TEMA Nº 7 → DISEÑOS CON MÁS DE DOS GRUPOS INDEPENDIENTES. ANÁLISIS DE VARIANZA CON DOS FACTORES COMPLETAMENTE ALEATORIZADOS

DISEÑOS FACTORIALES (ANOVA BIFACTORIAL)

*En los **diseños factoriales** se quiere estudiar el efecto que sobre una variable dependiente ejercen dos variables independientes. Dos factores A y B (variables independientes), completamente aleatorizados (los sujetos se asignan aleatoriamente a los distintos niveles de los factores), cada uno de ellos con igual número de niveles (categorías) de modo que todos los niveles de uno se combinan con los del otro.*

Con un diseño factorial se pueden analizar:

- ✓ ***Efectos principales** de cada una de las variables independientes (factores) sobre la VD. Se consideran tantos efectos principales como factores hay implicados en el diseño.*
- ✓ ***Efecto interacción entre los factores**. Significa que, además de la influencia que ejercen por separado cada uno de los dos factores, existe también una influencia debida a la combinación de los dos factores; es decir, cuando el efecto producido por los niveles de cada uno de los factores depende de los niveles del otro factor.*
- ✓ ***Efectos simples** (efectos más focalizados de cada factor con cada nivel del otro factor). Se trata de contrastar los tratamientos de un factor en cada nivel del otro factor.*

Teniendo en cuenta nuestro problema ejemplo:

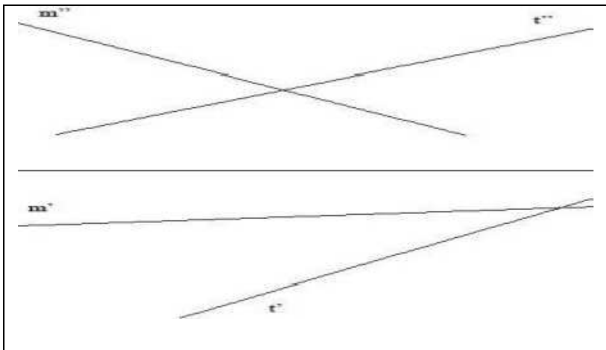
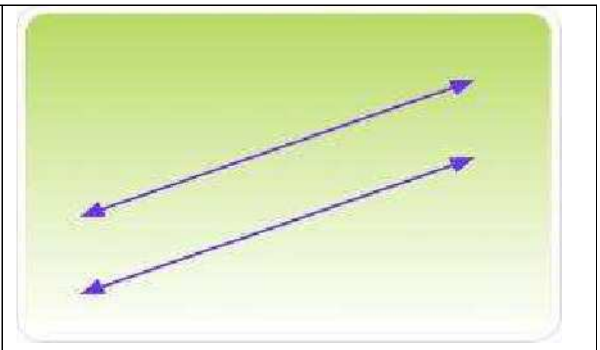
- ✓ ***Efectos principales**: Se puede contrastar, de forma independiente, el efecto de los métodos de aprendizaje en el rendimiento (al margen de los turnos) y el efecto de los turnos en el rendimiento (al margen de los métodos de aprendizaje). Se trata de realizar un ANOVA con cada factor, sin considerar el otro.*
- ✓ ***Efecto interacción**: se relaciona con el cruce de los tratamientos de cada uno de los factores. En nuestro caso no se produce interacción. Se considera interacción cuando una variable independiente no tiene un efecto constante en todos los niveles de la otra variable independiente.*
- ✓ ***Efectos simples**: considerar los métodos de aprendizaje solamente por la mañana, o por la tarde, o los turnos con respecto al método de aprendizaje A, o al C, etc. Se pueden contemplar tantos efectos simples como la suma de los niveles de cada factor.*

Procedimiento**: A partir de la tabla del ANOVA factorial, se realiza la **prueba ómnibus** (efectos principales y efecto interacción). Teniendo en cuenta el efecto **interacción:

*→ Si **rechazamos H_0**, (efecto interacción significativo), el efecto producido por los niveles de un factor depende de los niveles del otro factor.*

*→ Si **aceptamos H_0**, (efecto interacción no significativo), la interacción no existe, y por tanto el modelo es aditivo, es decir se convierte en un Anova factorial sin interacción.*

