Experimentos Factoriales

Maestría en Estadística – UNSAAC
Diseño de Experimentos

1. Introducción

En capítulos anteriores se estudiaron los diseños Completamente al Azar, de Bloques Completos al Azar y Cuadrado Latino, en los cuales se analizó un solo factor de tratamientos. Un experimento factorial es aquel en el que se estudian simultáneamente varios factores, de modo que los tratamientos se forman por todas las posibles combinaciones de los niveles de los factores. Un experimento factorial no constituye un nuevo diseño experimental, sino un diseño para la formación de los tratamientos. Los experimentos factoriales pueden ser conducidos bajo los lineamientos de cualquier diseño experimental tal como el DCA, DBCA o DCL.
Los experimentos factoriales son ampliamente utilizados y son de gran valor en el trabajo exploratorio cuando se sabe poco sobre los niveles óptimos de los factores o ni siquiera qué factores son importantes. Estos experimentos son útiles también en campos de estudio más complejos en las que se sabe que un factor no actúa independientemente sino en estrecha relación con otros factores. En este capítulo se tratará el caso de experimentos con dos factores y conducidos bajo los lineamientos de un DCA y DBCA.

2. Ventajas y desventajas

2.1. Ventajas:

- Permite obtener más información que en un experimento de un solo factor, ya que se estudian los efectos principales, los efectos simples, los efectos cruzados y de interacción entre los factores.

- Todas las unidades intervienen en la estimación de los efectos principales y de interacción, por lo que el número de repeticiones es elevado para estos casos.

2.2. Desventajas:

- Se requiere un mayor número de unidades experimentales que en los experimentos simples.

- Dado que todos los niveles de un factor se combinan con todos los niveles de los otros, por requerimientos del análisis estadístico, se tendrá que algunas combinaciones que no son de interés para el investigador, serán también incluidas en el experimento.

- El análisis estadístico y la interpretación de los resultados es más complicado que en los experimentos simples con un solo factor, y aumenta considerablemente conforme más factores son incluidos.

3. Notación y definiciones

3.1. Factor

Los factores son designados por letras mayúsculas. Por ejemplo en un experimento en que se evalúan 3 cantidades de semilla con 4 dosis de nitrógeno por parcela y 2 variedades de maíz destinado a chala, el factor cantidad de semilla se puede denotar por A, el factor dosis de nitrógeno por B y el factor variedad de maíz por C.

3.2. Niveles de un Factor

Los niveles de un factor son denotados por letras minúsculas con subíndices. Por ejemplo, las 3 cantidades de semilla podrían ser denotadas por a_1, a_2 y a_3, las 4 dosis de nitrógeno por b_1, b_2, b_3 y b_4, y las dos variedades de maíz por c_1 y c_2.

Una combinación de letras minúsculas con sus respectivos subíndices es utilizada para denotar una combinación de los niveles de los factores. Por ejemplo, la combinación $a_2b_2c_1$ denotará el tratamiento conformado por la aplicación de la cantidad a_2 de semilla con la dosis b_2 de nitrógeno y la variedad c_1 de maíz.

3.3. Tipos de Factores

Dependiendo de la naturaleza de los niveles de los factores, estos pueden ser cualitativos o cuantitativos. En el ejemplo, los factores A y B serían cuantitativos y el factor C cualitativo. En el caso de factores cuantitativos, estos pueden ser igualmente espaciados o no. Así, por ejemplo, para el factor B, niveles de 0, 10, 20 y 30 kg/parcela y de 10, 20, 40 y 80 kg/parcela constituirían niveles igualmente espaciados y no igualmente espaciados respectivamente.

Adicionalmente, los factores pueden ser fijos o al azar, dependiendo de la forma en que son seleccionados sus niveles. Un experimento factorial con todos sus factores fijos corresponderá a un **modelo I** o de efectos fijos; un experimento factorial con todos sus factores aleatorios corresponderá a un **modelo II** o de efectos aleatorios; y un experimento factorial con algunos factores fijos y otros aleatorios corresponderá a un **modelo III** o de efectos mixtos. En el desarrollo de este capítulo se considerará que todos los factores son fijos.

