Diseño de Experimentos - 3007340
DOE - Parte VI: Experimentos con dos factores de efectos fijos en un DCA

Nelfi González Alvarez

Profesora Asociada Escuela de Estadística

e-mail: ngonzale@unal.edu.co

Facultad de Ciencias, Universidad Nacional de Colombia Sede Medellín

Escuela de Estadística

Semestre 02 de 2021

Contenido I

Contenido

Experimentos factoriales de efectos fijos

- Dos o más factores contribuyen a los tratamientos.
- Los niveles de los factores son definidos a criterio del investigador y por tanto solo los tratamientos observados son de interés.
- **Propósito:** ***Estudiar los efectos individuales y de interacción de los factores de estudio, para ello es necesario replicar más de una vez los tratamientos.***

Definición 1.1

Tratamiento es cada combinación posible entre los niveles de los factores de estudio.

Tabla 1: Tratamientos A_iB_j en un experimento con dos factores A y B, cada uno con tres niveles.

	Niveles de B		
Niveles de A	B_1	B_2	B_3
A_1	A_1B_1	A_1B_2	A_1B_3
A_2	A_2B_1	A_2B_2	A_2B_3
A_3	A_3B_1	A_3B_2	A_3B_3

Contenido

Tipos de efectos fijos en el experimento factorial completo

En un experimento factorial con efectos fijos, tenemos

Definición 2.1

1. ***Efecto de un factor:*** *Cambio medio observado en la respuesta debido a un cambio de nivel del factor.*
2. ***Efectos principales:*** *Cambios en la respuesta promedio debidos a la acción individual de cada factor.*
3. ***Efectos de interacción:*** *Miden la magnitud y dirección de las interacciones entre factores. Dos factores interactúan significativamente cuando los efectos de un factor dependen del nivel en que está el otro, y viceversa.*
    - ***Interacciones dobles:*** *Involucran de a dos factores. Ejemplo, con tres factores A, B, C, las interacciones dobles serían AB, AC, BC, y asociadas a cada caso hay un grupo de efectos de interacción según niveles combinados.* ir a ejs. interacción
    - ***Interacciones triples:*** *Involucran de a tres factores.*
    - ***Interacciones de alto orden:*** *Involucrarían más de tres factores pero generalmente no son significativas y resulta poco económico tratar de estudiarlas.*

Nota 2.1

Con más de tres factores se utilizan:

- *Experimentos 2^k: k factores observados cada uno en dos niveles, si k no es muy grande pueden correrse completos los tratamientos en caso contrario se realizan experimentos fraccionados 2^{k-p}, en los que se corre una fracción 2^{-p} del total de 2^k tratamientos posibles.*
- *Experimentos 3^k: k factores cada uno con tres niveles (también existe versión fraccionada).*
- *Otros experimentos: Experimentos para optimizar procesos y experimentos para robustificar procesos, entre otros.*

A continuación veremos el caso de dos factores de efectos fijos en una estructura de diseño DCA.

Contenido

Características del experimento asociado

- Sean A, B los factores de estudio con a y b niveles, respectivamente, definidos según interés del investigador, **luego, la estructura de tratamientos es de efectos fijos.**
- Cada uno de los ab tratamientos es replicado n veces (experimento balanceado) en orden completamente al azar.
- En total se usan abn U.E homogéneas también escogidas y asignadas aleatoriamente a los tratamientos.
- **Dada las dos últimas características del experimento, la estructura de diseño es DCA.**

Ecuación del modelo

Sea $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, n$:

$$\begin{aligned} &Y_{ijk} = \mu + \underbrace{\alpha_i + \beta_j + (\alpha\beta)_{ij}}_{\tau_{ij} = \text{ efecto trat. } AiBj} + \varepsilon_{ijk}, \text{ con } \varepsilon_{ijk} \overset{iid}{\sim} N\left(0, \sigma^2\right), \\ &\text{sujeto a: } \sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = 0, \\ &\sum_{i=1}^{a} (\alpha\beta)_{ij} = 0, \text{para cada } j = 1, \ldots, b, \\ &\sum_{j=1}^{b} (\alpha\beta)_{ij} = 0 \text{ para cada } i = 1, \ldots, a. \end{aligned} \tag{1}$$

De las restricciones anteriores se tiene que $\sum_{i=1}^{a} \sum_{j=1}^{b} \tau_{ij} = 0$.

Variables, parámetros y datos experimentales

Variables y efectos en modelo ANOVA	
Y_{ijk}	Respuesta en la k-ésima réplica del tratamiento $AiBj$, $i = 1, 2, \ldots, a$, $j = 1, 2, \ldots, b$, $k = 1, \ldots, n$.
ε_{ijk}	Error aleatorio en la k-esima réplica del tratamiento $AiBj$.
α_i	Efecto principal fijo del nivel Ai sobre la media global.
β_j	Efecto fijo del principal del nivel Bj sobre la media global.
$(\alpha\beta)_{ij}$	Efecto fijo de la interacción AB en el tratamiento $AiBj$.
τ_{ij}	Efecto fijo del tratamiento $AiBj$, con $\tau_{ij} = \alpha_i + \beta_j + (\alpha\beta)_{ij}$

Medias poblacionales	
μ	Media global.
μ_{ij}	Respuesta media en el tratamiento $AiBj$, $\mu_{ij} = \mu + \tau_{ij}$.
$\mu_{i\bullet}$	Media en el nivel Ai del factor A, donde $\mu_{i\bullet} = \mu + \alpha_i$. Note que $\mu_{i\bullet} = \frac{1}{b}\sum_{j=1}^{b} \mu_{ij} = \mu + \alpha_i$ (prom. medias tratam. en nivel Ai de A).
$\mu_{\bullet j}$	Media en el nivel Bj del factor B donde $\mu_{\bullet j} = \mu + \beta_j$. Note que $\mu_{\bullet j} = \frac{1}{a}\sum_{i=1}^{a} \mu_{ij} = \mu + \beta_j$ (prom. medias tratam. en nivel Bj de B).