- ✓ *Si el efecto **interacción no es significativo** → Prueba de los **efectos principales** → Si ninguno es significativo (fin del contraste) → Si algún efecto es significativo (comparaciones múltiples habituales)*
- ✓ *Si el efecto **interacción es significativo** → Prueba de los **efectos simples** y comparación por pares de los efectos simples → Prueba de los efectos principales en función de los efectos simples.*

	
Gráfica con efecto interacción	***Gráfica sin efecto interacción***

PROBLEMA EJEMPLO →

Sea un grupo de 36 sujetos a los que distribuimos aleatoriamente en turnos de mañana, tarde y noche (factor A) y los sometemos a tres formas distintas de aprendizaje (factor B) ¿Cuál de las tres formas de aprendizaje produce mejores resultados? También se desea saber si los turnos influyen en los rendimientos. Sabemos que se cumplen los supuestos del Anova.

TURNOS (FACTOR A)	MÉTODOS DE APRENDIZAJE (FACTOR B)				
	A	B	C	**Σ**	**Σ^2**
MAÑANA	5 6 4 5	8 9 7 8	5 6 8 5	76	510
TARDE	4 7 6 7	9 6 6 7	4 7 6 3	72	462
NOCHE	3 7 6 4	7 6 5 6	8 6 7 7	72	454
Σ	64	84	72	**220**	
Σ^2	362	606	458		**1 26**

<u>Supuestos</u>: *Se trata de un ANOVA de dos factores**.** Al tener 3 niveles cada factor, el número de tratamientos es (3 x 3 = 9). Vamos a probar los efectos principales, la interacción y los efectos simples.*

<u>Cálculo de las sumas de cuadrados (nomenclatura clásica)</u>:

SC $_{TOTAL}$ **=** $\Sigma \Sigma Y^2 - \{(\Sigma \Sigma Y)^2 / N\} \rightarrow 1426 - (220)^2 / 36 =$ **81,55**

SC $_{FACTOR\ A}$ **=** $\Sigma (\Sigma Y)^2 / n - \{(\Sigma \Sigma Y)^2 / N\} \rightarrow (76)^2 / 12 + (72)^2 / 12 + (72)^2 / 12 =$ **0,87**

SC $_{FACTOR\ B}$ **=** $\Sigma (\Sigma X)^2 / n - \{(\Sigma \Sigma X)^2 / N\} \rightarrow (64)^2 / 12 + (84)^2 / 12 + (72)^2 / 12 =$ **16,9**

SC $_{INTRA\ (ERROR)}$ **=** $81,55 - (0,87 + 16,9) =$ **63,78**

SC$_{(AXB)}$ **=** $(20^2 + 32^2 + 24^2 + 24^2 + 28^2 + 20^2 + 20^2 + 24^2 + 28^2) / 4 - (0,87 - 16,9 - 1344,444) =$ **17,7** *(Al considerar la interacción, este valor se resta de la* **SC** $_{ERROR}$*)*

<u>Cálculo de las razones básicas y las sumas de cuadrados (nomenclatura abreviada)</u>:

[T] = $(\sum\sum N)^2 / a - s \rightarrow 220^2 / (3 \cdot 3 \cdot 4) =$ **1344´44**	**SC$_A$ =** $[A]-[T] = (1361´33 - 1344´44) =$ **16´9**
[Y] = $\sum\sum N^2 =$ **1426**	**SC$_B$ =** $[B]-[T] = (1345´33 - 1344´44) =$ **0´87**
[A] = $(\sum A^2) / s \rightarrow (64^2+84^2+72^2) / (3 \cdot 4) =$ **1361´33**	**SC$_{AB}$ =**$[AB]-[A]-[B]+[T] =$ **17´7**
[B] = $(\sum B^2) / s \rightarrow (76^2+72^2+72^2) / (3 \cdot 4) =$ **1345´33**	**SC$_{S \mid AB}$ =** $[Y]-[AB] = (1426 - 1380) =$ **46**
[AB] = $(20^2 + 32^2 + 24^2 + 24^2 + 28^2 + 20^2 + 20^2 +$ $+ 24^2 + 28^2) / 4 =$ **1380**	**SC$_T$ =** $[Y]-[T] = (1426 - 1344´44 =$ **81´5**

