Modelo de diseños anidados y cruzado-anidados

Modelo de diseños anidados

En algunas situaciones no se pueden combinar todos los niveles de un factor con todos los niveles de otro, es decir, no se pueden determinar todos los posibles tratamientos que aparecen al cruzar los factores.

Ejemplo.

Supongamos que en un centro de formación profesional se estudia el porcentaje de aprobados en una materia, en los grupos de mañana y de tarde. Por la mañana imparten la asignatura dos personas y por la tarde tres. Cada persona da clase a tres grupos y se supone que estos son réplicas (no son fuente de variación).

Así,

$$\begin{array}{ll}
\text{Factor A} \equiv & \text{Turno } (i = 1, 2) \\
\text{Factor B} \equiv & \text{Persona } (j = 1, \ldots, 5) \\
y_{ij} \equiv & \text{Porcentaje de aprobados}
\end{array}$$

$$\begin{array}{ccccc}
\multicolumn{3}{c}{\textbf{Mañana}} & \multicolumn{2}{c}{\textbf{Tarde}} \\
\swarrow & \downarrow & \searrow & \swarrow & \searrow \\
P1 & P2 & P3 & P4 & P5 \\
\text{g1} & \text{g1} & \text{g1} & \text{g1} & \text{g1} \\
\text{g2} & \text{g2} & \text{g2} & \text{g2} & \text{g2} \\
\text{g3} & \text{g3} & \text{g3} & \text{g3} & \text{g3}
\end{array}$$

Se dice que el factor B está anidado en el factor A, es decir $B \subset A$.

Modelo matemático

Se dice que un factor B está anidado en otro factor A (o que sus niveles están anidados en los de A) cuando cada nivel del factor B aparece asociado a un único nivel del factor A. Se denota como $B \subset A$.

	$\mathbf{A_1}$			$\mathbf{A_2}$		$\cdots$		$\mathbf{A_a}$	
$\swarrow$	$\downarrow$	$\searrow$	$\swarrow$	$\downarrow$	$\searrow$		$\swarrow$	$\downarrow$	$\searrow$
B_1	B_2	B_3	B_4	B_5	B_6	$\cdots$	B_{m-2}	B_{m-1}	B_m
obs. 1	obs. 1	obs. 1	obs. 1	obs. 1	obs. 1	$\cdots$	obs. 1	obs. 1	obs. 1
obs. 2	obs. 2	obs. 2	obs. 2	obs. 2	obs. 2	$\cdots$	obs. 2	obs. 2	obs. 2
obs. 3	obs. 3	obs. 3	obs. 3	obs. 3	obs. 3	$\cdots$	obs. 3	obs. 3	obs. 3

El modelo se expresa como

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

donde

$$\begin{array}{l} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{array}$$

y

$$\begin{aligned} \text{para cada } i, \qquad \sum_{j=1}^{b} \beta_{j(i)} &= 0 \\ \sum_{i=1}^{a} \alpha_i &= 0 \end{aligned}$$

Se observa que $\beta_{j(i)}$ representa el efecto medio adicional del nivel j-ésimo anidado en el nivel i.

Por otro lado, b es el número de niveles anidados en cada nivel i, de modo que el número total de niveles de B es $a \cdot b$ y la suma de los efectos del factor B dentro de cada nivel de A es 0.

Estimadores por mínimos cuadrados

Se tiene que

$$\min_{\mu,\alpha_i,\beta_{j(i)}} \phi = \min_{\mu,\alpha_i,\beta_{j(i)}} \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\mu-\alpha_i-\beta_{j(i)}\right)^2$$

Así,

$$\begin{aligned}
\frac{\partial\phi}{\partial\mu} &= -2\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\mu-\alpha_i-\beta_{j(i)}\right)=0 \implies\\
\hat{\mu} &= \bar{y}_{\cdots}
\end{aligned}$$