Nota 3.1

De las anteriores definiciones de medias y bajo las restricciones sobre los efectos, los efectos principales, de interacciones y de tratamientos se pueden expresar así:

$$\alpha_i = \mu_{i\bullet} - \mu \tag{2}$$

$$\beta_j = \mu_{\bullet j} - \mu \tag{3}$$

$$(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i\bullet} - \mu_{\bullet j} + \mu \tag{4}$$

$$\tau_{ij} = \mu_{ij} - \mu \tag{5}$$

Datos experimentales

	Niveles Factor B						
Niveles Factor A	$B1$	$B2$	$\ldots$	Bj	$\ldots$	Bb	Medias Factor A
$A1$	Y_{111} Y_{112} Y_{113} $\vdots$ Y_{11n}	Y_{121} Y_{122} Y_{123} $\vdots$ Y_{12n}	$\cdots$ $\cdots$ $\cdots$ $\ldots$ $\cdots$	Y_{1j1} Y_{1j2} Y_{1j3} $\vdots$ Y_{1jn}	$\cdots$ $\cdots$ $\cdots$ $\ldots$ $\cdots$	Y_{1b1} Y_{1b2} Y_{1b3} $\vdots$ Y_{1bn}	$\bar{Y}_{1\bullet\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
Ai	Y_{i11} Y_{i12} Y_{i13} $\vdots$ Y_{i1n}	Y_{i21} Y_{i22} Y_{i23} $\vdots$ Y_{i2n}	$\cdots$ $\cdots$ $\cdots$ $\ldots$ $\cdots$	Y_{ij1} Y_{ij2} Y_{ij3} $\vdots$ Y_{ijn}	$\cdots$ $\cdots$ $\cdots$ $\ldots$ $\cdots$	Y_{ib1} Y_{ib2} Y_{ib3} $\vdots$ Y_{ibn}	$\bar{Y}_{i\bullet\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
Aa	Y_{a11} Y_{a12} Y_{a13} $\vdots$ Y_{a1n}	Y_{a21} Y_{a22} Y_{a23} $\vdots$ Y_{a2n}	$\cdots$ $\cdots$ $\cdots$ $\ldots$ $\cdots$	Y_{aj1} Y_{aj2} Y_{aj3} $\vdots$ Y_{ajn}	$\cdots$ $\cdots$ $\cdots$ $\ldots$ $\cdots$	Y_{ab1} Y_{ab2} Y_{ab3} $\vdots$ Y_{abn}	$\bar{Y}_{a\bullet\bullet}$
Medias Factor B	$\bar{Y}_{\bullet 1\bullet}$	$\bar{Y}_{\bullet 2\bullet}$	$\ldots$	$\bar{Y}_{\bullet j\bullet}$	$\ldots$	$\bar{Y}_{\bullet b\bullet}$	Media global $\bar{Y}_{\bullet\bullet\bullet}$

Medias muestrales	
$\bar{Y}_{i\bullet\bullet}$	Media muestral de las nb réplicas en el nivel i del factor A, $\bar{Y}_{i\bullet\bullet} = \frac{1}{nb}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk} = \frac{1}{b}\sum_{j=1}^{b} \bar{Y}_{ij\bullet}$
$\bar{Y}_{\bullet j\bullet}$	Media muestral de las na réplicas en el nivel j del factor B, $\bar{Y}_{\bullet j\bullet} = \frac{1}{na}\sum_{i=1}^{a}\sum_{k=1}^{n} Y_{ijk} = \frac{1}{a}\sum_{i=1}^{a} \bar{Y}_{ij\bullet}.$
$\bar{Y}_{ij\bullet}$	Media muestral de las n réplicas en el tratamiento $AiBj$, $\bar{Y}_{ij\bullet} = \frac{1}{n}\sum_{k=1}^{n} Y_{ijk}.$
$\bar{Y}_{\bullet\bullet\bullet}$	Media muestral global de las abn observaciones, $\bar{Y}_{\bullet\bullet\bullet} = \frac{1}{abn}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk}.$

Significado de la interacción AB

Recuerde definición dada: volver a def. interacción Suponga A: Género; B: Edad; Y: Tiempo de aprendizaje de una tarea (min.). μ_{ij}, $\mu_{i\bullet}$, $\mu_{\bullet j}$, conocidas.

Escenario 1: No interacción entre A y B

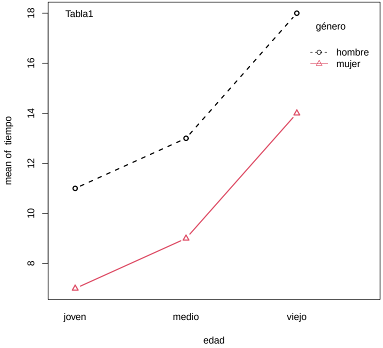
Tabla 1: Medias μ_{ij} en escenario 1

Género	Edad: joven	Edad: medio	Edad: viejo	$\mu_{i\bullet}$
hombre	11	13	18	14.00
mujer	7	9	14	10.00
$\mu_{\bullet j}$	9	11	16	12.00

Dif. Hombre - mujer en cada edad: $\mu_{1j}-\mu_{2j}$		
Edad		
joven	medio	viejo
4	4	4

Observe que las diferencias en medias entre hombres y mujeres en cada nivel de edad son iguales entre sí y coinciden con $\mu_{1\bullet}-\mu_{2\bullet}=4$.

Los perfiles de medias de tratamientos según género son paralelos:



Escenario 2: Interacción entre A y B

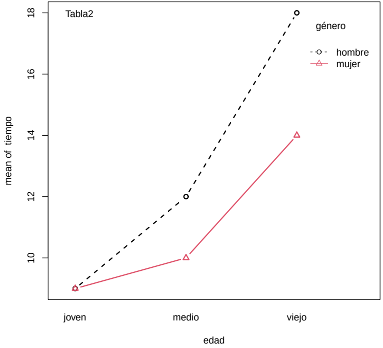
Tabla 2: Medias μ_{ij} en escenario 2

género	edad			$\mu_{i\bullet}$
	joven	medio	viejo	
hombre	9	12	18.00	13.00
mujer	9	10	14.00	11.00
$\mu_{\bullet j}$	9	11	16.00	12.00

Dif. Hombre - mujer en cada edad: $\mu_{1j} - \mu_{2j}$		
Edad		
joven	medio	viejo
0	2	4

Observe que las diferencias en medias entre hombres y mujeres en cada edad, aumentan con el nivel de edad, es decir, a medidas que envejecen, en promedio a los hombres les toma más tiempo aprender la tarea que a las mujeres. Además, no todas coinciden con $\mu_{1\bullet} - \mu_{2\bullet} = 2$.

