Capítulo

2

Diseño de Bloques Aleatorios y Cuadrados Latinos

2.1. Diseño de Bloques Aleatorios

Como se ha dicho antes, uno de los principales objetivos que se persigue al diseñar un experimento, es reducir el error aleatorio y de esta forma, incrementar la precisión de los resultados. En el diseño completamente aleatorio se supone que las unidades experimentales son relativamente homogéneas con respecto a factores que afectan la variable respuesta. Sin embargo algunas veces no tenemos disponibles suficientes unidades experimentales homogéneas. Por lo tanto, cualquier factor que afecte la variable respuesta y que varíe entre las unidades experimentales aumentará la varianza del error experimental, disminuyendo así la precisión de las comparaciones.

Por ejemplo, consideramos el problema de determinar si distintas máquinas exhiben diferente velocidad en el ensamblaje de un artículo. El gerente de una empresa desea comparar cuatro máquinas diferentes y tomar alguna decisión acerca de cual máquina adquirir de acuerdo a la velocidad de ensamblaje mostrada. El factor de interés es sólo la máquina, pero es importante tomar en cuenta que la operación de las máquinas requiere determinada destreza, pues operadores mas diestros pueden incidir en la disminución del tiempo de ensamblaje del artículo. Esto implica que la velocidad de ensamblaje no se debe sólo a la diferencia entre los cuatro tipos de maquinas sino también a la destreza de los operadores. En términos de variabilidad, la variación de los tiempos de ensamblaje no se debe sólo a la variación producida por las máquinas sino también a la variación producida por los operadores.

Para disminuir tal variabilidad, su utilizan mecanismos conocidos como control local. Uno de ellos es, disponer de unidades experimentales en varios grupos homogéneos, llamados generalmente bloques, los cuales admiten variación entre ellos.

En el ejemplo anterior se vio que hay dos factores que aportan sobre la variabilidad de la respuesta, el tipo de máquina y el operador, pero como solo es de interés el efecto que tiene la máquina, entonces es necesario controlar el efecto producido por los operadores, esto se logra colocando los operadores como bloques, es decir, cada operador debe usar las 4 máquinas, de esta manera la variabilidad producida por cada operario se deberá a la diferencia entre las máquinas.

Por lo tanto, los bloques se pueden definir como los valores de un factor que se piensa influye sobre la respuesta pero que no es de interés en el estudio.

Usar bloques estratifica a las unidades experimentales en grupos homogéneos. Una buena elección del criterio de bloqueo resulta en menor variación entre las unidades experimentales dentro de los bloques comparada con la variación de las unidades experimentales entre los bloques.

2.1.1. Tipos de diseños de bloques

Dependiendo del tamaño del bloque usado, existen dos tipos básicos de diseños de bloques aleatorizados:

1. **Diseño de bloque completo:** Cada bloque contiene todos los tratamientos. Esto es, el material experimental es dividido en b bloques de k unidades experimentales cada uno, donde k representa el número de tratamientos (Tabla 1)

2. **Diseño de bloque incompleto:** El tamaño de al menos un bloque es menor que el número de tratamientos en el experimento. Existen dos tipos de bloques aleatorizados incompletos:

   a) b.1. Balanceado: Todos los bloques tienen el mismo tamaño y el número de bloques en el que cualquier par de tratamientos aparece juntos, es constante. Si además, el número de tratamientos es igual al número de bloques, decimos que el diseño es simétrico. (Tabla 2)

   b) No Balanceado: El número de bloques que contiene cualquier par de tratamientos no es constante, puede diferir de un par a otro. (Tabla 3)

2.1.2. Diseño de bloques aleatorizados con bloques completos

Consideremos ahora en detalle el diseño de bloques aleatorizados con bloques completos. La aleatorización se da de la siguiente manera: Los tratamientos son primero numerados de 1 a k en cualquier orden. Las unidades en cada bloque son además numeradas, convenientemente de 1 a k. Los k tratamientos son asignados en forma aleatoria a las k unidades en cada bloque. La distribución aleatoria puede ser hecha o bien, consultando una tabla de números aleatorios, sorteos de lotes o el lanzamiento de una moneda como se describe en el diseño completamente aleatorizado.

