloque? (b) Suponga un DBCA, formule las hipótesis adecuadas y el modelo estadístico. c) ¿Hay diferencia entre los silos? d) ¿La temperatura de un día a otro es diferente? e) Revise residuales. ¿Hay algún problema evidente?

**EJERCICIO 33.** Se diseñó un experimento para estudiar el rendimiento de cuatro detergentes. Las siguientes lecturas de "blancura" se obtuvieron con un equipo especial diseñado para 12 cargas de lavado distribuidas en tres modelos de lavadoras:

		lavadora 1	lavadora 2	lavadora 3
Detergente	A	45	43	51
	B	47	44	52
	C	50	49	57
	D	42	37	49

- a) Señale el nombre del diseño experimental utilizado. b) Formule la hipótesis que se quiere probar, de acuerdo al problema. c) Realice el análisis estadístico apropiado a estos datos y obtenga conclusiones.

**EJERCICIO 34.** a) Conteste los tres incisos del problema anterior sin tomar en cuenta el efecto de lavadoras y obtenga conclusiones. b) ¿Hay diferencias en

las conclusiones anteriores y las del problema anterior? Explique. c) ¿Con cuáles conclusiones se queda? Explique.

**EJERCICIO 35.** Una de las variables críticas en el proceso de ensamble del brazo lector de un disco duro es el ángulo que forma con el cuerpo principal de la cabeza lectora. Se corre un experimento con el objetivo de comparar dos equipos que miden dicho ángulo en unidades de radianes. Se decide utilizar como factor de bloque a los operadores de los equipos. Los resultados se muestran en la siguiente tabla:

Operador	Equipo 1					Equipo 2				
1	1.328	0.985	1.316	1.553	1.310	1.273	0.985	1.134	1.412	0.917
	1.113	1.057	1.144	1.485	1.386	0.789	0.671	0.554	1.386	1.289
2	1.269	1.268	1.091	1.195	1.380	1.036	0.783	1.108	1.129	1.132
	1.093	0.984	1.087	1.482	1.442	0.201	0.900	0.916	1.434	1.223
3	1.440	1.079	1.389	1.611	1.445	1.454	1.063	1.219	1.602	1.583
	1.150	1.190	1.247	1.617	1.574	1.018	1.050	0.997	1.538	1.478

a) Plantee el modelo y las hipótesis adecuadas al problema. b) ¿Existen diferencias entre los equipos? Argumente estadísticamente. c) ¿Existen diferencias entre los operadores? d) Dibuje los diagramas de cajas simultáneos y las gráficas de medias para ambos factores. Interpretelas. e) Verifique los supuestos de normalidad e igualdad de varianza entre tratamientos, así como la posible presencia de puntos aberrantes.

**EJERCICIO 36.** Se quiere estudiar el efecto de cinco diferentes catalizadores (A, B, C, D y E) sobre el tiempo de reacción de un proceso químico. Cada lote de material sólo permite cinco corridas y cada corrida requiere aproximadamente de 1.5 horas, por lo que sólo se pueden realizar cinco corridas diarias. El experimentador decide correr los experimentos con un diseño en cuadro latino, para controlar activamente a los lotes y días. Los datos obtenidos son:

		Día				
		1	2	3	4	5
Lote	1	A=8	B=7	D=1	C=7	E=3
	2	C=11	E=2	A=7	D=3	B=8
	3	B=4	A=9	C=10	E=1	D=5
	4	D=6	C=8	E=6	B=6	A=10
	5	E=4	D=2	B=3	A=8	C=8

(a) ¿Cómo se aleatorizó el experimento? (b) Anote la ecuación del modelo y las hipótesis estadísticas correspondientes. (c) ¿Existen diferencias entre los tratamientos? ¿Cuáles tratamientos son diferentes entre sí? (d) ¿Los factores de ruido, lote y día afectan el tiempo de reacción del proceso? (e) Dibuje los gráficos de medias para los tratamientos, los lotes y los días. ¿Cuál tratamiento es mejor? (f) Verifique los supuestos del modelo, considerando que los datos se obtuvieron columna por columna, día a día.

**EJERCICIO 37.** En el problema anterior elimine el factor de bloque días, y conteste: (a) ¿Se justifica la eliminación? (b) Sin tomar en cuenta el día, señale el nombre del diseño, el modelo y las hipótesis adecuadas al problema. (c) Pruebe las hipótesis y obtenga conclusiones. (d) Compare el cuadro medio del error, en este caso con el del problema anterior. ¿Qué observa? ¿Cómo lo explica? (e) ¿Por qué



se obtienen las mismas conclusiones respecto a los tratamientos en este problema y el anterior?

EJERCICIO 38. Respecto a los dos últimos problemas, además de haber eliminado el factor día, ahora elimine el factor lote, y conteste (a) ¿Se justifica esta segunda eliminación respecto a ambos problemas? (b) Después de la doble eliminación, señale el nombre del diseño, el modelo y la(s) hipótesis adecuada(s) al problema. (c) Pruebe las hipótesis y obtenga conclusiones. (d) Compare el cuadro medio del error obtenido con los de los problemas anteriores. ¿Qué observa? ¿Cómo lo explica? (e) ¿Por qué se obtienen las mismas conclusiones respecto a los tratamientos en este problema y los dos anteriores? (f) ¿Cree que esta eliminación hubiese afectado si los factores de bloque hubieran sido significativos?

EJERCICIO 39. Se comprueba el peso en gramos de un material de tres proveedores: A, B y C, por tres diferentes inspectores: I, II y III, utilizando tres diferentes escalas: 1, 2 y 3. El experimento es llevado a cabo como el cuadro latino siguiente:

		Escala		
		1	2	3
Inspector	I	A=16	B=10	C=11
	II	B=15	C=9	A=10
	III	C=13	A=11	C=13

(a) ¿Hay diferencias entre los proveedores? (b) ¿Hay diferencias entre los inspectores y entre las escalas? (c) Si el peso debe ser 15g, ¿cuál proveedor es mejor? (d) Si algún factor de bloques no significativo, elimínelo, haga un análisis adecuado.

EJERCICIO 40. Un químico quiere probar el efecto de cuatro agentes químicos sobre la resistencia de un tipo particular de tela. Debido a que podría haber variabilidad de un rollo de tela a otro, el químico decide usar un diseño de bloques aleatorizados, con los rollos de tela considerados como bloques. Selecciona cinco rollos y aplica los cuatro agentes químicos de manera aleatoria a cada rollo. A continuación se presentan las resistencias tensión resultantes. (a) Analizar los datos de este experimento (utilizar  $\alpha = 0.05$ ) y sacar las conclusiones apropiadas. (b) Suponiendo que los tipos de agentes químicos y los rollos de tela son fijos, estimar los parámetros del modelo  $\tau_i$  y  $\alpha_j$ . (c) Suponga que falta la observación del agente químico 2 y el rollo 3 en el problema. Analizar el problema estimando el valor faltante. Realizar el análisis exacto y comparar los resultados.

		Rollo				
		1	2	3	4	5
Agente químico	1	73	68	74	71	67
	2	73	67	75	72	70
	3	75	68	78	73	68
	4	73	71	75	75	69

EJERCICIO 41. Se están comparando tres soluciones de lavado diferentes a fin de estudiar su efectividad para retardar el crecimiento de bacterias en contenedores de leche de 5 galones. El análisis se hace en un laboratorio y sólo pueden realizarse tres ensayos en un día. Puesto que los días podrían representar una fuente potencial de variabilidad, el experimentador decide usar un diseño de bloques aleatorizados. Se hacen observaciones en cuatro días, cuyos datos se muestran enseguida. Analizar

los datos de este experimento (utilizara  $\alpha = 0.05$ ) y sacar las conclusiones apropiadas. Analizar los datos del problema utilizando la prueba general de significación de la regresión.

		Días			
		1	2	3	4
Solución	1	13	22	18	39
	2	16	24	17	44
	3	5	4	1	22

EJERCICIO 42. En un artículo de *Fire Safety Journal* ("El efecto del diseño de boquillas en la estabilidad y el desempeño de surtidores de agua turbulenta", vol. 4) se describe un experimento en el que se determinó un factor de la forma para varios diseños diferentes de boquillas con seis niveles de la velocidad del flujo de salida del surtidor. El interés se centró en las diferencias potenciales entre los diseños de las boquillas, con la velocidad considerada como una variable perturbadora. Los datos se presentan a continuación.

		Velocidad del flujo de salida del surtidor (m/s)					
		11.73	14.37	16.59	20.43	23.46	28.74
Diseño	1	0.78	0.80	0.81	0.75	0.77	0.78
	2	0.85	0.85	0.92	0.86	0.81	0.83
	3	0.93	0.92	0.95	0.89	0.89	0.83
	4	1.14	0.97	0.98	0.88	0.86	0.83
	5	0.97	0.86	0.78	0.76	0.76	0.75

- (a) ¿El diseño de la boquilla afecta el factor de la forma? Comparar las boquillas con un diagrama de dispersión y con un análisis de varianza, utilizando  $\alpha = 0.05$ .  
 (b) Analizar los residuales de este experimento. (c) ¿Qué diseños de las boquillas son diferentes con respecto al factor de la forma? Trazar una gráfica del factor de la forma promedio para cada tipo de boquilla y compararla con una distribución  $t$  escalada. Comparar las conclusiones que se sacaron a partir de esta gráfica con las de la prueba del rango múltiple de Duncan.

EJERCICIO 43. El fabricante de una aleación maestra de aluminio produce refinadores de textura en forma de lingotes. La compañía produce el producto en cuatro hornos. Se sabe que cada horno tiene sus propias características únicas de operación, por lo que en cualquier experimento que se corra en la fundición en el que se use más de un horno, los hornos se considerarán como una variable perturbadora. Los ingenieros del proceso sospechan que la velocidad de agitación afecta la medida de la textura del producto. Cada horno puede operarse con cuatro diferentes velocidades de agitación. Se lleva a cabo un diseño de bloques aleatorizados para un refinador particular y los datos resultantes de la medida de la textura se muestran a continuación:

		Horno			
		1	2	3	4
velocidad de agitación	5	8	4	5	6
	10	14	5	6	9
	15	14	6	9	2
	20	17	9	3	6

- a) ¿Existe evidencia de que la velocidad de agitación afecta la medida de la textura?  
 b) Representar los residuales de este experimento en una gráfica de probabilidad

normal. Interpretar esta gráfica. c) Graficar los residuales contra el horno y la velocidad de agitación. ¿Esta gráfica proporciona alguna información útil? d) ¿Cuál sería la recomendación de los ingenieros del proceso con respecto a la elección de la velocidad de agitación y del horno para este refinador de textura particular si es deseable una medida de la textura pequeña?

EJERCICIO 44. Un ingeniero industrial está realizando un experimento sobre el tiempo de enfoque del ojo. Se interesa en el efecto de la distancia del objeto al ojo sobre el tiempo de enfoque. Cuatro distancias diferentes son de interés. Cuenta con cinco sujetos para el experimento. Debido a que puede haber diferencias entre los individuos, el ingeniero decide realizar el experimento en un diseño de bloques aleatorizados. Los datos obtenidos, se presentan a continuación. Analizar los datos de este experimento (utilizar  $\alpha = 0.05$ ) y sacar las conclusiones apropiadas.

		Sujeto				
		1	2	3	4	5
Distancia	4	10	6	6	6	6
	6	7	6	6	1	6
	8	5	3	3	2	5
	10	6	4	4	2	3

EJERCICIO 45. Se estudia el efecto de cinco ingredientes diferentes (A, B, C, D y E) sobre el tiempo de reacción de un proceso químico. Cada lote de material nuevo sólo alcanza para permitir la realización de cinco corridas. Además, cada corrida requiere aproximadamente hora y media, por lo que sólo pueden realizarse cinco corridas en un día. El experimentador decide realizar el experimento como un cuadrado latino para que los efectos del día y el lote puedan controlarse sistemáticamente. Obtiene los datos que se muestran enseguida. (a) Analizar los datos de este experimento (utilizar  $\alpha = 0.05$ ) y sacar conclusiones. (b) Suponga que en el problema falta la observación del lote 3 en el día 4. Estimar el valor faltante, y realizar el análisis utilizando este valor. (c) Suponga que en el problema los datos tomados en el día 5 se analizaron incorrectamente y fue necesario descartarlos. Desarrollar un análisis apropiado para los datos restantes.

		Día				
		1	2	3	4	5
Lote	1	A=8	B=7	D=1	C=7	E=3
	2	B=11	E=2	A=7	D=3	B=8
	3	C=4	A=9	C=10	E=1	D=5
	4	D=6	C=8	E=6	B=6	A=10
	5	E=4	D=2	B=3	A=8	C=8

EJERCICIO 46. Un ingeniero industrial investiga el efecto de cuatro métodos de ensamblaje (A, B, C y D) sobre el tiempo de ensamblaje de un componente de televisores a color. Se seleccionan cuatro operadores para el estudio. Además, el ingeniero sabe que todos los métodos de ensamblaje producen fatiga, de tal modo que el tiempo requerido para el último ensamblaje puede ser mayor que para el primero, independientemente del método. Es decir, se desarrolla una tendencia en el tiempo de ensamblaje requerido. Para tomar en cuenta esta fuente de variabilidad, el ingeniero emplea el diseño del cuadrado latino que se presenta a continuación.

Analizar de este experimento ( $\alpha = 0.05$ ) y sacar las conclusiones apropiadas.

		Operador			
		1	2	3	4
ensamblaje	1	C=10	D=14	A=7	B=8
	2	B=7	C=18	D=11	A=8
	3	A=5	B=10	C=11	D=9
	4	D=10	A=10	B=12	C=14

EJERCICIO 47. Suponga que en el problema anterior el ingeniero sospecha que los sitios de trabajo usados por los cuatro operadores pueden representar una fuente adicional de variación. Es posible introducir un cuarto factor, el sitio de trabajo ( $\alpha, \beta, \gamma, \delta$ ), y realizar otro experimento, de donde resulta el cuadrado grecolatino siguiente. Analizar los datos de este experimento (utilizar  $\alpha = 0.05$ ) y sacar conclusiones.

		Operador			
		1	2	3	4
ensamblaje	1	C $\beta$ =11	B $\gamma$ =10	D $\delta$ =14	A $\alpha$ =8
	2	B $\alpha$ =8	C $\delta$ =12	A $\gamma$ =10	D $\beta$ =12
	3	A $\delta$ =9	D $\alpha$ =11	B $\beta$ =7	C $\gamma$ =15
	4	D $\gamma$ =9	A $\beta$ =8	C $\alpha$ =18	B $\delta$ =6

EJERCICIO 48. El rendimiento de un proceso químico se midió utilizando cinco lotes de materia prima, cinco concentraciones del ácido, cinco tiempos de procesamiento (A, B, C, D y E) y cinco concentraciones del catalizador ( $\alpha, \beta, \gamma, \delta, \varepsilon$ ). Se usó el cuadrado grecolatino siguiente. Analizar los datos de este experimento (utilizar  $\alpha = 0.05$ ) y sacar conclusiones.

		concentración de ácido				
		1	2	3	4	5
lote	1	A $\alpha$ =26	B $\beta$ =16	C $\gamma$ =19	D $\delta$ =16	E $\varepsilon$ =13
	2	B $\gamma$ =18	C $\delta$ =21	D $\varepsilon$ =18	E $\alpha$ =11	A $\beta$ =21
	3	C $\varepsilon$ =20	D $\alpha$ =12	E $\beta$ =16	A $\gamma$ =25	B $\delta$ =13
	4	D $\beta$ =15	E $\gamma$ =15	A $\delta$ =22	B $\varepsilon$ =14	C $\alpha$ =17
	5	E $\delta$ =10	A $\varepsilon$ =24	B $\alpha$ =17	C $\beta$ =17	D $\gamma$ =14

## CHAPTER 4

# Diseños Factoriales

### 1. Competencias

Explicar cuando un diseño de experimentos es un diseño factorial, describiendo los conceptos básicos que estos involucran y mostrado cómo se hace tal experimentación. Desarrollar los diseños factoriales de dos y tres factores. Conocer el diseño factorial general y diferenciar los modelos de efectos fijos con los modelos de efectos aleatorios. Interpretar correctamente los análisis gráficos y el análisis de varianza en los diseños factoriales.

### 2. Introducción

Es frecuente que en muchos procesos existan varios factores de los que es necesario investigar de manera simultánea su influencia sobre una o varias variables de respuesta, donde cada factor tiene la misma importancia a priori desde el momento que se decide estudiarlo, y es poco justificable suponer de antemano que los factores no interactúan entre sí. Los diseños experimentales que permiten estudiar de manera simultánea el efecto de varios factores son los llamados *diseños factoriales*.

### 3. Conceptos básicos en diseños factoriales

El objetivo de un diseño factorial es estudiar el efecto de varios factores sobre una o varias respuestas o características de calidad y determinar una combinación de niveles de los factores en la cual el desempeño del proceso sea mejor que en las condiciones de operación actuales; es decir, encontrar nuevas condiciones de operación del proceso que eliminen o disminuyan ciertos problema de calidad en la variable de salida.

Los factores pueden ser de tipo *cualitativo* (máquinas, tipos de material, operador, la presencia o ausencia de una operación previa, etc.), o de tipo *cuantitativo* (temperatura, humedad, velocidad, presión, etc.). Para poder estudiar la manera en que influye cada factor sobre la variable respuesta, es necesario elegir al menos dos niveles de prueba para cada uno de ellos (tres máquinas, dos operadores, tres velocidades, dos temperaturas, etc). Con el diseño factorial completo se corren aleatoriamente en el proceso todas las posibles combinaciones que pueden formarse con los niveles seleccionados.

**3.1. Definición de experimento factorial.** Un diseño de experimentos factorial o *arreglo factorial* es el conjunto de puntos experimentales o tratamientos que pueden formarse considerando todas las posibles combinaciones de los niveles de los factores. Por ejemplo, con  $k = 2$  factores, ambos con dos niveles de prueba, se forma el diseño factorial  $2 \times 2 = 2^2$ , que consiste de cuatro combinaciones o puntos experimentales.

Considerando otra vez  $k = 2$  factores, pero ahora uno con tres niveles y el otro con dos niveles, se pueden construir  $3 \times 2$  combinaciones que dan lugar al diseño factorial  $3 \times 2$ . Observe que en el nombre del diseño factorial va implícito el número de tratamientos que lo componen. Para obtener el número de *corridas experimentales* se multiplica el número de tratamientos por el número de réplicas, donde una réplica se lleva a cabo cada vez que se repite el arreglo completo.

Más en general, la familia de diseños factoriales  $2^k$  consiste de  $k$  factores, todos con dos niveles de prueba; y la familia de diseños factoriales  $3^k$  consiste de  $k$  factores cada uno con tres niveles de prueba. Es claro que si los  $k$  factores no tienen la misma cantidad de niveles, entonces no se puede factorizar de esta forma, y debe escribirse el producto de manera más explícita: por ejemplo con  $k = 3$  factores, el primero con cuatro niveles y los dos restantes con dos niveles, se tiene el diseño factorial  $4 \times 2 \times 2$  o  $4 \times 2^2$ , que consiste de 16 combinaciones de niveles diferentes.

**EJEMPLO 9. (*Diseño factorial  $2^2$* )** Supongamos que se tienen dos factores  $A$ : tiempo y  $B$ : velocidad, cada uno con dos niveles (bajo y alto) denotados por  $A_1 = 3$  min,  $A_2 = 6$  min y  $B_1 = 600$  rpm,  $B_2 = 1000$  rpm, respectivamente. La respuesta de interés ( $Y$ ) es la cantidad de aditivo. En la tabla 1 se muestran los cuatro tratamientos o puntos del diseño factorial  $2^2$ . En el experimento original cada tratamiento se corrió tres veces (tres réplicas), lo que da un total de 12 corridas del proceso, pero por simplicidad en la última columna de la tabla 1 sólo se han anotado los resultados de la primera réplica.

Tabla 1

$A$ : Tiempo	$B$ : velocidad	$A$ : Tiempo	$B$ : velocidad	$Y$
$A_1$ (bajo)=3	$B_1$ (bajo)=600	$A_1$	$B_1$	17.10
$A_2$ (alto)=6	$B_1$ (bajo)=600	$A_2$	$B_1$	16.26
$A_1$ (bajo)=3	$B_2$ (alto)=1000	$A_1$	$B_2$	18.76
$A_2$ (alto)=6	$B_2$ (alto)=1000	$A_2$	$B_2$	18.16

**3.2. Efecto principal y efecto de interacción.** El efecto de un factor se define como el cambio observado en la variable de respuesta debido a un cambio de nivel de tal factor. En particular, *los efectos principales son los cambios en la media de la variable de respuesta debidos a la acción individual de cada factor*. Matemáticamente el efecto principal de un factor es la diferencia entre la respuesta media observada cuando tal factor estuvo en su nivel alto y la respuesta media observada cuando el factor estuvo en su nivel bajo. Por ejemplo, para los datos de la tabla 1, los efectos principales del tiempo y de la velocidad están dados por

$$\begin{aligned} \text{Efecto } A : \text{ tiempo} &= \frac{16.26+18.16}{2} - \frac{17.10+18.76}{2} = -0.72 \\ \text{Efecto } B : \text{ velocidad} &= \frac{18.76+18.16}{2} - \frac{17.10+16.26}{2} = 1.78. \end{aligned}$$

Por otra parte, se dice que dos factores interactúan entre sí o que tienen un efecto de interacción sobre la variable de respuesta, cuando el efecto de un factor depende del nivel en que se encuentra el otro. Por ejemplo los factores tiempo y velocidad interactúan si el efecto del tiempo es muy diferente en cada nivel de la velocidad, o viceversa, si el efecto de la velocidad depende del tiempo. Veamos esto con los datos de la tabla 1: el efecto del tiempo ( $A$ ) cuando la velocidad ( $B$ ) es baja, está determinado por

$$\text{Efecto } A \text{ (con } B \text{ bajo)} = 16.26 - 17.10 = -0.84$$

y cuando la velocidad es alta, el efecto de A es

$$\text{Efecto A (con B alto)} = 18.16 - 18.76 = -0.6$$

En la práctica, el efecto de interacción entre A y B, denotado por AB, se calcula como la diferencia entre la respuesta media cuando ambos factores se encuentran en el mismo nivel ((bajo, bajo), (alto, alto)), y la respuesta media cuando los factores se encuentran en niveles opuestos ((bajo, alto), (alto, bajo)). Para el ejemplo, el efecto de interacción tiempo  $\times$  temperatura está dado por

$$AB = \frac{17.10 + 18.16}{2} - \frac{16.26 + 18.76}{2} = 0.12$$

Los valores absolutos (sin importar el signo) de los efectos principales y el efecto de interacción son una medida de la magnitud de su efecto sobre la variable de respuesta. Así, entre más grande sea el valor absoluto de un efecto, mayor influencia tendrá sobre la variable de respuesta. Sin embargo, para saber si los efectos son estadísticamente significativos (diferentes de cero) se requiere el análisis de varianza (ANOVA).

#### 4. Diseños factoriales con dos factores

Considere los factores A y B con  $a$  y  $b$  ( $a, b \geq 2$ ) niveles de prueba, respectivamente. Con ellos se puede construir el arreglo o diseño factorial  $a \times b$ , que consiste de  $a \times b$  tratamientos. Se llama réplica cada repetición completa del arreglo factorial. Los diseños factoriales que involucran menos de cuatro factores se corren replicados para poder tener la potencia necesaria en las pruebas estadísticas sobre los efectos de interés, de tal forma que si se hacen  $n$  réplicas, el número total de corridas experimentales es  $n(a \times b)$ .

**4.1. Modelo estadístico.** Con un diseño factorial  $a \times b$  se pueden estudiar los dos efectos individuales y el efecto de interacción de ambos factores. En términos estadísticos, lo que se afirma es que el comportamiento de la respuesta  $Y$  en el experimento con  $k$  réplicas se podrá describir mediante el modelo de efectos,

$$(4.1) \quad \begin{aligned} Y_{ijk} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}; \\ i &= 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n, \end{aligned}$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto debido al  $i$ -ésimo nivel del factor A,  $\beta_j$  es el efecto del  $j$ -ésimo nivel del factor B,  $(\alpha\beta)_{ij}$  representa al efecto de interacción en la combinación  $ij$  y  $\varepsilon_{ijk}$  es el error aleatorio que supone sigue una distribución con media cero y varianza constante  $\sigma^2$  y son independientes entre sí. Para que la estimación de los parámetros en este modelo sea única, se introducen las restricciones:

$$\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = 0.$$

Es decir, *los efectos dados en el modelo son desviaciones respecto de la media global*. Como alternativa, suponga que ambos factores son cuantitativos, el comportamiento de la respuesta se puede describir con el modelo de regresión asociado dado por:

$$(4.2) \quad Y_{ijk} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \varepsilon_{ijk}$$

donde  $\beta_0$  representa a la media general,  $\beta_1$  es el efecto del factor  $X_1 = A$ ,  $\beta_2$  el efecto del factor  $X_2 = B$  y  $\beta_{12}$  representa el efecto de interacción  $X_1 X_2 = AB$ . Al igual que en el modelo de arriba, los  $\varepsilon_{ijk}$  son variables independientes e idénticamente distribuidas  $N(0, \sigma^2)$ . Se recomienda que las variables  $X_i$  se ajusten en la escala codificada:  $-1, 1$  para dos niveles y  $-1, 0, 1$  para tres niveles.

Cabe aclarar sobre los modelos (4.1) y (4.2) que aunque los correspondientes parámetros miden ambos los mismos efectos, no lo hacen en la misma escala. Los  $\beta_1$ , representan la pendiente de la superficie de respuesta que describe el modelo (4.1), y no desviaciones respecto a la media global como los  $\alpha_i$  y  $\beta_j$  del modelo (4.2). Sin embargo, al momento de hacer las pruebas estadísticas la significancia de los efectos correspondientes definidos vía uno u otro modelo debe ser igual. En algunos casos, como en la familia de factoriales  $2^k$ , es fácil establecer la relación entre los dos grupos de parámetros.

En cuanto a interpretación, los modelos (4.1) y (4.2) predicen la misma respuesta  $Y$  sobre los puntos experimentales; pero el segundo modelo debe usarse para predecir el comportamiento de la respuesta en cualquier punto dentro de los rangos de prueba, formando de esta manera una superficie de respuesta. Si los factores tienen al menos tres niveles de prueba el modelo de efectos (4.1) ) permite determinar el mejor tratamiento, mientras que el modelo (4.2) es útil para encontrar el punto óptimo de operación.

**4.2. Hipótesis a evaluar y análisis de varianza.** Como se observa en los modelos anteriores, con un diseño factorial  $a \times b$  interesa estudiar los tres efectos  $A$ ,  $B$  y  $AB$ . Así, en primera instancia se pueden plantear los tres pares de hipótesis siguientes:

$$H_0 : \text{Efecto } A = 0 \quad H_0 : \text{Efecto } B = 0 \quad H_0 : \text{Efecto } AB = 0$$

$$H_A : \text{Efecto } A \neq 0 \quad H_A : \text{Efecto } B \neq 0, \quad H_A : \text{Efecto } AB \neq 0.$$

También estas hipótesis se pueden plantear con los efectos descritos en los modelos (4.1) y (4.2).

$$\begin{array}{ll} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_a = 0 & H_0 : \beta_1 = \beta_2 = \dots = \beta_b = 0 \\ H_A : \alpha_i \neq 0 \text{ para algún } i & H_A : \beta_j \neq 0 \text{ para algún } j \end{array} \quad y$$

$$\begin{array}{l} H_0 : (\alpha\beta)_{ij} = 0 \text{ para todo } ij \\ H_A : (\alpha\beta)_{ij} \neq 0 \text{ para algún } ij. \end{array}$$

Estas hipótesis se prueban mediante la técnica de análisis de varianza. El ANOVA para un diseño factorial  $a \times b$  con  $n$  réplicas resulta de descomponer la variación total como

$$SCT = SCA + SCB + SCAB + SCE$$

donde los respectivos grados de libertad de cada una de ellas son

$$nab - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1).$$

El factor  $(n - 1)$  en los grados de libertad de la suma de cuadrados del error ( $SCE$ ) señala que se necesitan al menos dos réplicas del experimento para poder calcular este componente, y por ende poder construir una tabla de ANOVA. Toda



esa información se sintetiza en la tabla 1

Tabla 1					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Efecto A	SCA	a-1	CMA	$\frac{CMA}{CME}$	$P(F > F_0^A)$
Efecto B	SCB	b-1	CMB	$\frac{CMB}{CME}$	$P(F > F_0^B)$
Efecto AB	SCAB	(a-1)(b-1)	CMAB	$\frac{CMAB}{CME}$	$P(F > F_0^{AB})$
Error	SCE	ab(n-1)	CME		
Total	SCT	abn-1			

Si el *valor P* es menor al nivel de significancia  $\alpha$  prefijado, se concluye que la correspondiente hipótesis es significativa, es decir, ese efecto está o influye a la variable de respuesta; y si el efecto está activo entonces debe tomarse en cuenta en la interpretación de los resultados, buscando mejorar el desempeño de la respuesta  $Y$ .

Recordemos la notación de puntos para representar sumas y medias:  $Y_{...}$  es la suma de todas las observaciones,  $\bar{Y}_{...}$  es la media global,  $Y_{i..}$  es el total en el nivel  $i$  del factor A,  $\bar{Y}_{i..}$  es la media en el nivel  $i$  del factor A.  $Y_{.j.}$  es el total en el nivel  $j$  del factor B. Es decir:

$$\begin{aligned}
 Y_{...} &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} & \bar{Y}_{...} &= Y_{...}/abn \\
 Y_{i..} &= \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} & \bar{Y}_{i..} &= Y_{i..}/bn; i = 1, 2, \dots, a \\
 Y_{.j.} &= \sum_{i=1}^a \sum_{k=1}^n Y_{ijk} & \bar{Y}_{.j.} &= Y_{.j.}/an; j = 1, 2, \dots, b \\
 Y_{ij.} &= \sum_{k=1}^n Y_{ijk} & \bar{Y}_{ij.} &= Y_{ij.}/n \\
 SCT &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - Y_{...}^2/abn & SCA &= \sum_{i=1}^a Y_{i..}^2/bn - Y_{...}^2/abn \\
 SCB &= \sum_{j=1}^b Y_{.j.}^2/an - Y_{...}^2/abn \\
 SCAB &= \sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2/n - Y_{...}^2/abn - SCA - SCB
 \end{aligned}$$

y finalmente, al restar éstas del total, se obtiene la suma de cuadrado

$$SCE = SCT - SCA - SCB - SCAB.$$

**EJEMPLO 10. (Factorial  $4 \times 3$ )** Consideremos un experimento en el que se quiere estudiar el efecto de los factores velocidad de alimentación y profundidad de corte sobre el acabado de un metal. Aunque los factores son de naturaleza continua, en este proceso sólo se pueden trabajar en 3 y 4 niveles, respectivamente. Por ello, se decide correr un factorial completo  $4 \times 3$  con tres réplicas, que permitirá obtener toda la información relevante en relación al efecto de estos factores sobre el acabado. Aleatorizando las 36 pruebas se obtienen los datos en la tabla 2 El acabado ( $Y$ ) está en unidades de gramos e interesa minimizar su valor.

El modelo estadístico efectos para el diseño factorial  $4 \times 3$  que está en la tabla 2, está dado por

$$Y_{ijk} = \mu + \gamma_i + \delta_j + (\gamma\delta)_{ij} + \varepsilon_{ijk}; \quad i = 1, 2, 3, 4; \quad j = 1, 2, 3; \quad k = 1, 2, 3$$

donde  $\gamma_i$  es el efecto del nivel  $i$  factor  $A$ ,  $\delta_j$  representa el efecto del nivel  $j$  del factor  $B$ ,  $(\gamma\delta)_{ij}$  es la interacción de ambos en el tratamiento  $ij$  y  $k = 1, 2, 3$  denota las repeticiones de cada tratamiento, es decir, se hicieron  $n = 3$  réplicas este ejemplo. Este modelo es el caso particular del modelo dado en 4.1, tomando  $a = 4, b = 3$  y  $n = 3$ . Las hipótesis de interés para los tres factores en el modelo anterior son:

$$\begin{aligned} H_0 : & \text{Efecto de profundidad (A)} = 0 \\ H_A : & \text{Efecto de profundidad (A)} \neq 0 \\ H_0 : & \text{Efecto de velocidad (B)} = 0 \\ H_A : & \text{Efecto de velocidad (B)} \neq 0, \\ H_0 : & \text{Efecto de profundidad} \times \text{velocidad (AB)} = 0 \\ H_A : & \text{Efecto de profundidad} \times \text{velocidad (AB)} \neq 0 \end{aligned}$$

y los datos aparecen en la siguiente tabla:

Tabla 2	A: Profundidad (pulg.)					Y <sub>j.</sub>
	0.15	0.18	0.21	0.24		
B: veloc	74	79	82	99		
0.20	64   198	68   220	88   262	104   299		979
	60	73	92	96		
0.25	92	98	99	104		
	86   266	104   290	108   302	110   313		1171
	88	88	95	99		
0.30	99	104	108	114		
	98   299	99   298	110   317	111   332		1246
	102	95	99	107		
Y <sub>i.</sub>	763	808	881	944		Y...=3396

El ANOVA para los tres efectos  $A$ ,  $B$  y  $AB$  sin desglosar se obtiene observando que el factor  $A$  tiene cuatro niveles y el factor  $B$  tiene tres niveles, por lo que las correspondientes sumas de cuadrados tendrán 3, 2 y 6 grados de libertad, respectivamente. Para calcular las sumas de cuadrados se requiere primero obtener los totales por nivel de cada factor ( $Y_{i.}$  y  $Y_{.j.}$ ) y en cada combinación de niveles ( $Y_{ij.}$ ). Los primeros se muestran en los márgenes derecho e inferior de la tabla 2 de los datos y los segundos se encuentran encerrados en un cuadro dentro de cada casilla de esta misma tabla. Las sumas de cuadrados dadas

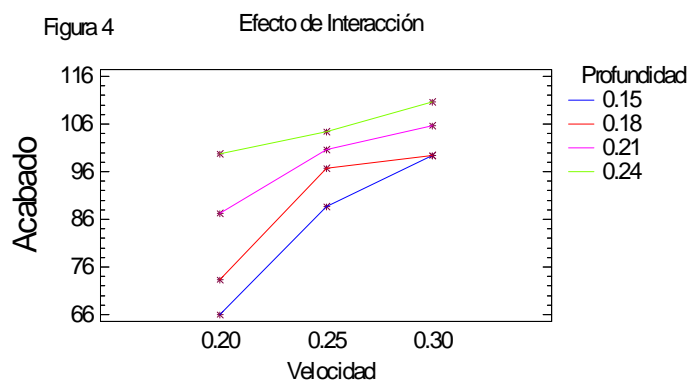
$$SCA = (763^2 + 808^2 + 881^2 + 944^2) / 3 \times 3 - 3396^2 / 4 \times 3 \times 3 = 2125.1$$

$$SCB = (979^2 + 1171^2 + 1246^2) / 4 \times 3 - 3396^2 / 4 \times 3 \times 3 = 3160.5$$

$$SCAB = (198^2 + 220^2 + \dots + 332^2) / 3 - 3396^2 / 4 \times 3 \times 3 - 2125.1 - 3160.5 = 557.07$$

$$SCT = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - Y_{...}^2 / abn = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - 3396^2 / 4 \times 3 \times 3 = 6532.0$$

$$SCE = SCT - SCA - SCB - SCAB = 6532.0 - 2125.1 - 3160.5 - 557.07 = 689.33$$



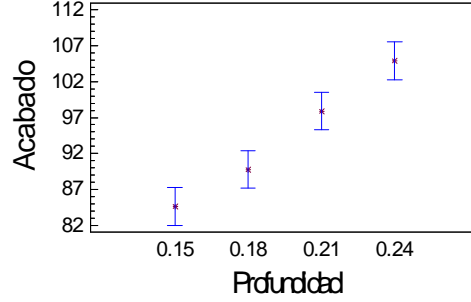
y la tabla ANOVA

Tabla 3. ANOVA					
<i>F. V</i>	<i>SC</i>	<i>GL</i>	<i>CM</i>	$F_0$	<i>valor-p</i>
<i>B: veloc</i>	3160.5	2	1580.25	55.02	0.0000
<i>A: profu</i>	2125.1	3	708.37	24.66	0.0000
<i>AB</i>	557.07	6	92.84	3.23	0.018
<i>Error</i>	689.33	24	28.72		
<i>Total</i>	6532.0	35			

Del ANOVA se concluye que los tres efectos  $A : vel$ ,  $B : prof$  y  $AB$  están activos o influyen en el acabado. Dado que el efecto de interacción  $AB$  resulta significativo, prácticamente toda la información relevante del experimento se puede apreciar en su representación gráfica (figura 4). Notese que aparecen tantas líneas como niveles tenga el factor que se dibuja en la parte de arriba, que en este caso es la profundidad con sus 4 niveles, que se denotan con una escala de  $-1$  a  $1$ . La significancia de la interacción detectada por el ANOVA se observa en el hecho de que las líneas en la figura 4 no se mantienen paralelas a lo largo de los niveles del factor  $A$  (tienen diferente pendiente). Como lo que interesa es minimizar la variable de respuesta, se observa que a mayor velocidad y profundidad hay una tendencia a obtener peores acabados, además se ve que cuando se tiene velocidad alta ( $A^+$ ) el efecto de profundidad es menor (vease la dispersión de las líneas en la figura cuando la velocidad es alta). Por tanto, las condiciones de operación o tratamiento que convienen es velocidad y profundidad bajas ( $A^-$ ,  $B^-$ ).

**4.3. Comparaciones Múltiples.** Estas técnicas se aplican en diseños factoriales en más de dos niveles en cada factor, para decidir en específico cuáles de los niveles probados son estadísticamente diferentes entre sí. Por facilidad, se notan los cuatro niveles de la profundidad ( $A$ ) como  $A_1$ ,  $A_2$ ,  $A_3$  y  $A_4$  y los cuatro niveles de la velocidad ( $B$ ) como  $B_1$ ,  $B_2$ , y  $B_3$ . Entonces, las de hipótesis para comparar

Figura 5a Gráficas de medias: Factor A



las medias del factor profundidad y velocidad son:

$$\begin{aligned}
 &H_0 : \mu_{A_1} = \mu_{A_2} \quad H_0 : \mu_{A_1} = \mu_{A_3} \quad H_0 : \mu_{A_1} = \mu_{A_4} \\
 &H_A : \mu_{A_1} \neq \mu_{A_2} \quad H_A : \mu_{A_1} \neq \mu_{A_3} \quad H_A : \mu_{A_1} \neq \mu_{A_4} \\
 &H_0 : \mu_{A_2} = \mu_{A_3} \quad H_0 : \mu_{A_2} = \mu_{A_4} \quad H_0 : \mu_{A_3} = \mu_{A_4} \\
 &H_A : \mu_{A_2} \neq \mu_{A_3} \quad H_A : \mu_{A_2} \neq \mu_{A_4} \quad H_A : \mu_{A_3} \neq \mu_{A_4} \\
 &H_0 : \mu_{B_1} = \mu_{B_2} \quad H_0 : \mu_{B_1} = \mu_{B_3} \quad H_0 : \mu_{B_2} = \mu_{B_3} \\
 &H_A : \mu_{B_1} \neq \mu_{B_2} \quad H_A : \mu_{B_1} \neq \mu_{B_3} \quad H_A : \mu_{B_2} \neq \mu_{B_3}
 \end{aligned}$$

Usando el método LSD, se prueban las hipótesis del factor  $A$  ignorando por el momento la interacción. La diferencia mínima significativa para comparar los niveles  $i$  y  $l$  del factor  $A$ , está dada por

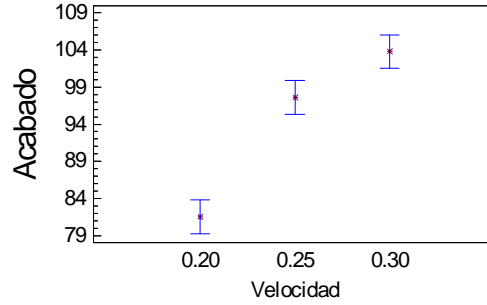
$$\begin{aligned}
 LSD(A) &= t_{\alpha/2, ab(n-1)} \cdot \sqrt{\left(\frac{1}{n_{A_i}} + \frac{1}{n_{A_l}}\right) CME} \\
 &= 2.064 \cdot \sqrt{\left(\frac{2}{9}\right) 28.72} = 5.21
 \end{aligned}$$

donde  $n_{A_i}, n_{A_l}$  son el total de observaciones en el nivel  $i$  y  $l$  del factor  $A$ , que se está comparando. como el diseño es balanceado con cuatro niveles para el factor  $A$ ,  $n_{A_i} = n_{A_l} = abn/4 = 9$

Tabla 4. LSD (sin interacción)			
Factor A	$n_i$	LS Mean	Grupos
0.15	9	84,7778	X
0.18	9	89,7778	X
0.21	9	97,8889	X
0.24	9	104,889	X

donde sólo la primera diferencia resulta no significativa, es decir, se acepta  $H_0 : \mu_{A_1} = \mu_{A_2}$  ; en las cinco comparaciones restantes se rechaza  $H_0$ . Las conclusiones, para ambos factores usando el método  $LSD$  se observan en las gráficas de medias de la figura 5a y figura 5b, donde no se toma en cuenta el efecto de interacción detectado en el ANOVA

Figura 5b Gráfica de medias: Factor B



**4.4. Tomando en cuenta la interacción.** Del análisis que ignora el efecto de interacción se concluyó que  $\mu_{A_2} \neq \mu_{A_3}$ ;  $\mu_{A_2} \neq \mu_{A_4}$ ;  $\mu_{A_3} \neq \mu_{A_4}$ , es decir, las tres últimas profundidades son diferentes entre sí. Sin embargo, esta conclusión cambia al tomar en cuenta que el factor velocidad interactúa con la profundidad. Si observamos el efecto de interacción en la figura 4, es fácil notar que las medias de las tres últimas profundidades están más cercanas entre sí cuando la velocidad está en su nivel intermedio que cuando está en su nivel bajo. Veámoslo de manera analítica en la velocidad intermedia (B), donde las medias muestrales del factor A : prof en la velocidad intermedia son (Tabla 2)  $\bar{Y}_{12\cdot} = \frac{266}{3} = 88.66$ ;  $\bar{Y}_{22\cdot} = \frac{290}{3} = 96.66$ ;  $\bar{Y}_{32\cdot} = \frac{302}{3} = 100.66$ ;  $\bar{Y}_{42\cdot} = \frac{313}{3} = 104.33$ .

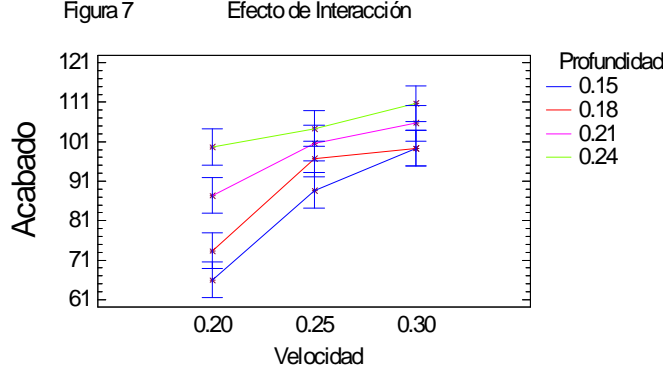
Entonces para comparar estas medias, la diferencia mínima significativa está dada por

$$\begin{aligned} LSDB_2(A) &= t_{\alpha/2, ab(n-1)} \cdot \sqrt{\left(\frac{1}{n} + \frac{1}{n}\right) CME} \\ &= 2.064 \cdot \sqrt{\left(\frac{2}{3}\right) 28.72} = 9.03 \end{aligned}$$

donde  $n$  es el número de réplicas de los tratamientos a comparar. Note que la diferencia entre esta expresión para la LSD y la que no toma en cuenta la interacción, está precisamente en el factor dentro de la raíz cuadrada que acompaña al CME, ya que ahí se anota el inverso del número de observaciones con los que se construyeron las medias muestrales con las que se calculan las diferencias. En el ejemplo  $n = 3$ .

Tabla 5. LSD (con interacción)			
Factor A	$n_i$	LS Mean	Grupos
0.15	9	88.66	X
0.18	9	96.66	X X
0.21	9	100.66	X
0.24	9	104.33	X

$$\begin{aligned} |\bar{Y}_{12\cdot} - \bar{Y}_{22\cdot}| &= |88.66 - 96.66| = 8 < LSDB_2(A) \\ |\bar{Y}_{12\cdot} - \bar{Y}_{32\cdot}| &= |88.66 - 100.66| = 12 > LSDB_2(A) \\ |\bar{Y}_{12\cdot} - \bar{Y}_{42\cdot}| &= |88.66 - 104.33| = 15.7 > LSDB_2(A) \end{aligned}$$



$$\begin{aligned} |\bar{Y}_{22\cdot} - \bar{Y}_{32\cdot}| &= |96.66 - 100.66| = 4 < LSDB_2(A) \\ |\bar{Y}_{22\cdot} - \bar{Y}_{42\cdot}| &= |96.66 - 104.33| = 7.7 < LSDB_2(A) \\ |\bar{Y}_{32\cdot} - \bar{Y}_{42\cdot}| &= |100.66 - 104.33| = 3.7 < LSDB_2(A) \end{aligned}$$

Por tanto, al tomar en cuenta el efecto de interacción AB se concluye que las tres profundidades mayores son estadísticamente iguales en nivel intermedio de la velocidad, lo contrario a lo que se había concluido en el análisis que ignora el efecto de interacción. La gráfica de interacción con intervalos LSD sobrepuestos se muestra en la figura 7. Note que en el nivel intermedio de la velocidad, los intervalos de confianza para las medias de las tres profundidades superiores se traslapan, indicio de que son estadísticamente iguales, como se acaba de concluir de manera analítica. Si se construyeran los intervalos con el método de Tukey, se traslaparían más por ser más anchos.

## 5. Diseños factoriales con tres factores

Cuando se tienen tres factores (A, B y C) y el número de niveles de prueba en cada uno de ellos son  $a$ ,  $b$  y  $c$ , se puede construir el arreglo factorial  $a \times b \times c$  que consiste de  $a \times b \times c$  tratamientos o puntos experimentales. Entre los arreglos de este tipo que se utilizan con frecuencia en aplicaciones diversas se encuentran: el factorial  $2^3$ , el factorial  $3^3$  y los factoriales mixtos con no más de cuatro niveles en dos de los factores, por ejemplo, el factorial  $4 \times 3 \times 3$  el factorial  $4 \times 4 \times 2$ , por mencionar dos de ellos.

**5.1. Modelo Estadístico.** En un diseño factorial  $a \times b \times c$  se supone que el comportamiento de la respuesta  $Y$  puede describirse mediante el modelo de efectos dado por

$$\begin{aligned} Y_{ijkl} &= \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl} \\ i &= 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, c; \quad l = 1, 2, \dots, n \end{aligned}$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto del nivel  $i$ -ésimo del factor A,  $\beta_j$  es el efecto del nivel  $j$  del factor B y  $\gamma_k$  es el efecto del nivel  $k$  en el factor C;  $(\alpha\beta)_{ij}$ ,  $(\alpha\gamma)_{ik}$ , y  $(\beta\gamma)_{jk}$  representan los efectos de interacción dobles (de dos factores) en los niveles

$ij, ik, jk$ , respectivamente, y  $(\alpha\beta\gamma)_{ijk}$  es el efecto de interacción triple en la combinación o punto  $ijk$ ;  $\varepsilon_{ijkl}$ , representa el error aleatorio en la combinación  $ijkl$  y  $l$  son las repeticiones o réplicas del experimento. Todos los efectos cumplen la restricción de sumar cero, es decir, son desviaciones respecto a la media general. De manera alternativa, se tiene el modelo de regresión dado por

$$\begin{aligned}
 Y_{ijkl} &= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2j} + \beta_3 X_{3k} + \beta_{12} X_{1i} X_{2j} + \beta_{13} X_{1i} X_{3k} \\
 &\quad + \beta_{23} X_{2j} X_{3k} + \beta_{123} X_{1i} X_{2j} X_{3k} + \varepsilon_{ijkl} \\
 i &= 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, c; \quad l = 1, 2, \dots, n
 \end{aligned}$$

**5.2. Hipótesis de interés.** El estudio factorial de tres factores (A, B y C) permite investigar los efectos: A, B, C, AB, AC, BC y ABC, donde el nivel de desglose o detalle con el que pueden estudiarse depende del número de niveles utilizado factor. Por ejemplo, si un factor se prueba en dos niveles, todo su efecto marginal (individual) es lineal, o sea que su efecto individual no se puede descomponer; pero si tuviera tres niveles, su efecto marginal se puede descomponer en una parte lineal y otra cuadrática pura. En resumen, se tienen siete efectos de interés sin considerar desglose, y con ellos se pueden plantear las siete hipótesis nulas  $H_0 : \text{Efecto } A = 0$ ,  $H_0 : \text{Efecto } B = 0$ ,  $H_0 : \text{Efecto } C = 0$ ,  $H_0 : \text{Efecto } AB = 0$ ,  $H_0 : \text{Efecto } AC = 0$ ,  $H_0 : \text{Efecto } BC = 0$ ,  $H_0 : \text{Efecto } ABC = 0$ , cada una aparejada con su correspondiente hipótesis alternativa. El ANOVA para probar estas hipótesis se muestra en la tabla 5.

Al efecto cuyo valor-p sea menor al valor especificado para  $\alpha$ , se declara estadísticamente significativo o se dice que está activo. El ANOVA de tres factores dado en la tabla 5. tiene cuatro renglones adicionales, por los nuevos cuatro efectos que pueden estudiarse. Las sumas de cuadrados son muy similares a las obtenidas para dos factores.

Tabla 5					
variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Efecto A	SCA	a-1	CMA	$\frac{CMA}{CME}$	$P(F > F_0^A)$
Efecto B	SCB	b-1	CMB	$\frac{CMB}{CME}$	$P(F > F_0^B)$
Efecto C	SCC	c-1	CMC	$\frac{CMC}{CME}$	$P(F > F_0^C)$
Efecto AB	SCAB	(a-1)(b-1)	CMAB	$\frac{CMAB}{CME}$	$P(F > F_0^{AB})$
Efecto AC	SCAC	(a-1)(c-1)	CMAC	$\frac{CMAC}{CME}$	$P(F > F_0^{AC})$
Efecto BC	SCBC	(b-1)(c-1)	CMBC	$\frac{CMBC}{CME}$	$P(F > F_0^{BC})$
Efecto ABC	SCABC	(a-1)(b-1)(c-1)	CMABC	$\frac{CMABC}{CME}$	$P(F > F_0^{ABC})$
Error	SCE	abc(n-1)	CME		
Total	SCT	abcn-1			

donde:

$$\begin{aligned}
SCT &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c \sum_{l=1}^n Y_{ijkl}^2 - Y_{....}^2 / N \\
SCB &= \sum_{j=1}^b Y_{.j..}^2 / acn - Y_{....}^2 / N \\
SCAB &= \sum_{i=1}^a \sum_{j=1}^b Y_{ij..}^2 / cn - Y_{....}^2 / N - SCA - SCB \\
SCBC &= \sum_{j=1}^b \sum_{k=1}^c Y_{.jk.}^2 / an - Y_{....}^2 / N - SCB - SCC \\
SCA &= \sum_{i=1}^a Y_{i...}^2 / bcn - Y_{....}^2 / N \\
SCC &= \sum_{k=1}^c Y_{..k.}^2 / acn - Y_{....}^2 / N \\
SCAC &= \sum_{i=1}^a \sum_{k=1}^c Y_{i.k.}^2 / bn - Y_{....}^2 / N - SCA - SCC \\
SCABC &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c Y_{ijk.}^2 / n - Y_{....}^2 / N - SCAB - SCAC - SCBC
\end{aligned}$$

EJEMPLO 11. **El experimento** Se desea investigar el efecto del tipo de suspensión (A), abertura de malla (B) y temperatura de ciclaje (C) en el volumen de sedimentación Y(%) de una suspensión. Para ello se decide correr un experimento factorial  $3 \times 2 \times 2$  con seis réplicas, y las observaciones obtenidas en las 72 corridas experimentales se muestran en la siguiente tabla:

Tabla 2	A <sub>1</sub>						A <sub>2</sub>						A <sub>3</sub>					
	B <sub>1</sub>			B <sub>2</sub>			B <sub>1</sub>			B <sub>2</sub>			B <sub>1</sub>			B <sub>2</sub>		
C <sub>1</sub>	60	75	75	67	73	73	62	68	65	71	80	80	70	71	75	75	75	75
	86	70	70	67	68	68	76	65	65	72	80	80	76	68	73	75	75	77
C <sub>2</sub>	55	53	53	52	52	57	44	44	45	60	60	60	52	51	50	56	55	57
	55	55	55	52	54	54	48	48	45	67	67	65	52	48	54	59	50	55

Los niveles de prueba para cada fator, tanto en unidades originales como en unidades codificadas se muestran en la siguiente tabla:



[2, 1, 1, , , c, ,]	
U. originales	
U. codificadas	
Factor	
<i>Bajo</i>	
<i>Medio</i>	
<i>Alto</i>	
<i>Bajo</i>	
<i>Medio</i>	
A: Tipo de suspensión	
$A_1$	
$A_2$	
$A_3$	
-1	
0	
B: Abertura de malla	
40	
-	
60	
-1	
-	
C: Temperatura	
0	
-	
30	
-1	

Tabla 5. ANOVA Volumen de sedimentación					
variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
A: Tipo de suspensi	13,8611	2	6.93056	0.49	0.6126
B: Abertura de mall	480,5	1	480.5	34.14	0.0000
C: Temperatura	60.8672	1	60.8672	432, 43	0.0000
Efecto AB	788,25	2	394,125	28,00	0.0000
Efecto AC	40,8611	2	20,4306	1,45	0,2421
Efecto BC	56,8889	1	56,8889	4,04	0,0487
Efecto ABC	31,0278	2	15,5139	1,11	0,3375
Error	841,667	60	14,0278		
Total	8339,78	71			

## 6. Ejercicios: Diseños Factoriales

EJERCICIO 49. ¿Qué es un experimento factorial completo?

EJERCICIO 50. ¿Cuántos efectos se pueden estudiar con un factorial  $4 \times 3 \times 2$ ? Bosqueje su tabla de análisis de varianza.

EJERCICIO 51. Mencione al menos tres ventajas de la experimentación factorial sobre la estrategia de mover un factor a la vez.

EJERCICIO 52. *¿Cuál es la implicación práctica de utilizar tres niveles de prueba en lugar de dos en un factor dado?*

EJERCICIO 53. *¿Por qué no tiene sentido utilizar el modelo de regresión cuando los factores son cualitativos? Si fueran cuantitativos, ¿qué se gana con el modelo de regresión en relación al modelo de efectos?*

EJERCICIO 54. *¿Cómo se construye la gráfica de un efecto de interacción doble? ¿Cómo se interpreta?*

EJERCICIO 55. *De los tres supuestos del modelo, ¿cuál puede afectar más el análisis en caso de no cumplirse?*

EJERCICIO 56. *En caso de no cumplirse los supuestos de normalidad y varianza constante, ¿qué se puede hacer para evitar problemas con el análisis y resultados obtenidos?*

EJERCICIO 57. *¿Con base en qué se puede encontrar una transformación adecuada de la respuesta, cuando no se cumplen los supuestos?*

EJERCICIO 58. *¿Cuáles son los supuestos del modelo en un diseño factorial y con cuáles gráficas de residuos se puede verificar cada uno de supuestos?*

EJERCICIO 59. *En la pregunta anterior, ¿cómo se vería en las gráficas un punto muy alejado o aberrante?*

EJERCICIO 60. *¿Qué significa que el modelo sea de efectos aleatorios? ¿En qué cambian las hipótesis de interés en factor aleatorio respecto de uno fijo?*

EJERCICIO 61. *A continuación se muestra parte del ANOVA para un diseño factorial  $3 \times 5$  con dos réplicas, el factor A con tres niveles y el B con cinco.*

Tabla 5. ANOVA Volumen de sedimentación					
variabilidad	SC	GL	CM	$F_0$	valor-p
A	800				
B	900				
AB	300				
Error	400				
Total					

(a) Suponga efectos fijos, anote el modelo estadístico apropiado y formule las hipótesis a probar para este experimento. (b) Agregue en esta tabla los grados de libertad, el cuadrado medio y la razón F para cada una de las fuentes de variación. (c) Con la información disponible se pueden hacer conjeturas sobre cuáles de las fuentes de variación son significativas estadísticamente? Argumente.

EJERCICIO 62. *Conteste todo el ejercicio anterior, pero ahora suponiendo que ambos factores son aleatorios.*

EJERCICIO 63. *Conteste todos los incisos del ejercicio 62 pero ahora suponga que el factor A es fijo y el factor B es aleatorio.*

EJERCICIO 64. En la tabla siguiente están los datos de un diseño factorial  $3 \times 2$  con tres réplicas, conteste:

	niveles de <b>A</b>		
<i>B</i>	<i>A</i> <sub>1</sub>	<i>A</i> <sub>2</sub>	<i>A</i> <sub>3</sub>
<i>B</i> <sub>1</sub>	44	49	52
	34	84	58
	30	43	62
<i>B</i> <sub>2</sub>	62	68	69
	56	70	75
	58	58	65

(a) Obtenga la tabla de análisis de varianza y anote las principales conclusiones. (b) Sin tomar en cuenta el posible efecto de interacción, obtenga la diferencia mínima significativa (LSD) para comparar los niveles de factor *A* y la LSD para comparar los niveles del factor *B*. (c) ¿Cuál sería la LSD exacta tomando en cuenta la interacción?

**EJERCICIO 65.** Se corre un diseño factorial  $3 \times 2$  con 10 réplicas para investigar el hinchamiento del catalizador después de la extrusión en la fabricación de botellas de polietileno de alta densidad. El catalizador se utiliza en la obtención de dicho polietileno. Los factores investigados son: *A*: catalizador (con tres niveles, *A*<sub>1</sub>, *A*<sub>2</sub> y *A*<sub>3</sub>) y *B*: molde (con dos niveles, *B*<sub>1</sub> y *B*<sub>2</sub>). Los datos obtenidos se muestran en tabla siguiente. (a) Plantee las hipótesis de interés en este problema, y el modelo estadístico correspondiente. (b) Construya la tabla de análisis de varianza y determine cuáles efectos están activos. (c) Dibuje las gráficas de medias para los dos efectos principales con los métodos LSD y de Tukey. Compare los resultados de ambos métodos. (d) Haga la gráfica de interacción con intervalos de confianza sobrepuestos. (e) Determine el mejor tratamiento. ¿Cuál es el hinchamiento predicho en el mejor tratamiento? (f) Verifique los supuestos de normalidad y varianza constante. (g) Utilice la gráfica de residuos contra factores para detectar posibles efectos sobre la dispersión del hinchamiento. ¿En cuál molde parece que es menor la dispersión?

	Catalizador: <i>A</i>											
<i>B</i>	<i>A</i> <sub>1</sub>				<i>A</i> <sub>2</sub>				<i>A</i> <sub>3</sub>			
<i>B</i> <sub>1</sub>	93	91	90	90	92	91	92	91	95	94	94	96
	92	92	91		94	90	92		94	97	95	
	90	91	93		90	91	92		94	96	94	
<i>B</i> <sub>2</sub>	88	88	87	87	90	88	88	89	91	90	90	91
	88	87	87		89	90	88		90	91	91	
	87	88	87		88	89	88		92	89	91	

EJERCICIO 66. Para mejorar la resistencia a la torsión de las adhesiones de componentes electrónicos sobre placas, se estudiaron dos tipos de pegamentos (*A* y *B*) y tres temperaturas de curado (60, 80 y 100°C). En cada combinación se analizaron dos componentes y los resultados obtenidos son los siguientes: a) Plantee las hipótesis de interés en este problema, y el modelo estadístico correspondiente. b) Construya el ANOVA y decida cuáles efectos están activos. c) Dibuje las gráficas de efectos y determine con ellas el mejor tratamiento. (d) Estime la resistencia a

la torsión en el mejor tratamiento. (e) Verifique residuos.

	Curado		
	60	80	100
Pegamento A	2.5	3.8	4.0
	2.8	3.4	4.2
Pegamento B	1.6	3.2	4.3
	1.22	2.8	4.7

EJERCICIO 67. Se desea investigar de qué manera afecta el tiempo de curado y el tipo de acelerante a la resistencia de caucho vulcanizado. Se realiza un experimento y se obtienen los siguientes datos:

Tiempo de cura a 14°C (minutos)	Curado		
	A	B	C
40	3900	4300	4100
	3600	3700	3700
60	4100	4200	4000
	3500	3900	3900
80	4000	4300	3600
	3800	3600	3800

(a) Señale el nombre del diseño de experimento utilizado y su modelo estadístico. (b) Formule claramente todas las hipótesis que se pueden probar (c) Realice el análisis estadístico apropiado para probar las hipótesis que formuló. (d) ¿Hay algún tiempo de cura que es mejor para aumentar la resistencia? Argumente. (e) ¿Algún acelerante es mejor? Explique. (f) Hay alguna combinación de tiempo y acelerante que sea mejor. (g) Explique gráficamente cómo se obtuvo en la computadora el y valor-p para tiempo de cura. (h) Verifique que se cumplan los supuestos. En caso de que no se cumpliera el supuesto de varianza constante para el tiempo de cura, ¿qué significaría eso y cómo pudiera corregirse?

EJERCICIO 68. En una fábrica de aceites vegetales comestibles la calidad se ve afectada por la cantidad de impurezas dentro del aceite, ya que éstas causan oxidación, y ello repercute a su vez en las características de sabor y color del producto final. El proceso de "blanqueo" que consiste básicamente en elevar la temperatura del aceite, agregar cierta cantidad de arcilla (arena), y agitar durante cierto tiempo para que las impurezas se adhieran a la arcilla; de esta manera al bajar la temperatura del aceite y dejarlo en reposo, la arcilla junto con las impurezas se depositen en el fondo del tanque de aceite. Una forma de medir la eficacia del proceso de blanqueo es midiendo el color del aceite. De acuerdo al método de medición que se aplica en esta planta, consideran que un aceite se le eliminó las suficientes impurezas si el color es mayor a 4.8 (a mayores valores más claro el aceite). Cuando los lotes del proceso de blanqueo no cumplen con los requerimientos de blancura, la solución del problema implica reprocesos, aumento del tiempo de ciclo, entre otras cosas. Para generar una primera aproximación a la solución del problema se decide estudiar la temperatura y el porcentaje de arcilla a nivel laboratorio inicialmente.

El diseño y los datos de las pruebas experimentales se muestran a continuación.

Temperatura	Porcentaje de arcilla			
	0.8	0.9	1.0	1.1
90	5.8	5.4	4.9	4.5
	5.9	5.5	5.1	4.4
100	5.0	4.8	4.6	4.1
	4.9	4.7	4.4	4.3
110	4.7	4.4	4.1	3.7
	4.6	4.4	4.0	3.6

(a) Construya el modelo estadístico y formule las hipótesis pertinentes. (b) ¿Cuál es el nombre del diseño utilizado? (c) Normalmente, a condiciones reales se utiliza 1.1% de arcilla y 100 grados de temperatura. ¿Por qué cree que se han elegido precisamente los niveles de prueba para el experimento? (d) Realice un análisis de varianza para probar las hipótesis y obtenga conclusiones. (e) Apoyándose en las gráficas de efectos, ¿cuál es la relación general entre el color y los factores controlados en su rango de experimentación? (f) A partir de la gráfica de interacciones, ¿cree que haya un efecto no lineal? Apoye su respuesta con el análisis de varianza apropiado. (g) Considerando que el nivel mínimo aceptable de blancura es de 4.8, ¿qué tratamiento utilizaría? (h) ¿Vale la pena ahora plantear el estudio a condiciones reales? (i) ¿Qué cambio le haría al experimento si lo corre a condiciones reales?

EJERCICIO 69. En un laboratorio de microbiología se realiza un experimento para investigar si influye el tipo de verdura (lechuga-L, cilantro-C, zanahoria-Z) y la temperatura (8 y 20° C) de almacenamiento en la sobrevivencia del vidrio colerae. Se hicieron varias réplicas. El porcentaje de sobrevivencia obtenido después de 24 horas de inoculado el alimento se muestra a continuación.

Alimento	Temperatura							
	20				8			
L	13.1	15.0	33.6	35.5	6.2	28.5	41.0	35.9
	42.0	11.1	12.8		25.0	23.8	79.0	41.6
C	19.0	19.0	66.6	66.6	84.3	68.7	68.7	30.5
	11.0	11.0	49.0	49.0	30.5	11.0	11.0	20.0
Z	1.20	1.20	0.20	0.10	25.8	21.8	16.0	16.0
	0.30	0.20	0.10	0.40	20.1	15.4	13.3	25.2
	0.20	0.30						

(a) Señale el nombre del diseño que se ha empleado y formule las hipótesis que pueden ser probadas. (b) Haga un análisis de varianza e interprételo con detalle. (c) Verifique el supuesto de igual varianza entre los tratamientos (varianza constante). ¿Se cumple satisfactoriamente? (d) En caso de que no se cumpla el supuesto anterior, ¿cómo afecta esto a sus conclusiones?

EJERCICIO 70. En una empresa alimenticia se desean evaluar cuatro antioxidantes, a través de su efecto en un aceite vegetal. El propósito es seleccionar el producto que retrase más la oxidación. Las pruebas se hacen a condiciones de estrés, midiendo como variable de respuesta al índice de peróxidos. Diferentes unidades

experimentales se evalúan a diferentes tiempos. Los datos obtenidos se muestran a continuación (en el control no se agrega ningún antioxidante).

Producto	Tiempo					
	4 horas		8 horas		12 horas	
Control	3.84	3.72	27.63	27.58	39.95	39.00
A	4.00	3.91	22.00	21.83	46.20	45.60
B	3.61	3.61	21.94	21.85	43.58	42.98
C	3.57	3.50	20.50	20.32	45.14	44.89
D	3.64	3.61	20.30	20.19	44.36	44.02

(a) Señale los factores controlados y la variable de respuesta. (b) Formule el modelo estadístico apropiado al problema y las hipótesis estadísticas que se pueden probar. (c) Haga un análisis de varianza y observe los aspectos más relevantes. (d) ¿Los supuestos del modelo se cumplen? (e) Considerando que a menor índice de peróxidos mejor es el producto, ¿hay algún producto que sea mejor estadísticamente?

EJERCICIO 71. Se cree que la adhesividad de un pegamento depende de la presión y de la temperatura al ser aplicado. Se realiza un experimento factorial con ambos factores fijos.

Presión (lb/pulg <sup>2</sup> )	Temperatura (°F)		
	250	260	270
120l	9.60	11.28	9.00
130	9.69	10.10	9.57
140	8.43	11.01	9.03
150	9.98	10.44	9.80

(a) Formule las hipótesis y el modelo estadístico que se desea probar. (b) Analice los datos y obtenga las conclusiones apropiadas. (c) ¿Se puede analizar si hay interacción entre los dos factores controlados? (d) Verifique residuos.

EJERCICIO 72. Se estudia el rendimiento de un proceso químico. Se piensa que las dos variables más importantes son la presión y la temperatura. Se seleccionan tres niveles de cada factor y se lleva a cabo un experimento factorial con dos réplicas. Los datos del rendimiento son:

Temperatura (°C)	Presión (psig)		
	200	215	230
150	90.4	90.7	90.2
	90.2	90.6	90.4
160	90.1	90.5	89.9
	90.3	90.6	90.1
170	90.5	90.8	90.4
	90.7	90.9	90.1

(a) Analizar los datos y sacar conclusiones. Utilizar  $\alpha = 0.05$ . (b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo. (c) ¿Bajo qué condiciones debería operarse este proceso? (d) Calcular la estimación de un intervalo de confianza de 95% de la diferencia media en la respuesta para velocidades de alimentación de 0.20 y 0.25 pulg/min.

EJERCICIO 73. Un ingeniero sospecha que el acabado superficial de una pieza metálica se afecta por la velocidad de alimentación y la profundidad de corte. Selecciona tres velocidades de alimentación y cuatro profundidades de corte. Después realiza un experimento factorial y obtiene los siguientes datos:

Velocidad (pulg/min)	Profundidad (pulg)			
	0.15	0.18	0.20	0.25
0.20	74	79	82	99
	64	68	88	104
	60	73	92	96
0.25	92	98	99	104
	86	104	108	110
	88	88	95	99
0.30	99	104	108	114
	98	99	110	111
	102	95	99	107

(a) Analizar los datos y sacar conclusiones. Utilizar  $\alpha = 0.05$ . (b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo. (c) Obtener estimaciones puntuales del acabado superficial promedio con cada velocidad de alimentación. (d) Encontrar los valores  $P$  para las pruebas del inciso (a).

EJERCICIO 74. En un artículo de Industrial Quality Control se describe un experimento para investigar el efecto del tipo de cristal y del tipo de fósforo sobre la brillantez de un cinescopio. La variable de respuesta es la corriente (en microamperes) necesaria para obtener un nivel de brillantez específico. Los datos son los siguientes:

Tipo de cristal	Tipo de fósforo		
	1	2	3
1	280	300	290
	290	310	285
	285	295	290
2	230	260	220
	235	240	225
	240	235	230

a) ¿Existe algún indicio de que alguno de los dos factores influye en la brillantez? Utilizar  $\alpha = 0.05$ . b) ¿Los dos factores interactúan? Utilizar  $\alpha = 0.05$ . c) Analizar los residuales de este experimento.

EJERCICIO 75. Johnson y Leone (Statistics and Experimental Design in Engineering and the Physical Sciences, John Wile) describen un experimento realizado para investigar la torcedura de placas de cobre. Los dos factores estudiados fueron la temperatura y el contenido de cobre de las placas. La variable de respuesta fue una medida de la cantidad de torcedura. Los datos fueron los siguientes:

Temperatura ( $^{\circ}\text{C}$ )	Contenido (%)			
	40	60	80	100
50	17,20	16,21	24,22	28,27
75	12, 9	18, 13	17, 12	27,31
100	16,12	18,21	25,23	30,23
125	21,17	23,21	23,22	29. 31

a) ¿Existe algún indicio de que alguno de los dos factores afecta la cantidad de torcedura? ¿Hay alguna interacción entre los factores? Utilizar  $\alpha = 0.05$ . b) Analizar los residuales de este experimento. c) Graficar la torcedura promedio con cada nivel del contenido de cobre y compararlas con una distribución  $t$  con la escala apropiada. Describir las diferencias en los efectos de los diversos niveles del contenido de cobre sobre la torcedura. Si es deseable una torcedura baja, ¿qué nivel del contenido de cobre debería especificarse? d) Suponga que no es sencillo controlar la temperatura en el medio ambiente donde van a usarse las placas de cobre. ¿Este hecho modifica la respuesta que se dio para el inciso c?

EJERCICIO 76. Se estudian los factores que influyen en la resistencia a la ruptura de una fibra sintética. Se eligen cuatro quinas de producción y tres operadores y se corre un experimento factorial utilizando fibra del mismo lote de producción. Los resultados son los siguientes:

Operador	Máquina			
	1	2	3	4
1	109	110	108	110
	110	115	109	108
2	110	110	111	114
	112	111	109	112
3	116	112	114	120
	114	115	119	117

(a) Analizar los datos y sacar conclusiones. Utilizar  $\alpha = 0.05$ . (b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo.

EJERCICIO 77. Un ingeniero mecánico estudia la fuerza de empuje desarrollada por una taladradora. Sospecha que la velocidad de taladrado y la velocidad de alimentación del material son los factores más importantes. Selecciona cuatro velocidades de alimentación y usa una velocidad de taladrado alta y otra baja elegidas para representar las condiciones de operación extremas. Obtiene los siguientes resultados. Analizar los datos y sacar conclusiones. Utilizar  $\alpha = 0.05$ .

Velocidad de taladrado	Velocidad de alimentación			
	0.015	0.030	0.045	0.060
125	2.70	2.45	2.60	2.75
	2.78	2.49	2.72	2.86
200	2.83	2.85	2.86	2.94
	2.86	2.80	2.87	2.88

EJERCICIO 78. Se realiza un experimento para estudiar la influencia de la temperatura de operación y tres tipos de placas de cubrimiento de cristal, en la salida luminosa de un tubo de osciloscopio. Se registraron los siguientes datos:



Tipo de cristal	Temperatura		
	100	125	150
1	580	1090	1392
	568	1087	1380
	570	1085	1386
	550	1070	1328
2	530	1035	1312
	579	1000	1299
	546	1045	867
3	575	1053	904
	599	1066	889

(a) Utilizar  $\alpha = 0.05$  en el análisis. ¿Existe un efecto de interacción significativo? ¿El tipo de cristal o la temperatura afectan la respuesta? ¿A qué conclusiones se llega? (b) Ajustar un modelo apropiado que relacione la salida luminosa con el tipo de cristal y la temperatura. (c) Analizar los residuales de este experimento. Comentar la adecuación de los modelos que se hayan considerado.

**EJERCICIO 79.** El porcentaje de la concentración de madera dura en la pulpa bruta, la presión de la cuba y el tiempo de cocción de la pulpa se investigan en cuanto a sus efectos sobre la resistencia del papel. Se seleccionan tres niveles de la concentración de madera dura, tres niveles de la presión y dos tiempos de cocción. Se lleva a cabo un experimento factorial con dos réplicas, obteniéndose los siguientes datos:

Coccion	3 horas			4 horas		
	Presión			Presión		
Porcentaje	400	500	650	400	500	650
2	196.6	197.7	199.8	198.4	199.6	200.6
	196.0	196.0	199.4	198.6	200.4	200.9
4	198.5	196.0	198.4	197.5	198.7	199.6
	197.2	196.9	197.6	198.1	198.0	199.0
8	197.5	195.6	197.4	197.6	197.0	198.5
	196.6	196.2	198.1	198.4	197.8	199.8

(a) Analizar los datos y sacar conclusiones. Utilizar  $\alpha = 0.05$ . (b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo. (c) ¿Bajo qué conjunto de condiciones debería operarse este proceso? ¿Por qué?

**EJERCICIO 80.** El departamento de control de calidad de una planta de acabados textiles estudia el efecto de varios factores sobre el teñido de una tela de algodón y fibras sintéticas utilizada para fabricar camisas para caballero. Se seleccionaron tres operadores, tres duraciones del ciclo y dos temperaturas, y se tiñeron tres ejemplares pequeños de la tela bajo cada conjunto de condiciones. La tela terminada se comparó con un patrón, y se le asignó una evaluación numérica. Los datos se presentan enseguida. Analizar los datos y sacar conclusiones. Comentar la adecuación del

*modelo.*

	<i>Temperatura</i>					
	<i>300°</i>			<i>350°</i>		
	<i>Operador</i>			<i>Operador</i>		
<i>Ciclo</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>1</i>	<i>2</i>	<i>3</i>
<i>40</i>	<i>23</i>	<i>27</i>	<i>31</i>	<i>24</i>	<i>38</i>	<i>34</i>
	<i>24</i>	<i>28</i>	<i>32</i>	<i>23</i>	<i>36</i>	<i>36</i>
	<i>25</i>	<i>26</i>	<i>29</i>	<i>28</i>	<i>35</i>	<i>39</i>
<i>50</i>	<i>36</i>	<i>34</i>	<i>33</i>	<i>37</i>	<i>34</i>	<i>34</i>
	<i>35</i>	<i>38</i>	<i>34</i>	<i>39</i>	<i>38</i>	<i>36</i>
	<i>36</i>	<i>39</i>	<i>35</i>	<i>35</i>	<i>36</i>	<i>31</i>
<i>60</i>	<i>28</i>	<i>35</i>	<i>26</i>	<i>26</i>	<i>36</i>	<i>28</i>
	<i>24</i>	<i>35</i>	<i>27</i>	<i>29</i>	<i>37</i>	<i>26</i>
	<i>27</i>	<i>34</i>	<i>25</i>	<i>25</i>	<i>34</i>	<i>24</i>

## Diseños factoriales $2^k$

### 1. Competencias

Aplicar los diseños factoriales  $2^2$ ,  $2^3$  y  $2^k$  y tomar decisiones acerca de cuándo se debe aplicar cada diseño. Diseñar un experimento factorial  $2^5$  no replicado para aplicarlo a diversos casos. Explicar las ventajas y desventajas de aplicar el diseño factorial  $2^k$  en bloques o con punto al centro.

### 2. Diseños factoriales $2^2$

Con un diseño factorial  $2^2$  se estudia el efecto de dos factores considerando dos niveles en cada uno. Cada réplica de este diseño consiste de  $2 \times 2 = 4$  combinaciones o tratamientos que se pueden denotar de diferentes maneras. En la tabla 1 se muestran diferentes maneras de escribir los cuatro tratamientos que conforman el diseño factorial  $2^2$ .

Tabla 1									
	A	B	A	B	A	B	A	B	Yates
Trat 1	bajo	bajo	A <sub>1</sub>	B <sub>1</sub>	A <sup>-</sup>	B <sup>-</sup>	-	-	(1)
Trat 2	alto	bajo	A <sub>2</sub>	B <sub>1</sub>	A <sup>+</sup>	B <sup>-</sup>	+	-	<i>a</i>
Trat 3	bajo	alto	A <sub>1</sub>	B <sub>2</sub>	A <sup>-</sup>	B <sup>+</sup>	-	+	<i>b</i>
Trat 4	alto	alto	A <sub>2</sub>	B <sub>2</sub>	A <sup>+</sup>	B <sup>+</sup>	+	+	<i>ab</i>

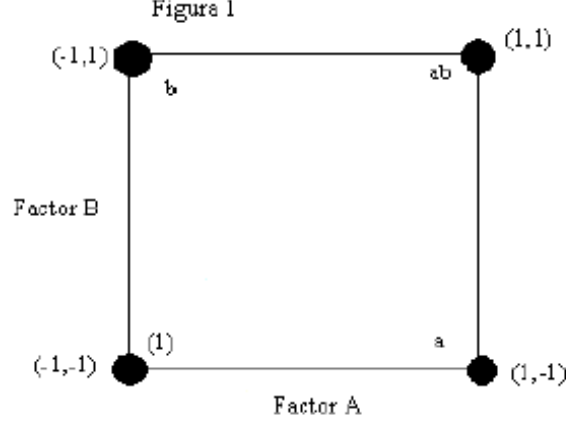
La notación de Yates [(1), *a*, *b*, *ab*] tiene un significado diferente a las demás: con ella se representa el total o suma de las observaciones en cada tratamiento. En específico, (1) es la suma de todos los datos obtenidos en el tratamiento (-1, -1); *a* es la suma de todas las mediciones hechas en la combinación (1, -1), y así sucesivamente.

**2.1. Representación geométrica.** El diseño factorial  $2^2$  se representa geométricamente por los vértices del cuadrado de la figura 1. Cada vértice representa un punto de diseño o tratamiento. El área limitada por este cuadrado se conoce como región experimental y en principio las conclusiones que se obtengan del experimento sólo tienen validez sobre esta región.

**2.2. Cálculo de los efectos .** En este diseño hay tres efectos de interés: los dos efectos principales (*A* y *B*) y el efecto de interacción (*AB*). Con el uso de la notación de *Yates* como los totales de las *n* repeticiones en cada punto de diseño, se tiene:

$$(2.1) \quad \text{Efecto } A = \frac{1}{2n}[a + ab - b - (1)] = \frac{1}{2n}[a + ab] - \frac{1}{2n}[b + (1)]$$

$$(2.2) \quad \text{Efecto } B = \frac{1}{2n}[b + ab - a - (1)] = \frac{1}{2n}[b + ab] - \frac{1}{2n}[a + (1)]$$



$$(2.3) \quad \text{Efecto } AB = \frac{1}{2n}[ab + (1) - a - b] = \frac{1}{2n}[ab - b] - \frac{1}{2n}[a - (1)]$$

Geoméricamente, el efecto  $A$  equivale a promediar los datos del lado derecho del cuadrado de la figura 1 y restarles el promedio de los datos del lado izquierdo; mientras que para el efecto  $B$  se promedian los datos del lado de arriba y se le resta la media de los datos del lado de abajo. Geométricamente, la interacción es la diferencia entre las medias de las diagonales del cuadrado de la figura 1.