Los perfiles de medias de tratamientos según género no son paralelos y se distancian más con la edad:



Escenario 3: Interacción entre A y B

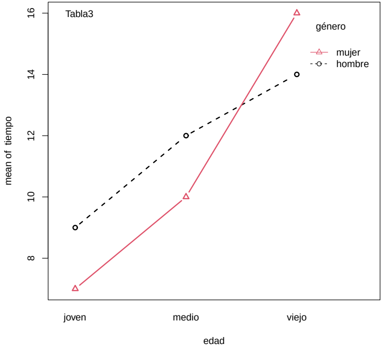
Tabla 3: Medias μ_{ij} en escenario 3

género	edad: joven	medio	viejo	$\mu_{i\bullet}$
hombre	9	12	14	11.67
mujer	7	10	16	11.00
$\mu_{\bullet j}$	8	11	15	11.33

Dif. Hombre - mujer en cada edad: $\mu_{1j} - \mu_{2j}$		
Edad		
joven	medio	viejo
2	2	-2

Observe que las diferencias en medias entre hombres y mujeres es 2, en los niveles de edad joven y medio, pero la diferencia se invierte en signo en el nivel edad viejo. Además, ninguna de estas diferencias concuerdan con $\mu_{1\bullet} - \mu_{2\bullet} = 0.67$.

Los perfiles de medias de tratamientos según género se cruzan cambiando el signo de las diferencias entre medias $\mu_{1j} - \mu_{2j}$:



Escenario 4: Interacción leve entre A y B

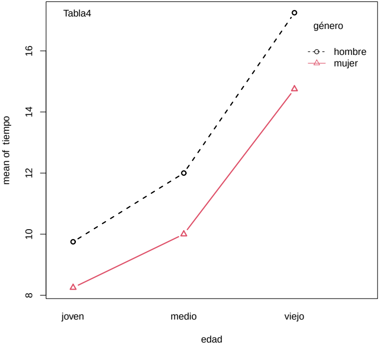
Tabla 4: Medias μ_{ij} en escenario 3

género	edad			$\mu_{i\bullet}$
	joven	medio	viejo	
hombre	9.75	12	17.25	13.00
mujer	8.25	10	14.75	11.00
$\mu_{\bullet j}$	9.00	11	16.00	12.00

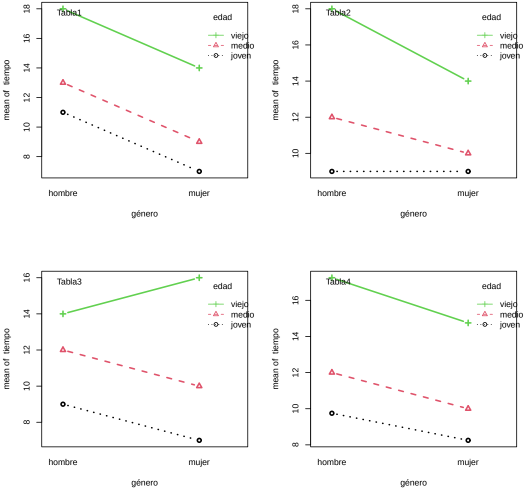
Dif. Hombre - mujer en cada edad: $\mu_{1j} - \mu_{2j}$		
Edad		
joven	medio	viejo
1.5	2	2.5

Observe que las diferencias en medias entre hombres y mujeres en cada nivel de edad, es positiva y crece levemente con la edad, mientras que $\mu_{1\bullet} - \mu_{2\bullet} = 2$.

Los perfiles de medias de tratamientos según género son levemente no paralelos:



También pueden evaluarse las diferencias de medias de tratamientos según edades, en cada nivel de género, es decir, $\mu_{i1}-\mu_{i2}$, $\mu_{i1}-\mu_{i3}$, $\mu_{i2}-\mu_{i3}$, *para cada* $i=1,2$.



Nota 3.2

En la práctica no conocemos las medias poblacionales, solo contamos con la información experimental, por tanto, para diagnosticar la posible presencia de interacción, analizamos los gráficos de perfiles de las medias muestrales de tratamientos $\bar{Y}_{ij\bullet}$:

1. *Si se observa que los gráficos de perfiles de medias muestrales de tratamientos ($\bar{Y}_{ij\bullet}$) no son paralelos.*
2. *Si la diferencia entre las respuestas medias para cualesquiera dos niveles del factor A no es la misma para todos los niveles de B.*
3. *Si la diferencia entre las respuestas medias para cualesquiera dos niveles del factor B no es la misma para todos los niveles de A.*
4. *Si las medias de tratamientos no son todas iguales a la suma de la media global más los respectivos efectos principales de los factores A y B.*

Nota 3.3

Lo anterior debe ser valorado estadísticamente, es decir, no es solo considerar diferencias numéricas entre medias muestrales, sino si estadísticamente pudieran ser significativas tales diferencias.

Estimaciones de mínimos cuadrados

Tabla 2: Estimadores de medias y efectos en un DCA con dos factores de efectos fijos, con interacción.

Medias	Estimador	s.e	I.C del $(1-\gamma)100\%$
μ	$\bar{Y}_{\bullet\bullet\bullet}$	$S_{\bar{Y}_{\bullet\bullet\bullet}} = \sqrt{\frac{MSE}{abn}}$	$\bar{Y}_{\bullet\bullet\bullet} \pm t_{\gamma/2,ab(n-1)} \times S_{\bar{Y}_{\bullet\bullet\bullet}}$
$\mu_{i\bullet}$	$\bar{Y}_{i\bullet\bullet}$	$S_{\bar{Y}_{i\bullet\bullet}} = \sqrt{\frac{MSE}{bn}}$	$\bar{Y}_{i\bullet\bullet} \pm t_{\gamma/2,ab(n-1)} \times S_{\bar{Y}_{i\bullet\bullet}}$
$\mu_{\bullet j}$	$\bar{Y}_{\bullet j\bullet}$	$S_{\bar{Y}_{\bullet j\bullet}} = \sqrt{\frac{MSE}{an}}$	$\bar{Y}_{\bullet j\bullet} \pm t_{\gamma/2,ab(n-1)} \times S_{\bar{Y}_{\bullet j\bullet}}$
μ_{ij}	$\bar{Y}_{ij\bullet}$	$S_{\bar{Y}_{ij\bullet}} = \sqrt{\frac{MSE}{n}}$	$\bar{Y}_{ij\bullet} \pm t_{\gamma/2,ab(n-1)} \times S_{\bar{Y}_{ij\bullet}}$