***<u>Modelo estadístico</u> (ANOVA con interacción) →** $Y = \mu + \alpha + \beta + (\alpha\beta) + \varepsilon$*

*Cualquier puntuación **(Y)** tiene cinco componentes: los niveles de los factores en que se encuentra **(α** y **β)** común a todos los componentes adscritos a esos niveles (promedio de los efectos del tratamiento en los niveles a y b), el efecto producido por la interacción de los factores en la celda **(αβ)**, el error experimental o factores no controlados en el experimento, asociado a cada puntuación **(ε)** y una constante **(μ)** que es la media total de la población.*

TABLA DEL ANOVA BIFACTORIAL CON INTERACCIÓN

FUENT DE VARIACIÓN	*SUMAS CUADRÁTICAS*	*GRADOS DE LIBERTAD*	*MEDIAS CUADRÁTICAS*	*ESTADÍSTICO DE CONTRASTE*
Factor A	*SC_A* *0,87*	*T – 1* *(3 – 1 = 2)*	*MC_A* *(0,87 / 2 = 0,43)*	*F_A = (0,43 / 1,7 = 0,25)*
Factor B	*SC_B* *16,9*	*I – 1* *(3 – 1 = 2)*	*MC_B* *(16,9 / 2 = 8,4)*	*F_B = (8,4 / 1,7 = 4,94)*
AxB	*SC_{AB}* *17,7*	*(T – 1) (I – 1)* *(2 · 2 = 4)*	*MC_{AB}* *(17,7 / 4 = 4,4)*	*F_{AB} = (4,4 / 1,7 = 2,6)*
Intra (S/AB)	*SC_{INTRA}* *46*	*TI (n – 1)* *9 · (4 –1) = 27*	*$MC_{S/AB}$* *(46 / 27 =1,7)*	
Total	*SC_{TOTAL}* *81,5*	*N – 1* *(36 – 1 = 35)*	*F_{TABLAS} → (2) y (27) gl. = 3,32* *F_{TABLAS} → (2) y (27) gl. = 3,32* *F_{TABLAS} → (4) y (27) gl. = 2,7*	

Se cumple → ***$SC_T = SC_A + SC_B + SC_{AB} + SC_{S/AB}$ → 81´5 = 0´87 + 16´9 + 17´7 + 46***

***Observaciones**: Cuando consideramos la tabla del ANOVA con interacción (AxB) la SC_{intra} disminuye en la cuantía que recoge la SC_{AxB} de la interacción. Igual ocurre con los grados de libertad. La consecuencia es que el valor F resulta más grande (disminuye la MC_{error})*

DECISIONES →** Comparando con los valores críticos F_{TABLAS} con 2 y 27 gl, (α = 0,05) → **3,32

Efectos principales:

Factor A ***0,25 < 3,32*** *Aceptamos Hipótesis nula*

Factor B ***4,94 > 3,32*** *Rechazamos Hipótesis nula.*

Efecto interacción:

*Interacción AxB: F_{TABLAS} → (4) y (27) gl.= **2,7.** Como la $F_{EXPERIMENTAL}$ = **2,6 < 2,7** → Aceptamos H_0 de la interacción (no hay interacción significativa) y consideramos el modelo aditivo.*

<u>Modelo estadístico</u> (ANOVA sin interacción) → Y = μ + α + β + ε

*Cualquier puntuación **(Y)** tiene cuatro componentes: los niveles de los factores en que se encuentra **(α y β)** común a todos los componentes adscritos a esos niveles (promedio de los efectos del tratamiento en los niveles a y b), el error experimental o factores no controlados en el experimento, asociado a cada puntuación **(ε)** y una constante **(μ)** que es la media total de la población.*

TABLA DEL ANOVA BIFACTORIAL SIN INTERACCIÓN

FUENTE DE VARIACIÓN	*SUMAS CUADRÁTICAS*	*GRADOS DE LIBERTAD*	*MEDIAS CUADRÁTICAS*	*ESTADÍSTICO DE CONTRASTE*
Factor A	*SC_A* *0,87*	*T – 1* *(3 – 1 = 2)*	*MC_A* *(0,87 / 2 = 0,43*	*F_A =* *(0,43 / 2,05 = **0,21**)*
Factor B	*SC_B* *16,9*	*I – 1* *(3 – 1 = 2)*	*MC_B* *(16,9 / 2 = 8,4)*	*F_B =* *(8,4 / 2,05 = **4,1**)*
Intra (S/AB)	*SC_{INTRA}* *46 + 17´7 = 63´7*	*N - T – I + 1* *(36 - 3 – 3 + 1 = 31)*	*$MC_{S/AB}$* *(63,7 / 31 = 2,05)*	
Total	*SC_{TOTAL}* *81,5*	*N – 1* *(36 – 1 = 35)*	*F_{TABLAS} → (2) y (31) gl. = 3,3* *F_{TABLAS} → (2) y (31) gl. = 3,3*	

