

Tema 31 Continuación ANOVA.

1. Introducción al diseño experimental. Uniel 176 ~~181~~
2. Modelo en bloques aleatorizados. Pete 510 / Uniel 177
3. Cuadrados latinos. Pete 513 / Uniel 182
4. Experimentos factoriales.

DISEÑO EN BLOQUES COMPLETOS CON TRATAMIENTOS ALEATORIZADOS

Este diseño consiste en asignar los tratamientos en forma completamente al azar a un grupo de parcelas, llamados bloques o repeticiones, con la condición que dicha bloque o conjunto de unidades experimentales sea lo más homogéneo posible. El objetivo de realizar un agrupamiento de unidades experimentales, es escoger aquellas más parecidas de manera de reducir hasta donde sea posible la variabilidad dentro de cada bloque y así evitar que los efectos de tratamiento se vean enmascarados o confundidos por la heterogeneidad de las unidades experimentales a las cuales se asignan los tratamientos de interés. Es por ello que el diseño en bloques completos con tratamientos aleatorizados DBA se considera más eficiente que el diseño completamente al azar (DCA) para controlar la variabilidad del material experimental. Generalmente, el número de bloques es igual al número de tratamientos. El adjetivo "completos" se refiere a que todos los tratamientos aparecen representados en cada uno de los bloques del experimento.

Este diseño ofrece menos grados de libertad para estimar el error, comparado con el DCA, esta reducción se debe a los grados de libertad necesarios para estimar el efecto de bloques. Es por ello que debe existir una razón real de bloqueo para usar este diseño, es decir entre mayor sea la variabilidad entre bloques más eficiente será el diseño para detectar diferencias entre tratamientos.

Aleatorización

En este diseño la aleatorización consiste en asignar los tratamientos a los grupos de unidades experimentales o bloques en forma completamente aleatoria, es decir con la restricción de aleatorizar dentro de bloques. Así por ejemplo, si se prueban cuatro tratamientos y se tienen cuatro bloques, dichos tratamientos se asignaran completamente al azar dentro de cada bloque.

Modelo estadístico asociado al diseño:

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \quad \begin{array}{l} i = 1, 2, 3, \dots, t \\ j = 1, 2, 3, \dots, r \end{array}$$

donde:

Y_{ij} = Variable respuesta en la j-ésima repetición del i-ésimo tratamiento

μ = Media general

τ_i = Efecto del tratamiento i.

β_j = Efecto del bloque j

ε_{ij} = Error aleatorio, donde $\varepsilon_{ij} \sim N(0, \sigma^2)$

Análisis de la Varianza para el modelo $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$

Ho: $\tau_1 = \tau_2 = \dots = \tau_t$

Ha: al menos un efecto de un tratamiento es diferente de los demás.

| Fuentes de Variación (F.V.) | Grados de Libertad (G.L.) | Suma de Cuadrados (S.C.) | Cuadrados Medios (C.M.) | F ₀ |
|-----------------------------|---------------------------|--|--------------------------------|------------------------------|
| Tratamientos | t-1 | $\sum_{i=1}^t \frac{Y_i^2}{b} - \frac{Y_{..}^2}{bt}$ | $\frac{S.C.TRAT.}{t-1}$ | $\frac{C.M.TRAT}{C.M.ERROR}$ |
| Bloques | b-1 | $\sum_{j=1}^b \frac{Y_j^2}{t} - \frac{Y_{..}^2}{bt}$ | | |
| Error | (t-1)(b-1) | $\sum_{i=1}^t \sum_{j=1}^b Y_{ij}^2 - \sum_{i=1}^t \frac{Y_i^2}{b} - \sum_{j=1}^b \frac{Y_j^2}{t} + \frac{Y_{..}^2}{bt}$ | $\frac{S.C.ERROR}{(t-1)(b-1)}$ | |
| Total | (bt-1) | $\sum_{i=1}^t \sum_{j=1}^b Y_{ij}^2 - \frac{Y_{..}^2}{bt}$ | | |

[Siguiente] [Arriba]

En el capítulo 3 se expuso el diseño de experimentos más sencillo, el modelo completamente aleatorizado, que tiene un factor tratamiento. Con el fin de reducir la variabilidad residual de este modelo se puede introducir en el mismo un factor-bloque para obtener el modelo de diseño en bloques completamente aleatorizados, primer modelo que se estudia en este capítulo. El siguiente modelo, un poco más complejo, es el modelo con dos factores tratamiento entre los que puede haber interacción. El estudio de estos modelos es fácilmente generalizable a modelos con más factores tratamiento y factores bloque. El último modelo que se estudia en este capítulo es el diseño fraccional de cuadrado latino, que es un buen ejemplo de diseño fraccional.

5.1 Concepto de bloque.

Al estudiar la influencia de un factor-tratamiento en una variable de interés puede ser importante eliminar (controlar) estadísticamente la influencia de un factor que puede influir en la variable respuesta. Para ello se utiliza el concepto de bloque, que se basa en seleccionar niveles de esta variable y aplicar en cada uno de ellos todos los niveles del factor principal, de esta forma disminuye la variabilidad residual o no explicada. Por tanto, un factor-bloque es un factor cuyo control puede reducir significativamente la variabilidad no explicada y que no interacciona con los factores principales.

El siguiente ejemplo ayuda a comprender estas ideas.

Ejemplo 5.1.

Una empresa fotográfica tiene que realizar una compra de impresoras de gran calidad que se van a utilizar en imprimir fotografías digitales. La empresa tiene ofertas de I marcas de impresoras de similares características y precio. Para la empresa fotográfica es muy importante la "*velocidad de impresión*" y por este motivo está interesada en saber si las I impresoras ofertadas tienen la misma velocidad o si hay una que es más rápida. Para responder a esta pregunta decide hacer un experimento que se puede plantear de dos formas:

[1] De los muchos ficheros de fotos digitales que tiene la empresa, elegir al azar I muestras de J fotos e imprimir en cada una de las impresoras una de las muestras, aleatorizando la asignación de muestras que se deben imprimir en cada impresora.

Esta estrategia es la del modelo de diseño de experimentos completamente aleatorizado que es perfectamente válido. En este ejemplo la variable de interés es la "*velocidad de impresión*" y el factor-tratamiento "*el tipo de impresora*".

Un inconveniente que puede tener esta estrategia es que exista una fuerte variabilidad en el tipo de fotos, esto es, que haya fotos que se impriman en poco tiempo y otras no, independientemente de la impresora utilizada. En este caso la variabilidad de la respuesta "*velocidad de impresión*" es debida

no solo al "*tipo de impresora*" sino también al "*tipo de fotos*" seleccionadas. Si la variabilidad debida al "*tipo de fotos*" es muy grande y no se tiene en cuenta, la variabilidad residual del modelo es grande y puede enmascarar la significatividad del factor de interés, el "*tipo de impresora*". Este problema se puede reducir en parte si el tamaño muestral es muy grande, aunque tiene el inconveniente de tener un mayor coste.

[2] Una estrategia alternativa es elegir una única muestra de J fotos e imprimirlas en cada una de las I impresoras, de esta forma se controla la variabilidad debida al "*tipo de fotos*". Esta estrategia es fuertemente recomendable si se supone que la variabilidad del "*tipo de fotos*" es alta.

Téngase en cuenta que el número de pruebas a realizar según las dos estrategias propuestas es el mismo: IJ .

La segunda propuesta conlleva el bloqueo de las unidades experimentales: *cada foto es un bloque*.

En este ejemplo se está interesado en estudiar la influencia del factor tratamiento "*tipo de impresora*" pero eliminando o controlando la posible influencia factor bloque "*tipo de foto*" en la variable respuesta "*velocidad de impresión*".

Los resultados del experimento se recogen en una tabla como la siguiente

| | Bloq.1 | Bloq.2 | | Bloq.J |
|--------|----------|----------|--|----------|
| Trat.1 | y_{11} | y_{12} | | y_{1J} |
| Trat.2 | y_{21} | y_{22} | | y_{2J} |
| | | | | |
| Trat.I | y_{I1} | y_{I2} | | y_{IJ} |

Del ejemplo anterior se deduce que

"Bloquear un experimento consiste en distribuir las unidades experimentales en grupos tales que unidades experimentales pertenecientes a un mismo grupo deben ser similares y pueden ser analizadas en condiciones experimentales semejantes, en tanto que unidades experimentales ubicadas en grupos distintos darán lugar, probablemente, a respuestas diferentes aún cuando sean asignadas a un mismo tratamiento.

Cada uno de los conjuntos de unidades experimentales similares se denomina **bloque**".

Del ejemplo anterior se deduce que:

"Bloquear un experimento consiste en distribuir las unidades experimentales en subgrupos tales que unidades experimentales pertenecientes a un mismo subgrupo deben ser similares y pueden ser analizadas en condiciones experimentales semejantes, en tanto que unidades experimentales ubicadas en subgrupos distintos darán lugar probablemente a respuestas diferentes aún cuando sean asignadas a un mismo tratamiento. Cada uno de estos conjuntos de unidades experimentales similares se denomina bloque."

Un diseño en bloques es apropiado cuando el objetivo del experimento es comparar los efectos de diferentes tratamientos promediados sobre un rango de condiciones experimentales distintas. Con los modelos de diseño de experimentos en bloques se quiere conseguir dos cosas:

1. evitar que grandes diferencias entre las unidades experimentales enmascaren diferencias reales entre los tratamientos,
2. medir los efectos de los tratamientos en condiciones experimentales distintas.

Un ejemplo de utilización de un diseño con bloques es el denominado de datos apareados para comparar dos tratamientos o medias de dos poblaciones (expuesto en el capítulo 1) cuando se aplican los dos tratamientos a los mismos individuos, en este caso cada individuo es un bloque.

[[Siguiete](#)] [[Arriba](#)]

[Siguiente] [Anterior] [Arriba]

5.2 Diseño en bloques completamente aleatorizados.

El modelo de diseño de experimentos con bloques más sencillo es el diseño de **bloques completamente aleatorizados**, con este diseño se quiere estudiar la influencia de un factor tratamiento ($T\alpha$) con I niveles en una variable de interés en presencia de una variable extraña, el factor bloque, $B\beta$, que tiene J bloques.

El motivo de la denominación de este modelo es la siguiente: se ha agrupado las unidades experimentales en J bloques, en función de $B\beta$, **aleatorizando** la forma de asignar los tratamientos dentro de cada bloque y es un diseño **completo** y equilibrado porque cada tratamiento se utiliza exactamente una vez dentro de cada bloque.

En este modelo, un **bloque** es un grupo de I unidades experimentales tan parecidas como sea posible con respecto a la variable $B\beta$, asignándose aleatoriamente cada tratamiento a una unidad dentro de cada bloque.