Efectos	Estimador	s.e	I.C del $(1-\gamma)100\%$
α_i	$\widehat{\alpha}_i = \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet}$	$S_{\widehat{\alpha}_i} = \sqrt{MSE\left(\frac{1}{bn} - \frac{1}{abn}\right)}$	$\widehat{\alpha}_i \pm t_{\gamma/2,ab(n-1)} \times S_{\widehat{\alpha}_i}$
β_j	$\widehat{\beta}_j = \bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet}$	$S_{\widehat{\beta}_i} = \sqrt{MSE\left(\frac{1}{an} - \frac{1}{abn}\right)}$	$\widehat{\beta}_i \pm t_{\gamma/2,ab(n-1)} \times S_{\widehat{\beta}_i}$
$(\alpha\beta)_{ij}$	$\widehat{(\alpha\beta)}_{ij} = \bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet j\bullet} + \bar{Y}_{\bullet\bullet\bullet}$	$S_{\widehat{(\alpha\beta)}_{ij}} = \sqrt{\frac{(a-1)(b-1)}{abn} MSE}$	$\widehat{(\alpha\beta)}_{ij} \pm t_{\gamma/2,ab(n-1)} \times S_{\widehat{(\alpha\beta)}_{ij}}$

De lo anterior, se tiene que

- La estimación de Y_{ijk} es la media muestral del respectivo tratamiento:

$$\widehat{Y}_{ijk} = \bar{Y}_{ij\bullet} \tag{6}$$

- Por tanto, el residuo de ajuste en la k-ésima observación del tratamiento $AiBj$, es

$$\widehat{\varepsilon}_{ijk} = Y_{ijk} - \bar{Y}_{ij\bullet} \tag{7}$$

Con los residuos estandarizados verificamos supuestos asumidos en los errores de ajuste.

Encubrimiento o enmascaramiento de efectos principales

Definición 3.1

Enmascaramiento de efectos principales ocurre cuando ante interacción significativa, los tests F no muestran significancia de los efectos principales de alguno de los factores que participan en la interacción.

Nota 3.4

Aún cuando no haya enmascaramiento de la significancia de los efectos principales cuando hay interacción significativa entre dos o más factores, es mejor no hacer inferencias marginales sobre los efectos o las medias asociadas a los niveles de un factor que participe en una interacción significativa.

Nota 3.5

Recuerde que si hay una interacción significativa, los efectos de un factor se componen de efectos principales más efectos de interacción, de modo que en el caso con dos factores A y B, el efecto del factor A en su nivel i dependerá también del nivel en el que se encuentre B. por ejemplo,

- *si B está fijo en el nivel B_2, el efecto de A en su nivel A_1 es: $\alpha_1 + (\alpha\beta)_{12}$*
- *pero si B está fijo en el nivel B_3, el efecto de A en su nivel A_1 es: $\alpha_1 + (\alpha\beta)_{13}$.*

Luego, bajo interacción entre A y B, no podemos generalizar que α_1 corresponde al efecto que el factor A tiene sobre la media de la respuesta, cuando se fija este factor en su nivel A_1.

Contenido

ANOVA y pruebas asociadas

La variabilidad total (SST) y sus grados de libertad, descomponen en:

$$\mathrm{SST} = \mathrm{SSA} + \mathrm{SSB} + \mathrm{SS(AB)} + \mathrm{SSE}$$

$$\underbrace{\mathrm{g.l(SST)}}_{abn-1} = \underbrace{\mathrm{g.l(SSA)}}_{(a-1)} + \underbrace{\mathrm{g.l(SSB)}}_{(b-1)} + \underbrace{\mathrm{g.l(SS(AB))}}_{(a-1)(b-1)} + \underbrace{\mathrm{g.l(SSE)}}_{ab(n-1)} \tag{8}$$

Donde

- SSA es la suma de cuadrados debida a efectos principales de A
- SSB es la suma de cuadrados debida a efectos principales de B
- SS(AB) es la suma de cuadrados debida a efectos de interacción
- SSE es la suma de cuadrados del error (variabilidad no explicada)

★★Ver en Tabla 8.4 en Notas de Clase - Diseño de Experimentos, Capítulo 8 las expresiones matemáticas de las sumas de cuadrados.

Tabla 3: ANOVA dos factores de efectos fijos con interacción en un DCA

Fuente	g.l	SC	CM	CME	**F**	Valor P
A	$a-1$	SSA	$MSA = \frac{SSA}{a-1}$	$\sigma^2 + \frac{bn\sum_{i=1}^{a}\alpha_i^2}{a-1}$	$\frac{MSA}{MSE}$	$P\left(f_{a-1,\text{dfe}} > F_0\right)$
B	$b-1$	SSB	$MSB = \frac{SSB}{b-1}$	$\sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$	$\frac{MSB}{MSE}$	$P\left(f_{b-1,\text{dfe}} > F_0\right)$
AB	dfi	$SS(AB)$	$MS(AB) = \frac{SS(AB)}{\text{dfi}}$	$\sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\alpha\beta)_{ij}^2}{\text{dfi}}$	$\frac{MS(AB)}{MSE}$	$P\left(f_{\text{dfi},\text{dfe}} > F_0\right)$
Error	dfe	SSE	$MSE = \frac{SSE}{\text{dfe}}$	σ^2		
Total	$abn-1$	SST				

dfi $= (a-1)(b-1)$, los grados de libertad de la interacción AB
dfe $= (abn-1)-(a-1)-(b-1)-(a-1)(b-1) = ab(n-1)$, los grados de libertad del error.