En este tipo de diseño, como se explico antes, sobre la variable respuesta influyen tres factores: el factor de interés a través de sus tratamientos, el factor que no es de interés estudiar y el error experimental el cual contiene todas aquellas variables que no han sido tomadas en cuenta. Por lo tanto la variable respuesta puede modelarse de la siguiente manera:

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij} \qquad \begin{array}{l} i = 1, 2, \cdots, b \\ j = 1, 2, \cdots, k \end{array}$$

donde

Y_{ij} es la observación del j-ésimo tratamiento en el i-ésimo bloque.

μ es la media general

β_i es el efecto del i-ésimo bloque

τ_j es el efecto de j-ésimo tratamiento

ε_{ij} es el error aleatorio, los cuales se suponen $N(0, \sigma^2)$ e independientes

En este modelo, $\beta_i = \mu_{i.} - \mu$ y $\tau_j = \mu_{.j} - \mu$

Además suponiendo que el modelo es de efectos fijos se cumple que

$$\sum_{i=1}^{b} \beta_i = \sum_{j=1}^{k} \tau_j = 0$$

Un supuesto adicional a los ya considerados, es que el efecto de cada tratamiento es el mismo en todos los bloques. Esto significa que no existe interacción entre bloques y tratamientos.

Ahora bien, los datos observados de un diseño en el cual los tratamientos son arreglados en b bloques completos pueden presentase como en la tabla **??**

Tabla 2.1: Tabla 1....

		Tratamiento				Total
		1	2	$\cdots$	k	bloque
	1	y_{11}	y_{12}	$\cdots$	y_{1k}	$y_{1.}$
Bloque	2	y_{21}	y_{22}	$\cdots$	y_{2k}	$y_{2.}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	
	b	y_{b1}	y_{b2}	$\cdots$	y_{bk}	$y_{b.}$
Total Trat		$y_{,1}$	$y_{,2}$	$\cdots$	$y_{.k}$	$y_{..}$

Al igual que en el diseño completamente aleatorizado la hipotésis a probar es:

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

que también se puede expresar como

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

Para probar dicha hipótesis se realiza un análisis de varianza

2.1.3. Análisis de Varianza

La hipótesis dada en (**??**) ó (**??**) se prueba con un *análisis de varianza con dos criterios de clasificación*; se utilizan los dos criterios porque se controlan dos fuentes de variación: el factor de tratamientos y el factor de bloque.

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

$$\sum_{j=1}^{k}\sum_{i=1}^{b}(y_{ij}-\bar{y..})^2 = \sum_{j=1}^{k}\sum_{i=1}^{b}[(\bar{y}_{i.}-\bar{y}_{..})+(\bar{y}_{.j}-\bar{y}_{..})+(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})]^2 \tag{2.3}$$

Al desarrollar el segundo miembro de la ecuación **??**, se obtiene

$$\begin{aligned}
\sum_{j=1}^{k}\sum_{i=1}^{b}(y_{ij}-\bar{y..})^2 &= \sum_{j=1}^{k}\sum_{i=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})^2+\sum_{j=1}^{k}\sum_{i=1}^{b}(\bar{y}_{.j}-\bar{y}_{..})^2+\sum_{j=1}^{k}\sum_{i=1}^{b}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})^2 \\
&+ 2\sum_{j=1}^{k}\sum_{i=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{.j}-\bar{y}_{..})+2\sum_{j=1}^{k}\sum_{i=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..}) \\
&= 2\sum_{j=1}^{k}\sum_{i=1}^{b}(\bar{y}_{.j}-\bar{y}_{..})(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})
\end{aligned}$$

Se puede probar que las sumatorias que contienen productos cruzados son iguales a cero. Por lo tanto,