Para cada i fijado

$$\begin{aligned}
\frac{\partial\phi}{\partial\alpha_i} &= -2\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\mu-\alpha_i-\beta_{j(i)}\right)=0 \implies\\
& \sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}-bn\bar{y}_{\cdots}-nb\alpha_i=0 \implies\\
\hat{\alpha}_i &= \bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}
\end{aligned}$$

Para cada i fijado y j fijado

$$\begin{aligned}
\frac{\partial\phi}{\partial\beta_{j(i)}} &= -2\sum_{k=1}^{n}\left(y_{ijk}-\mu-\alpha_i-\beta_{j(i)}\right)=0 \implies\\
& \sum_{k=1}^{n} y_{ijk}-n\bar{y}_{\cdots}-n\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)-n\beta_{j(i)}=0 \implies\\
\hat{\beta}_{j(i)} &= \bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}
\end{aligned}$$

De este modo,

$$\hat{y}_{ijk}=\bar{y}_{\cdots}+\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)+\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}\right)=\bar{y}_{ij\cdot}$$

El número total de observaciones es $a\cdot b\cdot n$ y el número total de parámetros a estimar es $1+(a-1)+a(b-1)=ab$, luego el número de grados de libertad total es $abn-ab=ab(n-1)$.

De este modo, la estima de la varianza es

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij\cdot}\right)^2}{ab(n-1)}.$$

Tabla ANOVA

Si se considera la suma de cuadrados total

$$SCT = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{\cdots})^2$$

sumando y restando los términos $\pm\bar{y}_{i\cdot\cdot}$ $\pm\bar{y}_{ij\cdot}$ se obtiene

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

entonces

$$SCT = SCA + SCB(A) + SCE$$

que puesto en términos de totales queda

	A1		**A2**		**A3**		
	B1	*B2*	*B3*	*B4*	*B5*	*B5*	
	y_{111}	y_{121}	y_{211}	y_{221}	y_{311}	y_{321}	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	y_{11n}	y_{12n}	y_{21n}	y_{22n}	y_{31n}	y_{32n}	
$y_{ij\cdot}$	$y_{11\cdot}$	$y_{12\cdot}$	$y_{21\cdot}$	$y_{22\cdot}$	$y_{31\cdot}$	$y_{32\cdot}$	
$y_{i\cdot\cdot}$	$y_{1\cdot\cdot}$		$y_{2\cdot\cdot}$		$y_{3\cdot\cdot}$		$y_{\cdots}$

$$\begin{aligned}SCT &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{1}{abn}y_{\cdots}^2 \\ SCA &= \frac{1}{bn}\sum_{i=1}^{a} y_{i\cdot\cdot}^2 - \frac{1}{abn}y_{\cdots}^2\end{aligned}$$

para cada nivel i fijado se tiene

$$SCB(A)_i = \frac{1}{n}\sum_{j=1}^{b} y_{ij\cdot}^2 - \frac{1}{bn}y_{i\cdot\cdot}^2$$

y como $SCB(A) = \sum_{i=1}^{a} SCB(A)_i$, entonces

$$\begin{aligned} SCB(A) &= \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij\cdot}^2 - \frac{1}{bn}\sum_{i=1}^{a} y_{i\cdot\cdot}^2 \\ SCE &= SCT - SCA - SCB(A) \end{aligned}$$

Los contrastes de hipótesis que se realizan son:

$$\begin{cases} H_0 : \alpha_1 = \cdots = \alpha_a \ = 0 \text{ (el factor } A \text{ no influye)} \\ H_1 : \text{algún } \alpha_i \neq 0 \ \text{ (el factor } A \text{ influye)} \end{cases}$$

en este caso

$$F_0 = \frac{\frac{SCA}{a-1}}{\frac{SCE}{ab(n-1)}} = \frac{MCA}{MCE}$$

de modo que se rechaza H_0 a nivel α si

$$F_0 > F_{(a-1),ab(n-1),\alpha}$$

La otra hipótesis que se contrasta es, $\forall i = 1, \ldots, a$

$$\begin{cases} H_0 : \beta_{1(i)} = \cdots = \beta_{b(i)} = 0 \\ H_1 : \text{algún } \beta_{j(i)} \neq 0 \end{cases}$$