Test sobre efectos de interacción	Estadístico de prueba:
$H_0 : (\alpha\beta)_{ij} = 0,\ \forall\ i = 1, \dots, a,\ j = 1, \dots, b$ $H_1 :$ algún $(\alpha\beta)_{ij} \neq 0,\ i = 1, 2, \dots a,\ j = 1, 2, \dots b$	$F_3 = \dfrac{MS(AB)}{MSE} \sim f_{(a-1)(b-1),ab(n-1)}$
Criterio de rechazo: A un nivel γ, si $F_3 > f_{\gamma,(a-1)(b-1),ab(n-1)}$, o si $P\left(f_{(a-1)(b-1),ab(n-1)} > F_3\right)$ es pequeño.	

Test sobre efectos principales de A	Estadístico de prueba:
$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$ $H_1 :$ algún $\alpha_i \neq 0,\ i = 1, 2, \dots a$	$F_1 = \dfrac{MSA}{MSE} \sim f_{a-1,ab(n-1)}$
Criterio de rechazo: A un nivel γ, si $F_1 > f_{\gamma,a-1,ab(n-1)}$, o si $P\left(f_{a-1,ab(n-1)} > F_1\right)$ es pequeño.	

Test sobre efectos principales de B	Estadístico de prueba:
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$ $H_1 :$ algún $\beta_j \neq 0,\ j = 1, 2, \dots b$	$F_2 = \dfrac{MSB}{MSE} \sim f_{b-1,ab(n-1)}$
Criterio de rechazo: A un nivel γ, si $F_2 > f_{\gamma,b-1,ab(n-1)}$, o si $P\left(f_{b-1,ab(n-1)} > F_2\right)$ es pequeño.	

Tener presente que:

- *La estimación de un efecto principal es un contraste entre todas las medias de los niveles del factor correspondiente, por ello en presencia de interacción, se están haciendo generalizaciones sobre los niveles del factor que hacen que dichas inferencias sean equivocadas.*
- *Una comparación generalizada entre las medias de un factor tiene sentido cuando no hay interacción. De lo contrario, lo mejor es realizar comparaciones y contrastes de medias de tratamientos μ_{ij} que ayuden a interpretar los efectos significativos de la interacción. Sin embargo, aplicar Tukey o LSD sobre todas las μ_{ij} puede no mostrar diferencias entre tratamientos, por pérdida de potencia cuando ab es grande.*
- *Ante interacción, es más conveniente analizar **los efectos simples de un factor**, que corresponden a las comparaciones entre los niveles de tal factor dentro de un solo nivel del otro factor.*

Efectos simples de un factor en cada nivel del otro

Para ilustrar, considere Factor A con $a = 3$ niveles, B con $b = 2$ niveles y $n = 3$ réplicas por tratamiento, entonces,

Efectos simples de A en el nivel B_1 de B

Niveles de A	Efectos simples	Estimador
A_1	$\mu_{11} - \mu_{\bullet 1} = \alpha_1 + (\alpha\beta)_{11}$	$\bar{Y}_{11\bullet} - \bar{Y}_{\bullet 1 \bullet}$
A_2	$\mu_{21} - \mu_{\bullet 1} = \alpha_2 + (\alpha\beta)_{21}$	$\bar{Y}_{21\bullet} - \bar{Y}_{\bullet 1 \bullet}$
A_3	$\mu_{31} - \mu_{\bullet 1} = \alpha_3 + (\alpha\beta)_{31}$	$\bar{Y}_{31\bullet} - \bar{Y}_{\bullet 1 \bullet}$
$SS(A\|B_1) = \sum_{i=1}^{3} \sum_{k=1}^{3} \left(\bar{Y}_{i1\bullet} - \bar{Y}_{\bullet 1 \bullet}\right)^2$		

Efectos simples de A en el nivel B_2 de B

Niveles de A	Efectos simples	Estimador
A_1	$\mu_{12} - \mu_{\bullet 2} = \alpha_1 + (\alpha\beta)_{12}$	$\bar{Y}_{12\bullet} - \bar{Y}_{\bullet 2 \bullet}$
A_2	$\mu_{22} - \mu_{\bullet 2} = \alpha_2 + (\alpha\beta)_{22}$	$\bar{Y}_{22\bullet} - \bar{Y}_{\bullet 2 \bullet}$
A_3	$\mu_{32} - \mu_{\bullet 2} = \alpha_3 + (\alpha\beta)_{32}$	$\bar{Y}_{32\bullet} - \bar{Y}_{\bullet 2 \bullet}$
$SS(A\|B_2) = \sum_{i=1}^{3} \sum_{k=1}^{3} \left(\bar{Y}_{i2\bullet} - \bar{Y}_{\bullet 2 \bullet}\right)^2$		

Tabla de datos

Niveles de A	Niveles de B	
	B_1	B_2
A_1	Y_{111} Y_{112} Y_{113} Prom: $\bar{Y}_{11\bullet}$	Y_{121} Y_{122} Y_{123} Prom: $\bar{Y}_{12\bullet}$
A_2	Y_{211} Y_{212} Y_{213} Prom: $\bar{Y}_{21\bullet}$	Y_{221} Y_{222} Y_{223} Prom: $\bar{Y}_{22\bullet}$
A_3	Y_{311} Y_{312} Y_{313} Prom: $\bar{Y}_{31\bullet}$	Y_{321} Y_{322} Y_{323} Prom: $\bar{Y}_{32\bullet}$
Promedios niveles B	$\bar{Y}_{\bullet 1\bullet}$	$\bar{Y}_{\bullet 2\bullet}$

$SS(A|B_1)$ es como el SSA de un experimento unifactorial con factor A, donde los datos corresponden a los de la columna sombreada en color magenta y $SS(A|B_2)$ es como el SSA de un experimento unifactorial con factor A, donde los datos corresponden a los de la columna sombreada en naranja.

