

# Diseños Factoriales

## 3.1. Introducción.

Existen situaciones en la que se está interesado en estudiar el efecto conjunto de dos o más factores sobre la respuesta. Si esta es la situación ninguno de los diseños hasta ahora tratados, es apropiado. En estos casos, un diseño en el que se investiguen todas las posibles combinaciones de los niveles de los factores en cada realización del experimento es el ideal. Es decir, un diseño que permita analizar tanto efectos principales como efectos de interacción, es el apropiado. A este tipo de diseño se le conoce con el nombre de Diseño Factorial.

Un diseño factorial es entonces, aquel en el que el conjunto de tratamientos está conformado por todas las posibles combinaciones de los distintos niveles de los factores involucrados. Supongamos que se desea estudiar el efecto conjunto de  $P$  factores sobre la variable respuesta, donde el  $r$ -ésimo factor tiene  $a_r$  niveles. Una realización de este experimento contiene  $a_1 * a_2 * \dots * a_l * \dots * a_p$  combinaciones de tratamientos.

Ahora bien, dependiendo de los recursos disponibles en cada situación en particular, el experimento puede o no ser replicado. Si el experimento no puede ser replicado,

solo se puede conducir un análisis de varianza para cada efecto principal. Si se desea analizar el efecto de interacción se requiere de más de una observación por combinación de tratamiento, es decir, es necesario replicar el experimento.

Los experimentos factoriales se usan en casi todos los campos de investigación. Son de gran valor en el trabajo exploratorio, cuando se sabe poco acerca de los niveles óptimos o combinación óptima de los factores. También, cuando se desconoce cuales de ellos son realmente importantes.

Una replica de un experimento factorial es una realización del mismo que contiene todas las  $a_1 * a_2, \dots, a_l, \dots, a_p$  combinaciones de tratamientos. En este diseño se asume que el orden de experimentación es aleatorizado.

## 3.2. Efecto de los Factores

El efecto de un factor se define como el cambio en la respuesta producido por un cambio en el nivel del factor. Con frecuencia se le llama **efecto principal** porque se refiere a los factores de interés primario en el experimento. Por ejemplo, considere el experimento sencillo en la parte (a) de la figura 3.2. Se trata de un experimento factorial de dos factores en el que los dos factores del diseño tienen dos niveles. A estos niveles se les ha denominado "bajo" y "alto" y se denotan como "-" y "+", respectivamente. El efecto principal de cada factor se puede calcular como:

- El efecto del factor A es la diferencia entre la respuesta promedio con el nivel bajo de A y la respuesta promedio con el nivel alto de A. Esto, es

$$A = \frac{40 + 52}{2} - \frac{20 + 30}{2} = 21$$

Es decir, cuando el factor  $A$  se incrementa del nivel bajo al nivel alto se produce un **incremento de la respuesta promedio** de 21 unidades.

- El efecto principal del factor  $B$  es la diferencia entre la respuesta promedio con el nivel bajo de  $B$  y la respuesta promedio con el nivel alto de  $B$ . Esto es,

$$B = \frac{30 + 52}{2} - \frac{20 + 40}{2} = 11$$

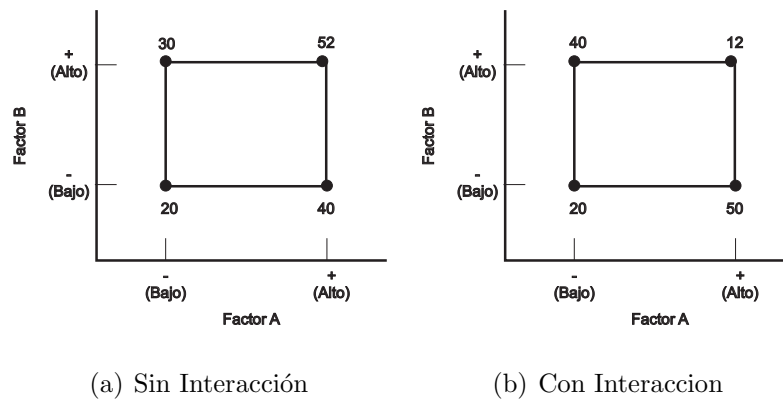


Figura 3.1: Experimento factorial de dos factores con la respuesta ( $y$ ) indicada por los vértices (a) y (b)

Cuando los factores tienen más de dos niveles, es necesario modificar el procedimiento anterior, ya que existen otras formas de definir el efecto de un factor.

En algunos experimentos puede encontrarse que la diferencia en la respuesta entre los niveles de un factor no es la misma para todos los niveles de los otros factores. Cuando esto ocurre, existe una **interacción** entre los factores. Por ejemplo, considere el experimento factorial de dos factores que se ilustra en la parte (b) de la figura 3.2. Con el nivel bajo del factor  $B$  (o  $B^-$ ), el efecto de  $A$  es

$$A = 50 - 20 = 30$$

y con el nivel alto del factor  $B$  (o  $B^+$ ), el efecto de  $A$  es

$$A = 12 - 40 = -28$$

Puesto que el efecto de  $A$  depende del nivel que se elige para el factor  $B$ , se observa que existe una interacción entre  $A$  y  $B$ . La magnitud del efecto de la interacción es la diferencia promedio de estos dos efectos de  $A$ , o  $AB = (-28 - 30)/2 = -29$ . Evidentemente, en este experimento la interacción es grande.

Estas ideas pueden ilustrarse gráficamente. En la parte (a) de la figura 3.2 se grafican los datos de las respuestas de la parte (a) de la figura 3.2 contra el factor  $A$  para ambos niveles del factor  $B$ . Observe que las rectas  $B^-$  y  $B^+$  son aproximadamente paralelas,