3.4. Efectos en los Experimentos Factoriales

- **Efecto Principal**: Es el efecto de un factor en promedio sobre los niveles de los otros factores.
- **Efecto Simple**: Es el efecto de un factor, en un nivel de los demás factores.
- **Efecto de Interacción**: Está dado por la variación que tiene un efecto simple de un factor al pasar de un nivel a otro de otro factor.
- **Efecto Cruzado**: Está dado por las combinaciones cruzadas de dos factores.

Ejemplo 1

A continuación se presentan datos para un experimento factorial 2×2.

	Niveles del Factor A (a_1)		Niveles del Factor A (a_2)	
Niveles del Factor B	b_1	b_2	b_1	b_2
Medias	54	38	45	56

Efectos Simples

- De A en b_1: $ES\big(A(b_1)\big) = a_1b_1 - a_2b_1 = 54 - 45 = 9$
- De A en b_2: $ES\big(A(b_2)\big) = a_1b_2 - a_2b_2 = 38 - 56 = -18$
- De B en a_1: $ES\big(B(a_1)\big) = a_1b_1 - a_1b_2 = 54 - 38 = 16$
- De B en a_2: $ES\big(B(a_2)\big) = a_2b_1 - a_2b_2 = 45 - 56 = -11$

Efectos Principales

- De A:

$$EP(A) = \tfrac{1}{2}\Big[ES\big(A(b_1)\big) + ES\big(A(b_2)\big)\Big] = \frac{9 + (-18)}{2} = \frac{-9}{2} = -4{,}5$$

- De B:

$$EP(B) = \tfrac{1}{2}\Big[ES\big(B(a_1)\big) + ES\big(B(a_2)\big)\Big] = \frac{16 + (-11)}{2} = \frac{5}{2} = 2{,}5$$

Efecto de Interacción

- De AB:

$$EI(AB) = \tfrac{1}{2\times 2}\Big[ES\big(A(b_1)\big) - ES\big(A(b_2)\big)\Big] = \frac{9 - (-18)}{4} = \frac{27}{4} = 6{,}75$$

$$EI(AB) = \tfrac{1}{2\times 2}\Big[ES\big(B(a_1)\big) - ES\big(B(a_2)\big)\Big] = \frac{16 - (-11)}{4} = \frac{27}{4} = 6{,}75$$

Efectos Cruzados

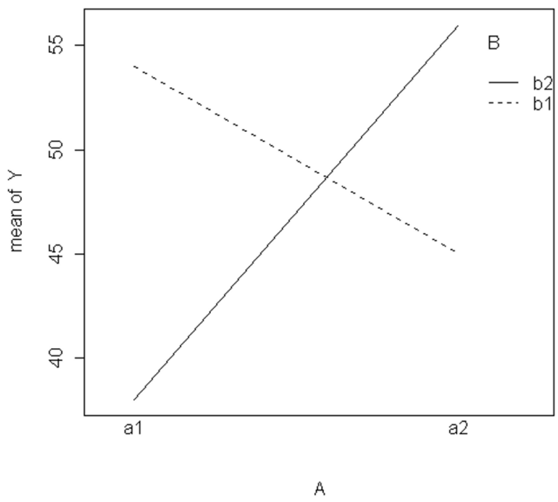
- Entre a_1b_1 y a_2b_2:

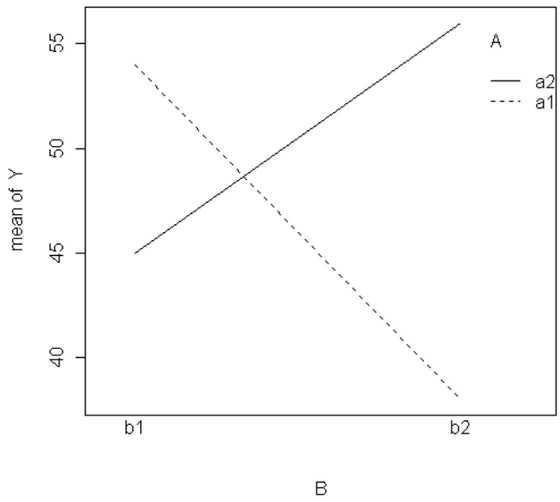
$$EC(a_1b_1 - a_2b_2) = a_1b_1 - a_2b_2 = 54 - 56 = -2$$