Efectos simples de B en el nivel A_1 de A

Niv. de B	Efectos simples	Estimador
B_1	$\mu_{11} - \mu_{1\bullet} = \beta_1 + (\alpha\beta)_{11}$	$\bar{Y}_{11\bullet} - \bar{Y}_{1\bullet\bullet}$
B_2	$\mu_{12} - \mu_{1\bullet} = \beta_2 + (\alpha\beta)_{12}$	$\bar{Y}_{12\bullet} - \bar{Y}_{1\bullet\bullet}$
$SS(B\|A_1) = \sum_{j=1}^{2}\sum_{k=1}^{3}\left(\bar{Y}_{1j\bullet} - \bar{Y}_{1\bullet\bullet}\right)^2$		

Efectos simples de B en el nivel A_2 de A

Niv. de B	Efectos simples	Estimador
B_1	$\mu_{21} - \mu_{2\bullet} = \beta_1 + (\alpha\beta)_{21}$	$\bar{Y}_{21\bullet} - \bar{Y}_{2\bullet\bullet}$
B_2	$\mu_{22} - \mu_{2\bullet} = \beta_2 + (\alpha\beta)_{22}$	$\bar{Y}_{22\bullet} - \bar{Y}_{2\bullet\bullet}$
$SS(B\|A_2) = \sum_{j=1}^{2}\sum_{k=1}^{3}\left(\bar{Y}_{2j\bullet} - \bar{Y}_{2\bullet\bullet}\right)^2$		

Efectos simples de B en el nivel A_3 de A

Niv. de B	Efectos simples	Estimador
B_1	$\mu_{31} - \mu_{3\bullet} = \beta_1 + (\alpha\beta)_{31}$	$\bar{Y}_{31\bullet} - \bar{Y}_{3\bullet\bullet}$
B_2	$\mu_{32} - \mu_{3\bullet} = \beta_2 + (\alpha\beta)_{32}$	$\bar{Y}_{32\bullet} - \bar{Y}_{3\bullet\bullet}$
$SS(B\|A_3) = \sum_{j=1}^{2}\sum_{k=1}^{3}\left(\bar{Y}_{3j\bullet} - \bar{Y}_{3\bullet\bullet}\right)^2$		

Tabla de datos

	Niveles de B		
Niv. de A	B_1	B_2	Prom. niv. A
A_1	Y_{111} Y_{112} Y_{113} Prom: $\bar{Y}_{11\bullet}$	Y_{121} Y_{122} Y_{123} Prom: $\bar{Y}_{12\bullet}$	$\bar{Y}_{1\bullet\bullet}$
A_2	Y_{211} Y_{212} Y_{213} Prom: $\bar{Y}_{21\bullet}$	Y_{221} Y_{222} Y_{223} Prom: $\bar{Y}_{22\bullet}$	$\bar{Y}_{2\bullet\bullet}$
A_3	Y_{311} Y_{312} Y_{313} Prom: $\bar{Y}_{31\bullet}$	Y_{321} Y_{322} Y_{323} Prom: $\bar{Y}_{32\bullet}$	$\bar{Y}_{3\bullet\bullet}$

$SS(B|A_1)$ es como el SSB de un experimento unifactorial con factor B, donde los datos corresponden a los de la fila sombreada en color magenta, $SS(B|A_2)$ es como el SSB de un experimento unifactorial con factor B, donde los datos corresponden a los de la fila sombreada en color verde y $SS(B|A_3)$ es como el SSB de un experimento unifactorial con factor B, donde los datos corresponden a los de la fila sombreada en naranja.

Evaluación de efectos simples

Sea factor A con a niveles y factor B con b niveles. *Evaluar significancia de los efectos simples de A en cada nivel de B, es equivalente a evaluar la igualdad de las medias de A en cada nivel de B:* Se hacen las siguientes b pruebas:

En el nivel B_1			
Test	SS$(A\|B_1)$	$F_{A\|B_1}$	VP
$H_0: \mu_{11} = \mu_{21} = \cdots = \mu_{a1}$ vs. $H_1:$ algún par $\mu_{i1} \neq \mu_{i'1}$ con $i \neq i'$	$\sum_{i=1}^{a}\sum_{k=1}^{n}\left(\bar{Y}_{i1\bullet} - \bar{Y}_{\bullet 1\bullet}\right)^2$	$\frac{SS(A\|B_1)/(a-1)}{MSE}$	$P\left(f_{a-1,ab(n-1)} > F_{A\|B_1}\right)$
En el nivel B_2			
Test	SS$(A\|B_2)$	$F_{A\|B_2}$	VP
$H_0: \mu_{12} = \mu_{22} = \cdots = \mu_{a2}$ vs. $H_1:$ algún par $\mu_{i2} \neq \mu_{i'2}$ con $i \neq i'$	$\sum_{i=1}^{a}\sum_{k=1}^{n}\left(\bar{Y}_{i2\bullet} - \bar{Y}_{\bullet 2\bullet}\right)^2$	$\frac{SS(A\|B_2)/(a-1)}{MSE}$	$P\left(f_{a-1,ab(n-1)} > F_{A\|B_2}\right)$
$\vdots$			
En el nivel B_b			
Test	SS$(A\|B_b)$	$F_{A\|B_b}$	VP
$H_0: \mu_{1b} = \mu_{2b} = \cdots = \mu_{ab}$ vs. $H_1:$ algún par $\mu_{ib} \neq \mu_{i'b}$ con $i \neq i'$	$\sum_{i=1}^{a}\sum_{k=1}^{n}\left(\bar{Y}_{ib\bullet} - \bar{Y}_{\bullet b\bullet}\right)^2$	$\frac{SS(A\|B_b)/(a-1)}{MSE}$	$P\left(f_{a-1,ab(n-1)} > F_{A\|B_b}\right)$

Evaluación de efectos simples

Sea factor A con a niveles y factor B con b niveles. *Evaluar significancia de los efectos simples de B en cada nivel de A, es equivalente a evaluar la igualdad de las medias de B en cada nivel de A:* Se hacen las siguientes a pruebas:

En el nivel A_1			
Test	$SS(B\|A_1)$	$F_{B\|A_1}$	VP
$H_0: \mu_{11} = \mu_{12} = \cdots = \mu_{1b}$ vs. $H_1:$ algún par $\mu_{1j} \neq \mu_{1j'}$ con $j \neq j'$	$\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{Y}_{1j\bullet} - \bar{Y}_{1\bullet\bullet}\right)^2$	$\frac{\frac{SS(B\|A_1)}{(b-1)}}{MSE}$	$P\left(f_{b-1,ab(n-1)} > F_{B\|A_1}\right)$
En el nivel A_2			
Test	$SS(B\|A_2)$	$F_{B\|A_2}$	VP
$H_0: \mu_{21} = \mu_{22} = \cdots = \mu_{2b}$ vs. $H_1:$ algún par $\mu_{2j} \neq \mu_{2j'}$ con $j \neq j'$	$\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{Y}_{2j\bullet} - \bar{Y}_{2\bullet\bullet}\right)^2$	$\frac{\frac{SS(B\|A_2)}{(b-1)}}{MSE}$	$P\left(f_{b-1,ab(n-1)} > F_{B\|A_2}\right)$
$\vdots$			
En el nivel A_a			
Test	$SS(B\|A_a)$	$F_{B\|A_a}$	VP
$H_0: \mu_{a1} = \mu_{a2} = \cdots = \mu_{ab}$ vs. $H_1:$ algún par $\mu_{aj} \neq \mu_{aj'}$ con $j \neq j'$	$\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{Y}_{aj\bullet} - \bar{Y}_{a\bullet\bullet}\right)^2$	$\frac{\frac{SS(B\|A_a)}{(b-1)}}{MSE}$	$P\left(f_{b-1,ab(n-1)} > F_{B\|A_a}\right)$