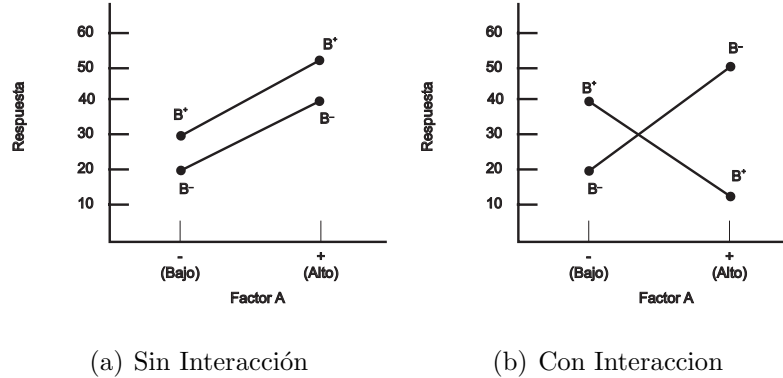


Figura 3.2: Experimento factorial (a) y (b)

lo cual indica la ausencia de interacción entre los factores  $A$  y  $B$ . De manera similar, en la parte (b) de la figura 3.2 se grafican los datos de las respuestas de la parte (b) de la figura 3.2. En este caso se observa que las rectas  $B^-$  y  $B^+$  no son paralelas. Esto indica una interacción entre los factores  $A$  y  $B$ . Gráficas como estas son de gran ayuda para interpretar las interacciones significativas y para reportar los resultados al personal sin interpretación estadística. Sin embargo, no deberán utilizarse como la única técnica

para el análisis de datos, ya que su interpretación es subjetiva y su apariencia con frecuencia es engañosa.

### 3.3. Ventajas y desventajas de un diseño factorial

Las ventajas de usar un diseño factorial son:

- Son más eficientes que los modelos unifactoriales. Producen ahorro y economía del recurso experimental así como mejoras en la inferencia.
- Todos los datos son usados en el cálculo de los efectos.
- El efecto de interacción puede ser analizado.

Las desventajas de usar un diseño factorial son:

- Al incrementarse el número de factores en el estudio, se incrementa rápidamente el número de combinaciones de tratamientos y así, el número de unidades experimentales necesarias para llevar a cabo una realización completa de un diseño factorial, llegando en algunas situaciones a ser imposible su realización.
- Cuando el número de niveles de algunos factores o el número de factores es grande, la variabilidad del experimento puede ser grande dado el número de posibles combinaciones de tratamientos. Esto puede hacer difícil la detección de efectos significativos en el experimento.

**Nota:** Considerar solo un subconjunto de todas las posibles combinaciones de tratamientos ó, considerar un número razonable de factores y restringir bajo algún criterio el número de niveles de cada uno de ellos, son dos formas de superar estas desventajas.

### 3.4. El Modelo

Suponga que se tiene un experimento con  $P$  factores, donde el  $P_r$ -esimo factor tiene  $a_r$  niveles,  $r = 1, \dots, p$ . El modelo matemático apropiado para describir esta situación esta dado por

$$Y_{ij\dots l} = \mu + \alpha_i + \beta_j + \dots + \eta_l + (\alpha\beta)_{ij} + \dots + (\alpha\beta\dots\eta)_{ij\dots l} + \varepsilon_{ij\dots l}; \quad (3.1)$$

donde  $\mu$  es el efecto general medio,  $\alpha_i$  el efecto del  $i$ -esimo nivel del factor  $A_1$ ,  $\beta_j$  el efecto del  $j$ -esimo nivel del factor  $A_2$ , ...,  $\eta_l$  el efecto del  $l$ -esimo nivel factor  $A_P$ ,  $(\alpha\beta)_{ij}$  el efecto de la interaccion entre  $\alpha_i$  y  $\beta_j$ ,  $(\alpha\beta\dots\eta)_{ij\dots l}$  el efecto de la interaccion entre  $\alpha_i$ ,  $\beta_j$ , ..., y  $\eta_l$  y  $\varepsilon_{ij\dots l}$  es un termino de error aleatorio tal que  $\varepsilon_{ij\dots l} \sim NID(0, \sigma^2)$ .

En general el número de efectos principales y de interacción presentes en el modelo pueden calcularse mediante la expresión  $\binom{P}{m}$ ,  $m = 1, 2, 3, 4, 5, \dots$ . Así, por ejemplo, un modelo asociado con experimento que involucre 5 factores contendrá  $\binom{5}{1} = 5$  efectos principales,  $\binom{5}{2} = 10$  efectos de interacciones dobles,  $\binom{5}{3} = 10$  efectos de interacciones triples,  $\binom{5}{4} = 5$  efectos de interacciones cuádruples y  $\binom{5}{5} = 1$  efectos de interacción de orden cinco.

El efecto principal de un factor, representa el cambio promedio en la respuesta cuando el es cambiando a los distintos niveles. El efecto de interacción, representa el cambio en la respuesta entre los niveles de un factor para los distintos niveles de algunos factores, promediado sobre los niveles de los factores restantes. Si se tienen por ejemplo tres factores, el efecto de interacción  $AB$ , mide el cambio en la respuesta entre los niveles de  $A$  para los distintos niveles de  $B$ , promediado sobre los niveles de  $C$ . Ahora bien, según se consideren los factores, efectos fijos o efectos aleatorios, se tienen tres variantes:

- Modelo de efectos fijos. Todos los factores se consideran fijos y los efectos de los tratamientos se definen como desviaciones de la media general.
- Modelo de efectos aleatorios. Todos los factores se consideran aleatorios. Esto es, los niveles de todos los factores se seleccionan en forma aleatoria.
- Modelo de efectos mixtos. Modelo en el cual unos factores son fijos y otros aleatorios.

En este curso solo se va a tratar el Modelo de efectos fijos.

### 3.5. Diseño Factorial de 2 Factores

Suponga que se tiene un experimento que involucra dos factores,  $A$  y  $B$ , con  $a$  y  $b$  niveles respectivamente. Se decide correr un diseño factorial de efectos fijos, el cual se repite  $n$  veces. Esto es, se tienen  $abn$  observaciones que son tomadas en forma aleatoria.

Sea  $Y_{ijk}$  la  $k$ -ésima observación tomada bajo el  $i$ -ésimo nivel del factor  $A$  y el  $j$ -ésimo nivel del factor  $B$ . El modelo estadístico matemático que describe esta observación está dado por

$$\begin{aligned}
 i &= 1, \dots, a \\
 Y_{ijk} &= \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}; \quad b = 1, \dots, b \\
 k &= 1, \dots, n
 \end{aligned} \tag{3.2}$$

donde  $\mu$  es el efecto general medio,  $\tau_i$  el efecto del  $i$ -ésimo nivel del factor  $A$ ,  $\beta_j$  el efecto del  $j$ -ésimo nivel del factor  $B$ ,  $(\tau\beta)_{ij}$  el efecto de la interacción entre  $\tau_i$  y  $\beta_j$ , y  $\varepsilon_{ijk}$  es

un termino de error aleatorio tal que  $\varepsilon_{ijk} \sim NID(0\sigma^2)$  y

$$\sum_{i=1}^a \tau_i = 0; \sum_{j=1}^b \beta_j = 0; \sum_{i=1}^a (\tau\beta)_{ij} = \sum_{j=1}^b (\tau\beta)_{ij} = 0 \quad (3.3)$$

Los datos de este modelo pueden arreglarse como en la tabla.