en este caso,

$$F_0 = \frac{\frac{SCB(A)}{a(b-1)}}{\frac{SCE}{ab(n-1)}} = \frac{MCB(A)}{MCE}$$

de modo que se rechaza H_0 a nivel α si

$$F_0 > F_{a(b-1),ab(n-1),\alpha}$$

En este caso, se contrasta la hipótesis de que todos los niveles del factor anidado B son iguales dentro del factor A donde están anidados.

Sin embargo, si se obtiene que son distintos a nivel global, es interesante contrastar, a continuación, si los niveles del factor B anidado en A son iguales entre sí, dentro de cada nivel i (de A) en el que están anidados.

Así, para cada nivel fijado de i, donde $i = 1, \ldots, a$ se contrasta si los niveles del factor anidado son iguales o no dentro de cada uno de los niveles del factor A en el que están anidados de manera individual

$$\begin{cases} H_0 : \beta_{1(i)} = \cdots = \beta_{b(i)} = 0 \\ H_1 : \text{algún } \beta_{j(i)} \neq 0 \end{cases}$$

en este caso,

$$F_0 = \frac{\frac{SCB(A)_i}{b-1}}{\frac{SCE}{ab(n-1)}} = \frac{MCB(A)_i}{MCE}$$

de modo que se rechaza H_0 a nivel α si

$$F_0 > F_{(b-1),ab(n-1)}, \alpha$$

La tabla ANOVA es

F. V.	**S. C.**	**G. L.**	**M. C.**	**F**
Factor A	SC_A	$a-1$	$MC_A = \frac{SC_A}{a-1}$	$F_A = \frac{MC_A}{MC_E}$
Factor B $(B \subset A)$	$SC_{B(A)}$	$a(b-1)$	$MC_{B(A)} = \frac{SC_{B(A)}}{a(b-1)}$	$F_{B(A)} = \frac{MC_{B(A)}}{MC_E}$
Residual	SCE	$ab(n-1)$	$MC_E = \frac{SCE}{ab(n-1)}$	
Total	SCT	$abn-1$		

Si $F_{B(A)} = \frac{MC_{B(A)}}{MC_E}$ resulta ser significativo, entonces el contraste se puede descomponer en $i = 1, \ldots, a$ contrastes individuales:

$$F_{B(A)_i} = \frac{MC_{B(A)_i}}{MC_E}.$$

Ejemplo.

Un geólogo estudia el contenido en trazas radiactivas de cinco tipos diferentes de suelo. Para ello recoge cuatro muestras de contenido en sustancias radiactivas en cuatro localidades diferentes que están situadas sobre cada tipo de suelo. Se obtienen los siguientes datos:

	A				**B**				**C**			
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
	6	13	1	7	10	2	4	0	0	10	8	7
	2	3	10	4	9	1	1	3	0	11	5	2
	0	9	0	7	7	1	7	4	5	6	0	5
	8	8	6	9	12	10	9	1	5	7	7	4
$y_{ij\cdot}$	16	33	17	27	38	14	21	8	10	34	20	18
$y_{i\cdot\cdot}$	93				81				82			
$y_{\cdot\cdot\cdot}$	402											

	D				**E**			
	13	*14*	*15*	*16*	*17*	*18*	*19*	*20*
	11	5	1	0	1	6	3	3
	0	10	8	8	4	7	0	7
	6	8	9	6	7	0	2	4
	4	3	4	5	9	3	2	0
$y_{ij\cdot}$	21	26	22	19	21	16	7	14
$y_{i\cdot\cdot}$	88				58			
$y_{\cdot\cdot\cdot}$	402							

$$
\begin{aligned}
SCT &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{1}{abn} y_{\cdots}^2 = \\
&= (6^2 + 13^2 + \cdots + 0^2) - \frac{1}{80} 402^2 = 969{,}95
\end{aligned}
$$