Ejemplo

Ver en Sección 8.5 de Notas de Clase, el ejemplo de dos factores con efectos fijos en un DCA balanceado, donde se presenta el fenómeno de enmascaramiento de los efectos principales. Este ejemplo es una adaptación del problema 19.16 de Kutner et. al. (2005).

Contenido

Modelo ANOVA dos factores de efectos fijos sin interacción

Si se ha probado que los efectos de interacción no son significativos, replanteamos el modelo, así: Sea $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, n$:

$$\begin{gathered} Y_{ijk} = \mu + \underbrace{\alpha_i + \beta_j}_{\tau_{ij}} + \varepsilon_{ijk}, \text{ con } \varepsilon_{ijk} \overset{iid}{\sim} N\left(0, \sigma^2\right), \\ \text{sujeto a: } \sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = 0, \end{gathered} \tag{9}$$

Lo anterior conduce a la siguiente descomposición:

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSE}$$

$$\underbrace{\text{g.l(SST)}}_{abn-1} = \underbrace{\text{g.l(SSA)}}_{(a-1)} + \underbrace{\text{g.l(SSB)}}_{(b-1)} + \underbrace{\text{g.l(SSE)}}_{abn-a-b+1} \tag{10}$$

⋆⋆Ver en Tabla 8.10 en Notas de Clase - Diseño de Experimentos, Capítulo 8 las expresiones matemáticas de las sumas de cuadrados.

Estimaciones de mínimos cuadrados

Tabla 4: Estimadores de medias y efectos en un DCA con dos factores de efectos fijos, sin interacción.

Medias	Estimador	s.e	I.C del $(1-\gamma)100\,\%$
μ	$\bar{Y}_{\bullet\bullet\bullet}$	$S_{\bar{Y}_{\bullet\bullet\bullet}} = \sqrt{\frac{MSE}{abn}}$	$\mu \in \bar{Y}_{\bullet\bullet\bullet} \pm t_{\gamma/2,abn-a-b+1} \times S_{\bar{Y}_{\bullet\bullet\bullet}}$
$\mu_{i\bullet}$	$\bar{Y}_{i\bullet\bullet}$	$S_{\bar{Y}_{i\bullet\bullet}} = \sqrt{\frac{MSE}{bn}}$	$\mu_{i\bullet} \in \bar{Y}_{i\bullet\bullet} \pm t_{\gamma/2,abn-a-b+1} \times S_{\bar{Y}_{i\bullet\bullet}}$
$\mu_{\bullet j}$	$\bar{Y}_{\bullet j\bullet}$	$S_{\bar{Y}_{\bullet j\bullet}} = \sqrt{\frac{MSE}{an}}$	$\mu_{\bullet j} \in \bar{Y}_{\bullet j\bullet} \pm t_{\gamma/2,abn-a-b+1} \times S_{\bar{Y}_{\bullet j\bullet}}$
μ_{ij}	$\bar{Y}_{i\bullet\bullet} + \bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet}$	$S_{\widehat{\mu}_{ij}} = \sqrt{\frac{MSE(a+b-1)}{abn}}$	$\mu_{ij} \in \widehat{\mu}_{ij} \pm t_{\gamma/2,abn-a-b+1} \times S_{\widehat{\mu}_{ij}}$

Efectos	Estimador	s.e	I.C del $(1-\gamma)100\,\%$
α_i	$\widehat{\alpha}_i = \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet}$	$S_{\widehat{\alpha}_i} = \sqrt{MSE\left(\frac{1}{bn} - \frac{1}{abn}\right)}$	$\alpha_i \in \widehat{\alpha}_i \pm t_{\gamma/2,abn-a-b+1} \times S_{\widehat{\alpha}_i}$
β_j	$\widehat{\beta}_j = \bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet}$	$S_{\widehat{\beta}_i} = \sqrt{MSE\left(\frac{1}{an} - \frac{1}{abn}\right)}$	$\beta_i \in \widehat{\beta}_i \pm t_{\gamma/2,abn-a-b+1} \times S_{\widehat{\beta}_i}$

De lo anterior, se tiene que

- La estimación de Y_{ijk} ya no es la media muestral en el tratamiento A_iB_j, sino que es
$$\widehat{Y}_{ijk} = \bar{Y}_{i\bullet\bullet} + \bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet} \tag{11}$$
- Por tanto, el residuo de ajuste en la k-ésima observación del tratamiento $AiBj$, es
$$\widehat{\varepsilon}_{ijk} = Y_{ijk} - \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet j\bullet} + \bar{Y}_{\bullet\bullet\bullet} \tag{12}$$

Con los residuos estandarizados verificamos supuestos asumidos en los errores de ajuste.

Tests ANOVA y otras inferencias cuando no hay interacción

Tabla 5: ANOVA en un DCA balanceado, con dos factores de efectos fijos sin interacción

Fuente	g.l	SC	CM	CME	F	Valor P
A	$a-1$	SSA	$MSA=\dfrac{SSA}{a-1}$	$\sigma^2+\dfrac{bn\sum_{i=1}^{a}\alpha_i^2}{a-1}$	$\dfrac{MSA}{MSE}$	$P\left(f_{a-1,abn-a-b+1}>F_0\right)$
B	$b-1$	SSB	$MSB=\dfrac{SSB}{b-1}$	$\sigma^2+\dfrac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$	$\dfrac{MSB}{MSE}$	$P\left(f_{b-1,abn-a-b+1}>F_0\right)$
Error	$abn-a-b+1$	SSE	$MSE=\dfrac{SSE}{abn-a-b+1}$	σ^2		
Total	$abn-1$	SST				

Test sobre efectos principales de A	Estadístico de prueba:
$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$ $H_1 :$ algún $\alpha_i \neq 0,\ i = 1, 2, \ldots a$	$F_1 = \frac{MSA}{MSE} \sim f_{a-1, abn-a-b+1}$
Criterio de rechazo: A un nivel γ, si $F_1 > f_{\gamma, a-1, abn-a-b+1}$, o si $P\left(f_{a-1, abn-a-b+1} > F_1\right)$ es pequeño.	