**2.3. Análisis de varianza.** Aunque los efectos calculados dados sean números distintos de cero, esto no implica que el efecto correspondiente sea estadísticamente diferente de cero. O si en su representación gráfica aparentan ser importantes, eso tampoco es suficiente para concluir que afectan de manera significativa la variable de respuesta. Para poder afirmar que tales efectos contribuyen a explicar el comportamiento de la respuesta, se debe hacer un análisis de varianza. Las sumas de cuadrados que componen el ANOVA se pueden calcular apartir de los efectos estimados.

**2.3.1. Definición de contraste.** Una combinación lineal de la forma  $C = \sum_{i=1}^{2^k} c_i Y_i$ , con  $\sum_{i=1}^{2^k} c_i = 0$  se llama contraste. La suma de cuadrados para cualquier contraste  $C$  está dada por

$$(2.4) \quad SCC = \sum_{i=1}^{2^k} c_i Y_i / n \sum_{i=1}^{2^k} c_i^2$$

la cual tiene sólo un grado de libertad. Por ejemplo, los contrastes correspondientes a los tres efectos  $A$ ,  $B$  y  $AB$  en el diseño factorial  $2^2$  están dados por

$$\begin{aligned} \text{Contraste } A &= [a + ab - b - (1)] \\ \text{Contraste } B &= [b + ab - a - (1)] \\ \text{Contraste } AB &= [ab + (1) - a - b] \end{aligned}$$

que como hemos visto, son las cantidades que definen a los efectos. Son contrastes por el hecho de que son combinaciones lineales donde los coeficientes suman cero ( $1+1-1-1=0$ ). Una vez obtenido el contraste, el efecto correspondiente se obtiene dividiéndolo entre la constante que lo convierta en una diferencia de medias; este número es la mitad de las observaciones hechas en el experimento [(véase ecuaciones (2.1), (2.2) y (2.3)]. Por ejemplo, en el factorial  $2^k$  con  $n$  réplicas, los contrastes se dividen por  $n2^{(k-1)}$  para estimar los efectos; en particular para el diseño  $2^2$  con  $n$  réplicas se divide por  $n2^{(2-1)} = 2n$ .

**2.3.2. Métodos para calcular contrastes.** Los contrastes de cualquier efecto, sea principal o de interacción en el diseño factorial  $2^k$ , se obtienen mediante el auxilio de tabla de signos, la cual se construye a partir de la matriz de diseño multiplicando las columnas que intervienen en la interacción que se quiera calcular. Por ejemplo, si se quiere calcular la interacción doble  $AB$  se multiplican la columna de signos  $A$  por la columna  $B$ , y el resultado son los signos del contraste  $AB$ , como se muestra en la siguiente tabla de signos del diseño factorial  $2^2$ .

$A$	$B$	$AB$	$Yates$
—	—	+	(1)
+	—	—	$a$
—	+	—	$b$
+	+	+	$ab$

Por ejemplo, *Contraste A* =  $[a + ab - b - (1)]$ . Lo mismo se hace para los demás efectos.

**2.3.3. Pasos para llegar al ANOVA .** Para investigar cuáles de los tres efectos están activos o son significativos se procede a probar las hipótesis dadas por

$$\begin{aligned} H_0 &: \text{Efecto } A = 0 \\ H_0 &: \text{Efecto } B = 0 \\ H_0 &: \text{Efecto } AB = 0 \end{aligned}$$

Cada una contra la alternativa de que el efecto en cuestión es diferente de cero. Estas hipótesis se prueban con el análisis de varianza, y para ello es necesario calcular las sumas de cuadrados que corresponden a los efectos  $A, B$  y  $AB$ , dados por las ecuaciones (2.1), (2.2) y (2.3), respectivamente.

$$(2.5) \quad SCA = \frac{1}{2^2 n} [a + ab - b - (1)]^2$$

$$(2.6) \quad SCB = \frac{1}{2^2 n} [b + ab - a - (1)]^2$$

$$(2.7) \quad SCAB = \frac{1}{2^2 n} [ab + (1) - a - b]^2$$

donde cada una de ellas tiene sólo un grado de libertad, debido a que cada factor sólo tiene dos niveles. La suma de cuadrados totales se calcula con la expresión:

$$SCT = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{l=1}^n Y_{ijl}^2 - \frac{1}{2^2 n} Y_{...}^2$$

y tiene  $n2^2 - 1$  grados de libertad. La suma de cuadrados del error se calcula por diferencia:  $SCE = SCT - SCA - SCB - SCAB$ , y tiene  $n2^2 - 1 - 3 = 4(n - 1)$

grados de libertad.

Tabla 2. ANOVA para el diseño $2^2$					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A	SCA	1	CMA	CMA/CME	$P(F > F_0^A)$
B	SCB	1	CMB	CMB/CME	$P(F > F_0^B)$
AB	SCAB	1	CMAB	CMAB/CME	$P(F > F_0^{AB})$
Error	SCE	$4(n-1)$	CME		
Total	SCT	$n2^2-1$			

Si el *valor  $p$*  es menor que el nivel de significancia prefijado  $\alpha$ , se concluye que el efecto correspondiente es estadísticamente distinto de cero, es decir, tal efecto está activo o influye de manera significativa sobre la respuesta y mientras más pequeño sea, más importante es tal efecto. Se recomienda correr el factorial  $2^2$  con al menos *tres* réplicas para poder estimar un *CME* confiable.

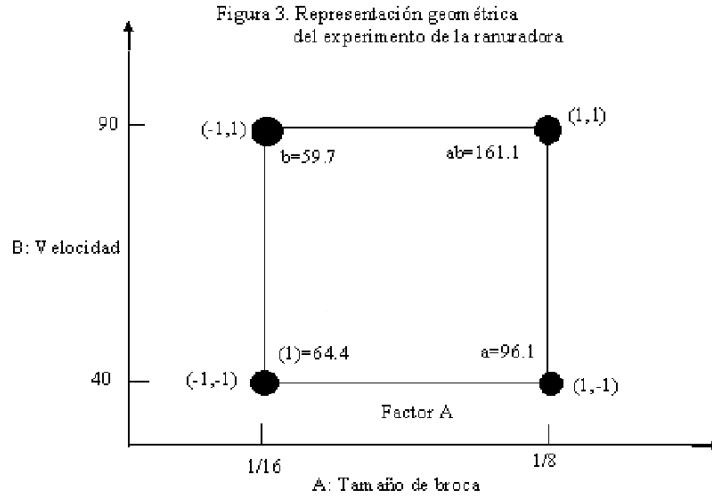
EJEMPLO 12. **Experimento  $2^2$**  *Interesa estudiar el efecto del tamaño de broca (factor A) y de la velocidad (factor B) sobre la vibración de la ranuradora (respuesta Y). Para ello se decide utilizar un diseño factorial  $2^2$  con cuatro réplicas, es decir, cuatro repeticiones en cada tratamiento, lo que da un total de  $4 \times 2^2 = 16$  corridas del proceso, que se realizan en orden aleatorio. El tamaño de la broca se prueba en  $1/16$  y en  $1/8$  de pulgada y la velocidad en 40 y 90 revoluciones por segundo, según se describe en la siguiente tabla:*

Tabla 3						Tabla 4						
Broca	Veloc.	Orden				A:Broca	B:Veloc.	Vibración				Totales
1/16	40	5	8	13	14	—	—	18.2	18.9	12.9	14.4	64.4 = (1)
1/8	40	1	6	10	12	+	—	27.2	24.0	22.4	22.5	96.1 = a
1/16	90	3	7	11	15	—	+	15.9	14.5	15.1	14.2	59.7 = b
1/8	90	2	4	9	16	+	+	41.0	43.9	36.3	39.9	161.1 = ab

Al usar el software Statgraphics se generan automáticamente cuatro bloques, uno por cada réplica, ya que considera que el experimento se correrá réplica por réplica, y entonces pudiera actuar algún factor de bloque en el transcurso de las réplicas. Cuando dicho factor de bloque existe y está identificado debe analizarse su posible efecto en el ANOVA; pero si no existe un factor a quién atribuir los bloques, éstos se eliminan a la hora del análisis activando la opción para que sean ignorados.

La representación geométrica del experimento se muestra en la figura 3. Observe la relación entre las unidades originales y las unidades codificadas y el significado de la notación de Yates.

Las preguntas fundamentales que se quieren responder con el experimento son: *¿la velocidad y el tamaño de la broca afectan la vibración de la ranuradora?*, *si la afectan, ¿cómo es tal efecto y cuál combinación de velocidad y tamaño de broca minimizan la vibración?*, *¿cuál es la vibración esperada en las condiciones óptimas?*, *¿se cumplen los supuestos del modelo?*



**Efectos estimados.** De acuerdo con las relaciones (2.1), (2.2) y (2.3), y a la última columna de la tabla .4, los efectos estimados están dados por

$$\begin{aligned} \text{Efecto } A &= \frac{1}{2(4)} [96.1 + 161.1 - 59.7 - 64.4] = 16.64 \\ \text{Efecto } B &= \frac{1}{2(4)} [59.7 + 161.1 - 96.1 - 64.4] = 7.54 \\ \text{Efecto } AB &= \frac{1}{2(4)} [161.1 + 64.4 - 96.1 - 59.7] = 8.71 \end{aligned}$$

Se observa que el efecto del tamaño de broca (factor  $A$ ) es más alto que los otros dos.

#### Análisis de varianza.

Las sumas de cuadrados de los efectos se calcula a partir de sus contrastes dados en las ecuaciones (2.5), (2.6) y (2.7) como:

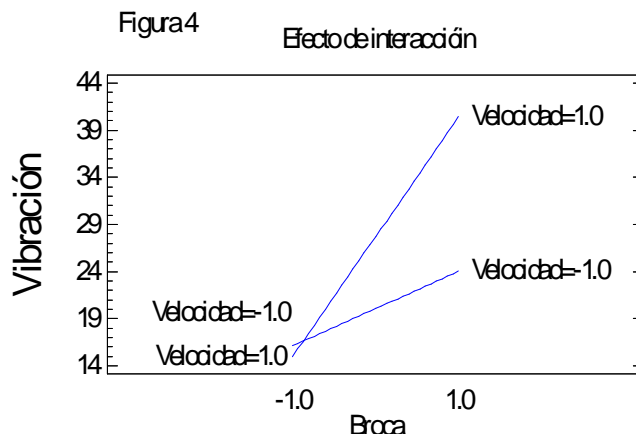
$$\begin{aligned} SCA &= \frac{1}{2^2(4)} [96.1 + 161.1 - 59.7 - 64.4]^2 = 1107.22 \\ SCB &= \frac{1}{2^2(4)} [59.7 + 161.1 - 96.1 - 64.4]^2 = 227.25 \\ SCAB &= \frac{1}{2^2(4)} [161.1 + 64.4 - 96.1 - 59.7]^2 = 303.63 \end{aligned}$$

cada uno tiene un grado de libertad. La suma de cuadrados totales es

$$SCT = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{l=1}^n Y_{ijl}^2 - \frac{1}{2^2 n} Y_{...}^2 = 1709.83$$

y tiene 15 grados de libertad. El cuadrado del error se calcula por diferencia:

$$\begin{aligned} SCE &= SCT - SCA - SCB - SCAB \\ &= 1709.83 - 1107.22 - 227.25 - 303.63 = 71.73 \end{aligned}$$



y le quedan  $15 - 3 = 12$  grados de libertad.

Tabla 5. ANOVA para el experimento de la ranuradora.					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A:Broca	1107.22	1	1107.22	185.25	0.0000
B:Velocidad	227.25	1	227.25	38.02	0.0000
AB	303.63	1	303.63	50.80	0.0000
Error	71.73	12	71.73		
Total	1709.83	15			

El análisis de varianza se muestra en la tabla 5. De acuerdo con la columna para el valor  $p$ , cuyas entradas son menores que 0.05, se rechazan las tres hipótesis nulas  $H_0 : efecto A = 0$ ,  $H_0 : efecto B = 0$  y  $H_0 : efecto AB = 0$ , y se concluye que  $H_1 : efecto A \neq 0$ ,  $H_1 : efecto B \neq 0$  y  $H_1 : efecto AB \neq 0$ . Es decir, los tres efectos son muy significativos.

#### Interpretación y conclusiones.

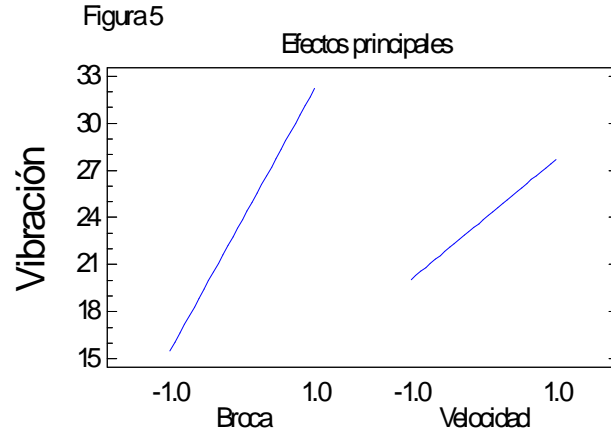
El objetivo es minimizar la vibración de la ranuradora y las interacciones tienen prioridad respecto a los efectos principales. La información relevante del experimento se encuentra en la gráfica de la interacción (figura 4). Se puede afirmar que:

- (1) Cuando la broca se encuentra en su nivel bajo la velocidad no afecta de manera significativa la vibración.
- (2) Cuando la broca se encuentra en su nivel alto, la velocidad tiene un efecto considerable sobre la vibración.

Como se quiere minimizar la vibración, entonces se puede utilizar el tratamiento  $(A^-, B^+)$  o el  $(A^-, B^-)$ , ambos prácticamente logran los mismos resultados. (figura 4)

En la figura 5 se muestra la gráfica de los efectos principales, se observa que no se debe aumentar la velocidad, ni aumentar el tamaño de broca si se quiere minimizar la vibración, conduciendo a utilizar la combinación  $(A^-, B^-)$ .





### Predicción sobre el mejor tratamiento y $R^2$ .

Con el uso de valores codificados para los dos factores  $(x_1, x_2)$ , como se muestra en la tabla 4, el modelo de regresión ajustado que describe el comportamiento de la vibración sobre cualquier punto, está dado por

$$(2.8) \quad \hat{Y} = 23.83125 + 8.32x_1 + 3.77x_2 + 4.35x_1x_2$$

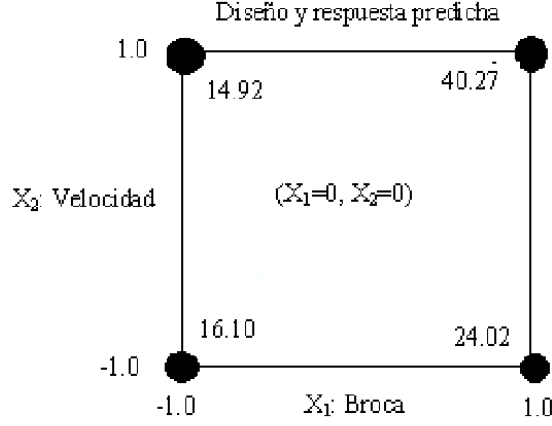
donde  $\hat{Y}$  es la respuesta predicha en el punto  $(x_1, x_2)$ , con  $x_1 = A$  (broca) y  $x_2 = B$  (velocidad).

NOTA 5. En el caso de diseños  $2^k$ , los coeficientes del modelo de regresión son iguales a los efectos estimados que resultaron significativos divididos entre dos. Así por ejemplo, el coeficiente de  $x_1$ , es igual a efecto de  $A \div 2$  ( $16.64 \div 2$ ). Esta división entre dos se hace para lograr una escala unitaria, que es la escala usual en regresión. Los efectos originales no se encuentran en una escala unitaria dado que el ancho de la región experimental es de dos unidades codificadas, de aquí que se divida entre dos. El término independiente  $\beta_0 = 23.83125 = \bar{Y} \dots$ , representa la vibración predicha en el centro de la región experimental ( $x_1 = 0, x_2 = 0$ ) (véase figura .6).

La predicción del comportamiento de la vibración sobre la combinación de niveles que se quiera dentro de la región experimental se puede realizar con modelo ajustado dado por la ecuación (2.8), evaluando el punto en la ecuación. El grado de credibilidad en la predicción obtenida depende de la calidad del modelo ajustado. A continuación vemos cómo medir la calidad del ajuste.

**2.4. Coeficientes de determinación  $R^2$  y  $R^2_{ajustado}$ .** Dos de los estadísticos más útiles para medir la calidad global del modelo de regresión múltiple es el coeficiente de determinación ( $R^2$ ), y el coeficiente de determinación ajustado

Figura 6.



( $\mathbf{R}_{ajustado}^2$ ), que están definidos de la manera siguiente:

$$\mathbf{R}^2 = \frac{SCT - SCE}{SCT} \times 100$$

$$\mathbf{R}_{ajustado}^2 = \frac{CMT - CME}{CMT} \times 100.$$

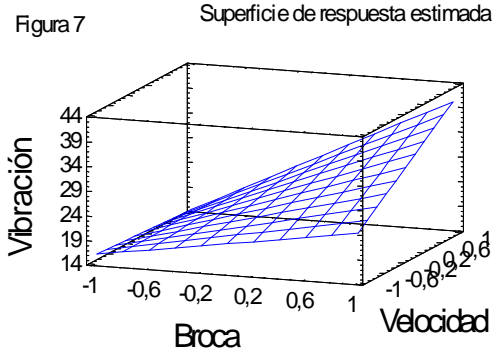
Estos coeficientes comparan la variabilidad explicada por el modelo contra la variación total, cuantificadas éstas a través de la suma de cuadrados o por el cuadrado medio. De esta forma, para interpretar estos coeficientes se cumple que  $0 \leq \mathbf{R}_{ajustado}^2 \leq \mathbf{R}^2 \leq 100$  y cuantifican el porcentaje de variabilidad presente en los datos que es explicada por el modelo; por lo que son deseables valores próximos a 100. Para fines de predicción se recomienda un coeficiente de determinación ajustado de al menos 70%. *Cuando hay muchos factores se prefiere el estadístico  $\mathbf{R}_{ajustado}^2$  en lugar del  $\mathbf{R}^2$  sin ajustar, puesto que este último es engañoso al incrementarse de manera artificial con cada término que se agrega al modelo, aunque sea un término que no contribuya en nada a la explicación de la respuesta.* En cambio, el  $\mathbf{R}_{ajustado}^2$  incluso baja de valor cuando el término que se agrega no aporta nada. Para el modelo de la vibración (experimento de la ranuradora) es muy fácil calcular estos coeficientes a partir de la tabla 5 de análisis de varianza.

Para el ejemplo 12

$$\mathbf{R}^2 = \frac{SCT - SCE}{SCT} \times 100 = \frac{1709.83 - 71.73}{1709.83} \times 100 = 95.8$$

$$\mathbf{R}_{ajustado}^2 = \frac{CMT - CME}{CMT} \times 100 = \frac{\frac{1709.83}{15} - \frac{71.73}{12}}{\frac{1709.83}{15}} \times 100 = 94.76$$

De esta manera, de acuerdo a  $\mathbf{R}_{ajustado}^2$  el modelo ajustado en la expresión, (2.8) explica el 94.76% de la variabilidad de la vibración observada en el experimento (algo similar nos dice el  $\mathbf{R}^2$ ). Esto significa que los factores estudiados (tipo de broca y velocidad de la ranuradora), junto con su interacción, son responsables o explican un alto porcentaje de la variabilidad observada en la variable de respuesta (vibración), y por tanto el efecto atribuible a factores no estudiado, sea que se



hayan mantenido en un nivel fijo o que hayan tenido pequeñas variaciones, más el efecto de errores experimentales, fueron pequeños comparados con el efecto de los factores estudiados.

**2.5. Hacer la predicción.** La respuesta predicha  $\hat{Y}$  en un punto dado es un estimador de la respuesta promedio en dicho punto y es una variable aleatoria, y como tal se puede estimar su intervalo de confianza. Por ejemplo, el valor promedio de la vibración predicha sobre uno de los mejores tratamientos,  $(-1, 1)$ , se obtiene al sustituir este punto en el modelo ajustado y el resultado es

$$\hat{Y}(-1, 1) = 23.83 + 8.32(-1) + 3.77(1) + 4.35(-1)(1) = 14.92.$$

**2.6. Gráficos de superficie.** Permiten tener una visualización de lo que significa el modelo ajustado dado en la ecuación 2.8. Este modelo representa una superficie de respuesta sobre la región experimental, que modela el comportamiento de la vibración y es útil cuando se tienen dos factores, dado que no se puede dibujar una superficie en cuatro o más dimensiones.

En la figura 7 se presenta el gráfico de superficie, que describe el comportamiento de la vibración sobre la región experimental, que es la cara inferior del cubo. Las curvas de nivel o isolíneas, dibujadas como líneas más gruesas sobre la región experimental son otra manera de representar la superficie. Cada curva de nivel representa puntos o combinaciones de broca y velocidad donde la vibración es constante.

La representación en curvas de nivel o gráfico de contornos se puede dibujar sin la superficie, como se muestra en la figura 8. Esta representación es mejor que la de superficie porque se pueden ver con bastante exactitud las coordenadas del punto con la vibración deseada. Los números que acompañan a las curvas de nivel son precisamente la altura de la superficie sobre toda la isolínea, y éstas se pueden dibujar con los valores que el experimentador quiera.

**2.7. Verificación de supuestos .** Los supuestos del modelo de análisis de varianza deben verificarse antes de dar por válidas las conclusiones de la etapa de

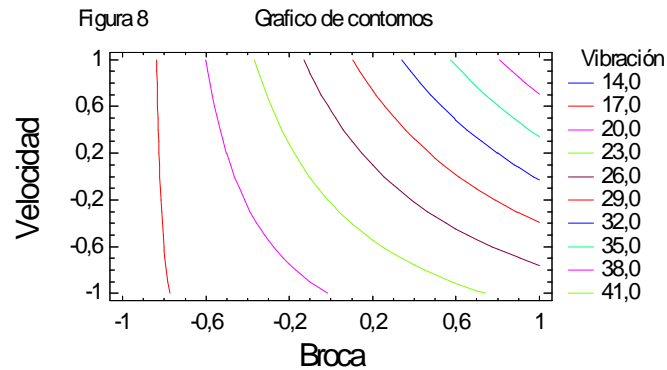
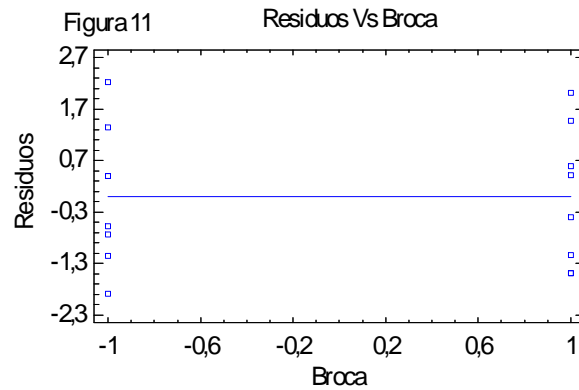
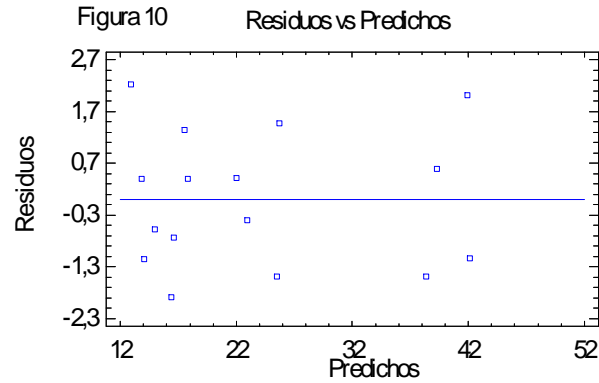


Figura 9 Valor de los residuos, experimento de la ranuradora

Velocidad	—	0.975 0.175 -0.425 -0.725 $S(-,+)=0.75$	0.725 -3.975 3.625 -0.375 $S(+,+)=3.14$
	+	2.1 -3.2 2.8 -1.7 $S(-,-)=2.91$	3.175 -1.625 -0.025 -1.525 $S(+,-)=2.24$
		— Broca	+

interpretación. La tabla de ANOVA supone que los residuos se distribuyen normales, independientes y con varianza constante. El supuesto de *varianza constante* se puede verificar graficando los *residuos vs los predichos*, y los puntos deben de caer aleatoriamente en el sentido vertical dentro de una banda horizontal, para concluir que el supuesto se cumple (véase figura 10). Los huecos en el sentido horizontal se deben a que sólo son cuatro puntos de diseño, y las predicciones en ellos difieren bastante. Aquí habría cierta evidencia para decir que el supuesto de varianza constante no se cumple, dada la dispersión más compacta de la primera columna de puntos en relación a las otras tres. Sin embargo, en este caso no es una situación grave que pudiera afectar las conclusiones del ANOVA, basadas en *valor - p* tan claramente pequeños. Además, la menor dispersión ocurre justo en el punto donde la vibración es mejor, lo que viene a reforzar las conclusiones obtenidas.

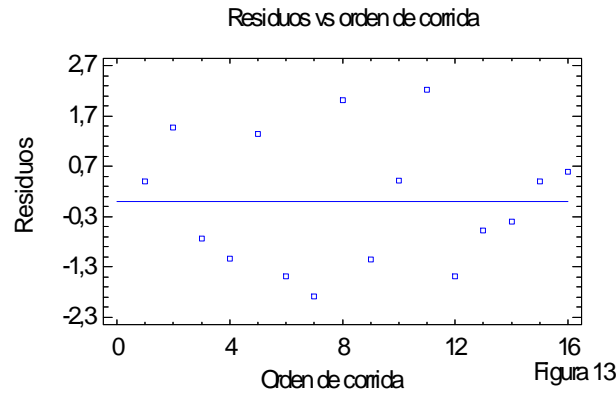
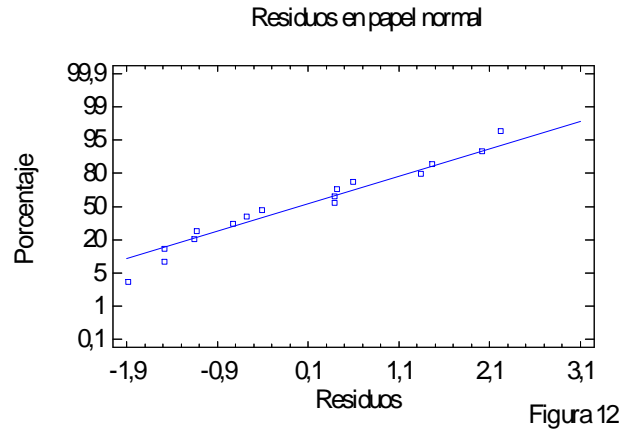
Otro gráfico que sirve para comprobar el supuesto de *varianza constante* es el de *residuos vs cada factor*. En la figura 11 se dibujan los residuos contra el factor



broca. Se observa que las columnas de puntos en los niveles bajo y alto del factor broca, tienen "más o menos" la misma dispersión, por tanto se cumple el supuesto de varianza constante.

En la figura 12 se grafican los residuos en papel probabilístico normal. Como los puntos se apegan a la línea colocada visualmente (no es línea de regresión), se concluye que no hay violaciones al supuesto de normalidad.

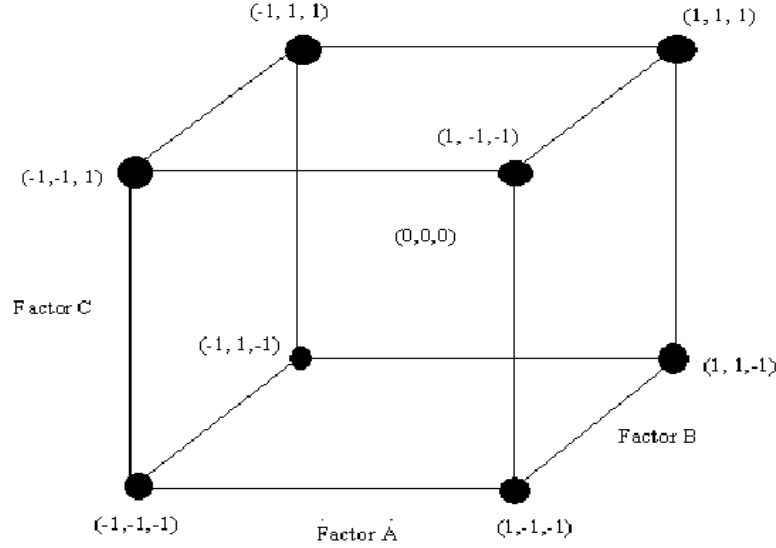
Para probar el supuesto de *independencia* se requiere capturar los datos en el orden en que fueron obtenidos, para que el software pueda graficar los *residuos vs orden de corrida* (figura 13). A1 no observarse ninguna tendencia en los puntos (que suban o bajen) se concluye que no existe problema con la declaración de independencia.



### 3. Diseño factorial $2^3$

En este diseño factorial se estudian tres factores en dos niveles cada uno, consta de  $2^3 = 2 \times 2 \times 2 = 8$  tratamientos diferentes. Los tratamientos del diseño  $2^3$  y su representación geométrica se muestran en la figura 14. La región experimental es ahora un cubo regular centrado en el origen  $(0, 0, 0)$ , cuyos vértices son los ocho tratamientos. La matriz de diseño se construye fácilmente alternando el signo menos y el signo más en la primera columna, dos menos y dos más en la segunda columna, y cuatro menos y cuatro más en la tercera; el diseño resulta acomodado en el orden estándar o de *Yates*

Con este diseño se pueden estudiar los  $2^3 - 1 = 7$  efectos: tres efectos principales  $A$ ,  $B$ ,  $C$ ; tres interacciones dobles  $AB$ ,  $AC$ ,  $BC$  y una interacción triple  $ABC$ . En los diseños factoriales  $2^k$  se estudian los efectos principales y las interacciones dobles; las interacciones de mayor orden (de tres o más factores) generalmente no influyen de manera significativa, motivo por el cual no es recomendable estudiarlas.

Figura 14 Diseño factorial  $2^3$  y su representación geométrica

**3.1. Análisis del diseño factorial  $2^3$ .** Sean  $A$ ,  $B$  y  $C$  los factores que se quieren estudiar y sean  $(1)$ ,  $a$ ,  $b$ ,  $ab$ ,  $c$ ,  $ac$ ,  $bc$  y  $abc$ , los totales observados en cada uno de los ocho tratamientos escritos en su orden estándar. Los contrastes para los siete efectos, dados por:

$$\begin{aligned}
 \text{Contraste } A &= [a + ab + ac + abc - (1) - b - c - bc] \\
 \text{Contraste } B &= [b + ab + bc + abc - (1) - a - c - ac] \\
 \text{Contraste } C &= [c + ac + bc + abc - (1) - a - b - ab] \\
 \text{Contraste } AB &= [ab - b - a + abc + (1) - bc - ac + c] \\
 \text{Contraste } AC &= [(1) - a + b - ab - c + ac - bc + abc] \\
 \text{Contraste } BC &= [(1) + a - b - ab - c - ac + bc + -abc] \\
 \text{Contraste } ABC &= [abc - bc - ac + c - ab + b + a - (1)].
 \end{aligned}$$

Si se hacen  $n$  réplicas de cada tratamiento, los efectos de un diseño  $2^3$  se estiman dividiendo los contrastes entre  $4n$ , así por ejemplo, el efecto principal de  $A$  se estima de la siguiente manera:

$$EfectoA = \frac{\text{Contraste } A}{n2^{k-1}}.$$

También las sumas de cuadrados de los efectos se calculan a partir de sus contrastes con la fórmula

$$SC_{efecto} = \frac{(\text{Contraste } efecto)^2}{n2^k}$$

La suma total de cuadrados se obtiene de la manera usual como:

$$SCT = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{l=1}^2 \sum_{m=1}^n Y_{ijlm}^2 - \frac{1}{2^k n} Y_{...}^2$$

y finalmente, la suma de cuadrados del error se calcula por sustracción

Tabla 6 de Signos para el diseño factorial $2^3$							
A	B	C	AB	AC	BC	ABC	Yates
−	−	−	+	+	+	−	(1)
+	−	−	−	−	+	+	a
−	+	−	−	+	−	+	b
+	+	−	+	−	−	−	ab
−	−	+	+	−	−	+	c
+	−	+	−	+	−	−	ac
−	+	+	−	−	+	−	bc
+	+	+	+	+	+	+	abc

Tabla 7. ANOVA para el diseño $2^3$					
Fuente	SC	gl	CM	$F_0$	valor-p
A	SCA	1	CMA	CMA/CME	$P(F > F_0^A)$
B	SCB	1	CMB	CMB/CME	$P(F > F_0^B)$
C	SCC	1	CMC	CMC/CME	$P(F > F_0^C)$
AB	SCAB	1	CMAB	CMAB/CME	$P(F > F_0^{AB})$
AC	SCAC	1	CMAC	CMAC/CME	$P(F > F_0^{AC})$
BC	SCBC	1	CMBC	CMBC/CME	$P(F > F_0^{BC})$
ABC	SCABC	1	CMABC	CMABC/CME	$P(F > F_0^{ABC})$
Error	SCE	$2^3(n-1)$	CME		
Total	SCT	$n2^3-1$			

**3.2. Ejemplo de un diseño de experimento en la Industria Electrónica.** Con este experimento se busca reducir los cambios bruscos de temperatura que se dan en una etapa del proceso de fabricación de dispositivos electrónicos, conocidos como *obleas*, ya que algunas obleas no resisten dichos cambios de temperatura y se rompen. *Una oblea de silicio es un dispositivo electrónico en el que vienen integrados microcircuitos para ser procesados juntos, es decir, la oblea es el medio que permite procesar al mismo tiempo cientos de dados o chips, los cuales deben cumplir ciertas propiedades eléctricas al final del proceso.* El proceso que siguen las obleas consta de más de 200 operaciones. En la figura 15 se muestra el aspecto que tiene una oblea típica con sus dados.

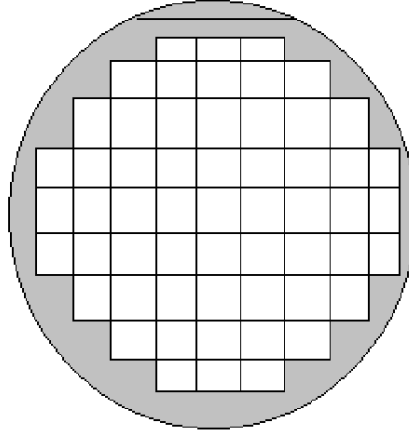
Las características de calidad más importantes al final del proceso son:

- (1) Ciertas propiedades eléctricas que debe tener cada dado.
- (2) Funcionalidad (rendimiento eléctrico), en el sentido de que cumplen las especificaciones que los clientes exigen.
- (3) Rendimiento mecánico, que se define como el porcentaje de obleas que soportan todo el proceso sin romperse. Mientras menos dados sean buenos y mientras más obleas se rompan, es menor la eficacia del proceso.

**3.2.1. Identificación del problema.** Existe un departamento donde se procesan millones de microcircuitos integrados en obleas de silicio. A lo largo del proceso se rompen muchas por diversas causas. La primera etapa consistió en localizar los *subprocesos más críticos*; identificado el subproceso de interés, se desplegaron todas las posibles causas del problema. Haciendo uso de *Diagramas de Pareto* en varios niveles se detectaron las causas primarias del rompimiento de obleas, los lugares del proceso donde esto ocurre con mayor frecuencia y las causas específicas del fenómeno en una operación particular. Se encontró que el *efecto térmico* es la principal causa específica del rompimiento de obleas en la *solución piranha*.



Figura 15 Dados sobre una oblea de silicio



Las obleas se rompen por diversas causas, pero se ha observado que cuando es por efecto térmico el borde de la falla es recto, y cuando se rompen por otras causas, como es manejo, la falla es irregular (figura 17). Es importante distinguir cuando la rotura es por efecto térmico, para no atribuir a éste obleas que se rompen por otros motivos durante las pruebas.

Un subproceso particularmente crítico en cuanto al número de obleas rotas por efecto térmico es el *grabado mesa*. En este subproceso se separan eléctricamente cada dispositivo en la oblea, para lo cual ésta debe pasar por varias sustancias que se encuentran a diferentes temperaturas (figura 18). Se ha registrado que 30 obleas de cada 1.000 en promedio no resisten los cambios bruscos de temperatura y se rompen al llegar a la solución llamada *piranha*, en el subproceso de la figura 18. Se procesan 18.000 obleas al mes, de las cuales se rompen 540, lo que representa \$17.280 dólares de pérdidas mensuales aproximadamente.

**3.2.2. El experimento.** Se utilizó un experimento factorial  $2^3$  (véase figura 19) con el objetivo de localizar una combinación de temperaturas en la cual se rompan un mínimo de obleas por efecto térmico. Se eligieron sólo tres temperaturas porque algunas sustancias, como por ejemplo el ácido nítrico, hacen su trabajo a cierta temperatura durante un periodo fijo, pero no se sabe el tiempo que requieren para hacer su trabajo a otra temperatura.

Se consideró que las tres temperaturas controladas eran las de mayor efecto térmico sobre la oblea. Los tres factores controlados y sus niveles en unidades originales son las temperaturas:

$T_1$ :	Temperatura de grabado	$(-3^\circ C, -1^\circ C)$
$T_2$ :	Temperatura de piranha	$(60^\circ C, 98^\circ C)$
$T_3$ :	Temperatura de agua	$(20^\circ C, 70^\circ C)$

La combinación  $(-3^\circ C, 98^\circ C, 20^\circ C)$  era el tratamiento usual o en operación antes del experimento. La respuesta medida a cada oblea procesada en el experimento

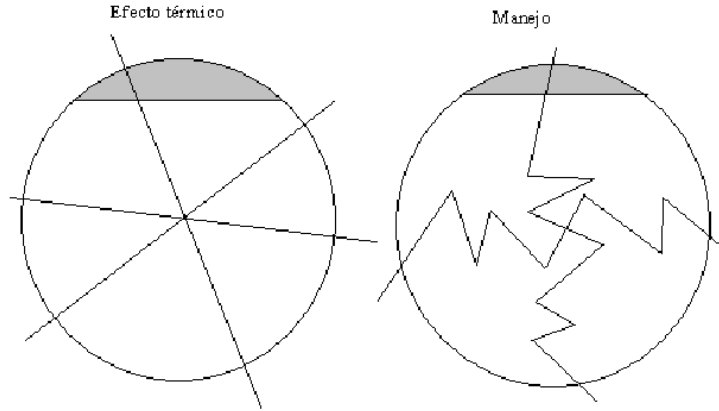


Figura 17 Roturas en obleas por diferentes causas

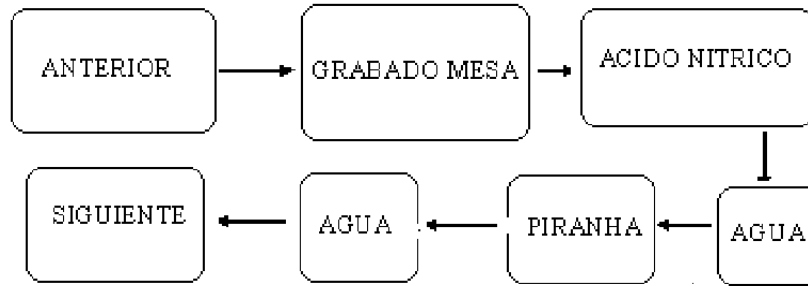


Figura 18. Subproceso objeto de estudio, en el ejemplo de obleas

toma los valores “*se rompe*” o “*no se rompe*”. Es decir, la respuesta es de tipo binaria con valor 1 si la oblea se rompe y 0 si no se rompe.

Para este tipo de variables, el tamaño de muestra se establece a partir de la estimación inicial de la magnitud del problema, de forma que en todas las corridas experimentales haya una alta probabilidad de reproducir el problema. Por ejemplo, en este caso se sabe que el número de obleas que se rompen en el tratamiento usual son 30 por cada 1000, lo que equivale a una proporción de  $p_0 = 0.03$ . La estimación del número de obleas a correr en *cada tratamiento o combinación* se calcula con la fórmula

$$(3.1) \quad m = (2.5)^2 \left( \frac{1 - p_0}{p_0} \right)$$

donde  $p_0$  es la proporción utilizada como base. En este experimento, se obtiene que  $m = 203$  obleas era suficiente, pero se decide utilizar  $m = 250$  obleas por tratamiento para poder detectar efectos un poco más pequeños con buena potencia.

Se corren dos réplicas, y las 250 obleas procesadas en cada réplica se reparten en diez lotes de tamaño 25. En total se utilizan 4000 obleas

NOTA 6. *Bisgaard y Fuller (1995) proporcionan tablas para estimar el tamaño de  $m$ , pero la regla dada en la ecuación (3.1) se ajusta bien a los tamaños de  $m$  que dichas tablas proporcionan.*

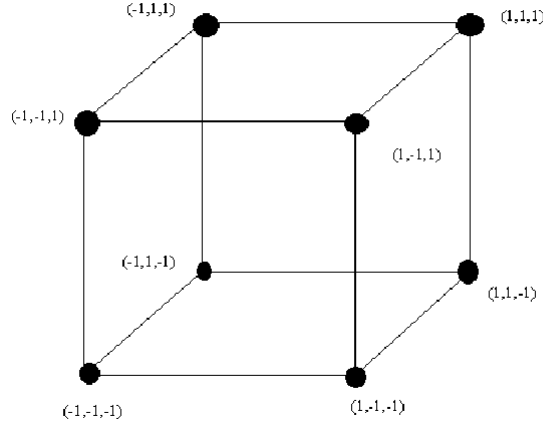
En la tabla 8 se muestran las frecuencias observadas de obleas rotas en cada lote, y se da la proporción de obleas rotas por cada 250 procesadas. Se observa a simple vista que el factor  $T_1$  parece tener un efecto importante: *en su nivel bajo se rompen notoriamente más obleas que en su nivel alto.*

Tabla 8. Obleas rotas por lote y proporción por tratamiento													
$T_1$	$T_2$	$T_3$	25	25	25	25	25	25	25	25	25	25	P
-1	-1	-1	2	0	3	0	0	3	0	0	2	0	0.040
1	-1	-1	0	0	1	0	0	0	0	2	0	0	0.012
-1	1	-1	1	0	2	0	0	2	2	0	2	0	0.036
1	1	-1	0	0	0	0	0	0	0	0	0	0	0.000
-1	-1	1	1	0	1	1	0	0	0	1	1	0	0.020
1	-1	1	0	0	0	0	0	0	0	0	0	0	0.000
-1	1	1	0	2	0	0	0	2	0	0	0	0	0.016
1	1	1	0	0	1	0	0	0	0	0	0	0	0.004
-1	-1	-1	1	0	0	3	0	0	3	0	1	0	0.032
1	-1	-1	0	0	0	1	0	0	0	0	0	1	0.008
-1	1	-1	1	0	0	1	0	3	0	0	2	0	0.028
1	1	-1	0	0	0	0	0	0	0	0	0	0	0.000
-1	-1	1	0	1	0	1	0	1	0	1	0	1	0.020
1	-1	1	0	1	0	0	1	0	0	0	1	1	0.016
-1	1	1	0	0	2	0	0	0	0	0	0	0	0.008
1	1	1	0	0	1	0	0	0	0	0	0	0	0.004

Al aplicar las fórmulas para estimar efectos con base en los contrastes, que se describieron en la sección anterior, se estima cada uno de los efectos y se obtiene la siguiente tabla:

Efecto	Estimación
A:T Grabado	-0,0195
B:T Piranha	-0,0065
C:T Agua	-0,0085
AB	-0,0005
AC	0,0095
BC	0,0005
ABC	0,0025

Estos efectos pueden graficarse en un diagrama de Pareto para así visualizar cuáles efectos tienen un mayor impacto sobre la variable de respuesta (ver figura 20). La figura muestra el *diagrama de Pareto estandarizado*, en el cual se representan los efectos divididos entre su error estándar. En general, para un diseño  $2^k$  con  $n$  réplicas la *estimación del error estándar de un efecto* ( $\hat{\sigma}$ ) y el efecto estandarizado

Figura 19. Diseño factorial  $2^3$  para obleas

( $ES$ ) vienen dado por:

$$\begin{aligned}\hat{\sigma} &= \sqrt{\frac{CME}{n2^{k-2}}} \\ ES &= Efecto / \hat{\sigma}\end{aligned}$$

El caso del efecto principal de A para el ejemplo de obleas rotas, tenemos que:

$$ES(A) = -0.0195 / \sqrt{\frac{0.000029}{2 \times 2^{3-2}}} = -7.2421$$

donde el  $CME$  se obtiene de la tabla de análisis de varianza con todos los efectos incluidos (véase tabla 9a). De manera similar se obtienen los efectos estandarizados para los demás efectos y se representan gráficamente en el diagrama de Pareto de la figura 20. El efecto estandarizado sirve de estadístico de prueba para probar la hipótesis:

$$\begin{aligned}H_0 &: \text{Efecto poblacional} = 0 \\ H_1 &: \text{Efecto poblacional} \neq 0.\end{aligned}$$

Se rechaza  $H_0$ , si  $|ES| > t_{\alpha/2, v}$ , donde  $\alpha$  es el nivel de significancia prefijado para la prueba (generalmente  $\alpha = 0.05$ ), y  $v$  son los grados de libertad asociados al error. En el ejemplo de las obleas, como el error tiene 8 grados de libertad y trabajando con  $\alpha = 0.05$ , se tiene que  $t_{0.25, 8} = 2.306$ . Por tanto, si el valor absoluto de la estimación del efecto estandarizado es mayor que 2.306, entonces el efecto poblacional correspondiente será estadísticamente diferente de cero. Si se agrega una línea en el diagrama de Pareto estandarizado a la altura del valor crítico como se muestra en la figura 20, los efectos cuyas barras superen tal línea serán los efectos significativos. De esta forma, para el caso de las obleas, los efectos significativos serán A, C y la interacción AC.

**3.2.3. El mejor ANOVA..** El ANOVA con todos los efectos se muestra en la tabla 9a, ahí se aprecia que los efectos que tienen un *valor - p* menor que 0.05, son: A, C y AC, que coincide con lo visto en el diagrama de Pareto estandarizado.

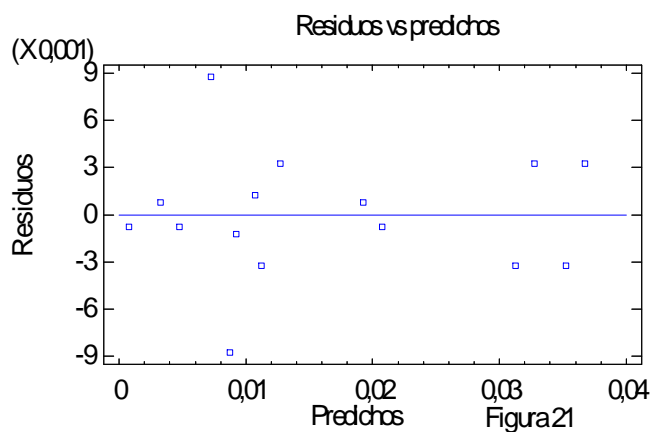
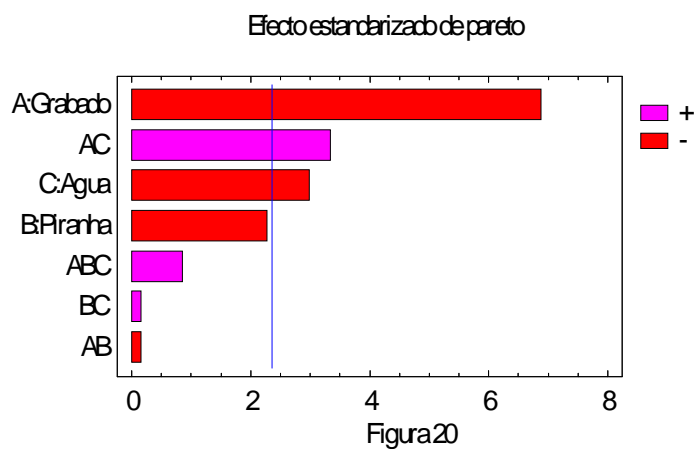
Los factores que no resultan significativo se excluyen del Anova para obtener un modelo final en el que sólo se incluyan términos significativos, construyendo el mejor ANOVA, en el que en una primer ronda se eliminan del análisis y se mandan al error a los efectos que claramente no son significativos. Después de esta primer ronda se revalora a los términos que estaban en una situación dudosa, como era el caso del efecto B en la tabla 9a. En una segunda o tercer ronda se eliminan los términos que no resultaron significativos después de la(s) ronda(s) inicial(es).

Tabla 9a. ANOVA completo para la proporción de obleas rotas					
Fuente	SC	gl	CM	F <sub>0</sub>	valor-p
A:T_GRAB	0,001521	1	0,001521	52,45	0,0001
B:T_PIRA	0,000169	1	0,000169	5,83	0,0422
C:T_AGUA	0,000289	1	0,000289	9,97	0,0135
AB	0,000001	1	0,000001	0,03	0,8573
AC	0,000361	1	0,000361	12,45	0,0078
BC	0,000001	1	0,000001	0,03	0,8573
ABC	0,000025	1	0,000025	0,86	0,3803
Error	0,000232	8	0,000029		
Total	0,00259915	15			
$R^2 = 91,0735$	$R^2_{ajustado} = 83,2628$				

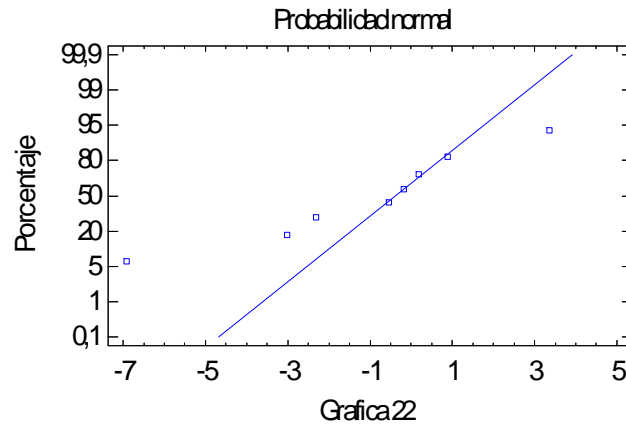
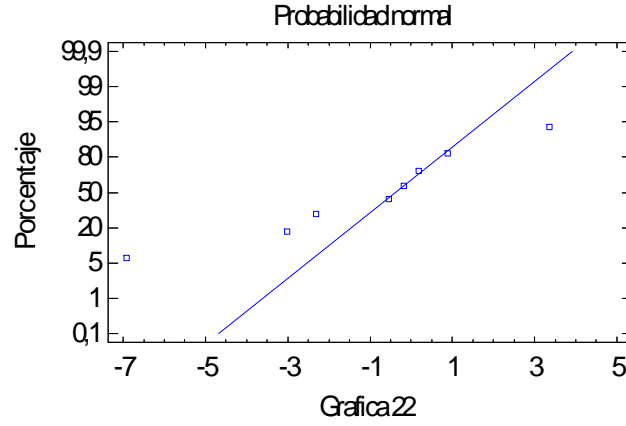
Al hacer lo anterior en el caso del ANOVA de la tabla 9a, es claro que los efectos *AB*, *BC* y *ABC* son no significativos, por lo que se eliminan y se mandan al error, con lo cual se obtiene el análisis de varianza de la tabla 9b. En la que se observa que ahora el efecto B sí es significativo. De esta manera, como este ANOVA sólo tiene términos significativos, podemos considerarlo como el mejor. También al final de este análisis se obtiene el coeficiente de determinación, que señala que un alto porcentaje de la variabilidad observada en el experimento (proporción de obleas rotas) es explicada por los efectos considerados en la tabla 9b.

Tabla 9b. El mejor ANOVA completo para la proporción de obleas rotas					
Fuente	SC	gl	CM	F <sub>0</sub>	valor-p
A:T_GRAB	0,001521	1	0,001521	64.60	0,0000
B:T_PIRA	0,000169	1	0,000169	7.18	0,0214
C:T_AGUA	0,000289	1	0,000289	12.27	0,0049
AC	0,000361	1	0,000361	15.33	0,0024
Error	0,000259	11	0,0000235455		
Total	0,0025999	15			
$R^2 = 90,0346$	$R^2_{ajustado} = 86,4109$				

En las figuras 21 y 22 se grafican los residuos contra predichos y los residuos en papel normal y se observa el cumplimiento de los supuestos de varianza constante y normalidad.



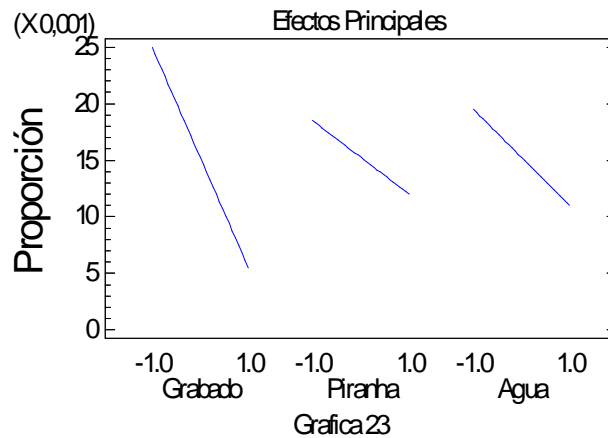
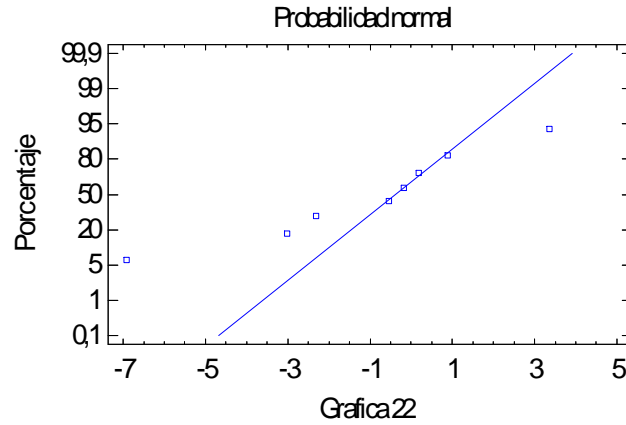
**3.2.4. Conclusiones e impacto económico.** Se detectaron como significativos los efectos  $A$ ,  $B$ ,  $C$  y  $AC$ . De acuerdo con la figura 24, se aprecia una interacción importante entre la *temperatura de grabado* y la *temperatura de agua*, en particular se observa que si se trabaja con temperatura alta de grabado, prácticamente da lo mismo trabajar con cualquiera de las dos temperaturas de agua. Además, en la figura 23 se muestran los efectos principales, por lo que de aquí y de la gráfica de interacción se concluye que hay dos condiciones de las tres temperaturas que minimiza el número de obleas rotas:  $(1, 1, 1) = (-1^{\circ}C, 98^{\circ}C, 70^{\circ}C)$ , y  $(1, 1, -1) = (-1^{\circ}C, 98^{\circ}C, 20^{\circ}C)$ . Esto también se puede apreciar en la gráfica de cubo de la figura 25, en la que se aprecia la respuesta predicha en cada punto del diseño. En este mismo cubo, comparando la respuesta predicha en la cara lateral derecha contra la cara de la izquierda del cubo, se nota que todos los tratamientos donde la temperatura de grabado ( $T - GRAB$ ) trabaja en su nivel alto, rompen



menos obleas que cuando esta temperatura se encuentra en su valor bajo. La respuesta predicha en los dos puntos óptimos es cercana a 0.0% de obleas rotas, lo que contrasta con la respuesta en la combinación de temperaturas que se utilizaban antes del experimento, que era  $(T_1, T_2, T_3) = (-1, 1, -1)$ , en donde se pronostica un 3% de obleas rotas.

$$\begin{aligned}\hat{Y}(1, 1, 1) &= 0.00275. \\ \hat{Y}(1, 1, -1) &= 0.00125.\end{aligned}$$

NOTA 7. Con base en lo anterior, se decidió implementar uno de los dos mejores tratamientos y se redujo el número de obleas rotas por efecto térmico, aunque la reducción no fue la pronosticada por el modelo de ANOVA. Esto se debió a que hay obleas que traen pequeñas fallas estructurales heredadas de procesos previos y se rompen como si la única causa fuera el efecto térmico. Las obleas utilizadas

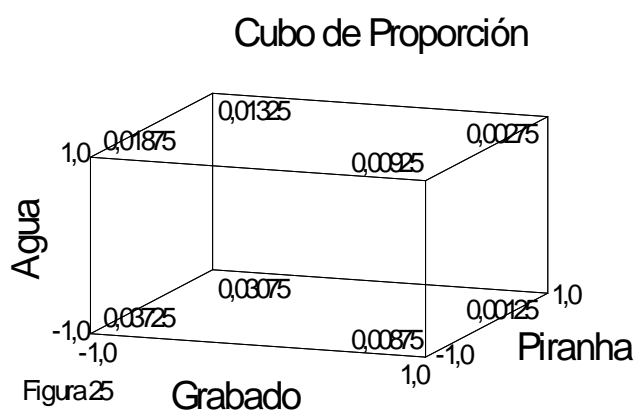
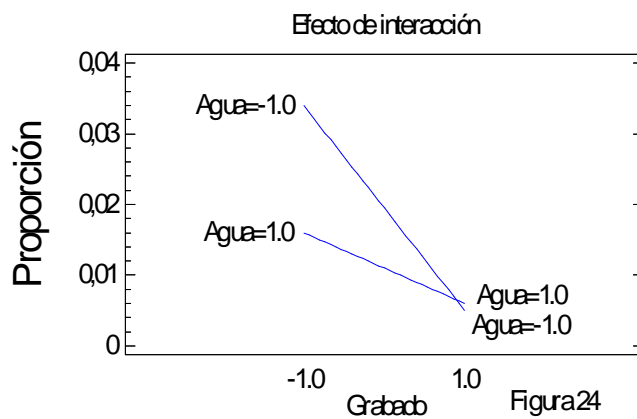


en el experimento no tenían este problema. Desde el punto de vista económico, es importante mencionar que antes del experimento se rompían 30 obleas por cada mil procesadas, lográndose reducir este número a 15 por cada mil (véase figura 26). Esta mejora representa un ahorro aproximado de U\$8000 dólares mensuales. Este resultado ha motivado el uso del diseño de experimentos en toda la empresa, no sólo en el área de obleas.

#### 4. Diseño Factorial General $2^k$

En el diseño  $2^k$  se consideran  $k$  factores con dos niveles cada uno, y se tienen  $2^k$  tratamientos o puntos de diseño. Las  $k$  columnas y  $2^k$  renglones que componen la matriz para este diseño, considerando una réplica, se pueden construir de la siguiente manera: en la primer columna, que corresponde a los niveles del factor A, se alternan signos  $+$  y  $-$ , empezando con  $-$ , hasta llegar a los  $2^k$  renglones;





en la segunda columna se alternan dos signos  $-$  con dos signos  $+$ , en la tercera cuatro signos  $-$  y cuatro signos  $+$ , y así sucesivamente, hasta la  $k$ -ésima columna compuesta por  $2^{k-1}$  signos  $-$  seguidos de  $2^{k-1}$  signos  $+$ .

Con el diseño factorial completo  $2^k$  se pueden estudiar en total los  $2^k - 1$  efectos siguientes:

$$\begin{aligned}
 \binom{k}{1} &= k, \text{ efectos principales} \\
 \binom{k}{2} &= \frac{k!}{2!(k-2)!} = \frac{k(k-1)}{2}, \text{ interacciones dobles} \\
 \binom{k}{3} &= \frac{k!}{3!(k-3)!}, \text{ interacciones triples} \\
 &\vdots \\
 \binom{k}{k} &= 1, \text{ interacciones de los } k \text{ factores}
 \end{aligned}$$

donde la operación  $\binom{k}{r} = \frac{k!}{r!(k-r)!}$  son las “*combinaciones de  $k$  en  $r$* ” y cuantifica el número de maneras diferentes en las que se pueden elegir  $r$  objetos de un grupo de  $k$  ( $r \leq k$ ). Aplicando esta operación se puede contar el número de interacciones de cada tipo que existen en el factorial completo  $2^k$ . Por ejemplo, el diseño factorial  $2^5$  tiene  $\binom{5}{1} = 5$  efectos principales,  $\binom{5}{2} = 10$  interacciones dobles,  $\binom{5}{3} = 10$  interacciones triples,  $\binom{5}{4} = 5$  interacciones cuádruples y una interacción quíntuple, lo que da un total de  $2^5 - 1 = 31$  efectos.

**4.1. Estimación de contrastes, efectos y sumas de cuadrados.** Cada uno de los efectos se estima a partir de su contraste, el cual a su vez se puede obtener construyendo la tabla de signos del diseño. Una vez construida la tabla de signos, el contraste de cada efecto se obtiene al multiplicar su columna de signos por la columna de totales expresados en la notación de Yates. Con los contrastes se procede a estimar los efectos mediante la fórmula

$$(4.1) \quad \text{Efecto } ABC \cdots K = \frac{1}{n2^{k-1}} [\text{Contraste } ABC \cdots K]$$

para la cual la suma de cuadrados con un grado de libertad esta dada por

$$(4.2) \quad SC_{ABC \cdots K} = \frac{1}{n2^k} [\text{Contraste } ABC \cdots K]^2$$

donde  $n$  es el número de réplicas del experimento.

**4.2. Diseño factorial  $2^5$ .** La tabla de signos para los 15 efectos potencialmente más importantes en el diseño factorial  $2^5$  se muestra en las tablas 10 y 11. Estos 15 efectos son los cinco efectos principales más las 10 interacciones dobles.

Tabla 10 Familia de diseños factoriales $2^k$ ( $k \leq 5$ )													
Tto	Yates	A	B	C	D	E	Tto	Yates	A	B	C	D	E
1	(1)	−	−	−	−	−	17	e	−	−	−	−	+
2	<i>a</i>	+	−	−	−	−	18	ae	+	−	−	−	+
3	<i>b</i>	−	+	−	−	−	19	be	−	+	−	−	+
4	<i>ab</i>	+	+	−	−	−	20	abe	+	+	−	−	+
5	<i>c</i>	−	−	+	−	−	21	ce	−	−	+	−	+
6	<i>ac</i>	+	−	+	−	−	22	ace	+	−	+	−	+
7	<i>bc</i>	−	+	+	−	−	23	bce	−	+	+	−	+
8	<i>abc</i>	+	+	+	−	−	24	abce	+	+	+	−	+
9	<i>d</i>	−	−	−	+	−	25	de	−	−	−	+	+
10	<i>ad</i>	+	−	−	+	−	26	ade	+	−	−	+	+
11	<i>bd</i>	−	+	−	+	−	27	bde	−	+	−	+	+
12	<i>abd</i>	+	+	−	+	−	28	abde	+	+	−	+	+
13	<i>cd</i>	−	−	+	+	−	29	cde	−	−	+	+	+
14	<i>acd</i>	+	−	+	+	−	30	acde	+	−	+	+	+
15	<i>bcd</i>	−	+	+	+	−	31	bcde	−	+	+	+	+
16	<i>abcd</i>	+	+	+	+	−	32	abcde	+	+	+	+	+

Tabla 11 Familia de diseños factoriales  $2^k$  ( $k \leq 5$ )

Tto	Yates	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
1	(1)	+	+	+	+	+	+	+	+	+	+
2	<i>a</i>	-	-	-	-	-	+	+	+	+	+
3	<i>b</i>	-	+	+	+	-	-	-	+	+	+
4	<i>ab</i>	+	-	-	-	-	-	-	+	+	+
5	<i>c</i>	+	-	+	+	-	+	+	-	-	+
6	<i>ac</i>	-	+	-	-	-	+	+	-	-	+
7	<i>bc</i>	-	-	+	+	+	-	-	-	-	+
8	<i>abc</i>	+	+	-	-	+	-	-	-	-	+
9	<i>d</i>	+	+	-	+	+	-	+	-	+	-
10	<i>ad</i>	-	-	+	-	+	-	+	-	+	-
11	<i>bd</i>	-	+	-	+	-	+	-	-	+	-
12	<i>abd</i>	+	-	+	-	-	+	-	-	+	-
13	<i>cd</i>	+	-	-	+	-	-	+	+	-	-
16	<i>acd</i>	-	+	+	-	-	-	+	+	-	-
15	<i>bcd</i>	-	-	-	+	+	+	-	+	-	-
16	<i>abcd</i>	+	+	+	-	+	+	-	+	-	-

Tabla 11 Factoriales  $2^k$  y sus efectos de interés ( $k \leq 5$ )

Tto	Y	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
17	<i>e</i>	+	+	+	-	+	+	-	+	-	-
18	<i>ae</i>	-	-	-	+	+	+	-	+	-	-
19	<i>be</i>	-	+	+	-	-	-	+	+	-	-
20	<i>abe</i>	+	-	-	+	-	-	+	+	-	-
21	<i>ce</i>	+	-	+	-	-	+	-	-	+	-
22	<i>ace</i>	-	+	-	+	-	+	-	-	+	-
23	<i>bce</i>	-	-	+	-	+	-	+	-	+	-
24	<i>abce</i>	+	+	-	+	+	-	+	-	+	-
25	<i>de</i>	+	+	-	-	+	-	-	-	-	+
26	<i>ade</i>	-	-	+	+	+	-	-	-	-	+
27	<i>bde</i>	-	+	-	-	-	+	+	-	-	+
28	<i>abde</i>	+	-	+	+	-	+	+	-	-	+
29	<i>cde</i>	+	-	-	-	-	-	-	+	+	+
30	<i>acde</i>	-	+	+	+	-	-	-	+	+	+
31	<i>bcde</i>	-	-	-	-	+	+	+	+	+	+
32	<i>abcde</i>	+	+	+	+	+	+	+	+	+	+

#### 4.3. ANOVA del diseño factorial $2^k$ .

(1) La suma de cuadrados totales (*SCT*) se calcula como:

$$SCT = \sum_{i=1}^{n2^k} Y_i^2 - \frac{1}{2^k n} Y^2$$

y tiene  $2^k n - 1$  grados de libertad, donde el subíndice  $i$  corre sobre el total de observaciones. La suma de cuadrados del error (*SCE*) se obtiene por diferencia y tiene  $2^k (n - 1)$  grados de libertad. Se prueba la hipótesis

$$H_0 : Efecto = 0 \text{ vs } H_1 : efecto \neq 0.$$

- (2) Cuando la tabla de ANOVA incluye todos los efectos que se pueden estimar con el factorial completo  $2^k$  es necesario realizar al menos dos réplicas para poder estimar *SCE*. En la mayoría de los casos sólo interesa estudiar a los efectos principales y las interacciones dobles cuando el número de factores es mayor o igual a cuatro, luego no es necesario realizar réplicas, ya que se puede construir una *SCE* aproximada utilizando las sumas de cuadrados de las interacciones triples o de mayor orden que generalmente son pequeñas.
- (3) Se supone que la respuesta es aproximadamente lineal en el rango de variación de cada uno de los factores estudiados. No es necesario suponer una linealidad perfecta, pero sí que no haya una curvatura muy grande. De esta manera, dado que cada factor se prueba en dos niveles, no es posible estudiar efectos de curvatura (efectos del tipo  $A^2, B^2$ , etc.), aunque ésta exista en el proceso.

### 5. Diseño factorial $2^k$ no replicado

A1 aumentar el número de factores en el diseño  $2^k$  crece rápido el número de tratamientos y por lo tanto el número de corridas experimentales. Si se realizan dos repeticiones en cada punto se tienen que hacer  $2 \times 2^k$  corridas experimentales, lo cual puede consumir muchos recursos para más de cuatro factores ( $k \geq 5$ ), ya que serían 64 o más corridas. El diseño  $2^4$  es quizá el factorial más grande que todavía se puede correr con dos réplicas, lo que implica hacer 32 corridas del proceso, pero incluso con una réplica (16 pruebas) de este diseño es muchas veces suficiente para estudiar los 10 efectos de interés (tabla 12). De acuerdo con esta tabla, cuando son cuatro factores se recomienda en primera instancia correr una sola réplica del diseño; cuando son cinco factores se recomienda correr una réplica de la mitad del diseño (fracción  $2^{5-1}$ ), y después de analizar esta primera mitad se decide completar a una réplica del factorial  $2^5$  completo; de seis factores en adelante el diseño siempre se corre fraccionado y sólo una réplica de la fracción elegida.

Tabla 12. Réplicas recomendadas en la familia de diseños $2^k$		
Diseño	Replicas recomendadas	Numero de corridas
$2^2$	3 ó 4	12, 16
$2^3$	2	16
$2^4$	1 ó 2	16, 32
$2^5$	fracción $2^{5-1}$ o 1	16, 32
$2^6$	fracción $2^{6-2}$ o fracción $2^{6-1}$	16, 32
$2^7$	fracción $2^{7-3}$ o fracción $2^{7-2}$	16, 32

Más aún, un máximo de 16 pruebas son suficientes para la mayoría de los problemas en una primera etapa de experimentación.

**5.1. ¿Cómo decidir cuáles efectos mandar al error?** Algunos aspectos a tener en cuenta cuando se corre sólo una réplica del experimento y se quiere aproximar al CME:

- (1) Si mandamos al error un efecto que no se debe al azar, que es un efecto real, éste puede inflar el CME lo que reduciría la potencia del ANOVA para detectar efectos significativos (sería como echar una cortina de humo

que no dejaría ver qué sucede con los efectos); por otro lado, si el error resulta exageradamente pequeño, se estarían detectando como significativos efectos que no lo son, lo cual puede llevar a decisiones incorrectas.

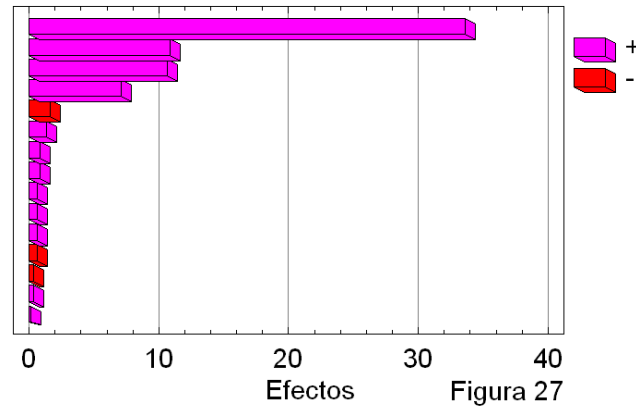
- (2) Si el cuadrado medio del error resulta muy diferente de la  $\sigma^2$  histórica de la misma respuesta, es un síntoma de que posiblemente no está bien estimado.
- (3) En cualquier experimento pueden existir tres tipos de efectos: *los que claramente afectan, los que claramente no afectan y los efectos intermedios*. Cuando existen sólo efectos de los dos primeros es fácil decidir con cuáles efectos conformar el error. El problema es que cuando existen efectos intermedios, no está nada claro si afectan o no a la respuesta.
- (4) Las técnicas para decidir qué efectos mandar al error no funcionan bien cuando los efectos que tiene el diseño son pocos. Por ejemplo, los diseños  $2^2$  y  $2^3$  no tendrían suficientes efectos para que estos métodos realmente tengan utilidad práctica. Pero estos diseños se corren frecuentemente con réplicas suficientes para construir el análisis de varianza directamente (tabla 12), evitando la necesidad de construir un error hasta cierto punto arbitrario. Las técnicas que se describen a continuación tienen mayor utilidad en diseños con cuatro factores en adelante, en los que existen 10 o más efectos a investigar. Estas técnicas se describen en seguida y se ejemplifica su uso en la siguiente sección.

5.1.1. *Gráfico de efectos en papel normal (Daniel's plot)*. Al usar los efectos como sumas de variables aleatorias (diferencia de medias), Daniel (1959) se dio cuenta que los efectos no significativos deben seguir una distribución normal con media igual a cero y varianza constante. Esto implica que si los efectos se grafican en *papel probabilístico normal*, los que no son significativos tenderán a formar una línea recta en esta gráfica ubicada a la altura del cero, lo que permite confirmar que tales efectos son efectivamente insignificantes. Por su parte, los efectos activos aparecerán alejados de la línea de normalidad, lo que indica que no se deben sólo al azar, sino a la existencia de efectos reales que influyen en la respuesta.

- Cuando se tienen *efectos positivos y negativos* puede ser mejor utilizar el papel *probabilístico medio normal (half normal)*, para tener una mejor perspectiva de cuáles efectos se alinean y cuáles no. Como su nombre lo indica, el papel medio normal utiliza sólo la parte positiva de la distribución normal estándar aprovechando su simetría y el hecho de que dos efectos de signo contrario y de la misma magnitud son igualmente importantes.
- Como se ha visto anteriormente, el papel probabilístico normal también sirve para verificar el cumplimiento del supuesto de normalidad de los residuos. La gráfica de efectos en papel normal tiene un objetivo muy diferente a esta gráfica de residuos.

5.1.2. *Diagrama de Pareto de efectos*. El diagrama de Pareto para los efectos sin estandarizar representa una manera práctica de ver cuáles efectos son los más grandes en cuanto a su magnitud. *El Pareto representa la realidad observada de los efectos de manera descriptiva sin considerar supuestos distribucionales*. En la gráfica de efectos en papel de probabilidad normal (gráfico de Daniel) es más difícil apreciar la importancia relativa de los efectos, pero es mejor que el Pareto para señalar cuáles efectos son activos. Lo mejor es utilizar ambas gráficas para decidir de la mejor manera cuáles efectos mandar al error. Muchas veces con el Pareto y

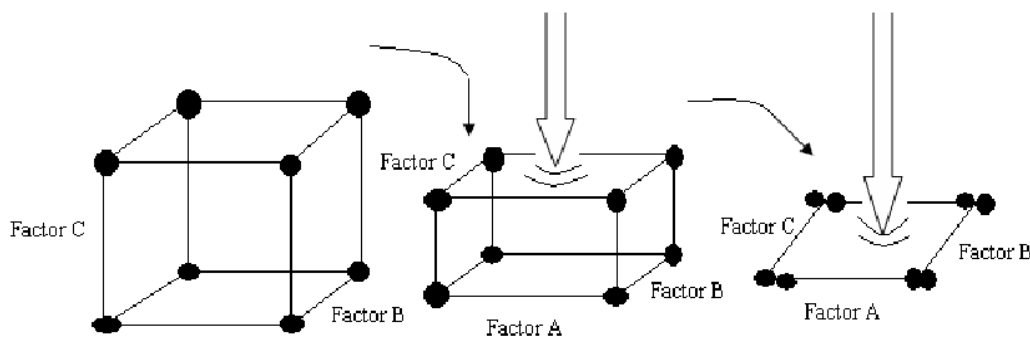
### Diagrama de Pareto



el gráfico de Daniel se logran detectar claramente los efectos significativos, y una vez que se construya el error el ANOVA sólo confirmará lo que ya se ha encontrado con estos gráficos (véase figura 27a).

5.1.3. *Otros criterios útiles.* Cuando se estiman varios efectos en un experimento, ocurre que algunos de ellos son significativos por su valor grande, otros son no significativos debido a su pequeña magnitud; pero cuando hay un tercer grupo de efectos cuya magnitud no es "ni muy grande ni muy chica", se vuelve un problema decidir si deben mandarse al error o no. En ocasiones, ni el gráfico de Daniel ni el diagrama de Pareto aclaran bien la situación de estos efectos de magnitud intermedia. En estos efectos se requiere utilizar otros criterios o aspectos adicionales, para poder decidir si mandarlos o no a engrosar las filas del error. Se recomienda fijarse en todos estos criterios y no solamente uno de ellos:

- 1.: **La magnitud del efecto.** Por experiencia en el proceso, la magnitud del efecto puede indicar si éste se debe sólo al azar, en cuyo caso se mandaría al error, o si se debe a un efecto real. Se puede comparar el efecto observado con el error estándar basado en una varianza histórica. Esto es, en el factorial  $2^k$  con una réplica se compara el efecto observado contra dos veces el error estándar del efecto.  $(\sigma/\sqrt{2^{k-2}})$ , y si el primero es más grande es porque puede ser un efecto real.
- 2.: Los efectos cuyas significancias en ANOVAS preliminares andan alrededor de 0.2 o menores no se excluyen del análisis necesariamente. Esta decisión es más confiable cuando dichos ANOVAS preliminares ya alcanzaron al menos 8 *grados de libertad para el error*. Si primero se excluyen los efectos que son claramente no significativos de acuerdo al gráfico de Daniel y al Pareto, se pueden lograr ANOVAS cuyas significancias dan información útil para excluir o no los efectos restantes.
- 3.: Los grados de libertad del error, deben ser al menos 8 para tener un ANOVA más confiable.

Figura 28 Acción de colapsar un factorial  $2^3$ 

4.: Cuando se van eliminando efectos que no son significativos, el estadístico  $R^2_{ajustado}$  crece. En el momento en que se elimina un efecto y este estadístico decrece, significa que posiblemente ese efecto no debe excluirse, aunque también se toma en cuanto la magnitud del decrecimiento. Se requiere que el decrecimiento mencionado sea de cuando menos 3% para que valga la pena incluir otra vez el efecto.

**5.2. Colapsación o proyección del diseño.** Cuando en el mejor ANOVA que se pudo determinar se detecta que un factor particular no es significativo, ya que su efecto principal y todas las interacciones en las que interviene no son importantes, entonces en lugar de mandar al error este factor y sus interacciones, otra posibilidad es colapsar o proyectar el diseño, que consiste en eliminar completamente del análisis a tal factor, con lo que el diseño factorial  $2^k$  original se convierte en un diseño completo con un factor menos ( $2^{k-1}$ ) y con dos repeticiones en cada punto. Al haber “repeticiones” en el diseño  $2^{k-1}$ , resultante de la colapsación, entonces se puede estimar el  $CME$  y construir la tabla de análisis de varianza de la manera usual. En general, si se pueden omitir  $h$  factores, los datos se convierten en un diseño factorial  $2^{k-h}$  con  $2^h$  repeticiones en cada punto. Por ejemplo, si se eliminan dos factores, el diseño  $2^k$  con una réplica, entonces el diseño resultante es un diseño factorial completo con  $k - 2$  factores y cuatro réplicas.

El efecto de colapsar un diseño factorial  $2^3$  se representa en la figura 28. Es como construir un cubo de cartón que representa el diseño  $2^3$ , colocarlo en el piso cuidando que el factor que no afecta quede en el sentido vertical, y brincar sobre él para aplastarlo. El resultado es un diseño factorial  $2^2$  con el doble de réplicas que tenía el diseño original. Si se tenía una réplica, colapsando un factor se tienen dos réplicas. Debemos aclarar que la acción de colapsar facilita la reproducción posterior del análisis del experimento, al eliminar del análisis los factores que se sabe no tienen influencia.

### 5.3. Ejemplo ilustrativo: Experimento $2^5$ no replicado. Problema.

Una planta donde se fabrican semiconductores quiere mejorar el rendimiento del proceso vía diseño de experimentos. De acuerdo con la experiencia del grupo de



mejora, los factores que podían tener mayor influencia sobre la variable de respuesta (rendimiento), así como los niveles de prueba utilizados, son los siguientes:

- $A$  = Nivel de la abertura (pequeña, grande)
- $B$  = Tiempo de exposición (20% abajo, 20% arriba)
- $C$  = Tiempo de revelado (30 seg, 45 seg)
- $D$  = Dimensión de la máscara (pequeña, grande)
- $E$  = Tiempo de grabado (14.5 min, 15.5 min)

Se corre un experimento  $2^5$  con una sola réplica para estudiar sus cinco factores. En la tabla 13 se muestra la matriz de diseño en orden aleatorio, construida con *Statgraphics*. Esta matriz debe guardarse en un archivo para no perder el orden aleatorio seleccionado y poder capturar los datos una vez hechas las corridas. Se hacen las 32 corridas y los resultados escritos en el orden estándar se muestran en la tabla 14.