$$
\begin{aligned}
SCA &= \frac{1}{bn}\sum_{i=1}^{a} y_{i\cdot\cdot}^2 - \frac{1}{abn} y_{\cdots}^2 = \\
&= \frac{1}{16}(93^2 + \cdots + 58^2) - \frac{1}{80} 402^2 = 45{,}75
\end{aligned}
$$

$$
\begin{aligned}
SCB(A) &= \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij\cdot}^2 - \frac{1}{bn}\sum_{i=1}^{a} y_{i\cdot\cdot}^2 = \\
&= \frac{1}{4}(16^2 + \cdots + 14^2) - \frac{1}{16}(93^2 + \cdots + 58^2) = 282{,}875
\end{aligned}
$$

$$
SCE = SCT - SCA - SCB(A) = 642
$$

La tabla ANOVA es

F. V.	**S. C.**	**G. L.**	**M. C.**	**F**
Factor A terreno	45,75	$a-1=4$	11,269	$F_A = \frac{MC_A}{MC_E} = 1{,}053$
Factor B(A) localidad	282,875	$a(b-1)=15$	18,858	$F_{B(A)} = \frac{MC_{B(A)}}{MC_E} = 1{,}762$
Residual	642	$ab(n-1)=60$	10,7	
Total	969,95	$abn-1=79$		

Se obtiene que

$$
F_{(a-1),ab(n-1),\alpha} = F_{4,60,0'1} = 2{,}04
$$

por lo cual se acepta $H_0 : \alpha_1 = \cdots = \alpha_a = 0$

Así no existen diferencias significativas entre los terrenos a nivel $\alpha = 0{,}1$.

Por otro lado,

$$F_{a(b-1),ab(n-1),\alpha} = F_{15,60,0'1} = 1{,}6$$

luego se rechaza $\forall i = 1, \ldots, a$ la hipótesis $H_0 : \beta_{1(i)} = \cdots = \beta_{b(i)} = 0$ a nivel $\alpha = 0{,}1$.

Estudiamos los contrastes individuales por nivel.

Se calcula para cada i

$$SCB(A)_i = \frac{1}{n}\sum_{j=1}^{b} y_{ij\cdot}^2 - \frac{1}{bn}y_{i\cdot\cdot}^2 = \frac{1}{4}\sum_{j=1}^{4} y_{ij\cdot}^2 - \frac{1}{16}y_{i\cdot\cdot}^2$$

De este modo,

$$\begin{array}{ll}
i=1 & SCB(A)_1 = \frac{1}{4}(16^2+33^2+17^2+27^2) - \frac{1}{16}93^2 = 50{,}18 \\
i=2 & SCB(A)_2 = \frac{1}{4}(38^2+14^2+21^2+8^2) - \frac{1}{16}83^2 = 126{,}18 \\
i=3 & SCB(A)_3 = \frac{1}{4}(10^2+34^2+20^2+18^2) - \frac{1}{16}82^2 = 74{,}75 \\
i=4 & SCB(A)_4 = \frac{1}{4}(21^2+26^2+22^2+19^2) - \frac{1}{16}88^2 = 6{,}5 \\
i=5 & SCB(A)_5 = \frac{1}{4}(21^2+16^2+7^2+14^2) - \frac{1}{16}58^2 = 25{,}25
\end{array}$$

La tabla ANOVA queda como

F. V.	**S. C.**	**G. L.**	**M. C.**	**F**
Factor A terreno	45,75	4	11,269	$F_A = 1{,}053$
Factor B(A) localidad	282,875	15	18,858	$F_{B(A)} = 1{,}762$
A(1)	50,18	3	16,726	$F_{B(A)_1} = 1{,}56$
B(2)	126,18	3	42,06	$F_{B(A)_2} = \mathbf{3{,}93}$
C(3)	74,75	3	24,92	$F_{B(A)_1} = \mathbf{2{,}33}$
D(4)	6,5	3	2,16	$F_{B(A)_1} = 0{,}202$
E(5)	25,25	3	8,41	$F_{B(A)_1} = 0{,}786$
Residual	642	60	10,7	
Total	969,95	79		

Como

$$F_{3,60,0'1} = 2{,}18$$

existen diferencias significativas en los niveles $B(2)$ y $C(3)$, es decir, respecto a los terrenos de tipo B y C las localidades tienen distinto nivel de sustancia radiactiva.