5.2.1 Modelo matemático.

La formulación matemática del modelo de diseño en bloques completamente aleatorizados con un factor principal (factor tratamiento), $T\alpha$, con I niveles y un factor secundario (factor bloque), $B\beta$, con J niveles o bloques es la siguiente:

Para cada $i = 1, \dots, I$; $j = 1, \dots, J$,

$$Y_{ij} = \overbrace{\mu + \alpha_i + \beta_j}^{\text{determinista}} + \underbrace{\varepsilon_{ij}}_{\text{aleatorio}}, \quad (5.1)$$

los ε_{ij} son v.a.i.i.d. según una $N(0, \sigma^2)$,

siendo,

* Y_{ij} el resultado del tratamiento i -ésimo, $i = 1, 2, \dots, I$ de $T\alpha$ al bloque j -ésimo, $j = 1, 2, \dots, J$.

* μ es la media de toda la población. Mide el nivel medio de todos los resultados.

* α_i es el efecto del tratamiento i -ésimo de $T\alpha$, $i = 1, 2, \dots, I$. Mide el efecto incremental del tratamiento del nivel i de $T\alpha$ sobre el efecto global. Se verifica que $\sum_{i=1}^I \alpha_i = 0$,

* β_j es el efecto del bloque j -ésimo, $j = 1, 2, \dots, J$, mide el efecto

incremental del tratamiento del factor secundario (bloque) sobre el efecto global (μ). Se verifica que $\sum_{j=1}^J \beta_j = 0$,
 ε_{ij} es el error experimental o perturbación, son variables aleatorias independientes idénticamente distribuidas (i.i.d.) con distribución $N(0, \sigma^2)$.

El número de observaciones es: $n = IJ$,

El problema básico que se plantea es contrastar la hipótesis nula de que el factor-tratamiento no influye,

$$H_0^{(\alpha)} \equiv \alpha_1 = \alpha_2 = \dots = \alpha_I = 0 \quad (5.2)$$

frente a la alternativa de que sí existen diferencias entre los valores medios de los distintos tratamientos.

En el estudio de este modelo debe de tenerse en cuenta que no existe interacción entre el factor-tratamiento y el factor-bloque y en el desarrollo el problema puede hacerse un segundo contraste acerca de si el factor-bloque es influyente o no. Este contraste es

$$H_0^{(\beta)} \equiv \beta_1 = \beta_2 = \dots = \beta_J = 0 \quad (5.3)$$

frente a la alternativa de que sí existen diferencias entre los valores medios de los distintos tratamientos del segundo factor. Sin embargo en el modelo tratamiento-bloque realizar este contraste carece de interés salvo para saber si ha sido conveniente bloquear o no.

Por ello en la práctica:

“Carece de interés plantearse la hipótesis nula de igualdad de los efectos bloque. El único objetivo puede ser el de concluir si bloquear el experimento resultó o no beneficioso”.

En efecto, si la suma de cuadrados medios atribuibles a los bloques es considerablemente mayor que la suma de cuadrados medios residual, habrá resultado útil bloquear en el sentido de que tal acción derivó en una reducción del tamaño del error experimental. En otro caso, bloquear es contraproducente.

5.2.2 Estimación de los parámetros.

El número de parámetros que hay que estimar en modelo (5.1) es

| Parámetros | Número |
|---|--------|
| μ | 1 |
| $\alpha_1, \alpha_2, \dots, \alpha_{I-1}$ | $I-1$ |
| $\beta_1, \beta_2, \dots, \beta_{J-1}$ | $J-1$ |
| σ^2 | 1 |

utilizando $n = IJ$ observaciones hay que estimar un número de parámetros

$$1 + (I - 1) + (J - 1) + 1 = I + J,$$

Se utiliza el método de mínimos cuadrados que se basa en minimizar la suma de los cuadrados de los residuos

$$\Psi(\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j) = \min_{\mu, \alpha_i, \beta_j} \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - (\mu - \alpha_i - \beta_j))^2, \quad (5.4)$$

se obtienen los siguientes estimadores:

$$\hat{\mu} = \bar{Y}_{..} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J Y_{ij}, \quad (5.5)$$

$$\hat{\alpha}_i = \bar{Y}_{i.} - \bar{Y}_{..}, \quad i = 1, \dots, I, \quad \text{con} \quad \bar{Y}_{i.} = \frac{1}{J} \sum_{j=1}^J Y_{ij}, \quad (5.6)$$

$$\hat{\beta}_j = \bar{Y}_{.j} - \bar{Y}_{..}, \quad j = 1, \dots, J, \quad \text{con} \quad \bar{Y}_{.j} = \frac{1}{I} \sum_{i=1}^I Y_{ij}, \quad (5.7)$$

Por tanto, la predicción en la casilla (i, j) es

$$\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j = \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}, \quad (5.8)$$

y los residuos son

$$e_{ij} = y_{ij} - \hat{y}_{ij}.$$

La suma de los residuos en cada fila y cada columna es cero, por tanto, hay $I + J - 1$ relaciones entre los IJ residuos y el número de grados de libertad es

$$g.l. = IJ - (I + J - 1) = (I - 1)(J - 1).$$

Razonando como en el modelo de diseño completamente aleatorizado se obtiene que el estimador de la varianza es la varianza residual

$$\hat{\sigma}_R^2 = \frac{1}{(I-1)(J-1)} \sum_{i=1}^I \sum_{j=1}^J e_{ij}^2 = \frac{SCR}{(I-1)(J-1)} \quad (5.9)$$

Propiedades de los estimadores.

La distribución de los estimadores anteriores es la siguiente,

$$\hat{\mu} = \bar{Y}_{..} \in N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\hat{\mu} - \mu}{\hat{\sigma}_R} \sqrt{n} \sim t_{(I-1)(J-1)} \quad (5.10)$$

$$\hat{\alpha}_i \sim N\left(\alpha_i, \sigma^2 \frac{I-1}{n}\right) \Rightarrow \frac{\hat{\alpha}_i - \alpha_i}{\hat{\sigma}_R} \sqrt{\frac{n}{I-1}} \sim t_{(I-1)(J-1)} \quad (5.11)$$

$$\hat{\beta}_j \sim N\left(\beta_j, \sigma^2 \frac{J-1}{n}\right) \Rightarrow \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_R} \sqrt{\frac{n}{J-1}} \sim t_{(I-1)(J-1)} \quad (5.12)$$

$$\frac{(I-1)(J-1)\hat{\sigma}_R^2}{\sigma^2} \sim \chi_{(I-1)(J-1)}^2 \quad (5.13)$$

Por tanto, los estimadores definidos son centrados y eficientes. Utilizando las distribuciones anteriores (la t y la χ^2) se pueden calcular intervalos de confianza de los parámetros del modelo.

Para calcular intervalos de confianza acerca de las medias de los niveles, las distribuciones de referencia son:

Para las medias de los niveles $(\mu + \alpha_i)$ del factor tratamiento $T\alpha$

$$\frac{(\mu + \alpha_i) - \bar{Y}_{i.}}{\hat{\sigma}_R} \sqrt{J} \sim t_{(I-1)(J-1)} \quad (5.14)$$

Para las medias de los bloques $(\mu + \beta_j)$ del factor bloque $B\beta$

$$\frac{(\mu + \beta_j) - \bar{Y}_{.j}}{\hat{s}_R} \sqrt{I} \sim t_{(I-1)(J-1)} \quad (5.15)$$

5.2.3 Análisis de la varianza.

Utilizando

$$e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) = \bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..},$$

se puede hacer la siguiente descomposición de las diferencias para cada $i = 1, \dots, I$;
 $j = 1, \dots, J$,

$$\begin{aligned} y_{ij} - \bar{y}_{..} &= (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \\ &= (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + e_{ij} = \hat{\alpha}_i + \hat{\beta}_j + e_{ij}, \end{aligned} \quad (5.16)$$

elevando al cuadrado en (5.16) y teniendo en cuenta que los dobles productos se anulan, la suma de cuadrados global se puede descomponer de la forma:

$$\begin{aligned} \underbrace{\overbrace{\sum_{i=1}^I \sum_{j=1}^J (y_{ij} - \bar{y}_{..})^2}_{g.l. = IJ-1}}_{\text{Suma de cuadrados Global (scG)}} &= \underbrace{\overbrace{J \sum_{i=1}^I (\bar{y}_{i.} - \bar{y}_{..})^2}_{g.l. = I-1}}_{\text{Suma de cuadrados Explicada por } T\alpha \text{ (scT)}} \\ &+ \underbrace{\overbrace{I \sum_{j=1}^J (\bar{y}_{.j} - \bar{y}_{..})^2}_{g.l. = J-1}}_{\text{Suma de cuadrados Explicada por } B\beta \text{ (scB)}} \\ &+ \underbrace{\overbrace{e_{ij}^2}_{g.l. = (I-1)(J-1)}}_{\text{Suma de cuadrados Residual (scR)}} \end{aligned}$$

esto es,

$$\sum_{i=1}^I \sum_{j=1}^J (y_{ij} - \bar{y}_{..})^2 = J \sum_{i=1}^I \hat{\alpha}_i^2 + I \sum_{j=1}^J \hat{\beta}_j^2 + \sum_{i=1}^I \sum_{j=1}^J e_{ij}^2, \quad (5.13)$$

escrito de otra forma

$$scG = scT\alpha + scB\beta + scR$$

de donde se deduce la siguiente tabla ANOVA

| CUADRO DEL ANÁLISIS DE LA VARIANZA – MODELO TRATAMIENTO-BLOQUE – | | | |
|---|---|------------------|--|
| Fuente de Variación | Suma de Cuadrados | g.l. | scm |
| Tratamientos | $scT = J \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$ | $I - 1$ | $\frac{scmT}{I - 1} = \frac{scT}{I - 1}$ |
| Bloques | $scB = I \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2$ | $J - 1$ | $\frac{scmB}{J - 1} = \frac{scB}{J - 1}$ |
| Residual | $scR = \sum_i \sum_j (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$ | $(I - 1)(J - 1)$ | $\frac{scmR}{(I - 1)(J - 1)} = \frac{scR}{(I - 1)(J - 1)}$ |
| Global | $scG = \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2$ | $IJ - 1$ | $\frac{scmG}{IJ - 1} = \frac{scG}{IJ - 1}$ |
| Rechazar $H_0^{(1)} : \alpha_1 = \alpha_2 = \dots = \alpha_I$, según $p = P\left(\frac{scmT}{scmR} < F_{I-1, (I-1)(J-1)}\right)$ | | | |
| Rechazar $H_0^{(2)} : \beta_1 = \beta_2 = \dots = \beta_J$, según $p = P\left(\frac{scmB}{scmR} < F_{J-1, (I-1)(J-1)}\right)$ | | | |

Tabla 5.1: Análisis de la varianza para un diseño en bloques completamente aleatorizados.