Tabla 13. Matriz de diseño en orden aleatorio													
Orden	A	B	C	D	E	Y	Orden	A	B	C	D	E	Y
1	1.0	1.0	-1.0	-1.0	-1.0	55	17	-1.0	-1.0	-1.0	1.0	1.0	-
2	-1.0	1.0	1.0	1.0	-1.0	44	18	-1.0	-1.0	-1.0	-1.0	1.0	-
3	1.0	1.0	1.0	1.0	-1.0	61	19	1.0	-1.0	-1.0	-1.0	1.0	-
4	-1.0	-1.0	-1.0	-1.0	-1.0	-	20	-1.0	-1.0	1.0	-1.0	-1.0	-
5	-1.0	1.0	-1.0	1.0	-1.0	-	21	1.0	-1.0	1.0	-1.0	1.0	-
6	1.0	-1.0	1.0	-1.0	-1.0	-	22	1.0	1.0	-1.0	1.0	-1.0	-
7	-1.0	1.0	1.0	-1.0	1.0	-	23	1.0	-1.0	-1.0	1.0	-1.0	-
8	-1.0	-1.0	-1.0	1.0	-1.0	-	24	-1.0	-1.0	1.0	1.0	1.0	-
9	1.0	1.0	-1.0	-1.0	1.0	-	25	1.0	1.0	1.0	-1.0	-1.0	-
10	-1.0	-1.0	1.0	-1.0	1.0	-	26	1.0	-1.0	-1.0	1.0	1.0	-
11	-1.0	1.0	1.0	-1.0	-1.0	-	27	-1.0	1.0	-1.0	1.0	1.0	-
12	-1.0	1.0	-1.0	-1.0	1.0	-	28	-1.0	1.0	-1.0	-1.0	-1.0	-
13	-1.0	-1.0	1.0	1.0	-1.0	-	29	-1.0	1.0	1.0	1.0	1.0	-
14	1.0	1.0	1.0	1.0	1.0	-	30	1.0	1.0	1.0	-1.0	1.0	-
15	1.0	-1.0	1.0	1.0	1.0	-	31	1.0	-1.0	1.0	1.0	-1.0	-
16	1.0	1.0	-1.0	1.0	-1.0	-	32	1.0	-1.0	-1.0	-1.0	-1.0	-

Tabla 14. Datos acomodados en el orden estandar.				Tabla 15. Efectos estimados.	
(1) = 7	$d = 8$	$e = 8$	$de = 6$	A: Abertu=11,75	AE=1,0
$a = 9$	$ad = 10$	$ae = 12$	$ade = 10$	B: T-expo=34	BC=0,125
$b = 34$	$bd = 32$	$be = 35$	$bde = 30$	C: T-revel= 9,75	BD=-0,75
$ab = 55$	$abd = 50$	$abe = 52$	$abde = 53$	D: máscara=-0,875	BE= 0,5
$c = 16$	$cd = 18$	$ce = 15$	$cde = 15$	E: T-grab=0,375	CD=0,75
$ac = 20$	$acd = 21$	$ace = 22$	$acde = 20$	AB= 7,875	CE=0,25
$bc = 41$	$bcd = 44$	$bce = 45$	$bcde = 41$	AC= 0,375	DE=-1,125
$abc = 60$	$abcd = 61$	$abce = 65$	$abcde = 63$	AD= 0,0	media global=30,5625

**Análisis de experimento.** Se estiman los efectos más importantes: cinco efectos principales y  $5(5-1)/2 = 10$  interacciones doble. Todos los efectos de las interacciones triples son pequeños, por ello en la tabla 15 sólo reporta los efectos principales y las interacciones dobles. El análisis de varianza preliminar de la tabla 16 sólo incluyen efectos principales y se han mandado al error las interacciones de tres factores en adelante, esto permite generar 16 grados de libertad del error

Estimación del error estándar de un efecto =  $\hat{\sigma}_{efecto} = \sqrt{CME/n2^{k-2}} = \sqrt{2.48/1 \times 2^{5-2}} = 0.557$

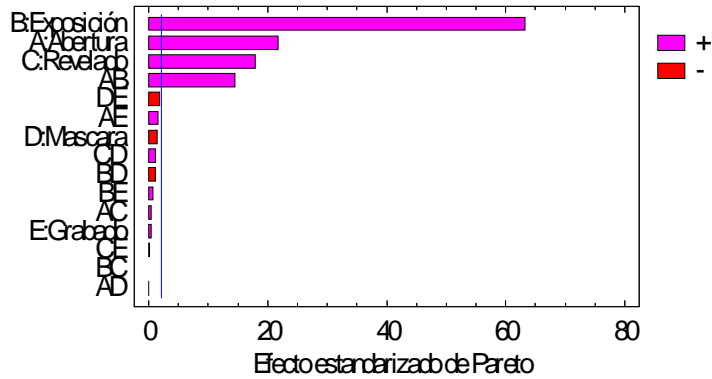
y recordemos que si un efecto es más grande que el doble del error estándar, es síntoma de que el efecto puede ser importante. De acuerdo con la tabla 15, en este caso sólo caen los efectos A, B, C, AB; y seguiría en duda la interacción DE.

Tabla 16. ANOVA preliminar para los semiconductores					
Efectos	SC	G:L	C:M	$F_0$	valor-p
A: Abertu	1104,5	1	104,5	482,51	0,0000
B: T-expo	9248,0	1	9248,0	4040,08	0,0000
C: T-revel	760,5	1	760,5	332,23	0,0000
D: máscara	6,125	1	6,125	2,68	0,1214
E: T-grab	1,125	1	1,125	0,49	0,4933
AB	496,125	1	496,125	216,74	0,0000
AC	1,125	1	1,125	0,49	0,4933
AD	0,0	1	0,0	0,00	1,0000
AE	8,0	1	8,0	3,49	0,0800
BC	0,125	1	0,125	0,05	0,8182
BD	4,5	1	4,5	1,97	0,1800
BE	2,0	1	2,0	0,87	0,3638
CD	4,5	1	4,5	1,97	0,1800
CE	0,5	1	0,5	0,22	0,6465
DE	10,125	1	10,125	4,42	0,0516
Error	36,625	16	2,28906		
Total	11683,9	31			

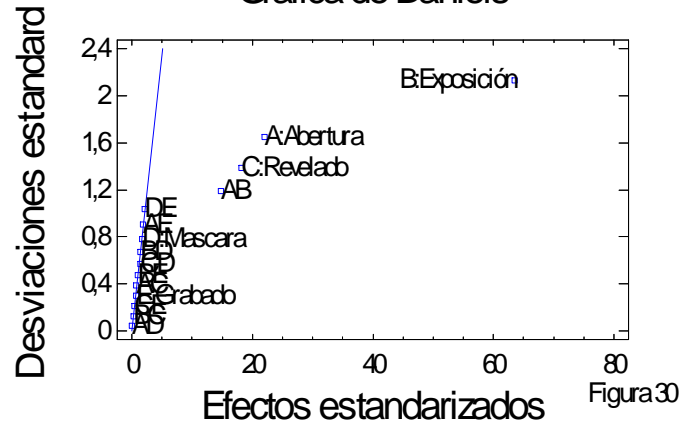
**5.4. Diagrama de Pareto y gráfico de Daniel.** Con la idea de despejar dudas y llegar al mejor ANOVA, lo que sigue es eliminar los efectos menos importantes en la tabla 15 y con valor-ps grandes en análisis de varianza preliminar de la tabla 16. Los gráficos de Pareto y de Daniel confirman claramente lo que ya se sabe: sólo hay cuatro efectos significativos. En la figura 29 se muestra el diagrama de Pareto para los 15 efectos estimados. Se observa que el gráfico hace un trabajo limpio en detectar a los efectos que están activos y a los efectos que deben eliminarse para obtener el mejor y definitivo análisis de varianza; los efectos activos son A, B, C y AB. De aquí se puede pasar al mejor ANOVA, pero antes veamos lo que dice el gráfico de Daniel.

Las mismas cuatro barras más grandes en el diagrama de Pareto corresponden a los cuatro puntos (efectos) que se alejan de la línea en el papel normal (véase figura 30) y que es señal de que dichos efectos no se deben a causas aleatorias, sino que son efectos reales en el proceso. Los efectos alineados se deben al azar, por ello se pueden mandar al error. Entonces ya está claro cuáles efectos eliminar para llegar al mejor ANOVA. Note que en este ejemplo todo lo que sucede con los efectos es claro desde el principio, cosa que no es común.

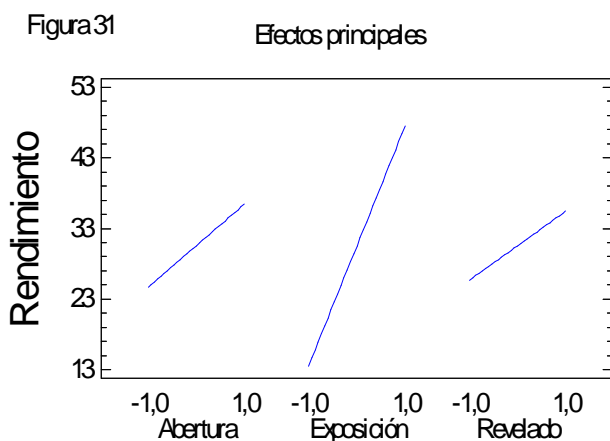
Figura 29



Grafica de Daniels



**5.5. El mejor ANOVA.** Al eliminar los efectos indicados por las gráficas anteriores se obtiene el mejor análisis de varianza dado en la tabla 17. Este análisis es el mejor porque, además de que detecta de manera contundente a los efectos significativos, es el modelo más simple posible para aplicar el comportamiento del rendimiento durante las pruebas. En el ANOVA preliminar (véase tabla 16) aparecía también la interacción  $DE$ , como un efecto significativo ( $\text{valor} - p = 0.0490$ ), y si la incluyésemos en este mejor análisis sigue apareciendo como significativo. Pero en realidad no debe incluirse porque su aportación a la explicación del rendimiento es mínima; los cuatro efectos explican, según el estadístico  $R^2_{ajustado}$ , el 99.22% de la variabilidad observada, y si se incluye en el ANOVA el efecto  $DE$ ,



el estadístico sube a 99.33%, es decir, es un aumento muy pequeño del 0.10%.

Tabla 17. El mejor ANOVA

FV	SC	G:L	C:M	F <sub>0</sub>	valor-p
A: Abertu	1116.28	1	1116.28	382.27	0.0000
B: T-expo	9214.03	1	9214.03	3155.34	0.0000
C: T-revel	750.78	1	750.78	257.10	0.0000
AB	504.03	1	504.03	172.61	0.0000
Error	78.84	27	2.92		
Total	11664.0	31			

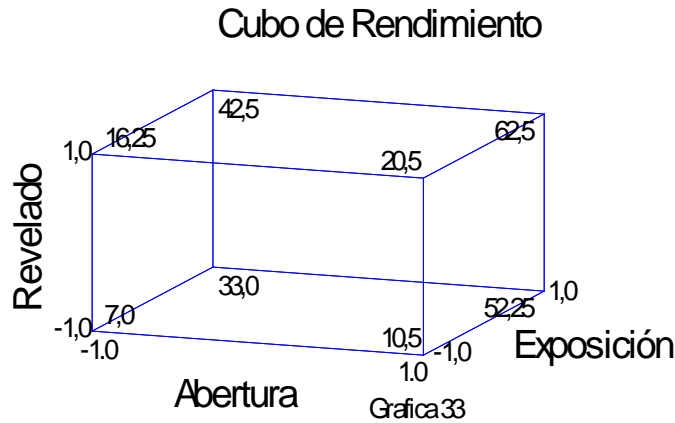
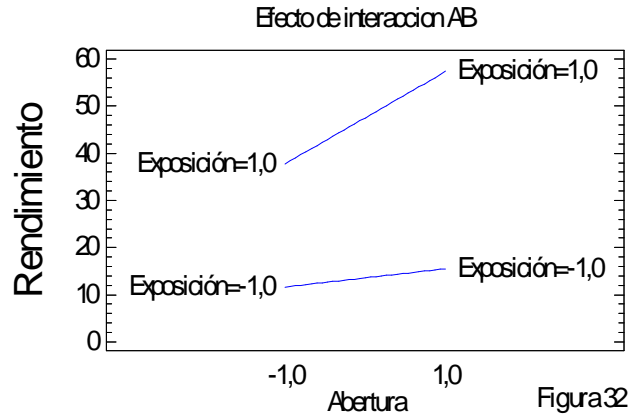
**Interpretación.** De los cuatro efectos que resultaron significativos  $A$ ,  $B$ ,  $C$  y  $AB$ . Los tres efectos principales se muestran en la figura 31, y de ellos sólo se interpreta el efecto de  $C$  (tiempo de revelado), puesto que  $A$  y  $B$  interactúan, lo cual tiene prioridad. Del efecto  $C$  se concluye que a mayor el tiempo de revelado mayor es el rendimiento, por tanto el tiempo de revelado debe fijarse en  $C^+$ .

El efecto de interacción  $AB$  se muestra en la figura 32. Se observa que el efecto del factor  $A$  es mayor cuando el factor  $B$  está en más, además el extremo de línea más alto en la escala del rendimiento corresponde a combinación  $(A^+, B^+)$ .

En conclusión, el mejor tratamiento es  $(A^+, B^+, C^+, D^s, E^s)$  y para  $D$  y  $E$  se eligen los niveles que resultan más convenientes desde el punto de vista económico, de productividad o de operabilidad. En este caso se decide utilizar la dimensión menor de la máscara ( $D^-$ ) y el menor tiempo de grabado ( $E^-$ ).

**Predicción.** El modelo de regresión ajustado relacionado al mejor ANOVA, está dado por

$$(5.1) \quad \hat{Y} = 30.53 + 5.91x_1 + 16.97x_2 + 5.84x_3 + 3.97x_1x_2$$



donde  $x_1$  es el factor  $A$ ,  $x_2$  es el factor  $B$  y  $x_3$  el factor  $C$ ;  $\hat{Y}$  es el rendimiento predicho en el punto  $(x_1, x_2, x_3)$  en unidades codificadas. Recordemos que los coeficientes del modelo al utilizar datos codificados son igual a la mitad del efecto estimado correspondiente. En la gráfica de cubo de la figura 33 se reporta el rendimiento predicho por el modelo en cada combinación de los tres factores incluidos. En particular en el mejor tratamiento se predice un rendimiento promedio de 62.5625%, y con una confianza del 95% el rendimiento futuro en ese proceso andará en el intervalo [60.2946%, 64.8304%]. Se observa que en el punto escogido ( $A^+, B^+, C^+, D^-, E^-$ ) la predicción es 61.325% y con una confianza del 95% el rendimiento futuro en ese proceso andará en el intervalo [59.0446%, 63.5804%]

**Verificación de supuestos.** La verificación de supuestos debe hacerse para el modelo que corresponde al mejor ANOVA. De violarse alguno de los supuestos,

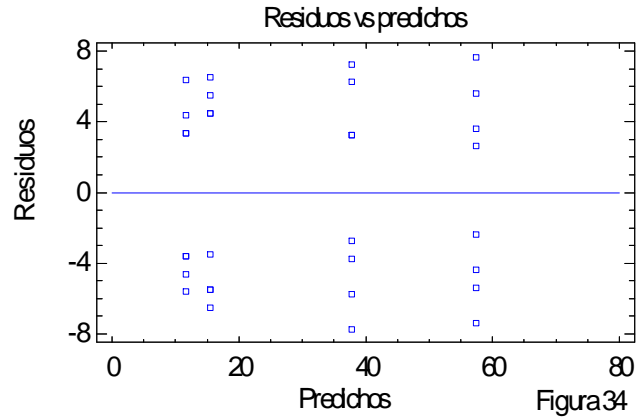


Figura 34

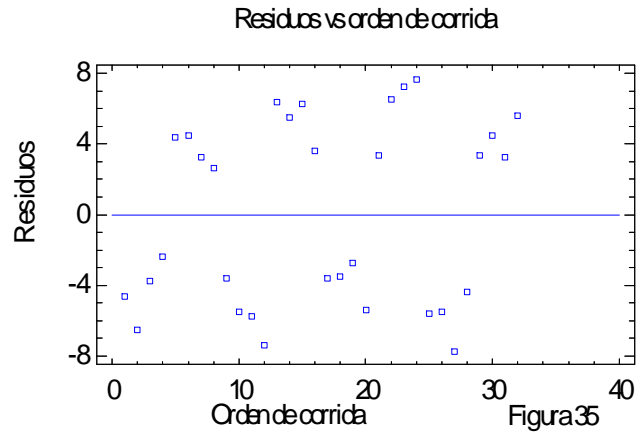
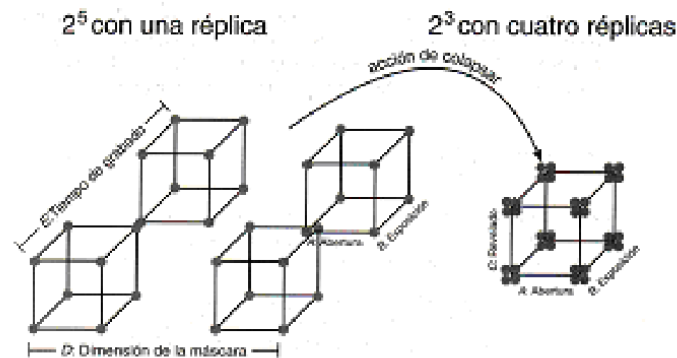
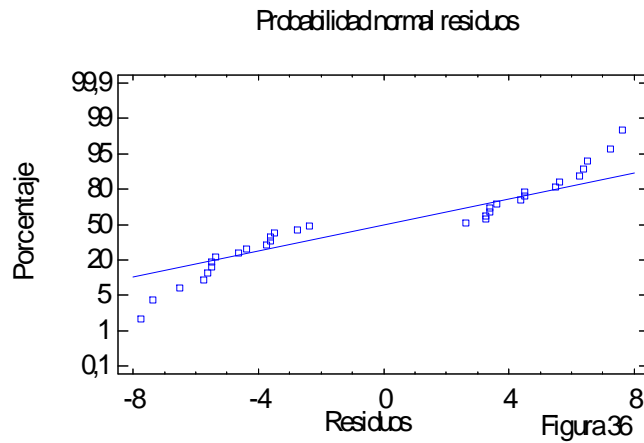


Figura 35

el análisis daría conclusiones incorrectas. En la figura 34 se grafican los *predichos* vs *residuos*, y se muestra una ligera violación al supuesto de varianza constante, que se nota en el patrón “tipo corneta” que siguen los residuos en la medida de que el predicho crece. Sin embargo, la violación no es tan fuerte que pueda tener algún impacto en las conclusiones.

En la figura 35 se grafican el *orden de corrida* vs *residuos*, de ahí se observa que el supuesto de independencia se cumple, ya que los residuos caen aleatoriamente en una banda horizontal. Asimismo, se cumple también el supuesto de normalidad de los residuos, ya que éstos se ajustan bien a una recta en el papel de probabilidad normal (véase figura 36).



**FIGURA 37** Colapsación en dos direcciones, ejemplo de semiconductores.

**Análisis alternativo: colapsación del diseño.** En el análisis del experimento sobre rendimiento de semiconductores una de las conclusiones es que no tuvieron ningún efecto los factores  $D$  (dimensión de la máscara) y  $E$  (tiempo de grabado). Este hecho da pie a colapsar el diseño en esas dos direcciones para convertirlo en un diseño factorial  $2^3$  con cuatro réplicas. Estas réplicas son más que suficientes para obtener un buen estimador del cuadrado medio del error en el ANOVA. La ganancia de la operación colapsar es la simplificación del análisis, al evitar la presencia en los resultados de los dos factores que ya se sabe no tienen ningún efecto; además de que ahora se tienen repeticiones que permiten estimar directamente un buen cuadrado medio del error. En la figura 37 se dibuja el diseño original y el resultado de colapsarlo en dos direcciones. Observe las cuatro réplicas del diseño resultante.

Varios software tienen la opción de colapsar, entre ellos el Statgraphic esto hace que no se requiera volver a capturar los datos. Si se quisieran los datos del diseño colapsado, la matriz de diseño se obtiene al ignorar las columnas de los factores que se quieren colapsar; en la notación de Yates se ignoran las letras que corresponden a dichos factores y surgen por si solas las réplicas de cada tratamiento. Por ejemplo, en algún lugar de la tabla 14 se observaron los resultados:

$$a = 9, ad = 10, ae = 12, ade = 10,$$

y al colapsar los factores D y E se eliminan las letras minúsculas correspondientes y estos resultados se convierten en:

$$a = 9, a = 10, a = 12, a = 10,$$

que son las cuatro réplicas en el tratamiento  $a \rightarrow (1, -1, -1)$  del diseño  $2^3$  resultante. Algo similar pasa con cada tratamiento.

El análisis de varianza directo del diseño colapsado (excluyendo sólo a la interacción triple), se muestra en la tabla.18. Nótese que este análisis está más cerca del mejor ANOVA dado en la tabla 17; de aquí sólo resta eliminar a los efectos AC y BC, para llegar al mejor análisis de varianza. Omitimos la interpretación de los efectos activos, así como la verificación de supuestos del modelo, porque se obtienen los mismos resultados que con el análisis sin colapsar, pero se deja como ejercicio al lector.

Tabla 18. ANOVA directo para los semiconductores					
FV	SC	G:L	C:M	$F_0$	valor-p
A: Abertu	1104.5	1	1104.5	375.68	0.0000
B: T-expo	9248.0	1	9248.0	3145.58	0.0000
C: T-revel	760.5	1	760.5	258.67	0.0000
AB	496.125	1	496.125	168.75	0.0000
AC	1.125	1	1.125	0.38	0.5418
BC	0.125	1	0.125	0.04	0.8383
Error	73.5	25	2,94		
Total	11683.9	31			

**5.6. Ejemplo ilustrativo: Cuando la significancia de los efectos es menos clara.** En la industria tequilera se muestra un experimento cuyo análisis es más “complicado”, ilustraremos las técnicas de separación de efectos en una situación menos ideal.

*El experimento consiste en mejorar el rendimiento en el área de fermentación de agave tequilero. Los factores que se controlan y sus niveles en unidades originales son los siguientes:*

A: Temperatura de carga ( $28^\circ C$ , $30^\circ C$ )
B: Inóculo de levadura (0.31%, 0.72%)
C: Temperatura medio ambiente ( $22^\circ C$ , $34^\circ C$ )
D: Temperatura de fermentación ( $34^\circ C$ , $42^\circ C$ )
E: Brix de carga (13.0°BX, 15.5°BX)

Se utiliza un diseño factorial  $2^5$  con sólo una corrida en cada tratamiento. El diseño y los rendimientos observados se muestran en la tabla 19



Tabla 19 Diseño factorial  $2^5$  con solo una corrida por tratamiento

Yates	A	B	C	D	E	Rend. (%)	Yates	A	B	C	D	E	Rend. (%)
(1)	-	-	-	-	-	60.30	e	-	-	-	-	+	79.44
a	+	-	-	-	-	73.20	ae	+	-	-	-	+	60.40
b	-	+	-	-	-	95.62	be	-	+	-	-	+	46.98
ab	+	+	-	-	-	79.41	abe	+	+	-	-	+	61.95
c	-	-	+	-	-	95.62	ce	-	-	+	-	+	57.05
ac	+	-	+	-	-	97.24	ace	+	-	+	-	+	42.09
bc	-	+	+	-	-	76.17	bce	-	+	+	-	+	47.65
abc	+	+	+	-	-	57.05	abce	+	+	+	-	+	53.69
d	-	-	-	+	-	90.76	de	-	-	-	+	+	61.58
ad	+	-	-	+	-	46.96	ade	+	-	-	+	+	27.00
bd	-	+	-	+	-	45.30	bde	-	+	-	+	+	61.15
abd	+	+	-	+	-	42.09	abde	+	+	-	+	+	57.05
cd	-	-	+	+	-	89.13	cde	-	-	+	+	+	56.12
acd	+	-	+	+	-	71.31	acde	+	-	+	+	+	55.36
bcd	-	+	+	+	-	76.17	bcde	-	+	+	+	+	29.30
abcd	+	+	+	+	-	61.58	abcde	+	+	+	+	+	28.59

5.6.1. *Análisis del experimento.* En este diseño el contraste de cada efecto tiene 32 términos, así que los cálculos a mano son engorrosos y es mejor utilizar un paquete estadístico. Por poner un ejemplo, el contraste del efecto  $A$  se obtiene multiplicando su columna de signos por los tratamientos en la notación de Yates. Haciendo los cálculos resulta que

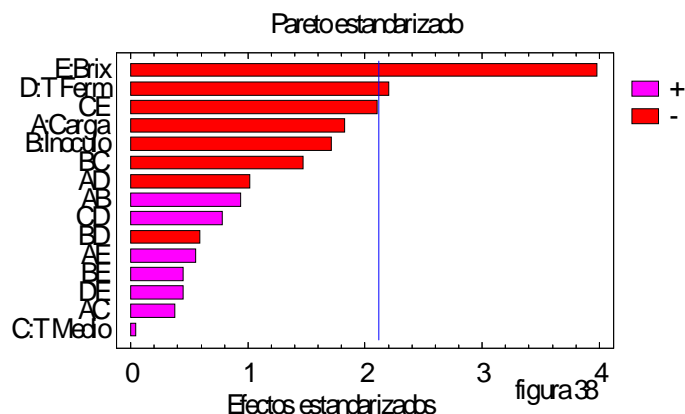
$$\begin{aligned}
 \text{Contraste } A &= a + ab + ac + abc + ad + abd + acd + abcd + ae + abe \\
 &\quad + ace + abce + ade + abde + acde + abede - (1) - b \\
 &\quad - c - bc - d - bd - ed - bcd - e - be - cebce \\
 &\quad - de - bde - cde - bcde \\
 &= 73.2 + 79.41 + 97.24 + 57.05 + 46.96 + 42.09 + 71.31 + 61.58 \\
 &\quad + 60.4 + 61.95 + 42.09 + 53.69 \\
 &\quad + 27.0 + 57.05 + 55.36 + 28.59 - 60.3 - 95.62 - 95.62 \\
 &\quad - 76.17 - 90.76 - 45.3 - 89.13 - 76.17 - 79.44 - 46.98 \\
 &\quad - 57.05 - 47.65 - 61.58 - 61.15 - 56.12 - 29.3 \\
 &= -153.37
 \end{aligned}$$

De acuerdo con la fórmula (4.1), se multiplica por  $1/16$  para obtener el efecto correspondiente, es decir:

$$Efecto A = \frac{1}{16} [\text{Contraste } A] = \frac{-153.37}{16} = -9.58563$$

Haciendo los cálculos para todos los efectos principales e interacciones de dos factores se obtienen los valores dados en la tabla 20. Se observa que los factores que impactan más al rendimiento son: *brix de carga* ( $E$ ) y *temperatura de fermentación* ( $D$ ); la *temperatura ambiental* afecta pero dependiendo del nivel que asume el *brix de carga*. Falta ver si estos efectos son importantes en el aspecto estadístico.

En la primera etapa del diagrama de Pareto y gráfica de efectos en papel de probabilidad normal se recomienda incluir a las interacciones triples porque ayudan a una mejor visualización de los gráficos y de paso se evita el no detectar alguna interacción triple grande que pase desapercibida, inflando fuertemente al error. *Puede ocurrir que una interacción triple, además de deberse a causas aleatorias,*



se deba a un pequeño efecto de la propia interacción, que hace que ésta no se deba enviar al error. En otras palabras, si una interacción triple tiene una significancia pequeña, aunque no menor que la significancia dada, debe dejarse en el ANOVA, aunque no se interpreta; si resulta significativa, debe interpretarse.

Tabla 20. Efectos estimados

media	=	61.9784	AD	=	-5.36062
A :Tiempo de carga	=	-9.58563	AE	=	2.94313
B:Inóculo	=	-8.98812	BC	=	-7.72687
C:Temp. medio ambiente	=	0.308125	BD	=	-3.13563
D:Temp.de fermentación	=	-11.5256	BE	=	2.40312
E:Brix de carga	=	-20.7819	CD	=	4.15062
AB	=	4.96938	CE	=	-11.0206
AC	=	2.04812	DE	=	2.38812

5.6.2. *Gráfica de efectos en papel normal.* En la figura 39 se grafican los 31 efectos en papel de probabilidad medio normal (half normal plot). Se elige el papel medio normal porque hace mejor separación visual de los efectos importantes, que el papel normal. En la figura se observan seis efectos alejados de la línea, que son los causantes de la mayoría de la variabilidad observada en el rendimiento y son los mismos efectos que el Pareto detecta en las primeras seis barras. Así, en apariencia los efectos *A*, *B*, *D*, *E*, *BC* y *CE* no se deben sólo al azar. La significancia de estos efectos se comprobará con el análisis de varianza.

5.6.3. *Análisis de varianza.* Se decide construir un ANOVA (tabla 21) en el la suma de cuadrados del error (*SCE*) está compuesta por 25 efectos pequeños y "no significativos" detectados con la ayuda del diagrama de Pareto y gráfico de Daniel (figuras 38 y 39), que son precisamente los grados de libertad que tiene esta suma. Éste es el mejor ANOVA de acuerdo a los criterios arriba señalados. Del ANOVA se concluye que son significativos los efectos *A*, *D*, *E* y *CE* a un nivel  $\alpha = 0.05$ . Los efectos *B* y *BC* no alcanzan a ser significativos a este nivel, pero no son del todo aleatorios, por eso no deben excluirse de la tabla. De hecho el efecto *B* tiene un valor  $-p = 0.064$  y se recomienda también considerarlo en la interpretación.

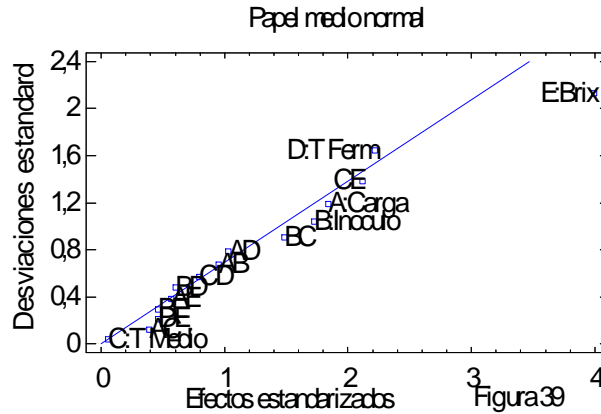


Tabla 21. ANOVA					
FV	SC	G:L	C:M	F <sub>0</sub>	valor-p
A:T-carga	735.074	1	735.074	5.27	0.0494
B:Inóculo	646.291	1	646.291	3.75	0.0641
D:T-fermen	1062.72	1	1062.72	6.17	0.0201
E:Brix	3455.09	1	3455.09	20.06	0.0001
BC	477.637	1	477.637	2.77	0.1083
CE	971.633	1	971.633	5.64	0.0255
Error	4305.84	25	172.234		
Total	11654.3	31			

5.6.4. *Verificación de supuestos.* De acuerdo a las figuras 40 y 41 se concluye que no hay problemas con los supuestos de varianza constante y de normalidad. En la primera, los puntos caen aleatorios sobre una banda horizontal, y en la segunda los residuos se ajustan bastante bien a una línea recta.

5.6.5. *Gráficas de efectos y conclusiones.* En la figura 42 se dibujan los efectos principales  $A$ ,  $D$ , y  $E$ , que resultaron significativos a un nivel  $\alpha = 0.05$  y también se dibuja el efecto  $B$ , que aun cuando no es muy significativo con  $\alpha = 0.05$ , si lo es para  $\alpha = 0.065$  (véase tabla 21). En otras palabras, asumiendo 1.5% más de riesgo tipo I, la conclusión sería que el factor  $B$  sí tiene efecto importante. Asumiendo este riesgo, es que decidimos también interpretar al efecto  $B$ .

Interesa encontrar la combinación de niveles en los cinco factores que da por resultado el mayor rendimiento. En la figura 42 se observa que conforme se aumenta el nivel de los factores  $A$ ,  $B$ ,  $D$  y  $E$  el rendimiento disminuye, por tanto los mejores niveles para estos factores son:  $A^-$ ,  $B^-$  y  $D^-$ , respectivamente. Por la prioridad de las interacciones sobre los efectos principales, los mejores niveles de los dos factores restantes se obtienen de la gráfica de interacción.

Residuos contra predichos

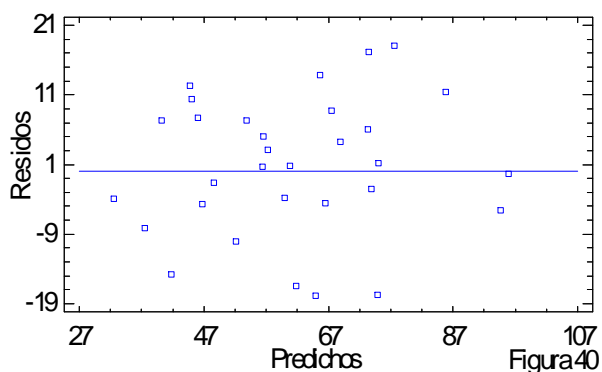


Figura 40

Residuos en papel normal

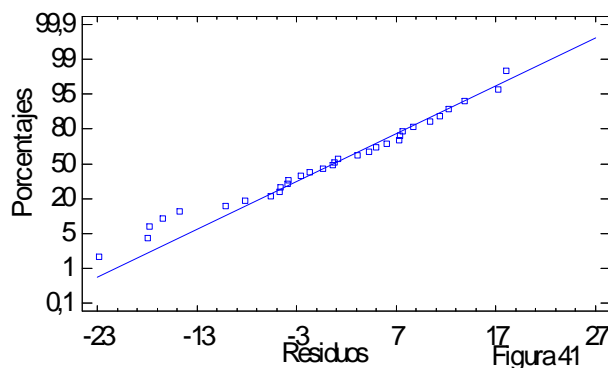
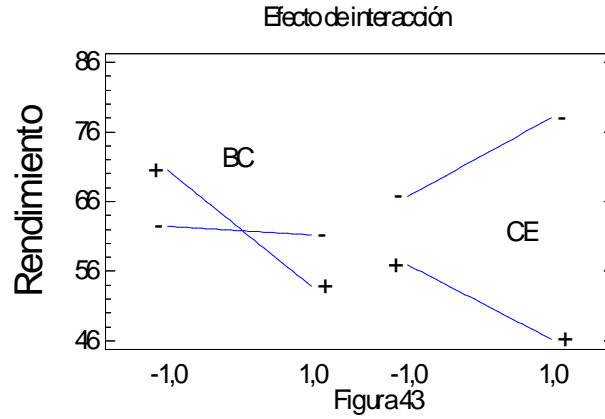
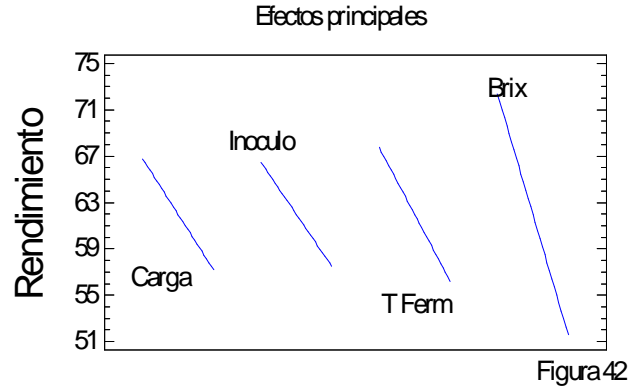


Figura 41

El efecto de interacción  $CE$  se representa en la figura 43, donde se observa que el efecto de incrementar  $C$  es positivo (aumenta el rendimiento) siempre y cuando  $E$  esté en su nivel  $-$ ; porque si  $E$  está en su nivel  $+$  ocurre un efecto negativo. De esta forma se busca el punto (extremo de línea) más alto respecto a la escala del rendimiento, y se obtiene que los mejores niveles para los factores  $C$  y  $E$  son:  $C^+$  y  $E^-$ . En resumen, el mejor tratamiento es cuando todos los factores, excepto  $C$ , se encuentran en su nivel bajo. Es decir, la temperatura de carga, el inóculo de levadura, los grados Brix y la temperatura de fermentación deben estar en su nivel bajo, mientras que la temperatura ambiental, cuyo efecto depende de los grados Brix, debe ubicarse en su nivel alto para obtener un mejor rendimiento en la fermentación. El efecto de interacción  $BC$  que también se representa en la figura, no se interpreta por no tener la suficiente importancia. Sin embargo, aunque se interpretara, la conclusión anterior no cambia.



5.6.6. *Predicción.* El modelo lineal asociado al análisis de varianza de la tabla 21 explica, de acuerdo al coeficiente de determinación  $R^2_{ajustado}$ , un 54.18% de la variabilidad en el rendimiento del proceso de fermentación del agave. Este valor tan bajo de  $R^2_{ajustado}$  no permite tener buena calidad en la predicción ya que el exceso de variabilidad no explicada hace que el intervalo de confianza para el rendimiento futuro sea demasiado amplio. Pero aun, con  $R^2$  bajos se obtienen conclusiones sólidas como en éste ejemplo; y lo que seguiría es poner el proceso a correr en la mejor condición encontrada con el análisis, y ver cómo esto mejora los rendimientos anteriores. Lo que es claro es que las predicciones con el modelo serán poco precisas, pero eso no quita que eso mejore el proceso respecto a las condiciones en las que operaba antes del experimento. Al respecto y en el ejemplo, puede verificarse que el modelo ajustado predice un rendimiento de 96.79% en el mejor tratamiento, es decir,

$$\hat{Y}(-1, -1, +1, -1, -1) = 96.79.$$

El intervalo de confianza al 95% para el rendimiento futuro sobre este punto está dado por  $[66.95, 126.63]$ . Por supuesto que de 100% nunca se va a pasar, y es factible observar rendimientos bajos como 67%. Sin embargo, se debe poner a trabajar el proceso en esa condición y ver cómo esto mejoró la situación prevaleciente del experimento. Si las mejoras aún no son suficientes, habrá que generar más conocimiento sobre el proceso y generar y confirmar nuevas conjeturas sobre las causas de la variabilidad del proceso.

## 6. Factoriales $2^k$ con punto al centro

Si los  $k$  factores del diseño factorial  $2^k$  admiten un nivel de prueba intermedio, se recomienda formar un tratamiento adicional formado por la combinación del nivel intermedio o medio de todos los factores. A tal tratamiento o combinación se le conoce como *punto al centro*.

Hay dos razones por las que es deseable correr el punto central con cierto número de réplicas:

- (1) Obtener grados de libertad adicionales para el error en la tabla de ANOVA, sin perjudicar el balance en la estimación de los efectos de interés. Es conveniente interpretar ANOVAS con al menos 8 grados de libertad en el error, condición que a veces es difícil de cumplir cuando por razones económicas el experimento se corre sin las réplicas suficientes. Por ejemplo, al correr el factorial  $2^3$  sin réplicas se tienen 7 grados de libertad totales, siendo entonces imposible construir un análisis de varianza con grados de libertad suficientes para el error. *Cuatro o cinco repeticiones* al centro agregarían esa cantidad de grados de libertad para el error, además de proveer un *estimador puro* (independiente de los efectos estimados) de la varianza en dicho punto. Es más factible obtener cuatro corridas en el centro que repetir los tratamientos del experimento completo.
- (2) Si los factores son de tipo cuantitativos, las repeticiones al centro permiten detectar a posible presencia de curvatura en al menos uno de los factores objetos de estudio. La curvatura son los efectos cuadráticos  $A^2, B^2, \dots$ . Una vez detectados este tipo de efectos, el experimento se aumenta con más puntos experimentales para poder estudiar dicha curvatura. No es conveniente utilizar de entrada un experimento que permita estudiar la curvatura, ya que de no existir ésta se estarían dilapidando recursos.

**6.1. Ejemplo de un diseño factorial  $2^3$  con punto al centro.** En un proceso de circuitos integrados (obleas) interesa minimizar la corriente de fuga, que se supone depende de la temperatura de quemado ( $A$ ), tiempo de quemado ( $B$ ) y porcentaje de nitrógeno ( $C$ ). Para ello se decide correr un experimento factorial  $2^3$

con cuatro repeticiones al centro. Los resultados obtenidos se muestran en seguida:

Temp.	Tiempo	% de N	Y= Corr. de fuga
-1	-1	-1	2.153, 1.843
+1	-1	-1	1.609, 2.018
-1	+1	-1	1.346, 1.766
+1	+1	-1	1.695, 2.051
-1	-1	+1	3.864, 5.041
+1	-1	+1	7.054, 5.574
-1	+1	+1	5.519, 4.181
+1	+1	+1	5.746, 6.088
0	0	0	2.490, 2.384
0	0	0	2.474, 1.778

Con el ANOVA para este experimento se obtiene la tabla 22a, en la cual se aprecia que hay claramente un efecto dominante ( $C$ ), y dos efectos que están en duda ( $A$  y  $AC$ ), y el resto son no significativos. Además los coeficientes  $R^2$  tienen un valor aceptable (86% y 78%). Para llegar al mejor ANOVA se eliminan los efectos no significativos ( $B$ ,  $AB$ ,  $BC$  y  $ABC$ ) y se obtiene que los efectos que estaban en duda ( $A$  y  $AC$ ) son importantes al 10% de significancia.

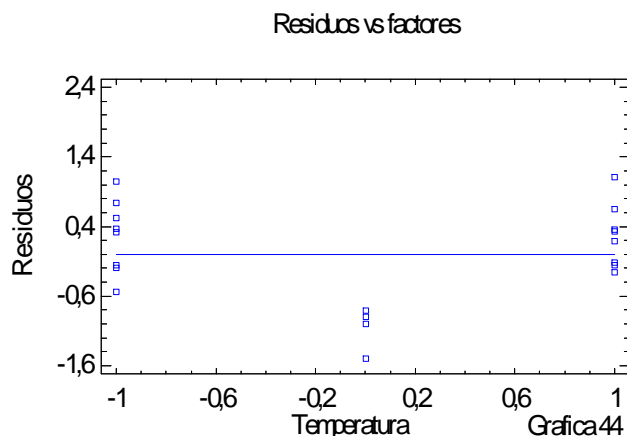
Tabla 22a. ANOVA para corriente de fuga de obleas

FV	SC	G.L	C.M	F <sub>0</sub>	valor-p
A:Temp.	2.342	1	2.342	3.16	0.1010
B:Tiem	0.0365	1	0.0365	0.05	0.8283
C:% de N	51.072	1	51.072	68.82	0.0000
AB	0.0214	1	0.0214	0.03	0.8678
AC	1.954	1	1.954	2.63	0.1306
BC	0.0366	1	0.0366	0.05	0.8278
ABC	0.4199	1	0.4199	0.57	0.4664
Error	8.9049	12			
Total	64.789	19	0.7421		

Pero dado que se corrió el punto al centro hay oportunidad de verificar curvatura. En la tabla 22b se muestra el mejor ANOVA y además el error se parte en los componentes de falta de ajuste (*lack-of fit*) y *error puro*, donde el error puro se obtiene de las repeticiones en el centro y en los puntos factoriales. La falta de ajuste resulta significativa ( $valor - p = 0.0003$ ) lo cual es un fuerte indicio de curvatura. Este indicio se comprueba observando la gráfica de residuos contra los niveles de los factores (figura 44 y 45), donde se aprecia que los residuos correspondientes a las observaciones en el centro del experimento *caen bastante abajo* respecto a los residuos en los extremos, lo cual es un indicativo de la presencia de curvatura en cada factor. *El modelo predice valores de la corriente de fuga mayores a los observados en el centro, por eso los residuos allí son negativos.*

Se observa que el factor  $B$  no afecta a la respuesta  $Y$ , por lo que el experimento se puede colapsar en un  $2^2 + \text{centro}$  con cuatro réplicas y el diseño colapsado presenta ANOVA.

La presencia de curvatura hace necesario aumentar el experimento con puntos adicionales para poder estudiar o estimar los efectos  $A^2$  y  $C^2$  y saber cuál de ellos o si ambos provocan la falta de ajuste. El modelo de regresión asociado al ANOVA



de la tabla 22b es

$$(6.1) \quad \hat{Y} = 3.33 + 0.38x_1 + 1.78x_3 + 0.35x_1x_3,$$

que no se ajusta bien a la respuesta observada en el punto al centro ya que le faltan los términos  $x_1^2$  y/o  $x_3^2$ . En resumen, después del análisis de varianza anterior (tabla 22b) que detecta curvatura, lo que sigue es correr puntos experimentales adicionales, y después de hacer el análisis del experimento aumentado considerando hasta los términos cuadráticos.

Tabla 22b. Mejor ANOVA y prueba de falta de ajuste para corr. de fuga

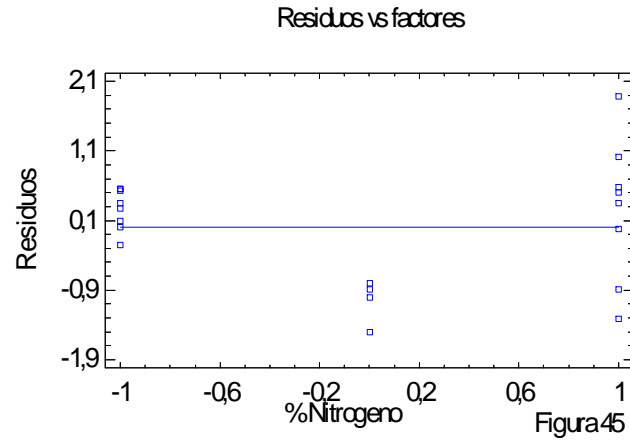
FV	SC	G.L	C.M	F <sub>0</sub>	valor-p
A:Temp.	2.342	1	2.342	9.05	0.0088
C:%deN	51.072	1	51.072	197.25	0.0000
AC	1.954	1	1.954	7.55	0.0150
Falta de ajuste	5.535	1	5.535	21.38	0.0003
Error puro	3.884	15	0.259		
Total	64.789	19			

## 7. Factoriales $2^k$ en bloques

Cuando no es posible correr todos los tratamientos de un diseño factorial  $2^k$  bajo las mismas condiciones experimentales, se puede pensar en factores de bloque. para incrementar la precisión del experimento, además que permite tener conclusiones válidas dentro de distintos bloques (días, máquinas, material, tipo de producto, etc.), que son inevitables y están siempre presentes en el proceso. Se puede pensar en hacer bloqueo en algunos de estos casos:

- (1) Si el diseño factorial completo  $2^k$  no se puede correr el mismo día, ya sea por lentitud de las corridas o del proceso de medición, o por la cantidad de corridas, y se considera que el factor día puede afectar los resultados del estudio, entonces se debe incorporar en el experimento como factor de





bloque. Si el proceso estudiado es sensible a los cambios de turno dentro del mismo día, los turnos debieran considerarse como el factor de bloque.

- (2) Cuando un lote de material no alcanza para hacer todas las corridas experimentales y se sospecha que las diferencias entre los lotes podrían sesgar los resultados, entonces es necesario repartir adecuadamente las corridas experimentales en varios lotes (bloques).
- (3) Cuando no es posible contar durante el experimento completo  $2^k$  con el mismo operador o con el mismo instrumento de medición, y si se sospecha que éstos pueden influir en el desempeño del proceso, entonces hay que considerarlos como factores de bloque.

Deben bloquearse aquellas fuentes de variación que puedan tener algún impacto importante sobre la respuesta de interés. Algunos posibles factores de bloque bastará con mantenerlos fijos durante las corridas experimentales. La repartición del arreglo en bloques se hace bajo el principio de jerarquía de efectos: son más importantes los efectos principales, seguidos por las interacciones dobles y luego las triples, cuádruples, etc.

**7.1. Ejemplo ilustrativo donde los bloques son las réplicas.** Considere otra vez el factorial  $2^2$  con cuatro réplicas discutido al inicio de este capítulo (ejemplo de ranuradora). Suponga que cada una de las cuatro réplicas se corrió en un día diferente. El experimento repartido en cuatro bloques (días) se muestra en la siguiente tabla:

broc	veloc	Bloque I	Bloque II	Bloque III	Bloque IV	Totales
-1.	-1	18.2	18.9	12.9	14.4	64.4=(1)
+1	-1	27.2	24.0	22.4	22.5	96.1=a
-1	+1	15.9	14.5	15.1	14.2	59.7=b
+1	+1	41.0	43.9	36.3	39.9	161.1=ab
Totales		$Y_{..1}=102.3$	$Y_{..2}=101.3$	$Y_{..3}=86.7$	$Y_{..4}=91.0$	$Y_{...}=381.3$

De los totales por bloque, dados en la parte inferior de la tabla, se tiene que la suma de cuadrados para los bloques es

$$\begin{aligned} \text{SCBLOQ} &= \sum_{k=1}^4 \frac{Y_{..k}^2}{2 \times 2} - \frac{Y_{...}^2}{N} \\ &= \frac{1}{4} (102.3^2 + 101.3^2 + 86.7^2 + 91.0^2) - \frac{381.3^2}{16} = 44.36 \end{aligned}$$

y tiene 3 grados de libertad porque son cuatro bloques o días. A partir de ésta se incluye un renglón en el ANOVA donde se prueba la hipótesis  $H_0 : \text{Efecto de bloque} = 0$  (tabla 23). Como el  $\text{valor} - p = 0.0280$ , se concluye que el efecto de bloque (debido a los días) es significativo. Note en la tabla de datos que en los días 3 y 4 se observó menos vibración que en los primeros dos días. Sin embargo, sabemos que el orden de las corridas fue completamente aleatorio y en un mismo día, así que este efecto de bloque no es tal y ocurre sólo por azar. Si el efecto fuera real, se interpreta con gráficas de medias y pruebas de comparaciones múltiples, para comprender mejor la manera en que los días afectan la respuesta.

Las conclusiones sobre el mejor tratamiento serían las mismas obtenidas anteriormente; el efecto de bloque significativo implicaría en este caso que la respuesta esperada sobre el mejor tratamiento cambia de manera significativa día con día.

Tabla 23. ANOVA con efecto de bloque					
FV	SC	G.L	C.M	F <sub>0</sub>	valor-p
A:Broca	1107.22	1	1107.22	364.21	0.0000
B:Velocidad	227.25	1	227.25	74.75	0.0000
AB	303.63	1	303.63	99.88	0.0000
Bloque (día)	44.36	3	14.79	4.86	0.0280
Error	27.36	9	3.04		
Total	1709.83	15			

**7.2. Generando los bloques con contrastes.** Si los bloques se generan con los contrastes de alguna(s) interacción(es), entonces los efectos correspondientes y sus productos aparecerán *confundidos* con los bloques. La *confusión* consiste en que al estimar el efecto de bloque también se está estimando la interacción o interacciones utilizadas en su construcción. Mientras los efectos utilizados en la generación de los bloques sean de interacciones triples en adelante, el efecto confundido se puede atribuir sin mayor trámite a los bloques.

Supongamos que se quiere correr un experimento  $2^3$  repartido en dos bloques. Éstos se definen de acuerdo a los signos en el contraste de la interacción triple  $ABC$ . En las tablas siguientes, los signos negativos de la columna  $ABC$  señalan el bloque 1 y los signos positivos señalan el bloque 2. Las cuatro combinaciones que conforman cada bloque se muestran en el lado derecho.

Al estimar los efectos, el de interacción triple estará confundido con el efecto de bloque, es decir, efecto  $ABC = \text{efecto del bloque}$ . El efecto observado se atribuye al bloque, porque es más probable que el bloque sea lo que influye. Además, se sabe que generalmente las interacciones de tres factores en adelante no son significativas.

Si este mismo diseño ( $2^3$ ) se quiere partir en cuatro bloques, se seleccionan los contrastes de dos efectos para generar los bloques, de manera que su producto (módulo 2) sea una interacción del más alto orden posible, los tres efectos de interacción estarán confundidos con bloques. La multiplicación módulo 2 de dos efectos

cualquiera significa simplemente que las letras repetidas se eliminan y prevalecen en el resultado las no repetidas. La selección más adecuada de los dos generadores es  $AB$  y  $AC$ .

A	B	C	AB	AC	ABC
-	-	-	+	+	-
+	-	-	-	-	+
-	+	-	-	+	+
+	+	-	+	-	-
-	-	+	+	-	+
+	-	+	-	+	-
-	+	+	-	-	-
+	+	+	+	+	+

$\Rightarrow$

bloque 1			bloque 2		
A	B	C	A	B	C
-	-	-	+	-	-
+	+	-	-	+	-
+	-	+	-	-	+
-	+	+	+	+	+

y

Bloque 1			Bloque 2			Bloque 3			Bloque 4		
A	B	C	A	B	C	A	B	C	A	B	C
-	-	-	+	-	-	-	+	-	+	+	-
+	+	+	-	+	+	+	-	+	-	-	+

Ahora, al estimar el efecto de bloque se estarán estimando también los efectos  $AB$ ,  $AC$  y  $BC$ , donde  $BC$  se obtiene del producto módulo 2 de los dos efectos iniciales. Es decir, estos efectos están confundidos con bloques y la suma de sus sumas de cuadrados es igual a la suma de cuadrados del efecto de bloque.

En general el factorial  $2^k$  se puede partir en  $2^b$  bloques, para lo cual se requieren  $b$  efectos generadores iniciales, cuyos signos definen las combinaciones que componen a cada bloque. Los generadores iniciales se seleccionan de las interacciones de mayor orden, de manera que todos sus posibles productos sean también interacciones del más alto orden. Los efectos iniciales así como sus posibles productos estarán confundidos con bloques, es decir, no se podrán estudiar. En la tabla 24 se proveen generadores iniciales adecuados para construir  $2^b$  bloques de un factorial completo  $2^k$ , para algunos valores de  $k$ . También se puede usar un software estadístico como Statgraphics para generar los bloques deseados para cualquier factorial completo o fraccionado.

Tabla 24. Efectos adecuados para generar bloques				
Factores $k$	Bloques $2^b$	Tamaño de bloques $2^{k-b}$	Efectos generadores	Efectos confundidos con bloques
3	2	4	ABC	ABC
4	2	8	ABCD	ABCD
4	4	4	ABC, ACD	ABC, ACD, BD
5	2	16	ABCDE	ABCDE
5	4	8	ABC, CDE	ABC, CDE, ABDE
5	8	4	ABE, BCE, CDE	ABE, BCE, CDE, AC
6	2	32	ABCDEF	ABCD, BD, ADE
6	4	16	ABCF, CDEF	ABCDEF
6	8	8	ABEF, ABCD, ACE	ABCF, CDEF, ABDE, ABEF, ABCD, ACE, CDEF, BCF, BDE, ADF

### 8. Ejercicios: Diseños factoriales $2^k$

EJERCICIO 81. Conteste las siguientes preguntas de índole general, respecto a lo, diseños factoriales. (a) Explique qué son los factores y qué es la(s) variable(s)

de respuesta. (b) ¿Cuál es el objetivo de un diseño factorial? (c) Ejemplifique y explique en qué consiste la estrategia de modificar o mover un factor a la vez, que es propia de la experimentación empírica. (d) Señale y argumente qué ventajas tienen los experimentos factoriales sobre la estrategia de mover un factor a la vez. (e) ¿Qué significa que un factor tenga un efecto significativo? ¿Sobre quién es el efecto? (f) ¿Todos los factores deben ser de tipo cuantitativo o es posible involucrar factores cualitativos, como por ejemplo dos tipos de máquinas, la presencia o ausencia de alguna sustancia?

EJERCICIO 82. Suponga un diseño factorial  $2^2$  cuyos factores y niveles son: temperatura (8, 20) y velocidad (4, 7). La variable de respuesta es rendimiento. Conteste las siguientes preguntas: (a) ¿Por qué este diseño recibe tal nombre? (b) Utilizando diferentes tipos de códigos, anote los diferentes tratamientos que forman este diseño. (c) Represente en forma geométrica al diseño y resalte la región de experimentación. (d) Explique cómo piensa que fue el proceso para seleccionar esos factores y esos niveles. (e) Defina qué son los efectos principales y cuál el efecto de interacción. (f) Señale los diferentes efectos que se pueden estudiar con este diseño y la forma que se calcula cada uno. (g) Describa en qué consiste la aplicación de los tres principios básicos del diseño de experimentos (capítulo 1), en este caso.

EJERCICIO 83. A continuación se muestran los resultados de un diseño factorial. Conteste los siguientes incisos sin utilizar un software computacional, es decir, haga las operaciones que se le piden de manera manual.

Réplica					
A	B	I	II	III	Total
–	–	82	80	84	(1) = 246
+	–	78	82	79	(a) = 239
–	+	71	70	66	(b) = 207
+	+	89	88	93	(ab) = 270

(a) ¿Qué nombre recibe este diseño y por qué? (b) ¿Cuántos tratamientos tiene este diseño, cuántas réplicas? (c) En total son 12 corridas experimentales las que se realizaron, señale en qué orden debieron correrse y explique por qué. (d) Señale los efectos que se pueden estudiar a través de este diseño. (e) Obtenga los contrastes para los efectos principales de A y B, y para la interacción. (f) Calcule los efectos principales y el efecto de interacción. (g) Haga las gráficas de los efectos principales de A y B, e interprételas. (h) Realice la gráfica de la interacción entre los factores A y B, e interprétela con detalle. (i) ¿Desde su punto de vista el factor B parece tener influencia sobre Y? Argumente.

EJERCICIO 84. Suponga un diseño factorial  $2^3$  y conteste las siguientes preguntas. (a) Utilizando la notación de (–, +) para los niveles de los factores, escriba todos los tratamientos que forman este diseño. (b) Represente en forma geométrica este diseño y resalte la región de experimentación. (c) ¿Cuáles son todos los posibles efectos que se pueden estudiar con este diseño? (d) Para cada uno de los efectos anteriores, obtenga su contraste. (e) Señale en forma específica cómo utilizaría los contrastes para calcular los efectos y la suma de cuadrados. (f) ¿Cómo aplicaría los tres principios básicos del diseño de experimentos (capítulo 1), en este caso para cada uno de los efectos anteriores?

EJERCICIO 85. A continuación se muestran los resultados obtenidos en un diseño factorial  $2^3$  no replicado. Conteste los siguientes incisos sin utilizar un software

computacional, es decir, haga las operaciones que se le piden de manera manual.

¿Código?	A	B	C	Y
	−	+	−	25
	+	+	+	12
	−	−	−	30
	+	−	+	10
	−	−	+	10
	+	+	−	14
	−	+	+	31
	+	−	−	17

(a) En la primera columna de la matriz de diseño especifique el código de cada uno de los tratamientos, de acuerdo a la notación de Yates. (b) Calcule los efectos principales de A y B. (c) Haga la gráfica de los efectos principales de A y B, e interprételas. (d) Calcule el efecto de la interacción de AB. (e) Realice la gráfica de la interacción entre los factores A y B; e interprétela con detalle. (f) ¿Qué tendría que hacer para saber si los efectos que calculó en los incisos anteriores, afectan de manera significativa a la variable de respuesta? (g) Calcule la suma de cuadrados para el efecto principal de A y para la interacción.

EJERCICIO 86. Suponga un diseño factorial  $2^4$ , y conteste las siguientes preguntas. (a) Anote la matriz de diseño, es decir, haga una lista de todos los tratamientos que forman este diseño. (b) ¿Por qué este diseño recibe tal nombre? (c) ¿Cuáles son todos los posibles efectos que se pueden estudiar con este diseño? (d) ¿Refiriéndose al análisis, en qué consiste y cuál es el objetivo de obtener el mejor ANOVA? (e) ¿Cómo se calculan los coeficientes de determinación  $R^2$  y  $R^2_{ajustado}$ ? (f) Si después de obtener el mejor ANOVA, se obtiene que estos  $R^2_{ajustado}$  coeficientes tienen un valor de alrededor de 90, ¿qué significa esto? (g) Si por el contrario tales coeficientes tienen un valor de alrededor de 20, ¿qué significa esto? (h) Obtenga el contraste para el efecto principal de D y para el efecto de interacción CD. (i) Señale en forma específica cómo utilizaría los contrastes para calcular los efectos y la suma de cuadrados. (j) ¿Puede darse el caso de que el efecto principal de A no sea significativo, y el efecto de la interacción AB sí lo sea?

EJERCICIO 87. En una fábrica de dientes se tiene problemas de calidad: porosidad (burbujas de aire dentro de los dientes), manchas blancas, dientes sucios, dientes quebrados. En los intentos por resolver los problemas han hecho cambios en algunos factores o variables del proceso. Siguiendo la metodología del DDE se decide correr un diseño de experimentos  $2^3$ . Los factores y niveles son temperatura de prensado (90, 130°C), tiempo de prensado (8 y 15 minutos) y tamaño de partícula (sin tamizar y con tamizado), la variable de respuesta fue porcentaje de diente bueno en cada corrida (un lote de producción). Los datos son los siguientes.

Temperatura	Tiempo	T. de partícula	Porcentaje de diente bueno
90	8	Sin tamizar	76.4, 76.2
130	8	Sin tamizar	76.8, 76.9
90	15	Sin tamizar	80.4, 81.0
130	15	Sin tamizar	77.9, 79.6
90	8	Con tamizado	84.4, 84.6
130	8	Con tamizado	84.7, 84.5
90	15	Con tamizado	82.7, 83.2
130	15	Con tamizado	85.0, 84.7

(a) Estime todos los posibles efectos, y diga cuáles son significativos. (b) Realice un análisis de varianza de estos datos y obtenga conclusiones generales. (c) Verifique residuos, ¿qué observa que sea notorio? (d) ¿Hay un tratamiento ganador? (e) ¿Qué condición de proceso sugiere emplear a futuro? Tome en cuenta, además, que a mayor tiempo y mayor temperatura, más costos. (f) Las condiciones que se utilizaban antes del experimento era temperatura de  $130^{\circ}\text{C}$  y tiempo de 15 minutos. ¿Por qué cree que se eligieron niveles inferiores de prueba para estos factores? (g) Dado los resultados, aunque parecen haber sido positivos, éstos no son suficientes, por tanto ¿qué sugiere usted, ¿explora más niveles de los factores ya estudiados?, ¿considerar otraQ causas? o ¿qué? Argumente.

**EJERCICIO 88.** En una empresa lechera se ha tenido problemas con la viscosidad de cierta bebida de chocolate. Se cree que tres ingredientes que se agregan en pequeñas cantidades son con los que se puede resolver este problema, por lo que es necesario explorar la situación para ello se corre un experimento  $2^3$  con dos réplicas. En seguida se aprecian los resultados obtenidos:

Ingrediente A	Ingrediente B	Ingrediente C	Viscosidad
-1	-1	-1	13.3, 13.7
+1	-1	-1	14.7, 14.4
-1	+1	-1	14.6, 14.5
+1	+1	-1	14.3, 14.1
-1	-1	+1	16.9, 17.2
+1	-1	+1	15.5, 15.4
-1	+1	+1	17.0, 17.1
+1	+1	+1	18.9, 19.0

(a) Estime todos los posibles efectos y diga cuáles son significativos. (b) Realice un análisis de varianza de estos datos y obtenga conclusiones generales. (c) ¿Hay un tratamiento ganador para minimizar? (d) Verifique residuos, ¿qué observa de destacado?

**EJERCICIO 89.** En una empresa del área electrónica una máquina toma componentes que le proporciona un alimentador, para montarlos o depositarlos en una tarjeta. Se ha tenido el problema de que la máquina falla en sus intentos por tomar el componente, lo cual causa paros de la máquina que detienen el proceso hasta que el operador se da cuenta y reinicia el proceso. Esto ocasiona tiempos muertos, aumento del tiempo de ciclo y baja productividad. Los intentos por corregir el problema han sido variar los parámetros de operación (incluso bajar la velocidad) sin tener la certeza de que la acción efectuada realmente redujo el problema. En este contexto, para diagnosticar mejor la situación, se decide correr un diseño de experimentos  $2^2$  con  $n = 2$  réplicas; en el que se tienen los siguientes factores y niveles  $(-, +)$ , respectivamente: A) velocidad de cam (70%, 100%), B) velocidad de

mesa (media, alta), C) orden o secuencia de colocación (continua, variable), D) alimentador (1, 2). Como el proceso es muy rápido, es necesario dejarlo operar en cada condición experimental el tiempo suficiente para reproducir el problema. Se consideró que esto se lograba con suficiente confianza con 500 componentes, por ello cada una de las corridas experimentales consistió en colocar 500 componentes, y se midieron dos variables de respuesta:  $Y_1$  = número de errores (o intentos fallidos), y  $Y_2$  = tiempo real (en segundos) para tomar y "colocar" los 500 componentes. Es evidente que se quiere minimizar ambas variables. Los datos obtenidos se muestran en la tabla de la página siguiente. (a) Al observar los datos obtenidos se deduce que hay algunos tratamientos que tienen pocos o ningún componente caídos, como por ejemplo el  $(-1, -1, +1, +1)$ , alguien muy "práctico" podría decidir poner la máquina a operar bajo estas condiciones, y olvidarse del análisis estadístico. De proceder así explique qué información se perdería. (b) Investigue qué efectos influyen de manera significativa sobre  $Y_1$  (apóyese en Pareto y ANOVA). (c) Obtenga el mejor ANOVA. (d) Si en el análisis anterior encuentra alguna interacción significativa, analice con detalle la más importante, e interprete en términos físicos. (e) ¿Qué tratamiento minimiza  $Y_1$ ? (f) Ahora investigue qué efectos influyen de manera relevante sobre  $Y_2$ . (g) ¿Qué tratamiento minimiza  $Y_2$ ? (h) Encuentre una condición satisfactoria tanto para minimizar  $Y_1$  y  $Y_2$ . (i) De los análisis de varianza para  $Y_1$  y  $Y_2$  observe el coeficiente R-cuadrada. ¿Qué concluye de ello? (j) Verifique residuos.

Factor A	Factor B	Factor C	Factor D	Réplica 1		Réplica 2	
				$Y_1$	$Y_2$	$Y_1$	$Y_2$
-1	-1	-1	-1	61	88	50	79
+1	-1	-1	-1	105	78	98	74
-1	+1	-1	-1	61	82	40	82
+1	+1	-1	-1	104	73	145	79
-1	-1	+1	-1	0	88	35	100
+1	-1	+1	-1	35	84	22	82
-1	+1	+1	-1	50	89	37	88
+1	+1	+1	-1	57	79	71	81
-1	-1	-1	+1	12	77	19	75
+1	-1	-1	+1	60	66	57	64
-1	+1	-1	+1	9	84	19	73
+1	+1	-1	+1	72	93	61	66
-1	-1	+1	+1	0	86	0	82
+1	-1	+1	+1	10	76	1	77
-1	+1	+1	+1	3	84	7	86
+1	+1	+1	+1	15	75	15	73

EJERCICIO 90. Un fabricante de bolsas de papel desea analizar la resistencia al rasgamiento ( $Y$ ), para lo cual utiliza una escala numérica. Examina tres factores, cada uno en dos niveles,  $x_1$  = papel,  $x_2$  = humedad,  $x_3$  = dirección rasguño. Decide obtener tres observaciones (réplicas) en cada combinación, las mismas que se muestran en la siguiente tabla:

$x_1$	$x_2$	$x_3$	Resistencia		
-	-	-	3.8	3.1	2.2
+	-	-	6.6	8.0	6.8
-	+	-	3.4	1.7	3.8
+	+	-	6.8	8.2	6.0
-	-	+	2.3	3.1	0.7
+	-	+	4.7	3.5	4.4
-	+	+	2.1	1.1	3.6
+	+	+	4.2	4.7	2.9

(a) Haga el análisis de varianza para estos datos. (b) Interprete los efectos significativos y encuentre el mejor tratamiento. (c) Verifique los supuestos del modelo.

EJERCICIO 91. En el área de SMT se busca reducir los defectos ocasionados por impresiones de soldadura en pasta inadecuada. Se corre un diseño  $2^4$  con dos réplicas y dos puntos centrales por réplica. Los factores son: altura de la mesa (A), velocidad de separación (B), velocidad de impresión (C), y presión de los escobillas (D). La variable de respuesta es la altura de la impresión de soldadura en pasta. El experimento se corrió en planta, pero como el proceso es muy rápido (la impresión de una tarjeta tarda menos de un minuto), entonces en estos casos se recomienda obtener más de un producto en cada condición experimental. Por ello se decidió que cada prueba experimental debería de consistir en dejar que el proceso se estabilizara y a partir de ahí imprimir diez tarjetas de manera consecutiva, a cada tarjeta se le midió la altura. Con estos diez datos se calculó la media y la desviación estándar, para así poder analizar el efecto de los factores sobre ambas. Una vez que se corre en orden aleatorio la primer réplica de todos los tratamientos, se deja de experimentar y al día siguiente se hace la segunda réplica de manera similar. Los datos se muestran a continuación:

Factor A	Factor B	Factor C	Factor D	Réplica 1		Réplica 2	
				Media	D. estan	Media	D. estan
-1	-1	-1	-1	6.8	0.17	6.3	0.18
+1	-1	-1	-1	6.9	0.28	6.6	0.51
-1	+1	-1	-1	6.4	0.17	5.8	0.41
+1	+1	-1	-1	6.6	0.29	6.6	0.19
-1	-1	+1	-1	6.8	0.27	6.5	0.19
+1	-1	+1	-1	8.7	0.80	7.3	0.75
-1	+1	+1	-1	6.7	0.16	6.4	0.21
+1	+1	+1	-1	7.8	0.64	7.1	0.60
-1	-1	-1	+1	5.5	0.28	5.3	0.15
+1	-1	-1	+1	5.8	0.51	5.4	0.24
-1	+1	-1	+1	5.8	0.14	5.3	0.21
+1	+1	-1	+1	5.5	0.19	5.4	0.13
-1	-1	+1	+1	6.1	0.29	6.0	0.34
+1	-1	+1	+1	6.6	0.38	6.2	0.50
-1	+1	+1	+1	6.6	0.26	5.6	0.25
+1	+1	+1	+1	6.7	0.22	6.3	0.37
0	0	0	0	6.5	0.25	6.0	0.53
0	0	0	0	6.4	0.27	5.8	0.5

(a) ¿Con qué finalidad se utilizan los puntos centrales?(b) Investigue qué efectos influyen de manera significativa sobre la altura promedio de la pasta (apóyese en



Pareto y ANOVA). (c) Si en el análisis anterior encuentra alguna interacción significativa, analice con detalle la más importante. (d) Si se quiere minimizar la altura de la pasta, ¿cuáles son las condiciones para lograrlo? (e) Ahora investigue qué efectos influyen de manera relevante sobre la variabilidad de la altura de la pasta. (f) Encuentre una condición satisfactoria tanto para minimizar la altura como para minimizar la variabilidad. (g) De los análisis de varianza para la media y la desviación estándar vea el coeficiente R-cuadrada. ¿Qué concluye de ello? (h) ¿Hay evidencia de curvatura? (i) Verifique residuos.

EJERCICIO 92. En la refinación de pulpa de madera interesa estudiar cómo se afecta la calidad de la fibra al introducirle cargas de material inorgánico con dos diferentes métodos o procesos (mecánico e in situ) y ver cómo interactúa al considerar otros factores. Para ello se decide correr un diseño  $2^4$  no replicado, con los siguientes factores y niveles.

Factor	Niveles (bajo, alto)	
A: Proceso	Mecánico	In situ
B: velocidad de agitación (rpm)	2000	3000
C: Tiempo (minutos)	30	60
D: Consistencia de la pulpa (%)	0.5	2.0

Se midieron cuatro variables de respuesta:  $Y_1$ : cenizas (%),  $Y_2$ : tensión (m),  $Y_3$ : blancura (%) y  $Y_4$ : opacidad (%). Los resultados para los 16 tratamientos se muestran en el cuadro siguiente, en el orden que se corrieron. (a) Bosqueje una gráfica de proceso en el que se muestren por un lado los factores controlados y por el otro las variables de respuesta. (b) Investigue qué efectos influyen significativamente sobre  $Y_1$  (apóyese en Pareto y ANOVA). (c) Si se considera que el factor de mayor interés es el A, comente la forma en que actúa sobre  $Y_1$ . (d) Comente con detalle la manera en que interactúa el factor A con los otros factores, sobre  $Y_1$ . (e) ¿Qué tratamiento maximiza  $Y_1$ ? (f) Verifique supuestos para la variable  $Y_1$ . (g) Repita los incisos anteriores para las otras variables de respuesta. (h) Encuentre condiciones satisfactorias para maximizar las cuatro variables de respuesta. (i) Haga un resumen del análisis hecho y destaque las principales conclusiones.

Tratamiento	$Y_1$	$Y_2$	$Y_3$	$Y_4$
c	0.48	579	86.17	77.85
bcd	1.46	692	86.5	76.28
bc	0.94	581	86.75	76.16
(1)	0.49	671	87.22	76.68
ad	1.55	653	88.55	79.3
b	1.49	867	87.62	77.15
a	7.49	496	88.77	79.75
ab	11.59	467	88.05	80.35
ac	13.23	437	87.3	78.72
bd	2.21	631	86.45	78.03
cd	5.06	565	85.75	79.42
abcd	7.78	549	88.45	80.81
abd	11.75	460	88.77	81.43
abc	12.57	462	89.42	81.55
d	0.72	620	87.67	77.87
acd	9.61	568	88.45	78.78

**EJERCICIO 93.** Se quiere aumentar el rendimiento de un proceso, y para ello se estudian tres factores con dos niveles cada uno. Se hacen tres repeticiones en cada tratamiento del diseño factorial  $2^3$  resultante. La variable de respuesta que se mide es rendimiento. Los datos son los siguientes:

Tratamiento	Repetición		
	1	2	3
(1)	22	31	25
a	32	43	29
b	35	34	50
ab	55	47	46
c	44	45	38
ac	40	37	36
bc	60	50	54
abc	39	41	47

(a) ¿Cuáles efectos están activos? (b) Si obtuvo una interacción importante, interprétela con detalle. c) Determine las condiciones de operación que maximizan el rendimiento. (d) ¿Cuál es la respuesta esperada en el mejor tratamiento? (e) Verifique los supuestos del modelo.

**EJERCICIO 94.** En una fábrica de componentes electrónicos, uno de los principales clientes ha reportado tener problemas con algunos de los productos (comportamiento eléctrico intermitente). Mediante el análisis de las muestras retornadas por el cliente, se identificó que el problema se relaciona con alambre mal colocado y ello podía obedecer a varias causas. Se decide correr una réplica de un experimento factorial  $2^5$ , utilizando los siguientes factores y niveles:

Factor	Niveles (bajo, alto)	
A: Patrón de reconocimiento	un punto	dos puntos
B: Sistema de luz	fibra	incandescente
C: Umbral (threshold)	725	850
D: Colocación del dado	girado	normal
E: Brillo de la oblea	brillo	normal

La respuesta a medir es el número de unidades con alambre mal colocado. Cada prueba se hizo en la línea de ensamble y consistió en colocar cierta cantidad de alambres, que por cierto lo hace un equipo automático. La cantidad de alambres a colocar en cada prueba, bajo cada tratamiento, se determinó de tal forma que tuviera alta probabilidad de detectar piezas con alambres mal colocados. Los datos son los siguientes:

(1)=105	d=0	e=34	de=0
a=0	ad=0	ae=3	ade=0
b=66	bd=0	be=18	bde=0
ab=7	abd=5	abe=2	abde=0
c=54	cd=25	ce=0	cde=0
ac=1	acd=1	ace=0	acde=0
bc=41	bcd=0	bce=49	bcde=0
abc=0	abcd=0	abce=4	abcde=0

(a) Dibuje el diagrama de Pareto y el gráfico de Daniel considerando todas las interacciones de alto orden. ¿Cuáles efectos parecen estar activos? (b) Determine el mejor análisis de varianza e interprételo. (c) Obtenga las gráficas de los efectos que resultaron importantes en el ANOVA e interprételas. (d) Determine el mejor

tratamiento. (e) Interprete con detalle las interacciones  $AD$  y  $AE$ , y considerando que el factor  $D$  es el resultado de un proceso anterior y que actualmente no es posible fijarlo en un nivel, y que el factor  $E$  es una característica de los materiales de un proveedor externo; entonces, ¿cuáles son sus recomendaciones para operar el proceso y por acciones de mejora a futuro? (f) Verifique los supuestos del modelo. ¿Que puede concluir del análisis de residuos? (g) ¿La forma especial de la gráfica de residuos contra predichos, afecta las conclusiones que ha hecho antes? (h) ¿Es pertinente colapsar este diseño en un factorial  $2^4$  con dos réplicas? Si la respuesta es posible hágalo. (i) ¿Se puede colapsar en un  $2^3$  con cuatro réplicas?

**EJERCICIO 95.** En una planta de fuerza se corrió un experimento factorial  $2^4$  con repeticiones al centro con el objetivo de hacer más eficaz la operación de la máquina de absorción. La eficacia de la máquina se mide en toneladas de refrigeración entre el flujo de vapor ( $tr/fv$ ). Los factores a controlar fueron: flujo de vapor ( $A$ ), temperatura de agua helada ( $B$ ), temperatura de agua de enfriamiento ( $C$ ), presión diferencial ( $D$ ). El diseño, escrito en orden aleatorio y en las unidades originales, es el que se muestra en la tabla siguiente. (a) Determine el mejor ANOVA para estos datos. (b) ¿Cuál porcentaje de la variación observada es explicado por el mejor ANOVA? (c) Realice la prueba de falta de ajuste. ¿Qué nos indica la presencia de efectos de curvatura? (d) Grafique los efectos significativos. (e) Determine el mejor tratamiento y haga la predicción de la eficacia esperada sobre él. (f) Verifique los supuestos de normalidad, varianza constante e independencia.

$A(t/h)$	$B(^{\circ}C)$	$C(^{\circ}C)$	$D(kg/cm^2)$	eficacia ( $tr/fv$ )
4.5	6.5	23	1.1	99
3.25	5.25	25	1.4	105
2.0	4.0	23	1.7	99
2.0	4.0	27	1.7	79
4.5	6.5	27	1.7	86
2.0	6.5	27	1.1	85
4.5	4.0	23	1.1	90
2.0	4.0	23	1.1	95
4.5	4.0	27	1.7	79
4.5	4.0	27	1.1	82
4.5	6.5	27	1.1	83
4.5	6.5	23	1.7	97
3.25	5.25	25	1.4	101
3.25	5.25	25	1.4	98
2.0	6.5	23	1.7	108
2.0	6.5	23	1.1	111
2.0	4.0	27	1.1	89
4.5	4.0	23	1.7	91
2.0	6.5	27	1.7	88

**EJERCICIO 96.** Una de las preocupaciones permanentes en la industria tequilera es obtener altos niveles de rendimiento, el cual puede depender de factores de control como: presión ( $A$ ), tiempo de cocimiento ( $B$ ), y tiempo de reposo ( $C$ ). Se decide realizar un experimento factorial a nivel piloto para investigar si estos factores tienen efecto sobre el rendimiento, medido éste por la eficiencia en mieles ( $EM$ ) y la eficiencia en agave cocido ( $EAC$ ). Los resultados obtenidos en una sola

*réplica del diseño se muestran a continuación:*

A	B	C	EM	EAC
-1	-1	-1	78.8	93.90
+1	-1	-1	94.0	97.30
-1	+1	-1	93.2	95.50
+1	+1	-1	95.9	96.91
-1	-1	+1	90.6	94.60
+1	-1	+1	97.9	98.60
-1	+1	+1	95.7	94.00
+1	+1	+1	90.5	95.80

(a) Haga una inspección visual de los datos y con base en ella establezca una conjetura sobre si los factores afectan o no a las variables de respuesta estudiadas. (b) ¿Qué diseño es el que se ha corrido? (c) Analice EM con el gráfico de Daniel (gráfica de efectos en papel normal) y el diagrama de Pareto. (d) A partir de lo anterior construya un error, obtenga el ANOVA para tal variable de respuesta. ¿Cuáles efectos están activos? (e) ¿Cómo afecta en el análisis el hecho de no haber replicado el experimento? (f) Determine el mejor tratamiento, si es que lo hay, para EM. Repita los incisos anteriores para la otra variable de respuesta, para cada una de las variables de rendimiento. (h) ¿Cuál es una posible causa de la poca variación observada en los datos de este experimento?