Tabla 3.1: Aspectos de los datos en un diseño factorial  $ab$

		Factor B			
		1	2	...	k
Factor A	1	$y_{111}, \dots, y_{11n}$	$y_{121}, \dots, y_{12n}$	...	$y_{1b1}, \dots, y_{1bn}$
	2	$y_{211}, \dots, y_{21n}$	$y_{221}, \dots, y_{22n}$	...	$y_{2b1}, \dots, y_{2bn}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	1	$y_{a11}, \dots, y_{a1n}$	$y_{a21}, \dots, y_{a2n}$	...	$y_{ab1}, \dots, y_{abn}$

Las posibles hipótesis a probar son las siguientes:

$$H_0 : \tau_i = 0; \forall i = 1, \dots, a \text{ Vs. } H_1 : \text{al menos un } \tau_i \neq 0 \quad (3.4)$$

$$H_0 : \beta_j = 0; \forall j = 1, \dots, b \text{ Vs. } H_1 : \text{al menos un } \beta_j \neq 0 \quad (3.5)$$

$$H_0 : (\tau\beta)_{ij} = 0; \forall i, j \text{ Vs. } H_1 : \text{al menos un } (\tau\beta)_{ij} \neq 0 \quad (3.6)$$

La primera hipótesis a probar es la referida al efecto de interacción. Si esta hipótesis resulta significativa, las hipótesis acerca de los efectos principales pierden importancia, dado que carecen de sentido si el efecto de un factor depende del otro factor. Para probar dichas hipótesis se realiza un análisis de varianza.



### 3.6. Análisis de Varianza

La técnica de análisis de varianza es una vez mas, el procedimiento a usar para analizar este tipo de problemas. Considere las siguientes cantidades:

$$Y_{ij.} = \sum_{k=1}^n Y_{ijk} = \text{Total correspondiente a la celda } ij\text{-esima en la tabla 1.}$$

$$Y_{i..} = \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} = \text{Total correspondiente a la fila } i \text{ de la tabla 1. Total asociado con el nivel } i\text{-esimo del factor A.}$$

$$Y_{.j.} = \sum_{i=1}^a \sum_{k=1}^n Y_{ijk} = \text{Total correspondiente a la columna } j \text{ de la tabla 1. Total asociado con el nivel } j\text{-esimo del factor B.}$$

$$Y_{...} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} = \text{Gran total.}$$

$$\bar{Y}_{ij.} = \frac{Y_{ij.}}{n}; \bar{Y}_{i..} = \frac{Y_{i..}}{bn}; \bar{Y}_{.j.} = \frac{Y_{.j.}}{an}; \bar{Y}_{...} = \frac{Y_{...}}{abn}$$

Ahora bien, el Análisis de Varianza en este caso, consiste en descomponer la variación total de los datos en cuatro porciones independientes atribuibles al factor  $A$ , factor  $B$ , interacción y error.

Una medida de la desviación de las observaciones con respecto a la media está dada por  $\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2$ , la cuál restándole y sumándole los promedios de los dos factores, la interacción y el promedio general y ordenándolo convenientemente se tiene que

$$\begin{aligned}
\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{...})^2 &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n [(\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.}) \\
&\quad + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})]^2
\end{aligned} \tag{3.8}$$

Al desarrollar el segundo miembro de la ecuación, y teniendo en cuenta que las sumatorias que contienen productos cruzados son iguales a cero se obtiene

$$\begin{aligned}
\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{...})^2 &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \bar{Y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{.j.} - \bar{Y}_{...})^2 \\
&\quad + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{ij.})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2
\end{aligned}$$

lo cual representa la descomposición de la suma de cuadrados total. Esta es la ecuación fundamental para el análisis de varianza asociado con el diseño factorial  $a*b$ . De aquí se tiene que

$$SC_T = SC_A + SC_B + SC_{AB} + SC_E \tag{3.9}$$

Para efecto de cálculos, las formulas anteriores usualmente se desarrollan y reescriben de la forma siguiente

$$\begin{aligned}
 SC_T &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - \frac{Y_{...}^2}{abn} \\
 SC_A &= \sum_{i=1}^a \frac{Y_{i..}^2}{bn} - \frac{Y_{...}^2}{abn} \\
 SC_B &= \sum_{j=1}^b \frac{Y_{.j.}^2}{an} - \frac{Y_{...}^2}{abn} \\
 SC_{subt} &= \sum_{i=1}^a \sum_{j=1}^b \frac{Y_{ij.}^2}{n} - \frac{Y_{...}^2}{abn} \\
 SC_E &= SC_T - SC_{subt}
 \end{aligned}$$

Los efectos principales tienen como grados de libertad el número de niveles que posee menos 1. Los grados de libertad de la interacción es la diferencia entre los grados de libertad de las celdas y los grados de libertad de los efectos principales.  $(ab - 1 - (a - 1) - (b - 1) = (a - 1)(b - 1))$ . Los grados de libertad del error se tienen de sumar los grados de libertad dentro de cada celda  $(ab(n - 1))$ . Por lo tanto, los grados de libertad están dados por

$$\begin{aligned}
 GL_T &= abn - 1 \\
 GL_A &= a - 1 \\
 GL_B &= b - 1 \\
 GL_{AB} &= (a - 1)(b - 1) \\
 GL_E &= GL_T - GL_A - GL_B - GL_{AB} = ab(n - 1)
 \end{aligned}$$

Cada suma de cuadrados dividida entre sus grados de libertad proveen un nuevo estadístico, llamado cuadrado medio.