De esta tabla ANOVA se deducen dos contrastes:

✚ Si $H_0^{(\alpha)}$ es cierto, el factor-tratamiento no influye, se verifica que

(5.14)

$$\frac{SCMT\alpha}{\sigma^2} \sim \chi_{I-1}^2 \Rightarrow$$

$$F_\alpha = \frac{SCMT\alpha}{SCMR} = \frac{\frac{SCT\alpha}{I-1}}{\frac{SCR}{(I-1)(J-1)}} \sim F_{(I-1), (I-1)(J-1)},$$

se rechaza $H_0^{(\alpha)}$ al nivel de significación α si $\hat{F}_\alpha = \frac{scmT}{scmR} > F_{(I-1), (I-1)(J-1)}(1 - \alpha)$

⚡ Si $H_0^{(\beta)}$ es cierto, el factor-bloque no influye, se verifica que

$$\frac{SCMB\beta}{\sigma^2} \sim \chi_{J-1}^2 \Rightarrow$$

$$F_\beta = \frac{SCMB\beta}{SCMR} \sim F_{(J-1), (I-1)(J-1)}, \quad (5.15)$$

se rechaza $H_0^{(\beta)}$ al nivel de significación α si

$$\hat{F}_\beta = (scmB\beta / scmR) > F_{(J-1), (I-1)(J-1)}(1 - \alpha)$$

Comentarios.

- ⚡ La eficacia de este diseño depende de los efectos de los bloques. Si éstos pequeños, es más eficaz el diseño completamente aleatorio ya que el error en la comparación de tratamientos tiene menos grados de libertad. Sin embargo, si los bloques influyen es mucho mejor y más eficaz este modelo, ya que disminuye la variabilidad no explicada. Por ello, es mejor estudiar primero el modelo de bloques completamente aleatorizados y, si los bloques no influyen, se pasa fácilmente al modelo de un factor sumando en la tabla ANOVA la fila del factor bloque con la de la variabilidad residual.
- ⚡ Se define el Coeficiente de Determinación como:

$$R^2 = \frac{scT\alpha + scB\beta}{scG} = R^2(T\alpha) + R^2(B\beta)$$

siendo $R^2(T\alpha)$ y $R^2(B\beta)$ los coeficientes de determinación parciales asociados al tratamiento y al factor-bloque, respectivamente. Representan el tanto por ciento de la variabilidad total explicada por los tratamientos y los bloques.

- ⚡ El tratamiento estadístico expuesto para el modelo de diseño de experimentos completamente aleatorizado con un factor tratamiento y un factor bloque es exactamente igual que el diseño de experimentos con dos factores tratando la interacción.
- ⚡ Si de la tabla ANOVA del modelo de diseño de experimentos completamente aleatorizado se deduce que existen diferencias entre los tratamientos, es decir, las diferencias $(\alpha_i - \alpha_k)$ se estiman por

$$\hat{\alpha}_i - \hat{\alpha}_k = \bar{Y}_{i.} - \bar{Y}_{k.}, \quad i, k = 1, \dots, I.$$

Se pueden obtener intervalos de confianza de $\alpha_i - \alpha_k$ a partir de la distrib

$$\frac{(\bar{Y}_i - \bar{Y}_k) - (\alpha_i - \alpha_k)}{\hat{\sigma}_R \sqrt{\frac{2}{J}}} \sim t_{(I-1)(J-1)}, \quad (5.16)$$

de forma análoga se puede hacer para las diferencias $\beta_j - \beta_l$.

- ⚡ La eficacia de este diseño depende de los efectos de los bloques. Si éstos son pequeños, es más eficaz el diseño completamente aleatorio ya que el error en la comparación de tratamientos (ver (5.13)) tiene menos grados de libertad. Sin embargo si los bloques influyen es mucho mejor y más eficaz este modelo ya que disminuye la variabilidad no explicada.

Por ello es mejor estudiar primero el modelo de bloques aleatorizados. Si los bloques no influyen, se pasa fácilmente al modelo de un solo factor sumando la fila del factor bloque con la de la variabilidad residual. Así, no existe una discusión acerca de si se puede pasar de un modelo a otro ya que la diferencia importante entre los dos modelos es que en un diseño completamente aleatorizado, los tratamientos y, equivalentemente, los niveles de los factores de tratamiento son asignados aleatoriamente a las unidades experimentales. En el contrario, en un diseño en bloques, aunque las observaciones son también sobre todas las combinaciones de tratamientos y bloques, sólo los niveles de los factores de tratamiento son asignados aleatoriamente a las unidades experimentales. Hay que tener en cuenta que la división de las unidades experimentales para formar bloques es determinista. Esto ha hecho que exista una fuerte controversia sobre si es apropiado o no contrastar la igualdad de los efectos de los bloques. Al final, en el diseño que se está analizando los bloques representan una fuente de variación "nuisance", esto es, no existe interés alguno en la posible significación del efecto bloque. Más aún, es poco probable que sea factible utilizar los bloques en una hipotética réplica del experimento.

Por todo ello, en la práctica

"Bloquear sin ser necesario conduce a pruebas de hipótesis menos potentes y a intervalos de confianza más amplios que aquellos que se obtendrían mediante un diseño completamente aleatorizado".

- ⚡ Si se tiene un diseño de experimentos con dos factores de tratamiento (T) y J bloques (B) entre los que no existe interacción, el tratamiento estadístico y desarrollo que se estudia en esta sección es válido para este modelo. En este caso el contraste de dos hipótesis acerca de la influencia de los efectos de los niveles de los factores de tratamiento $T\alpha$ y $T\beta$ dados en (5.2) y (5.3) son de gran interés, contrastes que se realizan a partir de la tabla ANOVA utilizando (5.13) y (5.14). En cualquier caso, al comparar el modelo de bloques completamente aleatorizado y el modelo de dos factores con interacción tienen un desarrollo matemático análogo, su planteamiento y resultados son diferentes.

5.2.4 Análisis de residuos.

Como en cualquier modelo estadístico hay que contrastar que se verifican las hipótesis del modelo. Esto se hace, básicamente, por medio del análisis de los residuos. Todo lo estudiado sobre este particular en el modelo de un solo factor (diseño completamente aleatorizado) sigue siendo válido para este modelo. Se contrastarán las hipótesis de:

- ✚ Normalidad de los residuos.
- ✚ Homocedasticidad: la varianza en los diferentes niveles de cada uno de los factores es constante.
- ✚ Independencia de los residuos.
- ✚ Homogeneidad de los datos, todos provienen de la misma distribución y no atípicos.
- ✚ No existe interacción entre los dos factores. El concepto de interacción se con mayor detalle en la sección siguiente. Intuitivamente y basándose en ejemplo 5.1. que no exista interacción entre el factor tratamiento $T\alpha$ y el bloque $B\beta$ significa que la velocidad de impresión de una determinada imagen es mayor (o menor) de la media global independientemente de la foto que se imprime (del bloque).

Si existe interacción entre $T\alpha$ y $B\beta$, el modelo de bloques completamente aleatorizado no es adecuado y hay que tratar el factor bloque como un factor de tratamiento ($T\beta$). Se tiene entonces un diseño de experimentos con dos factores (tratamiento) y el modelo matemático es

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, I; j = 1, \dots, J \quad (5.2C)$$

el parámetro $(\alpha\beta)_{ij}$ representa la interacción del nivel i del factor $T\alpha$ con el nivel j del factor $T\beta$. Este modelo se estudia en la sección 4 de este capítulo.

5.2.5 Análisis de un caso.

Se desarrolla el problema presentado en el Ejemplo 5.1. cuyo enunciado más concreto es el siguiente,

Ejemplo 5.1.b.

“Una empresa fotográfica tiene que realizar una compra de impresoras de gran calidad que se van a utilizar en imprimir fotografías digitales. La empresa tiene ofertas de $I = 5$ marcas de impresoras de similares características y precio. Para la empresa fotográfica es muy importante la “*velocidad de impresión*” y, por este motivo, está interesada en saber si las 5 impresoras ofertadas tienen la misma velocidad o hay una que es más rápida. Para responder a esta pregunta decide hacer un experimento que consiste en elegir una única muestra de $J = 4$ fotos e imprimirlas en las 5 impresoras.

Los resultados del experimento se recogen en la tabla adjunta”

| | Foto A | Foto B | Foto C | Foto D |
|--|--------|--------|--------|--------|

| | | | | |
|-------------|----|----|----|----|
| Impresora 1 | 89 | 88 | 97 | 94 |
| Impresora 2 | 84 | 77 | 92 | 79 |
| Impresora 3 | 81 | 87 | 87 | 85 |
| Impresora 4 | 87 | 92 | 89 | 84 |
| Impresora 5 | 79 | 81 | 80 | 88 |

Solución.

Estimación de los parámetros.

Se obtienen las siguientes estimaciones

Estimaciones

| | Foto A | Foto B | Foto C | Foto D | $\bar{y}_{.j}$ | $\hat{\alpha}_j$ |
|-----------------|--------|--------|--------|--------|---------------------|------------------|
| Impresora 1 | 89 | 88 | 97 | 94 | 92 | 6 |
| Impresora 2 | 84 | 77 | 92 | 79 | 83 | -3 |
| Impresora 3 | 81 | 87 | 87 | 85 | 85 | -1 |
| Impresora 4 | 87 | 92 | 89 | 84 | 88 | 2 |
| Impresora 5 | 79 | 81 | 80 | 88 | 82 | -4 |
| $\bar{y}_{.j}$ | 84 | 85 | 89 | 86 | | |
| $\hat{\beta}_j$ | -2 | -1 | 3 | 0 | $\bar{y}_{..} = 86$ | |

Las predicciones y residuos son :

Predicciones

| | F. A | F. B | F. C | F. D |
|-----|------|------|------|------|
| I.1 | 90 | 91 | 95 | 92 |
| I.2 | 81 | 82 | 86 | 83 |
| I.3 | 83 | 84 | 88 | 85 |
| I.4 | 86 | 87 | 91 | 88 |
| I.5 | 80 | 81 | 85 | 82 |

Residuos

| | F. A | F. B | F. C | F. D |
|-----|------|------|------|------|
| I.1 | -1 | -3 | 2 | 2 |
| I.2 | 3 | -5 | 6 | -4 |
| I.3 | -2 | 3 | -1 | 0 |
| I.4 | 1 | 5 | -2 | -4 |
| I.5 | -1 | 0 | -5 | 6 |

La varianza residual es

$$\hat{s}_R^2 = \frac{1}{(I-1)(J-1)} \sum_{i=1}^I \sum_{j=1}^J e_{ij}^2 = \frac{1}{12} 226 = 18'83 \Rightarrow \hat{s}_R = 4'34.$$

Intervalos de confianza.