**EJERCICIO 97.** Una de las fallas más importantes en la línea de empaque de un producto es la calidad de las etiquetas. Un equipo de mejora decide atacar este problema mediante diseño de experimentos. Para ello eligen una de las impresoras a la cual se le pueden manipular los factores: velocidad, temperatura, tensión y tipo de etiqueta. Los niveles utilizados con cada factor fueron:

Factor	Nivel bajo	Centro	Nivel alto
Velocidad	baja	media	alta
Temperatura	5	13	21
Tensión	4	8	12
Tipo de etiqueta	esmaltada	otra	mate

El diseño factorial utilizado fue un  $2^4$  con repeticiones al centro. En cada combinación del experimento se imprimieron 20 etiquetas y se contabiliza como variable de respuesta en número de impresiones rechazadas. Los resultados observados, listados en orden aleatorio, fueron los que se muestran en la página siguiente. (a) Utilice la notación de Yates y anote en la primer columna de la tabla el código correspondiente a cada una de las corridas, y asegúrese que se corrieron todos los tratamientos correspondientes al diseño empleado. (b) Encuentre el mejor ANOVA para estos datos. (c) Grafique los efectos significativos e interpréte los para determinar el tratamiento ganador. (d) Estime el número de etiquetas que se espera sigan sin pasar en el mejor tratamiento. (e) Verifique supuestos. ¿Hay algún problema potencial?

**EJERCICIO 98.** Se hace un experimento para mejorar el rendimiento de un proceso, controlando cuatro factores en dos niveles cada uno. Se corre una réplica de diseño factorial  $2^4$ , con los factores tiempo (A), concentración (B), presión (C) y temperatura (D), y los resultados son los siguientes:

	$A_0$				$A_1$			
	$B_0$		$B_1$		$B_0$		$B_1$	
	$C_0$	$C_1$	$C_0$	$C_1$	$C_0$	$C_1$	$C_0$	$C_1$
$D_0$	12	17	13	20	12	17	13	20
$D_1$	10	19	13	17	10	19	13	17

(a) Analice estos datos con el uso de todos los criterios existente para encontrar el mejor ANOVA. En las figuras considere de entrada los 15 efectos posibles. (b) ¿Cuáles efectos están activos? Determine el mejor tratamiento. (c) Prediga el rendimiento esperado en el mejor tratamiento y de un intervalo de confianza para el rendimiento futuro. Compruebe los supuestos del modelo. (e) ¿Puede este diseño colapsarse en uno  $2^3$  con dos réplicas? De ser posible, hágalo y repita los incisos anteriores para este nuevo diseño.

EJERCICIO 99. Se realiza un experimento para mejorar el rendimiento de un proceso químico. Se seleccionan cuatro factores y se hace sólo una repetición en cada tratamiento, con los siguientes resultados:

Trat	Rend	Trat	Rend
(1)	90	d	98
a	74	ad	72
b	81	bd	87
ab	83	abd	85
c	77	cd	99
ac	81	acd	79
bc	88	bcd	87
abc	73	abcd	80

(a) Incluyendo las interacciones de mayor orden, analice los efectos con gráfica de Pareto y con la gráfica de papel normal para efectos. ¿Qué efectos parecen significativos? (b) A partir de lo anterior construya un error y haga un análisis de varianza. ¿Qué efectos son significativos? (c) Si hay una interacción importante, interprétela con detalle. (d) ¿Cuál es el mejor tratamiento? (e) Verifique los supuestos del modelo. (f) Si algún factor no tiene ningún tipo de efecto, colapse el diseño en un  $2^3$  ( $n=2$ ). Repita los incisos previos y compare los resultados.

EJERCICIO 100. En el ejemplo de la sección "Cuando casi ningún efecto es significativo: un ejemplo" de este capítulo, uno de los análisis alternativos fue colapsar un diseño  $2^5$  en uno  $2^3$  con cuatro réplicas. Para este diseño: (a) Haga el análisis completo del diseño colapsado e, interprete los efectos activos y verifique supuestos. (b) ¿Obtiene las mismas conclusiones que se obtuvieron con el análisis del diseño sin colapsar? ¿A qué lo atribuye?

EJERCICIO 101. En una empresa del área electrónica se quieren minimizar los problemas generados en el proceso conocido como "Soldadora de ola". Los defectos que se quieren reducir son insuficiencias de soldadura en las tarjetas. Los factores y niveles que inicialmente se deciden estudiar son: velocidad de conveyor (4 y 7 pies/minuto), temperatura de precalentado (80 y 120° C), y temperatura de soldadura (470 y 500° C). Debido a que el proceso es muy rápido (se suelda una tarjeta cada 10 a 15 segundos) se decide soldar en cada condición de prueba 25 tarjetas. La variable de respuesta es la cantidad de insuficiencias detectadas en los diferentes puntos de soldadura de las 25 tarjetas. Se hicieron dos réplicas. La matriz de diseño y los datos obtenidos se muestran a continuación:

Velocidad	Precalentado	Soldadura	Insuficiencias	
4	80	470	29	25
7	80	470	110	110
4	120	470	23	27
7	120	470	77	59
4	80	500	12	44
7	80	500	146	162
4	120	500	51	35
7	120	500	42	48

(a) Haga un análisis completo y determine los efectos más importantes, el ANOVA y el análisis de residuos. (b) A1 parecer la interacción velocidad-precalentado es importante, de ser así haga una interpretación detallada en términos físicos de tal interacción. (c) ¿Cuáles serían las condiciones de operación del proceso que podrían utilizarse para reducir la cantidad de insuficiencias? Analice las opciones disponibles.

**EJERCICIO 102.** El tequila es una bebida que está sujeta a una norma oficial mexicana (NOM-006-SCFI-1993), y conforme a ésta se debe cumplir con ciertas especificaciones físico-químicas. Por ejemplo, el alcohol anhidro no debe rebasar los 300 mg/100 ml, y los alcoholes superiores tampoco deben exceder los 400 mg/100 ml. En un laboratorio de investigación mediante un diseño factorial  $2^5$  no replicado se estudió la influencia de diversos factores sobre la producción de alcoholes superiores en la etapa de fermentación (Pinal et al. 1997). Los factores estudiados y los niveles fueron: tipo de cepa A(1, 2), temperatura, B(30, 35°C), fuente de nitrógeno, C-( $\text{NH}_4$ ) $_2$ S $O_4$  y urea-, relación carbono/nitrógeno, D(62/1, 188/1) y porcentaje de inóculo, E(5 y 10%). En la siguiente tabla se muestran los resultados obtenidos en cuanto a alcohol isoamílico (mg/1), que es parte de los alcoholes superiores.

(1)=21.4	d=42.5	e=32.9	de=54.0
a=16.8	ad=21.0	ae=17.5	ade=21.8
b=29.3	bd=79.1	be=30.0	bde=79.9
ab=12.7	abd=20.0	abe=24.1	abde=31.5
c=27.5	cd=48.6	ce=26.7	cde=47.9
ac=22.9	acd=27.1	ace=11.4	acde=15.6
bc=35.4	bcd=85.2	bce=23.9	bcde=73.8
abc=18.8	abcd=26.1	abce=18.0	abcde=25.4

(a) Dibuje el diagrama de Pareto y el gráfico de Daniel considerando todas las interacciones de alto orden. ¿Cuáles efectos parecen estar activos? (b) Determine el mejor análisis de varianza e interprételo. (c) Obtenga las gráficas de los efectos que resultaron importantes en el ANOVA, e interprételas con detalle. (d) Determine los tratamientos que minimizan y maximizan la variable de respuesta. (e) Verifique los supuestos del modelo. ¿Qué puede concluir del análisis de residuos? (f) ¿Es pertinente colapsar este diseño en un factorial  $2^4$  con dos réplicas? Si la respuesta es positiva, hágalo.

**EJERCICIO 103.** Interesa estudiar el efecto de la temperatura y del tiempo de remojo sobre la corrosión en navajas de rasurar. Cada observación se obtiene de someter una navaja al tratamiento y posteriormente ponerla en una cámara con 98% de humedad por 48 horas; la respuesta medida es el porcentaje de superficie que exhibe corrosión. Los resultados obtenidos para ocho cuchillas se muestran en la siguiente tabla:

<i>Temperatura (°F)</i>	<i>Tiempo (min.)</i>	
	<i>5 min.</i>	<i>60 min</i>
<i>650</i>	<i>30%, 20%</i>	<i>25%,30%</i>
<i>690</i>	<i>75%, 85%</i>	<i>95%,90%</i>

(a) ¿El tiempo de remojo y la temperatura afectan la corrosión de las navajas? (b) Dibuje las gráficas de los efectos activos y obtenga el tratamiento ganador.





## CHAPTER 6

# Diseños factoriales fraccionados $2^{k-p}$

## 1. Competencias

*Identificar un diseño factorial fraccionado y conocer sus respectivas ventajas y desventajas en determinadas situaciones. Comprender y diferenciar los conceptos de resolución III, IV y V, y su aplicación en la elección de una fracción apropiada. Construir diseños fraccionados con cualquier grado de fraccionamiento. Seleccionar la fracción adicional más adecuada para aclarar ambigüedades heredadas de una primera fracción.*

## 2. Introducción

Al crecer el número de factores también crece rápido el número de tratamientos en los diseños factoriales completos  $2^k$ . Por ejemplo, para  $k = 6$  factores una sola réplica del diseño factorial completo  $2^6$  implica correr 64 pruebas, que corresponden al número de tratamientos del diseño; para  $k = 7$  son  $2^7 = 128$  puntos de diseño. Prácticamente no es posible hacer tantas corridas experimentales, porque además de que resulta costoso es innecesario. Una estrategia que reduce de manera importante el número de puntos experimentales y al mismo tiempo controla que se pierda el mínimo de información valiosa consiste en construir *diseños factoriales fraccionados* los cuales permiten reducir el número de corridas experimentales, y al mismo tiempo obtener la información acerca de los efectos considerados de antemano relevantes y se basa en una jerarquización de los efectos: *son más importantes los efectos principales, seguidos por las interacciones dobles, luego las triples, cuádruples, etcétera.*

El primer diseño factorial completo que genera un importante exceso de información es el factorial completo  $2^5$ , ya que este diseño permite estimar 31 efectos, de los cuales sólo 15 son potencialmente importantes (los 5 efectos principales más las 10 interacciones dobles) y los 16 restantes efectos, conformados por las interacciones de tres o más factores, prácticamente se pueden ignorar a priori. Mostraremos que con una fracción a la mitad del diseño factorial completo  $2^5$  ( $\frac{1}{2}2^5 = 2^{5-1}$ ), se pueden estimar limpiamente los 15 efectos potenciales importantes, sacrificando la información relativa a las 16 interacciones de alto orden que no interesan.

- *Con un diseño factorial fraccionado  $2^{5-1}$  se puede obtener esencialmente la misma información que con el factorial completo  $2^5$ , pero con la mitad del costo experimental.*

Cuando se tienen menos de cinco factores ( $k < 5$ ), los efectos potencialmente importantes superan en número a los efectos ignorables a priori, de aquí que si se fraccionan estos diseños, se pierde por fuerza información que puede ser relevante. Ahora, cuando  $k \geq 5$  el número de efectos ignorables supera el número de efectos no ignorables o potencialmente importantes, lo cual indica que estos diseños se pueden

fraccionar muchas veces sin perder información valiosa. Mientras más grande es el valor de  $k$ , el diseño admite un grado de fraccionamiento mayor.

Al correr sólo una fracción del diseño factorial completo ocurren dos hechos inevitables:

- (1) Pérdida información, ya que habrá efectos que no podrán estimarse y se tienen menos grados de libertad disponibles para el error. Los efectos que se pierden se espera que sean, en la medida de lo posible, interacciones de alto orden, las cuales se pueden ignorar de antemano con bajo riesgo.
- (2) Los efectos que sí se pueden estimar tienen al menos un *alias* (efectos que tienen el mismo contraste). *El que un efecto sea alias de otro significa que son en realidad el mismo efecto con nombres distintos, y al estimar a uno de ellos se estima al mismo tiempo el otro, de manera que no se pueden separar.* Cuando el experimentador elige una fracción en la que dos efectos potencialmente importantes son alias, debe contar de antemano con una estrategia de interpretación del efecto estimado.

Tabla 1. Efectos en los factoriales $2^k$ .			
Diseño $2^k$	Total de efectos	Efectos no ignorables	Efectos ignorables
$2^2$	3	3	0
$2^3$	7	6	1
$2^4$	15	10	5
$2^5$	31	15	16
$2^6$	63	21	42
$2^7$	127	28	99

### 3. Diseños factoriales fraccionados $2^{k-1}$

La notación  $2^{k-1}$  significa una fracción a la mitad del diseño factorial completo  $2^k$ ,  $k > 2$  ( $\frac{1}{2}2^k = 2^{k-1}$ ). No tiene sentido fraccionar el diseño factorial  $2^2$  porque prácticamente desaparece al tener sólo cuatro tratamientos, fraccionarlo a la mitad implicaría correr dos tratamientos y con ellos no se podrían estimar ni siquiera los dos efectos principales.

**3.1. Diseños factoriales fraccionados  $2^{3-1}$ .** El primer diseño que se puede fraccionar (aunque veremos que no se recomienda hacerlo) es el factorial completo  $2^3$ , el cual escrito en la notación estándar se muestra en la tabla 2. Si queremos fraccionar a la mitad este diseño, entonces es necesario seleccionar cuatro de entre los ocho tratamientos. De entrada sabemos que existen  $\binom{8}{4} = 70$  posibles maneras de elegir una fracción a la mitad del diseño factorial  $2^3$ , por lo que surge la interrogante sobre cuál o cuáles de esas 70 posibles fracciones son las más adecuadas. Lógicamente, es adecuada aquella fracción que permita estimar los efectos considerados importantes y sólo existen dos elecciones adecuadas.

Con el diseño  $2^3$  completo se pueden estimar siete efectos:  $A$ ,  $B$ ,  $C$ ,  $AB$ ,  $AC$ ,  $BC$  y  $ABC$ . De acuerdo a su jerarquía, el efecto menos importante a priori es la interacción triple  $ABC$ , así que éste es el efecto más sacrificable para generar la fracción a la mitad de manera que se pierda un mínimo de información.

La generación de la fracción se hace con base en los signos del contraste  $ABC$ , los signos “+” del contraste  $ABC$  señalan a los tratamientos que conforman la llamada fracción principal, y los signos “-” señalan la fracción complementaria

(tabla 2). Los dos diseños factoriales fraccionados  $2^{3-1}$  así generados proporcionan la misma calidad de información sobre los efectos potencialmente importantes y cualquier otra elección de cuatro tratamientos daría peor resultado.

En la tabla 3, la fracción 1 es la fracción principal generada por  $I = +ABC$  y la fracción 2 o complementaria se genera con  $I = -ABC$ . La letra  $I$  surge porque este efecto generador hace las veces de *identidad o neutro multiplicativo*. De aquí que el efecto no estimable  $ABC$  se llama *generador de la fracción*, puesto que su contraste es base para construir las dos fracciones.

Al correr alguna de las fracciones que se muestran en la tabla 3 no se podrá estimar el efecto  $ABC$ , puesto que no tiene contraste, debido a la manera en que se generan las fracciones. Por ejemplo, para la fracción 1, el contraste  $ABC$  tiene sólo signos  $+$ , luego no existe tal contraste. Al tener sólo signos positivos el “contraste  $ABC$ ”, podemos decir que se *confunde* o se *alía* con el total de los datos, o dicho de otro modo, el efecto  $ABC$  se *confunde* con la media global  $\mu$ .

Tabla 2. Diseño factorial completo $2^3$ y contraste ABC			
A	B	C	ABC
-1	-1	-1	-
1	-1	-1	+
-1	1	-1	+
1	1	-1	-
-1	-1	1	+
1	-1	1	-
-1	1	1	-
1	1	1	+

Tabla 3. Dos posibles diseños fraccionados $2^{3-1}$							
Fracción 1 ( $I = +ABC$ )				Fracción 2 ( $I = -ABC$ )			
A	B	C		A	B	C	
1	-1	-1	a	-1	-1	-1	(1)
-1	1	-1	b	1	1	-1	ab
-1	-1	1	c	1	-1	1	ac
1	1	1	abc	-1	1	1	bc

**3.2. Representación geométrica del diseño  $2^{3-1}$ .** En la *figura 1* se muestra la representación geométrica de las dos fracciones del diseño  $2^3$ , nótese que cada fracción tiende a cubrir toda la región experimental delimitada por el cubo.

**3.3. Estructura de alias del diseño  $2^{3-1}$  con  $I = ABC$ .** Al estimar los efectos potencialmente importantes con cualquiera de las fracciones dadas en la tabla 3, resulta que cada efecto estimado tiene un alias. Consideremos por ejemplo el diseño factorial fraccionado  $2^{3-1}$  dado por la fracción 1 de tabla 3. Este diseño se generó con  $I = +ABC$ , que en este caso también es la relación definidora, ya que define totalmente la estructura de *alias* del diseño. La estructura de *alias* consiste en escribir explícitamente cuáles son los alias de cada efecto, y esta estructura se deduce fácilmente del generador de la fracción, considerando el signo utilizado. Por ejemplo, de la tabla 3, el contraste del efecto A está dado por

$$\text{Contraste } A = a + abc - b - c,$$

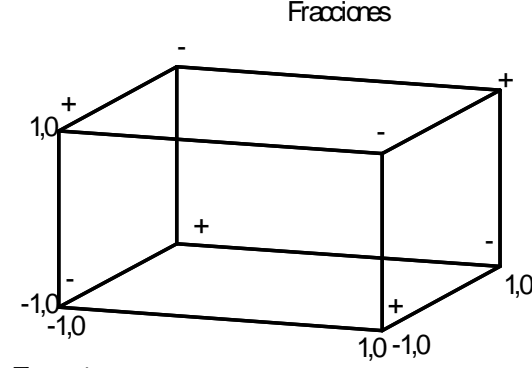


Figura1

mientras que al multiplicar las columnas  $B \times C$  se obtiene

$$\text{Contraste } BC = a + abc - b - c.$$

Observe que  $\text{Contraste } A = \text{Contraste } BC$ , lo que significa que los efectos  $A$  y  $BC$  son alias, porque al estimar el efecto  $A$  también se estima el efecto  $BC$ . Así, en realidad se estima la suma  $A + BC$  de ambos efectos y no se sabe con certeza cuál es el que predomina o si ambos afectan. Como hemos mencionado arriba, dos efectos alias son inseparables porque comparten el mismo contraste, y por ende, son dos nombres para el mismo efecto. De igual forma se puede ver que:

$$\text{Contraste } B = \text{Contraste } AC = b + abc - a - c$$

$$\text{Contraste } C = \text{Contraste } AB = c + abc - b - a$$

$B$  es alias de  $AC$  y  $C$  es alias de  $AB$ . En resumen, la estructura de alias del diseño factorial fraccionado  $2^{3-1}$  está dada por

$$(3.1) \quad \begin{aligned} &A + BC \\ &B + AC \\ &C + AB \end{aligned}$$

La estructura alias de un diseño factorial fraccionado se puede obtener de la relación definidora del diseño. En este caso la relación definidora coincide con el generador  $I = ABC$ . Multiplicando cada efecto por esta relación, con el uso de multiplicación módulo 2, se obtiene directamente estructura de alias. *La multiplicación módulo 2 significa que al multiplicar cualquier efecto por la identidad es igual al efecto, y al multiplicar un efecto por sí mismo es igual a la identidad*; por ejemplo, aplicando esto para el efecto  $A$ , tendríamos que  $A \times I = A$  y que  $A \times A \pmod{2} = A^2 = A^0 = I$ . Entonces, el alias de  $A$  se obtiene al multiplicar por  $A$  ambos lados de la relación  $I = ABC$ , y resulta que

$$A = A \times I = A \times ABC = A^2BC = BC$$

y también, los alias de  $B$  y  $C$  son, respectivamente

$$B = B \times I = B \times ABC = AB^2C = AC$$

y

$$C = C \times I = C \times ABC = ABC^2 = AB,$$

Con lo que resulta la estructura de alias que obtuvimos antes (véase 3.1) a partir de los contrastes.

3.3.1. *Interpretación de efectos alias.* Para interpretar los efectos alias o aliados es necesario suponer que sólo uno de ellos es el responsable del efecto observado y que los demás efectos son nulos. En general no es buena estrategia utilizar diseños fraccionados donde se alían dos efectos que son potencialmente importantes, como son los efectos principales y las interacciones dobles, sin embargo, habrá situaciones en las que no queda otra alternativa.

3.3.2. *Estructura de alias del diseño  $2^{3-1}$  con  $\mathbf{I} = -\mathbf{ABC}$ .* La estructura alias para el diseño  $2^{3-1}$  con relación definidora  $I = -ABC$  está dada por

$$(3.2) \quad \begin{aligned} &A - BC \\ &B - AC \\ &C - AB. \end{aligned}$$

En este caso, al estimar los efectos  $A$ ,  $B$  y  $C$  realmente se estiman  $A - BC$ ,  $B - AC$  y  $C - AB$ , respectivamente. Note que el signo del generador lo hereda la estructura de alias. Esto implica que si el experimentador se arrepiente de haber corrido el diseño  $2^{3-1}$  con generador  $I = ABC$ , por no estar seguro de cómo interpretar su estructura de alias, todavía puede correr la fracción complementaria con generador  $I = -ABC$  para completar el diseño  $2^3$ . Combinando las estructuras alias de ambas fracciones de manera apropiada se pueden separar los efectos alias; primero, para separar los efectos principales las estructuras se suman y se divide entre dos, y segundo, para estimar los efectos de interacción las estructuras se restan y se divide entre dos. Por ejemplo, el efecto  $A$  se obtiene de

$$\frac{(A + BC) + (A - BC)}{2} = A,$$

y la interacción doble  $BC$  se separa como

$$\frac{(A + BC) - (A - BC)}{2} = BC.$$

#### 4. El concepto de resolución

Bajo el supuesto de que los efectos principales son más importantes que las interacciones de dos factores, y éstas a su vez son más relevantes que las de tres, y así sucesivamente, es conveniente utilizar diseños factoriales fraccionados que tengan alta resolución.

4.0.3. *Definiciones de resolución.* Un diseño factorial fraccionado es de resolución  $R$  si los efectos formados por la interacción de  $P$  factores no son alias de efectos de interacción que tengan menos de  $R - P$  factores. A mayor resolución se observa más claramente lo que sucede con los efectos potencialmente importantes. Para fines prácticos basta entender las definiciones particulares de resolución III, IV y V.

- (1) *Diseños de resolución III.* En estos diseños los efectos principales no son alias entre ellos, pero existen efectos principales que son alias de alguna interacción doble. Por ejemplo, el diseño  $2^{3-1}$  con relación definidora  $I = ABC$  (o  $I = -ABC$ ) es de resolución III.

- (2) *Diseños de resolución IV.* En este diseño los efectos principales no están alias entre ellos ni con las interacciones dobles, pero algunas interacciones dobles están alias con otra interacción doble. Por ejemplo, el diseño  $2^{4-1}$  con relación definidora  $I = ABCD$  (o  $I = -ABCD$ ) es de resolución IV.
- (3) *Diseños de resolución V.* En estos diseños los efectos principales y las interacciones dobles están alias con interacciones triples o de mayor orden, es decir, los efectos principales e interacciones dobles están limpiamente estimados. Por ejemplo, el diseño  $2^{5-1}$  con relación definidora  $I = ABCDE$  (o  $I = -ABCDE$ ) es de resolución V.

En general, en los diseños factoriales fraccionados en dos niveles, la resolución está dada por la “palabra o efecto” de la relación definidora con el menor número de letras. Por ejemplo, en los diseños  $2^{k-1}$  la resolución es igual al número de letras del generador, ya que éste es al mismo tiempo la relación definidora. Así, las fracciones  $2^{3-1}$ ,  $2^{4-1}$  y  $2^{5-1}$  tienen resolución III, IV y V, respectivamente; porque sus correspondientes generadores se componen de 3, 4 y 5 letras. Tanto los efectos principales como las interacciones dobles están limpiamente estimados, es decir, sus alias son interacciones triples en adelante.

### 5. Construcción de fracciones $2^{k-1}$

Una manera de construir en dos pasos diseños fraccionados  $2^{k-1}$  con la más alta resolución posible es la siguiente:

- (1) Se lista el diseño factorial completo para  $k - 1$  factores, y de esta forma se tienen las primeras  $k - 1$  columnas de la fracción deseada.
- (2) La columna faltante (la  $k$ -ésima) se construye multiplicando entre sí las columnas previas. Si se quiere la fracción complementaria se cambian los signos de esta última columna. El diseño que resulta es un diseño factorial fraccionado  $2^{k-1}$  con resolución máxima  $R = k$ .

#### EJEMPLO 13. Construcción del diseño $2^{4-1}$

Estos dos pasos aplicados en la construcción del diseño factorial fraccionado  $2^{4-1}$  con resolución IV y con generador  $I = -ABCD$  quedan de la siguiente manera:

Primero se lista el diseño factorial completo  $2^{4-1} = 2^3$  dejando en blanco los espacios para los niveles del factor  $D$ . La columna faltante de niveles para el factor  $D$  se obtiene al multiplicar las columnas  $A$ ,  $B$  y  $C$  de acuerdo al generador. En este caso el generador indica que  $D = -ABC$ . Haciendo el producto  $-ABC$  se obtienen las tres columnas que conforman el diseño factorial fraccionado  $2^{4-1}$  con generador  $I = -ABC$  que se muestra en la siguiente tabla:

A	B	C	D		A	B	C	$D = -ABC$
—	—	—			—	—	—	+
+	—	—			+	—	—	—
—	+	—			—	+	—	—
+	+	—			+	+	—	+
—	—	+			—	—	+	—
+	—	+			+	—	+	+
—	+	+			—	+	+	+
+	+	+			+	+	+	—

Si se quisiera la fracción principal que tiene generador  $I = +ABC$ , el primer paso es el mismo, y en el segundo paso los niveles de  $D$  se obtienen con el producto positivo de las columnas ( $D = +ABC$ ). Una ventaja de la fracción complementaria que se acaba de construir respecto a la fracción principal es que no contiene las combinaciones de niveles más extremas  $(-, -, -, -)$  y  $(+, +, +, +)$ .

**5.1. Ejemplo ilustrativo de un diseño de experimento fraccionado  $2^{5-1}$ .** A continuación presentamos el ejemplo ilustrativo que habíamos titulado en el capítulo anterior como *Experimento  $2^5$  no replicado* con una sola réplica, pero esta vez sólo se correra a la mitad ( $2^{5-1}$ ), y se notará que las conclusiones son las mismas que se obtuvieron con el factorial completo, con la diferencia de que el costo experimental sería la mitad.

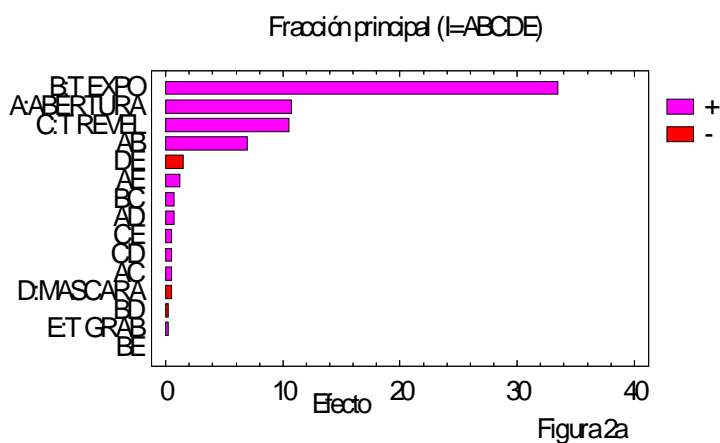
Recordemos que el experimento se desarrolló en una planta donde se fabrican semiconductores, en donde se quería mejorar el rendimiento del proceso mediante el diseño de experimentos. De acuerdo a la experiencia del grupo de mejora, los factores que podían tener mayor influencia sobre a variable de respuesta (rendimiento), así como los niveles de prueba utilizados, fueron los siguientes:

$A =$	Nivel de la abertura (pequeña, grande)
$B =$	Tiempo de exposición (20%, abajo, 20% arriba)
$C =$	Tiempo de revelado (30 s, 45 s)
$D =$	Dimensión de la máscara (pequeña, grande)
$E =$	Tiempo de grabado (14.5 min, 15.5 min)

Consideremos las fracciones a la mitad para este diseño  $2^5$ . La construcción de las fracciones se hace en dos pasos:

- (1) Se escribe el diseño  $2^{5-1}$  si fuera un factorial completo  $2^4$ .
- (2) La columna faltante es el producto de las cuatro primeras. Esta fracción corresponde a la fracción principal (generador  $I = ABCDE$ ). Si se toma el producto negativo de las cuatro columnas, se obtiene la fracción complementaria (generador  $I = -ABCDE$ ). Ambas fracciones juntas conforman el diseño factorial completo  $2^5$ .

$A$	$B$	$C$	$D$	$E = +ABCD$	$E = -ABCD$
-	-	-	-	+	-
+	-	-	-	-	+
-	+	-	-	-	+
+	+	-	-	+	-
-	-	+	-	-	+
+	-	+	-	+	-
-	+	+	-	+	-
+	+	+	-	-	+
-	-	-	+	-	+
+	-	-	+	+	-
-	+	-	+	+	-
+	+	-	+	-	+
-	-	+	+	+	-
+	-	+	+	-	+
-	+	+	+	-	+
+	+	+	+	+	-



Se generan las dos fracciones mencionadas y se buscan en la tabla de datos los valores que corresponden a cada tratamiento y en cada fracción. Las dos fracciones con sus respectivos datos se muestran en la tabla 4.

Tabla 4. Dos fracciones $2^{5-1}$			
Fracción principal $I = +ABCDE$		Fracción complementaria $I = -ABCDE$	
$e = 8.0$	$d = 8.0$	$(1) = 7.0$	$de = 6.0$
$a = 9.0$	$ade = 10.0$	$ae = 12.0$	$ad = 10.0$
$b = 34.0$	$bde = 30.0$	$be = 35.0$	$bd = 32.0$
$abe = 52.0$	$abd = 50.0$	$ab = 55.0$	$abde = 53.0$
$c = 16.0$	$cde = 15.0$	$ce = 15.0$	$cd = 18.0$
$ace = 22.0$	$acd = 21.0$	$ac = 20.0$	$acde = 20.0$
$bce = 45.0$	$bcd = 44.0$	$bc = 40.0$	$bcde = 41.0$
$abc = 60$	$abcde = 63.0$	$abce = 65.0$	$abcd = 61.0$

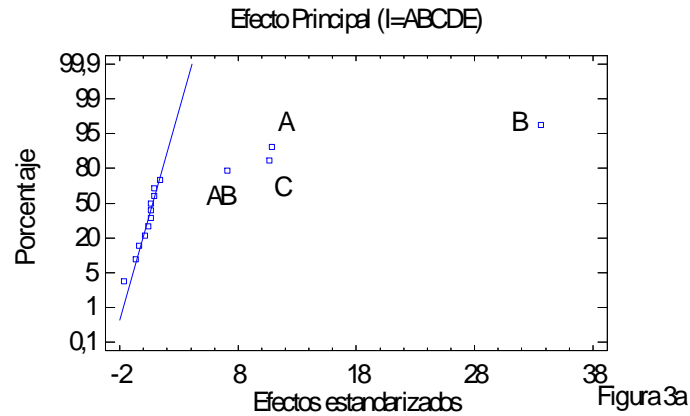
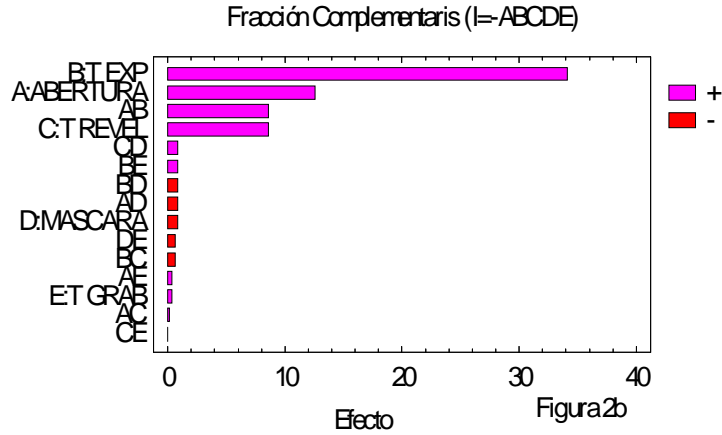
**Análisis de las dos fracciones.** En las figuras 2a y 2b se muestran los diagramas de Pareto para cada una de las fracciones analizadas por separado. Vemos que cualquiera de las fracciones detecta los efectos  $A$ ,  $B$ ,  $C$  y  $AB$  como los más importantes y los detecta en el mismo orden. Estos paretos son fundamentalmente iguales que en el análisis del diseño completo.

Con los gráficos de *Daniel* pasa lo mismo: ambas fracciones detectan los mismos efectos en apariencia significativos (véase figura 3).

Basta decir que los análisis de varianza resultan los mismos, debido a la claridad con la que trabajan los paretos y los gráficos de Daniel.

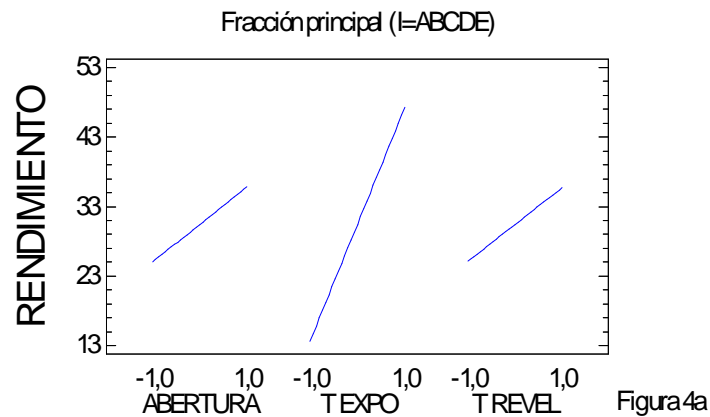
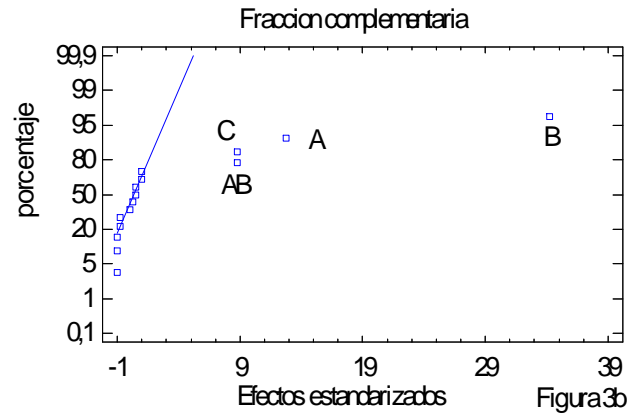
**Interpretación.** En las figuras 4a y 4b se muestran los efectos principales detectados en cada una de las fracciones y en las figuras 5a y 5b se presenta el efecto de interacción  $AB$  de manera que ha sido detectado por cada una de las fracciones. Las diferencias entre lo que se detecta con una fracción y otra es prácticamente imperceptible. Si comparamos estas gráficas contra las correspondientes del diseño





factorial completo en las figuras 5c y 5d, notamos que desde el punto de vista físico se detecta que pasa lo mismo a medida que se incrementan los niveles de los factores  $A$ ,  $B$  y  $C$ , el rendimiento se incrementa. Además el efecto del nivel de abertura ( $A$ ) es mayor cuando el tiempo de exposición es alto ( $B$ ). Por tanto, al igual que con el factorial completo, el mejor tratamiento es ( $A^+$ ,  $B^+$ ,  $C^+$ ,  $D^S$ ,  $E^S$ ). Los niveles de los factores  $D$  y  $E$  se eligen con el uso del criterio económico.

En resumen, el diseño factorial  $2^5$  sobre rendimiento de semiconductores se puede analizar de distintas maneras y los resultados son equivalentes. Se ha analizado como un  $2^5$  ( $n = 1$ ), como un  $2^3$  ( $n = 4$ ) usando la propiedad de colapsación, como una fracción  $2^{5-1}_V$  ( $I = ABCDE$ ) y su complementaria  $2^{5-1}_V$  ( $I = -ABCDE$ ). Algo importante a resaltar es que bastaba una fracción a la mitad, y por tanto, la mitad de los recursos invertidos, para obtener las mismas conclusiones. Se puede mostrar que incluso con una fracción a la cuarta parte del diseño original es posible llegar en este ejemplo a las mismas conclusiones.



## 6. Diseños factoriales fraccionados $2^{k-2}$

De acuerdo al número de factores y del costo de cada corrida experimental en ocasiones es necesario correr una fracción más pequeña del diseño factorial completo  $2^k$ . Un diseño factorial fraccionado  $2^{k-2}$  representa la cuarta parte del factorial original completo ( $\frac{1}{4}2^k = 2^{k-2}$ ). Para obtener este diseño se necesitan dos efectos generadores, de entre las interacciones del orden más alto orden. Estos efectos generadores (interacciones) deben ser elegidos de manera que su producto sea también

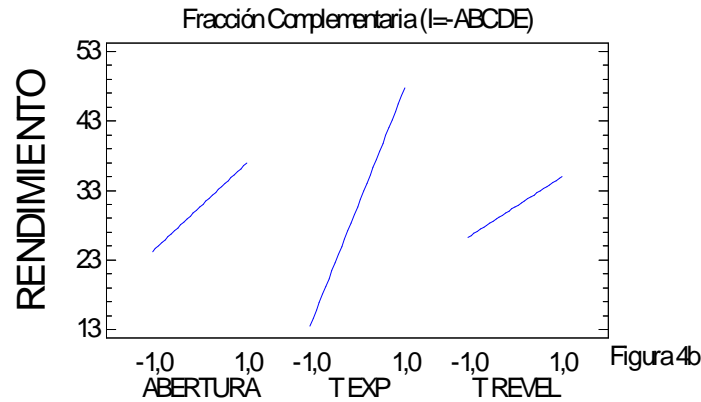


Figura 4b

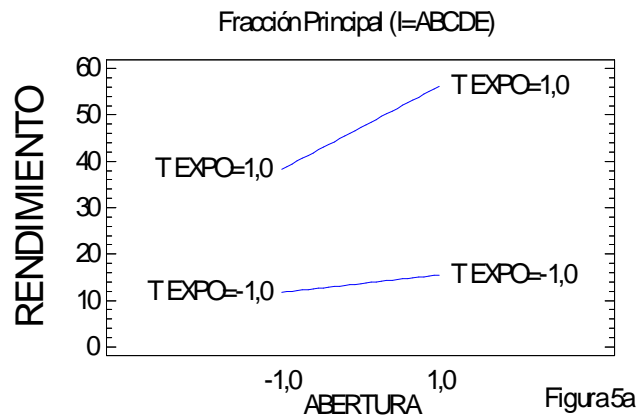
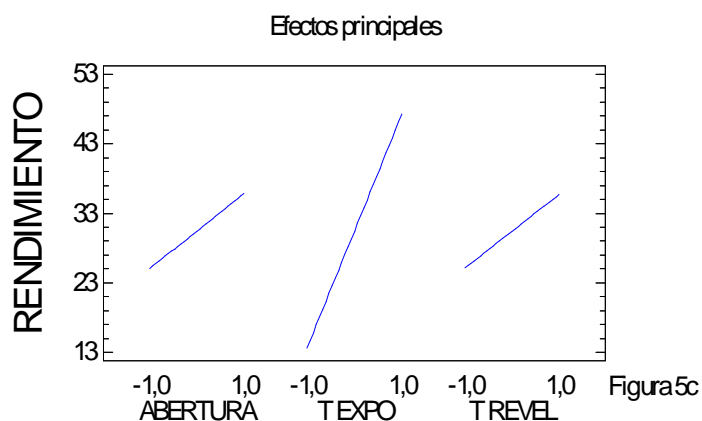
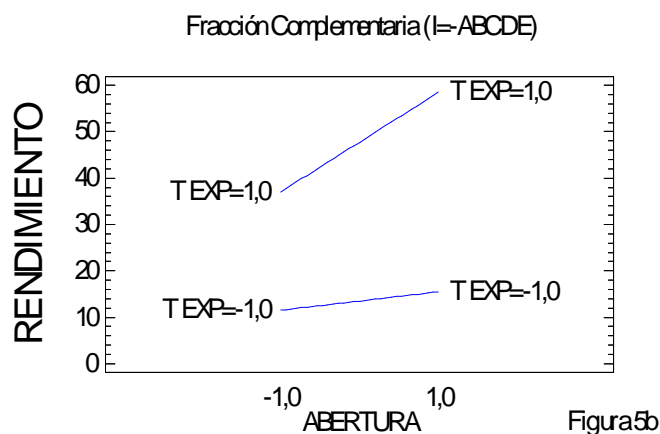


Figura 5a

una interacción del más alto orden posible. *Estos diseños tendrán tres generadores: los primeros dos que se seleccionaron más su producto entre sí, y ninguno será estimado. La estructura alias del diseño se obtiene de su relación definidora, la cual está dada por los dos generadores iniciales y su producto, de aquí que cada efecto tiene tres alias.* En general, el número de “palabras” de la relación definidora indica el número de alias que tendrá cada efecto, y multiplicando un efecto dado por esta relación se determinan sus alias. La palabra con menos letras en la relación definidora indica la resolución de la fracción.

**Construcción en dos pasos del diseño  $2^{k-2}$ .** El diseño factorial fraccionado  $2^{k-2}$  se puede construir en dos pasos:

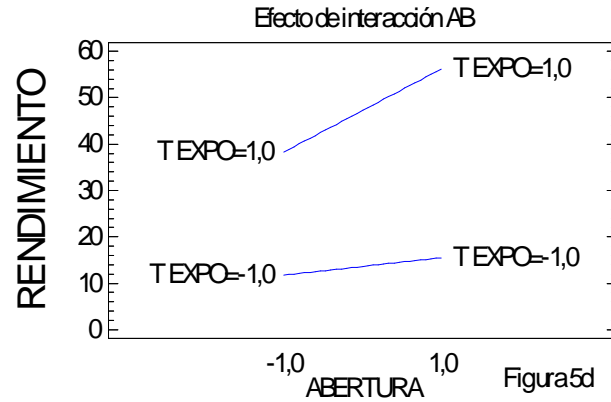
- (1) Se escribe el diseño  $2^{k-2}$  como si fuera un factorial completo en  $k - 2$  factores, y de esta forma se tienen los niveles de los primeros  $k - 2$  factores.



- (2) Los niveles que corresponden a los factores de las dos últimas columnas (factores  $k-1$  y  $k$ ) se obtienen multiplicando columnas previas de acuerdo a los generadores.

#### EJEMPLO 14. *Fracción* $2^{5-2}$

Supongamos que se tienen cinco factores y que sólo se tienen recursos para correr una cuarta parte del diseño factorial  $2^5$ , es decir, se quiere correr un diseño fraccionado  $2^{5-2}$ , que se construye en los dos pasos siguientes:



- (1) Se escribe el diseño factorial completo  $2^3$  para los tres primeros factores  $A$ ,  $B$  y  $C$ , dejando las columnas  $D$  y  $E$  en blanco:

Tabla 5. Diseño $2^{5-2}$ $I=ABD=ACE$				
A	B	C	D	E
-	-	-	+	+
+	-	-	-	-
-	+	-	-	+
+	+	-	+	-
-	-	+	+	-
+	-	+	-	+
-	+	+	-	-
+	+	+	+	+

- (2) Los niveles para los factores  $D$  y  $E$  se obtienen al seleccionar de manera adecuada generadores. En este caso se proponen como los generadores  $I = ABD$  e  $I = ACE$ , y el tercer generador es el producto  $ABD \times ACE = BCDE$ . Así, la relación definidora del diseño queda como:  $I = ABD = ACE = BCDE$ . Al reescribir los generadores en la forma:  $I = ABD \Rightarrow D = AB$  e  $I = ACE \Rightarrow E = AC$ , entonces con los productos  $AB$  y  $AC$  se generan los niveles del factor  $D$  y  $E$ , respectivamente. Haciendo los productos indicados se completa la tabla anterior y se tiene el diseño factorial que se muestra en la tabla 5, fraccionado  $2^{5-2}$  con resolución  $III$ , puesto que en la relación definidora los efectos más pequeños tienen tres letras.
- (3) La estructura de alias se obtiene al multiplicar cada efecto por la relación definidora dada por:

$$I = ABD = ACE = BCDE$$

y se muestra en la tabla 6. Los grupos de efectos alias tienen ahora cuatro elementos. Cada efecto principal tiene al menos una interacción doble como su alias, que es lo que implica la resolución  $III$  del diseño. Y no es posible construir una fracción  $2^{5-2}$  con resolución mayor que  $III$ .

La estructura de alias completa incorpora hasta los efectos de interacción del más alto orden, aunque no sean de interés. Toda la información importante está contenida en la estructura de alias reducida, que involucra solo hasta las interacciones dobles (véase tabla 7).

Cuando se alían efectos con la misma jerarquía, como es el caso de  $BE + CE$ , debe decidirse, con base en el conocimiento del proceso, a cuál interacción se atribuirá el efecto observado, en caso de que resulte significativo. Otro criterio es fijarse cuáles efectos principales resultaron significativos, ya que éstos tienen más probabilidad de estar activos también en sus interacciones. Con los generadores  $I = ABD$  e  $I = ACE$  se obtiene el diseño  $2^{5-2}$  llamado fracción principal que hemos presentado, pero tomando las otras combinaciones de signos en los generadores:

$$\begin{aligned} I &= -ABD, I = ACE \text{ ó} \\ I &= ABD, I = -ACE \text{ ó} \\ I &= -ABD, I = -ACE \end{aligned}$$

se obtienen otros tres posibles diseños  $2^{5-2}$  que pertenecen a la misma familia, y todos con la misma resolución. La decisión por una fracción en particular se hace evitando aquellas fracciones que contienen los tratamientos más extremos como la combinación  $(+, +, +, +, +)$ , que implica correr todos los factores en sus niveles altos. En muchos procesos puede ser riesgoso correr combinaciones extremas, aunque esto también depende del ancho de los niveles utilizados en cada factor

<p>Tabla 6. Estructura alias completa del diseño <math>2^{5-2}</math></p> <p>A + BD + CE + ABCDE  B + AD + ABCE + CDE  C + ABCD + AE + BDE  D + AB + ACDE + BCE  E + ABDE + AC + BCD  BC + ACD + ABE + DE  BE + ADE + ABC + CD  <math>\mu</math> + ABD + ACE + BCDE</p>	$\Rightarrow$	<p>Tabla 7. Estructura alias reducida del diseño <math>2^{5-2}</math></p> <p>A + BD + CE  B + AD  C + AE  D + AB  E + AC  BC + DE  BE + CE</p>
---	---------------	--

## 7. Diseño factorial fraccionado $2^{k-p}$

En general un diseño factorial fraccionado  $2^{k-p}$  es una fracción  $1/2^p$  del diseño factorial completo  $2^k$ . Para construir un diseño  $2^{k-p}$  se eligen  $p$  generadores iniciales, todos interacciones del más alto orden posible, de manera que todos sus productos también sean interacciones de alto orden. Una vez elegidos los  $p$  generadores, el diseño se puede construir en dos pasos a saber: (1). Se escribe el diseño  $2^{k-p}$  como si fuera el factorial completo para  $k - p$  factores. (2). Para los últimos  $p$  factores las columnas de signos se obtienen multiplicando las columnas que indican los generadores. De la relación definidora del diseño se obtiene la estructura alias y la resolución de la fracción resultante.

En todo el procedimiento descrito quizá lo más difícil es encontrar los mejores generadores de la fracción que se desea utilizar. Afortunadamente existen tablas de diseños factoriales fraccionados que incluyen su estructura de alias (véase McLean & Anderson, 1984), lo que permite al investigador elegir el diseño y saber, con anticipación, cuáles serían los alias de los efectos potencialmente importantes. Esta

selección adecuada del diseño factorial fraccionado también se puede hacer en software estadísticos como *Statgraphics*. Otra alternativa viable es utilizar tablas donde se proporcionan generadores adecuados para diferentes valores de  $k$  y  $p$ , lo cual ya es bastante ayuda (tabla 8). Lo más práctico es utilizar un software estadístico para generar la fracción deseada y su estructura de alias.

Tabla 8. Factoriales fraccionados con resolución IV, con máximo 64 corridas.					
Factores	Diseño	Corridas	Alias	Grado de fracción	Generadores
4	$2^{4-1}_{IV}$	8	1	1/2	$D = \pm ABC$
5	$2^{5-1}_{IV}$	16	1	1/2	$E = \pm ABCD$
6	$2^{6-1}_{VI}$	32	1	1/2	$F = \pm ABCDE$
6	$2^{6-2}_{IV}$	16	3	1/4	$E = \pm ABC, F = \pm BCD$
7	$2^{7-1}_{IV}$	64	1	1/2	$G = \pm ABCDEF$
7	$2^{7-2}_{IV}$	32	3	1/4	$F = \pm ABCD, G = \pm ABDE$
7	$2^{7-3}_{IV}$	16	7	1/8	$E = \pm ABC, F = \pm BCD$ $G = \pm ACD$
8	$2^{8-2}_{IV}$	64	3	1/4	$G = \pm ABCD, H = \pm ABEF$
8	$2^{8-3}_{IV}$	32	7	1/8	$F = \pm ABC, G = \pm ABD$ $H = \pm BCDE$
8	$2^{8-4}_{IV}$	16	15	1/16	$E = \pm BCD, F = \pm ACD$ $G = \pm ABC, H = \pm ABD$
9	$2^{9-3}_{IV}$	64	7	1/8	$G = \pm ABCD, H = \pm ACEF$ $J = \pm CDEF$
9	$2^{9-4}_{IV}$	32	15	1/16	$F = \pm BCDE, G = \pm ACDE$ $H = \pm ABDE, J = \pm ABCE$
10	$2^{10-4}_{IV}$	64	15	1/16	$G = \pm BCDF, H = \pm ACDF$ $J = \pm ABDE, K = \pm ABCE$
10	$2^{10-5}_{IV}$	32	31	1/32	$F = \pm ABCD, G = \pm ABCE,$ $H = \pm ABDE, J = \pm ACDE$ $K = \pm BCDE$

**Ejemplo de una fracción  $2^{7-4}$ .** Supongamos que se quieren estudiar  $k = 7$  factores y sólo se tienen recursos para correr una octava parte del diseño  $2^7$  completo, por lo que se decide utilizar un diseño factorial fraccionado  $2^{7-3}$ . En la tabla 8 se lee que los generadores  $E = \pm ABC, F = \pm BCD$  y  $G = \pm ACD$  son adecuados. Se pueden construir ocho fracciones diferentes con estos generadores, dependiendo de los signos que se toman. Por facilidad vamos a construir la fracción principal, que es la que se obtiene con los tres generadores con signo positivo.

*Construcción del diseño.* En un primer paso se escribe el diseño  $2^{7-3}$  como si fuera un factorial completo  $2^4$ . En el segundo paso se calculan las columnas faltantes  $E, F$  y  $G$  con base en los generadores. El diseño de tabla 9 es el factorial fraccionado  $2^{7-3}$  deseado. Antes de ir al proceso para obtener las mediciones en orden aleatorio, debe analizarse la estructura alias para poder asignar los siete factores de manera

óptima a las columnas, como se discute mas adelante.

Tabla 9. Diseño $2^{7-3}$ , con generadores						
I=ABCE, I=BCDF e I=ACDG						
A	B	C	D	E=ABC	F=BCD	G=ACD
-	-	-	-	-	-	-
+	-	-	-	+	-	+
-	+	-	-	+	+	-
+	+	-	-	-	+	+
-	-	+	-	+	+	+
+	-	+	-	-	+	-
-	+	+	-	-	-	+
+	+	+	-	+	-	-
-	-	-	+	-	+	+
+	-	-	+	+	+	-
-	+	-	+	+	-	+
+	+	-	+	-	-	-
-	-	+	+	+	-	-
+	-	+	+	-	-	+
-	+	+	+	-	+	-
+	+	+	+	+	+	+

*Estructura de alias.* La estructura de alias completa del diseño se obtiene multiplicando cada efecto por la relación definidora, que está dada por los generadores iniciales y todos sus posibles productos:

$$I = ABCE = ABFG = ACDG = ADEF = BCDF = BDEG = CEFG$$

Así, la estructura alias completa generada con Statgraphics se muestra en la *tabla 10*. En la estructura de alias completa cada efecto tiene siete alias, es decir, los grupos de efectos alias tienen ocho elementos. La estructura de alias completa tiene poca o nula utilidad, salvo tener la idea de la situación que se presenta con los efectos. Para fines prácticos basta conocer la estructura de alias reducida que se muestra en la *tabla 11*.

De acuerdo a esta estructura de alias, habría que cuidar en este diseño que dos interacciones de interés especial para el experimentador no aparezcan alias entre ellas. Esto se puede lograr si se asignan los factores a las letras de las columnas (tabla 9), de manera que las interacciones de interés aparezcan en grupos de alias diferentes. Por ejemplo, supongamos que cuatro de los siete factores son: temperatura, humedad, velocidad y ángulo, y es de interés poder estimar todas las interacciones dobles entre estos factores; supongamos además que de los otros tres factores sólo interesan los efectos principales. Se asignan la letra A al factor temperatura, B al factor humedad, C al factor velocidad y D al factor ángulo (véase tabla 11), se podrán estimar las seis interacciones de interés. En contraparte, si los primeros tres factores se asignan de igual manera pero el factor ángulo se asigna a la columna E, se tendrán varias interacciones de interés alias entre ellas, como son:  $AB + CE$ ,  $AC + BE$  y  $AE + BC$ . Hay que observar que un pequeño cambio en la asignación de los factores a las columnas del diseño, cambia radicalmente la estructura de alias en cuanto a los efectos de interés.



Note que si todas las interacciones de dos factores fueran igual de importantes, este diseño no hace un buen trabajo, y debe utilizarse otro con resolución al menos de V. Pero si de antemano se tienen identificadas sólo algunas de las interacciones como las potencialmente importantes, se puede buscar la manera de estudiarlas con una fracción de resolución IV.

Tabla 10. Estructura de alias completa del diseño $2_{IV}^{7-3}$
A + BCE + BFG + CDG + DEF + ABCDF + ABDEG + ACEFG
B + ACE + AFG + CDF + DEF + ABCDG + ABDEF + BCEFG
C + ABE + ADG + BDF + EFG + ABCFG + ACDEF + BCDEG
D + ACG + AEF + BCF + BEG + ABCDE + ABDFG + CDEFG
E + ABC + ADF + BDG + CFG + ABEFG + ACDEG + BCDEF
F + ABG + ADE + BCD + CEG + ABCEF + ACDFG + BDEFG
G + ABF + ACD + BDE + CEF + ABCEG + ADEFG + BCDFG
AB + CE + FG + ACDF + ADEG + BCDG + BDEF + ABCEFG
AC + BE + DG + ABDF + AEFG + BCFG + CDEF + ABCDEG
AD + CG + EF + ABCF + ABEG + BCDE + BDFG + ACDEFG
AE + BC + DF + ABDG + ACFG + BEFG + CDEG + ABCDEF
AF + BG + DE + ABCD + ACEG + BCEF + CDFG + ABDEFG
AG + BF + CD + ABDE + ACEF + BCEG + DEFG + ABCDFG
BD + CF + EG + ABCG + ABEF + ACDE + ADFG + BCDEFG
ABD + ACF + AEG + BCG + BEF + CDE + DFG + ABCDEFG
$\mu$ + ABCE + ABFG + ACDG + ADEF + BCEF + BDEF + CEFG

Tabla 11. Estructura de alias reducida del diseño  $2_{IV}^{7-3}$

Los efectos principales: A, B, C, D, E, F y G se estiman limpiamente, pero las iteraciones dobles se confunden entre ellas:
AB + CE + FG
AC + BE + DG
AD + CG + EF
AE + BC + DF
AF + BG + DE
AG + BF + CD
BD + CF + EG

**7.1. Estimación de efectos y sumas de cuadrados.** Los efectos y las sumas de cuadrados en los diseños factoriales fraccionados  $2^{k-p}$  se obtienen a partir de los contrastes, de manera similar a como se hace con los factoriales completos  $2^k$ . Se obtiene un contraste para cada grupo de efectos alias y se pondera por una constante apropiada para estimar el efecto correspondiente como una diferencia de medias. Así, el efecto de un grupo de efectos alias  $X$  se estima como

$$\text{Efecto } X = \frac{\text{Contraste } X}{2^{k-p-1}}$$

y su correspondiente suma de cuadrados es

$$SCX = \frac{(\text{Contraste } X)^2}{2^{k-p}}$$

la cual tiene 1 grado de libertad. Calculando la suma de cuadrados totales (SCT.) y la suma de cuadrados del error (SCE), compuesta por efectos pequeños excluidos

con base en el Pareto efectos y el gráfico de Daniel, se puede construir un ANOVA razonable. Cabe señalar que para eliminar un efecto del análisis se deben eliminar también todos sus alias.

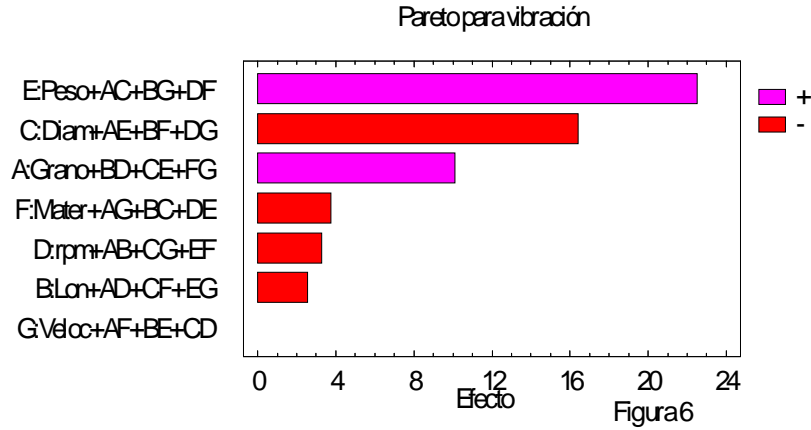
**EJEMPLO 15.** *Un ejemplo ilustrativo de un diseño de experimento fraccionado  $2^{7-4}$ . Una compañía había comprado un equipo de afilado, y después de una semana se da cuenta que el exceso de vibración en el proceso era un problema muy serio. Un equipo de mejora que involucra a los ingenieros responsables del proceso deciden tratar de reducir la vibración aplicando diseño de experimentos. El equipo identifica siete factores, todos asociados, con la herramienta de afilado, que pueden tener algo que ver en la cantidad de vibración: diámetro, longitud, tamaño de grano del material, velocidad de alimentación, revoluciones por minuto, estructura del material y peso de la precarga. Se seleccionan dos niveles para cada factor, en los cuales llevar a cabo el experimento. Como la puesta en marcha del experimento y el tiempo de corrida tienen un alto costo, el equipo decide utilizar un experimento de ocho corridas. Uno de los miembros del equipo de mejora que había trabajado antes con máquinas similares asegura que los efectos de interacción entre los factores se pueden considerar despreciables, lo cual hace factible el uso de una fracción que sólo permite el estudio de los efectos principales. Los factores, así como los niveles utilizados en unidades originales, se muestran en la tabla 12.*

Bajo el supuesto de que los efectos de interacción se pueden considerar despreciables, se decide correr un diseño altamente fraccionado (saturado), como lo es la fracción  $2^{7-4}_{III}$  que se muestra en la tabla 14. El diseño tiene resolución  $III$ , que implica efectos principales confundidos con interacciones dobles. La estructura de alias reducida se muestra en la tabla 13.

En la tabla 14 se muestra la matriz de diseño ( $2^{7-4}_{III}$ ), así como la vibración observada en el proceso en cada tratamiento.

Tabla 12. Factores y niveles utilizado, problema de vibración			Tabla 13. Estructura de alias reducida del diseño $2^{7-4}_{III}$	
Factor	Descripción (Unidades)	Niveles (bajo alto)		
C: diam	diámetro (pulgadas)	1.0, 1.5	A+BD+CE+FG B+AD+CF+EG C+AE+BF+DG D+AB+CG+EF E+AC+BG+DF F+AG+BC+DE G+AF+BE+CD	
B: long	longitud (pulgadas)	1.0, 2.0		
A: grano	tamaño de grano (/pulgada)	80,120		
G: alim	velocidad de alimentación (pulg/min)	2.0, 4.0		
D: rpm	rpm (x 1000)	15,20		
E: peso	peso de precarga (libras)	1.0, 4.0		
F: matest	estructura del material (onzas)	1.0, 4.0		

Tabla 14. Matriz de diseño y vibración observada							
Grano	Long.	Diam	RPM	Precar	Matest	Alim.	Vibra
-1.0	-1.0	-1.0	1.0	1.0	1.0	-1.0	77.4
1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	68.3
-1.0	1.0	-1.0	-1.0	1.0	-1.0	1.0	81.9
1.0	1.0	-1.0	1.0	-1.0	-1.0	-1.0	66.2
-1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	42.1
1.0	-1.0	1.0	-1.0	1.0	-1.0	-1.0	78.3
-1.0	1.0	1.0	-1.0	-1.0	1.0	-1.0	39.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	68.4

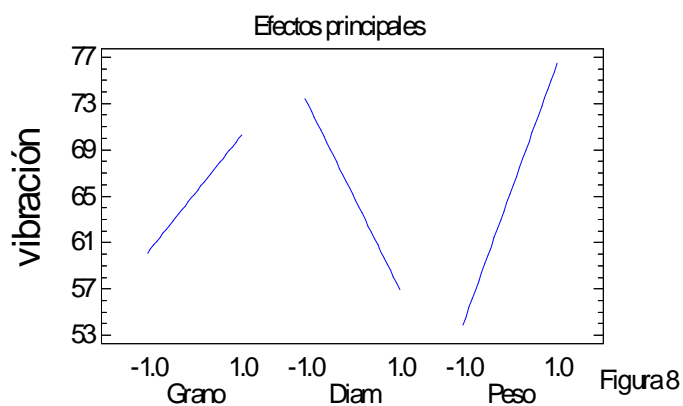
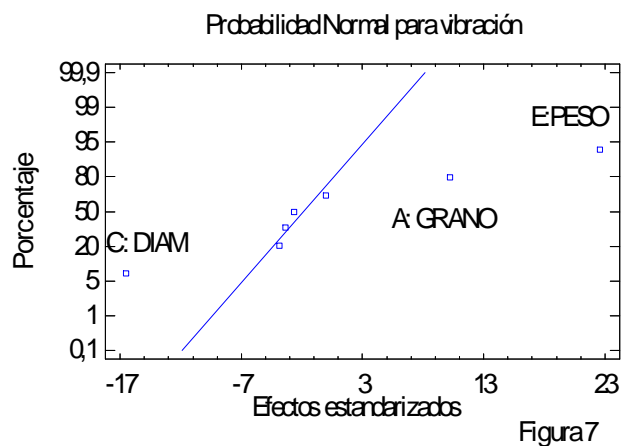


*Análisis del experimento.* El experimento tiene un total de siete grados de libertad, que se gastan en estimar sólo a los efectos principales, y quedan cero grados de libertad para el error en el ANOVA. De aquí la necesidad de recurrir al diagrama de Pareto y al gráfico de Daniel como paso previo antes de intentar un análisis de varianza; de hecho el ANOVA directo no se puede completar al no haber cuadrado medio del error.

El diagrama de Pareto de efectos hace un trabajo excelente al detectar tres efectos importantes y cuatro despreciables (véase figura 6). Esto es confirmado por el gráfico de Daniel de la figura 7, que muestra cuatro efectos sobre la línea (que no afectan) y tres efectos separados de la línea, que corresponden a las tres barras del Pareto. Con esta información se puede intentar un análisis de varianza, con inclusión del análisis a los cuatro efectos alineados. El resultado se muestra en la tabla 15. El estadístico  $R^2_{ajustado}$  que mide el grado de explicación de la variable de respuesta por el modelo es 93.71%. Aunque la suma de cuadrados del error tiene sólo cuatro grados de libertad, en este caso, parece que el cuadrado medio del error está razonablemente bien estimado.

Tabla 15. Mejor ANOVA					
<i>FV</i>	SC	G.L	C.M	$F_0$	<i>valor - p</i>
<i>A : grano</i>	208.08	1	208.08	12.58	0.0239
<i>C : diám.</i>	544.5	1	544.5	32.93	0.0046
<i>E : peso</i>	1021.52	1	1021.52	61.78	0.0014
<i>Error</i>	66.14	4	16.535		
<i>Total</i>	1840.24	7			

*Interpretación.* El objetivo del estudio es encontrar una combinación de los niveles de los factores que minimice el problema de la vibración de la máquina de afilado. En la figura 8 se muestran los efectos que están activos, y para cada uno se localiza el valor más bajo de la línea. De esta forma se aprecia que a mayor tamaño de precarga mayor vibración, a más diámetro menor vibración y a mayor tamaño



de grano más vibración. Por tanto, el mejor tratamiento es: *grano* en su nivel bajo, *diametro* en su nivel alto y *peso* en su nivel bajo; los niveles de los factores restantes se eligen con el criterio de economía (nivel más barato).

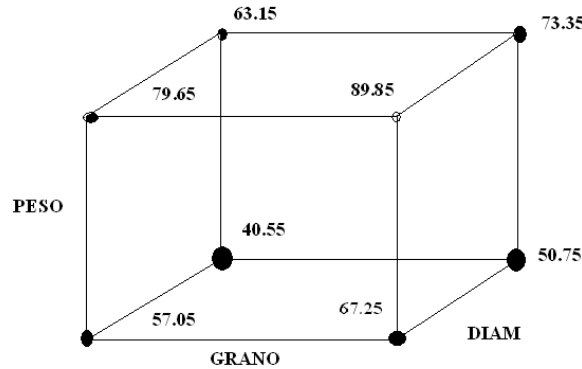
Como sólo hay tres factores activos tiene sentido dibujar la gráfica de cubo, que resume bien el comportamiento de la vibración predicha (véase *figura 9*). Se observa que la vibración mínima (40.55) ocurre en la combinación.

$(A : GRANO^{-}, C : DIAM^{+}, E : PESO^{-})$ .

### 8. Ejercicios: Diseños factoriales fraccionados $2^{k-p}$

EJERCICIO 104. ¿Qué es una diseño factorial fraccionado y cuándo se recomienda aplicarlo?

Figura 9



EJERCICIO 105. ¿Por qué se dice que un diseño factorial completo con muchos factores (cinco o más) genera un exceso de información?

EJERCICIO 106. Con el uso del concepto de resolución de un diseño, explique qué se pierde al correr diseños factoriales fraccionados.

EJERCICIO 107. Considere un diseño factorial fraccionado  $2^{5-1}$ . (a) ¿Cuántos factores se controlan? (b) En este diseño sólo se corren 16 de 32 posibles tratamientos, ¿qué se pierde y qué se gana al correr sólo la mitad? (c) ¿A partir de los 32 tratamientos posibles, se puede seleccionar cualquiera hasta completar 16 o cuál es el criterio de selección? (d) Elijiendo como generador a  $I = ABCDE$ , obtenga la matriz de diseño (los tratamientos que constituyen este diseño). (e) ¿En qué orden deben correrse los 16 tratamientos del diseño?

EJERCICIO 108. Respecto a un diseño factorial  $2^4$ : (a) Obtenga las dos posibles fracciones a la mitad para este diseño. (b) ¿Qué resolución tienen estas fracciones y por qué? (c) ¿En qué condiciones se recomienda correr cada fracción? (d) Obtenga la estructura de alias para ambas fracciones.

EJERCICIO 109. Considere un diseño factorial fraccionado  $2^{8-4}$ . (a) ¿Cuántos factores se controlan y cuántas corridas del proceso implica? (b) ¿Cuántos generadores independientes tiene? ¿Cuántas palabras o generadores tiene la relación definidora? (c) ¿Cuál es la resolución del diseño? ¿Qué implica esta resolución?

EJERCICIO 110. ¿Cuándo se dice que una fracción factorial es saturada?

EJERCICIO 111. Construya un diseño factorial fraccionado  $2_{III}^{7-4}$ . (a) ¿Qué significa que el diseño sea de resolución III, en términos prácticos? (b) Observe su estructura de alias completa. ¿Cuántos alias tiene cada efecto? (c) Describa una situación en la que es pertinente utilizar este diseño.

**EJERCICIO 112.** A continuación se muestran los tratamientos y los resultados obtenidos en un diseño factorial  $2^{5-1}$

(1) = 700	cd = 2031	de = 2515	ce = 580
ad = 2507	ac = 2247	ae = 1317	acde = 2314
be = 468	bc = 446	bd = 2247	bcde = 2262
ab = 424	abce = 468	abde = 2232	abcd = 2299

(a) ¿Cuál es el generador de esta fracción factorial? (b) ¿Cuál es la resolución de este diseño? ¿Qué significa? (c) Obtenga un diagrama de Pareto y la gráfica de Daniel para lo, efectos. ¿Cuáles efectos parecen ser activos? (d) Obtenga el mejor ANOVA. (e) Obtenga la gráfica de los efectos activos en el mejor ANOVA, e interprételos con detalle. (f) Si lo que se quiere es maximizar, ¿cuál es el mejor tratamiento?

**EJERCICIO 113.** Respecto al problema anterior: (a) ¿Puede colapsar el diseño factorial fraccionado en un factor completo? Argumente. (b) Elimine los factores que menos impacto tuvieron, colapse diseño  $2^{5-1}$  en un diseño  $2^3$ , ¿cuántas réplicas tiene el diseño colapsado? (c) Analice con detalle el diseño colapsado y obtenga conclusiones. (d) ¿Las conclusiones obtenidas con el análisis del diseño fraccionado y el colapsado son las mismas?

**EJERCICIO 114.** Una organización de manufactura produce partes de plástico con moldeo por inyección. Típicamente, las partes se encogen de manera excesiva, lo cual causa problemas en el ensamble posterior. Los ingenieros de calidad deciden utilizar diseño de experimentos para reducir el promedio de encogimiento de las partes y se espera reducir también la variabilidad de corrida a corrida. Se identifican siete factores para el estudio. Cuatro de esos factores son fácilmente controlables; temperatura del molde (A), velocidad del tornillo (B), tiempo de permanencia (C) y tamaño de compuerta (D). Tres de las variables no son fáciles de controlar durante la manufactura normal: tiempo de ciclo (E), contenido de mezcla (F) y presión en el molde (G), pero durante el experimento se controlarán. Debido a potenciales efectos no lineales en la relación entre el encogimiento y las variables de proceso, se deben considerar al menos tres niveles en cada factor. La alternativa que se considera es utilizar una fracción  $2^{7-3}$  con generadores  $I = ABCE$ ,  $I = BCDF$  e  $I = ACDG$ , que es una octava parte del factorial completo  $2^7$ . Se incluyen cuatro repeticiones al centro para poder detectar la posible presencia de curvatura. Los datos en unidades de encogimiento  $\times 10$  se muestran en seguida.

(1) = 6	acf = 15	bdeg = 34	abcdefg = 52
aeg = 10	bcb = 26	abd = 60	(0) = 25
bef = 32	abce = 60	cde = 16	(0) = 29
abfg = 60	dfg = 8	acd = 5	(0) = 24
cefg = 4	adef = 12	bcd = 37	(0) = 27

(a) ¿Cuál es la estructura de alias del diseño? (b) Encuentre el mejor ANOVA para estos datos. No olvide verificar la presencia de curvatura. (c) Proyecte el diseño en uno más simple si hay factores que no afectan, y haga el análisis. (d) Interprete con detalle los efectos activos y determine el mejor tratamiento para reducir el encogimiento promedio. (e) Determine el mejor tratamiento considerando que también interesa reducir la variabilidad.

EJERCICIO 115. En una empresa panificadora existen problemas con la simetría y el color del pan integral. Los responsables del proceso sospechan que el problema se origina desde la fase de fermentación. En ésta se combina agua, harina, cierta cantidad de levadura más una serie de ingredientes como fosfato, sal, etc. Al final de la fermentación se obtiene lo que llaman "esponja líquida" la cual debe cumplir una serie de parámetros de calidad: una acidez total titulable (ATT) mayor a 6.0 y un pH mayor a 4.8. Sin embargo, no se ha cumplido con dichas exigencias de calidad; se han hecho algunos intentos experimentando con un factor a la vez, pero los resultados han sido malos. En busca de una mejor estrategia experimental, se decide utilizar un diseño factorial fraccionado  $2^{6-2}$  para investigar el efecto de seis factores en las variables ATT y pH. Los primeros cinco factores se refieren a cierta cantidad que se agrega en la fermentación: A: levadura (17, 19), B: sal (2.5, 3.7), C: fosfato (2.0, 3.6), D: sulfato (1.5, 2.2) y E: cloruro (0.89, 1.20); el sexto factor es F: temperatura inicial del agua (22, 26). Los datos obtenidos se muestran en la tabla siguiente.

Trat	ATT	pH	Trat	ATT	pH	Trat	ATT	pH	Trat	ATT	pH
(1)	6.2	4.86	cef	5.7	4.94	df	5.3	4.81	cde	6.9	4.84
ae	5.6	4.86	acf	6.4	4.74	adef	6.6	4.81	acd	7.1	4.85
bef	5.8	4.85	bc	6.4	4.83	bde	5.2	4.98	bcd	6.7	4.96
abf	5.8	4.99	abce	6.6	4.85	abd	5.5	4.98	abcde	6.9	4.84

(a) Observe los datos con cuidado, sobretudo los correspondientes al pH, ¿qué observa de destacado? ¿A qué puede deberse eso? (b) ¿Cuáles la resolución de este diseño y qué significa ésta? Escriba la estructura alias reducida. ¿Cómo se pueden interpretar los efectos de interacción que son alias? (c) ¿Cuáles efectos explican el comportamiento de cada una de las respuestas? Encuentre el mejor ANOVA para cada respuesta e interprete utilizando  $\alpha = 0.05$ . (d) Determine las condiciones de operación que maximizan a ambas respuestas simultáneamente. ¿Es posible dar una solución simultánea al problema con los análisis individuales? Argumente. (e) Verifique los supuestos para cada variable de respuesta.

EJERCICIO 116. Considere un experimento  $2^{5-1}$  con  $I = ABCDE$  que fue utilizado para investigar los efectos de cinco factores sobre el color de un producto químico. Los factores son A = solvente /reactante, B = catalizador/ reactante, C = temperatura, D = pureza del reactante y E = acidez del reactante. Los resultados obtenidos son los siguientes:

e = -0.63	c = 2.06	d = 6.79	cde = 5.22
a = 2.51	ace = 1.22	ade = 5.47	acd = 4.38
b = -2.68	bce = -2.09	bde = 3.45	bcd = 4.30
abe = 1.66	abc = 1.93	abd = 5.68	abcde = 4.05

(a) Calcule los efectos y grafíquelos en Pareto y en papel normal. ¿Cuáles parecen significativos? (b) Obtenga el mejor análisis de varianza. ¿Con cuáles efectos se está construyendo el error? (c) Represente gráficamente cada efecto significativo e interprete con detalle. (d) Determine el mejor tratamiento y la respuesta predicha por el modelo. (e) Haga el análisis de residuos y comente los resultados. (f) Si hay algún factor que no tiene ningún efecto, colapse el diseño. ¿Qué diseño resultó?





## Metodología de superficie de respuesta

### 1. Competencias

*Comprender el concepto de optimización y su relación con la superficie de respuesta. Aplicar la metodología de superficie de respuesta y sus respectivos diseños y modelos. Conocer las técnicas de optimización y aplicarlas adecuadamente.*

### 2. Introducción a la Metodología de Superficie de Respuesta (MSR)

En algunos experimentos no se obtienen las respuestas buscadas o el nivel de mejoras logrado no es suficiente, por lo que es necesario experimentar de manera secuencial hasta encontrar el nivel de mejoras deseado. Después de una primera etapa experimental quizás sea necesario desplazar la región experimental (moverse de lugar) en una dirección adecuada, o bien, explorar en forma más detallada la región experimental inicial, este proceso forma parte de lo que se conoce como *la metodología de superficies de respuesta*, o **MSR**.

**DEFINICIÓN 2.** *La metodología de superficies de respuesta, o **MSR**, es una colección de técnicas matemáticas y estadísticas útiles en el modelado y el análisis de problemas en los que una respuesta de interés recibe la influencia de diversas variables y donde el objetivo es optimizar esta respuesta. Por ejemplo, suponga que un ingeniero quiere encontrar los niveles de temperatura ( $x_1$ ) y presión ( $x_2$ ) que maximicen el rendimiento ( $y$ ) de un proceso, el rendimiento del proceso es una función de los niveles de la temperatura y la presión:*

$$y = f(x_1, x_2) + \varepsilon$$

*donde  $\varepsilon$  representa el ruido o error observado en la respuesta  $y$ . Si la respuesta esperada se denota por*

$$E(y) = f(x_1, x_2) = \eta,$$

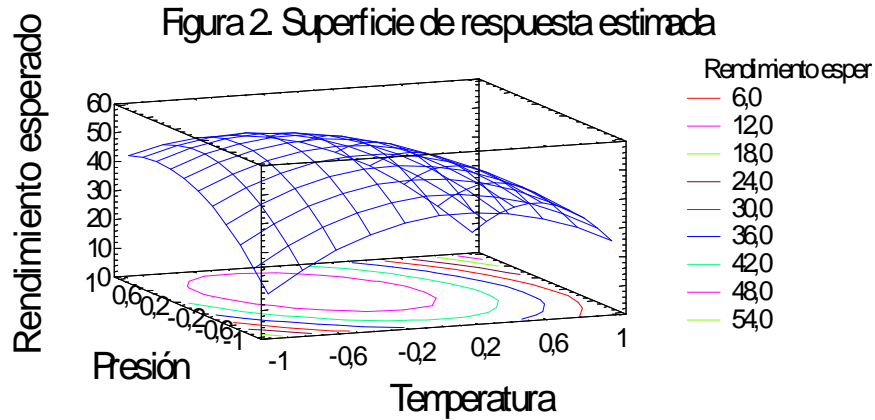
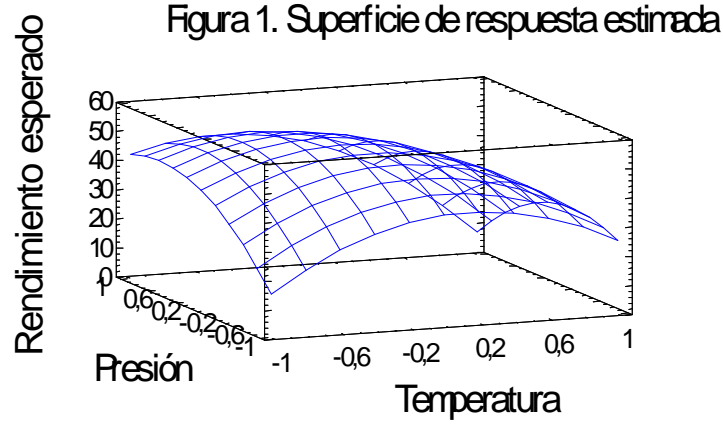
*entonces a la superficie representada por*

$$\eta = f(x_1, x_2)$$

*se llama superficie de respuesta.*

Por lo general la superficie de respuesta se representa gráficamente como en la figura 1, donde  $\eta$  se grafica contra los niveles de  $x_1$  y  $x_2$ . Para ayudar a visualizar la forma de una superficie de respuesta, con frecuencia se grafican los contornos de la superficie de respuesta, como se muestra en la figura 2. En la gráfica de contorno se trazan las líneas de respuesta constante en el plano  $x_1, x_2$ . Cada contorno corresponde a una altura particular de la superficie de respuesta.