Modelos anidados de efectos aleatorios y efectos mixtos

Modelo de efectos aleatorios

El modelo es

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

siendo

$$\begin{array}{l} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{array}$$

donde

$$\alpha_i \sim N\left(0, \sigma_\alpha^2\right)$$

para todo i,

$$\beta_{j(i)} \sim N\left(0, \sigma_\beta^2\right)$$

$$\varepsilon_{ijk} \sim N\left(0, \sigma^2\right)$$

donde todas las $v.a.$ son independientes.

Así,

$$y_{ijk} \sim N\left(\mu, \sigma_\alpha^2 + \sigma_\beta^2 + \sigma^2\right)$$

son $v.a.$ independientes.

Las **esperanzas de los cuadrados medios** son:

$$\left\{ \begin{array}{l} E(MCA) = \sigma^2 + n\sigma_\beta^2 + bn\sigma_\alpha^2 \\ E(MCB(A)) = \sigma^2 + n\sigma_\beta^2 \\ E(MCE) = \sigma^2 \end{array} \right.$$

Las **estimas de los componentes de la varianza** son:

$$\left\{ \begin{array}{l} \hat{\sigma}^2 = MCE \\ \hat{\sigma}_\beta^2 = \dfrac{MCB(A) - MCE}{n} \\ \hat{\sigma}_\alpha^2 = \dfrac{MCA - MCB(A)}{bn} \end{array} \right.$$

Los **contrastes de hipótesis** son:

Si $F_A = \dfrac{MCA}{MCB(A)} > F_{a-1,a(b-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \sigma_\alpha^2 = 0$

Si $F_{B(A)} = \dfrac{MCB(A)}{MCE} > F_{a(b-1),ab(n-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \sigma_\beta^2 = 0$

Modelo de efectos mixtos $B \subset A$, (A fijo, B aleatorio)

El modelo es

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

siendo

$$\begin{array}{l} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{array}$$

donde

$$\sum_{i=1}^{a} \alpha_i = 0$$

para todo i,

$$\beta_{j(i)} \sim N\left(0, \sigma_\beta^2\right)$$

$$\varepsilon_{ijk} \sim N\left(0, \sigma^2\right)$$

donde todas las $v.a.$ son independientes.

Así,

$$y_{ijk} \sim N\left(\mu + \alpha_i, \sigma_\beta^2 + \sigma^2\right)$$

son $v.a.$ independientes.

Las **esperanzas de los cuadrados medios** son:

$$\left\{ \begin{array}{l} E(MCA) = \sigma^2 + n\sigma_\beta^2 + \dfrac{bn \sum_{i=1}^{a} \alpha_i^2}{a-1} \\ E(MCB(A)) = \sigma^2 + n\sigma_\beta^2 \\ E(MCE) = \sigma^2 \end{array} \right.$$

Las **estimas de los componentes de la varianza** son:

$$\left\{ \begin{array}{l} \hat{\sigma}^2 = MCE \\ \hat{\sigma}_\beta^2 = \dfrac{MCB(A) - MCE}{n} \\ \hat{\alpha}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot} \end{array} \right.$$

Los **contrastes de hipótesis** son:

Si $F_A = \dfrac{MCA}{MCB(A)} > F_{a-1,a(b-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \alpha_i = 0, \forall i$

Si $F_{B(A)} = \dfrac{MCB(A)}{MCE} > F_{a(b-1),ab(n-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \sigma_\beta^2 = 0$

Modelo de efectos mixtos $B \subset A$, (A aleatorio, B fijo)

El modelo es

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

siendo

$$\begin{array}{l} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{array}$$

donde

$$\alpha_i \sim N\left(0, \sigma_\alpha^2\right)$$

para todo i,

$$\sum_{j=1}^{b} \beta_{j(i)} = 0$$

$$\varepsilon_{ijk} \sim N\left(0, \sigma^2\right)$$

donde todas las $v.a.$ son independientes.