Intervalos de confianza al 90% para los parámetros del modelo son:

Para σ^2 ,

$$\begin{aligned} \chi_{12}^2 (0'05) &= 5'23 \leq \frac{12 \cdot 18'83}{\sigma^2} \leq \chi_{12}^2 (0'95) = 21'03 \\ \Rightarrow 0'68 &\leq \sigma^2 \leq 21'03 \end{aligned}$$

Para μ (tiempo medio global)

$$\begin{aligned} \frac{\bar{y}_{..} - \mu}{\hat{s}_R} \sqrt{20} &\sim t_{12} \Rightarrow \\ \mu &\in 86 \mp \frac{4'34}{\sqrt{20}} t_{12} (0'95) = 86 \mp 0'97 \cdot 1'78 = 86 \mp 1'27. \end{aligned}$$

Para $\mu_b = \mu + \alpha_3$ (tiempo medio de imprimir de la impresora 3)

$$\begin{aligned} \frac{\bar{y}_{3.} - \mu_{13}}{\hat{s}_R} \sqrt{4} &\sim t_{12} \Rightarrow \\ \mu_{13} &\in 85 \mp \frac{4'34}{\sqrt{4}} t_{12} (0'95) = 85 \mp 3'86. \end{aligned}$$

Para α_3 (el efecto de la impresora 3)

$$\begin{aligned} \frac{\hat{\alpha}_3 - \alpha_3}{\hat{s}_R} \sqrt{\frac{20}{5-1}} &\sim t_{12} \Rightarrow \\ \alpha_3 &\in -1 \mp \frac{4'34}{\sqrt{5}} t_{12} (0'95) = 85 \mp 3'45. \end{aligned}$$

Para $\theta_{31} = \alpha_3 - \alpha_1 = \mu_{31} - \mu_{11}$ (diferencia entre la impresora 3 y la 1)

$$\frac{(\bar{y}_{3.} - \bar{y}_{1.}) - \theta_{31}}{\hat{s}_R} \sqrt{\frac{4}{2}} \sim t_{12} \Rightarrow$$

$$\theta_{31} \alpha \in -7 \mp \frac{4'34}{\sqrt{2}} t_{12} (0'95) = 85 \mp 5'46.$$

Trabajando al 90% se obtienen los siguientes grupos homogéneos de impresoras:

- Imp. 5 - Imp. 2 - Imp. 3
- Imp. 2 - Imp. 3 - Imp. 4
- Imp. 4 - Imp. 1

Tabla ANOVA.

| Fuentes de variación | Suma de cuadrados | Grados de libertad | Varianzas |
|----------------------|-------------------|--------------------|-----------|
| scT α | 264 | 4 | 66'00 |
| scB β | 70 | 3 | 23'33 |
| scExplicada | 334 | 7 | 47'71 |
| scR | 226 | 12 | 18'83 |
| scG | 560 | 19 | 29'47 |
| | | | |
| | | | |

A partir de esta tabla se obtienen los siguientes contrastes

$$H_0^{(\alpha)} \equiv \alpha_1 = \alpha_2 = \dots = \alpha_5 = 0 \text{ (el "tipo de impresora" no influye)}$$

$$\hat{F}_\alpha = \frac{scmT\alpha}{\hat{s}_R^2} = \frac{66}{18'83} = 3'504 \sim F_{I-1, (I-1)(J-1)} = F_{4,12}?$$

p-valor= 0'0407. Se rechaza $H_0^{(\alpha)}$ para $\alpha = 0'05$, pero el p-valor ofrece dudas acerca de la influencia del factor "tipo de impresora".

$$H_0^{(\beta)} \equiv \beta_1 = \beta_2 = \dots = \beta_4 = 0 \text{ (el "tipo de foto" no influye)}$$

$$\hat{F}_\beta = \frac{scmB}{\hat{s}_R^2} = \frac{23'33}{18'83} = 1'239 \sim F_{J-1, (I-1)(J-1)} = F_{3,12}?$$

p-valor= 0'3387. Se acepta $H_0^{(\beta)}$ para cualquier α razonable y el factor

bloque "*tipo de foto*" no influye, por tanto, no convenía bloquear.

Coeficientes de determinación.

Los coeficientes de determinación parciales son

$$R^2(\alpha) = \frac{264}{560} = 0'4714,$$

el factor "*tipo de impresora*" explica el 4714% de variabilidad.

$$R^2(\beta) = \frac{70}{560} = 0'1250,$$

el factor "*tipo de foto*" explica el 1250% de variabilidad.

El coeficiente de determinación total es la suma de los parciales,

$$R^2(\text{total}) = \frac{334}{560} = 0'5964,$$

el modelo explica el 5964% de variabilidad.

Un modelo alternativo.

En conclusión, parece razonable aceptar la influencia del factor-tratamiento "*tipo de impresora*" y la no influencia del factor bloque "*tipo de foto*". Se puede pasar fácilmente al modelo completamente aleatorizado, la tabla ANOVA con un solo factor que se obtiene de la anterior sumando las filas de $scB\beta$ y scR , obteniendo

Tabla ANOVA

| Resultados con un solo factor | | | |
|-------------------------------|-------------------|--------------------|-----------|
| Fuentes de Variación | Suma de Cuadrados | Grados de libertad | Varianzas |
| $scT\alpha$ | 264 | 4 | 66'00 |
| scR | 296 | 15 | 19'73 |
| scG | 560 | 19 | 29'47 |

Ahora el contraste $H_0^{(\alpha)} \equiv \alpha_1 = \alpha_2 = \dots = \alpha_5 = 0$ (el "*tipo de impresora*" no influye) es

$$\hat{F}_T = \frac{scmT\alpha}{s_R^2} = \frac{66'00}{19'73} = 3'34 \sim F_{I-1, I(J-1)} = F_{4,16}?$$

$p\text{-valor} = 0.038$. Se rechaza $H_0^{(\alpha)}$ para $\alpha > 0.038$. Se tiene mayor seguridad acerca de la influencia del factor "*tipo de impresora*" que en el modelo anterior.

[\[Siguiendo\]](#) [\[Anterior\]](#) [\[Arriba\]](#)

[[Siguiente](#)] [[Anterior](#)] [[Arriba](#)]

5.3 La interacción entre factores.

Se considera un diseño completamente aleatorizado con dos factores tratamiento $T\alpha$ y $T\beta$ cuyos niveles se cruzan. El factor $T\alpha$ tiene I niveles y el factor $T\beta$ tiene J niveles. Por tanto hay un total de IJ tratamientos distintos codificados por

$$11, 12, \dots, 1J, 21, 22, \dots, 2J, \dots, IJ.$$

Si para cada uno de los $n = IJ$ tratamientos se tiene una observación, el modelo matemático del diseño tiene la forma (dada en (5.20))

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, I; j = 1, \dots, J,$$

En este modelo el número de parámetros a estimar es

$$1 + (I - 1) + (J - 1) + (I - 1)(J - 1) + 1 = IJ + 1 > n,$$

mayor que el número de observaciones y que el número de grados de libertad de los residuos. Por tanto no es posible estimar el modelo. Para resolver este problema hay las siguientes alternativas:

[1] Aumentar el número de observaciones, se replica el diseño K veces y se tiene KIJ observaciones. El modelo con dos factores replicado se estudia en la sección siguiente.

[2] Disminuir el número de parámetros, suponiendo que algunas interacciones son nulas o que son funciones de unos pocos parámetros. Por ejemplo, suponer que las interacciones tienen una forma funcional multiplicativa (hipótesis de Tuckey)

$$(\alpha\beta)_{ij} = \gamma\alpha_i\beta_j \quad i = 1, \dots, I; j = 1, \dots, J$$

con esta hipótesis solo hay un parámetro adicional (γ) respecto al modelo sin interacción. El número de parámetros del modelo es $I + J + 1 < IJ = n$.

[2] En la práctica, cuando se ajustan modelos complejos donde intervienen muchos factores, se utiliza una combinación de las dos estrategias anteriores. Por ejemplo, se supone que las interacciones de orden superior a dos son nulas y se replica el experimento.

Test de Tuckey.

Para contrastar que no existe interacción entre el factor tratamiento y el factor bloque en el diseño por bloques completamente aleatorizado se puede utilizar el test de Tuckey que permite contrastar la hipótesis de que las interacciones son de la forma $(\alpha\beta)_{ij} = \gamma\alpha_i\beta_j$. Esto es, se quiere contrastar

$$H_0: \gamma = 0 \text{ frente a la alternativa } H_1: \gamma \neq 0$$

Al ajustar el modelo (5.1) si $\gamma \neq 0$ se verifica que

$$E(e_{ij}) = y_{ij} - \hat{y}_{ij} = \gamma\hat{\alpha}_i\hat{\beta}_j,$$

de esta expresión se deduce:

– Si α_i y β_j son valores grandes y del mismo signo, los residuos son altos, y las predicciones $\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$ están muy por arriba o por debajo de la media.

– Si α_i y β_j son valores pequeños y/o de distinto signo, los residuos son pequeños y las predicciones están en torno a la media.

En base a esto

“el contraste de Tuckey se basa en dibujar la gráfica de los residuos e_{ij} frente a las predicciones \hat{y}_{ij} , si existe una interacción de tipo multiplicativo la nube de puntos tiene forma parabólica”.

El contraste de Tuckey también se puede hacer de forma analítica. Se dibuja la nube de puntos de los residuos e_{ij} frente a la variable $x_{ij} = \hat{\alpha}_i\hat{\beta}_j$, y se ajusta una recta. Entonces contrastar la hipótesis $H_0: \gamma = 0$ es equivalente a contrastar que el coeficiente de correlación entre x_{ij} y e_{ij} es cero. Si no se acepta H_0 el estimador de γ es la pendiente de la recta ajustada.

En las Figura 5.1, 5.2. y 5.3. se presentan diferentes gráficos de e_{ij} frente a \hat{y}_{ij} , con diferentes posibilidades sobre la interacción multiplicativa.

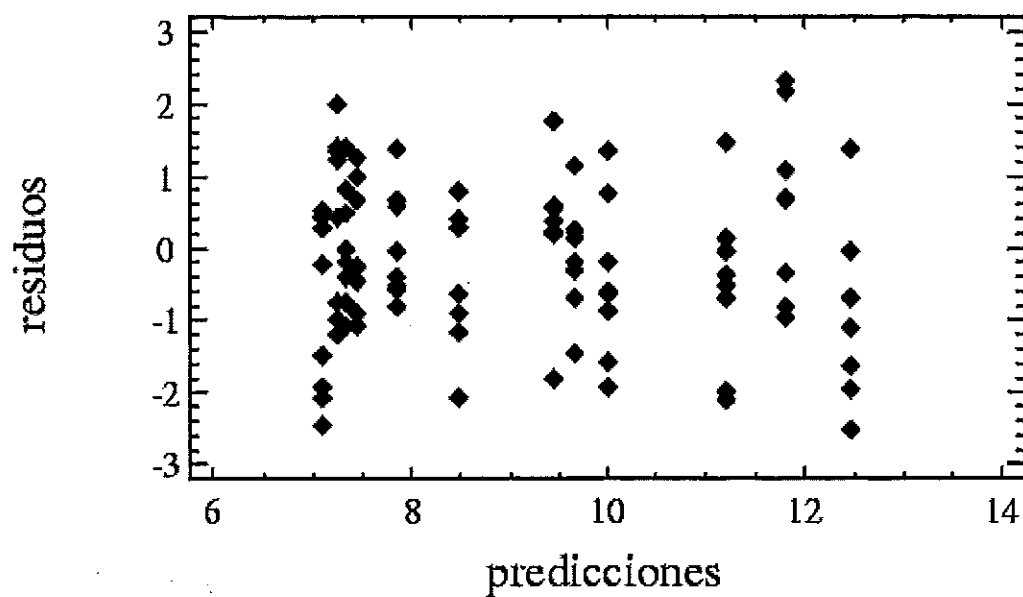


Figura 5.1. No hay indicios de que exista interacción multiplicativa.