*En la mayoría de los problemas **MSR**, la forma de la relación entre la respuesta y las variables independientes es desconocida; un primer paso de la **MSR** es encontrar una aproximación adecuada de la verdadera relación funcional entre  $y$*



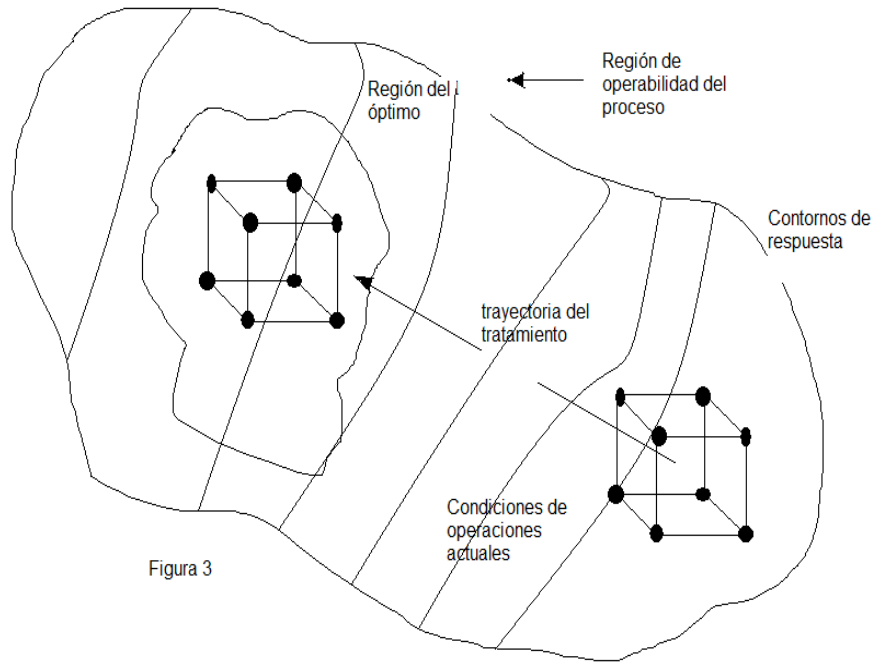
y el conjunto de variables independientes y por lo general, se emplea un polinomio de orden inferior en alguna región de las variables independientes. Si la respuesta está bien modelada por una función lineal de las variables independientes, entonces la función de aproximación es el *modelo de primer orden*

$$(2.1) \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

Si hay curvatura en el sistema, entonces debe usarse un polinomio de orden superior, tal como el *modelo de segundo orden*

$$(2.2) \quad y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \varepsilon$$

- El método de mínimos cuadrados, se usa para estimar los parámetros de los polinomios de aproximación, después se realiza el análisis de la superficie de respuesta utilizando la superficie ajustada.



- *La MSR es un procedimiento secuencial.* Cuando se está en un punto de la superficie de respuesta que está apartado del óptimo, como en el caso de las condiciones de operación actuales, el sistema presenta una curvatura moderada y el modelo de primer orden será apropiado. El objetivo en este caso es llevar al experimentador de manera rápida y eficiente por la trayectoria del mejoramiento hasta la vecindad general del óptimo. Una vez que se ha encontrado la región del óptimo, puede emplearse un modelo más elaborado, como el de segundo orden, y llevarse a cabo un análisis para localizar el óptimo. El análisis de una superficie de respuesta puede considerarse como “*el ascenso a una colina*”, donde la cima de ésta representa el punto de la respuesta máxima. Si el verdadero óptimo es un punto de respuesta mínima, entonces la situación puede considerarse como “*el descenso a un valle*”. (ver figura 3)
- *El objetivo último de la MSR es determinar las condiciones de operación óptimas del sistema o determinar una región del espacio de los factores en la que se satisfagan los requerimientos de operación.*

### 3. Método del ascenso más pronunciado

Cuando la estimación inicial de las condiciones de operación óptimas del sistema están lejos del *óptimo real*, el objetivo del experimentador es pasar con rapidez a la vecindad general del óptimo. *Cuando se está muy lejos del óptimo, por lo general se supone que un modelo de primer orden es una aproximación adecuada de la verdadera superficie en una región pequeña de las  $x$ .*

*El método del ascenso más pronunciado* es un procedimiento para moverse secuencialmente sobre la trayectoria del ascenso más pronunciado, es decir, en la

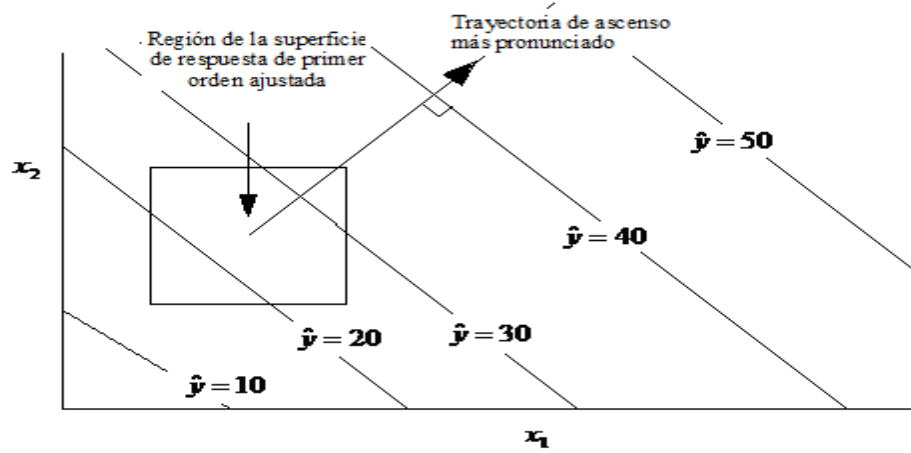


Figura 4

dirección del incremento máximo de la respuesta. Desde luego, si lo que se pretende es una minimización, entonces esta técnica se llama *método del descenso más pronunciado*. El modelo ajustado de primer orden es:

$$(3.1) \quad \hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i$$

y la superficie de respuesta de primer orden, es decir, los contornos de  $\hat{y}$ , es una serie de líneas paralelas como las que se muestran en la figura 4. La dirección del ascenso más pronunciado es aquella en la que  $\hat{y}$  se incrementa con mayor rapidez. Esta dirección es paralela a la normal de la *superficie de respuesta ajustada*. Por lo general, se toma como la *trayectoria del ascenso más pronunciado* a la recta que pasa por el centro de la región de interés y que es normal a la superficie ajustada. Por lo tanto, *los pasos sobre la trayectoria son proporcionales a los coeficientes de regresión*  $\{\hat{\beta}_i\}$ . El tamaño real del paso lo determina el experimentador con base en el conocimiento del proceso o de otras consideraciones prácticas.

- Se conducen experimentos sobre la trayectoria del ascenso más pronunciado hasta que deja de observarse un incremento adicional en la respuesta, luego, puede ajustarse un nuevo modelo de primer orden, determinarse una nueva trayectoria del ascenso más pronunciado y el procedimiento continúa. En última instancia, el experimentador llegará a la vecindad del óptimo. En general, la falta de ajuste del modelo de primer orden indica que se ha llegado a ella. En este momento se realizan experimentos adicionales para obtener una estimación más precisa del óptimo.

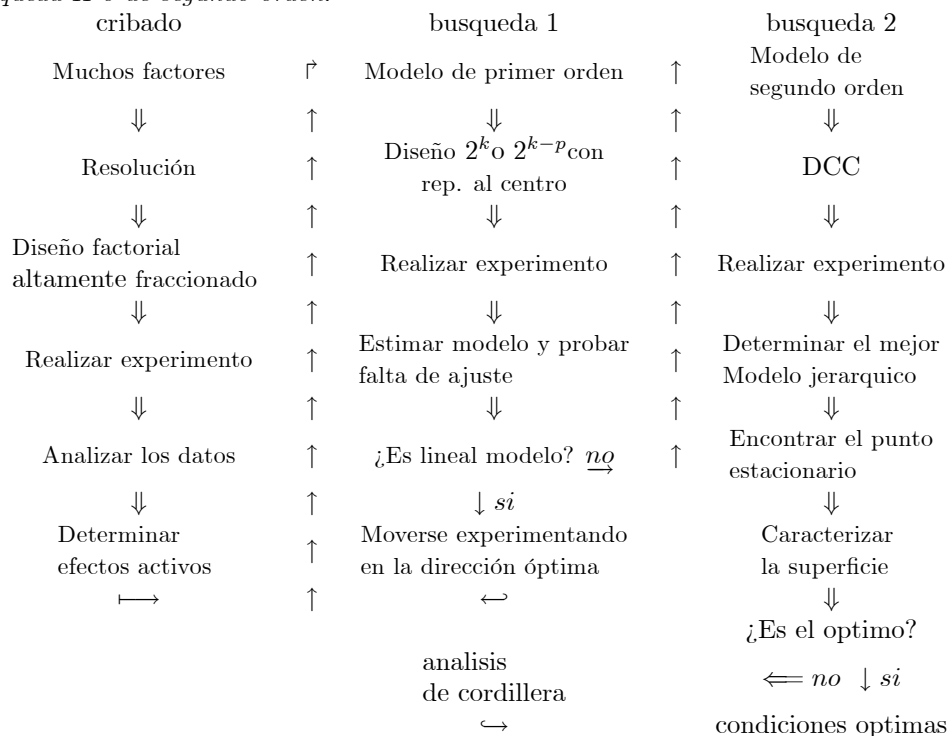
#### 4. Esquema de los elementos de la MSR en su contexto amplio

La metodología de superficie de respuesta implica tres aspectos: *diseño*, *modelo* y *técnica de optimización*. El diseño y el modelo se piensan al mismo tiempo, y

dependen del tipo de comportamiento que se espera en la respuesta. De manera específica, el modelo puede ser de primero o segundo orden (plano o con curvatura); por ello, el tipo de diseño utilizado y el método de optimización se clasifican, según sea el caso, como de primero o segundo orden.

El aspecto *diseño* implica que para optimizar un proceso se debe aplicar el diseño de experimentos, en particular aquellos que sirven para ajustar un modelo de regresión lineal múltiple. Más adelante se presentan algunos de estos diseños, conocidos genéricamente como diseños para superficie de respuesta. El aspecto del modelo utiliza el análisis de regresión lineal múltiple, junto con sus elementos básicos que son: parámetros del modelo, modelo ajustado, significancia del modelo, prueba de falta de ajuste, residuos, predichos, intervalos de confianza para predichos y coeficiente de determinación.

Por último, el aspecto de optimización está formado por algunas técnicas matemáticas que sirven para que dado un modelo ajustado explorarlo a fin de obtener información sobre el punto óptimo. Conviene recordar técnicas como: derivadas de funciones, multiplicadores de lagrange, operaciones con matrices, valores y vectores propios y sistemas de ecuaciones simultáneas. A continuación se presenta un esquema de la metodología de superficie de respuesta, donde se distinguen tres etapas en la búsqueda del punto óptimo, que son: *cribado*, *búsqueda I o de primer orden* y *búsqueda II o de segundo orden*.



A continuación se describe brevemente cada una de estas etapas,

- (1) *Cribado*. La optimización de un proceso se inicia con esta etapa cuando tiene muchos factores (más de 6 u 8) que influyen en la variable de interés. Por ejemplo, pensemos en una máquina que se puede manipular en 10 parámetros diferentes y que no se tiene una idea clara de cómo influye

cada uno de ellos; en primer lugar es preciso correr un experimento para identificar los pocos factores que tienen mayor influencia.

- (2) *Búsqueda I o de primer orden.* Esta etapa se aplica cuando se tienen pocos factores ( $k \leq 5$ ), y se sabe que éstos influyen en la variable de respuesta. En esta etapa se corre un diseño de primer orden que permita caracterizar en forma preliminar el tipo de superficie de respuesta y detectar la presencia de curvatura. Por lo general se utiliza un diseño factorial completo o fraccionado con repeticiones al centro.
- (3) *Búsqueda II o de segundo orden.* En el momento en que se detecta la presencia de curvatura, o bien, que la superficie es más complicada que un hiperplano, se corre o se completa un diseño de segundo orden para caracterizar mejor la superficie y modelar la curvatura. Con el modelo ajustado se determinan las condiciones óptimas de operación del proceso.

**4.1. Ejemplo Ilustrativo.** Un ingeniero químico está interesado en determinar las condiciones de operación que maximizan el rendimiento de un proceso. Dos variables controlables influyen en el rendimiento del proceso: *tiempo de reacción* y *temperatura de reacción*. El ingeniero opera actualmente el proceso con un tiempo de reacción de 35 minutos y una temperatura de  $155^\circ F$ , que dan como resultado rendimientos de cerca de 40%. Puesto que es improbable que esta región contenga el óptimo, el ingeniero ajusta un modelo de *primer orden* y aplica *el método del ascenso más pronunciado*. El ingeniero decide que la región de exploración para ajustar el modelo de primer orden deberá ser (30, 40) minutos de tiempo de reacción y  $(150, 160)^\circ F$ . Para simplificar los cálculos, las variables independientes se codificarán en el intervalo usual  $(-1, 1)$ . Por lo tanto, si  $\xi_1$  denota la variable natural tiempo,  $\xi_2$  la variable natural temperatura, entonces las variables codificadas son

$$x_1 = \frac{\xi_1 - 35}{5} \text{ y } x_2 = \frac{\xi_2 - 155}{5}$$

El diseño experimental se muestra en la tabla 1. Observe que el diseño usado para recabar estos datos es un factorial  $2^2$  aumentado con cinco puntos centrales. Las réplicas del centro se usan para estimar el error experimental y permitir la verificación de la adecuación del modelo de primer orden. Además, el está centrado alrededor de las condiciones de operación actuales del proceso.

Aplicando los métodos para diseños de dos niveles se obtiene el siguiente modelo en las variables codificadas:  $\hat{y} = 40.44 + 0.775x_1 + 0.325x_2$

Tabla 1.				
<i>Variables naturales</i>		<i>Variables codificadas</i>		respuesta
$\xi_1$	$\xi_2$	$x_1$	$x_2$	
30	150	-1	-1	39.3
30	160	-1	1	40.0
40	150	1	-1	40.9
40	160	1	1	41.5
35	155	0	0	40.3
35	155	0	0	40.5
35	155	0	0	40.7
35	155	0	0	40.2
35	155	0	0	40.6

**Análisis de la Varianza para Rendimiento**

Fuente	Suma de Cuadrados	Gl	Cuadrado medio	F-Ratio	P-Valor
A:Tiempo	2,4025	1	2,4025	55,87	0,0017
B:Temperatura	0,4225	1	0,4225	9,83	0,0350
AB	0,0025	1	0,0025	0,06	0,8213
Falta de ajuste	0,00272222	1	0,00272222	0,06	0,8137
Error Puro	0,172	4	0,043		
Total (corr.)	3,00222	8			

R-cuadrado = 94,1802 por ciento

R-cuadrado (ajustado para g.l.) = 90,6884 por ciento

Error Estándar de Est. = 0,207364

Error absoluto de la media = 0,103704

Estadístico Durbin-Watson = 2,89388 (P=0,1564)

Autocorrelación residual Lag 1 = -0,517267

**Coef. de regresión para Rendimiento**

constante	= 40,4444
A:Tiempo	= 0,775
B:Temperatura	= 0,325
AB	= -0,025

|  
.  
.  
.

$$\text{Rendimiento} = 40,4444 + 0,775 \cdot \text{Tiempo} + 0,325 \cdot \text{Temperatura} - 0,025 \cdot \text{Tiempo} \cdot \text{Temperatura}$$

Antes de explorar a lo largo de la trayectoria del ascenso más pronunciado, deberá investigarse la adecuación del modelo de primer orden. *El diseño 2<sup>2</sup> con puntos centrales permite al experimentador*

- Obtener una estimación del error.
- Verificar las interacciones (o términos de productos cruzados) del modelo.
- Verificar los efectos cuadráticos (curvatura).

Las réplicas del centro pueden usarse para calcular una estimación del error de la siguiente manera:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\text{error puro}}{g.l} = \frac{40.3^2 + 40.5^2 + 40.7^2 + 40.2^2 + 40.6^2 - 202.3^2/5}{4} \\ &= 0.043\end{aligned}$$

En el modelo de primer orden se supone que las variables  $x_1$  y  $x_2$  tienen un efecto aditivo sobre la respuesta. La interacción entre las variables se representaría por el coeficiente  $\beta_{12}$  del término de un producto cruzado  $x_1x_2$  sumado al modelo. La estimación de mínimos cuadrados de este coeficiente es simplemente la mitad del efecto de la interacción que se calcula como en un diseño factorial  $2^2$  ordinario, o

$$\begin{aligned}\hat{\beta}_{12} &= \frac{1}{4}[(1 \times 39.3) + (1 \times 41.5) + (-1 \times 40.0) + (-1 \times 40.9)] \\ &= \frac{1}{4}(-0.1) = -0.025\end{aligned}$$

La suma de cuadrados de la interacción con un solo grado de libertad es

$$SC_{\text{Interacción}} = \frac{1}{4}(-0.1)^2 = 0.0025$$

Al comparar  $SS_{\text{Interacción}}$  con  $\hat{\sigma}^2$  se obtiene el estadístico para la *interacción*

$$\begin{aligned}F &= \frac{SC_{\text{Interacción}}}{\hat{\sigma}^2} \\ &= \frac{0.0025}{0.043} = 0.058\end{aligned}$$

que es pequeño, lo cual indica que la interacción es insignificante.

Otra verificación de la adecuación del modelo de línea recta se obtiene aplicando la verificación del efecto de *curvatura cuadrática pura*. Recuerde que ésta consiste en comparar la respuesta promedio en los cuatro puntos de la porción factorial del diseño, por ejemplo  $\bar{y}_F = 40.425$ , con la respuesta promedio en el centro del diseño, por ejemplo  $\bar{y}_C = 40.46$ . Si existe curvatura cuadrática en la verdadera función de la respuesta, entonces  $\bar{y}_F - \bar{y}_C$  es una medida de esta curvatura. Si  $\beta_{11}$  y  $\beta_{22}$  son los coeficientes de los términos “cuadráticos puros”  $x_1^2$  y  $x_2^2$ , entonces  $\bar{y}_F - \bar{y}_C$  es una estimación de  $\beta_{11} + \beta_{22}$ . En el ejemplo tratado aquí, una estimación del término cuadrático puro es:

$$\begin{aligned}\hat{\beta}_{11} + \hat{\beta}_{22} &= \bar{y}_F - \bar{y}_C \\ &= 40.425 - 40.46 = -0.035\end{aligned}$$

La suma de cuadrados con un solo grado de libertad asociada con la hipótesis nula,  $H_0 : \beta_{11} + \beta_{22} = 0$ ,

$$\begin{aligned}SC_{\text{cuadrática pura}} &= \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} \\ &= \frac{(4)(5)(-0.035)^2}{4 + 5} = 0.0027\end{aligned}$$



donde  $n_F$  y  $n_C$  son el número de puntos de la porción factorial y el número de puntos centrales, respectivamente. Puesto que

$$\begin{aligned} F &= \frac{SC_{\text{cuadrática pura}}}{\hat{\sigma}^2} \\ &= \frac{0.0027}{0.043} = 0.063 \end{aligned}$$

es pequeño, no hay indicios de un efecto cuadrático puro. En la tabla 2

Tabla 2. ANOVA del modelo de primer orden

Fuente de variación	SC	g.l	CM	F <sub>0</sub>	valor p
Modelo ( $\beta_1, \beta_2$ )	2.8250	2	1.4125	47.83	0.0002
Residual	0.1772	6			
(Interacción)	0.0025	1	0.0025	0.058	0.8215
(Cuadrático puro)	0.0027	1	0.0027	0.063	0.8142
(Error puro)	0.1720	4	0.0430		
Total	3.0022	8			

En la tabla 2 se resume el análisis de varianza de este modelo. Las verificaciones de la interacción y la curvatura no son significativas, mientras que la prueba  $F$  de la regresión global es significativa. Además, el error estándar de  $\hat{\beta}_1$  y  $\hat{\beta}_2$  es:

$$\sigma_{\hat{\beta}_i} = \sqrt{\frac{CME}{4}} = \sqrt{\frac{\hat{\sigma}^2}{4}} = \sqrt{\frac{0.0430}{4}} = 0.10, i = 1, 2.$$

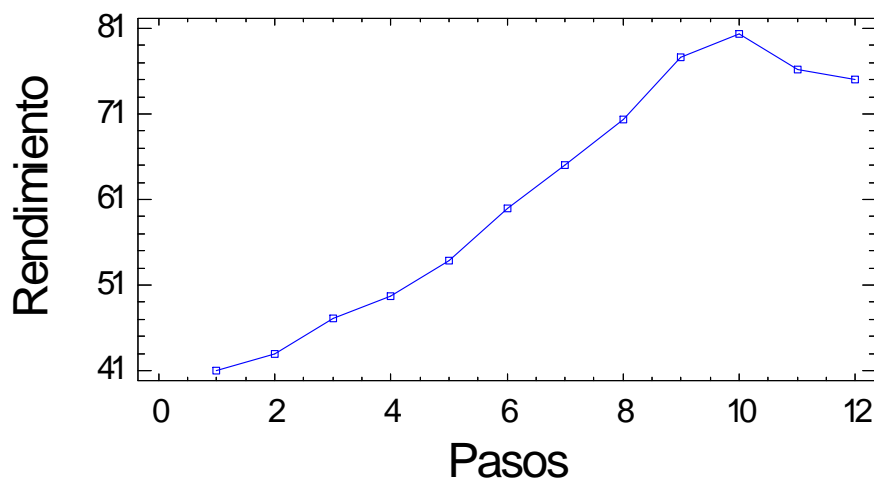
Ambos coeficientes de regresión  $\hat{\beta}_1 = 0.775$  y  $\hat{\beta}_2 = 0.325$  son grandes en comparación con sus errores estándar. En este punto no hay razón para cuestionar la adecuación del modelo de primer orden.

Para apartarse del centro del diseño (el punto  $(x_1 = 0, x_2 = 0)$ ) sobre la trayectoria del ascenso más pronunciado, se haría un movimiento de 0.775 unidades en la dirección  $x_1$  por cada 0.325 unidades en la dirección  $x_2$ . Por lo tanto, la trayectoria del ascenso más pronunciado pasa por el punto  $(x_1 = 0, x_2 = 0)$  y tiene pendiente  $0.325/0.775 = 0.42$ . El ingeniero decide usar 5 minutos de tiempo de reacción como tamaño básico del paso. Al utilizar la relación entre  $\xi_1$  y  $x_1$ , se observa que 5 minutos de tiempo de reacción es equivalente a un paso en la *variable codificada*  $x_1$ , de  $\Delta x_1 = 1$ . Por lo tanto, los pasos sobre la trayectoria del ascenso más pronunciado son

$$\Delta x_1 = 1 \text{ y } \Delta x_2 = (0.325/0.775)\Delta x_1 = 0.42$$

El ingeniero calcula puntos sobre esta trayectoria y observa los rendimientos en los mismos hasta que se nota un decremento en la respuesta. En la tabla 3 se muestran los resultados tanto en variables codificadas como naturales. Aun cuando la manipulación matemática de las variables codificadas es más sencilla, deben usarse las variables naturales cuando se corre el proceso. En la figura 5 se grafica el rendimiento en cada paso de la trayectoria del ascenso más pronunciado. Se observan incrementos de la respuesta hasta el décimo paso; sin embargo, todos los pasos después de este punto resultan en un decremento del rendimiento. Por lo tanto, deberá ajustarse otro modelo de primer orden en la vecindad general del punto  $(\xi_1 = 85, \xi_2 = 175)$ . La región de exploración para  $\xi_1$  es  $[80, 90]$  y para  $\xi_2$  es

Figura 5. Gráfico de Rendimiento frente a Pasos



[170, 180]. Por lo tanto, las variables codificadas son

$$x_1 = \frac{\xi_1 - 85}{5} \text{ y } x_2 = \frac{\xi_2 - 175}{5}$$

pasos	$x_1$	$x_2$	$\xi_1$	$\xi_2$	$y$
origen	0	0	35	155	
$\Delta$	1.00	0.42	5	2	
origen+ $\Delta$	1.00	0.42	40	157	41.0
origen+2 $\Delta$	2.00	0.84	45	159	42.9
origen+3 $\Delta$	3.00	1.26	50	161	47.1
origen+4 $\Delta$	4.00	1.68	55	163	49.7
origen+5 $\Delta$	5.00	2.10	60	165	53.8
origen+6 $\Delta$	6.00	2.52	65	167	59.9
origen+7 $\Delta$	7.00	2.94	70	169	65.0
origen+8 $\Delta$	8.00	3.36	75	171	70.4
origen+9 $\Delta$	9.00	3.78	80	173	77.6
origen+10 $\Delta$	10.00	4.20	85	175	80.3
origen+11 $\Delta$	11.00	4.62	90	179	76.2
origen+12 $\Delta$	12.00	5.04	95	181	75.1

$\xi_1$	$\xi_2$	$x_1$	$x_2$	$y$
80	170	-1	-1	76.5
80	180	-1	1	77.0
90	170	1	-1	78.0
90	180	1	1	79.5
85	175	0	0	79.9
85	175	0	0	80.3
85	175	0	0	80.0
85	175	0	0	79.7
85	175	0	0	79.8

De nuevo se usa un diseño  $2^2$  con cinco puntos centrales. El diseño experimental se muestra en la tabla 4. El ajuste del modelo de primer orden a las variables codificadas de la tabla 4 es  $\hat{y} = 78.97 + 1.00x_1 + 0.50x_2$ . En la tabla 5 se presenta el análisis de varianza de este modelo, incluyendo las verificaciones de la interacción y del término cuadrático puro. que implican que el modelo de primer orden no es una aproximación adecuada. Esta curvatura en la verdadera superficie puede indicar

Análisis de la Varianza para Rendimiento					
Fuente	Suma de Cuadrados	Gl	Cuadrado medio	F-Ratio	P-Valor
A:Tiempo	4,0	1	4,0	75,47	0,0010
B:Temperatura	1,0	1	1,0	18,87	0,0122
AB	0,25	1	0,25	4,72	0,0956
Falta de ajuste	10,658	1	10,658	201,09	0,0001
Error Puro	0,212	4	0,053		
Total (corr.)	16,12	8			
R-cuadrado = 32,5682 por ciento					
R-cuadrado (ajustado para g.l.) = 0,0 por ciento					
Error Estándar de Est. = 0,230217					
Error absoluto de la media = 1,08148					
Estadístico Durbin-Watson = 0,457452 (P=0,0123)					
Autocorrelación residual Lag 1 = 0,671241					

que el experimentador se encuentra cerca del óptimo. En este punto es necesario hacer análisis adicionales para localizar el óptimo con mayor precisión.

Tabla 5. ANOVA del segundo modelo de primer orden

Fuente de variación	SC	g.l	CM	F <sub>0</sub>	valor p
Regresión	5.0000	2			
Residual	11.120	6			
(Interacción)	0.2500	1	0.2500	4.72	0.0955
(Cuadrático puro)	10.658	1	10.658	201.09	0.0001
(Error puro)	0.2120	4	0.0530		
Total	16.120	8			

NOTA 8. Por el ejemplo ilustrativo se observa que la trayectoria del ascenso más pronunciado es proporcional a los signos y magnitudes de los coeficientes de regresión del modelo ajustado de primer orden  $\hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i$ . Es sencillo dar un algoritmo general para determinarlas coordenadas de un punto sobre la trayectoria del ascenso más pronunciado. Suponga que el punto  $x_1 = x_2 = \dots = x_k = 0$  es la base o punto de origen. Entonces, (1) Se elige el tamaño del paso en una de las variables del proceso, por ejemplo  $\Delta x_j$ . En general, se seleccionaría la variable de la que se tenga mayor información, o se seleccionaría la variable que tiene el coeficiente de regresión absoluto  $|\hat{\beta}_j|$  más grande. (2) El tamaño del paso de las otras variables es

$$\Delta x_i = \frac{\hat{\beta}_i}{\hat{\beta}_j / \Delta x_j}, i = 1, 2, \dots, k; i \neq j$$

---

**Coef. de regresión para Rendimiento**


---

constante = 78,9667  
 A:Tiempo = 1,0  
 B:Temperatura = 0,5  
 AB = 0,25

---



---

**El StatAdvisor**


---

Esta ventana muestra la ecuación de la regresión que se ajustado a los datos. La ecuación del modelo ajustado es

**Rendimiento = 78,9667 + 1,0\*Tiempo + 0,5\*Temperatura + 0,25\*Tiempo\*Temperatura**

---

(3) Se convierten las  $\Delta x_i$  de variables codificadas a variables naturales. Para ilustrar, considere la trayectoria del ascenso más pronunciado calculada en el ejemplo ilustrativo. Puesto que  $x_1$  tiene el coeficiente de regresión más grande, se selecciona el tiempo de reacción como la variable del paso 1 del procedimiento anterior. Cinco minutos de tiempo de reacción es el tamaño del paso (con base en el conocimiento del proceso). En términos de las variables codificadas, éste es  $\Delta x_1 = 1.0$ . Por lo tanto, por el lineamiento 2, el tamaño del paso de la temperatura es

$$\Delta x_2 = \frac{\hat{\beta}_2}{\hat{\beta}_1 / \Delta x_1} = \frac{0.325}{0.775 / 1.0} = 0.42.$$

Para convertir los tamaños de los pasos codificados ( $\Delta x_1 = 1.0$  y  $\Delta x_2 = 0.42$ ) a las unidades naturales de tiempo y temperatura, se usan las relaciones

$$\Delta x_1 = \frac{\Delta \xi_1}{5} \text{ y } \Delta x_2 = \frac{\Delta \xi_2}{5}$$

que dan como resultado

$$\begin{aligned} \Delta \xi_1 &= 5\Delta x_1 = 5(1.0) = 5 \text{ min} \\ \Delta \xi_2 &= 5\Delta x_2 = 5(0.42) = 2^\circ F \end{aligned}$$

### 5. Análisis de una superficie de respuesta de segundo orden

Cuando el experimentador se encuentra relativamente cerca del óptimo, por lo general se requiere un modelo que incorpore la curvatura para aproximar la respuesta. En la mayoría de los casos, el modelo de segundo orden

$$(5.1) \quad y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \varepsilon$$

es adecuado.

**5.1. Localización del punto estacionario.** Suponga que quieren encontrarse los niveles de  $\vec{\mathbf{x}}_s$  que optimizan la respuesta predicha. Este punto, en caso de existir, satisface  $\partial\hat{y}/\partial x_j = 0$   $j = 1, \dots, n$  y se le llama *punto estacionario*. El punto estacionario podría representar (1) un punto de respuesta máxima, (2) un punto de respuesta mínima, o (3) un punto silla. Es posible obtener una solución matemática general para la localización del punto estacionario, si se escribe el modelo de segundo orden en notación matricial, se tiene

$$(5.2) \quad \hat{y} = \hat{\beta}_0 + \mathbf{x}^t \mathbf{b} + \mathbf{x}^t \mathbf{B} \mathbf{x}$$

donde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \cdots & \hat{\beta}_{1k}/2 \\ & \hat{\beta}_{22} & \cdots & \hat{\beta}_{2k}/2 \\ & & \ddots & \vdots \\ \text{simétrica} & \cdots & & \hat{\beta}_{nn} \end{bmatrix}$$

Es decir,  $\mathbf{b}$  es un vector  $(k \times 1)$  de los coeficientes de regresión de primer orden y  $\mathbf{B}$  es una matriz simétrica  $(k \times k)$  cuyos elementos de la diagonal principal son los coeficientes cuadráticos puros  $(\hat{\beta}_{ii})$  y cuyos elementos que están fuera de la diagonal son la mitad de los coeficientes cuadráticos *mixtos*  $(\hat{\beta}_{ij}, i \neq j)$ . La derivada de  $\hat{y}$  con respecto a los elementos del vector  $\mathbf{x}$  igualada con  $\mathbf{0}$  es

$$(5.3) \quad \frac{\partial \hat{y}}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} = \mathbf{0}$$

El punto estacionario es la solución de la ecuación (5.3) o

$$(5.4) \quad \mathbf{x}_s = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b}$$

Además, al sustituirla la ecuación (5.4) en la (5.2), la respuesta predicha en el punto estacionario puede encontrarse como

$$(5.5) \quad \hat{y} = \hat{\beta}_0 + \frac{1}{2}\mathbf{x}_s^t \mathbf{b}$$

**5.2. Caracterización de la superficie de respuesta.** Una vez que se ha encontrado el punto estacionario, generalmente es necesario caracterizar la superficie de respuesta en la vecindad inmediata de este punto. Por caracterizar se entiende determinar si el punto estacionario es el punto de una respuesta máxima, mínima o un punto silla. Por lo general, también se desea estudiar la sensibilidad relativa de la respuesta a las variables  $x_1, x_2, \dots, x_k$ . Como ya se señaló, la forma más directa de hacer esto es examinando una *gráfica de contorno* del modelo ajustado. Si sólo hay dos o tres variables en el proceso (las  $x$ ), la construcción e interpretación de esta gráfica de contorno es relativamente sencilla. Sin embargo, incluso cuando hay un número relativamente grande de variables, un análisis más formal, llamado *análisis canónico*, puede ser útil. Es conveniente transformar primero el modelo en un nuevo sistema de coordenadas con el origen en el punto estacionario  $\mathbf{x}$ , y después hacer la rotación de los ejes de este sistema hasta que sean paralelos a los ejes principales de la superficie de respuesta ajustada. Puede demostrarse que se obtiene así el modelo ajustado

$$(5.6) \quad \hat{y} = \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_k w_k^2$$

donde: las  $\{w_i\}$  son las variables independientes transformadas,  $\hat{y}_s$  es el valor predicho por el modelo en el punto estacionario y las  $\{\lambda_i\}$  son constantes. A la ecuación (5.6) se le llama la *forma canónica* del modelo. Además, las  $\{\lambda_i\}$  son sólo *eigenvalores* o *raíces características* de la matriz B.

La naturaleza de la superficie de respuesta puede determinarse a partir del punto estacionario y de los *signos* y *magnitudes* de las  $\{\lambda_i\}$ . Primero suponga que el punto estacionario está dentro de la región de exploración para ajustar el modelo de segundo orden. Si todas las  $\{\lambda_i\}$  son positivas,  $\mathbf{x}_s$  es un punto de respuesta mínima; si todas las  $\{\lambda_i\}$  son negativas,  $\mathbf{x}_s$  es un punto de respuesta máxima; y si las  $\{\lambda_i\}$  tienen signos diferentes,  $\mathbf{x}_s$  es un punto silla.

**5.3. Ejemplo ilustrativo 2.** Se continuará el análisis del proceso químico del ejemplo ilustrativo. No es posible ajustar un modelo de segundo orden en las variables  $x_1$  y  $x_2$ , utilizando el diseño de la tabla 4. El experimentador decide aumentar este diseño con puntos suficientes para ajustar un modelo de segundo orden. Obtiene cuatro observaciones en  $(x_1 = 0, x_2 = \pm 1.414)$  y  $(x_1 = \pm 1.414, x_2 = 0)$ . El experimento completo se muestra en la tabla 6, y el diseño se ilustra en la figura 5. A este diseño se le llama *diseño central compuesto* (o DCC). En esta segunda fase del estudio, dos respuestas adicionales fueron de interés, la viscosidad y el peso molecular del producto. Las respuestas también se muestran en la tabla 6.

Tabla 6. Diseño Central Compuesto						
<i>Variables naturales</i>		<i>Variables codificadas</i>		respuestas		
$\xi_1$	$\xi_2$	$x_1$	$x_2$	$y_1$	$y_2$	$y_3$
80	170	-1	-1	76,5	62	2940
80	180	-1	1	77,0	60	3470
90	170	1	-1	78,0	66	3680
90	180	1	1	79,5	59	3890
85	175	0	0	79,9	72	3480
85	175	0	0	80,3	69	3200
85	175	0	0	80,0	68	3410
85	175	0	0	79,7	70	3290
85	175	0	0	79,8	71	3500
92.07	175	1.414	0	78.4	68	3360
77.93	175	-1.414	0	75.6	71	3020
85	182.07	0	1.414	78.5	58	3630
85	167.93	0	-1.414	77.0	57	3150

La atención se centrará en el ajuste de un modelo cuadrático para la respuesta rendimiento  $y_1$ . Con base en el *valor - p* pequeño de los términos cuadráticos, se decide ajustar un modelo de segundo orden a la respuesta rendimiento. La salida de computadora muestra el modelo final en términos tanto de las variables codificadas como de los niveles naturales o reales de los factores.

En las figuras 11a y 11b, se muestra la gráficas de la superficie de respuesta tridimensional y la gráfica contorno para la respuesta rendimiento en términos de las variables del proceso tiempo y temperatura. Es relativamente sencillo ver por el examen de estas figuras que el óptimo se encuentra muy cerca de  $175^\circ F$  y 85 minutos de tiempo de reacción y que la respuesta está en un máximo en este

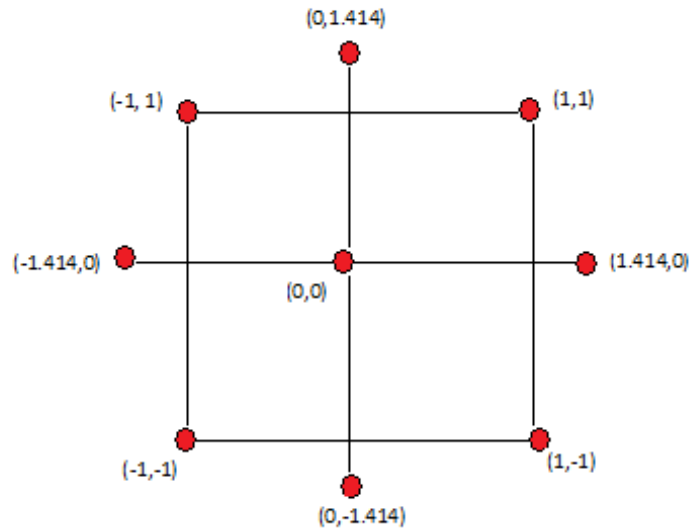
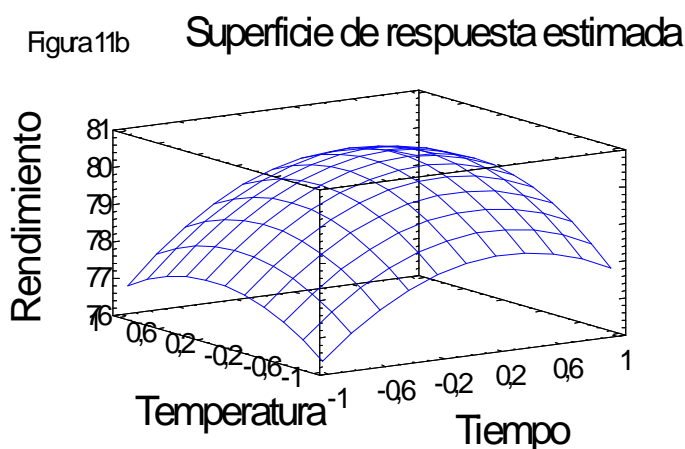
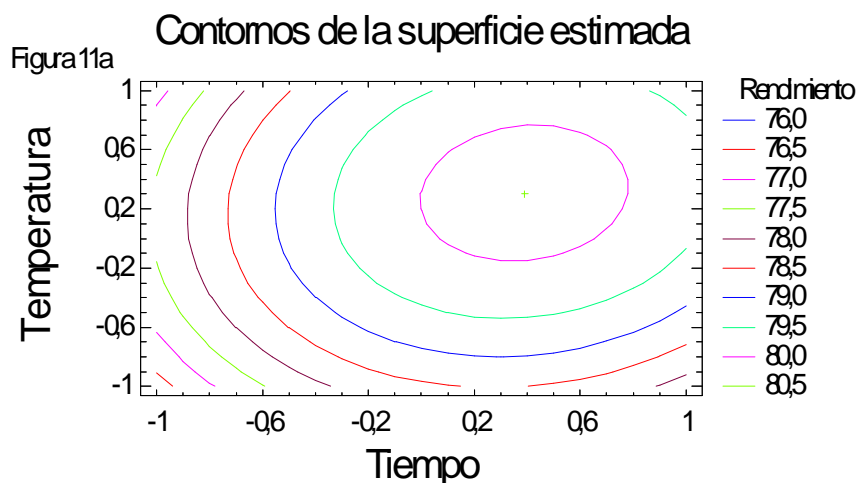


Figura 5

Analysis of Variance for Rendimiento					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
A:Tiempo	7,9198	1	7,9198	149,43	0,0003
B:Temperatura	2,12316	1	2,12316	40,06	0,0032
AA	13,1761	1	13,1761	248,61	0,0001
AB	0,25	1	0,25	4,72	0,0956
BB	6,97389	1	6,97389	131,58	0,0003
Lack-of-fit	0,28331	3	0,0944366	1,78	0,2897
Pure error	0,212	4	0,053		
Total (corr.)	28,7431	12			
R-squared = 98,2768 percent					
R-squared (adjusted for d.f.) = 97,0459 percent					
Standard Error of Est. = 0,230217					
Mean absolute error = 0,180004					
Durbin-Watson statistic = 1,00536 (P=0,0249)					
Lag 1 residual autocorrelation = 0,438098					

punto. Por el examen de la gráfica de contorno se observa que el proceso puede ser ligeramente más sensible a los cambios en el tiempo de reacción que a los cambios en la temperatura.



Es decir,  $x_{1,s} = 0.389595$  y  $x_{2,s} = 0.305137$ . En términos de las variables naturales, el punto estacionario es

$$0.389595 = \frac{\xi_1 - 85}{5}, \quad 0.305137 = \frac{\xi_2 - 175}{5}$$

de donde se obtiene  $\xi_1 = 86.948 \simeq 87$  minutos de tiempo de reacción y  $\xi_2 = 176.53 \simeq 176.5^\circ F$ . Este valor está muy cerca del punto estacionario que se encontró por examen visual en la gráfica de contorno de la figura 11. Al utilizar la ecuación (5.5), la respuesta predicha en el punto estacionario puede encontrarse como  $\hat{y}_s = 80.21$ .

El análisis canónico que se describe en esta sección también puede usarse para caracterizar la superficie de respuesta. Primero es necesario expresar el modelo ajustado en la forma canónica (ecuación 5.6). Los eigenvalores  $\lambda_1$  y  $\lambda_2$  son las



**Optimize Response****Goal: maximize Rendimiento****Optimum value = 80,2124**

Factor	Low	High	Optimum
Tiempo	-1,41421	1,41421	0,389595
Temperatura	-1,41421	1,41421	0,305137

**Regression coeffs. for Rendimiento**

constant	= 79,94
A:Tiempo	= 0,994976
B:Temperatura	= 0,515166
AA	= -1,37625
AB	= 0,25
BB	= -1,00125

raíces de la ecuación de determinantes

$$\begin{array}{rcl}
 |B - \lambda I| & = & 0 \\
 \left| \begin{array}{cc} -1.37625 - \lambda & 0.1250 \\ 0.1250 & -1.00125 - \lambda \end{array} \right| & = & 0
 \end{array}$$

que se reduce a

$$\lambda^2 + 2.3788\lambda + 1.369 = 0$$

Las raíces de esta ecuación cuadrática son  $\lambda_1 = -0.9634$  y  $\lambda_2 = -1.4141$ . Por lo tanto, la forma canónica del modelo ajustado es

$$\hat{y} = 80.21 - 0.9634w_1^2 - 1.4141w_2^2$$

Puesto que tanto  $\lambda_1$  como  $\lambda_2$  son negativas y el punto estacionario está en la región de exploración, se concluye que el punto estacionario es un máximo.

**5.4. Relación entre las variables canónicas.** En algunos problemas **MSR** puede ser necesario encontrar la relación entre las variables canónicas  $\{w_i\}$  y las variables del diseño  $\{x_i\}$ . Esto es particularmente cierto cuando es imposible operar el proceso en el punto estacionario. Como una ilustración, suponga que en el ejemplo ilustrativo 2 el proceso no pudo operarse en  $\xi_1 = 87$  minutos y  $\xi_2 = 176.5^\circ F$  debido a que esta combinación de factores resulta en un costo excesivo. Se quiere “regresar” ahora del punto estacionario a un punto con un costo menor sin incurrir en pérdidas considerables en el rendimiento. La forma canónica del modelo indica que la superficie es menos sensible a la pérdida de rendimiento en la dirección  $w_1$ . La exploración de la forma canónica requiere convertir los puntos del espacio  $(w_1, w_2)$  en puntos del espacio  $(x_1, x_2)$ .

En general, las variables  $\mathbf{x}$  se relacionan con las variables canónicas  $\mathbf{w}$  por:

$$\mathbf{w} = \mathbf{M}^t(\mathbf{x} - \mathbf{x}_s)$$

donde  $\mathbf{M}$  es una matriz ortogonal ( $k \times k$ ). Las columnas de  $\mathbf{M}$  son los eigenvectores normalizados asociados con  $\{\lambda_i\}$ . Es decir, si  $\mathbf{m}_i$  es la columna  $i$ -ésima de  $\mathbf{M}$ , entonces  $\mathbf{m}_i$  es la solución de

$$(5.7) \quad (\mathbf{B} - \lambda_i \mathbf{I}) \mathbf{m}_i = \mathbf{0}$$

para la que  $\sum_{j=1}^k m_{ji}^2 = 1$ .

El procedimiento se ilustra usando el modelo de segundo orden ajustado. Para  $\lambda_1 = -0.9641$ , la ecuación 5.7 queda como

$$\begin{bmatrix} -1.37625 + 0.9634 & 0.1250 \\ 0.1250 & -1.00125 + 0.9634 \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

o

$$\begin{aligned} -0.41285m_{11} + 0.1250m_{21} &= 0 \\ 0.1250m_{11} - 0.03785m_{21} &= 0 \end{aligned}$$

Quiere obtenerse la solución normalizada de estas ecuaciones, es decir, aquella para la que  $m_{11}^2 + m_{21}^2 = 1$ . Al no existir una solución única para estas ecuaciones, lo más conveniente es asignar un valor arbitrario a una de las incógnitas, resolver el sistema y normalizar la solución. Al hacer  $m_{21}^* = 1$ , se encuentra  $m_{11}^* = 0.302773$ . Para normalizar esta solución  $m_{11}^*$  y  $m_{21}^*$  se dividen entre

$$\sqrt{(m_{11}^*)^2 + (m_{21}^*)^2} = \sqrt{(0.302773)^2 + (1)^2} = 1.04483$$

Se obtiene así la solución normalizada

$$\begin{aligned} m_{11} &= \frac{m_{11}^*}{1.04483} = \frac{0.302773}{1.04483} = 0.289782 \\ m_{21} &= \frac{m_{21}^*}{1.0448} = \frac{1}{1.04483} = 0.957045 \end{aligned}$$

que es la primera columna de la matriz  $\mathbf{M}$ .

Utilizando  $\lambda_2 = -1.4141$ , y repitiendo el procedimiento anterior, se tiene

$$\begin{bmatrix} -1.37625 + 1.4141 & 0.1250 \\ 0.1250 & -1.00125 + 1.4141 \end{bmatrix} \begin{bmatrix} m_{12} \\ m_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

o

$$\begin{aligned} 0.3785m_{12} + 0.1250m_{22} &= 0 \\ 0.1250m_{12} + 0.41285m_{22} &= 0 \end{aligned}$$

Quiere obtenerse la solución normalizada de estas ecuaciones, es decir, aquella para la que  $m_{12}^2 + m_{22}^2 = 1$ . Al no existir una solución única para estas ecuaciones, lo más conveniente es asignar un valor arbitrario a una de las incógnitas, resolver el sistema y normalizar la solución. Al hacer  $m_{22}^* = 1$ , se encuentra  $m_{12}^* = -3.3028$ . Para normalizar esta solución  $m_{11}^*$  y  $m_{21}^*$  se dividen entre

$$\sqrt{(m_{12}^*)^2 + (m_{22}^*)^2} = \sqrt{(-3.3028)^2 + (1)^2} = 3.45087$$

Se obtiene así la solución normalizada

$$\begin{aligned} m_{12} &= \frac{m_{12}^*}{3,45087} = \frac{-3.3028}{3,45087} = -0.9571 \\ m_{22} &= \frac{m_{21}^*}{3,45087} = \frac{1}{3,45087} = 0.289782 \end{aligned}$$

que es la segunda columna de la matriz  $\mathbf{M}$ . Así que

$$\mathbf{M} = \begin{bmatrix} 0.2897 & -0.957045 \\ 0.9571 & 0.289782 \end{bmatrix}$$

La relación entre las variables  $\mathbf{w}$  y  $\mathbf{x}$  es

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0.289782 & -0.957045 \\ 0.957045 & 0.289782 \end{bmatrix} \begin{bmatrix} x_1 - 0.389595 \\ x_2 - 0.305137 \end{bmatrix}$$

Si quisiera explorarse la superficie de respuesta en la vecindad del punto estacionario, podrían determinarse los puntos apropiados en los cuales hacer las observaciones en el espacio  $(w_1, w_2)$  y usar después la relación anterior para convertir estos puntos en el espacio  $(x_1, x_2)$  para que puedan realizarse las corridas.

## 6. Sistemas de cordilleras

Muchas veces, el punto estacionario no es del tipo que se requiere y en esos casos la opción es encontrar el “mejor punto posible” dentro de la región experimental. *Este punto se ubica sobre la cordillera óptima a partir del centro del diseño, y es aquel que predice la mejor respuesta sobre la región.* Esta búsqueda se hace precisamente con el llamado análisis de cordilleras, que consiste en calcular el máximo o mínimo de la superficie de respuesta sobre esferas concéntricas al centro del diseño, empezando por una esfera de radio casi cero y posteriormente se abre la exploración incrementando el radio de la siguiente esfera. Así se continúa hasta llegar a la primera esfera que cubre los puntos experimentales. El mejor punto posible es aquel sobre el que se predice el óptimo desempeño de la variable de respuesta. Con frecuencia, este punto se ubica en la esfera de radio más grande. En el caso de  $k = 2$  factores, no son esferas sino circunferencias.

Ahora, brevemente veamos en forma matemática el análisis de cordillera. Consideremos el modelo ajustado de segundo orden escrito en su forma matricial

$$\hat{y} = \hat{\beta}_0 + \mathbf{x}^t \mathbf{b} + \mathbf{x}^t \mathbf{B} \mathbf{x}$$

Sea la esfera centrada en el origen con radio  $R_i$  cuyos puntos sobre ella cumplen la restricción:  $\sum_{i=1}^k x_i^2 = \mathbf{x}^t \mathbf{x} = R_i$ . El problema del análisis de cordillera es encontrar el punto sobre la esfera, donde la respuesta predicha por el modelo es máxima (o mínima). Para ello se plantea la función objetivo dada por

$$F = \hat{\beta}_0 + \mathbf{x}^t \mathbf{b} + \mathbf{x}^t \mathbf{B} \mathbf{x} - \lambda (\mathbf{x}^t \mathbf{x} - R_i)$$

donde  $\lambda$  es multiplicador de Lagrange. Derivando esta última relación con respecto al vector  $\mathbf{x}$  e igualando a cero, se obtiene

$$\frac{\partial F}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} - 2\lambda\mathbf{x} = \mathbf{0}$$

y de aquí se llega al sistema de ecuaciones

$$(\mathbf{B} - \lambda \mathbf{I}) \mathbf{x} = \frac{-\mathbf{b}}{2}$$

El punto  $(x_1, x_2, \dots, x_k)$  óptimo sobre una esfera particular se encuentra al sustituir un valor para  $\lambda$ , que no sea un valor propio de la matriz  $\mathbf{B}$  en esta última relación, y se resuelve el sistema de ecuaciones resultante. En general es mejor recurrir a un software para hacer el análisis de cordillera. Por ejemplo, en *Statgraphics*, una vez que el modelo ha sido ajustado y depurado, se elige la opción de Optimization y ahí, haciendo clic con el botón derecho del mouse, se puede indicar si se quiere minimizar, maximizar o elegir un valor objetivo. También se le especifica el rango de valores de los factores en donde se quiere la optimización.

A continuación presentamos un ejemplo en donde se aplica tanto la técnica de escalamiento ascendente como el análisis canónico.

### 7. Respuestas múltiples

Muchos problemas de superficies de respuesta incluyen el análisis de varias respuestas, como en el ejemplo ilustrativo 2, donde el experimentador midió tres. En dicho ejemplo, el proceso se optimizó únicamente con respecto a la respuesta rendimiento  $y_1$ . La consideración simultánea de respuestas múltiples requiere construir primero un modelo de superficie de respuesta apropiado para cada respuesta y después intentar encontrar un conjunto de condiciones de operación que optimice en cierto sentido todas las respuestas o que al menos las mantenga en los rangos deseados. En el ejemplo ilustrativo 2, pueden obtenerse modelos para las respuestas viscosidad y peso molecular ( $y_2$  y  $y_3$ , respectivamente) de la siguiente manera:

$$\begin{aligned}\hat{y}_2 &= 70.00 - 0.16x_1 - 0.95x_2 - 0.69x_1^2 - 6.69x_2^2 - 1.25x_1x_2 \\ \hat{y}_3 &= 3386.2 + 205.1x_1 + 17.4x_2\end{aligned}$$

En términos de los niveles naturales del tiempo ( $\xi_1$ ) y la temperatura ( $\xi_2$ ), estos modelos son

$$\begin{aligned}\hat{y}_2 &= -9030.74 + 13.393\xi_1 + 97.70\xi_2 - 2.75 \times 10^{-2}\xi_1^2 \\ &\quad - 0.26757\xi_2^2 - 5 \times 10^{-2}\xi_1\xi_2 \\ \hat{y}_3 &= -6308.8 + 41.025\xi_1 + 35.47\xi_2\end{aligned}$$

En las figuras 14 y 15 se presentan las graficas de contornos y superficies de respuestas para estos modelos.

### 8. Diseños de superficies de respuesta

Los diseños de superficie de respuesta se clasifican con base en el grado del modelo que se pretende utilizar. Estos diseños proporcionan los tratamientos a correr para generar datos que permitan ajustar un modelo que describa una variable de respuesta en una región experimental. Algunas propiedades deseables en los diseños para la MSR son:

- (1) Que genere una distribución satisfactoria de los puntos experimentales sobre la región experimental. Los diseños más utilizados son puntos distribuidos de manera uniforme sobre la región experimental, o cuando menos tienen alguna simetría con respecto al centro de ésta.

Figura 14a Contornos para la viscosidad

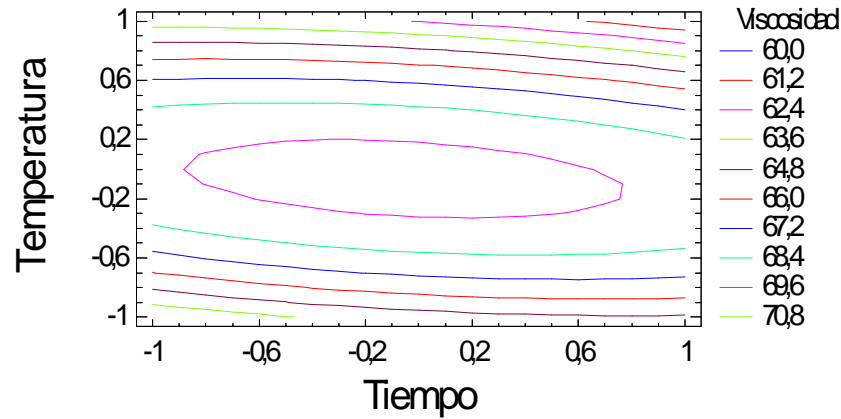
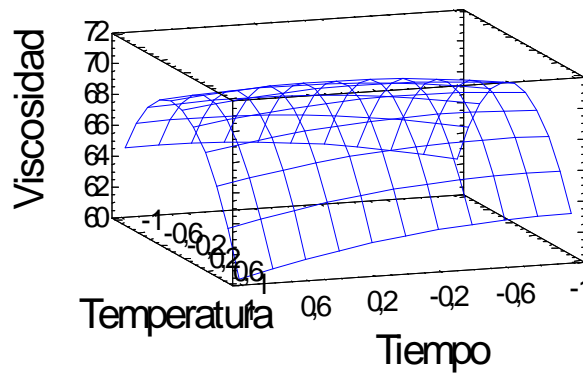


Figura 14b Superficie de respuesta para viscosidad



- (2) El diseño debe requerir un número mínimo de corridas experimentales, ya que en cada prueba realizada se gastan recursos que siempre son escasos.
- (3) El diseño debe permitir que otros diseños de orden mayor se construyan a partir de él. Esto permite que, cuando el comportamiento de la respuesta resulta ser más complicado de lo que se pensaba (por ejemplo, se detecta curvatura), se agregan puntos adicionales al diseño para tratar de explicar ese comportamiento.
- (4) El experimento debe permitir la detección de la falta de ajuste, para lo cual se requieren repeticiones al menos en el centro del diseño.
- (5) El diseño debe proporcionar un estimador puro de la varianza del error, lo cual se logra con repeticiones al menos en el punto central.

Otras dos propiedades deseables en los diseños para superficie de respuesta son la ortogonalidad y la rotabilidad. Estas propiedades aumentan la eficiencia de los diseños que las poseen, en el sentido de que facilitan la interpretación de los parámetros estimados en el modelo y de la superficie de respuesta.

**8.1. Ortogonalidad y rotabilidad.** Se considera que un *diseño es ortogonal* cuando los coeficientes estimados en el modelo ajustado no están correlacionados entre sí, lo cual hace que el efecto de cada término, representado por el parámetro correspondiente, se estime de manera más precisa. *Un experimento es ortogonal* si en la matriz de diseño todos los vectores columna son independientes entre sí. Es fácil verificar que en un diseño factorial completo  $2^k$  las columnas de su matriz de diseño son independientes: multiplique dos columnas cualesquiera, término a término usando la notación  $-1$  y  $+1$ , y el resultado es cero. *Un diseño se llama rotable* si la varianza de  $Y(x)$  sólo depende de la distancia del punto  $\mathbf{x}$  al centro del diseño y no de la dirección en la que se encuentra. Es decir, si pensamos en la variable  $\text{var}[\hat{Y}(x)]$  como otra respuesta, su gráfica de contornos tiene la forma de círculos concéntricos alrededor del centro del diseño. La rotabilidad del diseño asegura que la calidad de la predicción, medida por  $\text{var}[\hat{Y}(x)]$ , sea invariante a cualquier rotación del diseño alrededor del centro del mismo; de aquí se deriva el nombre de esta propiedad.

La importancia práctica de la rotabilidad en el problema de encontrar un punto óptimo es porque asegura que, con independencia de la dirección en que se explore a partir del centro del diseño, la calidad de predicción se comporta de igual manera.

**8.1.1. Relación modelo-diseño.** Existe una relación directa entre el tipo de modelo que se pretende ajustar y el tipo de diseño que se debe correr. No se debe exigir a un diseño experimental más información de la que puede dar. Por ejemplo, si se corre un factorial completo  $2^k$  sólo se podrán estimar e incluir en el modelo los efectos principales e interacciones dobles, asimismo, no es posible estimar términos cuadráticos puros (como  $x_i^2$ ). Si al diseño factorial se le agregan repeticiones al centro ( $2^k + \text{centro}$ ), en el modelo se puede incluir sólo uno de los términos cuadráticos puros, cualquiera de ellos, ya que son alias. Las repeticiones al centro no son suficientes para investigar cuál o cuáles de los términos cuadráticos están activos, pero sí permiten detectar la presencia de curvatura. Si el diseño se aumenta con puntos estrella o axiales es posible estudiar de manera separada los efectos cuadráticos puros e incluirlos a todos en el modelo ajustado, si fuera necesario.

**8.1.2. Diseños para ajustar el modelo de primer orden.** Suponga que quiere ajustarse el modelo de primer orden en  $k$  variables

$$(8.1) \quad y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \varepsilon$$

Hay una clase única de diseños que minimizan la varianza de los coeficientes de regresión  $\{\hat{\beta}_i\}$ , se trata de los diseños de primer orden ortogonales. Un diseño de primer orden es ortogonal si todos los elementos que están fuera de la diagonal de la matriz  $(\mathbf{X}^t \mathbf{X})$  son cero. Esto implica que la suma de los productos cruzados de las columnas de la matriz  $\mathbf{X}$  sea cero.

La clase de los diseños de primer orden ortogonales incluye los factoriales  $2^k$  y las fracciones de la serie  $2^k$  en las que los efectos principales no son alias entre sí.

Al usar estos diseños se supone que los niveles bajo y alto de los  $k$  factores están codificados en los niveles usuales  $\pm 1$ .

El diseño  $2^k$  no permite la estimación del error experimental a menos que se hagan réplicas de algunas corridas. Un método común de incluir las réplicas en el diseño  $2^k$  es aumentar el diseño con varias observaciones en el centro (el punto  $x_i = 0, i = 1, 2, \dots, k$ ). La adición de puntos centrales al diseño  $2^k$  no influye en las  $\{\hat{\beta}_i\}$  para  $i \geq 1$ , pero la estimación de  $\beta_0$  se convierte en el gran promedio de todas las observaciones. Además, la adición de puntos centrales no altera la propiedad de ortogonalidad del diseño.

8.1.3. *Diseños para ajustar el modelo de segundo orden.* Se llaman diseños de segundo orden aquellos que permiten ajustar un modelo de segundo orden para así estudiar, además de los efectos lineales y de interacción, los efectos cuadráticos o de curvatura pura. Por consiguiente, estos diseños se emplean cuando se quiere explorar una región que se espera sea más compleja o cuando se cree que el punto óptimo ya se encuentra dentro de la región experimental. El modelo de segundo orden está dado por:

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \varepsilon$$

tiene  $p = (k+1)(k+2)/2$  términos, por lo tanto se requiere al menos esa cantidad puntos de diseño. El diseño debe tener al menos tres niveles en cada factor para estimar la curvatura de la superficie en la dirección de cada factor. Es deseable que estos diseños sean ortogonales, pero a veces no es fácil que cumplan esta propiedad y se admite alguna dependencia entre las columnas de los contrastes. Los más utilizados tienen la propiedad de ser rotables. Se utilizó anteriormente una introducción informal del diseño central compuesto o DCC para ajustar un modelo de segundo orden. Se trata de la clase más popular de diseños usados para ajustar estos modelos. En general, el DCC consta de un factorial  $2^k$  (o de un factorial fraccionado de resolución V) con  $n_F$  corridas,  $2k$  corridas axiales o estrella y  $n_C$  corridas centrales. El despliegue práctico de un DCC surge con frecuencia a través de la experimentación secuencial. Es decir, se ha usado un diseño  $2^k$  para ajustar un modelo de primer orden, este modelo ha presentado falta de ajuste, y después se agregaron las corridas axiales para permitir la incorporación de los términos cuadráticos en el modelo. El DCC es un diseño muy eficiente para ajustar el modelo de segundo orden. Hay dos parámetros en el diseño que deben especificarse: la distancia  $\alpha$  de las corridas axiales al centro del diseño y el número de puntos centrales  $n_C$ . A continuación se analiza la elección de estos dos parámetros.

Rotabilidad. Es importante que el modelo de segundo orden proporcione buenas predicciones en toda la región de interés. Una manera de definir "buenas" es requerir que el modelo tenga una varianza razonablemente consistente y estable de la respuesta predicha en los puntos de interés  $\mathbf{x}$ . Recuerde que la varianza de la respuesta predicha en algún punto  $\mathbf{x}$  es

$$V[\hat{y}(x)] = \sigma^2 \mathbf{x}^t (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{x}$$

Box y Hunter propusieron que un diseño de superficie de respuesta de segundo orden debe ser *rotable*. Esto significa que la  $V[\hat{y}(x)]$  es la misma en todos los puntos  $\mathbf{x}$  que están a la misma distancia del centro del diseño. Es decir, la varianza de la respuesta predicha es constante en esferas. La rotabilidad es una base razonable

para la selección de un diseño de superficie de respuesta. Puesto que la finalidad de la **MSR** es la optimización, y la localización del óptimo se desconoce antes de correr el experimento, tiene sentido el uso de un diseño que proporcione una precisión de estimación igual en todas las direcciones (puede demostrarse que cualquier diseño de primer orden ortogonal es rotatable). Un diseño central compuesto se hace rotatable mediante la elección de  $\alpha$ . El valor de  $\alpha$  para la rotabilidad depende del número de puntos en la porción factorial del diseño; de hecho,  $\alpha = n_F^{1/4}$  produce un diseño central compuesto rotatable, donde  $n$ , es el número de puntos usados en la porción factorial del diseño

**El DCC esférico.** La rotabilidad es una propiedad esférica; es decir, tiene mayor sentido como criterio de diseño cuando la región de interés es una esfera. Sin embargo, no es importante tener una rotabilidad exacta para tener un buen diseño. De hecho, para una región esférica de interés, la mejor elección de  $\alpha$  desde el punto de vista de la varianza de predicción para el DCC es hacer  $\alpha = \sqrt{k}$ . Este diseño, llamado DCC esférico, coloca todos los puntos factoriales y axiales del diseño sobre la superficie de una esfera de radio  $\sqrt{k}$ .

**Corridas centrales en el DCC.** La elección de  $\alpha$  en el DCC está dictada principalmente por la región de interés. Cuando esta región es una esfera, el diseño debe incluir corridas centrales para proporcionar una varianza razonablemente estable de la respuesta predicha. En general, se recomiendan de tres a cinco corridas centrales.

**El diseño de Box-Behnken.** Box y Behnken han propuesto algunos diseños de tres niveles para ajustar superficies de respuesta. Estos diseños se forman combinando factoriales  $2^k$  con diseños de bloques incompletos. Los diseños resultantes suelen ser muy eficientes en términos del número requerido de corridas, y son rotatables o casi rotables.

En la tabla 11-8 se muestra el diseño de Box-Behnken para tres variables. El diseño también se ilustra geométricamente en la figura 11-22. Observe que el diseño de Box-Behnken es un diseño esférico, con dos los puntos localizados en una esfera de radio  $\sqrt{2}$ . Asimismo, el diseño de Box-Behnken no contiene

ningún punto en los vértices de la región cúbica creada por los límites superior e inferior de cada variable. Esto podría ser una ventaja cuando los puntos de los vértices del cubo representan combinaciones de los niveles de los factores cuya prueba es prohibitivamente costosa o imposible debido a restricciones físicas del proceso.

**Región cuboidal de interés.** Existen muchas situaciones en las que la región de interés es cuboidal en lugar de esférica. En estos casos, una variante útil del *diseño central compuesto* es el *diseño central compuesto con centros en las caras* o el *cubo con centros en las caras*, en el que  $\alpha = 1$ . En este diseño los puntos axiales o estrella se localizan en los centros de las caras del cubo, como se muestra en la figura 11-23 para  $k = 3$ . Esta variante del diseño central compuesto se usa en ocasiones debido a que sólo requiere tres niveles de cada factor, y en la práctica con frecuencia es difícil cambiar los niveles de los factores. Sin embargo, observe que los diseños centrales compuestos no son rotables.

El cubo con centros en las caras no requiere tantos puntos centrales como el DCC esférico. En la práctica  $n_C = 2$  o  $3$  es suficiente para proporcionar una buena varianza de predicción en toda la región experimental. Cabe señalar que en ocasiones se emplearán más corridas centrales para dar una estimación razonable del error experimental. En la figura 11-24 se muestra la raíz cuadrada de la varianza de



predicción  $V[\hat{y}(x)]$  del cubo con centros en las caras para  $k = 3$  con  $n_C = 3$  puntos centrales ( $x_3 = 0$ ). Observe que la desviación estándar de la respuesta predicha es razonablemente uniforme en una porción relativamente larga del espacio del diseño.

Otros diseños. Existen muchos otros diseños de superficie de respuesta que en ocasiones son útiles en la práctica. Para dos variables, podrían usarse diseños compuestos de puntos cuya separación en un círculo es igual y que forman polígonos regulares. Puesto que los puntos del diseño son equidistantes del origen, a estos arreglos con frecuencia se les llama *diseños equirradiales*.

Para  $k = 2$ , un *diseño equirradial* rotatable se obtiene combinando  $n_2 \geq 5$  puntos con una separación igual en un círculo con  $n_1 \geq 1$  punto en el centro del círculo. Diseños de particular utilidad para  $k = 2$  son el pentágono y el hexágono. Estos diseños se muestran en la figura 11-25. Otros diseños útiles incluyen el *diseño compuesto pequeño*, el cual consiste en un factorial fraccionado en el cubo de resolución III\* (los efectos principales son alias de las interacciones de dos factores y ninguna de las interacciones de dos factores es alias entre sí) y las corridas axiales y centrales usuales, y la clase de los *diseños híbridos*. Estos diseños pueden ser de valor considerable cuando es importante reducir el número de corridas tanto como sea posible.

En la tabla 11-9 se muestra un diseño compuesto pequeño para  $k = 3$  factores. Este diseño usa la fracción un medio estándar del diseño  $2^3$  en el cubo, ya que satisface los criterios de la resolución III\*. El diseño tiene cuatro corridas en el cubo y seis corridas axiales, y debe incluir al menos un punto central. Por lo tanto, el diseño tiene un mínimo de  $N = 11$  ensayos, y el modelo de segundo orden en  $k = 3$  variables tiene  $p = 10$  parámetros por estimar, por lo que se trata de un diseño muy eficiente con respecto al número de corridas. El diseño de la tabla 11-9 tiene  $n_C = 4$  corridas centrales. Se seleccionó  $\alpha = 1.73$  para obtener un diseño esférico debido a que el diseño compuesto pequeño no puede hacerse rotatable.

En la tabla 11-10 se muestra un diseño híbrido para  $k = 3$ . Algunos de estos diseños tienen niveles irregulares, y esto puede ser un factor limitante para su aplicación. Sin embargo, se trata de diseños muy pequeños, y poseen excelentes propiedades de la varianza de predicción.

Tabla 11-9. Diseño Compuesto pequeño para $k = 3$				Tabla 11-10. Diseño híbrido para $k = 3$			
orden estandar	$x_1$	$x_2$	$x_3$	orden estandar	$x_1$	$x_2$	$x_3$
1	1.00	1.00	-1.00	1	0.00	0.00	1.41
2	1.00	-1.00	1.00	2	0.00	0.00	-1.41
3	-1.00	1.00	1.00	3	-1.00	-1.00	0.71
4	-1.00	-1.00	-1.00	4	1.00	-1.00	0.71
5	-1.73	0.00	0.00	5	-1.00	1.00	0.71
6	1.73	0.00	0.00	6	1.00	1.00	0.71
7	0.00	-1.73	0.00	7	1.41	0.00	-0.71
8	0.00	1.73	0.00	8	-1.41	0.00	-0.71
9	0.00	0.00	-1.73	9	0.00	0.00	-0.71
10	0.00	0.00	1.73	10	0.00	0.00	-0.71
11	0.00	0.00	0.00	11	0.00	0.00	0.00
12	0.00	0.00	0.00				
13	0.00	0.00	0.00				
14	0.00	0.00	0.00				

8.1.4. *Formación de bloques en los diseños de superficie de respuesta.* Cuando se usan diseños de superficie de respuesta, con frecuencia es necesario considerar la formación de bloques para eliminar las variables perturbadoras. Por ejemplo, este problema puede ocurrir cuando un diseño de segundo orden se ensambla secuencialmente a partir de un diseño de primer orden, como se ilustró en los ejemplos 11-1 y 11-2. Puede transcurrir tiempo considerable entre que se corre el modelo primer orden y se corren los experimentos complementarios requeridos para construir un diseño de segundo orden, y durante este tiempo las condiciones de prueba pueden cambiar, haciendo necesaria la formación de bloques.

Se dice que un diseño de superficie de respuesta se forma de bloques ortogonales si se divide en bloques tales que sus efectos no afecten las estimaciones de los parámetros del modelo de superficie de respuesta.

Para hacer la formación de bloques ortogonales de un diseño de segundo orden, deben satisfacerse dos condiciones. Si hay  $n_b$  observaciones en el bloque  $b$ -ésimo, entonces estas condiciones son

- (1) Cada bloque debe ser un diseño ortogonal de primer orden; es decir,

$$\sum_{u=1}^{n_b} x_{iu}x_{ju} = 0, \quad i \neq j = 0, 1, \dots, k \text{ para toda } b$$

donde  $x_{iu}$  y  $x_{ju}$  son los niveles de las variables  $i$ -ésima y  $j$ -ésima en la corrida  $u$ -ésima del experimento con  $x_{0u} = 1$  para toda  $u$ .

- (2) La fracción de la suma de cuadrados total para cada variable con que contribuye cada bloque, debe ser igual a la fracción de las observaciones totales que están contenidas en el bloque; es decir,

$$\frac{\sum_{u=1}^{n_b} x_{iu}^2}{\sum_{u=1}^N x_{iu}^2} = \frac{n_b}{N}, \quad i = 0, 1, \dots, k \text{ para toda } b$$

donde  $N$  es el número de corridas del diseño.

Como un ejemplo de la aplicación de estas condiciones, considere un diseño central compuesto rotatable en  $k = 2$  variables con  $N = 12$  corridas. Los niveles  $x_1$  y  $x_2$  de este diseño pueden escribirse en la matriz del diseño

$$\mathbf{D} = \begin{bmatrix} x_1 & x_2 \\ -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1.414 & 0 \\ -1.414 & 0 \\ 0 & 1.414 \\ 0 & -1.414 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Observe que el diseño se ha dispuesto en dos bloques, con el primer bloque consistiendo en la porción factorial del diseño más dos puntos centrales y el segundo bloque consistiendo en los puntos axiales más dos puntos centrales adicionales. Es claro que la condición 1 se satisface; es decir, ambos bloques son diseños de primer orden ortogonales. Para investigar la condición dos, considere primero el bloque 1 y observe que

$$\begin{aligned}\sum_{u=1}^{n_1} x_{1u}^2 &= \sum_{u=1}^{n_1} x_{2u}^2 = 4 \\ \sum_{u=1}^N x_{1u}^2 &= \sum_{u=1}^N x_{2u}^2 = 8 \text{ y } n_1 = 6\end{aligned}$$

Por lo tanto,

$$\begin{aligned}\frac{\sum_{u=1}^{n_1} x_{iu}^2}{N} &= \frac{n_1}{N} \\ \frac{4}{8} &= \frac{6}{12}\end{aligned}$$

Así, la condición 2 se satisface en el bloque 1. Para el bloque 2 se tiene

$$\sum_{u=1}^{n_2} x_{1u}^2 = \sum_{u=1}^{n_2} x_{2u}^2 = 4 \text{ y } n_2 = 6$$

Por lo tanto,

$$\begin{aligned}\frac{\sum_{u=1}^{n_2} x_{iu}^2}{N} &= \frac{n_2}{N} \\ \frac{4}{8} &= \frac{6}{12}\end{aligned}$$

Puesto que la condición 2 también se satisface en el bloque 2, este diseño está formado de bloques ortogonales.

En general, el diseño central compuesto siempre puede construirse para hacer la formación de bloques ortogonales en dos bloques con el primer bloque consistiendo en  $n_F$  puntos factoriales más  $n_{CF}$  puntos centrales y el segundo bloque consistiendo en  $n_A = 2k$  puntos axiales más  $n_{CA}$  puntos centrales. La primera condición de la formación de bloques ortogonales se cumplirá siempre independientemente del valor que se use para  $\alpha$  en el diseño. Para que la segunda condición se cumpla,

$$(8.2) \quad \frac{\sum_{u=1}^{n_2} x_{iu}^2}{\frac{n_1}{n_2}} = \frac{n_A + n_{CA}}{n_F + n_{CF}}$$

El miembro izquierdo de la ecuación (8.2) es  $2\alpha^2/n_F$  y después de sustituir esta cantidad, la ecuación para el valor de  $\alpha$  que resultará en la formación de bloques ortogonales puede resolverse como la ecuación

$$(8.3) \quad \alpha = \left[ \frac{n_F (n_A + n_{CA})}{2(n_F + n_{CF})} \right]^{1/2}$$

Este valor de  $\alpha$  no dará como resultado, en general, un diseño rotatable o esférico. Si se requiere que el diseño también sea rotatable, entonces  $\alpha = (n_F)^{1/4}$  y

$$(8.4) \quad (n_F)^{1/2} = \frac{n_F (n_A + n_{CA})}{2(n_F + n_{CF})}$$

No siempre es posible encontrar un diseño que satisfaga exactamente la ecuación (8.4). Por ejemplo, si  $k = 3$ ,  $n_F = 8$  y  $n_A = 6$ , la ecuación (8.4) se reduce a

$$\begin{aligned} (8)^{1/2} &= \frac{8(6 + n_{CA})}{2(8 + n_{CF})} \\ 2.83 &= \frac{48 + 8n_{CA}}{16 + 2n_{CF}}. \end{aligned}$$

Es imposible encontrar valores de  $n_{CA}$  y  $n_{CF}$  que satisfagan exactamente esta última ecuación. Sin embargo, observe que si  $n_{CF} = 3$  y  $n_{CA} = 2$ , entonces el segundo miembro es

$$\frac{48 + 8(2)}{16 + 2(3)} = 2.91$$

por lo que el diseño se separa en bloques casi ortogonales. En la práctica podría relajarse un tanto el requerimiento de la rotabilidad o bien el de la formación de bloques ortogonales sin ninguna pérdida importante de información.

El diseño central compuesto es muy versátil en cuanto a su capacidad para incorporar la formación de bloques. Si  $k$  es lo suficientemente grande, la porción factorial del diseño puede dividirse en dos o más bloques. (El número de bloques factoriales debe ser una potencia de 2, con la porción axial formando un solo bloque). En la tabla 11-11 se presentan varias disposiciones útiles de la formación de bloques para el diseño central compuesto.

Tabla 11.11 Algunos diseños centrales compuestos rotables y casi rotables que se separan en bloques ortogonales

$k$	2	3	4	5	$\frac{5}{2}$ Rep.	6	$\frac{6}{2}$ Rep.	7	$\frac{7}{2}$ Rep.
Bloques factoriales									
$n_F$	4	8	16	32	16	64	32	128	64
No. de bloques	1	2	2	4	1	8	2	16	8
No. de puntos/bloque	4	4	8	8	16	8	16	8	8
No. total de puntos	3	2	2	2	6	1	4	1	1
	7	6	10	10	22	9	20	9	9
Bloque axial									
$n_F$	4	6	8	10	10	12	12	14	14
$n_{CA}$	3	2	2	4	1	6	2	11	4
No. de puntos/bloque	7	8	10	14	11	18	14	25	18
No. total de puntos	14	20	30	54	33	90	54	169	80
valores de $\alpha$									
Separación en bloques ortogonales	1.4142	1.6330	2.0000	2.3664	2.0000	2.8284	2.3664	3.3636	2.3664
Rotabilidad	1.4142	1.6818	2.0000	2.3784	2.0000	2.8284	2.3784	3.3333	2.3664

Cabe destacar dos puntos importantes acerca del análisis de varianza cuando el diseño de superficie de respuesta se ha corrido en bloques. El primero se refiere al uso de los puntos centrales para calcular una estimación del error puro. Sólo los puntos centrales que se corren en el mismo bloque pueden considerarse como réplicas, por lo que el término del error puro sólo puede calcularse dentro de cada bloque. Si la variabilidad es consistente en todos los bloques, entonces estas estimaciones del error puro podrían agruparse. El segundo punto se refiere al efecto de bloque. Si el diseño se forma de bloques ortogonales en  $m$  bloques, la suma de cuadrados de los bloques es

$$SS_{bloques} = \sum_{b=1}^m \frac{B_b^2}{n_b} - \frac{G^2}{N}$$

donde  $B_b$  es el total de las  $n_b$  observaciones en el bloque  $b$ -ésimo y  $G$  es el gran total de las  $N$  observaciones en los  $m$  bloques. Cuando los bloques no son exactamente ortogonales, puede usarse la prueba general de la significación de la regresión (el método de la "suma de cuadrados extra").