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

representa la descomposición de la suma de cuadrados total. La cual se puede expresar de la siguiente manera

$$SC_T = SC_B + SC_{Tr} + SC_E \tag{2.4}$$

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

$$\begin{aligned}
SC_T &= \sum_{j=1}^{k}\sum_{i=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{bk} \\
SC_B &= \sum_{i=1}^{b} \frac{y_{i.}^2}{k} - \frac{y_{..}^2}{bk} \\
SC_{Tr} &= \sum_{j=1}^{k} \frac{y_{.j}^2}{b} - \frac{y_{..}^2}{bk} \\
SC_E &= SC_T - SC_B - SC_{Tr}
\end{aligned}$$

Ahora bien, si el bloqueo es usado para reducir el error experimental, comparando los modelos para los diseños completamente aleatorizados y de bloques completamente aleatorizados para la i-esima observación bajo el j-esimo tratamiento se tiene que:

$$\varepsilon_{ij}^{*} = \varepsilon_{ij} + \beta_i \tag{2.5}$$

donde ε_{ij}^{*} es el error aleatorio del diseño completamente aleatorizado, ε_{ij} el error para el diseño de bloques completamente aleatorizado y β_i el efecto de bloque.

Esta igualdad implica que la suma de cuadrados del Error en el Diseño Completamente Aleatorizado es igual a la Suma de Cuadrados de Bloques mas la Suma de Cuadrados del Error en el Diseño de Bloques, es decir :

$$SC_E(DCA) = SC_B + SC_E(DBCA) \tag{2.6}$$

Las sumas de cuadrados dadas en (11), divididas por sus grados de libertad proveen otros estadísticos, llamados cuadrados medios.

$$CM_{Tr} = \frac{SC_{Tr}}{k-1} \qquad CM_B = \frac{SC_B}{b-1} \qquad CM_E = \frac{SC_E}{(k-1)(b-1)} \tag{2.7}$$

Si la varianza de los errores se supone constante, $sigma^2$, y además β_i y τ_j son fijos, entonces los valores esperados de estos cuadrados medios están dados por:

$$\begin{aligned} E(CM_E) &= \sigma^2 \\ E(CM_{Tr}) &= \sigma^2 + \frac{b\sum_{j=1}^{k}\tau_j^2}{k-1} \\ E(CM_B) &= \sigma^2 + \frac{k\sum_{i=1}^{n}\beta_i^2}{b-1} \end{aligned}$$

Existen dos posibles hipótesis a probar aquí, una sobre el efecto de los bloques y otra sobre el efecto de los tratamientos.

$$\begin{aligned} H_0^1 &: \tau_1 = \tau_2 = \cdots = \tau_k = 0 \text{ Igualdad de efectos de tratamientos} \\ H_0^2 &: \beta_i = \beta_2 = ... = \beta_b = 0 \text{ Igualdad de efectos de bloques} \end{aligned} \tag{2.8}$$

Si H_0 es verdadera, el valor esperado del cuadrado medio de los tratamientos esta dado por:

$$E(CM_{Tr}) = \sigma^2$$

Por consiguiente, bajo la hipótesis nula cierta, el cociente

$$F_c^1 = \frac{CM_{Tr}}{CM_E} \tag{2.9}$$

sigue una distribución F con $k-1$ y $(b-1)(k-1)$ grados de libertad.

De aquí podemos indicar que cuando la hipótesis nula H_0^1 no es verdadera, se espera que ocurra un valor grande para F_c^1, es decir, H_0^1 debe ser rechazada. F_c^1 es chequeado contra el valor crítico $F_{\alpha,k-1,(k-1)(b-1)}$; si F_c^1 es mayor que este valor crítico, rechazamos H_0^1.

Por razones similares, si H_0^2 es verdadera, el valor esperado del cuadrado medio de los bloques está dado por:

$$E(CM_B) = \sigma^2$$

Bajo H_0^2 cierta, el cociente

$$F_c^2 = \frac{CM_B}{CM_E} \tag{2.10}$$

y H_0^2 se rechaza si F_c^2 es mayor que el valor crítico $F_{\alpha,k-1,(k-1)(b-1)}$.