$SC_T = SC_A + SC_B + SC_{S/AB}$ → 81´5 = 0´87 + 16´9 + 63´7

***Observaciones:** al desaparecer la interacción aumenta la SC_{intra} en la cuantía que antes recogía la SC_{AxB} de la interacción. Igual ocurre con los grados de libertad. La consecuencia es que el valor F resulta más pequeño (aumenta la MC_{error})*

DECISIONES →** Comparando con los valores críticos F_{TABLAS} con 2 y 31 gl, (α = 0,05) → **3,3

Factor A ***0,21 < 3,3*** *Aceptamos Hipótesis nula*

Factor B ***4,2 > 3,3*** *Rechazamos Hipótesis nula.*

*El turno no produce efectos; el método de aprendizaje sí. Para ver entre que métodos existen diferencias significativas es preciso realizar **comparaciones múltiples.***

CONTRASTE DE LOS EFECTOS SIMPLES (A partir de las Razones Básicas)

	FACTOR A			
FACTOR B	**MAÑANA (a_1)**	**TARDE (a_2)**	**NOCHE (a_3)**	**SUMAS**
A (b_1)	**20**	**24**	**20**	**64**
B (b_2)	**32**	**28**	**24**	**84**
C (b_3)	**24**	**20**	**28**	**72**
SUMAS	**76**	**72**	**72**	**220**

*Á partir de la **razón básica** para calcular los efectos principales del factor A:*

$$[A] = (\sum A^2) / (b)(n) \quad y \quad [T] = T^2 / ((a)(b)(n)) \rightarrow SC_A = [A]-[T]$$

Deducimos la suma de cuadrados del factor A en el nivel b_1 del factor B:

$$[A\ en\ b_1] = (\sum A\ B^2{}_1) / (n) \quad y \quad [B_1] = B^2{}_1 / ((a)(n)) \rightarrow SC_{A\ en\ b1} = [A\ en\ b_1]-[B_1]$$

Aplicado a nuestro ejemplo:

$$SC_{A\ en\ b1} = \frac{(20^2 + 24^2 + 20^2)}{4} - \frac{64^2}{3 \cdot 4} = (344 - 341´33 = \mathbf{2´67})$$

$$SC_{A\ en\ b2} = \frac{(32^2 + 28^2 + 24^2)}{4} - \frac{84^2}{3 \cdot 4} = (596 - 588 = \mathbf{8})$$

$$SC_{A\ en\ b3} = \frac{(24^2 + 20^2 + 28^2)}{4} - \frac{72^2}{3 \cdot 4} = (440 - 432 = \mathbf{8})$$

$$SC_{B\ en\ a1} = \frac{(20^2 + 32^2 + 24^2)}{4} - \frac{76^2}{3 \cdot 4} = (500 - 481´33 = \mathbf{18´67})$$

$$SC_{B\ en\ a2} = \frac{(24^2 + 28^2 + 20^2)}{4} - \frac{72^2}{3 \cdot 4} = (440 - 432 = \mathbf{8})$$

$$SC_{B\ en\ a3} = \frac{(20^2 + 24^2 + 28^2)}{4} - \frac{72^2}{3 \cdot 4} = (440 - 432 = \mathbf{8})$$

TABLA RESUMEN DEL ANOVA (EFECTOS PRINCIPALES Y EFECTOS SIMPLES)