Diseños (óptimos) generados por computadora. Los diseños estándares de superficie de respuesta estudiados en las secciones anteriores, como el diseño central compuesto y el diseño de Box-Behnken y sus variantes (como el cubo con centros en las caras), son de uso generalizado porque son diseños bastante generales y flexibles. Si la región experimental es un cubo o una esfera, de manera típica existe un diseño de superficie de respuesta que será aplicable al problema. Sin embargo, ocasionalmente un experimentador se encuentra con una situación en la que el diseño estándar de superficie de respuesta puede no ser una elección obvia. Los diseños generados por computadora son una alternativa por considerar en estos casos.

Hay tres situaciones en las que puede ser apropiado algún tipo de diseño generado por computadora

- (1) *Una región experimental irregular.* Si la región de interés del experimento no es un cubo o una esfera, los diseños estándares quizá no sean la mejor elección. Las regiones de interés irregulares ocurren con bastante frecuencia. Por ejemplo, un experimentador está investigando las propiedades de un adhesivo particular. El adhesivo se aplica a dos piezas y después se cura a una temperatura elevada. Los dos factores de interés son la cantidad de adhesivo aplicada y la temperatura de curado. En los rangos de estos dos factores, tomados como  $-1$  a  $+1$  en la escala de la variable codificada usual, el experimentador sabe que si se aplica muy poco adhesivo y la temperatura de curado es muy baja, las piezas no se pegarán satisfactoriamente. En términos de las variables codificadas, esto lleva a una restricción sobre las variables de diseño, por ejemplo

$$-1.5 \leq x_1 + x_2$$

donde  $x_1$  representa la cantidad aplicada de adhesivo y  $x_2$  la temperatura. Además, si la temperatura es demasiado elevada y se aplica mucho adhesivo, las piezas resultarán dañadas por la fatiga térmica o bien ocurrirá un pegado inadecuado. Por lo tanto, hay otra restricción sobre los niveles de los factores

$$x_1 + x_2 \leq 1$$

En la figura 11-26 se muestra la región experimental que resulta de aplicar estas restricciones. Observe que las restricciones eliminan de hecho dos de los vértices del cuadrado, produciendo una región experimental irregular (en ocasiones a estas regiones irregulares se les llama "latas abolladas"). No existe ningún diseño de superficie de respuesta estándar que se ajuste exactamente a esta región.

- (2) *Un modelo no estándar.* Por lo general, un experimentador elige un modelo de superficie de respuesta de primer o de segundo orden, consciente de que este modelo empírico es una aproximación del verdadero mecanismo subyacente. Sin embargo, en ocasiones el experimentador puede tener un conocimiento o idea especial acerca del proceso bajo estudio que puede sugerir un modelo no estándar. Por ejemplo, el modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{112} x_1^2 x_2 + \beta_{1112} x_1^3 x_2 + \varepsilon$$

puede ser de interés. El experimentador estaría interesado en obtener un diseño eficiente para ajustar este modelo reducido de cuarto grado. Como otra ilustración, en ocasiones se encuentran problemas de superficie de respuesta en los que algunos de los factores del diseño son variables categóricas. No hay diseños de superficie de respuesta estándares para esta situación.

- (3) *Requerimientos inusuales para el tamaño de la muestra.* Ocasionalmente, un experimentador quizá necesite reducir el número de corridas requeridas en un diseño estándar de superficie de respuesta. Por ejemplo, suponga que se pretende ajustar un modelo de segundo orden en cuatro variables. El diseño central compuesto para esta situación requiere entre 28 y 30 corridas, dependiendo del número de puntos centrales seleccionados. Sin embargo, el modelo sólo tiene 15 términos. Si las corridas tienen un costo muy elevado o se llevan mucho tiempo, el experimentador querrá un diseño con menos ensayos. Aun cuando los diseños generados por computadora pueden usarse para este fin, por lo general se cuenta con enfoques mejores. Por ejemplo, puede construirse un diseño compuesto pequeño para cuatro factores con 20 corridas, incluyendo cuatro puntos centrales, y también se cuenta con un diseño híbrido con apenas 16 corridas. Éstas son en general elecciones superiores al uso de un diseño generado por computadora para reducir el número de ensayos.

8.1.5. *Diseño robusto.* A lo largo de éstas notas se ha hecho hincapié en la importancia del uso de experimentos diseñados estadísticamente en el proyecto, desarrollo y mejoramiento de productos y procesos. A partir de la década de 1980, los ingenieros y científicos han adquirido la conciencia creciente de los beneficios del uso de experimentos diseñados y, en consecuencia, ha habido muchas áreas de aplicaciones nuevas. Una de las más importantes de éstas es el diseño robusto, donde la atención se centra en uno o más de los siguientes puntos:

- (1) El diseño de sistemas (productos o procesos) que no sean sensibles a factores ambientales que puedan afectar el desempeño una vez que, el sistema se ha desplegado en el campo. Un ejemplo es la formulación de una pintura para exteriores que debe tener gran duración cuando se exponga a una variedad de condiciones climáticas. Puesto que las condiciones climáticas

no son del todo predecibles, y ciertamente no son constantes, el responsable de la formulación del producto quiere que éste sea robusto contra un amplio rango de factores de temperatura, humedad y precipitación pluvial que afectan el desgaste y acabado de la pintura.

- (2) El diseño de productos para que no sean sensibles a la variabilidad transmitida por los componentes del sistema. Un ejemplo es el diseño de un amplificador electrónico para que el voltaje de salida esté tan cerca como sea posible del valor nominal deseado, independientemente de la variabilidad de los parámetros eléctricos de los resistores, transistores y fuentes de poder que son los componentes del aparato.
- (3) El diseño de procesos para que el producto manufacturado esté tan cerca como sea posible de las especificaciones nominales, aun cuando sea imposible controlar con toda precisión algunas variables del proceso (como la temperatura) o las características de las materias primas.
- (4) Determinar las condiciones de operación de un proceso para que las características críticas del producto estén tan cerca como sea posible del valor objetivo deseado y la variabilidad en torno a este objetivo se minimice. Ejemplos de este tipo de problema ocurren con frecuencia. Uno de ellos sucede en la manufactura de semiconductores, donde sería deseable que el espesor del óxido de una oblea estuviera lo más cerca posible del espesor objetivo promedio, así como que la variabilidad del espesor a lo largo de la oblea (una medida de uniformidad) fuese lo más pequeña posible.

A principios de la década de 1980, el ingeniero japonés Genichi Taguchi introdujo un enfoque para resolver problemas de este tipo, a los que se hace referencia de manera conjunta como el problema del *diseño paramétrico robusto* (*RPD*, por sus siglas en inglés). Su enfoque se basó en la clasificación de las variables de un proceso o producto como variables de control (o controlables) y variables de ruido (o no controlables) para después encontrar los ajustes de las variables controlables que minimizan la variabilidad transmitida a la respuesta por las variables no controlables. Se establece el supuesto de que aun cuando los factores de ruido no son controlables en el sistema a gran escala, pueden controlarse para los fines de un experimento. Taguchi introdujo algunos métodos estadísticos novedosos y ciertas variantes de las técnicas establecidas como parte de este procedimiento *RPD*. Hizo uso de diseños factoriales altamente fraccionados y otros tipos de diseños fraccionados obtenidos a partir de arreglos ortogonales. Su metodología generó múltiples debates y controversias. Parte de la polémica surgió porque la metodología de Taguchi fue defendida en Occidente inicialmente (y principalmente) por empresarios, y la ciencia estadística subyacente había sido revisada adecuadamente por los especialistas. Para finales de la década de 1980, los resultados de una revisión muy completa indicaron que aun cuando los conceptos de ingeniería de Taguchi y el objetivo global del *RPD* tenían bases sólidas, había problemas de fondo con esta estrategia experimental y con los métodos para el análisis de datos. Para detalles específicos de estos temas, ver Box [12d], Box Bisgaard y Fung [14], Hunter [59a, b], Montgomery [80b], Myers y Montgomery [85a] y Pignatielli y Ramberg [94]. Muchas de estas preocupaciones se encuentran resumidas también en el amplio panel de en publicado en *Technometrics* (ver Nair, et al. [86]).





5	2	2	3	1	19.7	18.6	19.4	25.1	25.6	21.4	27.5
25.3	6	2	3	1	2	16.2	16.3	20.0	19.8	14.7	19.6
24.7	7	3	1	3	2	16.4	19.1	18.4	23.6	16.8	18.6
21.6	8	3	2	1	3	14.2	15.6	15.1	16.8	17.8	19.6
24.2	9	3	3	2	1	16.1	19.9	19.3	17.3	23.1	22.7
28.6											

El enfoque de la superficie de respuesta para el diseño robusto. Como se señaló en la sección anterior, las interacciones entre los factores controlables y los de ruido - son la clave en un problema de diseño robusto. Por lo tanto, es lógico usar un modelo de respuesta que incluya tanto a los factores controlables como a los factores de ruido y sus interacciones. Para ilustrar, suponga que se tienen dos factores controlables  $x_1$  y  $x_2$  y un solo factor de ruido  $z_1$ . Se supone que tanto los factores controlables como el de ruido se expresan como las variables codificadas usuales (es decir, tienen su centro en -ero y tienen límites inferior y superior  $\pm a$ ). Si quiere considerarse un modelo de primer orden que incluya las variables controlables, un modelo lógico es

$$(8.5) \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \gamma_1 z_1 + \delta_{11} x_1 z_1 + \delta_{21} x_2 z_1 + \varepsilon$$

Observe que este modelo incluye los efectos principales de ambos factores controlables, su interacción, el efecto principal de la variable de ruido y las dos interacciones entre las variables controlables y la de ruido. A este tipo de modelo, el cual incorpora a las variables controlables y las de ruido, suele llamársele *•: modelo de respuesta o de reacción*. Excepto cuando al menos uno de los coeficientes de regresión  $\delta_{11}$  y  $\delta_{21}$  sea diferente de cero, no habrá ningún problema de diseño robusto.

Una ventaja importante del enfoque del modelo de respuesta es que tanto los factores control como los factores de ruido pueden colocarse en un solo diseño experimental; es decir, puede evitarse la estructura de los arreglos interior y exterior del enfoque de Taguchi. Al diseño que contiene tanto los factores controlables como los de ruido suele llamársele *diseño de arreglo combinado*.

Se supone que las variables de ruido son aleatorias, aun cuando son controlables para los fines de un experimento. Específicamente, se supone que las variables de ruido están expresadas en unidades codificadas, que tienen valor esperado cero, varianza  $\sigma_z^2$ , y que si hay varias variables de ruido, tienen covarianzas cero. Bajo estos supuestos es sencillo encontrar un modelo respuesta media tomando el valor esperado de  $y$  en la ecuación (8.5). Se obtiene así

$$E_z(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

donde el subíndice  $z$  del operador expectativa es un recordatorio para tomar el valor esperado con respecto a ambas variables aleatorias de la ecuación (8.5),  $z_1$  y  $\varepsilon$ . Para encontrar un modelo de la varianza de la respuesta  $y$  se usa el enfoque de la transmisión del error. Primero, el modelo de respuesta de la ecuación (8.5) se expande en una serie de Taylor de primer orden alrededor de  $z_1 = 0$ . Se obtiene

así:

$$\begin{aligned} y &\cong y_{z=0} + \frac{dy}{dz_1} (z_1 - 0) + R + \varepsilon \\ &\cong \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + R + \varepsilon \end{aligned}$$

donde  $R$  es el término del residuo de la serie de Taylor. Como es común en la práctica, se ignorará el término del residuo. Ahora puede obtenerse la varianza de  $y$  aplicando el operador varianza en esta última expresión (sin  $R$ ). El modelo para la varianza resultante es:

$$V_z(y) = \sigma_z^2 (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2)^2 + \sigma^2.$$

De nueva cuenta se ha usado el subíndice  $z$  en el operador varianza como recordatorio de que tanto  $z_1$  como  $\varepsilon$  son variables aleatorias.

Se han derivado modelos simples para la media y la varianza de la variable de respuesta de interés. Observe lo siguiente:

- (1) Los modelos de la media y la varianza incluyen únicamente las variables controlables. Esto significa que es potencialmente posible fijar las variables controlables para alcanzar un valor objetivo de la media y minimizar la variabilidad transmitida por la variable de ruido.
- (2) Aun cuando en el modelo de la varianza intervienen sólo las variables controlables, incluye así mismo los coeficientes de regresión de la interacción entre las variables controlables y la de ruido. Es así como la variable de ruido influye en la respuesta.
- (3) El modelo de la varianza es una función cuadrática de las variables controlables.
- (4) El modelo de la varianza (dejando de lado  $\sigma^2$ ) es sólo el cuadrado de la pendiente del modelo de respuesta ajustado en la dirección de la variable de ruido.

Para dar un uso operacional a estos modelos sería necesario:

- (1) Realizar un experimento y ajustar un modelo de respuesta apropiado, tal como la ecuación (8.5)
- (2) Sustituir los coeficientes de regresión desconocidos en los modelos de la media y la varianza con sus estimaciones de mínimos cuadrados del modelo de la respuesta o de reacción, y sustituir la  $\sigma^2$  del modelo de la varianza con el cuadrado medio de los residuales que se encontró cuando se ajustó el modelo de respuesta.
- (3) Optimizar los modelos de la media y la varianza utilizando los métodos estándares de optimización de respuestas múltiples.

Es muy sencillo generalizar estos resultados. Suponga que hay  $k$  variables controlables y  $r$  variables de ruido. El modelo de respuesta general que incluye estas variables se escribirá como

$$(8.6) \quad y(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) + h(\mathbf{x}, \mathbf{z}) + \varepsilon$$

donde  $f(\mathbf{x})$  es la porción del modelo que incluye sólo las variables controlables y  $h(\mathbf{x}, \mathbf{z})$  son los términos que incluyen los efectos principales de los factores de ruido y las interacciones entre los factores controlables y los de ruido. De manera

típica, la estructura de  $h(\mathbf{x}, \mathbf{z})$  es:

$$h(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^r \gamma_i z_i + \sum_{i=1}^k \sum_{j=1}^r \delta_{ij} x_i z_j.$$

La estructura de  $f(\mathbf{x})$  dependerá de cuál sea el tipo de modelo que el experimentador considere apropiado para las variables controlables. Las elecciones lógicas son el modelo de primer orden con interacción y el modelo de segundo orden. Si se supone que las variables de ruido tienen media cero, varianza  $\sigma_z^2$  y covarianzas cero, y que las variables de ruido y los errores aleatorios  $\varepsilon$  tienen covarianzas cero, entonces el modelo de la media para la respuesta es:

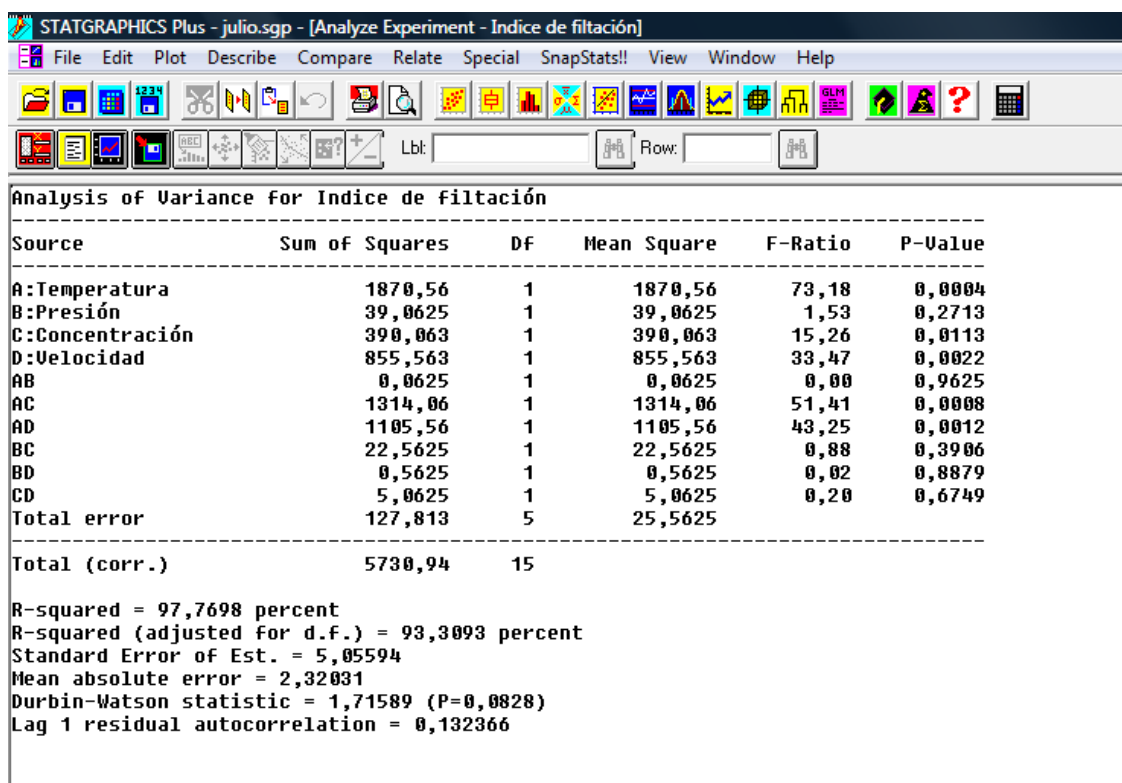
$$(8.7) \quad E_z [y(\mathbf{x}, \mathbf{z})] = f(\mathbf{x})$$

y el modelo de la varianza para la respuesta es:

$$(8.8) \quad V_z [y(\mathbf{x}, \mathbf{z})] = \sigma_z^2 \sum_{i=1}^r \left[ \frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial z_i} \right]^2 + \sigma^2$$

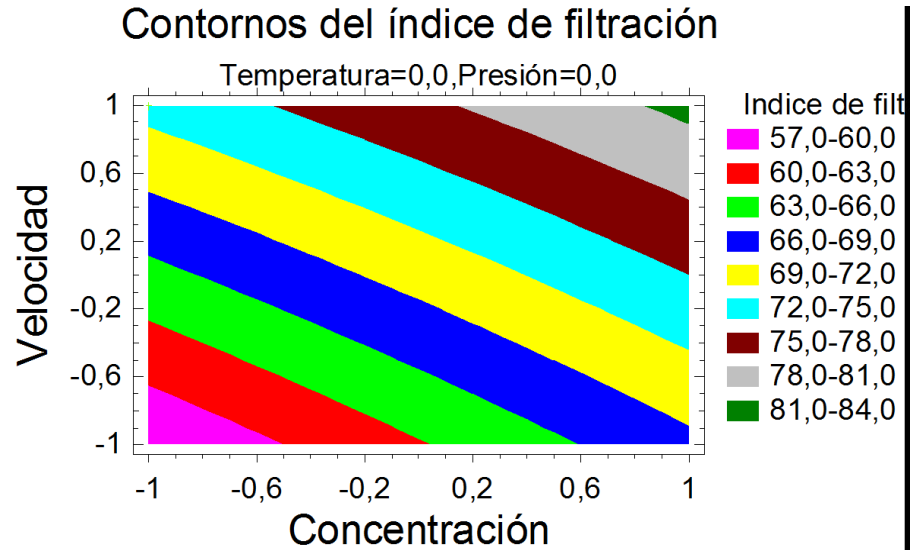
Myers y Montgomery [85a] presentan una forma un tanto más general de la ecuación (8.8) basada en la aplicación directa de un operador de varianza condicional al modelo de respuesta.

Ejemplo ilustrativo 1. Un producto químico se fabrica en un envase presurizado. Se lleva a cabo un experimento factorial la planta piloto para estudiar los factores que se piensan influyen en el índice de filtración de este producto. los cuatro factores son la temperatura (A), la presión (B), la concentración del formaldehído (C) y la velocidad de agitación (D). Cada factor está presente con dos niveles. la matriz de diseño y los datos de la respuesta obtenidos de una sola réplica del experimento  $2^4$  se muestra a continuación en la siguiente tabla. Las 16 corridas se hacen de manera aleatoria. El ingeniero del proceso está interesado en maximizar el índice de filtración, las condiciones actuales del proceso producen índices de filtración de alrededor de 75 gal/h. Asimismo, en el proceso actual la concentración de formaldehído, el factor C, se usa en un nivel alto. Al ingeniero le gustaría reducir la concentración de formaldehído lo más posible, pero no ha podido hecerlo por que siempre produce índices de filtración más bajos. Se supondrá que el factor A, la temperatura, es difícil de controlar en el proceso a gran escala, pero que puede controlarse durante el experimento (el cual se llevó a cabo en una planta piloto). Los otros tres factores, la presión (B), la concentración (C) y la velocidad de agitación (D), son fáciles de controlar. Por lo tanto, el factor de ruido  $z_1$  es la temperatura, y las variables controlables  $x_1, x_2$ , y  $x_3$  son la presión, la concentración y la velocidad de agitación, respectivamente. Puesto que tanto los factores controlables como el factor de ruido están en el mismo diseño, el diseño factorial  $2^4$  utilizado en experimento es un ejemplo de un *diseño de arreglo combinado*.



Experimento del índice de filtración						
Número de corrida	Factores				Etiqueta	índice de filtración
	A	B	C	D		
1	—	—	—	—	(1)	45
2	+	—	—	—	a	71
3	—	+	—	—	b	48
4	+	+	—	—	ab	65
5	—	—	+	—	c	68
6	+	—	+	—	ac	60
7	—	+	+	—	bc	80
8	+	+	+	—	abc	65
9	—	—	—	+	d	43
10	+	—	—	+	ad	100
11	—	+	—	+	bd	45
12	+	+	—	+	abd	104
13	—	—	+	+	bd	75
14	+	—	+	+	acd	86
15	—	+	+	+	bcd	70
16	+	+	+	+	abcd	96

El análisis inicial de los datos muestra en la ANOVA que el factor B y las interacciones AB, BC, BD y CD no son significativas y el factor supuesto de ruido A es altamente significativo.



Los coeficientes para el diseño de arreglo combinado, haciendo uso del mejor anova

Coeficientes de regresion para el índice de filtración

constant	70,0625	
A:Temperatura	10,8125	$z_1$
C:Concentración	4,9375	$x_2$
D:Velocidad	7,3125	$x_3$
AC	-9,0625	$x_2 z_1$
AD	8,3125	$x_3 z_1$

Utilizando los resultados encontrados, el modelo de respuesta es:

$$\hat{y}(\mathbf{x}, z_1) = 70,06 + 10,81z_1 + 4,94x_2 + 7,31x_3 - 9,06x_2z_1 + 8,3125x_3z_1$$

y loss modelos de la media y la varianza son:

$$\begin{aligned} E_z [y(\mathbf{x}, z_1)] &= 70,06 + 4,94x_2 + 7,31x_3 \\ V_z [y(\mathbf{x}, z_1)] &= \sigma_z^2 (10,81 - 9,06x_2 + 8,31x_3)^2 + \sigma^2 \\ V_z [y(\mathbf{x}, z_1)] &= \sigma_z^2 (116,91 + 82,08x_2^2 + 69,06x_3^2 - 195,88x_2 \\ &\quad + 179,66x_3 - 150,58x_2x_3) + \sigma^2 \end{aligned}$$

respectivamente. Suponga ahora que los niveles bajo y alto de la variable de ruido, temperatura, se corrieron a una desviación estándar a ambos lados de su valor típico o promedio, de tal modo que  $\sigma_z^2 = 1$ , y que se usa  $\hat{\sigma}^2 = 19,51$  (éste es el cuadrado medio de los residuales obtenido al ajustar el modelo de respuesta). Por lo tanto, el modelo de la varianza queda como

$$V_z [y(x, z_1)] = 136,42 - 195,88x_2 + 179,66x_3 - 150,58x_2x_3 + 82,08x_2^2 + 69,06x_3^2$$

A continuación se presenta la gráfica de contorno del paquete de software Statgraphics de los contornos de respuesta del modelo de la media. Para construir esta

gráfica se fijó el factor de ruido (temperatura) en cero y el factor controlable no significativo (presión) también en cero. Observe que la rapidez de filtración promedio se incrementa cuando tanto la concentración como la velocidad de agitación se incrementan. El ejemplo ilustra el uso de un modelo de primer orden con interacción como el modelo para los factores controlables,  $f(\mathbf{x})$ . Se presenta ahora un ejemplo adaptado de Montgomery [80b] que incluye un modelo de segundo orden.

Ejemplo ilustrativo 2. En una fábrica de semiconductores se realizó un experimento que incluyó dos variables controlables y tres variables de ruido. En la tabla a continuación se muestra el diseño de arreglo combinado utilizado por los experimentadores. Este diseño es una variante de 23 corridas de un diseño central compuesto que se creó empezando con un DCC estándar para cinco factores (la porción del cubo es un diseño  $2^{5-1}$ ) y eliminando las corridas axiales asociadas con las tres variables de ruido. Este diseño soportará un modelo de respuesta que tiene un modelo de segundo orden en las variables controlables, los efectos principales de las tres variables de ruido y las interacciones entre los factores controlables y los de ruido. El modelo de respuesta ajustado es

$$\begin{aligned}\hat{y}(\mathbf{x}, \mathbf{z}) = & 30.37 - 2.92x_1 - 4.13x_2 + 2.60x_1^2 + 2.18x_2^2 + 2.87x_1x_2 + 2.73z_1 - 2.33z_2 \\ & + 2.33z_3 - 0.27x_1z_1 + 0.89x_1z_2 + 2.58x_1z_3 + 2.01x_2z_1 - 1.43x_2z_2 + 1.56x_2z_3\end{aligned}$$

Los modelos de la media y la varianza son

$$\begin{aligned}E_z[y(\mathbf{x}, \mathbf{z})] &= 30.37 - 2.92x_1 - 4.13x_2 + 2.60x_1^2 + 2.18x_2^2 + 2.87x_1x_2 \\ V_z[y(\mathbf{x}, z_1)] &= \sigma_z^2(10, 81 - 9.06x_2 + 8, 31x_3)^2 + \sigma^2 \\ V_z[y(\mathbf{x}, z_1)] &= 19.26 + 3.20x_1 + 12.45x_2 + 7.52x_1^2 + 8.52x_2^2 + 2.21x_1x_2\end{aligned}$$

donde se han sustituido las estimaciones de los parámetros del modelo de respuesta ajustado en las ecuaciones de los modelos de la media y la varianza y, como en el ejemplo anterior, se supone que  $\sigma_z^2 = 1$ . En las figuras 11-45 y 11-46 se presentan las gráficas de contorno de la media generada a partir de estos modelos. En este problema es deseable mantenerla media del proceso abajo de 30. Al inspeccionar las figuras 11-45 y 11-46, es claro que se necesitará hacer un ajuste si se quiere hacer pequeña la varianza del proceso. Puesto que sólo hay dos variables controlables, una forma lógica de llegar a este arreglo es superponer los contornos de la respuesta media constante y la varianza constante, como se muestra en la figura 11-47. Esta gráfica muestra los contornos para los que la media del proceso es menor o igual que 30 y la desviación estándar del proceso es menor o igual que 5. La región delimitada por estos contornos representaría una zona de operación típica de respuesta media baja y varianza del proceso baja.

Experimento con arreglo combinado con dos variables variables controlables y tres variables de ruido						
corrida	$x_1$	$x_2$	$z_1$	$z_2$	$z_3$	$y$
1	-1.00	-1.00	-1.00	-1.00	1.00	44.2
2	1.00	-1.00	-1.00	-1.00	-1.00	30.0
3	-1.00	1.00	-1.00	-1.00	-1.00	30.2
4	1.00	1.00	-1.00	-1.00	1.00	35.4
5	-1.00	-1.00	1.00	-1.00	-1.00	49.8
6	1.00	-1.00	1.00	-1.00	1.00	36.3
7	-1.00	1.00	1.00	-1.00	1.00	41.3
8	1.00	1.00	1.00	-1.00	-1.00	31.4
9	-1.00	-1.00	-1.00	1.00	-1.00	43.5
10	1.00	-1.00	-1.00	1.00	1.00	36.1
11	-1.00	1.00	-1.00	1.00	1.00	22.7
12	1.00	1.00	-1.00	1.00	-1.00	16.0
13	-1.00	-1.00	1.00	1.00	1.00	43.2
14	1.00	-1.00	1.00	1.00	-1.00	30.3
14	-1.00	1.00	1.00	1.00	-1.00	30.1
16	1.00	1.00	1.00	1.00	1.00	39.2
17	-2.00	0.00	0.00	0.00	0.00	46.1
18	2.00	0.00	0.00	0.00	0.00	36.1
19	0.00	-2.00	0.00	0.00	0.00	47.4
20	0.00	2.00	0.00	0.00	0.00	31.5
21	0.00	0.00	0.00	0.00	0.00	30.8
22	0.00	0.00	0.00	0.00	0.00	30.7
23	0.00	0.00	0.00	0.00	0.00	31.0

### 9. Ejercicios: Metodología de superficie de respuesta

EJERCICIO 117. Un ingeniero químico recolectó los siguientes datos. la respuesta  $Y$  es el tiempo de filtración,  $x_1$  es la temperatura y  $x_2$  es la presión. Ajustar un modelo de segundo orden. (a) ¿Qué condiciones de operación recomendarían si el objetivo es minimizar el tiempo de filtración? (b) ¿Qué condiciones de operación recomendarían si el objetivo es operar el proceso con una velocidad de filtración

media muy próxima a 46?

$x_1$	$x_2$	$Y$
-1	-1	54
-1	1	45
1	-1	32
1	1	47
-1.414	0	50
1.414	0	53
0	-1.414	47
0	1.414	51
0	0	41
0	0	39
0	0	44
0	0	42
0	0	40

EJERCICIO 118. Describa la superficie de respuesta para las siguientes formas canónicas dadas para las superficies de respuestas cuadráticas:

$$\begin{aligned} (a) \hat{y} &= 100 - Z_1^2 - 2Z_2^2 \\ (b) \hat{y} &= 50 + 2Z_1^2 \\ (c) \hat{y} &= 75 + Z_1^2 - 2Z_2^2 \end{aligned}$$

EJERCICIO 119. Supongamos el modelo ajustado dado por:

$$\hat{Y} = 60 + 1.5x_1 + 2x_2$$

(a) Determine la trayectoria de ascenso máximo dado por el modelo. (b) Encuentre cinco puntos sobre ella, utilice un paso unitario en la variable de mayor influencia.

EJERCICIO 120. 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 acada tratamiento dietético y se registraron sus aumentos de peso después de 38 días. El aumento promedio de peso para los tratamientos es el



siguiente:

Factores originales		Factores codificados		
Metionina	Caroteno	$x_1$	$x_2$	Aumento de peso
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

(a) 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. (b) Pruebe la significancia del modelo cuadrático completo, las desviaciones cuadráticas del modelo lineal, la significancia del modelo y la falta de ajuste al modelo cuadrático. ¿Cuáles son sus conclusiones? (c) ¿La superficie de respuesta tiene un máximo dentro de las coordenadas del diseño? Determine las coordenadas. (d) Calcule la ecuación canónica y describa la superficie de respuesta.

EJERCICIO 121. En el área de desarrollo de una empresa se pretende obtener un nuevo polímero de bajo peso molecular, de lograrse esto, se obtendrá un polímero que funcione como dispersante en la industria cerámica. De acuerdo con los conocimientos técnicos que se tienen, se considera que los factores críticos son: persulfato de sodio (NaPS), ácido hiposforoso ( $H_3PO_2$ ) y el isopropanol (IPA) para encontrar las condiciones óptimas de estos tres factores, se corrió el diseño de superficie de respuesta (Box-Behnken): (a) Represente en un cubo los puntos del diseño empleado. (b) Realice un análisis estadístico completo para la variable peso molecular. Obtenga: los factores más importantes, un modelo depurado, la calidad del modelo y la verificación de supuestos. (c) ¿Cuáles son las condiciones que minimizan el peso molecular? Apóyese en gráficas de contornos y en análisis de cordilleras. (d) Haga lo que se pidió en el inciso (b), pero con la variable viscosidad. (e) Encuentre una condición que de manera simultánea minimice ambas respuestas. Hágalo

gráficamente.

Tratamiento	NaPS	H <sub>3</sub> PO <sub>2</sub>	IPA	Peso mol	Viscosidad
1	0	0	0	8392	1075
2	-1	-1	0	9895	2325
3	1	-1	0	9204	1575
4	-1	1	0	7882	690
5	1	1	0	7105	420
6	-1	0	-1	8939	1188
7	1	0	-1	8548	930
8	0	0	0	8598	920
9	-1	0	1	9152	1275
10	1	0	1	8992	860
11	0	-1	-1	10504	5600
12	0	1	-1	7462	540
13	0	-1	1	9368	1225
14	0	1	1	7772	620
15	0	0	0	8440	1015

EJERCICIO 122. Supongamos el modelo ajustado dado por:

$$\hat{Y} = 8.25 + 0.0396x_1 + 0.1884x_2 - 0.4313x_1^2 - 0.3563x_2^2 + 0.3x_1x_2$$

(a) Dibuje la superficie descrita por el modelo. (b) ¿Qué tipo de superficie es? (c) Encuentre el punto estacionario y diga de que tipo es. (d) Encuentre un punto sobre la dirección de mínima pérdida a partir del punto estacionario.

EJERCICIO 123. Supongamos el modelo ajustado dado por:

$$\hat{Y} = 10.46 - 0.57x_1 + 0.18x_2 + 0.45x_3 - 0.67x_1^2 + 0.56x_2^2 - 0.27x_3^2 - 0.67x_1x_2 + 1.18x_1x_3 + 0.23x_2x_3$$

(a) Encuentre el punto estacionario y diga de que tipo es. (b) Si se obtiene un punto silla, realice un análisis de cordillera.

EJERCICIO 124. El diseño hexagonal que se presenta a continuación se usa en un experimento que tiene como objetivo ajustar un modelo de segundo orden. (a) Ajustar el modelo de segundo orden. (b) Efectuar el análisis canónico. ¿Qué tipo de superficie se ha encontrado? (c) ¿Qué condiciones de operación para  $x_1$  y  $x_2$  llevan al punto estacionario? (d) ¿Dónde se correría este proceso si el objetivo es obtener una respuesta que esté tan cerca de 65 como sea posible?

$x_1$	$x_2$	$Y$
1	0	68
0,5	$\sqrt{0,75}$	74
-0,5	$\sqrt{0,75}$	65
-1	0	60
-0,5	$-\sqrt{0,75}$	63
0,5	$-\sqrt{0,75}$	70
0	0	58
0	0	60
0	0	57
0	0	55
0	0	69

EJERCICIO 125. Se realizó un experimento  $2^2$  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  $\mu\text{l}$

Reactivo		% de droga recuperada	
A	B		
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  $\beta_{12}$  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.

EJERCICIO 126. Se ha corrido un experimento en un proceso que aplica un material de recubrimiento a una oblea. En cada corrida del experimento se fabricó una oblea y se midió varias veces el espesor del recubrimiento en varios sitios de la misma. Después se obtuvo la media  $Y_1$  y la desviación estándar  $Y_2$  de la medición

del espesor. Los Datos se muestran en la tabla siguiente:

Corrida	velocidad	presión	distancia	$Y_1$	$Y_2$
1	-1	-1	-1	24	12.5
2	0	-1	-1	120.3	8.4
3	1	-1	-1	213.7	42.8
4	-1	0	-1	86	3.5
5	0	0	-1	136.6	80.4
6	1	0	-1	340.7	16.2
7	-1	1	-1	112.3	27.6
8	0	1	-1	256.3	4.6
9	1	1	-1	271.7	23.6
10	-1	-1	0	81	0
11	0	-1	0	101.7	17.7
12	1	-1	0	357	32.9
13	-1	0	0	171.3	15
14	0	0	0	372	0
15	1	0	0	501.7	92.5
16	-1	1	0	264	63.5
17	0	1	0	427	88.6
18	1	1	0	730.7	21.1
19	-1	-1	1	220.7	133.8
20	0	-1	1	239.7	23.5
21	1	-1	1	422	18.5
22	-1	0	1	199	29.4
23	0	0	1	485.3	44.7
24	1	0	1	673.7	158.2
25	-1	1	1	176.7	55.5
26	0	1	1	501.0	138.9
27	1	1	1	1010.0	142.4

(a) ¿Qué tipo de diseño usaron los experimentadores? ¿Es ésta una buena elección del diseño para ajustar un modelo cuadrático? (b) Construir los modelos para ambas respuestas. (c) Encontrar un conjunto de condiciones óptimas que resulten en una media tan grande como sea posible con la desviación estandar menor que 60

EJERCICIO 127. Un experimentador corrió un diseño de Box-Behnken y obtuvo los siguientes resultados donde la variable de respuesta es la viscosidad de un polímero:

		Velocidad				
Nivel	Temperatura	de agitación	presión	$x_1$	$x_2$	$x_3$
Alto	200	10.0	25	+1	+1	+1
Intermedio	175	7.5	20	0	0	0
Bajo	150	5.0	15	-1	-1	-1

<i>Corrida</i>	$x_1$	$x_2$	$x_3$	$Y_1$
1	-1	-1	0	535
2	+1	-1	0	580
3	-1	+1	0	596
4	+1	+1	0	563
5	-1	0	-1	645
6	+1	0	-1	458
7	-1	0	+1	350
8	+1	0	+1	600
9	0	-1	-1	595
10	0	+1	-1	648
11	0	-1	+1	532
12	0	+1	+1	656
13	0	0	0	653
14	0	0	0	599
15	0	0	0	620

(a) Ajustar el modelo de segundo orden. (b) Efectuar el análisis canónico. ¿Qué tipo de superficie se ha encontrado? (c) ¿Qué condiciones de operación para  $x_1$ ,  $x_2$  y  $x_3$  llevan al punto estacionario? (d) ¿Dónde se correría este proceso si el objetivo es obtener una respuesta que esté tan cerca de 600 como sea posible?

EJERCICIO 128. Los datos que se presentan en la siguiente tabla fueron recopilados en un experimento para optimizar el crecimiento de cristales en función de las tres variables  $x_1, x_2, x_3$ . Se buscan valores altos de  $Y$  (rendimiento en gramos). (a) ¿Qué diseño se empleó? (b) Ajuste un modelo de segundo orden y determine los componentes significativos y más importantes. (c) ¿El modelo describe adecuadamente la superficie? Argumente. (d) Considere las dos variables independientes más importantes en el modelo y obtenga la superficie de respuesta y de contornos en función de éstas. (e) ¿Con la superficie que se encontró en el inciso anterior se puede hacer algún tipo de inferencia? Argumente. (f) Verifique los supuestos para el modelo. ¿Observa algo relevante? (g) Observe las respuestas obtenidas en los puntos centrales del diseño. ¿Esto le ayuda a explicar lo que ocurrió? Argumente.

(h) ¿A que cree que se debe el mal ajuste del modelo?

$x_1$	$x_2$	$x_3$	$Y$
-1	-1	-1	66
1	-1	-1	70
-1	1	-1	78
1	1	-1	60
-1	-1	1	80
1	-1	1	70
-1	1	1	100
1	1	1	75
0	0	-1.68	100
0	0	1.68	80
0	-1.68	0	68
0	1.68	0	63
0	0	-1.68	65
0	0	1.68	82
0	0	0	113
0	0	0	100
0	0	0	118
0	0	0	88
0	0	0	100
0	0	0	85

## Diseño Robusto

**0.1. Competencias.** Diferenciar los diseños robustos de Taguchi de los diseños clásicos, identificando y valorando las ventajas y desventajas en cada uno de ellos. Determinar la combinación de niveles de los factores controlables, en la cual los factores de ruido no afectan al proceso teniendo en cuenta que estos últimos no se controlan.

### 1. Introducción

El diseño robusto tiene su origen en las ideas del ingeniero japonés *Gen'ichi Taguchi*, quien desarrolló su propia filosofía y métodos de ingeniería de la calidad desde la década de los años cincuenta. Fue a partir del éxito de los japoneses en industrias tan importantes como la automotriz y la electrónica que Occidente comienza a fijarse en los métodos utilizados por ellos. De esta manera, los métodos de *Taguchi*, y de otros autores japoneses, se introducen en Estados Unidos a partir de la década de los años ochenta.

Taguchi hace contribuciones que van desde la definición de calidad pasando por el control estadístico de procesos (*online quality control*), hasta el diseño de experimentos para la mejora y optimización de procesos y productos. Es precisamente en el diseño de experimentos donde hace sus aportaciones más importantes, con la introducción de lo que él llama *diseño de parámetros* (Taguchi, 1987), que a la postre viene a ser lo que ahora se conoce en Occidente como *diseño robusto*.

### 2. Enfoques de *Gen'ichi Taguchi* en el diseño de experimentos

Las aplicaciones industriales del diseño experimental tuvieron inicialmente mucho mayor éxito en el Japón que en las economías occidentales. Se afirma que en gran parte el éxito de calidad de los productos japoneses se debe a estas aplicaciones, y los experimentos multifactoriales parecen ser una parte integral del proceso de diseño de las compañías de ese país (*Box et al* (1988)). En empresas occidentales también están difundiéndose las técnicas de *Taguchi*, que prestan especial atención a las consideraciones de ahorro de costos: en efecto, según algunos autores como Sullivan (1987), empresas de la talla de ITT, que han capacitado a más de mil ingenieros en estos métodos, miden el *éxito de la calidad en sus estudios de casos en términos de ahorro y no de número de unidades defectuosas*.

**2.1. La función de pérdida y el concepto de calidad.** *Gen'ichi Taguchi* realizó un gran esfuerzo para llevar a un terreno práctico el diseño experimental, introdujo, además, conceptos revolucionarios que afectaron la forma de medir la calidad y su costo. Para Taguchi, la *calidad*, antes que por la satisfacción de especificaciones, debe medirse en términos de la llamada *función de pérdida*, que establece la pérdida que la sociedad sufre como consecuencia de la mala calidad. Un producto

de calidad es para el cliente aquél que cumple con las expectativas de performance o rendimiento cada vez que lo utiliza, sin fallas y en cualquier condición o circunstancia. Los productos que no cumplen con dichas expectativas causan pérdidas, tanto para los clientes y los productores, como para, eventualmente, el resto de la sociedad.

*Para Taguchi, la calidad debe medirse en función de la pérdida que causa: mientras mayor es la pérdida que se produce, menor es la calidad. Las especificaciones también son clave para Taguchi, y calidad significa conformidad con las especificaciones. Apartarse de las especificaciones equivale a ocasionar al cliente y, en última instancia, a la sociedad, una pérdida. Taguchi se apartó de la sabiduría convencional, que suponía que calidad equivalía a producir dentro de los márgenes de tolerancia, y postuló que el costo de la mala calidad se incrementa con el alejamiento del valor de diseño, produciendo una pérdida para el cuerpo social.*

La función de pérdida vale cero cuando el desvío con respecto al parámetro objetivo es nulo y se incrementa cuadráticamente cuando los valores de los productos fabricados se acercan a los límites de tolerancia. En otras palabras, los productos cercanos a los límites de tolerancia son productos casi defectuosos y los gerentes deben trabajar para reducir la variabilidad de sus procesos de producción. La función de pérdida se define como:

$$L(y) = k(y - t)^2$$

donde:

- $L(y)$  indica la pérdida (en unidades monetarias) que sufre la sociedad;
- $k$  es una constante específica de cada caso considerado;
- $t$  es un valor objetivo que la dimensión de interés debe tener ( $t$  mide la calidad nominal o de diseño); y
- $y$  es el apartamiento que la dimensión de interés presenta con respecto al valor objetivo,  $t$ .

En contraste con el pensamiento tradicional sobre la calidad, que solo penaliza los valores de  $y$  que superan los límites inferior o superior de especificaciones (**EI**, **ES**) de tolerancia, *Taguchi* considera que todo apartamiento del valor objetivo es un costo para la sociedad y como tal debe ser penalizado.

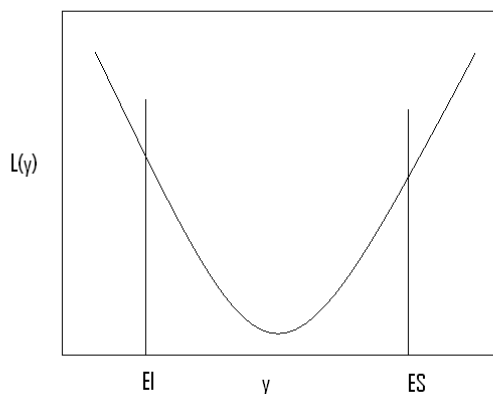
**2.2. Importancia del diseño robusto.** Para *Taguchi*, es posible incorporar la calidad en los productos desde su diseño, sin aumentar su costo; los problemas deben eliminarse en el laboratorio de diseño, no en la fábrica o en el campo. Según esta perspectiva, es necesario diseñar productos robustos que toleren variaciones en el proceso de producción y durante el servicio de mantenimiento. Los métodos estadísticos deben seleccionar los factores importantes que afectan el diseño. Taguchi establece su metodología para:

- *Diseñar productos y procesos robustos a las condiciones ambientales;*
- *Diseñar y desarrollar productos robustos a la variación en sus componentes;*
- *Minimizar la variación alrededor de un valor objetivo.*

La ingeniería de la calidad de *Taguchi* combina métodos estadísticos y de ingeniería para optimizar los procesos de diseño y fabricación de modo que aumente la calidad y se reduzcan los costos de los productos. El diseño de experimentos juega un papel esencial en el enfoque de Taguchi, pues ayuda a identificar los factores



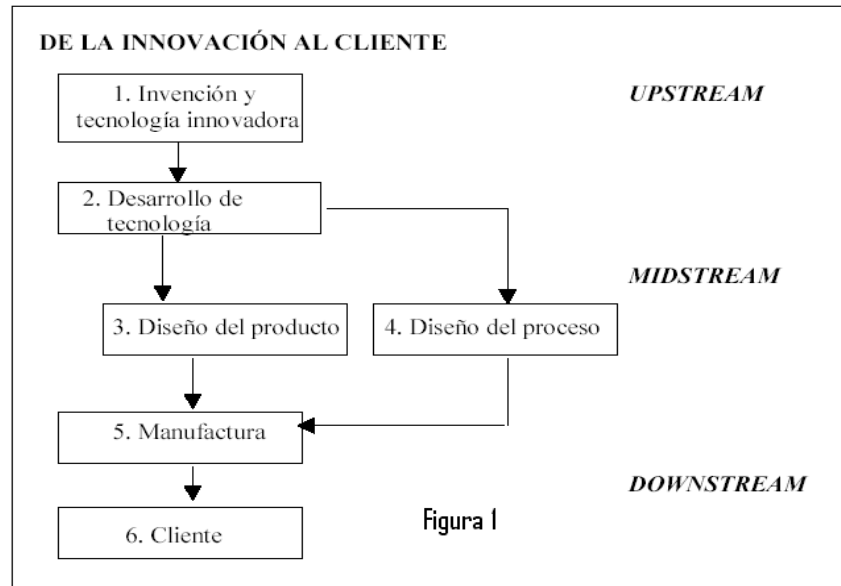
Figura 1. Función de pérdida de Taguchi.



que más intervienen en la generación de problemas de calidad o, alternativamente, los factores que más contribuyen a lograr resultados positivos. A lo largo de este proceso, se fortalece la cooperación entre diversos niveles y áreas de la empresa.

La Figura 1 muestra un esquema de actividades de las firmas desde la invención del producto hasta su llegada al mercado. El flujo se divide en tres etapas, conocidas por las expresiones inglesas *upstream*, *midstream*, y *downstream*. La etapa de *upstream* consiste en los bloques 1 y 2, la de *midstream*, en los bloques 2, 3, 4, y 5, y la de *downstream*, en los bloques 5 y 6. El bloque 2, de desarrollo de tecnología, actúa como perno entre el *upstream* y el *midstream*. El bloque 5, manufactura, cumple el mismo papel entre el *midstream* y el *downstream*. La ingeniería de la calidad enseña que las actividades para la mejora de la calidad deberían ser empujadas hacia el *upstream*, alejándolas del proceso de manufactura. El diseño experimental, y las técnicas de *Taguchi* en particular, apuntan a este fin.

La Figura 1 nos permite reflexionar sobre la responsabilidad de los gerentes en los temas de calidad. Por ejemplo, es importante destacar su papel en el diseño y la manufactura de los productos, pues allí se concentran las consecuencias del costo y la calidad de los procesos. Además, *dado que los acontecimientos que ocurren en el upstream tiene consecuencias en las etapas inferiores, cuando se intenta mejorar la calidad, y en definitiva ahorrar en terminos de costos, se debe mejorar la tecnología genérica, en vez de estudiar las formas de mejorar producto por producto*. De esta manera, los avances tecnológicos son aplicables a cualquier tipo de producto que se quiera fabricar, y no derrochan recursos innovando en cada producto nuevo que se lance al mercado. Los ingenieros, en suma, deben desarrollar tecnologías robustas en la etapa de investigación y desarrollo, tecnologías que sean fácilmente transferibles a nuevos procesos y productos. Entonces, las innovaciones más valiosas son las que se producen *upstream*; no hay que descuidar la prevención de problemas de calidad *downstream* pues es aquí y no en el laboratorio donde aparecen los factores no controlables.



**2.3. El enfoque de Taguchi.** Taguchi presenta tres etapas en el diseño de un producto o de un proceso:

1. *Diseño del sistema*
2. *Diseño de los parámetros y*
3. *Diseño de las tolerancias.*

- En el *diseño del sistema* se determina la configuración básica de los componentes. Por ejemplo, en una línea de acondicionamiento de comprimidos, el diseño del sistema incluye la determinación de los materiales y el diseño del sistema de líneas que realizarán el empaque, con todos sus componentes.
- En el *diseño de los parámetros*, se determinan los niveles o valores de los factores controlables para minimizar el efecto de los factores incontrolables en las características del producto terminado, es decir, en nuestro ejemplo, los comprimidos envasados, con su prospecto y caja.
- El *diseño de las tolerancias* apunta a reducir la varianza en las características del producto terminado cuando la reducción lograda en el diseño de los parámetros no es suficiente.

El diseño experimental debería aplicarse fundamentalmente al *diseño de los parámetros y al diseño de las tolerancias*. Fue un logro de Taguchi el haber destacado la importancia de aplicar el diseño experimental en las etapas *upstream* del proceso de creación, fabricación y puesta en el mercado de un nuevo producto. También es clave en su enfoque (aunque no haya sido inventado por él) el concepto de *robustez*; *un producto o un proceso es robusto cuando su funcionamiento sea consistente al exponerse a las condiciones cambiantes del medio*.

**2.4. Métodos recomendados por Taguchi.** Taguchi divide los factores de un experimento en factores controlables y factores incontrolables, o ruido; y recomienda seleccionar dos diseños experimentales, uno para los factores controlables y otro para el ruido en un esquema de dos componentes:

- *el arreglo de los factores controlables (arreglo interior);*
- *y el arreglo de los factores no controlables (arreglo exterior).*

Para el análisis de datos, *Taguchi* recomienda evaluar en el arreglo interior la respuesta promedio de cada corrida del experimento y analizar la variación de los resultados con un cociente *señal/ruido* apropiado. Estos cocientes se derivan de la función de pérdida cuadrática presentada anteriormente. Se consideran óptimos los niveles de los factores que maximicen un cociente *señal/ruido* adecuado. Estos cocientes difieren según que el objetivo del experimento sea reducir la variabilidad alrededor de un valor objetivo determinado, o producir un efecto resultante lo mayor posible, o inversamente, producir el menor efecto posible.

Para *Taguchi*, el uso de los cocientes *señal/ruido* elimina en general, la necesidad de examinar las interacciones entre los factores controlables y los factores de ruido. Finalmente, dado que los arreglos no siempre se corren completos, por razones de economía, es posible realizar experimentos confirmatorios.

## 2.5. Crítica de los diseños experimentales y el análisis de datos.

- (1) *Complejidad de sus diseños experimentales.* Los enfoques de *Taguchi* para el diseño de parámetros utilizan diseños ortogonales, muchos de los cuales son experimentos factoriales fraccionales de dos niveles. Otros diseños, sin embargo, son del tipo factorial fraccional de tres niveles (*alto-medio-bajo*) y tienen estructuras muy complejas.
- (2) Falta de un mecanismo para tratar con la interacción potencial entre los factores controlables de un proceso.
- (3) La estructura de arreglos internos y externos lleva generalmente a experimentos muy grandes que, con los enfoques tradicionales, podrían realizarse más eficientemente. Autores como Montgomery (1991) sostienen que el enfoque de arreglos internos y externos es innecesario, y podría reemplazarse con un único diseño que incluya tanto a los factores controlables como a los no controlables. El diseño único reduce en general el tamaño del experimento y, además, puede aportar una mejor comprensión del proceso subyacente.
- (4) El uso del cociente *señal/ruido* ha sido criticado por su ineficiencia para detectar los efectos de dispersión, aunque pueda servir para identificar los factores que afectan al promedio de una variable de interés.

**2.6. Comparación entre el diseño experimental clásico y los métodos de Taguchi.** Tanto el diseño de experimentos como los métodos de Taguchi sirven para optimizar procesos que tienen entradas controlables y salidas medibles. Ambas metodologías trabajan con múltiples entradas que afectan a la salida por medir. Su diferencia fundamental radica en la forma en que cada método maneja las interacciones entre los factores de entrada. Mientras en el diseño experimental clásico se prueba con todas las combinaciones de niveles de entrada, o una fracción simétrica de las mismas, con el método de Taguchi se trabaja con una porción pequeña de las posibles combinaciones de factores de entrada, pero de una manera que permite calcular los efectos de todos los factores de entrada en el resultado de interés.

- **Métodos de Taguchi**

- (1) Suponen que los usuarios de estos métodos poseen cierto entendimiento del proceso estudiado y de las interacciones subyacentes entre las entradas.
- (2) Fueron inventados "por ingenieros para ingenieros".
- (3) Según Taguchi, algunas variables están bajo control, y otras no (factores de ruido). Uno de los primeros pasos en los métodos de Taguchi es el uso de los conocimientos previos del proceso para identificar los factores controlables y de ruido más significativos.
- (4) Taguchi recomienda un cociente *señal/ruido* para representar la robustez, (aunque un simple análisis de la varianza o el desvío standard bastaría).
- (5) En contraste con el diseño de experimentos clásico, Taguchi sugiere una ronda final de experimentos de confirmación.
- (6) Mientras que el diseño de experimentos clásico puede ignorar o no considerar explícitamente los factores de ruido, el diseño Taguchi los usa para contrastar la robustez del sistema, así como para encontrar las entradas óptimas.

- **Diseño de experimentos clásicos**

- (1) Supone que los usuarios no tienen un conocimiento profundo de los mecanismos que gobiernan el proceso investigado.
- (2) Fue inventado "por científicos para científicos".
- (3) Presupone que todos los inputs deben interactuar entre sí. Por esto, cuando no se tiene definido de antemano qué interacciones son las relevantes, el diseño experimental es la opción que conviene elegir.
- (4) En el diseño de experimentos, no sólo se estudian todas las interacciones, sino que ellas se estudian al mismo tiempo, en una misma ronda de pruebas.

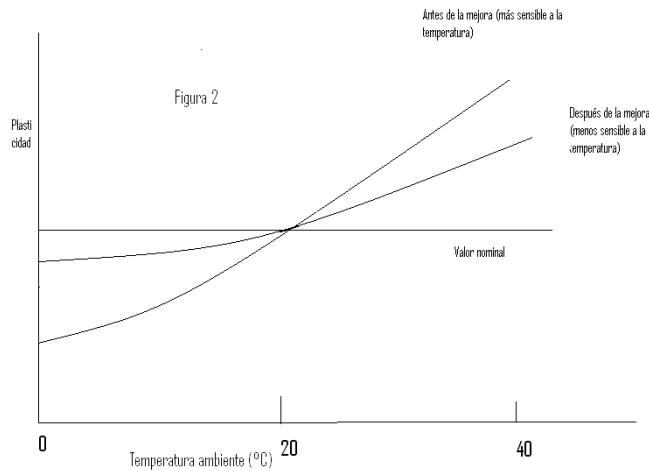
### 3. El concepto de robustez

Por *diseño robusto* se debe entender un experimento en el cual existen factores de ruido (no controlables), considerados de manera explícita o implícita, cuyo efecto se pretende minimizar de manera indirecta (o sea sin controlarlos directamente), encontrando la combinación de niveles de los factores de proceso que sí se pueden controlar, en la cual el efecto de dichos factores de ruido es mínimo.

- *En un experimento robusto se trata de lograr que el producto/proceso tenga el desempeño deseado sin que le afecten las fuentes de variación no controlada.*

EJEMPLO 16. *Consideremos una copiadora. Al momento de sacar copias, el usuario desea que la máquina funcione bien sin importar el tipo de papel usado y de la humedad ambiental, por mencionar sólo dos factores de ruido que el fabricante de las copadoras no puede controlar. Si la máquina trabaja bien en esas condiciones se puede decir que es robusta al papel y a la humedad ambiental. Además, el usuario también desea poder elegir reducción o aumento en tamaño de la copia, así como lo oscuro de la tinta, y que la copia siempre salga con la calidad deseada.*

EJEMPLO 17. *Si pensamos en una tostadora de pan, al usuario le gustaría que fuera robusta a la marca de pan, a la humedad ambiental, a la temperatura ambiental, a las variaciones en las dimensiones del pan, etc. Se desea seleccionar*



el grado de tostado y observar que efectivamente el pan salga con el color deseado, en toda su superficie y por ambos lados.

**EJEMPLO 18.** En una fábrica de dulces se tenía el problema de que el caramelo era altamente dependiente de la temperatura ambiental, de manera que cuando hacía mucho calor se escurría sobre las manos del consumidor final. El problema entonces era formular un dulce robusto a la temperatura ambiental. Se encontró, mediante un experimento robusto, una nueva formulación de caramelo más resistente a los cambios de temperatura. El efecto de la temperatura ambiental sobre la plasticidad del caramelo, antes y después de la mejora, se muestra en la figura 2. Cada curva representa a plasticidad en función de la temperatura ambiental. La función con menor pendiente obtenida después del experimento indica un efecto menor de la temperatura sobre la plasticidad.

#### 4. Factores de control, de ruido y de señal

En el Diseño robusto existen básicamente dos tipos de factores: controlables y no controlables (o de ruido). Los factores controlables que pueden influenciar el proceso se clasifican en cuanto a su efecto sobre la media y la variabilidad de la respuesta de interés, si:

- 1. Afecta la media y la variabilidad (factor clase *I*, figura 3a)
- 2. Afecta sólo la variabilidad (factor clase *II*, figura 3b)
- 3. Afecta sólo la media (factor clase *III*, figura 3c)
- 4. No afecta la media ni la variabilidad (factor clase *IV*, figura 3d).

Cuando se afirma en el diseño clásico que un factor tiene efecto sobre la respuesta, generalmente se quiere decir que el factor tiene efecto sobre la media de la característica de calidad. En cambio en el diseño robusto se tiene más presente que el efecto de un factor puede ser también sobre la variabilidad, o sobre la media y la variabilidad de manera simultánea. En el eje horizontal se muestra el efecto sobre la media y en el vertical el efecto sobre la variabilidad.

El factor clase *I* tiene una relación curva con la variable de respuesta. Considerando dos niveles ( $A_1$  y  $A_2$ ) de este factor, es claro de la gráfica que una misma

Figura 3

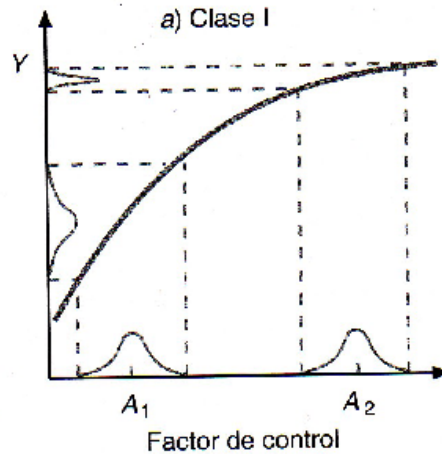
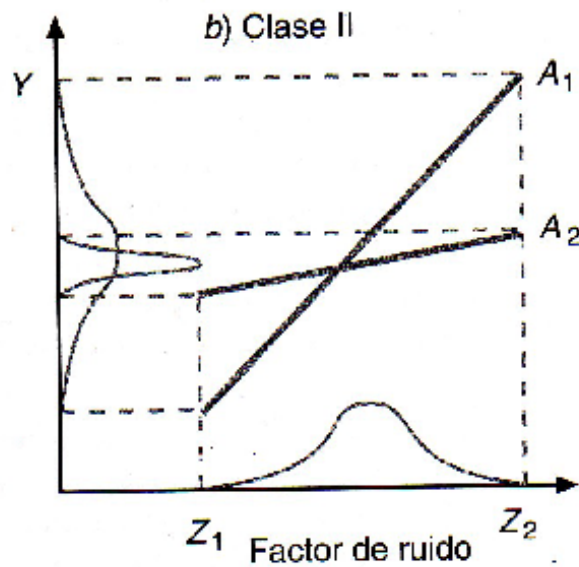
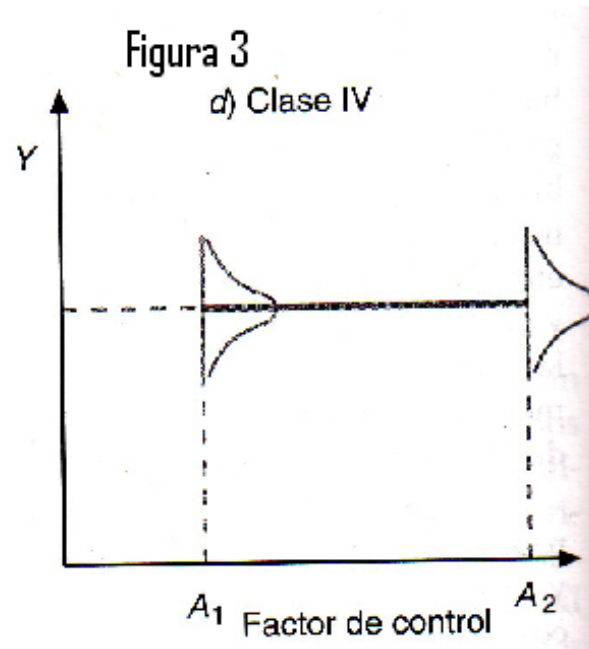
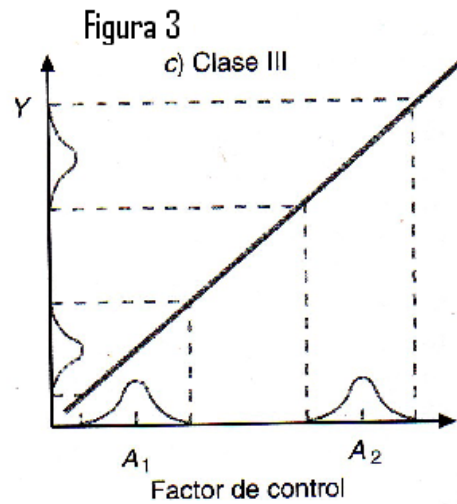


Figura 3



oscilación o variación del factor sobre cada uno de estos niveles (representada por las curvas sobre ellos) tiene un efecto distinto sobre la variable de respuesta, debido a la relación entre el factor y la respuesta. En el nivel  $A_2$  la respuesta  $Y$  tendría una variabilidad menor que en el nivel  $A_1$ , es decir, en el nivel  $A_2$  se tiene un comportamiento más robusto del proceso a las posibles oscilaciones del factor. Es claro que al mismo tiempo en  $A_2$  la media de la característica es mayor que en  $A_1$ .

El factor de control clase *II* se representa interactuando con un factor de ruido. Esta interacción tiene efecto más que nada en la variabilidad, ya que el factor de



ruido no se controla y seguirá variando entre sus valores extremos  $Z_1$  y  $Z_2$  de manera que en cualquiera de los niveles ( $A_1$  y  $A_2$ ) la media de la característica es la misma. Lo relevante con esta interacción es que en el nivel  $A_1$  del factor de control se minimiza el efecto del factor de ruido sobre la variabilidad de respuesta  $Y$ . Se puede afirmar entonces que en el nivel  $A_1$  el proceso es más *robusto* al efecto del factor de ruido. Este tipo de relación entre un factor de control y un factor de ruido es la idea fundamental que permite hacer diseño robusto. Si en un experimento dado los factores de control actuaran independientemente de los factores de ruido sería

imposible hacer diseño robusto, ya que no existiría una combinación de los primeros que minimizara el efecto de los segundos.

El factor clase *III* afecta sólo la media de la característica de calidad. La variabilidad que pasaría a la respuesta ( $Y$ ) por la oscilación de este factor es la misma en cualquier lugar de su rango. *Este tipo de factor es útil para llevar la media a su valor nominal* una vez elegidos los niveles de los factores clases *I* y *II* que minimizan la variabilidad, de aquí que al factor clase *III* se le llame *factor de ajuste*. Esto es, con los factores clases *I* y *II* es posible elegir las condiciones más robustas, reduciendo la variabilidad de la respuesta, pero la media se habrá movido de su valor deseable; con el factor clase *III* ésta se regresa a su valor nominal sin afectar la variabilidad. Finalmente, el factor clase *IV* no tiene efecto ni en la media ni en la variabilidad y de éste se elige su nivel más económico, como el mejor.

En la medida que se conoce la relación entre los factores controlables y no controlables con la variable de respuesta se está en posición de determinar mejores condiciones de operación del proceso. En el diseño robusto se trata de sacar ventaja principalmente de los factores de control clase *II* que interactúan con factores de ruido:

- Se trata de elegir el nivel del factor controlable que hace al proceso más insensible al ruido. Posteriormente se busca ajustar la media al valor nominal con un *factor de ajuste* (clase *III*).

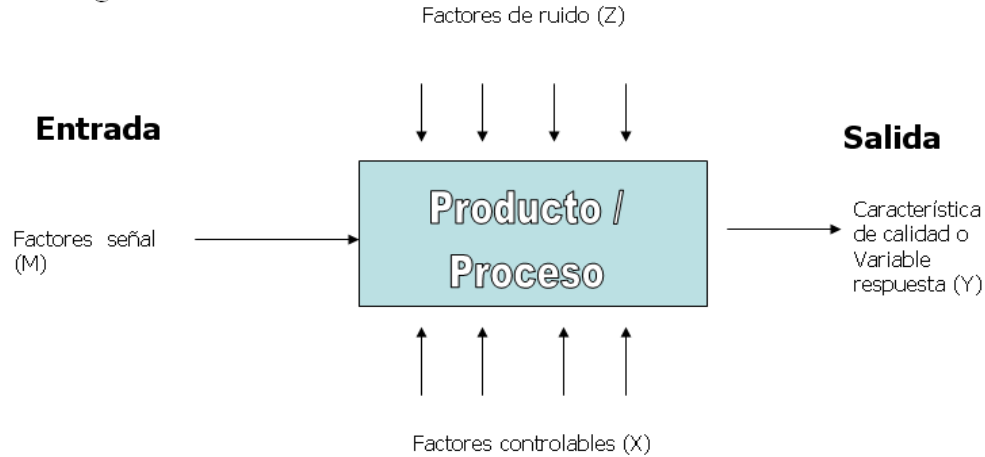
**4.1. Factor señal.** Se llama *factor señal* al dispositivo que permite cambiar el nivel de operación de acuerdo a los deseos del usuario. El usuario puede elegir la señal (o valor promedio de la respuesta) que desea en un momento dado del producto. Por ejemplo, en una *tostadora de pan* el factor señal es el mecanismo que permite seleccionar el *grado de tostado* deseado; en el caso de una *copiadora*, un factor señal es el mecanismo para elegir la *oscuridad deseada* de la impresión. En el caso del *limpiaparabrisas* de automóvil el factor señal son las diferentes *velocidades* que el conductor puede elegir de acuerdo al clima imperante. Es decir, el producto debe ser robusto en cada nivel del factor señal. Por ejemplo, el factor de control clase *III* de la figura 3, que sólo afecta la media, podría hacer las veces de factor señal, puesto que permitiría cambiar el valor de la media sin que cambie la variabilidad de la respuesta.

**4.2. Factores de ruido.** Los factores de ruido que actúan sobre el producto o sobre el proceso se pueden clasificar como de ruido externo, ruido interno y de deterioración.

- El *ruido externo* se refiere al ambiente en el cual el proceso (o producto) se desempeña y a la carga de trabajo a que es sometido. Por ejemplo, es ruido externo la humedad ambiental, el polvo o los errores en la operación del equipo.
- El *ruido interno* se refiere a la variación de unidad a unidad, debida a la diversidad de sus componentes. En el proceso es la falta de uniformidad en su manera de operar.
- El *deterioro* se refiere a efectos que aparecen poco a poco con el tiempo, que pueden causar la aparición de fallas en el *proceso/producto*. Por ejemplo, piezas o herramientas que se van gastando por el mismo uso.



Figura 4



**4.3. Diagrama esquemático.** En la figura 4 se muestran los diferentes tipos de factores que intervienen en el diseño robusto.

## 5. Arreglos ortogonales

Los arreglos ortogonales son diseños propuestos por *Taguchi* que tienen la propiedad de ortogonalidad, misma que también poseen los diseños factoriales clásicos. De hecho son diseños factoriales completos, fraccionados o mixtos, dependiendo del número de factores a estudiar en un arreglo ortogonal particular. Por ejemplo, el arreglo ortogonal  $L_8$  tiene ocho corridas experimentales, y con él se pueden estudiar desde dos hasta siete factores en dos niveles cada uno (figura 5).

Si se estudian siete factores equivale a un diseño factorial fraccionado  $2_{III}^{7-3}$  mientras que con dos factores sería un factorial completo  $2^2$  con dos réplicas. Para menos de siete factores es necesario saber a cuáles columnas del arreglo deben asignarse los factores, para tener la estructura *alias* que permita estudiar a esos factores lo más claro posible. Por ejemplo, estudiar cuatro factores con el arreglo  $L_8$ , éstos se deben asignar a las columnas 1, 2, 4 y 7, para obtener la información con la calidad de la fracción  $2_{IV}^{4-1}$  con generador  $I = ABCD$ .

Con el arreglo ortogonal  $L_{16}$  se pueden estudiar desde cuatro hasta 15 factores (para menos de cuatro factores mejor se utiliza un  $L_8$ ). Si se estudiaran cuatro factores, éstos se deben asignar a las columnas 1, 2, 4 y 8 para poder estimar de manera separada a las interacciones, y en este caso equivaldría a un diseño factorial completo  $2^4$ .

Arreglo $L_4$ ( $2^{3-1}$ )				Arreglo $L_9$ ( $3^{4-2}$ )					Arreglo $L_8$ ( $2^{7-4}$ )														
No. de col				No. de col					No. de columna														
Corrida		1 2 3		Corrida		1 2 3 4			Corrida		1 2 3 4 5 6 7												
1		1 1 1		1		1 1 1 1			1		1 1 1 1 1 1 1												
2		1 2 2		2		1 2 2 2			2		1 1 1 2 2 2 2												
3		2 1 2		3		1 3 3 3			3		1 2 2 1 1 2 2												
4		2 2 1		4		2 1 2 3			4		1 2 2 2 2 1 1												
2 factores: 1, 2				5		2 2 3 1			5		2 1 2 1 2 1 2												
3 factores: 1, 2, 3				6		2 3 1 2			6		2 2 1 1 2 2 1												
				7		3 1 3 2			7		2 2 1 1 2 2 1												
				8		3 2 1 3			8		2 2 1 2 1 1 2												
				9		3 3 2 1			9		2 2 1 2 1 1 2												
				2 factores: 1, 2					2 factores: 1, 2														
				3 factores: 1, 2, 3					3 factores: 1, 2, 4														
				4 factores: 1, 2, 3, 4					4 factores: 1, 2, 4, 7														
									5 factores: 1, 2, 4, 7, 6														
									6 factores: 1, 2, 4, 7, 6, 5														
									7 factores: las siete columna														
Arreglo $L_9$ ( $2 \times 3^{7-5}$ )												Arreglo $L_{12}$ (Plackett-Burman para $k = 11$ )											
		No. de columna												No. de columna									
Corrida		1 2 3 4 5 6 7 8										Corrida		1 2 3 4 5 6 7 8									
1		1 1 1 1 1 1 1 1										1		1 1 1 1 1 1 1 1									
2		1 1 2 2 2 2 2 2										2		1 1 2 2 2 2 2 2									
3		1 1 3 3 3 3 3 3										3		1 1 3 3 3 3 3 3									
4		1 2 1 1 2 2 3 3										4		1 2 1 1 2 2 3 3									
5		1 2 2 2 3 3 1 1										5		1 2 2 2 3 3 1 1									
6		1 2 3 3 1 1 2 2										6		1 2 3 3 1 1 2 2									
7		1 3 1 2 1 3 2 3										7		1 3 1 2 1 3 2 3									
8		1 3 2 3 2 1 3 1										8		1 3 2 3 2 1 3 1									
9		1 3 3 1 3 2 1 2										9		1 3 3 1 3 2 1 2									
10		2 1 1 3 3 2 2 1										10		2 1 1 3 3 2 2 1									
11		2 1 2 1 1 3 3 2										11		2 1 2 1 1 3 3 2									
12		2 1 3 2 2 1 1 3										12		2 1 3 2 2 1 1 3									
13		2 2 1 2 3 1 3 2										13		2 2 1 2 3 1 3 2									
14		2 2 2 3 1 2 1 3										14		2 2 2 3 1 2 1 3									
15		2 2 3 1 2 3 2 1										15		2 2 3 1 2 3 2 1									
16		2 3 1 3 2 3 1 2										16		2 3 1 3 2 3 1 2									
17		2 3 2 1 3 1 2 3										17		2 3 2 1 3 1 2 3									
18		2 3 3 2 1 2 3 1										18		2 3 3 2 1 2 3 1									
												1 factor con dos niveles: col 1											
												factores con tres columnas se asignan a las columnas restantes											

Figura 5. Arreglo $L_{16} (2^{15-11})$															
Corrida	No. de columna														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1
4 fac (1,2,4,8); 5 fac (1,2,4,8,15); 6 fac (1,2,4,8,15,14); 7 fac (1,2,4,8,15,14,13)															
8 fac (1,2,4,8,15,14,13,11); 9 factores (1,2,4,8,15,14,13,11,7);															
10 factores (1,2,4,8,15,14,13,11,7,12); 11 factores (1,2,4,8,15,14,13,11,7,12,10);															
12 fac (1,2,4,8,15,14,13,11,7,12,10,9); 13 fac (1,2,4,8,15,14,13,11,7,12,10,9,6);															
14 fac (1,2,4,8,15,14,13,11,7,12,10,9,6,5); 14 fac (1,2,4,8,15,14,13,11,7,12,10,9,6,5,3)															

Si se estudian 15 factores el arreglo  $L_{16}$  equivaldría a una fracción  $2_{III}^{15-11}$ . Ocurre que en la medida de que se estudian más factores, se tienen menos grados de libertad para estudiar interacciones; pero

- *Taguchi no enfatiza en el estudio de las interacciones, prefiere saturar lo más posible los arreglos y estudiar sólo los efectos principales de cada factor.*

Note que Taguchi acomoda las columnas de los arreglos ortogonales en un orden diferente al orden de *Yates* visto en los diseños factoriales. La primera columna de cada arreglo ortogonal es aquella donde los niveles aparecen lo más agrupados posible, de manera que el factor correspondiente se cambia de nivel un mínimo número de veces si el arreglo se corre en este orden. *Taguchi recomienda asignar a la primera columna aquel factor que sea más difícil de manipular durante el experimento*; es decir, aquel factor que sea difícil cambiarle su nivel de una prueba a otra. Así, aquellos factores que cambian lentamente o que por alguna otra razón no es fácil cambiar de una corrida a otra durante el experimento, se recomienda asignarlos a la primera columna de los arreglos ortogonales. De lo anterior se concluye que

- *Taguchi no enfatiza la necesidad de correr el experimento en orden aleatorio como se recomienda en diseño clásico, sino más bien presupone las complicaciones prácticas que se han señalado, y estructura el orden de las columnas conforme a esas dificultades.*

**5.1. Tipos de estudios de robustez.** Los estudios de robustez se clasifican utilizando como criterios al tipo de variable de respuesta y la existencia o ausencia de factores de señal. Una variable de respuesta puede ser de tres tipos:

- Entre más pequeña mejor.
- Entre más grande mejor.
- Nominal es lo mejor.

En cuanto al factores señal, se dice que el estudio es estático si no hay factor señal y es dinámico en el otro caso. De acuerdo a esto, los cuatro estudios de robustez de mayor interés práctico son los siguientes:

- Estático del tipo valor nominal es lo mejor.
- Estático del tipo más grande es lo mejor.
- Estático del tipo más pequeño es lo mejor.
- Dinámico con función ideal.

## 6. Diseño con arreglo interno y externo (diseño de parámetros)

La condición fundamental para que un diseño experimental sea de tipo robusto es que exista al menos un factor de ruido para el cual se busca hacer que el proceso o producto sea insensible a su efecto, sin pretender controlar dicho factor de ruido. Éste seguirá actuando como siempre en el proceso después del experimento, pero su efecto sobre el mismo será menor.

Un diseño experimental propuesto por Taguchi para determinar condiciones de operación robustas a uno o varios factores de ruido es el *diseño con arreglo interno y externo*. Una vez identificados los factores de control y los factores de ruido con los que se quiere experimentar, se construyen los arreglos ortogonales, uno para cada tipo de factores. El diseño consiste en sobreponer ambos arreglos de manera que en cada combinación de los factores controlables se prueban todas las combinaciones de los factores de ruido (figura 6). Note que en cada combinación de los factores controlables (arreglo interno) se están “simulando” distintos tipos de ruido que pueden ocurrir (arreglo externo). La mejor combinación de los factores de control es aquella donde los ruidos tienen el menor efecto (causan mínima variación) y al mismo tiempo la media se encuentra más cerca valor nominal.

Para ilustrar el diseño consideremos como arreglo interno, el arreglo ortogonal  $L_8$  (factorial fraccionado  $2_{III}^{7-4}$ ) y como arreglo externo el  $L_9$  (factorial fraccionado  $3_{III}^{4-2}$ ), cada uno de ellos con la cantidad máxima de factores. Esto es, se tienen siete factores de control y cuatro de ruido. El diseño completo, que consta de 72 corridas experimentales, se muestra en figura 6.

Una desventaja del diseño con arreglo interno y externo es que requiere una cantidad grande de corridas experimentales, aun utilizando los arreglos ortogonales más pequeños. De aquí que algunos autores (Grize, 1995) occidentales propongan como alternativa (con menos corridas):

- *Utilizar diseños factoriales completos o fraccionados clásicos incorporando los factores de ruido como factores de control, cuidando que el diseño permita estudiar la posible interacción entre unos y otros. Al interpretar dichas interacciones se selecciona el nivel del factor de control en el cual*

el impacto negativo del factor de ruido sea menor.