Todo este desarrollo lo podemos resumir como se muestra en la Tabla **??**. Esta tabla recibe el nombre de Tabla de Análisis de Varianza para el Diseño de Bloques Completamente Aleatorizado.

Ejemplo 2.1 *Se realiza un experimento para determinar el efecto que tiene el grado de trabajo (vueltas por pulgada) en la resistencia del algodón. Se decide utilizar cinco niveles para el grado de trabajo; 150, 163, 169, 178 y 10 vueltas por pulgadas. Se sabe que además de este factor, existen otras posibles fuentes de variación, como las máquinas, operadores, material experimental, entre otros. Después de una discusión se*

Tabla 2.2: Tabla de Análisis de Varianza para el DBCA

Fuente de Variación	Suma de Cuadrados	Grado de Libertad	Cuadrado Medio	F
Tratamiento	SC_{Tr}	k-1	CM_{Tr}	F_c^1
Bloque	SC_B	b-1	CM_B	F_c^2
Error	SC_E	(k-1)(b-1)	CM_E	
Total	SC_T	N-1		

decide ignorar el efecto de estos factores, excepto el factor máquinas, el cual será controlado. La variable respuesta medida es el número de roturas por cada cien libras de material. La tabla 3-1 muestra los resultados obtenidos:

Tabla 2.3: Número de Rupturas por cada cien libras

		Grados de Trabajo				
		10	163	169	178	190
Máquina	1	9	24	42	29	68
	2	12	27	23	49	34
	3	12	22	22	17	60
	4	31	16	47	45	50
	5	22	25	17	39	57
	6	10	24	23	44	37

Probar la hipótesis de que no existen efectos de tratamientos a un nivel de significación del 5 %.

Solución:

El diseño utilizado en esta investigación es un diseño de bloques completos, donde:

- *Tratamientos: Grados de Trabajo*

- *Bloques: Maquinas*
- *Variable Respuesta: Numero de roturas por cada cien libras de material*

La primera hipótesis a probar es:

$$H_0^1 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = 0$$

Bajo el supuesto de normalidad se puede hacer uso de la técnica de análisis de varianza para probar dicha hipótesis.

Tabla 2.4: Tabla de totales

		Tratamiento					Total bloque
		10	163	169	178	190	
Bloque	1	9	24	42	29	68	172
	2	12	27	23	49	34	145
	3	12	22	22	17	60	133
	4	31	16	47	45	50	189
	5	22	25	17	39	57	160
	6	10	24	23	44	37	138
Total Tratamiento		96	138	174	224	305	937

Además, $\displaystyle\sum_{j=1}^{6}\sum_{i=1}^{5} Y_{ij}^2 = 36475$. *Entonces*

$$\begin{aligned}
SC_T &= 36475 - \frac{937^2}{30} = 7209,367 \\
SC_B &= 29732,600 - \frac{937^2}{30} = 466,967 \\
SC_{Tr} &= 33{,}650,167 - \frac{937^2}{30} = 4384,533 \\
SC_E &= 7209,367 - 466,967 - 4384,533 = 2357,867
\end{aligned}$$

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

Tabla 2.5: Tabla de Análisis de Varianza para el DBCA

Fuente de Variación	Suma de Cuadrados	Grado de Libertad	Cuadrado Medio	F
Tratamiento	4384,533	4	1096,133	9,298
Bloque	466,967	5	93,393	0,792
Error	2357,867	20	93,393	
Total	7209,367	29		

de significación del 5 %, el valor crítico es $F_{4,20,0,05} = 2,87$ *y dado que 9,298 es mayor que 2,87, entonces se rechaza la hipótesis, es decir, se concluye que existe efectos del grado de trabajo sobre el número de roturas.*

Al observar el valor de la F_c *asociada con los bloques, 0.792, podemos concluir que no existen diferencias significativas entre las máquinas, lo que implica que el diseño de bloques no se justifica.*