Para determinar las expresiones de los estadísticos de prueba para cada una de las hipótesis que pueden ser planteadas en este caso, debemos simplemente observar el valor esperado de los cuadrados medios asociados con cada fuente de variación. Esto nos permitirá construir las relaciones apropiadas.

$$\begin{aligned}
 E(CM_E) &= \sigma^2 \\
 E(CM_A) &= \sigma^2 + \frac{bn}{a-1} \sum_{i=1}^a \tau_i^2 \\
 E(CM_B) &= \sigma^2 + \frac{an}{b-1} \sum_{j=1}^b \beta_j^2 \\
 E(CM_{AB}) &= \sigma^2 + \frac{N}{(A-1)(B-1)} \sum_{i=1}^a \sum_{j=1}^b (\tau\beta)_{ij}^2
 \end{aligned}$$

Bajo la hipótesis  $H_0 : (\tau\beta)_{ij} = 0; \forall ij$  cierta, el valor esperado de el  $CM_{AB}$  es igual a  $\sigma^2$ . Esto nos indica que la relación apropiada entre los cuadrados medios para probar esta hipótesis es  $\frac{CM_{AB}}{CM_E}$ . Para valores grandes de este cociente se debe tomar la decision de rechazar la hipótesis acerca de los efectos de interacción. Esto es, un valor grande de  $\frac{CM_{AB}}{CM_E}$  nos indicaría que existe efecto de interacción.

Bajo análisis similares a este, se puede concluir que los cocientes  $\frac{CM_A}{CM_E}$  y  $\frac{CM_B}{CM_E}$ , son los cocientes apropiados para probar las hipótesis  $H_0 : (\tau)_i = 0; \forall i$  y  $H_0 : (\beta)_j = 0; \forall j$ , respectivamente. Para valores grandes de estos cocientes se rechazaría la hipótesis correspondiente.

El gran dilema ahora es, como decidir cuando un cociente es lo suficientemente grande como para rechazar la hipótesis nula, es decir, a partir de que valor del cociente debemos

rechazar  $H_0$ . Para ello debemos identificar las propiedades distribucionales de estos cocientes y entonces, seleccionar el valor asociado con dicha distribución y que permitan tomar las decisiones correspondientes. A ese valor se le denomina valor crítico.

Bajo  $H_0$  cierta,

$$\frac{SC_A}{\sigma^2} \sim \chi_{(a-1)}^2 \quad \frac{SC_B}{\sigma^2} \sim \chi_{(b-1)}^2 \quad \frac{SC_{AB}}{\sigma^2} \sim \chi_{(a-1)(b-1)}^2 \quad \frac{SC_E}{\sigma^2} \sim \chi_{ab(n-1)}^2$$

De esta forma, podemos indicar que

$$F_A = \frac{CM_A}{CM_E} \sim F_{(a-1);ab(n-1)} \quad F_B = \frac{CM_B}{CM_E} \sim F_{(b-1);ab(n-1)} \quad F_{AB} = \frac{CM_{AB}}{CM_E} \sim F_{(a-1)(b-1);ab(n-1)}$$

Luego, para rechazar a un nivel de significación del  $100(1 - \alpha)\%$ , en cada caso debe cumplirse que:

- Rechazamos  $H_0 : (\tau\beta)_{ij} = 0 \forall i, j$ , si y sólo si  $F_{AB} = \frac{CM_{AB}}{CM_E} \sim F_{(a-1)(b-1);ab(n-1)}$
- Rechazamos  $H_0 : \tau_i = 0 \forall i = 1, \dots, a$ , si y sólo si  $F_A = \frac{CM_A}{CM_E} \sim F_{(a-1);ab(n-1)}$
- Rechazamos  $H_0 : \beta_j = 0 \forall j = 1, \dots, b$ , si y sólo si  $F_B = \frac{CM_B}{CM_E} \sim F_{(b-1);ab(n-1)}$

Todo lo anteriormente expuesto puede resumirse como en la tabla 3.2, la cual representa la tabla de análisis de varianza para un diseño factorial  $(a*b)$ . Obsérvese que si tenemos una sola replica no es posible estimar la varianza, lo que imposibilita la realización de las pruebas antes planteadas.

**Ejemplo 3.1** *Se desea estudiar la supervivencia de cuatro animales a los que se ha asignado al azar tres venenos y cuatro tratamientos. Ya que ambos factores, venenos y tratamientos son de igual interés, podría interesar si existe interacción entre ellos. A*

Tabla 3.2: Tabla de Análisis de Varianza para el Diseño Factorial

Fuente de Variación	Suma de Cuadrados	Grado de Libertad	Cuadrado Medio	F
A	$SC_A$	a-1	$CM_A$	$F_A$
B	$SC_B$	b-1	$CM_B$	$F_B$
AB	$SC_{AB}$	(a-1)(b-1)	$CM_{AB}$	$F_{AB}$
Error	$SC_E$	ab(n-1)	$CM_E$	
Total	$SC_T$	$abn - 1$		

continuación se muestra el Tiempo de supervivencia (unidad igual a 10 horas) de los animales con 4 repeticiones.

Veneno	Tratamientos			
	A	B	C	D
I	0.31	0.82	0.43	0.45
	0.45	1.10	0.45	0.71
	0.46	0.88	0.63	0.66
	0.23	0.72	0.76	0.62
II	0.36	0.92	0.44	0.56
	0.29	0.61	0.35	1.02
	0.40	0.49	0.31	0.71
	0.23	1.24	0.40	0.38
III	0.22	0.30	0.23	0.30
	0.21	0.37	0.25	0.36
	0.18	0.38	0.24	0.31
	0.23	0.29	0.22	0.33

### **Solución:**

El diseño utilizado en esta investigación es un diseño factorial de dos factores con tres y cuatro niveles respectivamente, donde:

- Factor A: Tratamiento con 4 niveles (A,B,C y D)
- Factor B: Veneno con tres niveles (I, II y III)

- *Variable Respuesta: Tiempo de supervivencia*

*Las hipótesis a probar son:*

$$H_0 : \tau_i = 0; \forall i = 1, 2, 3, 4 \text{ Vs. } H_1 : \text{al menos un } \tau_i \neq 0 \quad (3.10)$$

$$H_0 : \beta_j = 0; \forall j = 1, 2, 3 \text{ Vs. } H_1 : \text{al menos un } \beta_j \neq 0 \quad (3.11)$$

$$H_0 : (\tau\beta)_{ij} = 0; \forall i, j \text{ Vs. } H_1 : \text{al menos un } (\tau\beta)_{ij} \neq 0 \quad (3.12)$$

*Para hallar las sumas de cuadrados se procede de la siguiente manera:*

1. *De la tabla de los datos se obtiene que*

$$\blacksquare \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 = 0,45^2 + 1,10^2 + \dots + 0,33^2 = 13,9035$$

$$\blacksquare \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} = 0,45 + 1,10 + \dots + 0,33 = 22,81$$