Figura 6. Diseño con arreglos interno $L_8$ y externo $L_9$																			
<div>Arreglo interno</div> <div>Factores controlables</div>								Arreglo externo										<div>media</div> <div><math>S/R</math></div>	
							$N$	1	2	3	3	1	2	2	3	1			
							$M$	1	2	3	2	3	1	3	1	2			
							$L$	1	2	3	1	2	3	1	2	3			
							$K$	1	1	1	2	2	2	3	3	3			
$A$	$B$	$C$	$D$	$E$	$F$	$G$													
1	1	1	1	1	1	1	$y_{11}$	$y_{12}$	·	·	·					$y_{19}$	$\bar{y}_1$ .	$S/R_1$	
1	1	1	2	2	2	2	$y_{21}$	$y_{22}$	·	·	·					$y_{29}$	$\bar{y}_2$ .	$S/R_2$	
1	2	2	1	1	2	2	·	·								·	$\bar{y}_3$ .	$S/R_3$	
1	2	2	2	2	1	1	·	·								·	$\bar{y}_4$ .	$S/R_4$	
2	1	2	1	2	1	2	·	·								·	$\bar{y}_5$ .	$S/R_5$	
2	1	2	2	1	2	1											$\bar{y}_6$ .	$S/R_6$	
2	2	1	1	2	2	1											$\bar{y}_7$ .	$S/R_7$	
2	2	1	2	1	1	2	$y_{81}$	$y_{82}$	·	·	·					$y_{89}$	$\bar{y}_8$ .	$S/R_8$	

## 7. El cociente señal/ruido

Para el análisis del diseño con arreglo interno y externo, *Taguchi* propone una estrategia de dos pasos basada en cierto estadístico de desempeño que le llama cociente *señal/ruido* (*signal to noise ratio*).

- (1) Se elige un estadístico diferente dependiendo de si se busca maximizar, minimizar o lograr cierto *valor objetivo* de la característica de calidad. Una vez que se calcula el valor del cociente *señal/ruido* en cada combinación de los factores controlables (figura 6), se analiza como cualquier variable de respuesta, y se determina la combinación más robusta de los niveles de los factores controlables, que es aquella que maximiza la variable *señal/ruido*.
- (2) En un segundo paso, con los factores que sólo afectan la media de la característica de calidad, se lleva ésta a su valor objetivo.

De acuerdo con el tipo de característica de calidad, el estadístico cociente *señal/ruido* se define de acuerdo a la tabla 1. Se obtiene un valor de cociente *señal/ruido* en cada renglón del arreglo interior de los factores controlables, a lo largo de las distintas combinaciones de los factores de ruido. De aquí que las sumas necesarias para hacer los cálculos de estos estadísticos corran hasta  $n$ , que es el número de combinaciones en el arreglo exterior.

- *El hecho de que se saque logaritmo en los estadísticos tiene que ver con buscar favorecer que los efectos de los factores controlables sean aditivos, es decir, que se minimice la posibilidad de efectos de interacción entre ellos. Se multiplican por 10 para trabajar en una escala más grande y el signo que lo antecede se escoge de manera que el problema siempre sea maximizar el valor del estadístico para obtener las condiciones de operación más robustas.*

La diferencia entre los dos estadísticos para respuestas del tipo *nominal es lo mejor*, es el tipo de características de calidad para las que se recomiendan:

- *El estadístico tipo I es para respuestas no negativas que van de cero a infinito, que tienen valor objetivo diferente de cero y que la varianza es*

cero cuando la respuesta es cero o bien cuando la media y la desviación estándar tienen una relación lineal.

- El estadístico tipo II es para respuestas que pueden tomar tanto valores positivos como negativos, y el cero puede ser el valor nominal o bien cuando la media y la desviación estándar son independientes (Fowlkes y Creveling, 1995).
- Para verificar el tipo de relación entre la media y la desviación estándar se realiza un diagrama de dispersión con los puntos  $(\bar{Y}_i, S_i)$  calculados en los renglones del arreglo interior. Si los puntos caen en una banda horizontal a lo largo del eje entonces la media y la desviación estándar son independientes. Si los puntos muestran alguna relación lineal, o de otro tipo, entonces son dependientes.

Los estadísticos *señal/ruido* fueron de los aspectos del método propuesto por Taguchi que más polémica generaron (Box, 1988). Se llegó a demostrar que estos estadísticos pueden dar resultados subóptimos puesto que confunden el efecto sobre la media con el efecto en la variabilidad. Una alternativa que se propuso es transformar los datos de manera que se logre un comportamiento independiente de la media y la variabilidad, y luego analizarlas de manera separada para los datos transformados (Nair y Pregibon, 1986). Otra alternativa más directa y práctica, que da buenos resultados en muchos casos, es utilizar con cualquier tipo de respuesta continua el estadístico del caso nominal es lo mejor el tipo II  $(-10 \log(S^2))$ .

Tabla 1. Tipos de razones señal/ruido	
Tipo de columna	Razón señal/ruido (S/R)
Mientras más pequeña es mejor	$-10 \log \left[ \frac{1}{n} \sum_{i=1}^n Y_i^2 \right]$
Mientras más grande es mejor	$-10 \log \left[ \frac{1}{n} \sum_{i=1}^n Y_i^{-2} \right]$
Su valor nominal es lo mejor (tipo I)	$-10 \log \left( \frac{\bar{Y}^2}{S^2} \right)$
Su valor nominal es lo mejor (tipo II)	$-10 \log (S^2)$
Proporción de defectuosos	$-10 \log \left( \frac{p}{(1-p)} \right)$

**7.1. Optimización en dos pasos.** La estrategia de análisis del experimento con arreglos interior y exterior se resume en los dos pasos siguientes:

- (1) Se determinan los factores controlables que tienen efecto sobre el estadístico razón *señal/ruido* ( $S/R$ ), que fue seleccionado de acuerdo al tipo de característica de calidad que se tiene (véase tabla 1). Con los efectos activos se determinan las condiciones de operación más robustas, que son las que maximizan el valor de la respuesta  $S/R$ . Esto se hace mediante las gráficas de efectos, como se ilustrará en el ejemplo siguiente.
- (2) Se realiza el análisis para la media  $\bar{Y}$ . Los factores que sólo afectan a la media se utilizan como factores de ajuste para llevar la media su valor nominal. Si todos los factores que afectan a la media también afectaran al estadístico  $S/R$  será necesario encontrar una solución de compromiso, utilizando como factor de ajuste el de más efecto en la media y con menos efecto en la dispersión o variabilidad.

- *Adicional a los pasos anteriores, es útil para cualquier respuesta continua también realizar estos mismos dos pasos con el estadístico general  $S/R = -10\log(S^2)$ , con el propósito de tener otra visión del análisis. Si los resultados de este otro análisis llegaran a diferir de los resultados del estadístico recomendado por Taguchi, se deben analizar con detenimiento las dos soluciones encontradas para determinar cuál es finalmente la mejor. Este segundo análisis puede proteger al experimentador de una solución subóptima que algunas veces generan las señal/ruido originales (Logothetis y Wynn, 1994).*

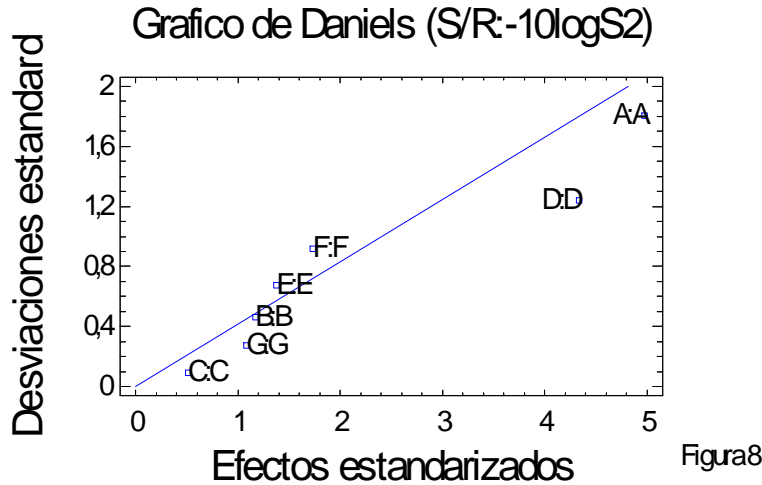
### 8. Un ejemplo ilustrativo de un experimento robusto

Una de las características importantes en el proceso de producción de un pigmento es su color. El problema que se tenía en este proceso era el de variación del color del pigmento. Un grupo de mejora decide utilizar diseño robusto para tratar de hacer el proceso menos sensible al efecto de factores de ruido difíciles de controlar durante la producción. Se identificaron seis factores de control y tres de ruido, los cuales se muestran en la tabla 2. Se decide utilizar un arreglo ortogonal  $L_8$  para los factores de control y un  $L_4$  para los factores de ruido; con lo que el diseño resultante tiene 32 corridas (pruebas) a nivel proceso. El diseño y los datos obtenidos muestran en la figura 7.

TABLA 2 Factores de control y de ruido.	
Factores de control (fáciles de controlar)	Factores de ruido (difíciles de controlar)
A: Tiempo de carga de materiales B: Tiempo de amasado C: Exceso de sal D: Temperatura de amasado E: Orden de introducción de materiales F: Velocidad de amasado	K: Calidad de la sal L: Aspecto de la resina M: Temperatura del agua para enfriar

Figura 7. Diseño con arreglos interno $L_8$ y externo $L_4$														
Arreglo interno  Factores controlables							$M$ $L$ $K$	Arreglo externo						
								1	2	2	1			
								1	2	1	2			
								1	1	2	2			
$A$	$B$	$C$	$D$	$E$	$F$	$G$					media	$S$	$-10 \log (S^2)$	
1	1	1	1	1	1	1	36	26	24	15	25.25	8.61	-18.71	
1	1	1	2	2	2	2	32	62	24	32	37.50	16.76	-24.49	
1	2	2	1	1	2	2	34	16	25	12	21.75	9.81	-19.83	
1	2	2	2	2	1	1	10	30	26	32	24.50	9.98	-19.98	
2	1	2	1	2	1	2	33	31	27	23	28.50	4.43	-12.93	
2	1	2	2	1	2	1	34	48	26	39	36.75	9.21	-19.29	
2	2	1	1	2	2	1	26	27	18	20	22.75	4.42	-12.92	
2	2	1	2	1	1	2	28	40	21	32	30.25	7.93	-17.99	

Se procede a calcular los estadísticos de interés en cada combinación de niveles del arreglo interior o de factores de control. Los valores de la media, desviación estándar y del estadístico  $S/R$  nominal *tipo II* se muestran en las tres últimas columnas de la figura 7. Se utiliza este estadístico *señal/ruido* porque el color del pigmento del tipo nominal es mejor y además la desviación estándar se comporta



independiente de la media. A continuación aplicamos los dos pasos descritos en la sección anterior para el análisis de los datos y optimización de la respuesta.

- (1) Se analiza la variable  $S/R = -10 \log(S^2)$  y se obtienen los efectos y sus gráficas como se explicó en los diseños factoriales. El gráfico de Daniel al igual que como se hace en el diseño clásico detecta efectos activos que más inciden en a la razón  $S/R$ . (Vea figura 8), donde se observan que los factores  $A$  y  $D$  son los más significativos.
- (2) Las gráficas de efectos principales se muestran en la figura 9. Aquí sólo presentamos las gráficas de efectos, y de ellas se observa que los efectos  $A$  y  $D$  afectan de manera significativa a la  $S/R$ . Es decir, los factores  $A$  y  $D$  influyen bastante sobre la variación del color del pigmento. De aquí que lo que se recomiende sea utilizar el factor  $A$  en su nivel alto y el factor  $D$  en su nivel bajo, este tratamiento hará más robusto al proceso (menos sensible al efecto de los factores de ruido que se han estudiado). Note que estos niveles ( $A = 2$ ,  $D = 1$ ) corresponden al renglón 5 y 7 del arreglo interno de la figura 7, y es en estas combinaciones de niveles donde se observa la menor variación;  $S = 4.43$  y  $S = 4.42$ , respectivamente.
- (3) Se procede al análisis de la media y las gráficas de Daniels para los efectos se muestran en la figura 10. Se observa que los factores  $B$  y  $D$  tienen efecto significativo sobre la media. La figura 11 de los efectos principales, muestra que el factor  $B$  tiene efecto sobre la media, pero no tiene efecto sobre el estadístico  $S/R$ , de manera que puede utilizarse como *factor de ajuste* para llevar la media a su valor nominal. Por ejemplo, si el valor objetivo del color es 23 conviene elegir el nivel alto del factor  $B$ . Por tanto, la solución propuesta es

$$(A = 2, B = 2, C = \$, D = 1, E = \$, F = \$, G = \$)$$



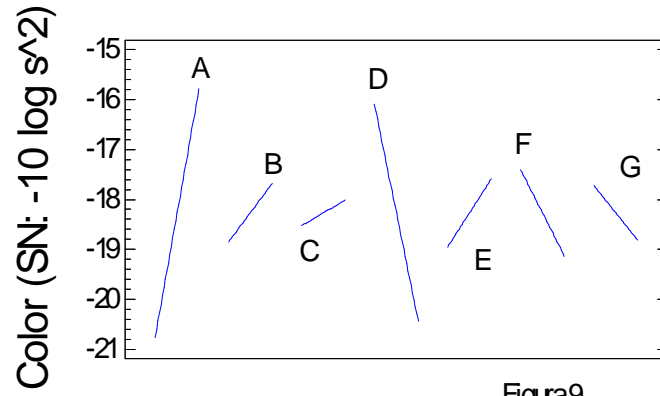


Figura9

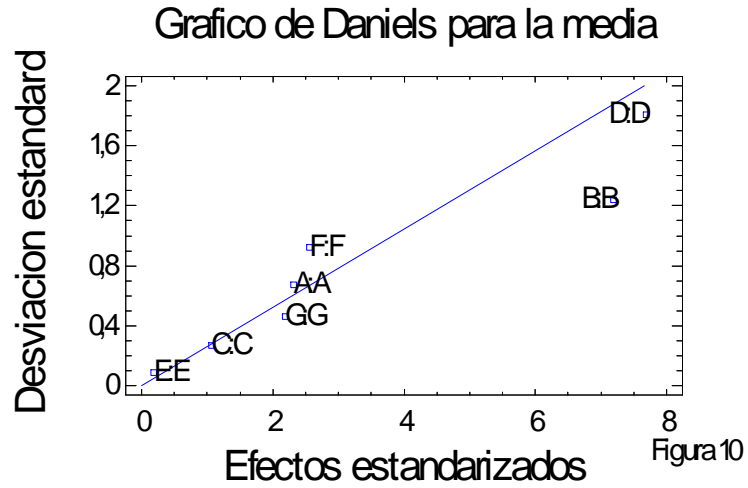


Figura 10

- (4) En la figura 12 se observa el diagrama de Pareto sin estandarizar que confirma las conclusiones para el efecto  $B$ .

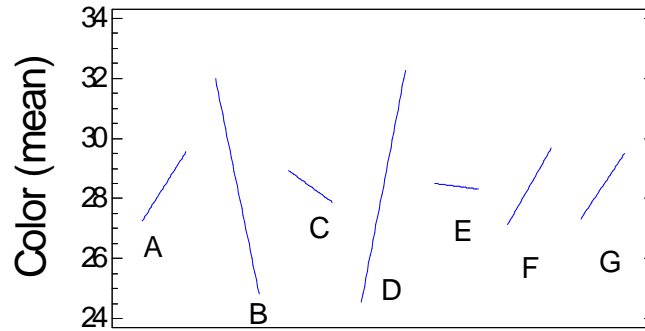


Figura11

### Grafico de Pareto para la media

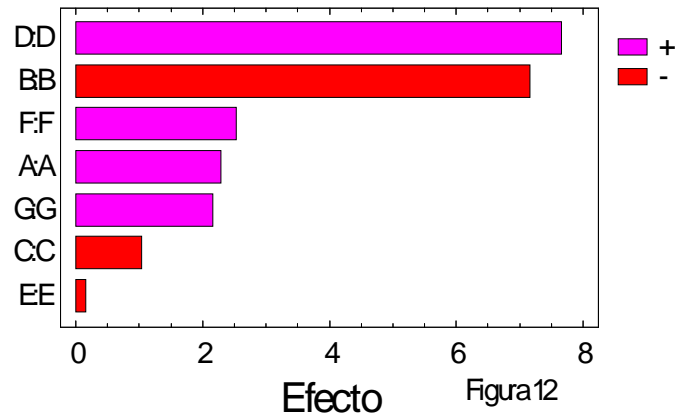


Figura12

- (5) A continuación se muestra el mejor ANOVA. Con base a lo que ya hemos estudiado durante todo el curso

Tala 3. ANOVA pare el color-Diseño robusto					
FV	SC	G.L	C.M	F <sub>0</sub>	valor-p
A	10,6953	1	10,6953	9,19	0,0938
B	103,32	1	103,32	88,76	0,0111
D	118,195	1	118,195	101,54	0,0097
F	13,1328	1	13,1328	11,28	0,0784
G	9,57031	1	9,57031	8,22	0,1032
Error	2,32813	2	1,16406		
Total	257,242	7			
R <sup>2</sup> := 99,095%					
R <sup>2</sup> -ajustado: = 96,8324%					

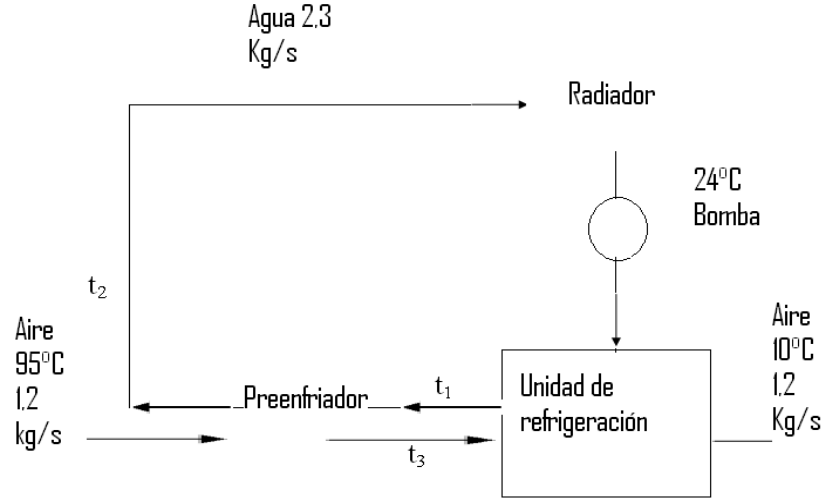


Figura1. Sistema de refrigeración de aire comprimido

### 9. Otro ejemplo ilustrativo de un diseño robusto

- (1) **Identificar la función principal.** La función principal del sistema representado en la Figura 1, es enfriar el aire desde 95 hasta los 10 °C. El proceso de enfriamiento se realiza en dos etapas. En la primera etapa, el aire comprimido entra en el preenfriador a 95°C y sale con una temperatura de  $t_3$ °C, posteriormente se vuelve a enfriar en la unidad de refrigeración hasta alcanzar la temperatura deseada de 10°C. El caudal de aire comprimido es de 1,2 Kg/s y el de agua 2,3 Kg/s, siendo 24°C la temperatura que se espera tenga el agua a la salida del radiador.
- (2) **Identificación de la característica de calidad a optimizar. Definición de los objetivos.** El sistema se diseña para que el costo total sea mínimo. Este costo es la suma de los costos asociados a la unidad de refrigeración, preenfriador y radiador. Las ecuaciones paramétricas que relacionan los costos  $X_i$  con las temperaturas de salida son las siguientes:

$$\begin{aligned}
 X_1 &= 1,2a(t_3 - 10) \\
 X_2 &= \frac{1,2b(95 - t_3)}{(t_3 - t_4)}, \quad t_3 > t_4 \\
 X_3 &= 9,637c(t_2 - 24)
 \end{aligned}$$

Donde :  $X_1$  := Costo asociado a la unidad de refrigeración (\$),  $X_2$  := Costo asociado al preenfriador (\$),  $X_3$  := Costo asociado al radiador (\$). Stoecker [7] da los siguientes valores para los parámetros :  $a = 48$ ,  $b = 50$ ,  $c = 25$ . En este ejemplo la característica de calidad a optimizar sería el costo total

$$CT = X_1 + X_2 + X_3.$$

El objetivo es lograr un diseño que cumpla la función al mínimo costo ,sujeto a las restricciones de balance de masas, energía, compresión y potencia y teniendo en cuenta los factores de ruido.

- (3) **Identificación de los factores de control , factores de ruido y de sus niveles de experimentación.** En la Tabla 2 se muestran los factores y niveles que determinan la región de experimentación.

Tabla 2. Niveles de los factores en la experimentación			
FACTORES DE CONTROL	NIVEL 1	NIVEL 2	NIVEL 3
t <sub>1</sub> = Temperatura del agua a la entrada del preenfriador	25 °C	28 °C	32 °C
t <sub>2</sub> =Temperatura del agua a la salida del preenfriador	34 °C	39 °C	42 °C
t <sub>3</sub> = Temperatura del aire a la salida del preenfriador	35 °C	38 °C	43 °C
FACTORES DE RUIDO			
R <sub>1</sub> =Temperatura de salida del agua en el radiador	24 °C	27 °C	
R <sub>2</sub> =Temperatura del aire a la entrada del preenfriador	95 °C	120 °C	

- (4) **Planificación y ejecución del experimento.** Como se tienen en cuenta tres factores de control a tres niveles, el arreglo ortogonal para la matriz interna es el mostrado en la Figura 2 . En este arreglo, las columnas son mutuamente ortogonales. Esto significa que, por cada par de columnas, se encuentran todas las combinaciones posibles de los niveles en un mismo número de veces. Este arreglo se denomina diseño L<sub>9</sub>, el 9 indica el número de filas, combinaciones o prototipos a ensayar bajo los niveles de los factores de control.
- (5) **Análisis de los datos.** Por cada combinación de niveles de los factores de control se calculan la media y los valores de la  $S/R$  (Carácterística de calidad tipo “Menor – Mejor”). Posteriormente se calculan los valores medios de los niveles y se representan gráficamente los efectos (ver Figura

Figura 2. Diseño con arreglos interno L <sub>9</sub> y externo L <sub>4</sub>										
Arreglo interno	Arreglo externo						más pequeña es mejor			
	R <sub>2</sub>	1	2	1	2					
	R <sub>1</sub>	1	1	2	2					
t <sub>1</sub>	t <sub>2</sub>	t <sub>3</sub>					media	S	$-10 \log \left[ \frac{1}{n} \sum_{i=1}^n Y_i^2 \right]$	
1	1	1	4209	4359	3486	3636	3922,863	426,186	-71,910	
1	2	2	5490	5605	4767	4882	5186,057	422,578	-74,318	
1	3	3	6411	6494	5688	5771	6091,063	420,059	-75,709	
2	1	3	4518	4618	3795	3895	4206,663	421,269	-72,511	
2	2	1	5568	5782	4845	5060	5313,916	435,248	-74,530	
2	3	2	6291	6441	5569	5719	6005,063	426,186	-75,587	
3	1	2	4592	4842	3869	4119	4355,663	441,552	-72,814	
3	2	3	5798	5935	5076	5212	5505,106	424,656	-74,835	
3	3	1	6977	7477	6254	6754	6865,263	507,413	-76,751	

- (6) **Tablas de las medias y gráficos de los efectos.** El gráfico nos revela que el factor de control t<sub>2</sub> es el más significativo y tiene un gran efecto sobre el costo. Claramente el nivel 1 es el nivel más robusto para t<sub>1</sub> y t<sub>2</sub>. El nivel óptimo para t<sub>3</sub> es el nivel 2 pero su efecto es casi despreciable. La combinación óptima de niveles para el diseño del sistema es: t<sub>1</sub> = 25°C, t<sub>2</sub> = 34°C, t<sub>3</sub> = 38°C. Bajo estas condiciones, las predicciones son las

## Efectos S/R (menor es mejor)

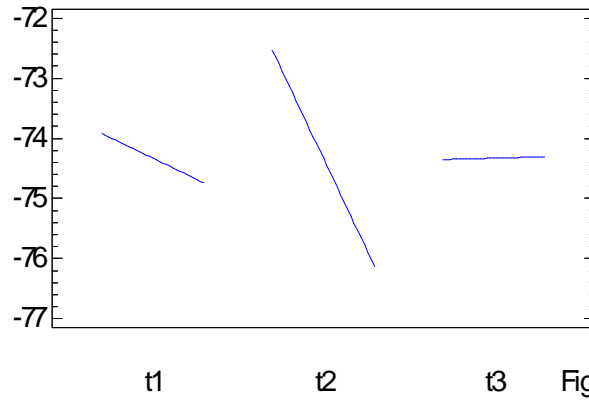


Figura 3a

## Efectos Media

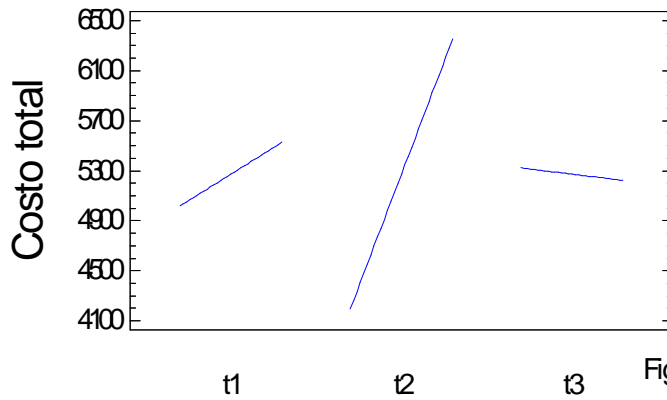
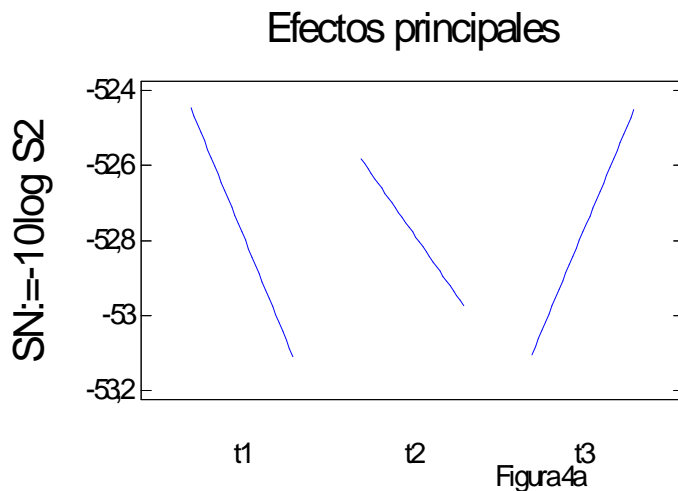


Figura 3b

siguientes :  $Costo = \$3981$ ,  $S/R = -72,03Db$ .

	$S/R$			$MEDIA$		
	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$
Nivel 1	-73,979	-72,412	-74,397	5066,66	4161,73	5367,35
Nivel 2	-74,209	-74,561	-74,240	5175,21	5335,03	5182,26
Nivel 3	-74,800	-76,016	-74,352	5575,34	6320,46	5267,61

- (7) **Realización del experimento confirmatorio.** Antes de dar definitivamente por bueno los resultados anteriores. Taguchi propone realizar una serie de prototipos con la combinación óptima obtenida para garantizar la fiabilidad de las predicciones.



- (8) **Nominal es lo mejor tipo II.** Realizaremos independientemente el análisis con el estadístico nominal es lo mejor tipo II. La grafica de efectos principales 4a. muestra que los tres factores estudiados afectan considerablemente la variabilidad, de manera que el mejor tratamiento se obtiene en el punto de diseño  $(1, 1, 3) = (t_1 = 25^{\circ}C, t_2 = 34^{\circ}C, t_3 = 38^{\circ}C)$ , que coincide con lo visto en el análisis de la señal entre más pequeña es lo mejor. La diferencia sustancial consiste en que el análisis inicial los factores  $t_1, t_3$  no afectaban la variabilidad.
- (9) **Conclusiones.** En esta comunicación se refleja que los métodos de Taguchi permiten a los técnicos encontrar los valores óptimos de los parámetros del diseño a través de la investigación sobre la variabilidad, o más específicamente, sobre la sensibilidad de estos parámetros con respecto a los factores que causan de dicha variabilidad. La idea es reducir la variabilidad sin incrementar los costos. En lugar de eliminar las causas de la variación, lo que resultaría excesivamente caro, se minimizan los efectos de estas causas. En el ejemplo expuesto, se trasladan las consideraciones sobre el costo a las etapas del diseño de producto donde los beneficios que se pueden obtener son mayores.

**EJEMPLO 19.** *Se conduce un experimento para encontrar una manera económica de ensamblar un conector elástico a un tubo de nylon, con la fuerza de unión requerida. Son dos los objetivos básicos que se tienen: por un lado minimizar el esfuerzo necesario para hacer el ensamble y por el otro maximizar la fuerza del ensamble hecho. Los investigadores identificaron cuatro factores de control y tres de ruido, que ellos pensaban podían estar afectando la fuerza del ensamble. Los factores y sus niveles fueron los siguientes:*

**Factores de control:**

A: Interferencia (baja, media, alta)

B: Grosor de pared (delgada, media, gruesa)

C: Profundidad de inserción (superficial, media, honda)

D: Porcentaje de adhesivo (bajo, medio, alto)

**Factores de ruido:**

M: Tiempo de acondicionamiento (24 h, 12 h)

N: Temperatura de acondicionamiento (72° F, 150° F)

O: Humedad relativa de acondicionamiento (25%, 75%)

Como se aprecia, se decide probar cada factor controlable en tres niveles y variar cada factor de ruido en dos niveles. Aunque no es posible controlar los factores de ruido durante la producción, sí hay manera de hacerlo para fines experimentales. Se seleccionó un arreglo ortogonal  $L_9$  para los factores de control y un arreglo  $L_8$  para los factores de ruido. El diseño con arreglos interno y externo resultante, así como los datos obtenidos, se muestran en la figura 13. Estos datos representan la fuerza necesaria para desunir el ensamble, el cual se busca maximizar. (a) De acuerdo con las recomendaciones de la figura 5, en el arreglo externo ( $L_8$ ), se han elegido los renglones 1, 2 y 4 como los que definen los niveles de los factores M, N y O, respectivamente; para cada una de las ocho corridas. Tomando en cuenta esto, note que cada combinación de los factores de control se prueban en todas las combinaciones de los factores de ruido. Límitese a los renglones referidos antes y anote los niveles reales de estas combinaciones de los factores de ruido. (b) De acuerdo con lo que hizo en el inciso anterior, compruebe que las ocho condiciones o combinaciones de los factores de ruido conforman un diseño  $2^3$ . (c) ¿Cuál es la razón señal/ruido adecuada para la fuerza del ensamble? Argumente. (d) Calcule la media, la desviación estándar y la razón señal/ruido en cada combinación de los factores de control, a lo largo de las combinaciones de los factores de ruido. (e) Realice el análisis de la razón señal/ruido y grafique los efectos principales de los factores de control y determine las condiciones de operación más robustas. (f) Haga el análisis para la media, grafique los efectos de los factores controlados y

determine las mejores condiciones de operación para el proceso de ensamble.

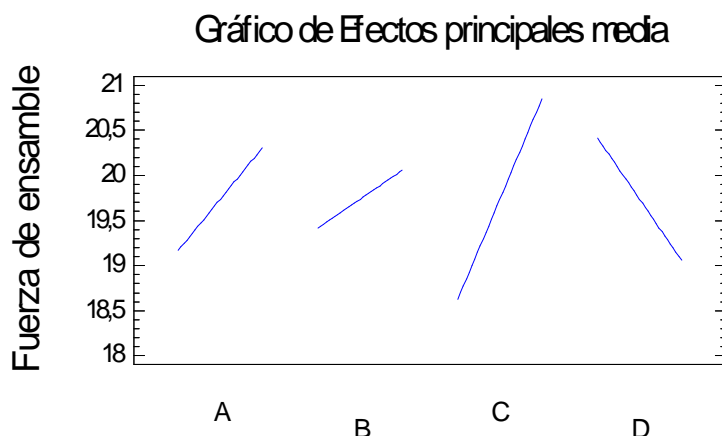
Figura 13				Diseño con arreglos interno $L_9$ y externo $L_8$											
Arreglo interno Factores controlables	$M$	2	2	2	2	1	1	1	1						
	$N$	2	2	1	1	2	2	1	1						
	$MN$	1	1	2	2	2	2	1	1						
	$O$	2	1	2	1	2	1	2	1						
	$MO$	1	2	1	2	2	1	2	1						
	$NO$	1	2	2	1	1	2	2	1						
	$E$	2	1	1	2	1	2	2	1						
$A$	$B$	$C$	$D$										media	$S$	$-10 \log (S^2)$
1	1	1	1	19,1	20,0	19,6	19,6	19,9	16,9	9,5	15,6	17,525	3,61250	-11,15634947	
1	2	2	2	21,9	24,2	19,8	19,7	19,6	19,4	16,2	15,0	19,475	2,90652	-9,267466314	
1	3	3	3	20,4	23,3	18,2	22,6	15,6	19,1	16,7	16,3	19,025	2,88333	-9,197887013	
2	1	2	3	24,7	23,2	18,9	21,0	18,6	18,9	17,4	18,3	20,125	2,59766	-8,291646139	
2	2	3	1	25,3	27,5	21,4	25,6	25,1	19,4	18,6	19,7	22,825	3,42751	-10,69957461	
2	3	1	2	24,7	22,5	19,6	14,7	19,8	20,0	16,3	16,2	19,225	3,37967	-10,57748593	
3	1	3	2	21,6	24,6	18,6	16,8	23,6	18,4	19,1	16,4	19,8875	3,0498	-9,685427202	
3	2	1	3	24,4	23,2	19,6	17,8	16,8	15,1	15,6	14,2	18,3375	3,7747	-11,53764882	
3	3	2	1	28,6	22,6	22,7	23,1	17,3	19,3	19,9	16,1	21,2	3,94787	-11,92725687	

- (1) Variable de interés: Fuerza de Emsamble
- (2) (b) De acuerdo con lo que hizo en el inciso anterior, compruebe que las ocho condiciones o combinaciones de los factores de ruido conforman un diseño  $2^3$ . En efecto, las combinaciones de los factores de ruido coincide con el diseño  $2^3$  clasico que ya hemos estudiado.
- (3) (c) ¿Cuál es la razón señal/ruido adecuada para la fuerza del ensamble? Argumente. Como el objetivo es de maximizar la fuerza de ensamble, la razon señal/ruido segun Taguchi es mientras más grande es lo mejor, sinembargo, se recomienda su valor nominal es lo mejor  $-10 \log(S^2)$ , presupone independencia entre la media y la desviación
- (4) (d) Calcule la media, la desviación estándar y la razón señal/ruido en cada combinación de los factores de control, a lo largo de las combinaciones de los factores de ruido. Ver arriba
- (5) (e) Realice el análisis de la razón señal/ruido y grafique los efectos principales de los factores de control y determine las condiciones de operación más robustas. Se observa que en el nivel que en el nivel 3 del Factor D,  $-10 \log(S^2)$  es mayor, los niveles superiores de los otros tres factores al parecer presentan la misma contribución a la variabilidad.
- (6) (f) Haga el análisis para la media, grafique los efectos de los factores controlados y determine las mejores condiciones de operación para el proceso de ensamble. Se observa que las combinaciones que maximizan la fuerza de ensamble en cuanto a su valor medio son: nivel alto de A, nivel alto de B, nivel alto de C. El tramiento que optimiza el diseño es nivel alto de A, nivel alto de B, nivel alto de C y nivel alto del factor D.

## 10. Ejercicios

EJERCICIO 129. Algunas de las diferencias básicas entre los diseños de experimentos clásico y el robusto se da en torno a diferentes énfasis que cada uno hace





respecto a la media y a la varianza. Explique con detalle en qué consisten estas diferencias.

EJERCICIO 130. Comente algunas características de la función de pérdida de Taguchi y sus implicaciones.

EJERCICIO 131. ¿El diseño robusto propone controlar los factores de ruido o qué es lo que propone al respecto? Dé ejemplos.

EJERCICIO 132. Apoyándose en los ejemplos vistos en el presente capítulo, dé un par de ejemplos de productos que ante ciertas circunstancias se desearía que fueran robustos.

EJERCICIO 133. Señale las cuatro categorías de un factor controlable respecto a su efecto sobre la media y la variabilidad.

EJERCICIO 134. Explique qué es un factor señal y qué es un factor de ruido.

EJERCICIO 135. Los arreglos ortogonales son cierto tipo de diseños de experimentos, explique en qué consisten éstos y dé un par de ejemplos de este tipo de diseños.

EJERCICIO 136. ¿Para qué se recomienda emplear un arreglo  $L_{16}$ ?

EJERCICIO 137. ¿En qué consiste un arreglo  $L_9$  y cuándo se recomienda emplearlo?

EJERCICIO 138. Explique en qué consiste un diseño con arreglo interno y externo.

EJERCICIO 139. ¿Con qué propósito la primera columna de un arreglo ortogonal tiene los niveles bastante agrupados, no aleatorizados? ¿Qué ventaja práctica tiene esto?

EJERCICIO 140. ¿Qué es una razón o cociente señal/ruido?

EJERCICIO 141. El análisis de un diseño robusto con arreglo interno y externo, se hace en dos pasos. Explique con detalle estos pasos.

EJERCICIO 142. *¿Qué se sugiere hacer cuando haya conflictos entre la solución del análisis del estadístico  $S/R$  y el análisis de la media?*

EJERCICIO 143. *¿Cuáles inconvenientes pueden surgir del análisis de los estadísticos  $S/R$  propuestos por Taguchi?*

EJERCICIO 144. *Diga al menos una desventaja que tiene el diseño con arreglo interno y externo.*

EJERCICIO 145. *¿Para hacer diseño robusto es estrictamente necesario utilizar un diseño con arreglo interno y externo? Argumente.*

EJERCICIO 146. *En el proceso de fabricación de muelles para automóviles, una variable de salida importante es la altura sin carga, cuyo valor nominal es 8 pulgadas. La altura sin carga de un muelle se determina durante el tratamiento de calor en el cual se forma la curvatura del muelle. El problema es determinar las condiciones de este tratamiento que den por resultado la altura de carga deseada. Se corre un experimento con cuatro factores de control y un factor de ruido. Los cuatro factores controlables son: (B) temperatura del horno, (C) tiempo de quemado, (D) tiempo de transferencia y (E) tiempo de permanencia (tiempo en el cual se forma la curvatura en una pieza caliente). Los ingenieros también estaban interesados en estudiar los efectos de interacción BC, BD y CD. El factor de ruido es la temperatura del lubricante (O), la cual es difícil de controlar durante la producción. Los niveles utilizados en cada factor se muestran en la siguiente tabla:*

<b>Factores de control</b>	<b>Niveles</b>	
	<i>Bajo</i>	<i>Alto</i>
<i>B: Temperatura del horno (<math>^{\circ}F</math>)</i>	1840	1880
<i>C: Tiempo de quemado (seg.)</i>	25	23
<i>D: Tiempo de transferencia (seg.)</i>	12	10
<i>E: Tiempo de permanencia (seg.)</i>	2	150-170
<i>O: Temperatura del lubricante (<math>^{\circ}F</math>)</i>	130-150	

*Note que como niveles del factor de ruido se utilizan rangos de temperatura, dada la dificultad de controlarlo de manera puntual. Se utiliza un arreglo ortogonal  $L_8$  para los factores controlables. Los datos obtenidos se muestran en la tabla 3. (a) Analice la matriz de diseño para los factores de control y diga a qué diseño factorial fraccionado corresponde. (b) ¿Qué resolución tiene tal fracción factorial? Escriba su estructura alias reducida. (c) ¿Cuál es el estadístico  $S/R$  más apropiado al tipo de variable de respuesta que se tiene? Argumente. (d) Analice con detalle la respuesta  $S/R$  y determine el tratamiento más robusto. (e) ¿Tienen efecto sobre la razón señal/ ruido las interacciones que les interesaban a los ingenieros? (f) Analice la respuesta promedio  $\bar{Y}$ . (g) Si hay algún factor que no afectó al cociente  $S/R$  pero que sí afecta la media, utilícelo como factor de ajuste para acercar la media a su valor objetivo. (h) Repita el análisis utilizando el estadístico  $S/R = -10 \log(S^2)$  y*

comente las diferencias observadas.

B	C	BC	D	BD	CD	E	Factor de ruido					
							$O^-$			$O^+$		
-	-	+	-	+	+	-	7,78	7,78	7,81	7,50	7,25	7,12
+	-	-	-	-	+	+	8,15	8,18	7,88	7,88	7,88	7,44
-	+	-	-	+	-	+	7,50	7,56	7,50	7,50	7,56	7,50
+	+	+	-	-	-	-	7,59	7,56	7,75	7,63	7,75	7,56
-	-	+	+	-	-	+	7,94	8,00	7,88	7,32	7,44	7,44
+	-	-	+	+	-	-	7,69	8,09	8,06	7,56	7,69	7,62
-	+	-	+	-	+	-	7,56	7,62	7,44	7,18	7,18	7,25
+	+	+	+	+	+	+	7,56	7,81	7,69	7,81	7,50	7,59

EJERCICIO 147. En 1951 en la Ina Tile Company se tenía el problema de que el horno quemaba disperejo debido a variación de la temperatura en diferentes partes del horno, y esto causaba defectos en las lozas que se fabricaban. Una posibilidad de solución (imposible en ese momento) era cambiar el horno por otro que no tuviera ese problema. Otra posibilidad era reformular las lozas de manera que fueran robustas al funcionamiento “disperejo” del horno. Esto último fue lo que se decidió hacer, utilizando los siguientes niveles de prueba en siete factores de la formulación de la loza:

Factor	Nivel 1	Nivel 2
A: Aditivo de cal	$A_1=5\%$	$A_2=1\%$ (actual)
B: Granularidad del aditivo	$B_1=$ tosca (actual)	$B_2=$ fin
C: Contenido de algamatolite	$C_1=43\%$	$C_2=53\%$ (actual)
D: Tipo de algamatolite	$D_1=$ mezcla actual	$D_2=$ más barata
E: Cantidad de carga	$E_1=1\ 300\text{ kg}$	$E_2=1\ 200\text{ kg}$ (actual)
F: Contenido de reciclado	$F_1=0\%$	$F_2=4\%$ (actual)
G: Contenido de feldespat	$G_1=0\%$	$G_2=5\%$ (actual)

Note que uno de los niveles de prueba para cada uno de los factores corresponde al nivel que se utilizaba hasta ese momento. Se tomó una muestra de 100 lozas en cada uno de los ocho tratamientos y se obtiene el porcentaje de lozas defectuosas. Los resultados obtenidos se muestran en la siguiente tabla. (a) ¿Por qué este experimento es un diseño robusto? (b) Analice con detalle los datos: efectos principales y efectos activos. (c) Obtenga la mejor formulación de las lozas. Asigne el nivel más económico a los factores que no tienen efecto sobre el porcentaje de defectuosos. (d) ¿Cuál es la proporción de loza defectuosa esperada en el tratamiento elegido? (e) Estime la diferencia entre la proporción de loza esperada en el tratamiento anterior (actual) y el tratamiento nuevo sugerido por el estudio.

Núm. de Corrida	A	B	C	D	E	F	G	% de lozas defectuosas
1	1	1	1	1	1	1	1	16
2	1	1	1	2	2	2	2	17
3	1	2	2	1	1	2	2	12
4	1	2	2	2	2	1	1	6
5	2	1	2	1	2	1	2	6
6	2	1	2	2	1	2	1	68
7	2	2	1	1	2	2	1	42
8	2	2	1	2	1	1	2	26



## Diseños anidados y diseños en parcelas divididas

### 1. Competencias

- (1) *Distinguir entre factores anidados y cruzados identificando situaciones donde el diseño anidado es la opción experimental adecuada y no el diseño de factores cruzados.*
- (2) *Realizar el análisis de diseños anidados que involucran o no factores cruzados.*
- (3) *Identificar diferentes situaciones prácticas que pueden llevar a utilizar un experimento en parcelas divididas.*
- (4) *Distinguir diferentes estructuras de diseño que pueden presentar los factores en las parcelas y las subparcelas.*
- (5) *Comprender por qué algunos experimentos con arreglo interno y externo son en realidad parcelas divididas.*
- (6) *Analizar e interpretar correctamente el diseño en parcelas divididas comprendiendo como un análisis incorrecto podría cambiar las conclusiones.*

### 2. Introducción

Existen situaciones experimentales que involucran a varios factores y donde surgen consideraciones que se traducen en restricciones a la completa aleatorización de las corridas experimentales, o bien, hay situaciones en donde los niveles de prueba de un factor no son los mismos al cambiar los niveles de otro factor. *Estos escenarios causan que el experimento deje de ser un factorial completamente aleatorizado para convertirse en un diseño en parcelas divididas (Split-plot Design) o en un diseño anidado (Nested Design).* Recordemos que los diseños que usan bloques también surgen de restricciones a la aleatorización, y en este caso se incluyen los factores de bloques no porque sean de interés en sí, sino para incrementar la precisión del experimento. En cambio, en los diseños anidados y de parcelas divididas ésa no es la situación. Por ejemplo, en el caso de este último tipo de diseño, la restricción es impuesta por factores que son de interés para el experimentador, y se implementa para facilitar y hacer más eficiente la ejecución del experimento.

El error típico que se comete en las situaciones experimentales de los diseños anidados y de los diseños en parcelas divididas es que, a menudo, el experimentador los analiza como si fueran diseños factoriales, y esto puede llevar a conclusiones erróneas, como lo ilustraremos a través de ejemplos.

### 3. Diseños anidados

En capítulos previos vimos que cuando se tiene un diseño factorial con *dos factores cruzados A y B*, se corren en orden aleatorio todas las posibles combinaciones de niveles de los dos factores. Ahí, los niveles de cada factor se pueden combinar

en cualquier momento con los niveles del otro factor, y en este caso los niveles de un factor son exactamente los mismos que en cada nivel del otro factor. Por otra parte, *cuando se dice que el factor B está anidado en el factor A significa que los niveles del factor B no son los mismos en cada nivel del factor A*. Es decir, hay una especie de relación *padre-hijo* entre los niveles del factor A (padres) y los niveles del factor B (hijos).

Un diseño anidado es un diseño que posee dos o más factores tal que:

- (1) Hay un sólo factor, el cual no se encuentra anidado con ningún otro factor, sea este  $F_1$ , que representa la primera jerarquía entre todos los factores que existen.
- (2) Existe un segundo factor  $F_2$ , el cual se encuentra dentro de  $F_1$ . Diremos  $F_2$  es el factor que representa la segunda jerarquía entre todos los factores que existen.
- (3) El próximo factor en caso que exista ( $F_3$ ), debe encontrarse anidado dentro del factor  $F_2$ . Diremos  $F_3$  es el factor que representa la tercera jerarquía entre todos los factores que existan.

En este sentido, los diseños anidados también se conocen como *diseños jerárquico*.

#### 4. Modelo y análisis estadístico de diseños anidados

**4.1. Modelo de diseños anidados: anidamiento en dos etapas.** En algunas situaciones no se pueden combinar todos los niveles de un factor con todos los niveles de otro, es decir, no se pueden determinar todos los posibles tratamientos que aparecen al cruzar los factores.

**EJEMPLO 20.** *Supongamos que en un centro de formación profesional se estudia el porcentaje de aprobados en una materia, en los grupos de mañana y de tarde. Por la mañana imparten la asignatura dos personas y por la tarde tres. Cada persona da clase a tres grupos y se supone que estos son réplicas (no son fuente de variación). Así: Factor A: Turno ( $i = 1, 2$ ), Factor B: Persona ( $j = 1, \dots, 5$ ),  $Y_{ij}$  = Porcentaje de aprobados*

TURNO				
MAÑANA			TARDE	
$P_1$	$P_2$	$P_3$	$P_4$	$P_5$
$g_1$	$g_1$	$g_1$	$g_1$	$g_1$
$g_2$	$g_2$	$g_2$	$g_2$	$g_2$
$g_3$	$g_3$	$g_3$	$g_3$	$g_3$

*Se dice que el factor B está anidado en el factor A, es decir  $B \subset A$ .*

##### 4.1.1. Modelo matemático para un diseño anidado en dos etapas de efectos fijos.

En el diseño anidado más simple se tienen dos factores A y B, con los niveles de B anidados en cada nivel de A. Consideremos inicialmente el diseño anidado en dos

etapas balanceado, como se ilustra en la siguiente tabla:

Factor A												
A <sub>1</sub>				A <sub>2</sub> ...					A <sub>a</sub>			
Factor B				Factor B ...					Factor B			
B <sub>1(1)</sub>	B <sub>2(1)</sub>	...	B <sub>b(1)</sub>	B <sub>1(2)</sub>	B <sub>2(2)</sub>	...	B <sub>b(2)</sub>	...	B <sub>1(a)</sub>	B <sub>2(a)</sub>	...	B <sub>b(a)</sub>
obs <sub>1</sub>	obs <sub>1</sub>	...	obs <sub>1</sub>	obs <sub>1</sub>	obs <sub>1</sub>	...	obs <sub>1</sub>	...	obs <sub>1</sub>	obs <sub>1</sub>	...	obs <sub>1</sub>
obs <sub>2</sub>	obs <sub>2</sub>	...	obs <sub>2</sub>	obs <sub>2</sub>	obs <sub>2</sub>	...	obs <sub>2</sub>	...	obs <sub>2</sub>	obs <sub>2</sub>	...	obs <sub>2</sub>
⋮	⋮	...	⋮	⋮	⋮	...	⋮	...	⋮	⋮	...	⋮
obs <sub>n</sub>	obs <sub>n</sub>	...	obs <sub>n</sub>	obs <sub>n</sub>	obs <sub>n</sub>	...	obs <sub>n</sub>	...	obs <sub>n</sub>	obs <sub>n</sub>	...	obs <sub>n</sub>

El modelo estadístico que describe el comportamiento esperado de la respuesta en este diseño está dado por:

$$(4.1) \quad Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}; \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n.$$

donde hay  $a$  niveles de  $A$ ,  $b$  niveles del factor  $B$  anidados bajo cada nivel de  $A$  y  $n$  réplicas. Además,  $\mu$  es la media general,  $\alpha_i$  es el efecto del  $i$ -ésimo nivel del factor  $A$ ;  $\beta_{j(i)}$  es el efecto del  $j$ -ésimo nivel del factor  $B$ , que está anidado en el nivel  $i$  del factor  $A$ ; y  $\varepsilon_{ijk}$  es el error aleatorio con los supuestos usuales de normalidad, varianza constante e independencia. Note que el modelo no incluye el concepto de interacción, ya que el factor  $B$  no está cruzado sino anidado en el factor  $A$ .

Además, para cada  $i$ ,  $\sum_{j=1}^b \beta_{j(i)} = 0$ ,  $\sum_{i=1}^a \alpha_i = 0$ . Por otro lado,  $b$  es el número de niveles anidados en cada nivel  $i$ , de modo que el número total de niveles de  $B$  es  $ab$  y la suma de los efectos del factor  $B$  dentro de cada nivel de  $A$  es 0.

4.1.2. *Estimadores por mínimos cuadrados y ANOVA.* Se tiene que

$$\min_{\mu, \alpha_i, \beta_{j(i)}} \phi = \min_{\mu, \alpha_i, \beta_{j(i)}} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n \left( y_{ijk} - \mu - \alpha_i - \beta_{j(i)} \right)^2$$

Así,

$$\frac{\partial \phi}{\partial \mu} = -2 \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n \left( y_{ijk} - \mu - \alpha_i - \beta_{j(i)} \right) = 0 \Rightarrow \hat{\mu} = \bar{Y} \dots$$

Para cada  $i$  fijo

$$\frac{\partial \phi}{\partial \alpha_i} = -2 \sum_{j=1}^b \sum_{k=1}^n \left( y_{ijk} - \mu - \alpha_i - \beta_{j(i)} \right) = 0 \Rightarrow \hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y} \dots$$

Para cada  $i$  fijo y  $j$  fijo

$$\frac{\partial \phi}{\partial \beta_{j(i)}} = -2 \sum_{k=1}^n \left( y_{ijk} - \mu - \alpha_i - \beta_{j(i)} \right) = 0 \Rightarrow \hat{\beta}_{j(i)} = \bar{Y}_{ij.} - \bar{Y}_{i..}$$

De este modo

$$\hat{Y}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_{j(i)} = \bar{Y}_{ij.}$$

El número total de observaciones es  $N = abn$  y el número total de parámetros a estimar es  $1 + (a - 1) + a(b - 1) = ab$ , luego el número de grados de libertad total

es  $abn - ab = ab(n - 1)$ . De este modo, la estima de la varianza es

$$\hat{\sigma}^2 = CME = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{ij\cdot})^2}{ab(n-1)}$$

Las hipótesis de interés expresadas en la notación del modelo son:

$$(4.2) \quad \begin{cases} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_a = 0 \\ H_1 : \alpha_i \neq 0, \text{ para algún } i \end{cases},$$

$$\begin{cases} H_0 : \beta_{1(1)} = \beta_{2(1)} = \dots = \beta_{b(1)} = 0 \\ H_1 : \beta_{j(1)} \neq 0 \text{ para algún } j \end{cases}$$

$$\vdots$$

$$\begin{cases} H_0 : \beta_{1(a)} = \beta_{2(a)} = \dots = \beta_{b(a)} = 0 \\ H_1 : \beta_{j(a)} \neq 0 \text{ para algún } j \end{cases}$$

Observe que las hipótesis sobre el factor anidado se expresan dentro de cada nivel del factor  $A$  usando la notación  $j(i)$ , que significa nivel  $j$  anidado en el nivel  $i$ . Estas hipótesis se prueban mediante un ANOVA apropiado, el cual toma en cuenta si los factores son fijos o aleatorios. La mayoría de las sumas de cuadrados que se deben calcular para el ANOVA son las mismas que se calculan en el diseño factorial con dos factores, salvo la suma de cuadrados del efecto anidado  $B(A)$  que se denota por  $SC_{B(A)}$ , y que se obtiene con respecto a la media dentro de cada nivel del factor  $A$  y no con relación a la media global como se hace con el efecto  $A$  no anidado. Específicamente, la suma de cuadrados totales está dada por:

$$\begin{aligned} SCT &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{...})^2 \\ &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n ((\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{ij\cdot} - \bar{Y}_{i..}) + (Y_{ijk} - \bar{Y}_{ij\cdot}))^2 \\ &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \bar{Y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{ij\cdot} - \bar{Y}_{i..})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{ij\cdot})^2 \\ &= SC_A + SC_{B(A)} + SCE \\ &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - \frac{Y_{...}^2}{N} \end{aligned}$$

donde  $N = abn$  es el total de observaciones en el experimento. Las sumas de cuadrados de los efectos  $A$  y  $B(A)$  son:

$$SC_A = \frac{1}{bn} \sum_{i=1}^a Y_{i..}^2 - \frac{Y_{...}^2}{N}$$

$$SC_{B(A)} = \sum_{i=1}^a \left( \sum_{j=1}^b \frac{Y_{ij\cdot}^2}{n} - \frac{Y_{i..}^2}{bn} \right)$$

y la suma de cuadrados del error, por diferencia, es:  $SC_E = SC_T - SC_A - SC_{B(A)}$  donde los grados de libertad correspondientes cumplen esta misma relación,  $ab(n - 1) =$



$(abn - 1) - (a - 1) - a(b - 1)$ . A continuación se presenta la tabla anova para este diseño anidado

Tabla Anova para el Diseño Anidado con efectos fijos					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor - p
Efecto A	SCA	a-1	CMA	$\frac{CMA}{CME}$	$P(F > F_0^A)$
Efecto B(A)	SC <sub>B(A)</sub>	a(b-1)	CM <sub>B(A)</sub>	$\frac{CM_{B(A)}}{CME}$	$P(F > F_0^{B(A)})$
Error	SCE	ab(n-1)	CME		
Total	SCT	abn-1			

4.1.3. *Expresiones para el modelo lineal en un diseño anidado balanceado.* Al igual que en los diseños factoriales, los valores esperados de los cuadrados medios en un experimento anidado dependen de si ambos factores se consideran fijos o aleatorios; o uno fijo y el otro aleatorio. En la práctica se suele abusar del supuesto de factores fijos, quizá porque los paquetes estadísticos tienden a suponer lo mismo, o porque es más fácil comprender las hipótesis de interés expresadas en términos de los efectos promedio, como se muestran en las ecuaciones (4.2). Cabe recordar que cuando los efectos son aleatorios, las hipótesis de interés se expresan en términos de las varianzas de los mismos ( $H_0 : \sigma_{efecto}^2 = 0$ ). En resumen, para el caso anidado con dos factores, en la tabla 1 se muestran los valores esperados de los cuadrados medios para las tres combinaciones de interés de factores fijos y aleatorios. La combinación A aleatorio y B fijo usualmente no es de interés práctico, por eso no se incluye.

Factores	Modelo lineal	Hipotesis
A fijo, B fijo	$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$	$H_0: \alpha_i = 0, \text{ para todo } i$ $H_0: \beta_{j(i)} = 0, \forall i, j$ $\varepsilon_{ijk} \approx NID(0, \sigma^2)$
A fijo, B aleatorio	$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$	$H_0: \alpha_i = 0, \text{ para todo } i$ $\beta_j \approx NID(0, \sigma_\beta^2)$ $\varepsilon_{ijk} \approx NID(0, \sigma^2)$
A aleatorio, B aleatorio	$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$	$\alpha_i \approx NID(0, \sigma_\alpha^2)$ $\beta_j \approx NID(0, \sigma_\beta^2)$ $\varepsilon_{ijk} \approx NID(0, \sigma^2)$

Tabla 1. Valores esperados de los CM en el diseño anidado con dos factores			
$E(CM)$	A fijo, B fijo	A fijo, B aleatorio	A aleatorio, B aleatorio
$E(CM_A)$	$\sigma^2 + bn \sum \alpha_i^2 / (a - 1)$	$\sigma^2 + n\sigma_\beta^2 + bn \sum \alpha_i^2 / (a - 1)$	$\sigma^2 + n\sigma_\beta^2 + bn\sigma_\alpha^2$
$E(CM_{B(A)})$	$\sigma^2 + n \sum \sum \beta_{j(i)}^2 / a(b - 1)$	$\sigma^2 + n\sigma_\beta^2$	$\sigma^2 + n\sigma_\beta^2$
$E(CM_E)$	$\sigma^2$	$\sigma^2$	$\sigma^2$
		$\hat{\sigma}^2 = CM_E$	$\hat{\sigma}^2 = CM_E$
		$\hat{\sigma}_\beta^2 = \frac{CM_{B(A)} - CM_E}{n}$	$\hat{\sigma}_\beta^2 = \frac{CM_{B(A)} - CM_E}{n}$
		$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}...$	$\hat{\sigma}_\alpha^2 = \frac{CM_A - CM_{B(A)}}{bn}$

La información que provee esta tabla es útil para saber, en cada caso, cómo se deben construir los estadísticos de prueba para las hipótesis de interés y, por lo tanto, para saber cómo se calculan las dos últimas columnas de la tabla de ANOVA. Por

ejemplo, en el caso (fijo, fijo) los estadísticos se construyen con los cocientes:

$$(4.3) \quad F_0^A = \frac{CM_A}{CM_E}, \quad F_0^{B(A)} = \frac{CM_{B(A)}}{CM_E}$$

que serían igual a uno bajo las respectivas hipótesis nulas; en la tabla 1, observe que cuando los efectos  $A$  y  $B(A)$  son nulos o iguales a cero, los estadísticos de (4.3) se reducirían a  $\sigma^2/\sigma^2 = 1$ . De manera similar, de la columna (A fijo, B aleatorio) se ve que para esta combinación los estadísticos correspondientes son:

$$(4.4) \quad F_0^A = \frac{CM_A}{CM_{B(A)}}, \quad F_0^{B(A)} = \frac{CM_{B(A)}}{CM_E}$$

Observe que en los estadísticos de prueba (ecuaciones 4.3 y 4.4) cambia el denominador, dependiendo de lo que se supone con respecto a los factores. La información hasta aquí presentada es fácil de bosquejar en la tabla de ANOVA para cada caso. Veámoslo, pero en el contexto de un ejemplo.

4.1.4. *Ejemplo: Anidamiento en dos etapas.* Un geólogo estudia el contenido en trazas radiactivas de cinco tipos diferentes de suelo. Para ello recoge cuatro muestras de contenido en sustancias radiactivas en cuatro localidades diferentes que están situadas sobre cada tipo de suelo. Se obtienen los siguientes datos:

	A: Tipos de suelo																			
	K				L				M				N				O			
	B: Localidad				B: Localidad				B: Localidad				B: Localidad				B: Localidad			
	L <sub>1</sub>	L <sub>2</sub>	L <sub>3</sub>	L <sub>4</sub>	L <sub>5</sub>	L <sub>6</sub>	L <sub>7</sub>	L <sub>8</sub>	L <sub>9</sub>	L <sub>10</sub>	L <sub>11</sub>	L <sub>12</sub>	L <sub>13</sub>	L <sub>14</sub>	L <sub>15</sub>	L <sub>16</sub>	L <sub>17</sub>	L <sub>18</sub>	L <sub>19</sub>	L <sub>20</sub>
	6	13	1	7	10	2	4	0	0	10	8	7	11	5	1	0	1	6	3	3
	2	3	10	4	9	1	1	3	0	11	5	2	0	10	8	8	4	7	0	7
	0	9	0	7	7	1	7	4	5	6	0	5	6	8	9	6	7	0	2	4
	8	8	6	9	12	10	9	1	5	7	7	4	4	3	4	5	9	3	2	0
$y_{ij.}$	16	33	17	27	38	14	21	8	10	34	20	18	21	26	22	19	21	16	7	14
$y_{i..}$	93				81				82				88				58			
$y_{...}$	402																			

- (1) *Variable respuesta.*  $Y_{ijk}$  = contenido en trazas radiactivas de la  $k$  sustancia de la localidad  $j$  en el suelo  $i$
- (2) *Dos Factores:* Tipos de suelo (5 niveles), Localidad (4 niveles).
- (3) *Unidades experimentales:* cada uno de las sustancias a los que se le mide el contenido en trazas radiactivas (80)
- (4) *Módulo matemático asociado:*  $Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$  con  $i = 1, 2, 3, 4, 5$ ;  $j = 1, 2, 3, 4$  y  $k = 1, 2, 3, 4$ . Donde  $\mu$  = media poblacional de la variable respuesta,  $\alpha_i$  = efecto medio adicional debido al tipo de suelo  $i$ ;  $\beta_{j(i)}$  = efecto medio adicional debido a la localidad  $j$  del tipo de suelo  $i$ ,  $\varepsilon_{ijk}$  es el error aleatorio, son parámetros a estimar, con las condiciones  $\sum_{i=1}^4 \alpha_i = 0$ ,

$$\forall i \sum_{j=1}^4 \beta_{j(i)} = 0, \quad \varepsilon_{ijkl} \sim N(0, \sigma^2) \text{ independientes.}$$

- (5) *Hipótesis del modelo:*  $\left\{ \begin{array}{l} H_0 : \alpha_i = 0 \\ H_1 : \alpha_i \neq 0, \text{ para algún } i \end{array} \right\}, \left\{ \begin{array}{l} H_0 : \beta_{j(i)} = 0 \\ H_1 : \beta_{j(i)} \neq 0 \text{ para algos } i, j \end{array} \right\}$

(6) *Fuente de variación*

$$\begin{aligned}
 SCT &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - \frac{Y_{...}^2}{N} = (6^2 + 13^2 + \dots + 0^2) - \frac{1}{80} 402^2 = 969.95 \\
 SC_A &= \frac{1}{bn} \sum_{i=1}^a Y_{i..}^2 - \frac{Y_{...}^2}{N} = \frac{1}{16} (93^2 + \dots + 58^2) - \frac{1}{80} 402^2 = 45.75 \\
 SC_{B(A)} &= \sum_{i=1}^a \left( \sum_{j=1}^b \frac{Y_{ij.}^2}{n} - \frac{Y_{i..}^2}{bn} \right) = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2 - \frac{1}{bn} \sum_{i=1}^a Y_{i..}^2 \\
 &= \frac{1}{4} (16^2 + \dots + 14^2) - \frac{1}{16} (93^2 + \dots + 58^2) = 282.875 \\
 SCE &= SCT - SC_A - SC_{B(A)} = 642
 \end{aligned}$$

(7) La tabla Anova es

Tabla Anova para el Diseño Anidado con efectos fijos					
Variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Efecto A	45.75	4	11.269	1.053	0.387699
Efecto B(A)	282.875	15	18.858	1.762	0.0626039
Error	642	60	10.7		
Total	969.95	79			

Con un nivel de significancia de  $\alpha = 0.10$  se concluye que no existen diferencias significativas entre los terrenos.

4.1.5. *Ejemplo ilustrativo de un diseño en dos etapas.* Se desea hacer un estudio acerca del número de días que permanecen hospitalizados los enfermos de cardiología. Para ello, se realiza un estudio en los cuatro hospitales más importantes de la ciudad y, dentro de cada uno de los hospitales, se consideran los tres médicos de la especialidad en cardiología. A continuación, se eligen aleatoriamente tres pacientes tratados por cada médico, y se observa el número de días que permanecieron ingresados estos pacientes. Los datos se presentan en la siguiente tabla:

Hospital											
1			2			3			4		
Médico			Médico			Médico			Médico		
M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
7	11	12	9	14	14	8	11	17	6	9	8
10	9	11	9	11	13	12	14	20	7	8	10
8	12	15	12	12	10	13	14	19	6	9	6

- (1) Identificaremos los distintos elementos (factores, unidades experimentales y variable de respuesta) que intervienen en el estudio y plantearemos el modelo matemático detallado y las condiciones del mismo.
- (2) *Variable respuesta.*  $Y_{ijk}$  = número de días que permanece ingresado el paciente  $k$  tratado por el médico  $j$  en el hospital  $i$
- (3) *Dos Factores:* Hospital ( 4 niveles), Médico (3 niveles). Los tratamientos son cada uno de los médicos adscritos a cada hospital (12 tratamientos)
- (4) *Unidades experimentales:* Cada uno de los pacientes atendidos por cada médico en cada Hospital.

- (5) *Módulo matemático asociado:*  $Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$  con  $i = 1, 2, 3, 4$ ;  $j = 1, 2, 3$ ;  $k = 1, 2, 3$ . Donde  $\mu$  =media poblacional de la variable respuesta,  $\alpha_i$  =efecto medio adicional debido al hospital  $i$ ;  $\beta_{j(i)}$  =efecto medio adicional debido al médico  $j$  del hospital  $i$ ,  $\varepsilon_{ijk}$  es el error aleatorio, son parámetros a estimar, con las condiciones  $\sum_{i=1}^4 \alpha_i = 0$ ,  $\forall i \sum_{j=1}^3 \beta_{j(i)} = 0$ ,  $\varepsilon_{ijkl} \sim N(0, \sigma^2)$  independientes.
- (6) *Hipótesis del modelo:*  $\begin{cases} H_0 : \alpha_i = 0 \\ H_1 : \alpha_i \neq 0, \text{ para algún } i \end{cases}, \begin{cases} H_0 : \beta_{j(i)} = 0 \\ H_1 : \beta_{j(i)} \neq 0 \text{ para algos } i, j \end{cases}$
- (7) Analizaremos si el número de días que los pacientes permanecen ingresados en cada hospital es similar. En caso negativo, analizar cuál o cuáles son los hospitales en los que los enfermos están menos días ingresados usando el método de Duncan. Se construye la tabla Anova para días de permanencia

Tabla 1. Anova para días de permanencia					
variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
Modelo	340,667	11	30,9697	10,42	0.0000
Residuo	71,3333	24	2,97222		
Total (corregido)	412,0	35			

Tabla 2. Anova para días de permanencia					
variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
HOSPITAL	198,0	3	66,0	22,21	0.0000
MEDICO(HOSPITAL)	142,667	8	17,8333	6,00	0,0003
Residuo	71,3333	24	2,97222		
Total (corregido)	412,0	35			

De la Tabla Anova 2, se concluye que los hospitales difieren muy significativamente en cuanto al número de días medio que los pacientes permanecen ingresados en los hospitales ya que el  $p$ -valor asociado es 0.0000. Utilizando el método de Duncan, estudiaremos en qué hospital u hospitales están menos días ingresados los pacientes. La tabla 3, muestra que no existen diferencias significativas entre los hospitales 1 y 2, pero si entre el hospital 4 y el resto, así como el hospital 3 y el resto.

Tabla 3. Método de Duncan para días de permanencia					
Hospital	recuento	media	desviacion	Grupo Homogeneo	
4	9	7,66667	0,574671	X	
1	9	10,5556	0,574671		X
2	9	11,5556	0,574671		X
3	9	14,2222	0,574671		X
Hospital con menor duración: 4					

- (8) ¿Se puede asegurar que los médicos influyen en el número de días que los pacientes permanecen ingresados en el hospital? De la Tabla Anova 2, se concluye que los Médicos difieren muy significativamente en cuanto al número de días medio que los pacientes permanecen ingresados en los hospitales ya que el  $p$ -valor asociado es 0.0003.

### 5. Modelo y análisis estadístico de diseños cruzado-anidados

Se dice que dos factores están completamente cruzados cuando aparecen todas las posibles cobinaciones de los niveles de cada factor, como es el caso de un diseño bifactorial. Se dice que dos factores están cruzados cuando ninguno de ellos está anidado en el otro, es decir, ni  $A \subset B$  ni  $B \subset A$ . Por ejemplo en el ejemplo previo de turnos y personal docente, puede ser que alguno de estos dé clase en los dos turnos a la vez, con lo que se tienen dos factores cruzados aunque no completamente cruzados.

Los diseños cruzado-anidados se caracterizan por tener tanto factores cruzados como anidados. No existe un único modelo matemático, ya que depende de la disposición de los factores en el diseño.

**5.1. Modelo 1.** Se trata de estudiar el tiempo de montaje de una serie de piezas de relojería que ha de hacerse a mano. Se consideran 3 posiciones diferentes para montar las piezas y cuatro tamaños diferentes de las mismas. El montaje lo efectua una serie de personas, de modo que se ocupan dos personas distintas para cada montaje con cada tamaño y posición. Se tiene así, como factores,

$$\begin{aligned} A &\equiv \text{Posición } (i = 1, 2, 3) \\ B &\equiv \text{Tamaño } (j = 1, 2, 3, 4) \\ C &\equiv \text{Individuo } (k = 1, 2) \end{aligned}$$

El esquema es el siguiente:

Factor A											
A <sub>1</sub>				A <sub>2</sub>				A <sub>3</sub>			
Factor B				Factor B				Factor B			
B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>
C <sub>1</sub>	C <sub>3</sub>	C <sub>5</sub>	C <sub>7</sub>	C <sub>9</sub>	C <sub>11</sub>	C <sub>13</sub>	C <sub>15</sub>	C <sub>17</sub>	C <sub>19</sub>	C <sub>21</sub>	C <sub>23</sub>
C <sub>2</sub>	C <sub>4</sub>	C <sub>6</sub>	C <sub>8</sub>	C <sub>10</sub>	C <sub>12</sub>	C <sub>14</sub>	C <sub>16</sub>	C <sub>18</sub>	C <sub>20</sub>	C <sub>22</sub>	C <sub>24</sub>

Se observa que todos los niveles del factor posición se cruzan con todos los niveles del factor tamaño y que el factor individuo tiene niveles distintos para cada uno de los cruces, es decir, trabajan personas diferentes en cada caso. El esquema es  $A \times B$  y  $C \subset (A \times B)$ .

5.1.1. *Modelo matemático.* En el modelo se tienen que incluir:

- (1) Los efectos principales de  $A$  y  $B$ .
- (2) Las interacciones entre  $A$  y  $B$ .
- (3) Los efectos de cada nivel  $k$  del factor  $C$  anidado en la combinación de  $(i, j)$ .

El modelo estadístico para este tipo de diseño está dado por:

$$\begin{aligned} Y_{ijkl} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{k(ij)} + \varepsilon_{l(ijk)} \\ i &= 1, \dots, a; j = 1, \dots, b; k = 1, \dots, c; l = 1, \dots, n \end{aligned}$$

sujeto a las restricciones  $\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = 0, \forall (i, j)$

$\sum_{k=1}^c \gamma_{k(ij)} = 0$ , ya que el factor  $C$  tiene  $c$  niveles anidados en cada combinación de los niveles de  $A$  y  $B$ . La tabla ANOVA es

Tabla 2. ANOVA para el diseño $2^2$					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A	SCA	a-1	CMA	CMA/CME	$P(F > F_0^A)$
B	SCB	b-1	CMB	CMB/CME	$P(F > F_0^B)$
AB	SCAB	(a-1)(b-1)	CMAB	CMAB/CME	$P(F > F_0^{AB})$
C(AB)	$SC_{C(AB)}$	ab(c-1)	$CM_{C(AB)}$	$CM_{C(AB)}/CME$	$P(F > F_0^{C(AB)})$
Error	SCE	abc(n-1)	CME		
Total	SCT	$abcn - 1$			

**5.2. Modelo 2.** En una serie de establecimientos de una cadena de tiendas de ropa se contabilizan las ventas realizadas de ropa de verano y de ropa de invierno. Se anotan, además los empleados que realizan las ventas.

Hay tres factores:  $A \equiv$  Tienda,  $B \equiv$  Tipo de ropa,  $C \equiv$  Empleado

El esquema es el siguiente:

Factor A											
A <sub>1</sub>				A <sub>2</sub>				A <sub>3</sub>			
Factor B				Factor B				Factor B			
B <sub>1</sub>		B <sub>2</sub>		B <sub>1</sub>		B <sub>2</sub>		B <sub>1</sub>		B <sub>2</sub>	
C <sub>1</sub>	C <sub>2</sub>	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>5</sub>	C <sub>6</sub>

Se observa, aquí, que el factor  $C$  (empleado) no está anidado en el cruce de los factores  $A$  y  $B$  porque no está combinado con una única combinación  $(i, j)$ , ya que se supone que están los mismos empleados en verano e invierno. Así,

$$\begin{aligned}
 \text{Empleados} &\subset \text{Tiendas} \Rightarrow C \subset A \\
 \text{Empleados} \times \text{Ropa} &\Rightarrow C \times B \\
 \text{Tiendas} \times \text{Ropa} &\Rightarrow A \times B
 \end{aligned}$$

5.2.1. *Modelo matemático.* El modelo estadístico para este tipo de diseño está dado por:

$$\begin{aligned}
 Y_{ijkl} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{k(i)} + (\gamma\beta)_{k(i)j} + \varepsilon_{l(ijk)}; \\
 i &= 1, \dots, a; \quad j = 1, \dots, b; \quad k = 1, \dots, c; \quad l = 1, \dots, n
 \end{aligned}$$

sujeto a las restricciones  $\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = 0, \forall (i)$   
 $\sum_{k=1}^c \gamma_{k(i)} = 0, \sum_{k=1}^c (\gamma\beta)_{k(i)j} = \sum_{j=1}^b (\gamma\beta)_{k(i)j} = 0.$  La tabla ANOVA es

Tabla 2. ANOVA modelo 2					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A	SCA	a-1	CMA	CMA/CME	$P(F > F_0^A)$
B	SCB	b-1	CMB	CMB/CME	$P(F > F_0^B)$
AB	SCAB	(a-1)(b-1)	CMAB	CMAB/CME	$P(F > F_0^{AB})$
C(A)	$SC_{C(A)}$	a(c-1)	$CM_{C(A)}$	$CM_{C(A)}/CME$	$P(F > F_0^{C(A)})$
BC(A)	$SC_{BC(A)}$	a(b-1)(c-1)	$CM_{BC(A)}$	$CM_{BC(A)}/CME$	$P(F > F_0^{BC(A)})$
Error	SCE	abc(n-1)	CME		
Total	SCT	abcn - 1			

### 5.2.2. Ejemplo: Factores cruzados y anidados.

- (1) **Problema.** Consideremos el caso de un ingeniero industrial que quiere comparar dos métodos de ensamble (factor A), utilizando dos arreglos diferentes del área de trabajo (factor B) y cuatro operadores (factor C). Como las áreas de trabajo donde se implementa cada arreglo están físicamente separadas, se deben utilizar cuatro operadores diferentes en cada arreglo. En otras palabras, los operadores están anidados en los arreglos. Mientras que en los métodos de ensamble están cruzados con los arreglos y con los operadores. En este diseño, conocido como 5,6factorial-anidado, los datos se pueden acomodar como en la figura 2, la variable de respuesta es el tiempo de ensamble.

Ensamble 1								Ensamble 2							
Arreglo 1				Arreglo 2				Arreglo 1				Arreglo 2			
Operador				Operador				Operador				Operador			
o1	o2	o3	o4	o5	o6	o7	o8	o9	o10	o11	o12	o13	o14	o15	o16
5.6	4.5	6.0	4.3	6.3	4.7	5.1	5.2	5.4	5.0	6.1	5.5	6.6	4.9	4.8	5.5
4.9	4.7	6.3	4.2	6.5	4.7	4.7	5.4	5.6	5.3	6.5	5.7	6.7	5.3	5.4	6.0
5.4	5.1	5.2	4.0	6.4	5.0	4.8	6.4	4.9	5.4	5.8	5.9	6.6	5.3	5.3	5.8
Figura 2. Diseño factorial anidado								Figura 2. Figura 2. Diseño factorial anidado							

- (2) *Variable respuesta.*  $Y_{ijkl}$  = tiempo de ensamble en la medición  $l$ , con el método  $i$  en el arreglo  $j$  utilizado por el trabajador  $k$
- (3) *Factores:* Hay tres factores diferenciados. Factor A: el método de ensamble con dos niveles ( $a = 2$ ). Factor B: el tipo de arreglo con dos niveles ( $b = 2$ ). Factor C: los operadores que realizan los ensambles en cada arreglo, hay cuatro trabajadores ( $c = 4$ ) por cada arreglo.
- (4) *Relación entre los factores:* El factor A se cruza con todos los niveles del factor B y con los distintos niveles del factor C. Además el factor C está anidado en el factor B ya que los operadores que realizan los ensambles por arreglos son distintos. Luego los tratamientos son cada uno de los operadores en cada arreglo cruzado con cada método de ensamble, es decir, cuarenta y ocho tratamientos.
- (5) *Unidades experimentales:* Cada uno de los ensamblejes, en total 48.
- (6) *Módelo matemático asociado:*

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{k(j)} + (\alpha\gamma)_{ik(j)} + \varepsilon_{l(ijk)};$$

$$i = 1, 2; j = 1, 2; k = 1, 2, 3, 4; l = 1, 2, 3.$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto medio adicional debido al ensamble  $i$ ,  $\beta_j$  es el efecto medio adicional debido al arreglo  $j$ ,  $(\alpha\beta)_{ij}$  es el efecto medio adicional debido al método de ensamble  $i$  y al arreglo  $j$ ,  $\gamma_{k(j)}$  es el efecto medio adicional debido al operador  $k$  anidado en el  $j$ -ésimo arreglo,  $(\alpha\gamma)_{ik(j)}$  es efecto medio adicional debido al método de ensamble  $i$  con el  $k$  operador anidado en el  $j$ -ésimo arreglo y  $\varepsilon_{l(ijk)}$  es el error aleatorio asociado con la observación  $l$ -ésima del tratamiento  $ijk$ .

Sujeto a las restricciones  $\sum_{i=1}^2 \alpha_i = \sum_{j=1}^2 \beta_j = \sum_{i=1}^2 (\alpha\beta)_{ij} = \sum_{j=1}^2 (\alpha\beta)_{ij} = 0$ ,

$\forall (i, j) \sum_{k=1}^4 \gamma_{k(ij)} = 0, \sum_{k=1}^c (\gamma\beta)_{k(i)j} = \sum_{j=1}^b (\gamma\beta)_{k(i)j} = 0, \varepsilon_{ijkl} \sim N(0, \sigma^2)$   
independientes

(7) *Hipótesis del modelo:* se prueban las hipótesis

$$\begin{aligned} H_0 : & \alpha_i = 0 \\ H_0 : & \beta_j = 0 \\ H_0 : & (\alpha\beta)_{ij} = 0 \\ H_0 : & (\alpha\gamma)_{ik(j)} = 0 \end{aligned}$$

contra las alternativas usuales.