Capítulo

$\boxed{3}$

3.1. Diseño en Cuadrado Latino

En la sección anterior se vio que una manera de reducir el error experimental era tomando en cuenta otros factores que se piensan influyen sobre la variable respuesta, En ese caso, se supuso que sólo existía un sólo factor al cual se le llamo factor de bloqueo, pues su función era bloquear la variabilidad que se producía sobre la variable respuesta. En esta sección vamos a estudiar situaciones en las que consideramos dos factores de bloqueo, en cuyo caso decimos que se esta realizando un diseño en cuadrado latino.

En este tipo de diseño se tiene que sobre la variable respuesta influyen cuatro factores: el factor de interés a través de sus tratamientos, 2 factores de bloqueo y el error experimental el cual contiene todas aquellas variables que no han sido tomadas en cuenta. Los 2 factores de bloque son conocidos como bloque columna y bloque fila, o simplemente, columna y fila, estos deben tener el mismo numero de niveles, es por ello que se llama cuadrado. Los tratamientos se denotan con las letras latinas, razón por la cual se llama latino, y solo aparece uno por cada combinación de fila-columna, por lo tanto el número de tratamientos es igual al número de filas y columnas y en cada fila y columna deben estar presentes todos los tratamientos, los mismos son asignados de manera aleatoria en cada fila o columna.

Por lo tanto la variable respuesta, de un diseño con k tratamientos, k filas y k columnas, puede modelarse de la siguiente manera:

$$Y_{ijr} = \mu + \tau_i + \beta_j + \gamma_r \varepsilon_{ij}; (i, j, r) = 1, 2, \cdots, k \tag{3.1}$$

donde

Y_{ijr} es la observación del tratamiento i, en la fila j y la columna r.

μ es la media general

τ_i es el efecto de i-ésimo tratamiento

β_j es el efecto del j-ésimo nivel del factor fila.

γ_r es el efecto del r-ésimo nivel del factor columna.

ε_{ijr} es el error aleatorio, los cuales se suponen $N(0, \sigma^2)$ e independientes

Suponiendo que el modelo es de efectos fijos se cumple que

$$\sum_{i=1}^{b} \tau_i = \sum_{j=1}^{k} \beta_j = \sum_{j=1}^{k} \gamma_k = 0$$

Un supuesto adicional a los ya considerados, es que los factores afectan los resultados en forma independiente, uno de otro. Esto es, la interacción no es importante o, no existe.

Ahora bien, los datos observados de un diseño en el cual los k tratamientos son arreglados en k filas y k columnas pueden presentase como en la tabla **??**

Al igual que en los diseños anteriores la hipótesis a probar es:

Tabla 3.1: Aspectos de los datos en un diseño de cuadrado latino

		Columna				
		1	2	3	$\cdots$	k
	1	$A = y_{111}$	$B = y_{212}$	$C = y_{313}$	$\cdots$	$K = y_{k1k}$
	2	$B = y_{221}$	$C = y_{322}$	$D = y_{423}$	$\cdots$	$A = y_{12k}$
Fila	3	$C = y_{331}$	$D = y_{432}$	$E = y_{533}$	$\cdots$	$B = y_{23k}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	k	$K = y_{kk1}$	$A = y_{1k2}$	$B = y_{2k3}$	$\cdots$	$J = y_{jkk}$

$$H_0 \;:\; \tau_1 = \tau_2 = \cdots = \tau_k = 0 \tag{3.2}$$

$$H_1 \;:\; \tau_i \neq 0 \text{ para algún } i \tag{3.3}$$

Para probar dicha hipótesis se realiza un análisis de varianza.

Análisis de Varianza

La hipótesis dada se prueba con un *análisis de varianza con tres criterios de clasificación*; se utilizan los tres criterios porque se controlan tres fuentes de variación: el factor de tratamientos y los dos factores de bloque.