Test sobre efectos principales de B	Estadístico de prueba:
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$ $H_1 :$ algún $\beta_j \neq 0,\ j = 1, 2, \ldots b$	$F_2 = \frac{MSB}{MSE} \sim f_{b-1, abn-a-b+1}$
Criterio de rechazo: A un nivel γ, si $F_2 > f_{\gamma, b-1, abn-a-b+1}$, o si $P\left(f_{b-1, abn-a-b+1} > F_2\right)$ es pequeño.	

Nota 5.1

Los efectos principales de cada factor corresponden también a los efectos globales de cada uno de ellos sobre la media de la respuesta, desde que ya no hay interacción. Tenga en cuenta además que la no significancia de efectos α_i, es equivalente a la igualdad de medias $\mu_{i\bullet}$, $i = 1, 2, \ldots, a$, y la no significancia de efectos β_j, es equivalente a la igualdad de medias $\mu_{\bullet j}$, $j = 1, 2, \ldots, b$

Tabla 6: Resumen métodos LSD y Tukey Sobre las medias de A: $H_0: \mu_{i\bullet} = \mu_{j\bullet}$ vs. $H_1: \mu_{i\bullet} \neq \mu_{j\bullet}$

	Método LSD	Método de Tukey
Estadístico de prueba	$D_{ij} = \left\|\bar{Y}_{i\bullet\bullet} - \bar{Y}_{j\bullet\bullet}\right\|$	$D_{ij} = \left\|\bar{Y}_{i\bullet\bullet} - \bar{Y}_{j\bullet\bullet}\right\|$
Valor crít. para D_{ij}	$LSD_A = t_{\gamma/2,abn-a-b+1}\sqrt{\frac{2MSE}{nb}}$	$HSD_A = q_\gamma(a, abn-a-b+1)\sqrt{\frac{MSE}{nb}}$
Rechazo de H_0	si $D_{ij} > LSD_A$	si $D_{ij} > HSD_A$
Usando I.C de nivel $(1-\gamma)100\%$ para $\mu_{i\bullet} - \mu_{j\bullet}$		
Rechazo de H_0	si $0 \notin (\bar{Y}_{i\bullet\bullet} - \bar{Y}_{j\bullet\bullet}) \pm LSD_A$	si $0 \notin (\bar{Y}_{i\bullet\bullet} - \bar{Y}_{j\bullet\bullet}) \pm HSD_A$

$q_\gamma(a, abn-a-b+1)$ es el valor crítico de nivel γ del rango estudentizado para a medias muestrales de una $N(0,\sigma^2)$ y con $abn-a-b+1$ grados de libertad en el estimador de σ^2.

Tabla 7: Resumen métodos LSD y Tukey Sobre las medias de B: $H_0: \mu_{\bullet i} = \mu_{\bullet j}$ vs. $H_1: \mu_{\bullet i} \neq \mu_{\bullet j}$

	Método LSD	Método de Tukey
Estadístico de prueba	$D_{ij} = \left\lvert \bar{Y}_{\bullet i \bullet} - \bar{Y}_{\bullet j \bullet} \right\rvert$	$D_{ij} = \left\lvert \bar{Y}_{\bullet i \bullet} - \bar{Y}_{\bullet j \bullet} \right\rvert$
Valor crít. para D_{ij}	$LSD_B = t_{\gamma/2, abn-a-b+1} \sqrt{\frac{2MSE}{na}}$	$HSD_B = q_\gamma(b, abn-a-b+1) \sqrt{\frac{MSE}{na}}$
Rechazo de H_0	si $D_{ij} > LSD_B$	si $D_{ij} > HSD_B$
Usando I.C de nivel $(1-\gamma)100\%$ para $\mu_{\bullet i} - \mu_{\bullet j}$		
Rechazo de H_0	si $0 \notin (\bar{Y}_{\bullet i \bullet} - \bar{Y}_{\bullet j \bullet}) \pm LSD_B$	si $0 \notin (\bar{Y}_{\bullet i \bullet} - \bar{Y}_{\bullet j \bullet}) \pm HSD_B$

$q_\gamma(b, abn-a-b+1)$ es el valor crítico de nivel γ del rango estudentizado para b medias muestrales de una $N(0, \sigma^2)$ y con $abn-a-b+1$ grados de libertad en el estimador de σ^2.

Nota 5.2

Otras inferencias, bien sea por tests de hipótesis o intervalos de confianza,

- *Contrastes entre medias $\mu_{i\bullet}$ del factor A*
- *Contrastes entre medias $\mu_{\bullet j}$ del factor B*
- *Contrastes entre medias de tratamientos μ_{ij}*

Es necesario deducir apropiadamente las varianzas de diferencias entre medias muestrales y de los contrastes estimados, según el tipo de medias involucradas en cada caso, así como los cuantiles críticos de las distribuciones t student y del rango crítico estudentizado (en los casos de comparaciones de Tukey).

Dean, A., Voss, D., and Draguljić, D. (2017). *Design and Analysis of Experiments*, 2$^{\text{nd}}$ Edition. Springer.

Gutiérrez Pulido, H. y de la Vara Salazar, R. (2012). *Análisis y Diseño de Experimentos*, 3ª Edición. McGraw-Hill.

Kuehl, R. O. (2001). *Diseño de Experimentos. Principios Estadísticos de Diseño y Análisis de Investigación*, 2ª Edición. Thomson Learning.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*, 5$^{\text{th}}$ Edition. McGraw-Hill Irwin.

Montgomery, D. C. (2020). *Design and Analysis of Experiments*, 10$^{\text{th}}$ Edition. John Wiley & Sons, Inc.