Así,

$$y_{ijk} \sim N\left(\mu + \beta_{j(i)}, \sigma_\alpha^2 + \sigma^2\right)$$

son $v.a.$ independientes.

Las **esperanzas de los cuadrados medios** son:

$$\begin{cases} E(MCA) = \sigma^2 + bn\sigma_\alpha^2 \\ E(MCB(A)) = \sigma^2 + \dfrac{n\sum_{i=1}^{a}\sum_{j=1}^{b}\beta_{j(i)}^2}{a(b-1)} \\ E(MCE) = \sigma^2 \end{cases}$$

Las **estimas de los componentes de la varianza** son:

$$\begin{cases} \hat{\sigma}^2 = MCE \\ \hat{\sigma}^2_\alpha = \dfrac{MCA - MCE}{bn} \\ \hat{\beta}_{j(i)} = \bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} \end{cases}$$

Los **contrastes de hipótesis** son:

Si $F_A = \dfrac{MCA}{MCE} > F_{a-1,a(b-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \sigma^2_\alpha = 0$

Si $F_{B(A)} = \dfrac{MCB(A)}{MCE} > F_{a(b-1),ab(n-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \beta_{i(i)} = 0, \forall i$

Modelos de diseños cruzado-anidados

Se dice que dos factores están completamente cruzados cuando aparecen todas las posibles cobinaciones de los niveles de cada factor, como es el caso de un diseño bifactorial.

Se dice que dos factores están cruzados cuando ninguno de ellos está anidado en el otro, es decir, ni $A \subset B$ ni $B \subset A$.

Por ejemplo en el ejemplo previo de turnos y personal docente, puede ser que alguno de estos dé clase en los dos turnos a la vez, con lo que se tienen dos factores cruzados aunque no *completamente* cruzados.

Los diseños cruzado-anidados se caracterizan por tener tanto factores cruzados como anidados. No existe un único modelo matemático, ya que depende de la disposición de los factores en el diseño.

Modelo 1.

Se trata de estudiar el tiempo de montaje de una serie de piezas de relojería que ha de hacerse a mano. Se consideran 3 posiciones diferentes para montar las piezas y cuatro tamaños diferentes de las mismas. El montaje lo efectua una serie de personas, de modo que se ocupan dos personas distintas para cada montaje con cada tamaño y posición.

Se tiene así, como factores,

$P \equiv$ Posición $(i = 1, 2, 3)$
$T \equiv$ Tamaño $(j = 1, 2, 3, 4)$
$I \equiv$ Individuo $(k = 1, 2)$

$$\begin{array}{cccccccccc} & \mathbf{P_1} & & & & & \mathbf{P_2} & & & \cdots \\ \swarrow & \downarrow & \downarrow & \searrow & & \swarrow & \downarrow & \downarrow & \searrow & \cdots \\ T_1 & T_2 & T_3 & T_4 & & T_1 & T_2 & T_3 & T_4 & \cdots \\ \downarrow & \downarrow & \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\ I1 & I3 & I5 & I7 & & I9 & I11 & I13 & I15 & \cdots \\ I2 & I4 & I6 & I8 & & I10 & I12 & I14 & I16 & \cdots \end{array}$$

Se observa que todos los niveles del factor *posición* se cruzan con todos los niveles del factor *tamaño* y que el factor individuo tiene niveles distintos para cada uno de los cruces, es decir, trabajan personas diferentes en cada caso. El esquema es $P * T$ y $I \subset (P * T)$.