Una medida de la desviación de las observaciones con respecto a la media está dada por $\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(y_{ijr} - \bar{y_{\ldots}})^2$, la cuál restándole y sumándole los promedios de los tratamientos, de las filas, las columnas y el promedio general y ordenándolo convenientemente se tiene que

$$\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(y_{ijk} - \bar{y_{\ldots}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}[(\bar{y}_{i..} - \bar{y}_{\ldots}) + (\bar{y}_{.j.} - \bar{y}_{\ldots}) + (\bar{y}_{..r} - \bar{y}_{\ldots}) \tag{3.4}$$

$$+ \; (y_{ij} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..r} + 2\bar{y}_{..})]^2 \tag{3.5}$$

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

$$\begin{aligned}\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(y_{ijk}-\bar{y_{...}})^2 &= \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(\bar{y}_{i..}-\bar{y}_{...})^2+\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(\bar{y}_{.j.}-\bar{y}_{...})^2 \\ &+ \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(\bar{y}_{..r}-\bar{y}_{...})^2+\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}(y_{ij}-\bar{y}_{i..}-\bar{y}_{.j.}-\bar{y}_{..r}+2\bar{y}_{..})^2\end{aligned}$$

lo cual representa la descomposición de la suma de cuadrados total. Esta ecuación se puede expresar de la siguiente manera

$$SC_T = SC_{Tr} + SC_F + SC_C + SC_E \tag{3.6}$$

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

$$\begin{aligned}SC_T &= \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k}y_{ijr}^2-\frac{y_{...}^2}{N}\\ SC_{Tr} &= \sum_{i=1}^{k}\frac{y_{i..}^2}{k}-\frac{y_{...}^2}{N}\\ SC_F &= \sum_{j=1}^{k}\frac{y_{.j.}^2}{k}-\frac{y_{...}^2}{N}\\ SC_C &= \sum_{r=1}^{k}\frac{y_{..r}^2}{k}-\frac{y_{...}^2}{N}\\ SC_E &= SC_T-SC_{Tr}-SC_F-SC_C\end{aligned}$$

los grados de libertad están dados por

$$
\begin{aligned}
GL_T &= N-1 \\
GL_{Tr} &= k-1 \\
GL_F &= k-1 \\
GL_C &= k-1 \\
GL_E &= GL_T - GL_{Tr} - GL_F - GL_C = (k-2)(k-1)
\end{aligned}
$$

los cuadrados medios en este caso son

$$
CM_{Tr} = \frac{SC_{Tr}}{k-1} \qquad CM_F = \frac{SC_F}{k-1} \qquad CM_C = \frac{SC_C}{k-1} \qquad CM_E = \frac{SC_E}{(k-2)(k-1)}
$$

cuyos valores esperados son

$$
\begin{aligned}
E(CM_E) &= \sigma^2 \\
E(CM_{Tr}) &= \sigma^2 + \frac{k}{k-1}\sum_{j=1}^{k}\tau_j^2 \\
E(CM_F) &= \sigma^2 + \frac{k}{k-1}\sum_{i=1}^{n}\beta_i^2 \\
E(CM_C) &= \sigma^2 + \frac{k}{k-1}\sum_{r=1}^{k}\gamma_r^2
\end{aligned}
$$

y el estadístico de prueba es

$$
F_c = \frac{CM_{Tr}}{CM_E} \tag{3.7}
$$

La regla de decisión es rechazar H_0 si $F > F_{\alpha;GL_{Tr};GL_E}$.

Otras hipótesis que pueden ser de interés son las siguientes

- No existe efecto de fila

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \cdots = \beta_k = 0 \\ H_1 &: \beta_j \neq 0 \text{ para algún } j \end{aligned} \tag{3.8}$$

- No existe efecto de columna

$$\begin{aligned} H_0 &: \gamma_1 = \gamma_2 = \cdots = \gamma_k = 0 \\ H_1 &: \gamma_r \neq 0 \text{ para algún } r \end{aligned} \tag{3.9}$$

cuyos estadísticos de prueba son respectivamente

$$F_c^F = \frac{CM_F}{CM_E} \qquad F_c^C = \frac{CM_C}{CM_E} \tag{3.10}$$

y las reglas de decisión es rechazar H_0 si $F^F > F_{\alpha;GL_F;GL_E}$ y $F^C > F_{\alpha;GL_C;GL_E}$.