- (8) En la tabla 3 se muestra la tabla de ANOVA de este modelo declarando el factor operador aleatorio. Es muy importante no olvidarse de especificar si los factores son aleatorios o fijos, porque eso puede cambiar de manera radical las conclusiones. En esta tabla se observa que son significativos los efectos operador (arreglo), arreglo (Ensamble $\times$ Operador) y prácticamente ensamble (dado el nivel tan pequeño del valor- $p$ ). La razón F se obtiene considerando al factor operador aleatorio, con lo que la mayoría de los cuadrados medios se divide entre el  $CM$  del efecto anidado de menor jerarquía:  $C(AB)$  con excepción de  $A : Arreg$  que se divide con el  $CM$  de  $C(B)$  y el propio  $C(AB)$  que se contrasta con el  $CM$  del error. Por supuesto que es más fácil usar un software con una opción de ANOVA que considere factores anidados y cruzados.

Tabla 3. Anova anidado para tiempo de ensamble. Ejemplo 5.2.2					
variabilidad	SC	GL	CM	F <sub>0</sub>	valor-p
A:Ensamble	2.04187	1	2.04187	5.94	0.0507
B:Arreglo	0.775208	1	0.77520	0.34	0.5793
AB	0.316875	1	0.316875	0.92	0.3741
C(B):Oper(Arreg)	13.5529	6	2.25882	6.57	0.0187
C(AB)	2.06292	6	0.343819	3.38	0.0107
Error	3.2533	32	0.1017		
Total	22.0031	47			

Si no se cuenta con un software estadístico que obtenga directamente el ANOVA de la tabla 3, éste se puede obtener a partir de la tabla de ANOVA del análisis factorial que ignora el anidamiento, pero considerando que los



términos anidado, son sumas de dos términos factoriales, específicamente,

$$(5.1) \quad \begin{aligned} SC_{Oper(Arreg)} &= SC_{Oper} + SC_{Oper \times Arreg} \\ SC_{Ensa \times Oper(Arreg)} &= SC_{Ensa \times Oper} + SC_{Ensa \times Oper \times Arreg} \end{aligned}$$

5.2.3. *Ejemplo ilustrativo de un diseño anidado cruzado: posibilidades múltiples.* Una cadena de alimentación dispone de dos tiendas en una ciudad. El horario de atención al público en ambas es de 10:00 a 13:30 horas en turno de mañana y de 17:00 a 20:30 en turno de tarde. las dos tiendas estan abiertas de lunes a sábado. Se tiene interés en estudiar como se distribuyen las ventas a lo largo de la semana, así como en comparar las dos tiendas, distinguiendo los dos turnos, y se obtuvieron los siguientes datos:

Primera tienda						
Turno	Lunes	Martes	Miercoles	Jueves	Viernes	Sábado
Mañana	1300	1465	1283	2221	3804	5643
	2645	1642	1643	1842	3462	6265
	2201	1301	1469	2062	4241	5823
Tarde	1405	985	2003	2465	3660	2421
	1646	786	1626	3226	3182	2823
	1828	1221	1587	2782	4425	2085

Segunda tienda						
Turno	Lunes	Martes	Miercoles	Jueves	Viernes	Sábado
Mañana	2383	1664	2026	2403	3360	6027
	2068	2225	1665	2364	3824	5664
	1986	2405	1384	2183	4465	6302
Tarde	969	1048	2243	2247	4606	3186
	1327	1220	1684	2601	4385	3227
	788	1163	1507	2887	4165	2468

Se pide:

- (1) Escribir un modelo lineal que refleje la situación planteada y la aportación a las ventas de las distintas fuentes de variación. Obtener la tabla ANOVA y, a partir de los resultados obtenidos en ella, contestar las siguientes preguntas: (a) ¿Se puede decir que en la primera tienda las ventas son menores que en la segunda? (b) ¿Dependen las ventas realizadas del día de la semana? ¿Cuál es el día de mayor venta? ¿Y el de menor venta? ¿Son iguales las ventas en los dos turnos? ¿Dependen las diferencias de ventas en los turnos, del día de la semana?  
En cada tienda trabajan varios vendedores y se quiere completar el estudio comparando las ventas efectuadas por ellos. Con este fin se contabilizan las operaciones realizadas por tres vendedores de cada tienda, durante

ocho semanas distintas. los datos recogidos son:

Primera tienda			Segunda tienda		
Javier	Ángel	Miguel	Felipe	Carlos	Emilio
745	975	831	1094	1062	1033
700	804	915	1021	964	1015
752	1090	888	963	881	981
801	973	998	974	990	1064
703	795	899	903	814	685
680	843	784	910	845	720
659	915	765	834	763	752
773	900	770	802	782	744

- (2) Identificar los factores involucrados en el estudio y la relación entre ellos.  
¿Qué se puede decir de la eficacia de los vendedores?

A la vista del resultado obtenido en el apartado anterior y con la información adicional de que las cuatro primeras semanas correspondían al mes de diciembre y las cuatro últimas al mes de enero, se replanteó el experimento en el sentido de considerar la posible influencia del factor mes sobre la variable respuesta. En esa línea se seleccionaron aleatoriamente cuatro días del mes de diciembre y otros cuatro del mes de enero y se recogieron los datos de las ventas de los seis vendedores anteriores en esos días. Dichos datos se presentan en la tabla siguiente:

Mes	Primera tienda			Segunda tienda		
	Javier	Ángel	Miguel	Felipe	Carlos	Emilio
Diciembre	124	162	138	182	177	172
	116	134	152	170	161	169
	125	182	148	160	147	163
	133	162	166	162	165	177
Enero	117	132	150	150	136	114
	113	140	131	152	141	120
	110	152	127	139	127	125
	129	150	128	134	130	124

- (3) ¿Cuál es ahora el modelo asociado al experimento? Analizarlo detalladamente.  
(4) ¿Qué resultados se obtienen en el modelo anterior si los vendedores se hubieran elegidos aleatoriamente entre todos los vendedores de las tiendas?

Análisis del modelo trifactorial.

- (1) *Variable respuesta.*  $Y_{ijkl}$  =  $l$ -ésima venta en la tienda  $i$  en el día  $j$  en el turno  $k$   
(2) *Factores:* Hay tres factores diferenciados. Factor A: tienda con dos niveles ( $a = 2$ ). Factor B: día con seis niveles ( $b = 6$ ). Factor C: turno con dos niveles ( $c = 2$ )  
(3) *Módulo trifactorial de efectos fijos:*

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

$$i = 1, 2; j = 1, 2, 3, 4, 5, 6; k = 1, 2; \text{ y } l = 1, 2, 3$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto medio adicional debido a la tienda  $i$ ,  $\beta_j$  es el efecto medio adicional debido al día  $j$ ,  $\gamma_k$  efecto del  $k$ -ésimo turno,  $(\alpha\beta)_{ij}$  es el efecto medio adicional debido a la interacción entre tienda y día,  $(\alpha\gamma)_{ik}$  es el efecto medio adicional debido a la interacción entre tienda y turno,  $(\beta\gamma)_{jk}$  es el efecto medio adicional debido a la interacción entre día y turno,  $(\alpha\beta\gamma)_{ijk}$  efecto de interacción entre los tres factores y  $\varepsilon_{ijkl}$  es el error aleatorio asociado con la observación  $l$ -ésima del tratamiento  $ijk$ . Sujeto a las restricciones

$$\begin{aligned}\sum_{i=1}^2 \alpha_i &= \sum_{j=1}^6 \beta_j = \sum_{k=1}^2 \gamma_k = \sum_{i=1}^2 (\alpha\beta)_{ij} = \sum_{j=1}^6 (\alpha\beta)_{ij} = \sum_{i=1}^2 (\alpha\gamma)_{ik} = 0 \\ &= \sum_{k=1}^2 (\alpha\gamma)_{ik} = \sum_{j=1}^6 (\beta\gamma)_{jk} = \sum_{k=1}^2 (\beta\gamma)_{jk} = 0 \\ &= \sum_{i=1}^2 (\alpha\beta\gamma)_{ijk} = \sum_{j=1}^6 (\alpha\beta\gamma)_{ijk} = \sum_{k=1}^2 (\alpha\beta\gamma)_{ijk} = 0, \\ \varepsilon_{ijkl} &\approx N(0, \sigma^2) \text{ independientes}\end{aligned}$$

(4) *Hipótesis del modelo:* se prueban las hipótesis

$$\begin{aligned}H_0 : \alpha_i = 0 \quad H_0 : \gamma_k = 0 \quad H_0 : (\alpha\gamma)_{ik} = 0 \quad H_0 : (\alpha\beta\gamma)_{ijk} = 0 \\ H_0 : \beta_j = 0 \quad H_0 : (\alpha\beta)_{ij} = 0 \quad H_0 : (\beta\gamma)_{jk} = 0\end{aligned}$$

(5) A continuación se presenta en la tabla 1, el Anova para este diseño, donde se muestra que los efectos ABC, AC y AB son pocos significativos a un nivel de  $\alpha = 0.05$  por lo cual el modelo debe ser modificado eliminando estos tres factores.

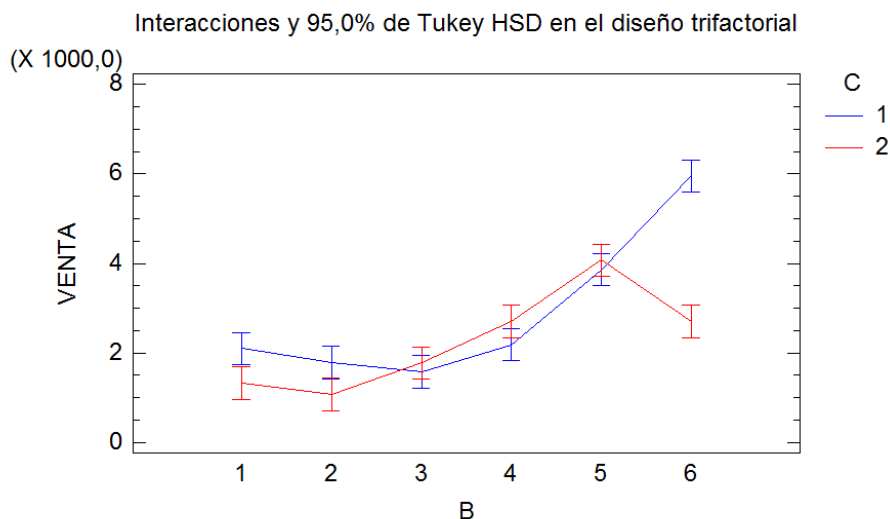
Tabla 1. ANOVA para el diseño trifactorial					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A	443525	1	443525	3,64	0,0626
B	9,46802E7	5	1,893E7	155,2	0,0000
C	7,24092E6	1	7,24092E6	59,35	0,0000
AB	881239	5	176248	1,44	0,2256
AC	88270	1	88270	0,72	0,3992
BC	2,88644E7	5	5,77288E6	47,31	0,0000
ABC	1,06321E6	5	212641	1,74	0,1429
Error	5,85665E6	48	122014		
Total	1,39118E8	71			

(6) En la tabla 2, aparece el mejor Anova para este diseño, hay evidencia en la muestra para afirmar que las ventas promedio difieren significativamente

por los días y los turnos a un nivel de  $\alpha = 0.05$ .

Tabla 2. Mejor Anova para el diseño trifactorial					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A	443525	1	443525	3,32	0,0736
B	9,46802E7	5	1,893E7	141,61	0,0000
C	7,24092E6	1	7,24092E6	54,15	0,0000
BC	2,88644E7	5	5,77288E6	43,71	0,0000
Error	7,88937E6	59	133718		
Total	1,39118E8	71			

- (7) El gráfico de interacción para el diseño trifásico revela que el día sábado a la jornada de la mañana se obtienen las ventas más altas y el día de menor venta es el lunes en la jornada de la tarde sin importar la tienda.



- (8) El Modelo corregido para este diseño es el siguiente:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \varepsilon_{ijkl}$$

$$i = 1, 2; j = 1, 2, 3, 4, 5, 6; k = 1, 2; \text{ y } l = 1, 2, 3$$

Modelo de factores anidados.

- (1) *Variable respuesta:*  $Y_{ijk}$  = venta en la semana  $k$  por el vendedor  $j$  en la tienda  $i$  en el día.
- (2) *Factores:* Hay dos factores diferenciados. Factor A: tienda con dos niveles ( $a = 2$ ). Factor B: vendedor con seis niveles ( $b = 6$ ). Pero todos los niveles del factor tienda no están cruzados con todos los niveles del factor vendedor, sino que cada tienda se cruza sólo con los niveles de vendedor correspondiente a esa tienda. Es decir, el factor vendedor está anidado en el factor tienda. Por otro lado, podríamos considerar el factor semana (Factor C), pero el número de observaciones de que disponemos no nos

permite hacer el estudio adecuado. Sin embargo es conveniente tener en cuenta esta fuente de variación, por lo que vamos a considerar la semana como efecto bloque ( $c = 8$ )

(3) *Módulo de factores anidados en bloque*

$$\begin{aligned} Y_{ij(i)k} &= \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ij(i)k} \\ i &= 1, 2; j = 1, 2, 3; k = 1, \dots, 8 \end{aligned}$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto medio adicional debido a la tienda  $i$ ,  $\beta_{j(i)}$  es el efecto medio adicional debido al vendedor  $j$  en la tienda  $i$ ,  $\gamma_k$  efecto bloque de la  $k$ -ésima semana y  $\varepsilon_{ij(i)k}$  es el error aleatorio asociado con la observación  $l$ -ésima del tratamiento  $ijk$ . Sujeto a las restricciones

$$\begin{aligned} \sum_{i=1}^2 \alpha_i &= 0, \quad \forall i, \sum_{j=1}^6 \beta_{j(i)} = 0, \quad \sum_{k=1}^2 \gamma_k = 0 \\ \varepsilon_{ij(i)k} &\approx N(0, \sigma^2) \text{ independientes} \end{aligned}$$

(4) *Hipótesis del modelo:* se prueban las hipótesis

$$\begin{aligned} H_0 : \quad \alpha_i &= 0 \\ H_0 : \quad \beta_{1(1)} &= \beta_{2(1)} = \dots = \beta_{3(2)} = 0 \end{aligned}$$

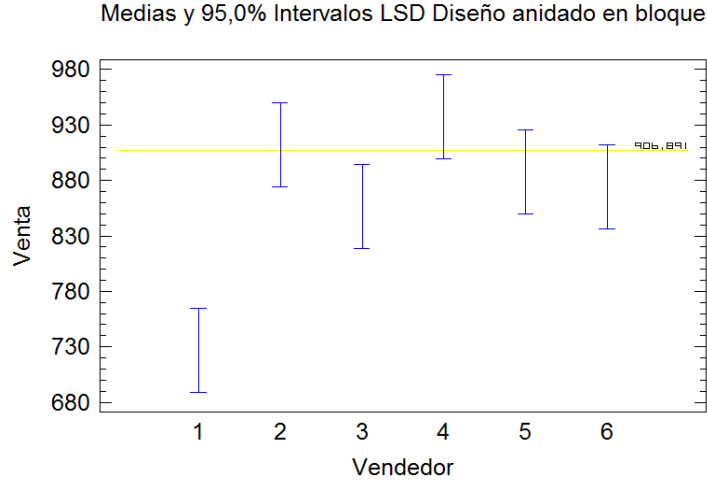
(5) Para analizar el modelo, obtenemos la tabla Anova que se presenta en la tabla 3, a un nivel de  $\alpha = 0.05$  se prueba que haber considerado semana como factor de bloque fue un acierto por lo cual el modelo debe estar bien planteado. Existen diferencias significativas entre las tiendas y entre los vendedores en cada tienda.

Tabla 3. ANOVA para el diseño anidado en bloque					
Fuente	SC	gl	CM	$F_0$	valor - p
A:Tienda	55896,8	1	55896.8	10.00	0,0032
B(A):Vendedor(Tienda)	162426	4	40606.2	7.27	0,0002
C:Semana	267839	7	38262.7	6.85	0,0000
Error	195571	35	5587.74		
Total	681732	47			

(6) El siguiente gráfico nos muestra que existe diferencia significativa entre los vendedores de la tienda 1, mientras que no existe diferencia entre los vendedores de la tienda 2.

Modelo de factores anidados y cruzados. A partir del resultado anterior, se plantea la posibilidad de modificar el modelo en el sentido de considerar el factor mes con dos niveles, en lugar del bloque semana.

- (1) *Variable respuesta.*  $Y_{ijkl}$  = venta en el día  $l$  del  $k$ -ésimo mes por el vendedor  $j$  en la tienda  $i$ .
- (2) *Factores:* Hay tres factores diferenciados. Factor A: tienda con dos niveles ( $a = 2$ ). Factor B: día con tres niveles en cada tienda ( $b = 3$ ). Factor C: mes con dos niveles ( $c = 2$ ). El factor vendedor está anidado en el factor tienda. Los niveles del factor mes se cruzan con todos los niveles de tienda y de vendedor, por lo que estamos ante un modelo de factores anidados y cruzados.



(3) *Módulo factores anidados y cruzados.*

$$Y_{ij(i)kl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + (\gamma\beta)_{kj(i)} + \varepsilon_{ij(i)kl}$$

$$i = 1, 2; j = 1, 2, 3; k = 1, 2; \text{ y } l = 1, 2, 3, 4$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto medio adicional debido a la tienda  $i$ ,  $\beta_{j(i)}$  es el efecto medio adicional debido al vendedor  $j$  en la tienda  $i$ ,  $\gamma_k$  efecto del  $k$ -ésimo mes,  $(\alpha\gamma)_{ik}$  es el efecto medio adicional debido a la interacción entre tienda y mes, y  $\varepsilon_{ijkl}$  es el error aleatorio. Sujeto a las restricciones

$$\sum_{i=1}^2 \alpha_i = 0, \sum_{j=1}^3 \beta_{j(i)} = 0, \text{ para cada } i \sum_{k=1}^2 \gamma_k = 0$$

$$\sum_{i=1}^2 (\alpha\gamma)_{ik} = \sum_{k=1}^2 (\alpha\gamma)_{ik} = 0$$

$$\varepsilon_{ijkl} \approx N(0, \sigma^2) \text{ independientes}$$

(4) *Hipótesis del modelo:* se prueban las hipótesis

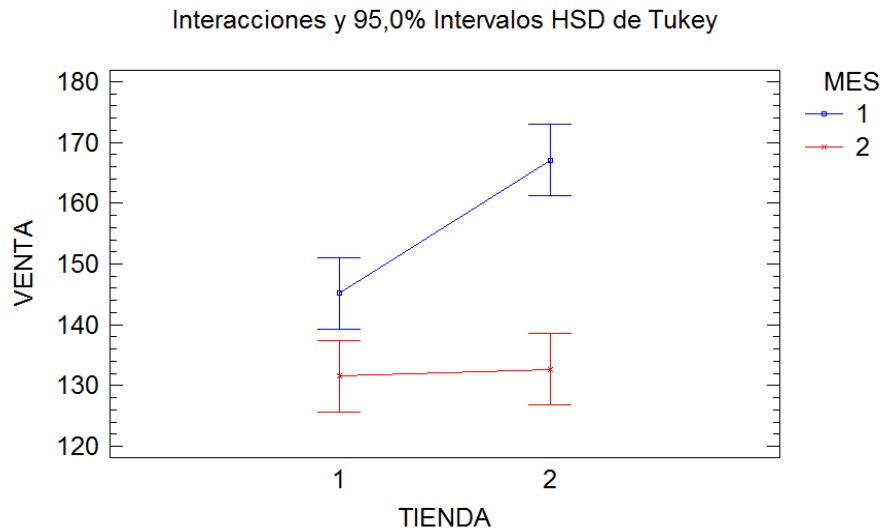
$$H_0: \alpha_i = 0, \forall i$$

$$H_0: \forall j, \beta_{j(i)} = 0, \forall i$$

$$H_0: \forall k, \gamma_k = 0$$

$$H_0: \forall i, k (\alpha\gamma)_{ik} = 0$$

(5) Para analizar el modelo, obtenemos la tabla Anova que se presenta en la tabla 4, a un nivel de  $\alpha = 0.05$  se prueba que haber considerado mes como tercer factor fue un acierto por lo cual el modelo debe estar bien planteado. Existen diferencias significativas entre las tiendas y entre los



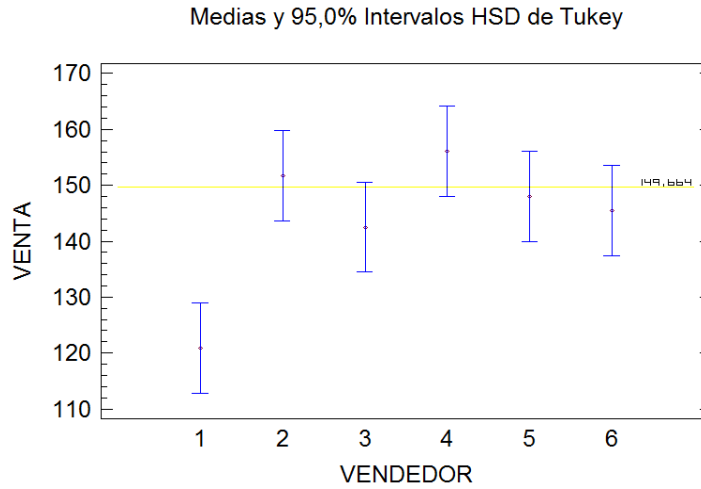
vendedores en cada tienda.

Tabla 4. ANOVA para el diseño anidado en cruzado					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A:Tienda	1587.0	1	1587.0	13.70	0,0006
B(A):Vendedor(Tienda)	4511.0	4	1127.75	9.74	0,0000
C:Mes	6912.0	1	6912.0	59.67	0,0000
A*C	1302.08	1	1302.08	11.24	0,0018
C*B(A):Mes×Vendedor(Tienda)	821.167	4	205.292	1.94	0.1251
Error	3812.0	36	105.889		
Total	18945,3	47			

Tabla 4.1. Mejor ANOVA para el diseño anidado en cruzado					
Fuente	SC	gl	CM	$F_0$	valor - $p$
A:Tienda	1587.0	1	1587.0	13.70	0,0006
B(A):Vendedor(Tienda)	4511.0	4	1127.75	9.74	0,0000
C:Mes	6912.0	1	6912.0	59.67	0,0000
A*C	1302.08	1	1302.08	11.24	0,0018
Error	4633,17	40	115.829		
Total	18945,3	47			

- (6) Los análisis gráficos nos muestran que la tienda 2 en el mes de diciembre tiene mayores venta que la tienda 1, no hay diferencia en las ventas en la tienda 1 en diciembre ni enero. Entre los vendedores de la tienda 2, no hay diferencia significativa alguna, en cambio que si la hay entre los distintos vendedores. El vendedor de la tienda 1 es el de menor ventas comparados con el resto de vendedores de ambas tiendas.

Modelo de efectos mixtos. Consideremos ahora que los niveles del factor vendedor se eligen aleatoriamente en cada tienda, entre los vendedores de esa tienda. Describiremos completamente el proceso como antes para que se observe las diferencias.



- (1) *Variable respuesta.*  $Y_{ijkl}$  = venta en el día  $l$  del  $k$ -ésimo mes por el vendedor  $j$  en la tienda  $i$ .
- (2) *Factores:* Hay tres factores diferenciados. Factor A: tienda con dos niveles ( $a = 2$ ). Factor B: día con tres niveles en cada tienda ( $b = 3$ ). Factor C: mes con dos niveles ( $c = 2$ ). El factor vendedor está anidado en el factor tienda. Los niveles del factor mes se cruzan con todos los niveles de tienda y de vendedor, por lo estamos ante un modelo de factores anidados y cruzados.
- (3) *Módulo factores anidados y cruzados.*

$$\begin{aligned}
 Y_{ij(i)kl} &= \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + (\gamma\beta)_{kj(i)} + \varepsilon_{ij(i)kl} \\
 i &= 1, 2; j = 1, 2, 3; k = 1, 2, \text{ y } l = 1, 2, 3, 4
 \end{aligned}$$

donde  $\mu$  es la media general,  $\alpha_i$  es el efecto medio adicional debido a la tienda  $i$ ,  $\beta_{j(i)}$  es el efecto medio adicional debido al vendedor  $j$  en la tienda  $i$ ,  $\gamma_k$  efecto del  $k$ -ésimo mes,  $(\alpha\gamma)_{ik}$  es el efecto medio adicional debido a la interacción entre tienda y mes, y  $\varepsilon_{ijkl}$  es el error aleatorio. Sujeto a las restricciones

$$\begin{aligned}
 \sum_{i=1}^2 \alpha_i &= 0, \quad \sum_{j=1}^3 \beta_{j(i)} = 0, \quad \text{para cada } i \quad \sum_{k=1}^2 \gamma_k = 0 \\
 \sum_{i=1}^2 (\alpha\gamma)_{ik} &= \sum_{k=1}^2 (\alpha\gamma)_{ik} = 0 \\
 \varepsilon_{ijkl} &\approx N(0, \sigma^2) \text{ independientes}
 \end{aligned}$$



- (4) *Hipótesis del modelo:* se prueban las hipótesis

$$H_0: \alpha_i = 0, \forall i$$

$$H_0: \beta_j \approx N(0, \sigma_\beta^2)$$

$$H_0: \forall k, \gamma_k = 0$$

$$H_0: \forall i, k (\alpha\gamma)_{ik} = 0$$

- (5) Para analizar el modelo, obtenemos la tabla Anova que se presenta en la tabla 5, a un nivel de  $\alpha = 0.05$  se prueba que al haber considerado al factor vendedor como factor aleatorio no se observan diferencias entre las tiendas, pero si las hay entre los distintos vendedores en cada tienda y entre los meses.

Tabla 5. ANOVA para el diseño anidado en cruzado					
Fuente	SC	gl	CM	$F_0$	valor - p
A:Tienda	1587.0	1	1587.0	1.41	0,3012
B(A):Vendedor(Tienda)	4511.0	4	1127.75	5.49	0,0638
C:Mes	6912.0	1	6912.0	33.67	0,0044
AC	1302.08	1	1302.08	6.32	0,0655
C*B(A):Mes×Vendedor(Tienda)	821.167	4	205.292	1.94	0.1251
Error	4633,17	40	115.829		
Total	18945,3	47			

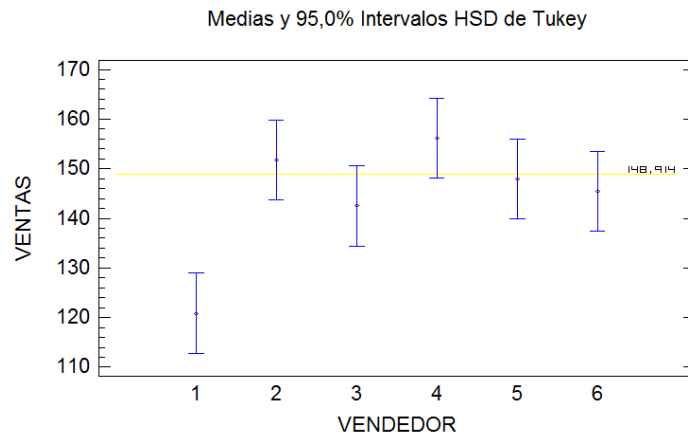
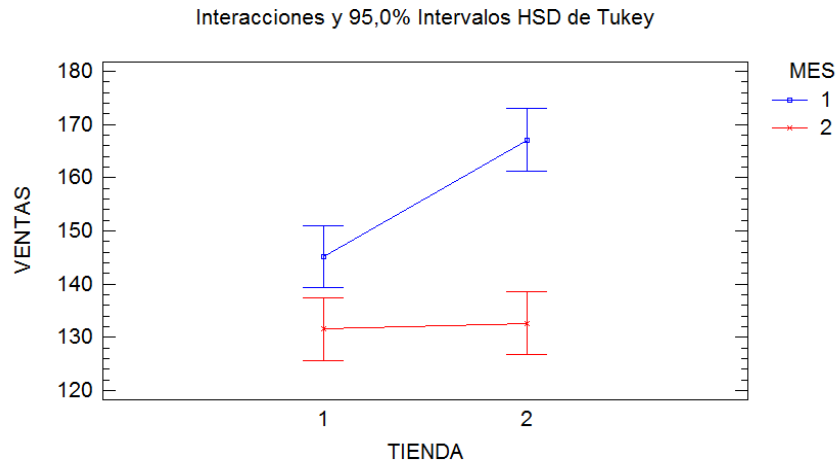
Tabla 5. ANOVA para el diseño anidado en cruzado					
Fuente	SC	gl	CM	$F_0$	valor - p
A:Tienda	1587.0	1	1587.0	1.41	0,3012
B(A):Vendedor(Tienda)	4511.0	4	1127.75	9.74	0,0000
C:Mes	6912.0	1	6912.0	59.67	0,0000
AC	1302.08	1	1302.08	11.24	0,0018
Error	4633,17	40	115.829		
Total	18945,3	47			

- (6) Los análisis gráficos nos muestran que en diciembre la tienda 2 tiene las mayores venta que la tienda 1, no hay diferencia en las ventas en la tienda 1 en diciembre ni enero. Entre los vendedores de la tienda 2, no hay diferencia significativa alguna, en cambio que si la hay entre los distintos vendedores. El vendedor de la tienda 1 es el de menor ventas comparados con el resto de vendedores de ambas tiendas.

## 6. Diseño en parcelas divididas (*split-plot*)

Este es un diseño experimental combinado que resulta útil cuando al estudiar simultáneamente varios factores, alguno o algunos de ellos deben ser aplicados sobre unidades experimentales relativamente grandes, pudiéndose aplicar el otro o los otros en unidades experimentales menores, dentro de las unidades mayores. El caso más sencillo es aquél en el que se tienen sólo dos factores, asignando los niveles de uno de ellos a las unidades mayores y los niveles del otro a las subunidades. A las unidades experimentales mayores suele llamárseles *parcelas grandes* o *parcelas principales* y a las unidades experimentales menores se le llama *subparcelas* o *subunidades*.

EJEMPLO 21. Suponga que se quieren comparar varios fertilizantes, cada uno de ellos se aplica a una parcela completa o grande, pero dicha parcela a su vez



se divide en parcelas pequeñas o subparcelas para estudiar diferentes variedades de un cultivo. De esta manera, se tiene el factor de parcela (fertilizante) y el factor de subparcela (variedad). El hecho relevante es que los fertilizante, se aleatorizan al asignarlos a las parcelas y las variedades al asignarlas a las subparcela (pero dentro de cada parcela); por ello, no se tiene, ni es conveniente una aleatorización completa de fertilizantes y variedades, como se haría con un diseño factorial. Una representación gráfica de las parcelas y subparcelas en este experimento agrícola, considerando la asignación aleatoria de tres fertilizantes (F1, F2 y F3), cuatro variedades (V1, V2, V3 y V4) en tres réplicas (bloques I, II, III) se muestra en

la figura 4. Así cada columna es una parcela.

Figura 4. Diseño en parcela divididas								
Bloque I			Bloque II			Bloque III		
F2	F1	F3	F1	F2	F3	F2	F3	F1
V1	V3	V4	V2	V2	V1	V3	V4	V4
V3	V4	V2	V4	V1	V3	V2	V1	V1
V4	V1	V3	V1	V3	V2	V4	V3	V4
V2	V2	V1	V3	V4	V4	V1	V1	V3

Situaciones similares a la descrita se presentan en otras áreas experimentales, donde las parcelas y subparcelas ya no son físicamente pedazos de tierra, pero la idea es la misma: *habrá tratamientos que se tienen que aplicar de manera aleatoria a unidades experimentales grandes (parcelas), y habrá otros tratamientos que se pueden aplicar al azar en unidades experimentales chicas (subparcelas), que resultan al dividir las primeras.* Así, los primeros tratamientos son combinaciones de niveles de los factores en las parcelas y los segundos son combinaciones de niveles de los factores en las subparcelas. *A menudo, los tratamientos en la parcela se aplican antes que los tratamientos en la subparcela.* Estos diferentes tamaños de unidades experimentales provocan la aparición de dos estructuras de error que deben tomarse en cuenta al momento del análisis estadístico.

Se debe notar que además de que los niveles de los diferentes factores son asignados a unidades experimentales de diferentes tamaños, está implícito también un número diferente de repeticiones. El número de repeticiones para el factor asignado a las subunidades es  $r \times a$ , siendo  $r$  el número de repeticiones del factor asignado a las unidades principales y  $a$  su número de niveles.

El factor correspondiente a las parcelas principales puede asignarse a éstas utilizando cualquiera de los esquemas de aleatorización básicos: *Completamente al Azar*, en *Bloques al Azar* o en *Cuadro Latino*. El factor correspondiente a las subparcelas se asigna al azar dentro de cada parcela principal; en tal sentido, las parcelas principales son análogas a bloques, solo que por asignarse a éstas los niveles de un efecto fijo y por existir repeticiones de las mismas, es posible evaluar tanto los efectos principales del factor asignado a las mismas como su posible interacción con en el otro factor.

En adelante nos concentraremos en el caso más sencillo de un Diseño Parcelas Divididas, es decir aquél con sólo dos factores. Todos los resultados obtenidos son generalizables a casos más complejos en los que los tratamientos asignados a las unidades principales, a las subunidades o a ambas estén conformados a su vez por las combinaciones de los niveles de dos o más factores.

Supóngase que se quiere realizar un experimento que involucre dos factores: un factor  $A$  con tres niveles ( $A_1, A_2, A_3$ ) y un factor  $B$  con dos ( $B_1, B_2$ ). Se muestran a continuación posibles esquemas de aleatorización considerando los diseños más comunes (Parcelas Divididas y otros), suponiendo que ambos factores tienen igual importancia relativa y que se desean evaluar tanto sus efectos principales como su posible interacción, es decir, descartando la opción de tomar alguno de éstos como factor de bloqueo. Para facilitar la ilustración de los posibles esquemas de aleatorización, se considerarán sólo dos repeticiones, excepto donde el diseño exige tantas repeticiones como tratamientos (Cuadro Latino).

(1) DCA

$A_2B_2$	$A_2B_1$	$A_1B_1$	$A_2B_1$
$A_1B_1$	$A_1B_2$	$A_3B_1$	$A_1B_2$
$A_3B_1$	$A_3B_2$	$A_2B_2$	$A_3B_2$

(2) DBCA

Bloque I	Bloque II
$A_1B_2$	$A_1B_1$
$A_3B_1$	$A_3B_2$
$A_2B_2$	$A_1B_2$
$A_1B_1$	$A_2B_2$
$A_2B_1$	$A_3B_1$
$A_3B_2$	$A_2B_1$

(3) Diseño Cuadro Latino (si el número de combinaciones de tratamientos no es muy alto)

$A_1B_1$	$A_2B_1$	$A_3B_2$	$A_2B_2$	$A_1B_2$	$A_3B_1$
$A_3B_2$	$A_2B_2$	$A_1B_2$	$A_3B_1$	$A_1B_1$	$A_2B_1$
$A_2B_1$	$A_3B_2$	$A_2B_2$	$A_1B_2$	$A_3B_1$	$A_1B_1$
$A_3B_1$	$A_1B_1$	$A_2B_1$	$A_3B_2$	$A_2B_2$	$A_1B_2$
$A_2B_2$	$A_1B_2$	$A_3B_1$	$A_1B_1$	$A_2B_1$	$A_3B_2$
$A_1B_2$	$A_3B_1$	$A_1B_1$	$A_2B_1$	$A_3B_2$	$A_2B_2$

(4) Diseño Parcelas Divididas

**6.1. Cuándo utilizar el diseño de parcelas divididas.** El usuario típico no suele distinguir cuándo y cómo puede utilizar el diseño en parcelas divididas y aprovechar sus ventajas, o cuando menos hacer el análisis correcto de una situación experimental que cae en este tipo de diseño. Existen diferentes motivaciones para considerar el diseño en parcelas divididas como el más apropiado. Por ejemplo, algunas situaciones específicas donde es recomendable este tipo de diseño son las siguientes:

- (1) *Cuando en un estudio experimental se tienen factores que son difíciles de cambiar de un nivel a otro, y también se tienen factores que son fáciles de mover.* En general, los factores difíciles de manipular dificultan el trabajo experimental en un diseño factorial, ya que se tienen que correr en forma aleatoria los tratamientos, lo cual obliga a cambiar con frecuencia los niveles de tales factores. La solución mas efectiva es aplicar un diseño en parcelas divididas, donde los factores difíciles de mover se asignen a las parcelas y los fáciles de mover a las subparcelas, minimizando de esta manera los cambios de nivel de los factores difíciles durante el experimento. Un ejemplo típico de esta situación se da en los procesos de horneado (de pasteles, de metal, etc.) donde se dificulta estar cambiando el nivel de temperatura del horno; una mejor estrategia es elegir al azar uno de los niveles de temperatura con los que se desea experimentar y correr en ella todas las combinaciones de los factores restantes. De esta manera, el factor temperatura define las parcelas grandes y el resto de los factores serían las parcelas chicas.
- (2) *Por su naturaleza, algunos factores requieren que la unidad experimental sea de gran tamaño mientras que otros factores se pueden manejar*

a menor escala. Este caso se mostró en la figura 3, donde se quieren comparar varios fertilizantes pero es técnicamente impráctico aplicarlos a parcela pequeñas, es decir, cada fertilizante sólo se puede aplicar en parcelas grandes. Éstas se dividen en subparcelas para estudiar diferentes combinaciones de otros factores, como variedades, prácticas de manejo, etc. Otro ejemplo de un factor de parcela grande es el fumigante, suponiendo que éste se tiene que aplicar desde una avioneta que no puede maniobrar en una parcela chica sin contaminar a la parcela de al lado.

- (3) *En ocasiones se requiere incluir un factor adicional en aras de un mayor alcance de las conclusiones del experimento, pero la importancia de este factor adicional es menor.* Por ejemplo, supongamos que se quieren comparar cuatro medicamentos para una enfermedad, y además controlar la dosis y el tiempo de aplicación en dos niveles cada uno. Pero si adicionalmente se considera que en estudios de este tipo la hora del día en la que se aplica el tratamiento puede influir en los resultados, entonces se incluye la hora en el experimento, dividiendo cada día en cuatro parcelas. Así, se aleatoriza a cuál de las cuatro parcelas (hora de aplicación) se asigna cada uno de los cuatro medicamentos. Dentro de cada parcela se aleatorizan las cuatro combinaciones de dosis y tiempo (subparcela) que corresponden a cada medicamento. Y esto se puede repetir varios días (*bloque*).
- (4) *Otro caso es cuando se espera de antemano, por conocimiento previo, que algunos factores tengan efectos grandes sobre la respuesta, mientras que para otros factores se espera un efecto pequeño.* En estas circunstancias, los niveles de los factores con efectos grandes se asignan de manera aleatoria a las parcelas y los niveles de los factores con efectos pequeños se asignan aleatoriamente a la subparcelas. En un diseño en parcelas divididas los efectos de los factores en las parcelas se estiman con menor precisión que los efectos de los factores en las subparcelas, ya que existe más variabilidad entre las primeras dado su mayor tamaño. De aquí que el factor con efecto grande se asigne a la parcela grande, puesto que no se requiere mucha precisión para detectar un efecto de este tipo.
- (5) *Muchos experimentos robustos que usan el (diseño con arreglo interno y externo propuesto por Taguchi son, por la manera en que fueron realizados, diseños en parcelas divididas.* Típicamente, cada combinación de los factores de proceso en el arreglo interno es una parcela, y cada combinación de los factores de ruido en el arreglo externo es una subparcela.

*Para complementar lo anterior, en un diseño en parcelas divididas es de suma importancia decidir cuáles factores se irán a la parcela y cuáles a la subparcela, ya que de esto depende el tamaño de la parcela que será usada con cada factor y la precisión con que se estimarán los efectos.* Todo esto, sin olvidar las prácticas de manejo y facilidad para mover niveles, que son criterios a tomar en cuenta. Otras consideraciones que deben asignarse a las subparcelas son: si un factor es muy importante, sus niveles se mueven fácilmente o requiere poca cantidad de material experimental.

**6.2. Ejemplo: Caso con dos factores.** Potcner y Kowalski (2004) describen un experimento que tiene que ver con la propiedad de resistencia de la madera al agua. En el experimento se estudian dos tipos de pretratamientos y cuatro marcas de barniz. La variable de respuesta es la resistencia. El experimento se realizó de

la siguiente manera: primero se selecciona al azar un pretratamiento y se aplica a un panel de madera. Después, el panel se parte en cuatro pedazos y a cada trozo se aplica una marca de barniz elegida al azar. Luego se aplica el segundo pretratamiento a otro panel, que se divide a su vez en cuatro pedazos donde se aplican los barnices al azar. Esto se hace tres veces. La resistencia obtenida en cada combinación de pretratamiento y barniz se muestra en la tabla 4. Por la manera en que se corrió el experimento, cada panel corresponde a una parcela y cada trozo de panel es una subparcela. De modo que en total se tienen seis parcelas y 24 subparcelas. Cada renglón son los datos observados en una parcela y cada casilla representa el dato en una subparcela.

Si no se supiera cómo se obtuvieron los datos, éstos se podrían tomar como un diseño en bloques completos al azar, o como un diseño factorial  $4 \times 2$  con tres réplicas, pero de acuerdo con el procedimiento experimental descrito, se trata de un diseño en parcelas divididas. Existen varios aspectos característicos del diseño en parcelas divididas que lo delatan como tal: los factores, tienen dos tamaños de unidades experimentales: el panel completo para los pretratamientos y los pedazos de panel para los barnices. El procedimiento de aleatorización consta de dos pasos. Además, el número de repeticiones no es el mismo para cada factor, ya que pretratamiento tiene tres repeticiones y barniz seis. Más adelante se analizan los datos de la tabla 4.

Tabla 4. Datos de resistencia de la madera al agua					
	Barnices				
Pretratamiento	B1	B2	B3	B3	Media
1	43.0	51.8	40.8	45.5	45.28
1	57.4	60.9	51.1	55.3	56.18
1	52.8	59.2	51.7	55.3	54.75
2	46.6	53.5	35.4	32.5	42.00
2	52.2	48.3	45.9	44.6	47.75
2	32.1	34.4	32.2	30.1	32.20

Los diseños en parcelas divididas presentan complicaciones dependiendo de la situación experimental. Los problemas más comunes ocurren cuando:

- (1) Los niveles del factor parcela están metidos en una estructura de diseño completamente al azar (como en el ejemplo 6.3), de bloques completos al azar o de cuadro latino. Mientras que los niveles del factor subparcela también tienen diferentes arreglos dentro de cada parcela.
- (2) Existen más de dos factores tanto a nivel de parcela como a nivel de subparcela, como ocurre en muchos diseños con arreglo interno y externo propuestos por Taguchi (Diseño robusto y ejemplo 6.2). Cuando son muchos los factores en la subparcela es necesario utilizar factoriales fraccionados, pero la estructura de confusión resultante es diferente a la fracción por tratarse de un diseño en parcelas divididas (Bingham et al., 2004).
- (3) La subparcela se parte en subsubparcelas para dar lugar a otro factor de interés, en cuyo caso el experimento se llama *diseño en parcelas doblemente divididas* (*Split-split-plot design*).

**6.3. Ejemplo: Caso con varios factores.** Consideremos un ejemplo del tipo diseño robusto con arreglos interno y externo, donde el objetivo del experimento

es diseñar una harina preparada para pastel que sea robusta a las variaciones de temperatura y tiempo de preparación por parte de los consumidores. Para ello se plantea el experimento de la tabla 5, donde se tienen tres factores de diseño (control): harina ( $H$ ), grasa ( $G$ ) y polvo de huevo ( $P$ ) acomodados en un factorial  $2^3$ . Los factores de ruido son temperatura del horno ( $T$ ) y tiempo de horneado ( $Ti$ ). La variable de respuesta es el resultado de la evaluación promedio de un panel de expertos en una escala de 1 a 7, donde siete es la máxima evaluación positiva. Para sintetizar, nos referimos a la variable como el sabor.

El experimento de la tabla 5 es un diseño en parcelas divididas porque se corrió de la siguiente manera: se amasó una cantidad suficiente de harina en cada combinación de los factores de diseño, y luego ésta se dividió en cuatro lotes con los cuales se prepararon cuatro pasteles que fueron horneados en las diferentes combinaciones de temperatura y tiempo. Para ser un experimento factorial se tendrían que haber preparado en orden aleatorio los 32 pasteles uno por uno. Pero se trata de un diseño en parcelas divididas porque sólo se prepararon en orden aleatorio ocho mezclas de ingredientes (parcelas), y cada una se dividió en cuatro partes (subparcelas) que se asignaron en orden aleatorio a una condición de horneado. Note que el horno se tuvo que usar 32 veces, pero seguramente existen otras maneras más eficientes de realizar las corridas en este experimento (véase Box y Jones, 2000; Bisgaard y Kulahci, 2001). Más adelante analizamos los datos de la tabla 5.

Tabla 5. Experimentos sobre el sabor de pasteles							
Arreglo interno			Arreglo externo				
H	G	P	T	-1	+1	-1	+1
			Ti	-1	-1	+1	+1
-1	-1	-1		1.1	1.4	1.0	2.9
+1	-1	-1		1.8	5.1	2.8	6.1
-1	+1	-1		1.7	1.6	1.9	2.1
+1	+1	-1		3.9	3.7	4.0	4.4
-1	-1	+1		1.9	3.8	2.6	4.7
+1	-1	+1		4.4	6.4	5.2	6.6
-1	+1	+1		1.6	2.1	2.3	1.9
+1	+1	+1		4.9	5.5	5.2	5.7

## 7. Modelo y análisis estadístico de los diseños en parcelas divididas

**7.1. El caso de dos factores.** El caso más simple de parcelas divididas con dos factores es similar al ejemplo 6.2 de la tabla 4, donde se estudian dos pretratamientos y cuatro barnices, los primeros como parcelas y los segundos como subparcelas. En este caso los pretratamientos se aplican en un esquema completamente al azar, mientras que los barnices se aplican al azar dentro de cada pretratamiento. Este acomodo de los factores de parcela y subparcela da lugar a un análisis particular. Si en general los factores se denotan por  $A$  (parcela) y  $B$  (subparcela) con  $a$  y  $b$  niveles, respectivamente, y si consideramos  $n$  repeticiones de cada nivel del factor de parcela, las cuales están acomodadas en un diseño completamente al azar, el modelo estadístico que describe esta situación es:

(7.1)

$$Y_{ijk} = \mu + \alpha_i + \delta_{ik} + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}; i = 1, 2, \dots, a; j = 1, 2, \dots, b; k = 1, 2, \dots, n.$$

donde  $\alpha_i$  es el efecto del  $i$ -ésimo nivel del factor de parcela y  $\delta_{ik}$  es el error de parcela. Sin considerar la media general  $\mu$ , estos primeros dos efectos se refieren a la parcela. Los últimos tres corresponden a la subparcela, donde  $\beta_j$  es el efecto del  $j$ -ésimo nivel del factor B,  $(\alpha\beta)_{ij}$  es el efecto de interacción  $AB$  en la combinación  $ij$  y  $\varepsilon_{ijk}$  es el error de la subparcela. Se espera que el error de subparcela siempre sea menor que el error de parcela. El objetivo del análisis es probar las hipótesis usuales:

$$\begin{cases} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_a = 0 \\ H_1 : \alpha_i \neq 0, \text{ para algún } i \end{cases},$$

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \dots = \beta_b = 0 \\ H_1 : \beta_j \neq 0 \text{ para algún } j \end{cases}$$

$$\begin{cases} H_0 : (\alpha\beta)_{ij} = 0 \text{ para todo } (i, j) \\ H_1 : (\alpha\beta)_{ij} \neq 0 \text{ para algún } (i, j) \end{cases}$$

Para ello es necesario obtener la tabla de ANOVA, que en los diseños en parcelas divididas es como calcular dos ANOVA; el primero para el efecto del factor de parcela y el segundo para el efecto del factor de subparcela y su interacción con el factor de parcela. De aquí que en el modelo aparezcan dos términos error, uno para cada tamaño de unidad experimental. Son precisamente los diferentes tamaños de las unidades experimentales utilizadas con cada factor lo que da lugar a las dos estructuras de error.

Tabla 6. Valores esperados de cuadrados medios del diseño en parcelas			
	variabilidad	GL	$E(CM)$
Parcela	$\alpha_i : (A)$	$a - 1$	$\sigma_\varepsilon^2 + b\sigma_\delta^2 + bn\phi_A$
	$\delta_{ik} : (\text{Error parcela})$	$a(n - 1)$	$\sigma_\varepsilon^2 + b\sigma_\delta^2$
Subparcela	$\beta_j : B$	$b - 1$	$\sigma_\varepsilon^2 + an\phi_B$
	$(\alpha\beta)_{ij} : (AB)$	$(a - 1)(b - 1)$	$\sigma_\varepsilon^2 + n\phi_{AB}$
	$\varepsilon_{ijk} (\text{error subparcela})$	$a(n - 1)(b - 1)$	$\sigma_\varepsilon^2$
	Total	$abn - 1$	

Para calcular las sumas de cuadrados de los efectos en el modelo se utilizan las mismas expresiones que se presentaron en los diseños con bloques y en diseños factoriales, pero no las vamos a reescribir aquí. En este momento es más importante revisar la tabla de los valores esperados de cuadrados medios (tabla 6), donde queda en evidencia la necesidad de las dos estructuras de error mencionadas. En esta tabla también se incluyen los grados de libertad, y el factor de parcela se acomoda como en un diseño completamente al azar, con  $n$  repeticiones en cada parcela. Los factores  $A$  y  $B$  se consideran fijos y las repeticiones aleatorias. Las reglas para obtener los valores esperados de cuadrados medios se pueden consultar en Hicks (1993).

De la tabla 6 se concluye que la hipótesis sobre el factor  $A$  se prueba con el estadístico  $F_0^A = \frac{CM_A}{CM_\delta}$ , donde llamarnos  $CM_\delta$  al cuadrado medio del error de parcela completa para distinguirlo del cuadrado medio del error en la subparcela ( $CM_{ES}$ ). La hipótesis acerca de  $B$  se prueba con  $F_0^B = \frac{CM_B}{CM_{ES}}$ ; mientras que la hipótesis sobre la interacción  $AB$  se verifica con  $F_0^{AB} = \frac{CM_{AB}}{CM_{ES}}$ .

7.1.1. *Análisis del ejemplo 6.2.* Donde se quiere estudiar el efecto de dos pre-tratamientos (parcela) y cuatro marcas de barniz (subparcela) sobre la resistencia de la madera al agua. Vamos a proceder por pasos para apoyarnos en los ANOVA



que se obtienen con cualquier software estadístico. Primero, para obtener el error de parcela sólo se toman en cuenta los promedios de cada réplica de los pretratamientos de la tabla 4, y se hace un ANOVA de la manera usual para comparar los pretratamientos. A las sumas de cuadrados resultantes se les multiplica por cuatro (tamaño de parcela, número de datos utilizados para calcular el promedio), para volver a la escala original de los datos. Los resultados de este primer paso se muestran en la tabla 7.

Tabla 7. ANOVA con promedio para el factor de parcelas					
variabilidad	SC	GL	$CM$	$F_0$	Valor- $p$
Pretratamiento	780.90	1	780.90	4.02	0.1154
Error (parcela)	776.72	4	194.18		
Total	1557.62	5			

Enseguida, para comparar las subparcelas (barnices), primero se obtiene el ANOVA para los datos de la tabla 4 como si el diseño fuera factorial  $2 \times 4$  con tres réplicas, véase tabla 8. Este ANOVA no es apropiado para el diseño de parcelas divididas, sin embargo, ayuda en la obtención del ANOVA apropiado.

Tabla 8. ANOVA del ejemplo 6.2 como factorial $2 \times 4$					
variabilidad	SC	GL	$CM$	$F_0$	Valor- $p$
Pretratamiento	782.04	1	782.04	13.49	0.002
Barniz	266.00	3	88.67	1.53	0.245
$Petra \times Barniz$	62.79	3	20.93	0.36	0.782
Error (parcela)	927.88	16	57.99		
Total	2038.72	23			

El ANOVA final del diseño en parcelas divididas se muestra en la tabla 9. Éste se obtiene a partir de la información de las dos tablas anteriores. De manera específica, al tomar como base la tabla 8, se le agrega el error de parcela obtenido en la tabla 7. El error de subparcela se obtiene restando al error de la tabla 7 el error de la tabla 8:

$$SC_{subparcela} = SC_E - SC_{parcela} = 927.88 - 776.72 = 151.06$$

que tiene  $16 - 4 = 12$  grados de libertad. Con estos datos se obtiene la tabla 9, donde los  $F_0$  se consiguen dividiendo los cuadrados medios entre el error correspondiente de parcela o subparcela, según corresponda, como se desprende de la tabla 6. Los valores- $p$  también se pueden obtener con el apoyo de un software. De la tabla 9 se concluye que los barnices son diferentes estadísticamente y que los pretratamientos y su interacción con los barnices no tienen un efecto considerable. De manera gráfica también es posible determinar cuál de los barnices es mejor. Vale la pena mencionar que si el experimento se hubiera analizado de manera inapropiada como si fuera un diseño factorial, de la tabla 8, se hubiese concluido erróneamente que los pretratamientos son importantes y que los barnices no lo son. Esto nos permite reconocer la trascendencia del diseño en parcelas divididas y su análisis apropiado.

Tabla 9. ANOVA completo del diseño en parcelas divididas del ejemplo 6.2						
	variabilidad	SC	GL	$CM$	$F_0$	Valor- $p$
Parcela	Pretratamiento	782.04	1	782.04	4.03	0.1353
	Error ( <i>parcela</i> )	776.72	4	194.18	15.28	
Subparcela	Barniz	266.00	3	88.67	7.04	0.0054
	Petra $\times$ Barniz	62.79	3	20.93	1.66	0.228
	Error (subparcela)	151.06	12	12.59		
	Total	2038.61	23			

**7.2. Parcelas acomodadas en bloques.** Si en lugar de considerar a las parcelas acomodadas en un diseño completamente al azar como en el ejemplo anterior, se consideran en un diseño en bloques completos al azar, los estadísticos de prueba cambian drásticamente. En la figura 3 se presenta un esquema de esta situación experimental y en los ejercicios se muestra un experimento donde además de los factores parcela (variedad) y subparcela (espaciamento de surco) está presente un factor de bloque en cuyos niveles se hacen las repeticiones.

Supongamos que en general hay dos factores de interés,  $A$  (parcela) y  $B$  (subparcela) con  $a$  y  $b$  niveles, pero ahora las  $r$  repeticiones de las parcelas son los niveles de un factor de bloques que denotamos por  $R$ . Con estas consideraciones el modelo estadístico queda como:

$$(7.2) \quad Y_{ijk} = \mu + \alpha_i + r_k + (\alpha r)_{ik} + \beta_j + (\beta r)_{jk} + (\alpha\beta)_{ij} + (\alpha\beta r)_{ijk} + \varepsilon_{ijk}$$

$$i = 1, 2, \dots, a; j = 1, 2, \dots, b; k = 1, 2, \dots, r.$$

donde  $\alpha_i$  es el efecto del  $i$ -ésimo nivel del factor A (parcela),  $r_k$  es el efecto del  $k$ -ésimo nivel del factor  $R$  (bloque),  $(\alpha r)_{ik}$  es el efecto de interacción  $AR$  en la combinación  $ik$ ,  $\beta_j$  es el efecto del factor B en su nivel  $j$ ,  $(\beta r)_{jk}$  representa la interacción  $BR$  en  $jk$ ,  $(\alpha\beta)_{ij}$  es el efecto de interacción  $AB$  en la combinación  $ij$ ,  $(\alpha\beta r)_{ijk}$  es la interacción  $ABR$  en la combinación  $ijk$  y  $\varepsilon_{ijk}$  es el error aleatorio en la casilla  $ijk$ .

Observe cómo un cambio aparentemente leve en la manera de acomodar las parcelas impacta bastante los modelos (ecuaciones 7.1 y 7.2). Ahora, en lugar de tener el error de parcela se tiene el efecto del factor bloque y la interacción  $AR$ . Esta última será usada como error de parcela. Obviamente esto impacta los valores esperados de los cuadrados medios para estimar el efecto de cada componente del modelo, como se aprecia en la tabla 10.

Tabla 10. Valores esperados de cuadrados medios del diseño en parcelas acomodadas en bloques			
	variabilidad	GL	$E(CM)$
Parcela	$R : r_k$	$r - 1$	$\sigma_\varepsilon^2 + ab\phi_R$
	$A : \alpha_i$	$a - 1$	$\sigma_\varepsilon^2 + b\sigma_{r\alpha}^2 + br\phi_A$
	$AB : (\alpha r)_{ik} : (\text{Error parcela})$	$(a - 1)(r - 1)$	$\sigma_\varepsilon^2 + b\sigma_{r\alpha}^2$
Subparcela	$B : \beta_j$	$b - 1$	$\sigma_\varepsilon^2 + a\sigma_{r\beta}^2 + ar\phi_B$
	$BR : (\beta r)_{jk}$	$(b - 1)(r - 1)$	$\sigma_\varepsilon^2 + a\sigma_{r\beta}^2$
	$AB : (\alpha\beta)_{ij}$	$(a - 1)(b - 1)$	$\sigma_\varepsilon^2 + a\sigma_{r\alpha\beta}^2 + r\phi_{AB}$
	$ABR : (\alpha\beta r)_{ijk}$	$(a - 1)(b - 1)(r - 1)$	$\sigma_\varepsilon^2 + \sigma_{r\alpha\beta}^2$
	$\varepsilon_{ijk}$ (error subparcela)	0	$\sigma_\varepsilon^2$ (no estimable)
	Total	$abr - 1$	

En la tabla 10 queda claro que los estadísticos de prueba apropiados ahora son:

$$F_0^A = \frac{CM_A}{CM_{AR}}, F_0^B = \frac{CM_B}{CM_{BR}}, F_0^{AB} = \frac{CM_{AB}}{CM_{ABR}}$$

donde todos los denominadores son cuadrados medios de interacciones que involucran las réplicas (bloques), mientras que el  $CM_E$  ni siquiera es estimable debido a que sólo hay un dato en cada combinación de (bloque,  $A$ ,  $B$ ). Por lo general, se combinan las sumas de cuadrados de las interacciones que involucran las repeticiones o bloques ( $R$ ) para conformar un error de subparcela con  $a(b-1)(r-1)$  grados de libertad ( $SC_E = SC_{BR} + SC_{ABR}$ ).

**7.3. El caso con más de dos factores.** Como se mencionó antes, existen situaciones experimentales que llevan a una estructura más compleja de parcelas divididas, dependiendo principalmente de cómo se arreglen las parcelas y las subparcelas y de cuántos factores de proceso o factores de bloque se incluyan en el experimento. Una situación especial se presenta con los experimentos robustos con arreglo interno y externo, cuando éstos en realidad son corridos como parcelas divididas, como en el ejemplo 5.2.2. Muchas veces, estos diseños no se corren completamente aleatorizados porque el arreglo interno (de factores de proceso) está ordenado en función de factores que son difíciles de mover, y en este caso no es conveniente hacer las corridas en orden aleatorio. De tal forma que, una vez que se fija una combinación de niveles de los factores de proceso, se obtienen en orden aleatorio los datos para todas las combinaciones de niveles de los factores de ruido. De manera que cada combinación de niveles de los factores de proceso es una parcela y, dentro de ésta, cada combinación de niveles de los factores de ruido es una subparcela.

Análisis del ejemplo 5.2.2. Veamos el modelo y el análisis del ejemplo 5.2.2, donde se quiere diseñar una harina robusta, con tres factores de proceso y dos factores de ruido. Por la manera en que se corrió este experimento se trata de un diseño en parcelas divididas, en el cual los factores de proceso definen las parcelas y los factores de ruido las subparcelas. El modelo estadístico para este tipo de situaciones está dado por:

$$\begin{aligned} (Y_{ijklm})_{klm} = & \mu + \alpha_i + \beta_j + \phi_k + (\alpha\beta)_{ij} + (\alpha\phi)_{ik} + (\beta\phi)_{jk} + (\alpha\beta\phi)_{ijk} + \delta_{ijk} \\ & + \varphi_l + \gamma_m + (\varphi\gamma)_{lm} + (\alpha\varphi)_{il} + (\beta\varphi)_{jl} + (\phi\varphi)_{kl} + (\alpha\gamma)_{im} + (\beta\gamma)_{jm} \\ & + (\phi\gamma)_{km} + \varepsilon_{ijklm}, \end{aligned}$$

con  $i, j, k, l, m = 1, 2$ . Los términos en el primer renglón de la ecuación son los efectos en las parcelas y los del segundo renglón son los efectos en las subparcelas. Note que los efectos de la interacción de factores de proceso con factores de ruido, además de que son de particular interés para lograr la robustez del producto, forman parte de los efectos de subparcela. En el caso del ejemplo de la harina no hay réplicas en el experimento, por lo tanto, los términos de error de parcela  $\delta_{ijk}$  y de subparcela  $\varepsilon_{ijklm}$  no son estimables en forma directa; más bien, se estimarán con los efectos no activos y con los de alto grado. Por ejemplo, estaría conformado por las 15 interacciones de alto orden que faltan en el modelo, como ocurre en los diseños factoriales sin réplicas. Interesa probar la significancia de los 16 efectos potencialmente importantes incluidos en el modelo 7.3 pero como no hay réplica, no tiene sentido hacer directamente el ANOVA de parcela grande, como se hizo en el ejemplo anterior. En este caso, en primer lugar se recomienda estimar todos

los efectos analizando los datos como si fueran un factorial completo  $2^5$ . En el lado izquierdo de la **figura 5** se muestran los efectos estimados; a continuación, los efectos estimados se separan en efectos de parcela y de subparcela de acuerdo con el modelo, y cada grupo se grafica por separado en papel de probabilidad normal (gráfico de Daniel). En la **figura 5** se aprecia que los efectos de parcela que no se alinean y que parecen importantes son los efectos  $H$  (harina) y  $P$  (polvo de huevo), mientras que los efectos subparcela son  $T$  (temperatura),  $Ti$  (tiempo),  $GT$  (grasa $\times$ temperatura) y posiblemente  $HT$  (harina $\times$ temperatura).

Con base en lo anterior, los efectos no significativos en cada unidad experimental se combinan para construir los errores y así obtener ANOVA aproximados. En el caso de parcelas, como sólo  $H$  y  $P$  parecen significativos, de la tabla 5 se calculan los promedios obtenidos en cada una de las combinaciones de  $H$  y  $P$ , y se obtiene el ANOVA como si se tratara de un diseño  $2 \times 2$  con dos réplicas. Las sumas de cuadrados obtenidas se multiplican por el número de datos con los que se obtuvo cada promedio (en este caso 4), para conservar la escala. De esta manera se obtiene el ANOVA para los efectos en la parcela completa (factores de proceso) de la tabla 11. Asimismo, se aprecia que, efectivamente, los efectos harina ( $H$ ) y polvo de huevo ( $P$ ) resultan significativos.

Tabla 11. ANOVA para los efectos de parcela					
variabilidad	SC	GL	$CM$	$F_0$	Valor- $p$
H (Harina)	52.78	1	52.78	61.09	0.000
P (polvo de huevo)	11.64	1	11.64	13.47	0.014
Error (parcela)	4.32	5	0.86		
Total	68.74	7			

Sólo para facilitar los cálculos en la tabla 12, el ANOVA del experimento se muestra como si fuera un factorial completo  $2^5$ , donde el término error se estima con los 24 efectos más pequeños detectados por dicho análisis. Este análisis no detecta la interacción  $HT$  como significativa, pero se decide dejarlo en la tabla porque el análisis gráfico lo muestra como candidato a estar activo. En cambio, la interacción  $GP$  se detecta como significativa, pero en el análisis gráfico no aparece como tal. El resto de los efectos de la tabla también se detectaron activos en el análisis gráfico. El ANOVA de la tabla 12 no es apropiado para el diseño porque el término de error no es correcto. De hecho, es típico que el error del análisis factorial subestime el error de parcela y sobreestime el error de subparcela.

Tabla 12. ANOVA como diseño factorial (análisis incorrecto)					
variabilidad	SC	GL	$CM$	$F_0$	Valor- $p$
H (harina)	52.78	1	52.78	205.41	0.0000
P (polvo de huevo)	11.64	1	11.64	45.29	0.0000
T (temperatura)	9.79	1	9.79	38.09	0.0000
Ti (tiempo)	2.26	1	2.26	8.79	0.0067
GT	6.75	1	6.75	26.26	0.0000
GP	1.76	1	1.76	6.85	0.0151
HT	0.75	1	0.75	2.92	0.1004
Error (parcela)	6.16	24	0.257		
Total	91.90	31			

Para obtener el análisis de varianza para parcelas divididas se usa como base la tabla 12, pero se corrige agregando el error de parcela estimado en la tabla 11, y también se estima el error de subparcela. Para esto, al error de la tabla 12 se le resta y agregan términos de la siguiente manera:

$$SC_{subparcela} = SC_E - SC_{parcela} + SC_{GP} = 6.16 - 4.32 + 1.76 = 3.6$$

que tiene 20 grados de libertad. La interacción  $GP$  se mandó al error porque en el análisis gráfico se detectó que no era significativa. En la tabla 13 se muestra el ANOVA completo para el diseño de parcelas divididas. Cada efecto se contrasta con el error que le corresponde. Ahora, la interacción  $HT$ , al ser contrastada con el error de subparcela, aumenta su significancia hasta casi 0.05, por lo que valdría la pena interpretarla. El resto de los efectos mantienen sus características. Cabe aclarar que al no existir repeticiones genuinas en el experimento se tuvieron que hacer las gráficas de efectos para construir los errores en forma apropiada mediante la suma de cuadrados de efectos pequeños. En caso de que haya repeticiones se procede en forma similar pero todo se obtiene directamente de los ANOVA.

Tabla 13. ANOVA completo del diseño en parcelas divididas del ejemplo						
	variabilidad	SC	GL	CM	$F_0$	valor $p$
Parcela	H (harina)	52.78	1	52.78	61.37	0.000
	P (polvo de huevo)	11.64	1	11.64	13.53	0.014
	Error parcela	4.32	5	0.86	4.78	
Subparcela	T (temperatura)	9.79	1	9.79	54.39	0.000
	Ti (tiempo)	2.26	1	2.26	12.55	0.002
	GT	6.75	1	6.75	37.5	0.0000
	HT	0.75	1	0.75	4.17	0.054
	error subparcela	3.6	20	0.18		
	Total	91.90	31			

Como siempre, la interpretación de los efectos activos se hace de manera gráfica como en los diseños factoriales. En la **figura 14.6** sólo representamos las interacciones que resultaron significativas en las subparcelas, y que involucran un factor de proceso y un factor de ruido. Con estas interacciones es posible hacer un intento para lograr la robustez de la harina, eligiendo el nivel del factor de control que minimiza el efecto del factor de ruido. De la interacción  $CT$  se aprecia que en el nivel alto de  $G$  (grasa) hay poco efecto de la variación de la temperatura sobre el sabor de los pasteles. En cambio, la interacción  $HT$  es muy leve y no hay mucha diferencia en el efecto de  $T$  en los diferentes niveles de  $H$ . Por ello, se recomienda un nivel alto de  $H$  para atender el sabor. Para fijar un nuevo nivel del otro factor de proceso significativo ( $P$ ) se recomienda ver su gráfica de efectos principales en donde se observa que lo más conveniente es el nivel alto.

## 8. Ejercicios: Diseños anidados y diseños en parcelas divididas

EJERCICIO 148. *Explique en qué consiste un diseño anidado y proporcione un ejemplo.*

EJERCICIO 149. *¿Por qué en un diseño anidado con dos factores no se puede estudiar la interacción de éstos?*

EJERCICIO 150. *Bosqueje la tabla de ANOVA para un diseño con tres factores A, B y C, con B anidado en A y C anidado en B. Escriba el modelo estadístico asociado.*

EJERCICIO 151. *¿Qué es un diseño en parcelas divididas? Escriba un ejemplo.*

EJERCICIO 152. *Describe un par de situaciones prácticas que llevan a utilizar un diseño en parcelas divididas.*

EJERCICIO 153. *Explique las diferencias entre un diseño en parcelas divididas y un diseño anidado, ambos con dos factores.*

EJERCICIO 154. *Con respecto a los datos del ejemplo 4.1.4 dados en la figura 1:*  
a) *Analice los datos como si fuera un estudio R&R cruzado o factorial. Compare la tabla de ANOVA resultante con la tabla 2, la cual representa el análisis anidado correcto. Verifique los cálculos de la  $SC_{B(A)}$  usando la relación dada en la ecuación ??.*  
b) *Si el estudio se analiza como un diseño factorial, ¿empeora o mejora de la calidad del sistema de medición? Para cada análisis obtenga los componentes de varianza y calcule qué porcentaje de la variabilidad total se atribuye al sistema de medición.*

### 8.1. Ejercicios de diseños anidados.

EJERCICIO 155. *Con respecto a los datos del ejemplo 5.2.2 dados en la figura 2.*  
a) *Analice los datos considerando el factor operador fijo, después compare los resultados con el análisis dado en la tabla 3. Por último, explique las diferencias observadas en los valores-p.*  
b) *Haga otra vez el análisis pero ahora como si fuera un experimento factorial incluyendo hasta la interacción triple. A partir de este análisis obtenga las sumas de cuadrados anidadas de la tabla 3, usando las ecuaciones (5.1).*

EJERCICIO 156. *Suponga que le muestran la tabla de ANOVA que aparece a continuación, en la cual un diseño anidado se analizó como un diseño factorial (cruzado). Reconstruya la tabla de ANOVA correcta asumiendo que 8 está anidado en A (A es fijo, 8 es aleatorio).*

variabilidad	SC	GL	CM	$F_0$	Valor-p
A	15	2	7.5	2.50	0.10
B	24	3	8.0	2.67	0.07
AB	42	6	7.0	2.33	0.06
Error	72	24	3.0		
Total	153	35			

EJERCICIO 157. *Analice otra vez los datos del experimento anidado que se presentó en la figura 1, pero ahora considere ambos factores fijos. Obtenga la tabla de ANOVA del diseño anidado a partir de la tabla de ANOVA del diseño factorial. Compare los resultados con los reportados en la tabla 2.*

EJERCICIO 158. *Beckman et al. (1987) describen un experimento cuyo objetivo es determinar si el aerosol estándar utilizado para probar los filtros de respiración puede reemplazarse por un aerosol alternativo (aerosol 2). También interesaba investigar la variabilidad de los filtros de dos fabricantes. Los datos obtenidos, donde*

la respuesta es el porcentaje de penetración del aerosol, se muestran a continuación:

Fabricante	Filtro	Aerosol 1			Aerosol 2		
1	1	0.750	0.770	0.840	1.120	1.100	1.120
	2	0.082	0.076	0.077	0.150	0.120	0.120
2	1	0.600	0.680	0.870	0.910	0.830	0.950
	2	1.000	1.800	2.700	2.170	1.520	1.580

a) Escriba el modelo que plantea este experimento considerando el anidamiento que ocurre. b) Haga el análisis y saque conclusiones. c) Considere que interesa minimizar la respuesta, ¿recomendaría usted el aerosol 2 como alternativa para probar los filtros?

EJERCICIO 159. Se realiza un experimento para estudiar las concentraciones de calcio en plantas de nabo. Para ello, se eligen cuatro plantas al azar, de cada planta se seleccionan al azar tres hojas y de cada hoja se sacan dos muestras a las que se les mide la concentración de calcio, con los siguientes resultados:

	Planta 1			Planta 2		
Muestra	Hoja 1	Hoja 2	Hoja 3	Hoja 1	Hoja 2	Hoja 3
M1	3.28	3.52	2.88	2.46	1.87	2.19
M2	3.09	3.48	2.80	2.44	1.92	2.19
	Planta 3			Planta 4		
Muestra	Hoja 1	Hoja 2	Hoja 3	Hoja 1	Hoja 2	Hoja 3
M1	2.77	3.74	2.55	3.78	4.07	3.31
M2	2.66	3.44	2.55	3.87	4.12	3.31

a) Los tres factores (planta, hoja y muestra) están completamente anidados. Explique por qué. b) Analice los datos como un diseño completamente anidado con los tres factores aleatorios y saque conclusiones. ¿Hay diferencias entre las plantas? ¿Hay diferencias entre las hojas?

EJERCICIO 160. Un genetista colecta tres semillas de dos árboles en cada uno de los tres bosques seleccionados. Las semillas se siembran en un vivero y se mide su crecimiento. Los datos en pulgadas son los siguientes:

	Bosque		
Árbol	A	B	C
1	15.8	18.5	12.3
	15.6	18.0	13.0
	16.0	18.4	12.7
2	13.9	17.9	14.0
	14.2	18.1	13.1
	13.5	17.4	13.5

a) Plantee un modelo estadístico que considere el anidamiento entre los dos factores. b) Analice los datos y saque conclusiones. Explique y pruebe las hipótesis de interés. c) Analice los datos como diseño factorial y compare los resultados con el análisis del inciso anterior.

## 8.2. Ejercicios de parcelas divididas.

EJERCICIO 161. En Steel y Torrie (1980) se publicó el siguiente experimento, que estudia el efecto de cuatro diferentes espaciamentos entre surcos sobre el rendimiento

de dos variedades de soya. Los rendimientos en fanegas por acre se muestran en la siguiente tabla. Una fanega es igual 55.5 litros, y un acre es igual a 4047m<sup>2</sup>.

Variedad	Espacio en pulgadas	Bosque					
		1	2	3	4	5	6
1	18	33.6	37.1	34.1	34.6	35.4	36.1
	24	31.1	34.5	30.5	32.7	30.7	30.3
	30	33.0	29.5	29.2	30.7	30.7	27.9
	36	28.4	29.9	31.6	32.3	28.1	26.9
	42	31.4	28.3	28.9	28.6	18.5	33.4
2	18	28.0	25.5	28.3	29.4	27.3	28.3
	24	23.7	26.2	27.0	25.8	26.8	23.8
	30	23.5	26.8	24.9	23.3	21.4	22.0
	36	25.0	25.3	25.6	26.4	24.6	24.5
	42	25.7	23.2	23.4	25.6	24.5	22.9

a) Describa mediante un dibujo la manera de aleatorizar este experimento. Considere a la variedad como el factor de parcela, que está acomodado en un diseño en bloques, y al espaciamiento entre surcos como el factor de subparcela. b) Escriba el modelo estadístico y las hipótesis de interés. c) Obtenga la tabla de análisis de varianza e interprete los resultados.

EJERCICIO 162. Potcner y Kowalski (2004) describen un experimento con un factor difícil de mover ( $D$ ) y tres factores fáciles de mover ( $A$ ,  $B$  y  $C$ ), cada uno de éstos con dos niveles. Los niveles del factor  $D$  son las parcelas que se repiten dos veces, y dentro de cada parcela se tienen ocho subparcelas que son las combinaciones de niveles de los factores  $A$ ,  $B$  y  $C$ . El experimento y los datos obtenidos se muestran en la siguiente tabla:

$D$	$A$	$B$	$C$	$Y$	$D$	$A$	$B$	$C$	$Y$
1	-1	1	1	108.4	1	-1	1	1	100.8
1	1	-1	1	131.6	1	1	1	-1	114.4
1	-1	-1	-1	124.0	1	1	-1	1	132.8
1	1	-1	-1	134.9	1	1	-1	-1	131.4
1	-1	1	-1	103.7	1	-1	-1	-1	118.4
1	1	1	-1	112.9	1	-1	1	-1	104.4
1	1	1	1	113.4	1	1	1	1	111.7
1	-1	-1	1	122.3	1	-1	-1	1	121.1
-1	-1	-1	-1	119.3	-1	1	1	-1	116.7
-1	1	1	-1	120.9	-1	-1	1	-1	112.8
-1	1	1	1	123.0	-1	-1	1	1	112.2
-1	1	-1	1	127.9	-1	1	-1	1	127.7
-1	-1	1	1	117.3	-1	-1	-1	-1	118.4
-1	-1	-1	1	120.9	-1	1	1	1	120.9
-1	1	-1	-1	129.9	-1	1	-1	-1	127.0
-1	-1	1	-1	115.4	-1	-1	-1	1	119.4

a) Describa el procedimiento de aleatorización que se debe seguir con este diseño. b) Escriba el modelo estadístico y bosqueje la tabla de ANOVA. c) Obtenga la tabla de ANOVA para el efecto en la parcela analizando el promedio por parcela. Ajuste las sumas de cuadrados a la escala de los efectos originales. d) Haga el análisis



incorrecto como si fuera un diseño factorial e interprételo. e) Combinando de forma manual los análisis de los incisos c) y d), obtenga el análisis como un diseño en parcelas divididas. f) Interprete y compare los resultados con el análisis incorrecto del inciso d). ¿Qué efectos se pueden malinterpretar cuando el experimentador sólo realiza ese análisis? g) Interprete los efectos activos suponiendo que interesa maximizar la respuesta.

EJERCICIO 163. Considerando el experimento descrito en el ejemplo 6.3. Conste lo siguiente. a) Analice los datos pero ahora considerando a los factores de ruido como las parcelas y a los factores de proceso como las subparcelas. b) ¿Cómo se habría corrido el experimento en el caso planteado en el inciso a)? c) Compare sus resultados con el análisis de la tabla 13.

EJERCICIO 164. En un laboratorio se usan tres hornos para correr un experimento cuyo fin es estudiar el efecto de la temperatura y la orientación sobre la resistencia de aleaciones de acero. Cada horno tiene diferente temperatura; dentro de cada uno se colocan orientadas al azar (orientación 1) dos muestras de cada una de las tres aleaciones, y otras dos muestras de las mismas aleaciones se colocan alineadas (orientación 2). Es decir, dentro de cada horno se colocan en lugares elegidos al azar 12 muestras, seis orientadas al azar y seis alineadas. Los datos obtenidos se muestran en la siguiente tabla.

	Temperatura					
	675			700		
Orientación	Alea 1	Alea 2	Alea 3	Alea 1	Alea 2	Alea 3
1	12	15	20	27	35	48
	19	28	26	40	39	55
2	15	25	25	48	48	48
	23	31	33	55	62	60
	Temperatura					
	725					
Orientación	Alea 1	Alea 2	Alea 3			
1	47	55	62			
	55	63	58			
2	48	63	68			
	60	68	62			

a) ¿Cuál es el factor de parcela y cuáles son los factores de subparcela? ¿En dónde radica la restricción a la aleatorización en este experimento? b) Escriba un modelo estadístico para este experimento. c) Analice los datos y saque conclusiones. ¿Por qué no es posible probar el efecto individual de la temperatura? d) Determine la mejor combinación de temperatura, orientación y aleación pensando que interesa maximizar la resistencia. Utilice gráficas.

EJERCICIO 165. Considere un experimento de horneado donde se quiere estudiar el efecto de tres temperaturas y tres tiempos de horneado sobre el sabor del

*producto; el arreglo propuesto es el siguiente:*

	<i>Tiempo</i>		
<i>Temperatura</i>	<i>5 min</i>	<i>7 min</i>	<i>9 min</i>
<i>175 C</i>			
<i>190 C</i>			
<i>205 C</i>			

*a) Si el experimento se corriera como un diseño factorial, explique cómo se debe obtener los datos para llenar la tabla. b) Si se corriera como diseño en parcelas divididas, explique la manera de obtener datos. ¿Cuál de los dos factores tomaría usted como factor de parcela? ¿Qué se ganaría con respecto al diseño factorial? c) En los dos diseños mencionados bosqueje las tablas de ANOVA correspondiente.*

## Bibliography

- [MONTGOMERY] MONTGOMERY C. DOUGLAS. Diseño y Análisis de Experimentos. Segunda Edición. LIMUSA WILEY
- [GUTIERREZ] GUTIERREZ H y DE LA VARA R. Análisis y Diseño de Experimentos. Segunda Edición. Mc Graw Hill.
- [KUEHL] KUEHL ROBERT. Diseño de Experimentos. Segunda Edición. Thomson.
- [VICENTE] VICENTE M, GIRON P, NIETO C, PÉREZ T. Diseño de Experimentos. Pearson Prentice Hall.
- [DEVORE] DEVORE JAY. Probabilidad y Estadística para Ingeniería y Ciencias. Sexta Edición.
- [WALPOLE] WALPOLE MYERS. Probabilidad y Estadística. Cuarta Edición. Mc Graw Hill.