Modelo matemático. En el modelo se tienen que incluir:

– Los efectos principales de P y T.

– Las interacciones entre P y T.

– Los efectos de cada nivel k del factor I anidado en la combinación de (i, j).

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{k(ij)} + \varepsilon_{ijkl}$$

donde
$i = 1, \ldots, a$
$j = 1, \ldots, b$
$k = 1, \ldots, c$
$l = 1, \ldots, n$
sujeto a las restricciones

$$\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = \sum_{i=1}^{a} (\alpha\beta)_{ij} = \sum_{j=1}^{b} (\alpha\beta)_{ij} = 0$$

$$\forall (i, j) \text{ fijo } \sum_{k=1}^{c} \gamma_{k(ij)} = 0$$

ya que el factor I tiene c niveles anidados en cada combinación de los niveles de P y T.

Se minimiza la suma de cuadrados de los errores para obtener los estimadores, derivando con respecto a cada uno de los parámetros e igualando a 0. Se obtiene:

$$
\begin{aligned}
\hat{\mu} &= \bar{y}_{....} \\
\hat{\alpha}_i &= \bar{y}_{i...} - \bar{y}_{....} \\
\hat{\beta}_j &= \bar{y}_{.j..} - \bar{y}_{....} \\
(\widehat{\alpha\beta})_{ij} &= \bar{y}_{ij..} - \bar{y}_{i...} - \bar{y}_{.j..} + \bar{y}_{....}
\end{aligned}
$$

Del mismo modo, para $\forall i, j, k$ fijados

$$
\begin{aligned}
\frac{\partial}{\partial \gamma_{k(ij)}} \sum_{i,j,k,l} \left(y_{ijkl} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\widehat{\alpha\beta})_{ij} - \gamma_{k(ij)} \right)^2 &= 0 \Longrightarrow \\
-2 \sum_{l=1}^{n} \left(y_{ijkl} - \bar{y}_{ij..} - \gamma_{k(ij)} \right) &= 0 \Longrightarrow \\
\hat{\gamma}_{k(ij)} &= \bar{y}_{ijk\cdot} - \bar{y}_{ij..}
\end{aligned}
$$

Las respectivas sumas de cuadrados y grados de libertad son:

Sumas de Cuadrados		**grados de libertad**
$SCT = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} (y_{ijkl} - \bar{y}_{....})^2$	$\Longrightarrow$	$abcn - 1$
$SCA = bcn \sum_{i=1}^{a} (\bar{y}_{i...} - \bar{y}_{....})^2$	$\Longrightarrow$	$a - 1$
$SCB = acn \sum_{j=1}^{b} (\bar{y}_{.j..} - \bar{y}_{....})^2$	$\Longrightarrow$	$b - 1$
$SCAB = cn \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij..} - \bar{y}_{i...} - \bar{y}_{.j..} + \bar{y}_{....})^2$	$\Longrightarrow$	$ab - a - b + 1 = (a-1)(b-1)$
$SCC(AB) = n \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} (\bar{y}_{ijk\cdot} - \bar{y}_{ij..})^2$	$\Longrightarrow$	$abc - ab = ab(c-1)$
$SCE = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} (y_{ijkl} - \bar{y}_{ijk\cdot})^2$	$\Longrightarrow$	$abcn - abc = abc(n-1)$

La tabla ANOVA es

F. V.	S. C.	G. L.	M. C.	F
Factor A	SC_A	$a-1$	$MC_A = \frac{SC_A}{a-1}$	$F_A = \frac{MC_A}{MC_E}$
Factor B	SC_B	$b-1$	$MC_B = \frac{SC_B}{b-1}$	$F_B = \frac{MC_B}{MC_E}$
Interacción A*B	SC_{AB}	$(a-1)(b-1)$	$MC_{AB} = \frac{SC_{AB}}{(a-1)(b-1)}$	$F_{AB} = \frac{MC_{AB}}{MC_E}$
Factor C $(C \subset A*B)$	$SC_{C(AB)}$	$ab(c-1)$	$MC_{C(AB)} = \frac{SC_{C(AB)}}{ab(c-1)}$	$F_{C(AB)} = \frac{MC_{C(AB)}}{MC_E}$
Residual	SCE	$abc(n-1)$	$MC_E = \frac{SCE}{abc(n-1)}$	
Total	SCT	$abcn-1$		