2. *También a partir de la tabla se construye la tabla de los totales de los factores y de los subtotales, la cual está dada por donde  $1,45 = Y_{11.} = \sum_{k=1}^n Y_{ijk} = 0,31 +$*

Veneno	Tratamientos				$Y_{i..}$
	A	B	C	D	
I	1.45	3.52	2.27	2.44	9.68
II	1.28	3.26	1.5	2.67	8.71
III	0.84	1.34	0.94	1.30	4.42
$Y_{.j.}$	3.57	8.12	4.71	6.41	22.81

$$0,45 + 0,46 + 0,23.$$

3. Por lo tanto las sumas de cuadrados son

$$\begin{aligned}
 SC_{To} &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - \frac{Y_{...}^2}{abn} = 13,9035 - \frac{22,81^2}{48} = 3,064 \\
 SC_V &= \sum_{i=1}^a \frac{Y_{i..}^2}{bn} - \frac{Y_{...}^2}{abn} = \frac{9,68^2 + 8,71^2 + 4,42^2}{16} - \frac{22,81^2}{48} = 0,9794 \\
 SC_{Tr} &= \sum_{j=1}^b \frac{Y_{.j.}^2}{an} - \frac{Y_{...}^2}{abn} = \frac{3,57^2 + 8,12^2 + 4,71^2 + 6,41^2}{12} - \frac{22,81^2}{48} = 0,9898 \\
 SC_{subt} &= \sum_{i=1}^a \sum_{j=1}^b \frac{Y_{ij.}^2}{n} - \frac{Y_{...}^2}{abn} = \frac{1,45^2 + 3,52^2 + \dots + 4,42^2}{4} - \frac{22,81^2}{48} = 2,2403 \\
 SC_{V*Tr} &= SC_{subt} - SC_V - SC_{Tr} = 2,2403 - 0,9794 - 0,9898 = 0,2711 \\
 SC_E &= SC_T - SC_{subt} = 3,064 - 2,2403 = 0,8237
 \end{aligned}$$

De aquí se puede construir la siguiente tabla de Análisis de Varianza: Si se usa un

Tabla 3.3: Tabla de Análisis de Varianza para el Diseño Factorial

Fuente de Variación	Suma de Cuadrados	Grado de Libertad	Cuadrado Medio	F
Veneno (A)	0.9794	2	0.3299	14.42
Tratamiento(B)	0.9898	3	0.4897	21.40
A*B	0.2711	6	0.0452	1.97
Error	0.8237	36	0.0229	
Total	3.064	47		

nivel de significación del 5 %, se obtiene que los valores críticos son

- Para la interacción:  $F_{6,36,0,05} = 2,36$
- Para Veneno:  $F_{2,36,0,05} = 3,26$
- Para tratamiento:  $F_{3,36,0,05} = 2,87$



*Si se analiza primero la interacción, se observa que no se rechaza la hipótesis de que la interacción no tiene efecto, por lo tanto es necesario considerar los efectos principales. Al comparar los estadísticos calculador con los tabulados se concluye que ambos factores influyen en el tiempo de supervivencia de los animales pero de manera independiente.*

## 3.7. Diseño factorial $2^K$

### 3.7.1. Introducción

Como se dijo en el capítulo anterior, los diseños factoriales son ampliamente utilizados en experimentos en los que se tienen varios factores y se desea estudiar el efecto conjunto de estos sobre la respuesta.

Existen varios casos especiales del diseño factorial general de gran importancia, dado que son ampliamente utilizados en el trabajo de investigación y constituyen la base para otros diseños de gran valor práctico. Estos casos ocurren cuando se tienen  $k$  factores cada uno con dos o, tres niveles. En el primer caso, cada factor con dos niveles, una réplica completa de tal diseño requiere de  $2^k$  observaciones. Este diseño se conoce como diseño factorial  $2^k$ . En el segundo caso, el cual se conoce como diseño factorial  $3^k$ , cada factor tiene tres niveles y se requiere de  $3^k$  observaciones para llevar a cabo una réplica completa de tal diseño.

El diseño factorial  $2^k$  es de gran utilidad en las primeras fases del trabajo experimental, fundamentalmente cuando hay muchos factores por investigar.

### 3.7.2. Diseño factorial $2^k$

#### Diseño factorial $2^2$

Consideremos el caso más sencillo del diseño factorial  $2^k$ , un experimento que involucra 2 factores, cada uno con dos niveles. Los niveles de cada factor pueden denotarse como alto y bajo. El número de condiciones experimentales o tratamientos para el cual se necesita coleccionar información es igual a  $2^2$ . Las cuatro condiciones experimentales o tratamientos, se listan en la tabla 3.4. En la tabla 3.4, se muestra el total de las  $n$

Tabla 3.4: Condiciones experimental de un diseño  $2^2$

Tratamiento	A	B	AB
1	$a_0$	$b_0$	$a_0b_0$
2	$a_1$	$b_0$	$a_1b_0$
3	$a_0$	$b_1$	$a_0b_1$
4	$a_1$	$b_1$	$a_1b_1$

observaciones correspondientes a cada tratamiento.  $a_0b_0$  representa la condición experimental en la que los dos factores están en su nivel bajo,  $a_1b_0$  la condición experimental en la que el factor  $A$  esta en su nivel alto y  $B$  en su nivel bajo,  $a_0b_1$  la condición en la que el factor  $A$  se encuentra en su nivel bajo y el  $B$  en su nivel alto y  $a_1b_1$ , aquella en la que los dos factores están en su nivel alto.