Todo este desarrollo lo podemos resumir como se muestra en la Tabla **??**. Esta tabla recibe el nombre de Tabla de Análisis de Varianza para el Diseño de Cuadrados Latinos.

Ejemplo 3.1 *Se sospecha que cualquier estímulo produce cambios en la sensibilidad del ojo humano adaptado a la oscuridad. Para investigar esto, se diseñó un experimento el cual consistió en someter a cinco individuos durante cinco días seguidos a cinco estímulos diferentes una vez que sus ojos se adaptaron a la oscuridad. Se registró como resultado, sus sensibilidades a la prueba de bajo contraste de Luckiesh-Moss.*

Tabla 3.2: Tabla de Análisis de Varianza para el DCL

Fuente de Variación	Suma de Cuadrados	Grado de Libertad	Cuadrado Medio	F
Tratamiento	SC_{Tr}	k-1	CM_{Tr}	F_c
Fila	SC_F	k-1	CM_F	F_c^F
Columna	SC_C	k-1	CM_C	F_c^C
Error	SC_E	(k-1)(k-2)	CM_E	
Total	SC_T	k^2-1		

Tabla 3.3: Tabla de Análisis de Varianza para el DCL

		Días				
		1	2	3	4	5
Sujetos	1	A=22	B=21	D=22	C=20	E=22
	2	C=23	D=22	A=16	E=23	B=19
	3	D=20	A=14	E=14	B=23	C=24
	4	B=28	E=29	D=24	C=24	A=24
	5	E=4	C=2	B=3	A=8	D=8

3.2. Preguntas y Ejercicios

3.3. Preguntas y Ejercicios

1. ¿Qué es un diseño de bloques completamente aleatorios?

2. ¿Cuándo es apropiado utilizar un diseño de bloques completamente aleatorios?.

3. ¿Cuál es el modelo de un diseño de bloques completamente aleatorios?

4. ¿Qué diferencia hay entre un diseño completamente aleatorizado y uno de bloques

completos?.

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

6. Explique por que el adjetivo aleatorios en el nombre del diseño de bloques completamente aleatorios

7. Demuestre que los productos cruzados obtenidos en la partición de las sumas de cuadrados son iguales a cero.

8. Demuestre que

$$\begin{aligned} E(CM_E) &= \sigma^2 \\ E(CM_{Tr}) &= \sigma^2 + \frac{b\displaystyle\sum_{j=1}^{k} \tau_j^2}{k-1} \\ E(CM_B) &= \sigma^2 + \frac{k\displaystyle\sum_{i=1}^{n} \beta_i^2}{b-1} \end{aligned}$$

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

Fuente de Variación	Suma de Cuadrados	Grado de Libertad	Cuadrado Medio	F calculado
Tratamiento	600			
Bloque	850			
Error	500			
Total				

a) Complete la tabla.

b) Escriba el modelo estadístico y las hipótesis pertinentes.

c) Apoyándose en tablas de la distribución F, decida i se aceptan o se rechazan las hipótesis.

10. Realice el problema anterior, pero ahora suponga que no se hay bloqueado. ¿Se hubiese obtenido las mismas conclusiones?. Argumente.

11. Una compañía farmacéutica realizó un experimento para comprobar los tiempos promedio (en días) necesarios para que una persona se recupere d elos 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 contacto 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 como aplicaría la idea de bloqueo para controlar tal fuente de variabilidad.