FUENTE DE VARIACIÓN	**SUMAS CUADRÁTICAS**	**GRADOS DE LIBERTAD**	**MEDIAS CUADRÁTICAS**	**ESTADÍSTICO DE CONTRASTE (F)**
SC_A	**0´87**	(3 – 1 = 2)	MC_A (0,87 / 2 = 0,43)	F_A = (0,43/1,7 = 0,25)
SC_B	**16´9**	(3 – 1 = 2)	MC_B (16,9 / 2 = 8,4)	**F_B = (8,4/1,7 = 4,94) ****
SC_{AB}	**17´7**	(2 · 2 = 4)	MC_{AB} = 17,7/4 = 4,4	F_{AB} = (4,4 / 1,7 = 2,6)
SC_{AB1}	2´67	2	2´67/2 =1´335	1´33/1´7 = 0´78
SC_{AB2}	8	2	8/2 = 4	4/1´7 = 2´35
SC_{AB3}	8	2	8/2 = 4	4/1´7 = 2´35
SC_{BA1}	18´67	2	18´67/2 = 9´335	**9´33/1´7 = 5´49 ****
SC_{BA2}	8	2	8/2 = 4	4/1´7 = 2´35
SC_{BA3}	8	2	8/2 = 4	4/1´7 = 2´35
$SC_{S/AB}$	46	TI (n – 1) 9 · (4 –1) = 27	$MC_{S/AB}$ (46 / 27 =1,7)	
SC_{TOTAL}	**81´5**	N – 1 (36 – 1 = 35)	F_{TABLAS} → (2) y (27) gl. = 3,32 F_{TABLAS} → (2) y (27) gl. = 3,32 F_{TABLAS} → (4) y (27) gl. = 2,7	

- *Se debe cumplir →* **$SC_{A\ en\ b1} = SC_A + SC_{AxB}$**
  (2´67 + 8 + 8) = (0´87 + 17´7) → 18´6 = 18´6 → *(Teniendo en cuenta los redondeos)*
- *Los efectos principales son independientes de la interacción (ortogonales)*
- *Los grados de libertad (efectos simples) son 2 → (a – 1) y b – 1)*
- ***** Resultados significativos con α = 0´05 →** En la tabla se observan resultados significativos respecto al factor **métodos de aprendizaje** (F = 4´94); se trata de un efecto principal por lo que habría que proceder a realizar comparaciones múltiples (determinar entre qué grupos se producen diferencias). De igual modo se aprecian resultados significativos al considerar la **interacción entre el factor B y el turno de mañana** (efecto simple)*

Comparaciones por pares dentro de los efectos simples *(se realizan para los efectos simples significativos):* **$SC_{B\ en\ a1}$** *(F = 5´49) → Factor B (métodos) considerando el turno de mañana.*

Medias de las celdas a comparar → (20/4 = 5; 32/4 = 8; 24/4 = 6)
Coeficientes de los contrastes → (1; 0; -1)

Diferencias entre medias: $\boldsymbol{\psi_{B\ en\ a1} = \sum (c_i)(\overline{Y}_{AB})}$ →

$\psi_{B1\text{-}3\ en\ a1} = (-1 \cdot 5) + (0 \cdot 8) + (1 \cdot 6) = \mathbf{(1)}$
$\psi_{B1\text{-}2\ en\ a1} = (1 \cdot 5) + (-1 \cdot 8) + (0 \cdot 6) = \mathbf{(-3)}$
$\psi_{B2\text{-}3\ en\ a1} = (0 \cdot 5) + (1 \cdot 8) + (-1 \cdot 6) = \mathbf{(2)}$

Sumas de cuadrados: $\boldsymbol{SC_{B1\text{-}3\ en\ a1} = n \cdot (\psi)^2 / \sum c^2_i}$ →

$SC_{B1\text{-}3\ en\ a1} = 4 \cdot (-1)^2 / (1)^2 + (0)^2 + (-1)^2 = \mathbf{2}$
$SC_{B1\text{-}2\ en\ a1} = 4 \cdot (-3)^2 / (1)^2 + (0)^2 + (-1)^2 = \mathbf{18}$
$SC_{B2\text{-}3\ en\ a1} = 4 \cdot (2)^2 / (1)^2 + (0)^2 + (-1)^2 = \mathbf{8}$

Razón $\boldsymbol{F = (SC_{B1\text{-}3\ en\ a1}) / SC_{S/AB}}$ → *Resultados no significativos* (F_{TABLAS} = con 1 y 27 gl **→ 4´19**)

$F_{B1\text{-}3\ en\ a1} = 2 / 46 = \mathbf{0´043}$
$F_{B1\text{-}2\ en\ a1} = 18 / 46 = \mathbf{0´39}$
$F_{B2\text{-}3\ en\ a1} = 8 / 46 = \mathbf{0´17}$