Nosotros podemos estimar el efecto de pasar el factor  $A$  de un nivel a otro en cada nivel del factor  $B$ . De esta forma tenemos que,

- Efecto de  $A$  en el nivel bajo de  $B$  es igual a  $a_1b_0 - a_0b_0$
- Efecto de  $A$  en el nivel alto de  $B$  es igual a  $a_1b_1 - a_0b_1$

Como una aproximación del efecto de  $A$ , podemos promediar los dos resultados anteriores, obteniéndose

$$A = \frac{1}{2n}[(a_1b_1 - a_0b_1) + (a_1b_0 - a_0b_0)] = \frac{1}{2n}(a_1 - a_0)(b_1 + b_0) \quad (3.13)$$

De la misma forma el efecto de  $B$  estaría dado por

$$B = \frac{1}{2n}[(a_1b_1 - a_1b_0) + (a_0b_1 - a_0b_0)] = \frac{1}{2n}(a_1 + a_0)(b_1 - b_0) \quad (3.14)$$

El efecto de interacción de  $A$  y  $B$ , usando el mismo procedimiento es entonces,

$$AB = \frac{1}{2n}[\text{efecto de } A \text{ en el nivel alto de } B - \text{efecto de } A \text{ en el nivel bajo de } B] \quad (3.15)$$

o,

$$AB = \frac{1}{2n}[\text{efecto de } B \text{ en el nivel alto de } A - \text{efecto de } B \text{ en el nivel bajo de } A] \quad (3.16)$$

Es decir, el efecto de la interacción se define como la diferencia promedio entre el efecto del factor  $A$  ( $B$ ) al nivel alto del factor  $B$  ( $A$ ) y el efecto de  $A$  ( $B$ ) al nivel bajo del factor  $B$  ( $A$ ).

Ahora bien, por lo general se usa la siguiente notación para expresar las condiciones experimentales.

Los factores  $A$  y  $B$  en su nivel bajo = (1)

El factor  $A$  en su nivel alto y  $B$  en su nivel bajo =  $a$

El factor  $A$  en su nivel bajo y  $B$  en su nivel alto =  $b$

Los factores  $A$  y  $B$  en su nivel alto =  $ab$

Usando esta notación, las expresiones anteriores entonces, se pueden reescribir de la siguiente manera:

$$A = \frac{1}{2n}[(ab - b) + (a - (1))] \quad (3.17)$$

$$B = \frac{1}{2n}[(ab - a) + (b - (1))] \quad (3.18)$$

$$AB = \frac{1}{2n}[(ab - b) - (a - (1))] = \frac{1}{2n}[(ab - a) - (b - (1))] \quad (3.19)$$

La interacción puede darse si se da una variación en la magnitud de la respuesta o, por un cambio en la dirección de la respuesta.

Si la interacción es significativa indica que los factores no actúan en forma independiente. Los efectos simples de un factor difieren y la magnitud de un efecto simple depende del nivel del otro factor involucrado en la interacción. Análisis sobre efectos principales carecen por lo tanto, de interés.

Ahora bien, si la interacción no es significativa, los factores son independientes entre si. Los efectos simples de un factor son los mismos para los distintos niveles del otro factor. Analizar los efectos principales, es lo más apropiado.

Obsérvese que los numeradores de las ecuaciones (3.17), (3.18) y (3.19) se pueden expresar de la siguiente manera:

$$[(ab-b)+(a-(1))]=ab+a-b-(1)=(a-(1))(b+(1))$$

$$[(ab-a)+(b-(1))]=ab+b-a-(1)=(a+(1))(b-(1))$$

$$[(ab-b)-(a-(1))]=[(ab-a)-(b-(1))]=ab+(1)-b-a=(a-(1))(b-(1))$$

Recordemos que  $a, b, ab$  y  $(1)$  representan el total de las  $n$  observaciones para condición experimental. Este numerador por lo tanto, no es más que un contraste. Esto es, para

estimar  $A$ ,  $B$  y  $AB$  se usan contrastes, los cuales son ortogonales y se denotan por  $L_A$ ,  $L_B$  y  $L_{AB}$ .

En general, el contraste asociado con el efecto de interacción  $AB.....K$  esta dado por

$$L_{AB...K} = (a \pm (1))(b \pm (1))...(k \pm (1)) \quad (3.20)$$

Donde, el signo es negativo si el factor se encuentra en la interacción y positivo en caso contrario. El efecto de interacción está dado por

$$AB...K = \frac{1}{n2^{k-1}}L_{AB...K} \quad (3.21)$$

De acuerdo a lo anteriormente expuesto, las ecuaciones (3.17), (3.18) y (3.19) están dadas por:

$$A = \frac{1}{2n}L_A \quad (3.22)$$

$$B = \frac{1}{2n}L_B \quad (3.23)$$

$$AB = \frac{1}{2n}L_{AB} \quad (3.24)$$

Ordenando las condiciones experimentales en forma estándar y considerando su signo en cada efecto se obtiene la tabla 3.5, la cual puede ser fácilmente generalizada a el caso de tres o más factores. Obsérvese que el efecto de interacción es obtenido multiplicando las columnas correspondientes a los efectos de  $A$  y  $B$ .

En general el producto de cualquier par de columnas produce otra columna perteneciente a la tabla, simplemente tomando modulo 2 sobre el exponente de tales productos. Si se tiene por ejemplo, un diseño  $2^4$ , el cual contiene 16 combinaciones de tratamientos, el producto de las columnas correspondientes a los efectos  $ABC$  y  $BCD$ , dan como

Tabla 3.5: Tabla de signos para calcular los efectos en un Diseño  $2^2$ 

Condición Experimental	Efecto		
	A	B	AB
(1)	-	-	+
a	+	-	-
b	-	+	-
ab	+	+	+

resultado la columna correspondiente a el efecto  $AD$ . Esto es,

$$ABC * BCD = AB2C2D = AD(2 \text{ modulo } 2 = 0, 1 \text{ modulo } 2 = 1)$$

### 3.8. Método de Yates

Es una técnica de cálculo eficaz propuesta por Yates (1937), la cual se usa para estimar los efectos principales e interacciones y determinar las correspondientes sumas de cuadrados en un diseño factorial  $2^k$ . La técnica consiste en los siguientes pasos:

1. Diseñar una tabla de  $k + 5$  columnas ( $k$  es el numero de factores).
2. En la primera columna, llamada tratamiento, se presentan los tratamientos en el orden estándar o normalizado. Este orden es el siguiente:

- Para 2 factores: (1) a b ab
- Para 3 factores: (1) a b ab c ac bc abc
- Para 4 factores: (1) a b ab c ac bc abc d ad bd abd cd acd bcd abcd

El orden estándar se obtiene por la tabla de signos. Para el primer factor los signos se alternan de - a +, Para el siguiente factor se alternan de - a + pero en

parejas, es decir dos negativos dos positivos, luego dos negativos y dos positivos: para el tercer factor se alternan cada cuatro, y así sucesivamente.