- Entre a_1b_2 y a_2b_1:

$$EC(a_1b_2 - a_2b_1) = a_1b_2 - a_2b_1 = 38 - 45 = -7$$

La interacción entre dos factores puede también analizarse gráficamente. El gráfico de la interacción presenta las medias de los niveles de un factor en cada uno de los niveles del otro. Con los datos del ejemplo se tienen los siguientes gráficos:

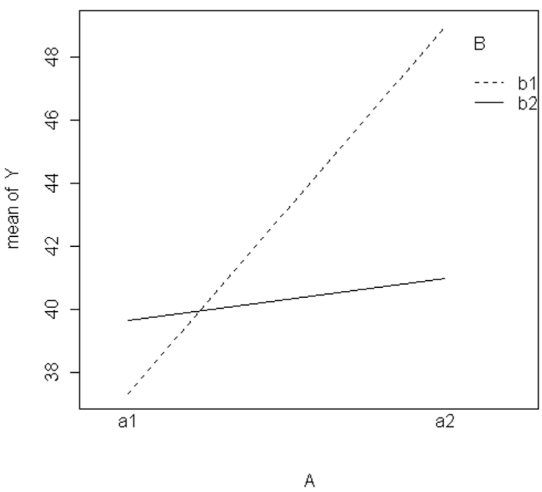


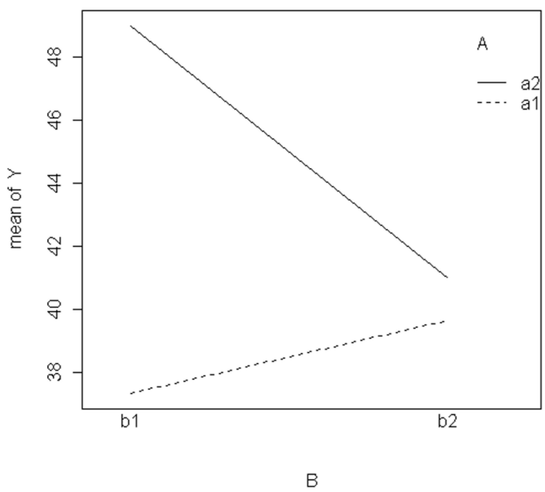


En un experimento factorial, cada línea corresponde a un *efecto simple*, y la *interacción* se manifiesta cuando las líneas tienen pendientes diferentes; es decir, cuando el efecto simple de un factor no es el mismo en todos los niveles del otro.

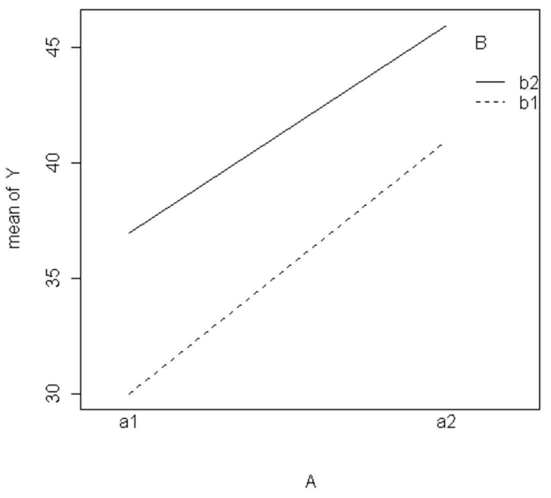
En el primer gráfico se muestran los efectos simples de A en b_1 (línea punteada) y de A en b_2 (línea continua), siendo evidente que el efecto simple de A depende del nivel de B (lo cual indica la existencia de interacción).