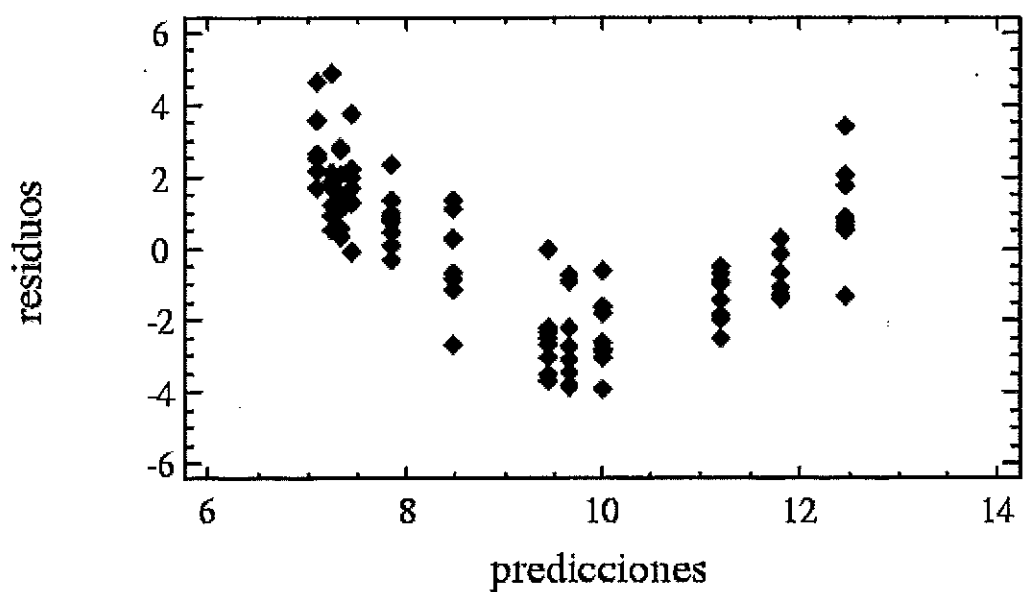


Figura 5.2. Existe interacción entre los factores.

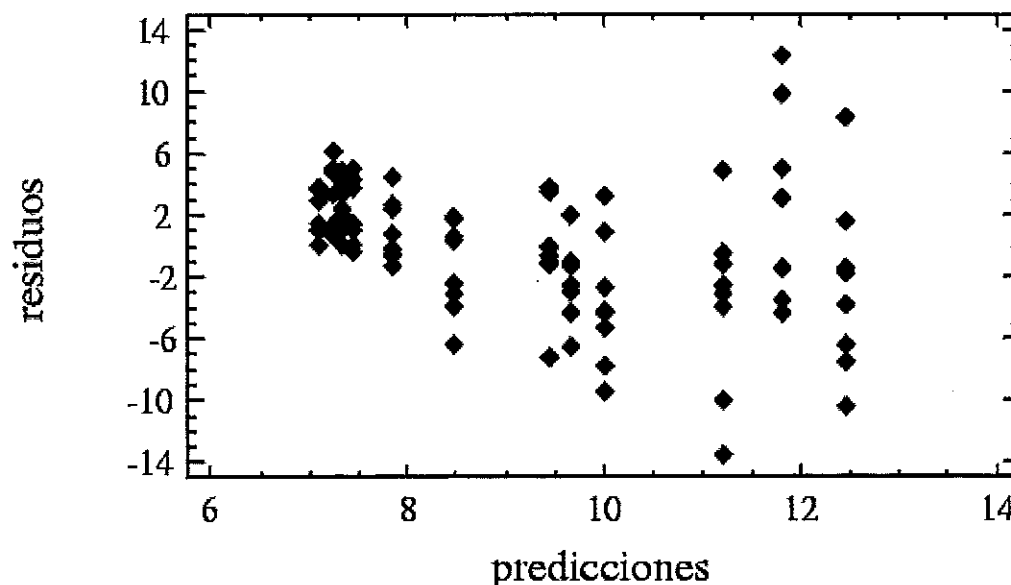


Figura 5.3. Existe interacción y heterocedasticidad.

El significado de la interacción

Considérese el modelo con dos factores tratamiento $T\alpha$ y $T\beta$ con I y J niveles, respectivamente. El diseño completo se ha replicado K veces, esto es, para cada tratamiento (casilla) ij se tienen K observaciones. Se denota y_{ijk} a la k -ésima observación del tratamiento ij , con $k = 1, 2, \dots, K$. El tamaño del experimento es $n = IJK$, el modelo asociado es

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}, \quad i = 1, \dots, I; j = 1, \dots, J, k = 1, 2, \dots, K$$

Entonces la falta de interacción entre los factores $T\alpha$ y $T\beta$ se interpreta como sigue:

"Se dice que no existe interacción entre los dos factores tratamiento $T\alpha$ y $T\beta$ cuando sus efectos sobre la respuesta son aditivos. En otros términos, la diferencia de las respuestas medias teóricas en dos niveles cualesquiera de un factor es constante en todos los niveles del otro factor y viceversa:

$$\text{Para todo } i, s = 1, \dots, I, \quad \mu_{ij} - \mu_{sj} = \mu_{it} - \mu_{st}, \quad \text{con } j, t = 1, \dots, J.$$

Un gráfico ilustrativo de la posible existencia de interacción es el denominado gráfico de interacción. Para construir este gráfico se marcan en el eje de abscisas los niveles de uno de los dos factores tratamiento, por ejemplo el A , y se dibuja la nube de puntos

$$\left\{ (i, \bar{y}_{ij.}) , \text{ para } i = 1, \dots, I, j = 1, \dots, J, \text{ siendo } \bar{y}_{ij.} = \frac{1}{K} \sum_{k=1}^K y_{ijk} \right\},$$

uniendo a continuación con segmentos las medias muestrales $\bar{y}_{ij.}$ con igual j .

Aunque los gráficos de interacción son muy intuitivos y útiles, pueden conducir a interpretaciones peligrosas debido a que en ellos *no se refleja el tamaño del error experimental*. Esto puede llevar a deducir del gráfico la existencia de interacción y, sin embargo, el error experimental ser lo suficientemente grande como para que el análisis de la varianza no detecte la interacción como significativa (y viceversa). Por tanto, se debe ser muy prudente con las conclusiones que se derivan de un gráfico de este tipo.

Frecuentemente el interés del diseño radica fundamentalmente en evaluar la contribución individual de cada factor tratamiento sobre la respuesta observada. A los efectos marginales de cada factor se les denomina **efectos principales**. Sin embargo, cuando el efecto interacción entre ambos factores es importante, puede ser imposible examinar por separado cada uno de los efectos principales.

Ejemplo 5.2.

En base a las calificaciones en pruebas escritas se evalúan los resultados obtenidos con dos métodos de enseñanza distintos (método 1 y método 2) y con tres profesores (profesores 1, 2 y 3). Las Figuras 5.4 y 5.5 muestran ocho posibles gráficos de interacción reflejando otros tantos resultados posibles para el experimento.

En las cuatro situaciones descritas en la Figura 5.4. la interacción no existe. Las líneas resultantes de unir las medias obtenidas con cada método son paralelas o coincidentes y, por ello, las diferencias (cuando existen) entre los niveles de cada factor son estables a través de los niveles del otro factor.