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

		Tratamiento		Total bloque
	A	B	C	
Bloque 1	3	7	4	$y_{1.} =$
2	4	9	6	$y_{2.} =$
3	2	3	3	$y_{b.} =$
4	6	10	7	$y_{b.} =$
Total Trat	$y_{,1} =$	$y_{,2} =$	$y_{,3} =$	$y_{..} =$

a) Complete los totales que se4 piden en la tabla anterior.

b) Calcule las sumas de cuadrados correspondientes.

c) Obtenga la tabla de análisis de varianza y anote las principales conclusiones.

d) De ser necesario realice el análisis a posteriori usando el método de la diferencia mínima significativa.

13. 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.

Marca de atomizador	Número de réplica (día)					
	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 las hipótesis adecuadas y el modelo estadístico.

b) ¿Existe diferencia entre la efectividad promedio de los atomizadores?.

c) ¿Hay algún atomizador mejor?.

d) ¿Hay diferencias significativas en los resultados de diferentes días en que se realizó el experimento?. Argumente.

14. En una empresa lechera se tienen varios silos para almacenar leche (ciusternas 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.

Silo	Día				
	Lunes	Martes	Miércoles	Jueves	Viernes
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) Formule las hipótesis adecuadas y el modelo estadístico

c) ¿Hay diferencia entre los silos?

d) La temperatura de un día a otro es diferente?

15. Se diseño un experimento para estudiar el rendimiento de cuatro detergentes. Las siguientes lecturas de ”blancura”se obtuvieron con un equipo especial diseñado para doce cargas de lavado distribuidas en tres modelos de lavadoras.

Detergente	Lavadora 1	Lavadora 2	Lavadora 3
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.

16. Una química desea probar el efecto que tienen cuatro agentes químicos sobre la resistencia de un tipo particular de tela. Como puede haber variabilidad entre un rollo de tela y otro, decide utilizar un diseño aleatorizado por bloques, considerando los rollos de tela como bloques. Ella selecciona 5 rollos y les aplica los cuatro agentes químicos en orden aleatorio. A continuación, se proporcionan los resultados de la resistencia a la tensión. Analice estos datos y haga las conclusiones apropiadas.

Agente químico	Rollo de tela				
	1	2	M3	4	5
1	73	68	74	71	67
2	73	67	75	72	70
3	75	68	78	73	68
4	73	71	75	75	69

17. ¿Qué es un diseño de cuadrados latinos?

18. ¿Cuándo es apropiado utilizar un diseño de cuadrados latinos ?.

19. ¿Cuál es el modelo de un diseño de de cuadrados latinos?

20. ¿Qué diferencia hay entre un diseño completamente aleatorizado, uno de bloques completos y uno de cuadrados latinos?.

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

22. Explique por que el el nombre de cuadrados latinos.

23. Demuestre que los productos cruzados obtenidos en la partición de las sumas de cuadrados son iguales a cero.

24. Una compañía de mensajería está interesada en determinar cuál marca de llantas tiene mayor duración, la medida ésta en términos del desgaste. Para ella se planea un experimento en el que se comparan las cuatro marcas de llantas sometiéndolas a una prueba de 32.000 kilómetros de recorrido, utilizando cuatro diferentes tipos de auto y las cuatro posiciones posibles de las llantas en el auto.

	Carro			
Posición	1	2	3	4
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

$a)$ Anote la ecuación del modelo y las hipótesis estadísticas correspondientes.

b) ¿Existen diferencias entre los tratamientos? ¿Cuáles tratamientos son diferentes entre sí?.

c) ¿Los factores de marca de llanta, tipo de auto y posiciones influyen en la duración?

25. 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 cuadrado latino, para controlar activamente a los lotes y días. Los datos obtenidos son:

		Día				
		1	2	3	4	5
	1	A=8	B=7	D=1	C=7	E=3
	2	C=11	E=2	A=7	D=3	B=8
Lote	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) Anote la ecuación del modelo y las hipótesis estadísticas correspondientes.

b) ¿Existen diferencias entre los tratamientos? ¿Cuáles tratamientos son diferentes entre sí?.

c) ¿Los factores de ruido, lote y día afectan el tiempo de reacción del proceso?