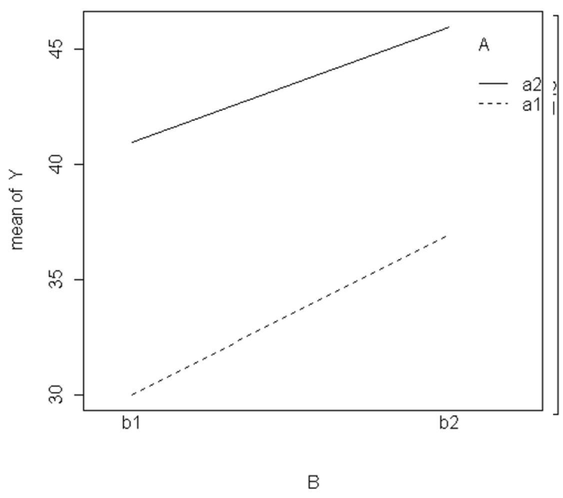
A continuación se presenta otro ejemplo de interacción:





Note que no es necesario que las líneas se crucen para evidenciar una interacción entre ambos factores. Un caso de factores sin interacción aparente sería el siguiente:





4. Experimento Factorial pxq

4.1. Modelo Aditivo Lineal

En un DCA, el modelo aditivo lineal está dado por:

$$Y_{ij} = \mu + \tau_i + e_{ij} \qquad i = 1, \dots, t$$

Ahora, dado que los tratamientos son generados por las combinaciones entre los niveles de dos factores, el efecto de los tratamientos se descompone en el efecto del factor A, el efecto del factor B y el efecto de la interacción entre los dos factores. Así, el modelo aditivo lineal para un factorial pxq en DCA será:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \qquad i = 1, \dots, p \quad j = 1, \dots, q \quad k = 1, \dots, r_{ij}$$

donde:

- Y_{ijk} es el valor o rendimiento observado con el i-ésimo nivel del factor A, j-ésimo nivel del factor B, k-ésima repetición.
- μ es el efecto de la media general.
- α_i es el efecto del i-ésimo nivel del factor A.
- β_j es el efecto del j-ésimo nivel del factor B.
- $(\alpha\beta)_{ij}$ es el efecto de la interacción en el i-ésimo nivel del factor A y j-ésimo nivel del factor B.
- e_{ijk} es el error experimental en el i-ésimo nivel del factor A, j-ésimo nivel del factor B, k-ésima repetición.
- p es el número de niveles del factor A.
- q es el número de niveles del factor B.
- r_{ij} es el número de repeticiones en el i-ésimo nivel del factor A, j-ésimo nivel del factor B.

En el caso de un experimento factorial en DBCA, el modelo aditivo lineal es:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + e_{ijk} \qquad i = 1, \ldots, p \quad j = 1, \ldots, q \quad k = 1, \ldots, b$$

donde:

- γ_k es el efecto del k-ésimo bloque.
- b es el número de bloques.

Los supuestos del modelo serán los mismos que para el DCA o DBCA con un solo factor vistos en capítulos anteriores. Los cálculos y procedimientos presentados de aquí en adelante corresponderán al caso del experimento factorial pxq en DBCA. El caso del experimento factorial en DCA es similar al del DBCA y será ilustrado mediante un ejemplo.

Ejemplo 2 (Experimento Factorial en DBCA):

Se realizó un experimento con un arreglo factorial $2A3B$ en DBCA en 4 campos de cultivo, para evaluar el efecto en el rendimiento de maíz obtenido con dos tipos de abono (a_1 y a_2) y tres dosis ($b_1 = 20$, $b_2 = 30$ y $b_3 = 40$ kg/ha). Los resultados obtenidos (en TM/ha) se presentan a continuación:

Campos	a_1			a_2		
	b_1	b_2	b_3	b_1	b_2	b_3
1	1.9	1.8	2.7	1.8	2.9	3.0
2	2.3	2.1	2.4	2.2	2.7	3.2
3	2.0	2.4	2.9	2.0	3.2	2.9
4	2.1	2.9	2.8	2.4	3.5	3.4
Total	8.3	9.2	10.8	8.4	12.3	12.5