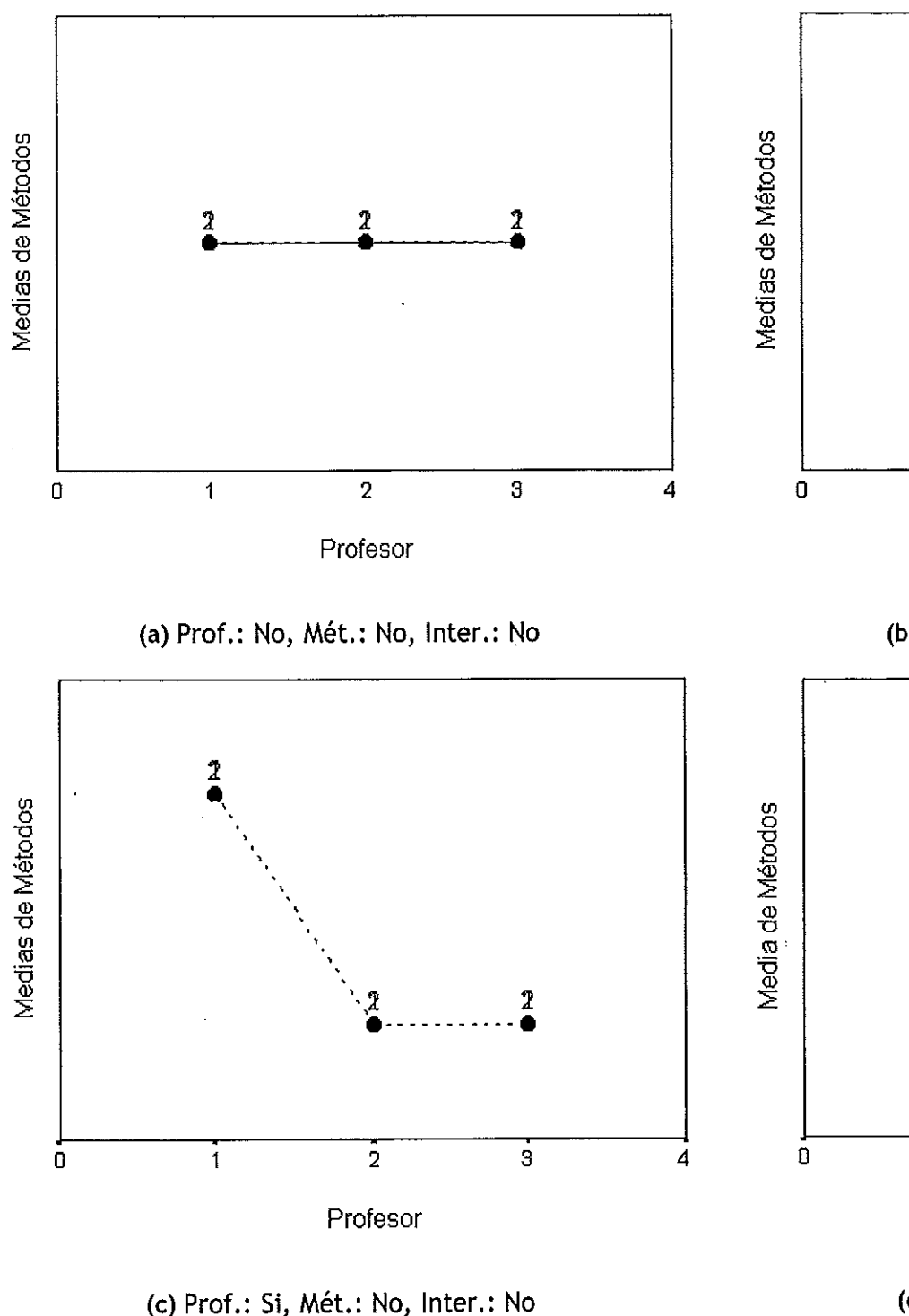
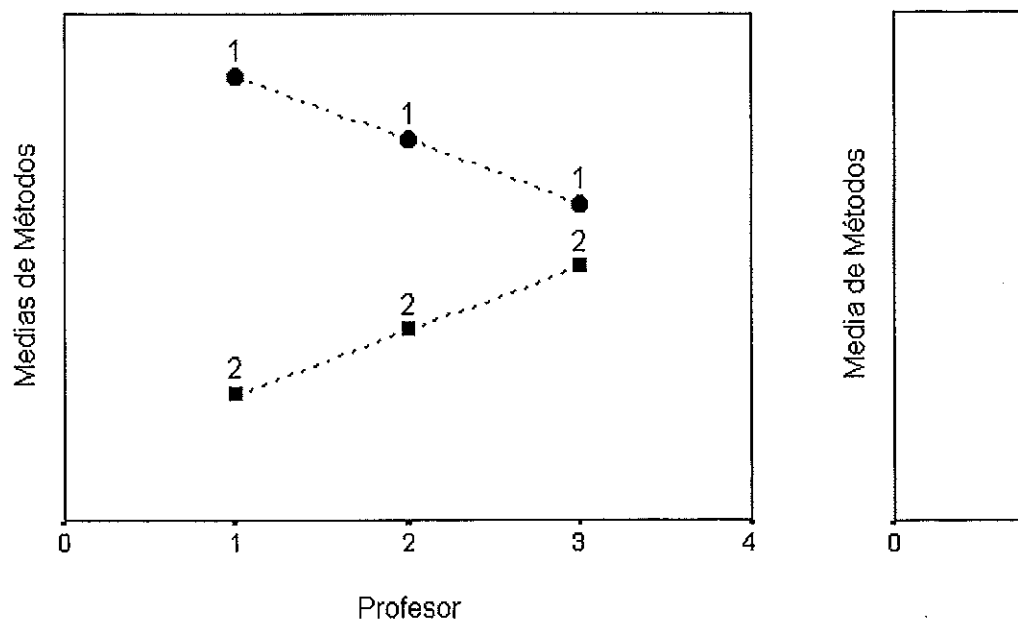


Figura 5.4.: Cuatro posibles configuraciones de los efectos sin presencia de interacción

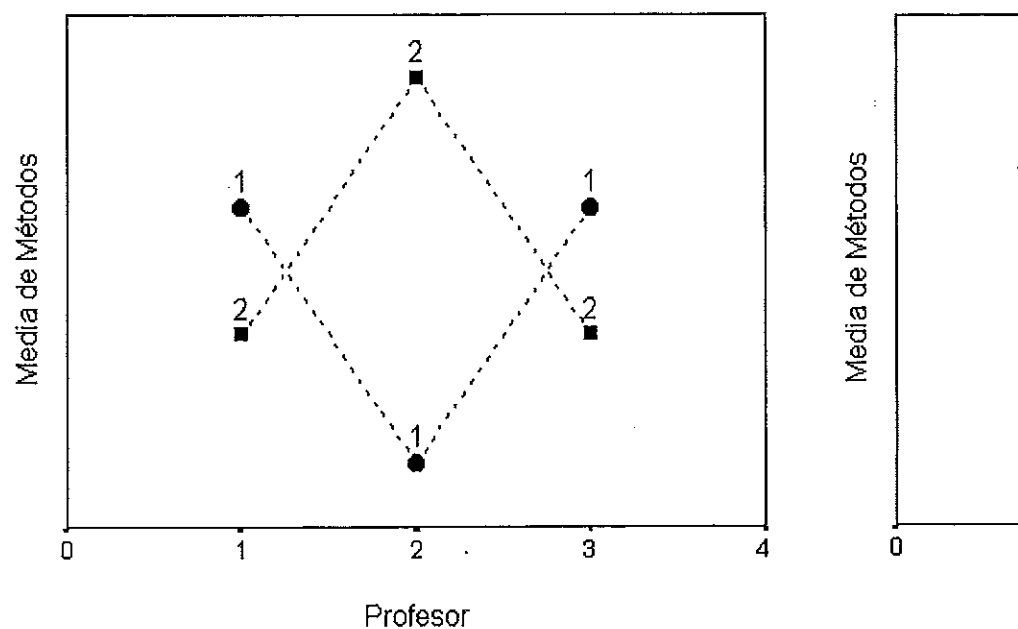
En la Figura 5.5. en todos los gráficos el efecto interacción es significativo. En el gráfico (a) todos los profesores obtienen las calificaciones más altas con el método de enseñanza 1; ahora bien, las diferencias son muy grandes en el caso del profesor 1 y muy pequeñas con el profesor 3. Es claro que existen diferencias entre los métodos de enseñanza pero, ¿existen diferencias entre

los profesores? (obsérvese que los promedios muestrales de todos ellos son idénticos). Por el contrario, en el gráfico (c), está claro que existen diferencias entre los profesores. Dos obtienen mejores puntuaciones con el método 1 y uno con el método 2. Sin embargo, si se hubiesen ignorado los métodos, las diferencias no existirían pues las puntuaciones medias de los tres son idénticas. Más aún, un ANOVA no detectaría diferencias significativas.



(a) Prof.: No?, Mét.: Si, Inter.: Si

(b)



(c) Prof.: No?, Mét.: No?, Inter.: Si

(d)

Figura 5.5: Cuatro posibles configuraciones de los efectos con presencia de interacción

De este ejemplo se sigue la siguiente conclusión:

“Si la interacción es significativa, será complejo examinar los efectos de cada factor tratamiento por separado. Por ejemplo, la presencia de interacción significativa podría encubrir diferencias reales entre los niveles de algún factor, de modo que no se detectasen diferencias significativas entre ellos en el análisis estadístico”.

[\[Siguiete\]](#) [\[Anterior\]](#) [\[Arriba\]](#)

[Anterior] [Arriba]

5.6 Fracciones factoriales. El cuadrado latino.

Los modelos de diseño de experimentos expuestos en las secciones previas son diseños *completos o equilibrados*. En estos diseños se obtienen pruebas cruzando los niveles de los factores de todas las formas posibles, por ello, en estos diseños los factores son *ortogonales*.

El concepto de ortogonalidad de factores.

En un diseño de experimentos los factores $T\alpha$, con I niveles, y $T\beta$, con J niveles, son *ortogonales* si en las pruebas del diseño en cada uno de los niveles i del factor $T\alpha$ aparecen en idénticas proporciones los J niveles del factor $T\beta$.

La propiedad de ortogonalidad permite separar los efectos de cada uno de los factores sobre la variable de interés.

Si los efectos simples de todos los factores estudiados en el diseño de experimentos son ortogonales, la estimación ($\hat{\alpha}_i$) del efecto del nivel i del factor $T\alpha$ se obtiene como la diferencia entre la media de los resultados obtenidos cuando el factor $T\alpha$ está al nivel i y la media general de todos los resultados.

$$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..}$$

Las estimaciones así obtenidas para los efectos de un factor no están afectadas por los efectos de los otros factores, lo que permite separar los efectos simples de todos los factores estudiados.

En los diseños equilibrados el número de pruebas que hay que realizar crece muy rápidamente con el número de factores, aún en el caso de que se supongan nulas las interacciones y no sea necesario replicar el diseño. En estas situaciones son de gran utilidad los diseños de experimentos denominados **fracciones factoriales**, que permitan estudiar la influencia de los factores sin necesidad de realizar todas las pruebas pero manteniendo la propiedad de ortogonalidad de los efectos a estudiar. Como ejemplo de este tipo de modelos se expone a continuación la fracción factorial denominada cuadrado latino.

5.6.1 El modelo de cuadrado latino.

En un diseño de experimentos completo de tres factores, todos ellos con K niveles, necesita K^3 observaciones, número elevado si K es grande. Un diseño más eficaz que solo utiliza K^2 observaciones para el mismo problema es el *cuadrado latino*. Este modelo se basa en aprovechar la simetría del experimento factorial seleccionando un conjunto de condiciones experimentales con la condición de que cada nivel de un factor aparezca una

vez con cada uno de los niveles de los otros factores. Por tanto, el diseño de cuadrado latino se puede utilizar si se verifican las siguientes condiciones:

1. Es un diseño de experimentos con tres factores.
2. Los tres factores tienen el mismo número de niveles: K .
3. No hay interacciones entre los tres factores.

El diseño en cuadrado latino está especialmente indicado para estudiar un factor-tratamiento con K niveles y con dos factores-bloque de K bloques cada uno. Este diseño se basa en el concepto de cuadrado latino que es el siguiente

"Un *cuadrado latino* $K \times K$ es una disposición de K letras en una matriz $K \times K$ de forma que todas las letras aparecen una vez en cada fila y una vez en cada columna.

Por ejemplo, un cuadrado latino 3×3 es el siguiente

| | | |
|---|---|---|
| A | B | C |
| B | C | A |
| C | A | B |

Tabla 5.5. Cuadrado latino 3×3 .

Un cuadrado latino es un **cuadrado latino estándar** cuando las letras de la primera fila y de la primera columna están dispuestas en orden alfabético.

Un cuadrado latino es un **cuadrado latino cíclico** si las letras de cada fila se generan cíclicamente de la anterior según el orden alfabético.

El cuadrado latino 3×3 de la Tabla 5.5 es estándar y cíclico.

Existe un único cuadrado latino 3×3 estándar, sin embargo hay cuatro cuadrados latinos 4×4 estándar que se presentan en la Tabla 5.6.

| Cuadro 1 | Cuadro 2 | Cuadro 3 | Cuadro 4 |
|----------|----------|----------|----------|
| A B C D | A B C D | A B C D | A B C D |
| B C D A | B A D C | B A D C | B D A C |
| C D A B | C D A B | C D B A | C A D B |
| D A B C | D C B A | D C A B | D C B A |

Tabla 5.6: Cuatro posibles cuadrados latinos 4×4 estándar.