3. En la columna 2, llamada respuesta, se introduce el resultado o la respuesta observada para cada tratamiento de la columna 1.
4. En la mitad superior de la columna 3, llamada (1), introducir por orden las sumas de los pares consecutivos de los valores de la columna 2. En la mitad inferior, introducir por orden las diferencias entre esos mismos pares consecutivos de valores, es decir, segundo valor menos primero, cuarto menos tercero, y así sucesivamente.
5. Repetir el paso anterior en las columnas siguientes a la (1) llamadas (2), (3), ..., (k).
6. Diseñar la columna llamada efecto, la cual contiene los símbolos que representan los efectos principales y las interacciones en el orden estándar. (la fila (1) se representa por I).
7. En la siguiente columna, llamada estimación del efecto, dividir los valores de la columna (k) por  $n * 2^{k-1}$
8. La última columna contiene las sumas de cuadrados, la cuál está dada por el valor de la columna (k) elevado al cuadrado dividido por  $n * 2^k$

Para la comprobación de los cálculos podemos observar que:

1. La suma de los valores de la segunda columna (columna de respuestas) es igual al primer elemento de la columna k.
2. Para cualquier efecto principal los valores de la columna k es igual a la suma de las respuestas en las que ese factor está a su nivel más alto menos las respuestas en las que ese factor está a su nivel más bajo.

**Ejemplo 3.8.1** Los datos en la siguiente tabla trata sobre tratamientos retardadores del fuego para tejidos. El experimento tiene 4 factores, cada uno con dos niveles, es decir, se trata de un experimento factorial  $2^4$ .

		$A_0$		$A_1$	
		$B_0$	$B_1$	$B_0$	$B_1$
$C_0$	$D_0$	4.2	4.5	3.1	2.9
	$D_1$	4.0	5.0	3.0	2.5
$C_1$	$D_0$	3.9	4.6	2.8	3.2
	$D_1$	4.0	5.0	2.5	2.3

Los factores y niveles experimentales son los siguientes:

<b>Factor</b>	<b>Nivel</b>
$A$ : Tejido	Satén Paño
$B$ : Tratamiento	$X$ $Y$
$C$ : Condiciones de lavado	Antes de un lavado Después de un lavado
$B$ : Sentido de la prueba	Urdimbre Trama

La tabla de Yates es la siguiente:



<i>tratamiento</i>	<i>Respuesta</i>	<i>(1)</i>	<i>(2)</i>	<i>(3)</i>	<i>(4)</i>	<i>Efecto</i>	<i>Efec. Est.</i>	<i>S.C.</i>
<i>(1)</i>	<i>4.2</i>	<i>7.3</i>	<i>14.7</i>	<i>29.2</i>	<i>57.5</i>	<i>I</i>	<i>7.1825</i>	<i>206.64</i>
<i>a</i>	<i>3.1</i>	<i>7.4</i>	<i>14.5</i>	<i>28.3</i>	<i>-12.9</i>	<i>A</i>	<i>-1.6125</i>	<i>10.40</i>
<i>b</i>	<i>4.5</i>	<i>6.7</i>	<i>14.5</i>	<i>-5.2</i>	<i>2.5</i>	<i>B</i>	<i>0.3125</i>	<i>0.39</i>
<i>ab</i>	<i>2.9</i>	<i>7.8</i>	<i>13.8</i>	<i>-7.7</i>	<i>-3.5</i>	<i>AB</i>	<i>-0.4375</i>	<i>0.77</i>
<i>c</i>	<i>3.9</i>	<i>7.0</i>	<i>-2.7</i>	<i>1.2</i>	<i>-0.9</i>	<i>C</i>	<i>-0.1125</i>	<i>0.05</i>
<i>ac</i>	<i>2.8</i>	<i>7.5</i>	<i>-2.5</i>	<i>1.3</i>	<i>-0.5</i>	<i>AC</i>	<i>-0.0625</i>	<i>0.02</i>
<i>bc</i>	<i>4.6</i>	<i>6.5</i>	<i>-3.5</i>	<i>-0.8</i>	<i>1.3</i>	<i>BC</i>	<i>0.1625</i>	<i>0.11</i>
<i>abc</i>	<i>3.2</i>	<i>7.3</i>	<i>-4.2</i>	<i>-2.7</i>	<i>0.5</i>	<i>ABC</i>	<i>0.0625</i>	<i>0.02</i>
<i>d</i>	<i>4.0</i>	<i>-1.1</i>	<i>0.1</i>	<i>-0.2</i>	<i>-0.9</i>	<i>D</i>	<i>-0.1125</i>	<i>0.05</i>
<i>ad</i>	<i>3.0</i>	<i>-1.6</i>	<i>1.1</i>	<i>-0.7</i>	<i>-2.5</i>	<i>AD</i>	<i>-0.3125</i>	<i>0.39</i>
<i>bd</i>	<i>5.0</i>	<i>-1.1</i>	<i>0.5</i>	<i>0.2</i>	<i>0.1</i>	<i>BD</i>	<i>0.0125</i>	<i>0.00</i>
<i>abd</i>	<i>2.5</i>	<i>-1.4</i>	<i>0.8</i>	<i>-0.7</i>	<i>-1.9</i>	<i>ABD</i>	<i>-0.2375</i>	<i>0.23</i>
<i>cd</i>	<i>4.0</i>	<i>-1.0</i>	<i>-0.5</i>	<i>1</i>	<i>-0.5</i>	<i>CD</i>	<i>-0.0625</i>	<i>0.02</i>
<i>acd</i>	<i>2.5</i>	<i>-2.5</i>	<i>-0.3</i>	<i>0.3</i>	<i>-0.9</i>	<i>ACD</i>	<i>-0.1125</i>	<i>0.05</i>
<i>bcd</i>	<i>5.0</i>	<i>-1.5</i>	<i>-1.5</i>	<i>0.2</i>	<i>-0.7</i>	<i>BCD</i>	<i>-0.0875</i>	<i>0.03</i>
<i>abcd</i>	<i>2.3</i>	<i>-2.7</i>	<i>-1.2</i>	<i>0.3</i>	<i>0.1</i>	<i>ABCD</i>	<i>0.0125</i>	<i>0.00</i>

### 3.9. Ejercicios

- Se lleva a cabo un experimento para determinar el crecimiento de plantas de frijol de acuerdo a tres factores: la profundidad de plantado de la semilla (.5 pulgadas y 1.5 pulgadas), la frecuencia de riego (una o dos veces diarias) y el tipo de frijol sembrado (pequeño o grande). Se hacen tres réplicas del experimento y se obtienen los siguientes datos:

Tipo de semilla	Frecuencia de riego	Profundidad	Rep 1	Rep 2	Rep 3
Pequeña	Una vez	0.5	6	7	7
		1.5	4	5	5
	Dos veces	0.5	10	9	8
		1.5	7	7	6
Grande	Una vez	0.5	4	5	4
		1.5	3	3	1
	Dos veces	0.5	8	7	7
		1.5	5	5	4

Calcule los efectos correspondientes a los factores y a las interacciones. Interprete los resultados obtenidos y sugiera condiciones ideales para sembrar.