El modelo aditivo lineal para este ejemplo será:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + e_{ijk} \qquad i = 1, \ldots, p \quad j = 1, \ldots, q \quad k = 1, \ldots, b$$

donde:

- Y_{ijk} es el rendimiento de maíz en Tm/Ha obtenido con el i-ésimo tipo de abono, j-ésima dosis, k-ésimo campo de cultivo.
- μ es el efecto de la media general.
- α_i es el efecto del i-ésimo tipo de abono.
- β_j es el efecto de la j-ésima dosis de abono.
- $(\alpha\beta)_{ij}$ es el efecto de la interacción en el i-ésimo tipo de abono, j-ésima dosis.
- γ_k es el efecto del k-ésimo campo de cultivo.
- e_{ijk} es error experimental en el i-ésimo tipo de abono, j-ésima dosis, k-ésimo campo de cultivo.
- $p = 2$ es el número de niveles del factor A.
- $q = 3$ es el número de niveles del factor B.
- $b = 4$ es el número de bloques.

4.2. Estimación de los Efectos

Los efectos del modelo, μ, α_i, β_j, $(\alpha\beta)_{ij}$ y γ_k son estimados de modo que se minimice la siguiente expresión (Método de Mínimos Cuadrados):

$$Q = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{b} e_{ijk}^2 = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{b} \left(Y_{ijk} - \mu - \alpha_i - \beta_j - (\alpha\beta)_{ij} - \gamma_k\right)^2$$

teniendo en cuenta las siguientes restricciones:

$$\sum_{i=1}^{p}\alpha_i = 0, \qquad \sum_{j=1}^{q}\beta_j = 0, \qquad \sum_{i=1}^{p}(\alpha\beta)_{ij} = 0, \qquad \sum_{j=1}^{q}(\alpha\beta)_{ij} = 0, \qquad \sum_{k=1}^{b}\gamma_k = 0$$

La aplicación de este método da los siguientes resultados para la estimación de los parámetros:

$$\hat{\mu} = \overline{Y}_{...}, \qquad \hat{\alpha}_i = \overline{Y}_{i..} - \overline{Y}_{...}, \qquad \hat{\beta}_j = \overline{Y}_{.j.} - \overline{Y}_{...},$$
$$\hat{\gamma}_k = \overline{Y}_{..k} - \overline{Y}_{...}, \qquad (\hat{\alpha\beta})_{ij} = \overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...}$$

Y el residual

$$\hat{e}_{ijk} = Y_{ijk} - \overline{Y}_{ij\bullet} - \overline{Y}_{\bullet\bullet k} + \overline{Y}_{\bullet\bullet\bullet}$$

4.3. Análisis de Variancia

En este modelo la variabilidad total se descompone de la siguiente manera:
Variabilidad (Total) = Var (Tratamientos) + Var (Bloques) + Var (Error)
donde a su vez, la variabilidad correspondiente a los tratamientos se descompone en:
Variabilidad (Tratamientos) = Var (Factor A) + Var (Factor B) + Var (Interacción AB)
La variabilidad total es cuantificada por la suma de cuadrados total:

Fuentes de Variación	Grados de Libertad (GL)	Suma de Cuadrados (SC)	Cuadrados Medios (CM)	Fc
Bloques	$b-1$	SC(Bloques)	$\frac{SC(\text{Bloques})}{GL(\text{Bloques})}$	
A	$p-1$	$SC(A)$	$\frac{SC(A)}{GL(A)}$	$\frac{CM(A)}{CM(\text{Error})}$
B	$q-1$	$SC(B)$	$\frac{SC(B)}{GL(B)}$	$\frac{CM(B)}{CM(\text{Error})}$
AB	$(p-1)(q-1)$	$SC(AB)$	$\frac{SC(AB)}{GL(AB)}$	$\frac{CM(AB)}{CM(\text{Error})}$
Error Experimental	$(pq-1)(b-1)$	SC(Error)	$\frac{SC(\text{Error})}{GL(\text{Error})}$	
Total	$pqb-1$	SC(Total)		

Cuadro 1: Tabla ANOVA factorial en bloques