"Un **diseño en cuadrado latino** es un diseño de un factor tratamiento con K niveles y K^2 unidades experimentales agrupadas en K bloques fila y K bloques columna, de forma que unidades experimentales de un mismo bloque fila son semejantes, unidades experimentales de un mismo bloque columna son semejantes y unidades experimentales de distintos bloques fila y distintos bloques columna son sustancialmente diferentes".

Para cualquier número de tratamientos K existe siempre al menos un diseño en cuadrado latino estándar cíclico.

Obsérvese que si en un diseño en cuadrado latino se ignora el bloque

columna se tiene un diseño en bloques completamente aleatorizado (el bloque fila es el factor bloque) y, análogamente, si se ignora el bloque fila se tiene un diseño en bloques completamente aleatorizado (el bloque columna es el factor bloque). Además se trata de un diseño equirreplicado: cada tratamiento aparece un mismo número K de veces en el diseño.

Modelo matemático.

Se tiene un diseño en cuadrado latino de dos factores bloque y un factor tratamiento, el primer factor bloque se denota por $B\alpha$ y se coloca en filas, el segundo factor bloque se denota por $B\beta$ y se coloca en columnas, el factor tratamiento se denota por $T\gamma$ y sus niveles se colocan según el cuadrado latino. Por tanto, el cuadrado latino *condiciona* el nivel de $T\gamma$ que se utiliza en la casilla ij (bloque i de $B\alpha$ y bloque j de $B\beta$) y este nivel no se elige.

La formulación matemática del modelo es la siguiente:

para cada $i = 1, \dots, K$, $j = 1, \dots, K$, (el índice $k \in \{1, \dots, K\}$ lo impone el diseño en cuadrado latino) se tiene

$$\underbrace{Y_{ij(k)}}_{\text{aleatorio}} = \overbrace{\mu + \alpha_i + \beta_j + \gamma_k}^{\text{determinista}} + \underbrace{\varepsilon_{ij(k)}}_{\text{aleatorio}}$$

con $\varepsilon_{ij(k)}$ v.a. independientes con distribución $N(0, \sigma^2)$,

donde,

* $Y_{ij(k)}$ es el resultado del bloque i -ésimo, $i = 1, \dots, K$ del factor bloque $B\alpha$ y del bloque j -ésimo, $j = 1, \dots, J$ del factor-bloque $B\beta$, y del nivel k -ésimo del factor $T\gamma$. Se denota la k entre paréntesis, para indicar que este índice no se elige sino que viene condicionado por el par ij .

* μ es el efecto global que mide el nivel medio de todos los resultados,

* α_i es el efecto (positivo o negativo) sobre la media global debido al bloque i de $B\alpha$. Se verifica que $\sum_{i=1}^I \alpha_i = 0$,

* β_j es el efecto (positivo o negativo) sobre la media global debido al bloque j de $B\beta$. Se verifica que $\sum_{j=1}^J \beta_j = 0$,

* γ_k es el efecto (positivo o negativo) sobre la media global debido al nivel k del factor $T\gamma$. Se verifica que $\sum_{k=1}^K \gamma_k = 0$,

* $\varepsilon_{ij(k)}$ es el error experimental, son variables aleatorias i.i.d. con

distribución $N(0, \sigma^2)$.

Estimación de los parámetros.

La técnica de mínimos cuadrados proporciona los siguientes estimadores:

$$\hat{\mu} = \bar{y}_{..} = \frac{1}{K^2} \sum_{i=1}^K \sum_{j=1}^K y_{ij(k)}$$

$$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..}, \quad i = 1, \dots, K, \quad \text{con} \quad \bar{y}_{i.} = \frac{1}{K} \sum_{j=1}^K y_{ij(k)},$$

$$\hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..}, \quad j = 1, \dots, K, \quad \text{con} \quad \bar{y}_{.j} = \frac{1}{K} \sum_{i=1}^K y_{ij(k)},$$

$$\hat{\gamma}_k = \bar{y}_{..(k)} - \bar{y}_{..}, \quad k = 1, \dots, K, \quad \text{con} \quad \bar{y}_{..(k)} = \frac{1}{K} \sum_{i=1}^K \sum_{j=1}^K y_{ij(k)}$$

Los residuos son

$$\begin{aligned} e_{ij(k)} &= y_{ij(k)} - \hat{y}_{ij(k)} \\ &= y_{ij(k)} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k), \quad i, j, k = 1, \dots, K \end{aligned}$$

que verifican las siguientes restricciones

$$\sum_{i=1}^K e_{ij(k)} = 0, \quad \text{con} \quad j = 1, \dots, K,$$

$$\sum_{j=1}^K e_{ij(k)} = 0, \quad \text{con} \quad i = 1, \dots, K,$$

$$\sum_{i=1}^K \sum_{j=1}^K e_{ij(k)} = 0, \quad \text{con} \quad k = 1, \dots, K,$$

en total hay $3K - 2$ restricciones, y los residuos tienen $K^2 - (3K - 2) = (K - 1)(K - 2)$ grados de libertad.

Tabla ANOVA.

De la descomposición de la variabilidad se obtiene la tabla ANOVA (Tabla 5.7.) de donde se deducen los siguientes contrastes:

[1] Si la hipótesis nula $H_0^{(\gamma)}: \gamma_1 = \gamma_2 = \dots = \gamma_K = 0$, (el factor F^γ no influye, el más importante porque es el factor-tratamiento en el que se está interesado) es cierta, se verifica que

$$\frac{SCMT\gamma}{\sigma^2} \sim \chi^2_{(K-1)} \Rightarrow F_\gamma = \frac{SCMT\gamma}{SCMR} \sim F_{(K-1), (K-1)(K-2)} \quad (1.36)$$

se rechaza $H_0^{(\gamma)}$ al nivel de significación α si $\hat{F}_\gamma = scmT\gamma / scmR > \hat{F}_{(K-1), (K-1)(K-2)}(1 - \alpha)$.

[2] Aunque de menor interés también se pueden hacer contrastes acerca de la influencia de los bloques fila y columna para saber si ha sido conveniente bloquear o no.

Si la hipótesis nula $H_0^{(\alpha)} : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, (el bloque fila no influye) es cierta, se verifica que

$$\frac{SCMB\alpha}{\sigma^2} \sim \chi^2_{(K-1)} \Rightarrow F_\alpha = \frac{SCMB\alpha}{SCMR} \sim F_{(K-1), (K-1)(K-2)}, \quad (1.37)$$

se rechaza $H_0^{(\alpha)}$ al nivel de significación α si $\hat{F}_\alpha = scmB\alpha / scmR > \hat{F}_{(K-1), (K-1)(K-2)}(1 - \alpha)$.

[3] Si la hipótesis nula $H_0^{(\beta)} : \beta_1 = \beta_2 = \dots = \beta_k = 0$, (el factor columna no influye) es cierta, se verifica que

$$\frac{SCMB\beta}{\sigma^2} \sim \chi^2_{(K-1)} \Rightarrow F_\beta = \frac{SCMB\beta}{SCMR} \sim F_{(K-1), (K-1)(K-2)}, \quad (1.38)$$

se rechaza $H_0^{(\beta)}$ al nivel de significación α si $\hat{F}_\beta = scmB\beta / scmR > \hat{F}_{(K-1), (K-1)(K-2)}(1 - \alpha)$.

| CUADRO DEL ANÁLISIS DE LA VARIANZA – MODELO CUADRADO LATINO – | | | | |
|--|------------------------------------|---------|-----------------------------------|---|
| Fuente de Variación | Suma de Cuadrados | g.l. | scm | |
| Bloques (B α) | $scBF = K \sum_i \hat{\alpha}_i^2$ | $K - 1$ | $scmB\alpha = \frac{scBF}{K - 1}$ | $\hat{F}_{(K-1), (K-1)(K-2)}$ (scmB α) |
| Bloques (B β) | $scBC = K \sum_j \hat{\beta}_j^2$ | $K - 1$ | $scmB\beta = \frac{scBC}{K - 1}$ | $\hat{F}_{(K-1), (K-1)(K-2)}$ (scmB β) |

| | | | | |
|--|---|--------------|-----------------------------------|------------------------------------|
| Factor T (T γ) | $scTL = K \sum_k \hat{\gamma}_{k2}$ | $K - 1$ | $scmT\gamma = \frac{scTL}{K - 1}$ | $\hat{\gamma}$ (scmB γ) |
| Residual | $scR = \sum_i \sum_j e_{ij(k)}^2$ | $(K-1)(K-2)$ | $scmR = \frac{scR}{(K-1)(K-2)}$ | $\frac{s}{(K-1)}$ |
| Global | $scG = \sum_i \sum_j (y_{ij(k)} - \bar{y}_{...})^2$ | $K^2 - 1$ | $scmG = \frac{scG}{K^2 - 1}$ | |
| Rechazar $H_0^{(\alpha)} : \alpha_1 = \alpha_2 = \dots = \alpha_K$, según $p = P(\hat{F}_\alpha < F_{K-1, (K-1)(K-2)})$ | | | | |
| Rechazar $H_0^{(\beta)} : \beta_1 = \beta_2 = \dots = \beta_K$, según $p = P(\hat{F}_\beta < F_{K-1, (K-1)(K-2)})$ | | | | |
| Rechazar $H_0^{(\gamma)} : \gamma_1 = \gamma_2 = \dots = \gamma_K$, según $p = P(\hat{F}_\gamma < F_{K-1, (K-1)(K-2)})$ | | | | |

Tabla 5.6. Cuadro del análisis de la varianza para un diseño de cuadrado latino.

Extensiones de los modelos de diseños experimentales.

Siguiendo la metodología expuesta en los diseños estudiados es fácil generalizar el diseño de cuadrado latino y tienen interés los siguientes modelos:

Cuadrado latino replicado. Si se replica el modelo del cuadrado latino, aún manteniendo las mismas condiciones de experimentación, es posible que exista cierta heterogeneidad entre las réplicas por lo que es conveniente considerar las réplicas como bloques. El modelo matemático de este diseño es:

$$y_{ij(k)r} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_r + \varepsilon_{ij(k)r}$$

donde δ_r ($r = 1, \dots, R$) es el efecto réplica que se estimará por la diferencia entre la media de cada réplica completa y la media general.

Cuadrado greco-latino. Si se aumenta el número de factores-bloque, la extensión del cuadrado latino es el greco-latino, que permite con K^2 observaciones estudiar cuatro factores de K niveles sin interacciones (un factor-tratamiento y tres factores bloque), si se utilizase el diseño completo es necesario utilizar K^4 observaciones. En el diseño en cuadrado greco-latino se superponen dos cuadrados latinos, resultando el siguiente modelo matemático:

$$y_{ij(kh)} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_h + \varepsilon_{ij(kh)}$$

El inconveniente de este modelo es que su utilización es muy restrictiva. Además pueden no existir cuadrados latinos de determinadas

condiciones.