- Un ingeniero está interesado en el efecto de la velocidad de corte (A), la dureza del metal (B) y el ángulo de corte (C) sobre la duración de una herramienta de corte. Para ello se eligen dos niveles para cada factor y se corren dos réplicas del diseño factorial  $2 \times 3$ . La tabla siguiente presenta los datos de tiempo de duración (en horas) de la herramienta.

- Calcule los efectos correspondientes a cada uno de los factores.
- Determine cuáles de estos efectos son importantes usando la tabla de análisis de varianza. Reporte las conclusiones obtenidas.

Combinación de tratamientos	Réplica	
	I	II
(1)	221	311
a	325	435
b	354	348
ab	552	472
c	440	453
ac	406	377
bc	605	500
abc	392	419

3. Un producto químico es elaborado en un tanque presión. Un experimento factorial fue realizado en la planta piloto para estudiar los factores que se piensa pueden influenciar la tasa de infiltración de este producto: la temperatura (A), la presión (B), la concentración del reactante (C) y la tasa de mezclado (D). Los resultados de un experimento 24 completo y sin réplicas se presentan a continuación:

	$A^-$				$A^+$			
	$B^-$		$B^+$		$B^-$		$B^+$	
	$C^-$	$C^+$	$C^-$	$C^+$	$C^-$	$C^+$	$C^-$	$C^+$
$D^-$	45	68	48	80	71	60	65	65
$D^+$	43	75	45	70	100	86	104	96

- a) Calcule los efectos y decida cuales son significativos.
- b) ¿Cuáles son las condiciones que maximizan la tasa de infiltración, y cuál es la respuesta promedio que se espera obtener? ¿Cuál sería su respuesta si la concentración del reactante debe mantenerse en nivel alto?
4. Se estudia el rendimiento de un proceso químico. Se cree que las dos variables más importantes son la presión y la temperatura. se seleccionan tres niveles de cada factor y se realiza un experimento factorial con dos réplicas. Se recopilan los siguientes datos:

Temperatura	Presión		
	200	215	230
Baja	90.4	90.7	90.2
	90.2	90.6	90.4
Media	90.1	90.5	89.9
	90.3	90.6	90.1
Alta	90.5	90.8	90.4
	90.7	90.9	90.1

- a) Analice los datos y obtenga las conclusiones.
- b) Prepare las gráficas apropiadas de los residuos y comente la adecuación del modelo
- c) ¿En qué condiciones debe operarse el proceso?.
5. Un ingeniero de manufactura sospecha que el acabado superficial de una pieza metálica depende de la alimentación y de la profundidad de corte. Selecciona tres niveles de alimentación y cuatro niveles de profundidad del corte. A continuación, realiza un experimento factorial y recopila los siguientes datos:

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

- a) Analice los datos y obtenga las conclusiones.
- b) Prepare las gráficas apropiadas de los residuos y comente la adecuación del modelo

6. Se realizó un experimento para investigar el efecto de tipo de vidrio y el tipo de fósforo sobre la brillantez de un cinescopio de televisor. La variable respuesta es la corriente necesaria (en micro amperios) para obtener un nivel de brillantez especificado. los datos son como sigue:

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

- a) ¿Existe indicación de que algunos de los factores influya en la brillantez?
- b) ¿Interactúan entre sí ambos factores?
- c) Analice los residuos de este experimento.
7. Se están estudiando los factores que influyen en la resistencia de ruptura de una fibra sintética. Para ello se tienen cuatro máquinas y tres operadores y se realiza un experimento usando fibras de un mismo lote de producción. Los resultados se muestran a continuación.

Operario	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) Analice los datos y obtenga las conclusiones apropiadas.

- b) Prepare las gráficas apropiadas de los residuos y comente la adecuación del modelo
8. Un ingeniero está interesado en los efectos de la velocidad de corte ( $A$ ), la geometría de la herramienta ( $B$ ) y el ángulo de corte ( $C$ ) sobre la vida (en horas) de una máquina. Se eligen dos niveles de cada factor y se corren tres réplicas. Los resultados fueron los siguientes:

A	B	C	Combinación de Tratamientos	Replica		
				I	II	III
-	-	-	(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) Estimar los efectos de los factores. ¿Qué efectos parecen ser grandes?
- b) Usar el análisis de varianza para confirmar las conclusiones del inciso a.
- c) Analizar los residuales. ¿Hay algún problema evidente?
- d) Con base en el análisis de las gráficas de los efectos principales y las interacciones, ¿cuáles serían los niveles de  $A$ ,  $B$  y  $C$  que se recomendaría utilizar?.
9. En un artículo de AT&T Technical Journal (vol 65, pp. 39-50) se describe la aplicación de diseños factoriales de dos niveles en la fabricación de circuitos integrados. Un paso básico del procesamiento es hacer crecer una capa epitaxial sobre obleas de silicio pulidas. Las obleas se montan en un susceptor, se colocan en el interior de una campana de cristal y se introducen vapores químicos. El susceptor se hace girar y se aplica calor hasta que la capa epitaxial tiene el espesor

suficiente. Se corrió un experimento utilizando dos factores: rapidez de flujo de arsénico ( $A$ ) y tiempo de deposición ( $B$ ). Se corrieron cuatro replicas y se midió el espesor de la capa epitaxial (en  $\mu\text{m}$ ). Los datos se muestran a continuación:

A	B	Replicas			
		I	II	III	IV
-	-	14,037	16,165	13,972	13,907
+	-	13,880	13,860	14,032	13,914
-	+	14,821	14,757	14,843	14,878
+	+	14,888	14,921	14,415	14,932

Los niveles de los factores son

Niveles de factores		
	Bajo(-)	Alto(+)
A	55 %	59 %
B	Corto (10 min)	Largo (15 min)

- Estimar los efectos de los factores.
- Conducir un análisis de varianza. ¿Qué factores son importantes?
- Analizar los residuales. ¿Se observa algún residual que debiera causar preocupación?