5.6.2 Análisis de un caso

En este apartado se desarrolla un problema de diseño de experimentos de cuadrado latino. El enunciado del problema es el siguiente:

Ejemplo 5.4.

"Se quiere estudiar la posible influencia de los "aditivos de combustible" (factor tratamiento, $T\gamma$) en la "reducción de óxidos de nitrógeno en las emisiones de los automóviles" (variable respuesta) controlando la influencia del "conductor" (factor-bloque $B\alpha$) y del "tipo de coche" (factor-bloque, $B\beta$).

Se consideran cuatro conductores: C1, C2, C3, C4. (efecto α)

Cuatro tipos de coche: Seat, Ford, Opel, Renault. (efecto β)

Cuatro aditivos de combustible: A1, A2, A3, A4. (efecto γ)

Los resultados del experimento diseñado según la técnica del cuadrado latino son los de la tabla adjunta, también se presenta el cuadrado latino utilizado. ¿Qué conclusiones se deducen del experimento?"

| | Seat | Ford | Opel | Renault |
|----|-------|-------|-------|---------|
| C1 | 21 A1 | 26 A2 | 20 A4 | 25 A3 |
| C2 | 23 A4 | 26 A3 | 20 A1 | 27 A2 |
| C3 | 15 A2 | 13 A4 | 16 A3 | 16 A1 |
| C4 | 17 A3 | 15 A1 | 20 A2 | 20 A4 |

| C. Latino | | | |
|-----------|---|---|---|
| 1 | 2 | 4 | 3 |
| 4 | 3 | 1 | 2 |
| 2 | 4 | 3 | 1 |
| 3 | 1 | 2 | 4 |

Solución.

Estimación de los parámetros. Se obtienen los siguientes estimadores:

Estimaciones

| $\bar{y}_{i.}$ | $\hat{\alpha}_i$ | $\bar{y}_{..i}$ | $\hat{\beta}_j$ | $\bar{y}_{..k}$ | $\hat{\gamma}_k$ |
|---------------------------------|------------------|-----------------|-----------------|-----------------|------------------|
| 23 | 3 | 19 | -1 | 18 | -2 |
| 24 | 4 | 20 | 0 | 22 | 2 |
| 15 | -5 | 19 | -1 | 21 | 1 |
| 18 | -2 | 22 | 2 | 19 | -1 |
| $\bar{y}_{..} = \hat{\mu} = 20$ | | | | | |

Los residuos del modelo son:

Residuos

| | Seat | Ford | Opel | Renault |
|----|-------|-------|-------|---------|
| C1 | 1 A1 | 1 A2 | -1 A4 | -1 A3 |
| C2 | 1 A4 | 1 A3 | -1 A1 | -1 A2 |
| C3 | -1 A2 | -1 A4 | 1 A3 | 1 A1 |
| C4 | -1 A3 | -1 A1 | 1 A2 | 1 A4 |

Tabla ANOVA. Utilizando las estimaciones y residuos obtenidos se obtiene la siguiente tabla ANOVA

Tabla ANOVA

| Fuentes de variación | Suma de cuadrados | Grados de libertad | Varianza | \hat{F} | p - valor |
|-------------------------|-------------------|--------------------|----------|-----------|-------------|
| Factor <i>conductor</i> | 216 | 3 | 72 | 27 | 0'0007 |
| Factor <i>coche</i> | 24 | 3 | 8 | 3 | 0'1117 |
| Factor <i>aditivo</i> | 40 | 3 | 13'33 | 5 | 0'0452 |
| Variab. Explicada | 280 | 9 | | | |
| Residual | 16 | 6 | 2'66 | | |
| Global | 296 | 15 | 19'73 | | |

De esta tabla se deducen los siguientes contrastes:

[1] El contraste de la hipótesis: "el factor γ (*aditivo*) no influye". Se

realiza por el estadístico

$$\hat{F}_\gamma = \frac{13'33}{2'667} = 4'998 \in F_{3,6} \Rightarrow p_\gamma - \text{valor} = 0'0452,$$

se tienen dudas acerca de si aceptar o no esta hipótesis ya que su p -valor $\geq 0'05$. Es el contraste más interesante ya que se contrasta la posible influencia del factor tratamiento en el que se está interesado.

[2] El contraste de la hipótesis: "el factor α (conductor) no influye".

$$\hat{F}_\alpha = \frac{72}{2'667} = 27 \sim F_{3,6} \Rightarrow p_\alpha - \text{valor} = 0'0007,$$

se rechaza esta hipótesis de no influencia del factor "conductor".

[3] El contraste de la hipótesis: "el factor β (coche) no influye".

$$\hat{F}_\beta = \frac{24}{2'667} = 8 \sim F_{3,6} \Rightarrow p_\beta - \text{valor} = 0'1170,$$

se acepta, a un nivel inferior razonable ($< 0'11$) la no influencia del factor "coche".

Los coeficientes de determinación de los tres factores son:

$$R^2(\text{"aditivo"}) = \frac{40}{296} = 0'1351,$$

$$R^2(\text{"conductor"}) = \frac{216}{296} = 0'7297,$$

$$R^2(\text{"coche"}) = \frac{24}{296} = 0'0811,$$

$$R^2(\text{"total"}) = \frac{280}{296} = 0'9459.$$

De los contrastes anteriores se deduce que ha sido conveniente bloquear el tipo de "conductor" pero no conviene bloquear el tipo de "coche". Se puede eliminar el factor "coche", basta con sumar la fila correspondiente al factor "coche" con la fila de la variabilidad residual, aunque se pueden hacer críticas al diseño resultante. Se obtiene la siguiente tabla ANOVA

Tabla ANOVA 2

| Fuentes de Variación | Suma de Cuadrados | Grados libertad | Varianza | \hat{F} | p - valor |
|----------------------|-------------------|-----------------|----------|-----------|-------------|
| Factor conductor | 216 | 3 | 72'00 | 16'20 | 0'0006 |

| | | | | | |
|-----------------------|-----|----|-------|------|--------|
| Factor <i>aditivo</i> | 40 | 3 | 13'33 | 3'00 | 0'0877 |
| Variab. Exp. Total | 256 | 6 | | | |
| Residual | 40 | 9 | 4'44 | | |
| Global | 296 | 15 | 19'73 | | |

Trabajando con un nivel de significación de $\alpha = 0'05$ se **acepta** la no influencia del factor tratamiento "*tipo de aditivo*".

[Anterior] [Arriba]

V.- EXPERIMENTOS FACTORIALES.

Los experimentos factoriales son aquellos en los que se prueban varios niveles de dos o más factores. El número de tratamientos es el resultado de combinar los diferentes niveles de los factores. Un factor es un ingrediente que interviene en un tratamiento, mientras que el nivel es cada una de las dosis o categorías de cada factor. Por ejemplo, el factor A podría ser RAZA con cuatro categorías (las razas de ganado Holstein, Suizo, Charolais y Cebú) y el factor B, DIETA con tres categorías (dieta testigo, una dieta con Canavalia y otra con pollinaza), en este experimento el número de tratamientos es de 12. Todo experimento con dos o más factores tiene un arreglo de tratamientos y un diseño experimental; así, p.e.; hablamos de un experimento con un arreglo factorial en un diseño completamente al azar o en un diseño de bloques al azar.

La aleatorización de los tratamientos se lleva a cabo de acuerdo con el diseño experimental de que se trate. Recuerde que aquí el número de tratamientos es la combinación de los niveles de cada uno de los factores.

Las tres razones principales para realizar Experimento Factorial son:

- 1.- Para obtener información de los efectos medios de todos los factores de un experimento simple de tamaño moderado.
- 2.- Para ampliar la base de las inferencias de un factor para probarlo bajo condiciones variadas de otros.
- 3.- Para evaluar la manera en la cual los efectos de los factores interactúan con cada uno.

Obviamente las razones no son independientes y el énfasis varía con el tipo de experimento.

En general un experimento factorial es más completo porque se puede obtener más información y un grado de precisión mayor del mismo número de observaciones. Suponga que desea realizar dos experimentos. En el primero se compararan tres niveles de proteína usando 36 cerdos en total, en un DCA. En el segundo experimento se desea comparar tres niveles de energía también en un DCA con 36 cerdos. El total de cerdos sería de 72 y en cada experimento la varianza de las medias de las dietas (tratamientos) sería igual al cuadrado medio entre 12 (hay 12 cerdos por dieta). Si se hiciera un experimento factorial (de dos vías) con los 72 cerdos, en cada una de las combinaciones habrían 8 cerdos, y por eso en cada nivel de energía y de proteína un total de 24 observaciones. La varianza de las medias de las dietas sería el cuadro medio

entre 24, es decir, se reduce la varianza de las medias de interés de CM/12 a CM/24 simplemente por combinar los dos experimentos, y al mismo tiempo, se generaliza el experimento debido a la estimación de algunas interacciones entre niveles de proteína y niveles de energía.

5.1 Modelo estadístico.

a) El modelo estadístico para un experimento factorial, con dos factores A y B, en un **diseño completamente al azar** sería,

$$Y_{ijk} = \mu + \alpha_i + \tau_j + \alpha\tau_{ij} + e_{ijk}$$

donde:

Y_{ijk} = es la ijk -ésima observación en el i -ésimo nivel del factor A y j -ésimo nivel del factor B; μ = es la media general; α_i = es el efecto del j -ésimo nivel del factor A; τ_j = es el efecto del k -ésimo nivel del factor B; $\alpha\tau_{ij}$ = es la interacción del i -ésimo nivel del factor A con el j -ésimo nivel del factor B; y e_{ijk} = es el error aleatorio NID $(0, \sigma^2)$.

b) El modelo estadístico para un experimento factorial, con dos factores A y B, en un **diseño en bloques al azar** sería,

$$Y_{ijk} = \mu + \beta_i + \alpha_j + \tau_k + \alpha\tau_{jk} + e_{ijk}$$

donde:

Y_{ijk} = es la ijk -ésima observación en el i -ésimo bloque que contiene el j -ésimo nivel del factor A y el k -ésimo nivel del factor B; μ = es la media general; β_i = es el efecto del i -ésimo bloque; α_j = es el efecto del j -ésimo nivel del factor A; τ_k = es el efecto del k -ésimo nivel del factor B; $\alpha\tau_{jk}$ = es la interacción del j -ésimo nivel del factor A con el k -ésimo nivel del factor B; y e_{ijk} = es el error aleatorio NID $(0, \sigma^2)$.

5.2 Ejemplo de un Diseño de Bloques al Azar con un arreglo factorial 2 x 3.

Se desea determinar el efecto de dos niveles de proteína (alto y bajo) y tres fuentes de proteína (res, cerdo y vegetal) sobre el aumento de peso (g) de ratas.

Los resultados se presentan a continuación,