Así los contrastes que se establecen son:
Si $F_A > F_{a-1,abc(n-1),\alpha}$ se rechaza $H_0 \equiv \alpha_i = 0$ a nivel α.
Si $F_B > F_{b-1,abc(n-1),\alpha}$ se rechaza $H_0 \equiv \beta_j = 0$ a nivel α.
Si $F_{AB} > F_{(a-1)(b-1),abc(n-1),\alpha}$ se rechaza $H_0 \equiv (\alpha\beta)_{ij} = 0$ a nivel α.
Si $F_{C(AB)} > F_{ab(c-1),abc(n-1),\alpha}$ se rechaza $H_0 \equiv \gamma_{k(ij)} = 0$ a nivel α.

Modelo 2.

En una serie de establecimientos de una cadena de tiendas de ropa se contabilizan las ventas realizadas de ropa de verano y de ropa de invierno. Se anotan, además los empleados que realizan las ventas.

Hay tres factores:

$A \equiv$ Tienda
$B \equiv$ Tipo de ropa
$C \equiv$ Empleado

El esquema es el siguiente

$$
\begin{array}{cccccccccc}
 & \multicolumn{2}{c}{\textbf{Tienda 1}} & & & \multicolumn{2}{c}{\textbf{Tienda 2}} & & & \cdots \\
 & \swarrow & & \searrow & & \swarrow & & \searrow & & \cdots \\
 & V & & I & & V & & I & & \cdots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & & \cdots \\
E_1 & E_2 & E_1 & E_2 & E_3 & E_4 & E_3 & E_4 & & \cdots
\end{array}
$$

Se observa, aquí, que el factor C (empleado) no está anidado en el cruce de los factores A y B (A$*B$) porque no está combinado con una única combinación (i, j), ya que se supone que están los mismos empleados en verano e invierno.

Así,
$Empleados \subset Tiendas \implies C \subset A$
$Empleados * Ropa \implies C * B$
$Tiendas * Ropa \implies A * B$

Modelo matemático.

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{k(i)} + (\gamma\beta)_{k(i)j} + \varepsilon_{ijkl}$$

donde
$i = 1, \ldots, a$
$j = 1, \ldots, b$
$k = 1, \ldots, c$
$l = 1, \ldots, n$
sujeto a las restricciones

$$\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = \sum_{i=1}^{a} (\alpha\beta)_{ij} = \sum_{j=1}^{b} (\alpha\beta)_{ij} = 0$$

y para cada i

$$\begin{aligned} \sum_{k=1}^{c} \gamma_{k(i)} &= 0 \\ \sum_{k=1}^{c} (\gamma\beta)_{k(i)j} &= \sum_{j=1}^{b} (\gamma\beta)_{k(i)j} = 0 \end{aligned}$$

Se minimiza la suma de cuadrados de los errores para obtener los estimadores, derivando con respecto a cada uno delo parámetros e igualando a 0. Se obtiene:

$$\begin{aligned} \hat{\mu} &= \bar{y}_{\cdots\cdot} \\ \hat{\alpha}_i &= \bar{y}_{i\cdots} - \bar{y}_{\cdots\cdot} \\ \hat{\beta}_j &= \bar{y}_{\cdot j\cdot\cdot} - \bar{y}_{\cdots\cdot} \\ (\widehat{\alpha\beta})_{ij} &= \bar{y}_{ij\cdot\cdot} - \bar{y}_{i\cdots} - \bar{y}_{\cdot j\cdot\cdot} + \bar{y}_{\cdots\cdot} \\ \hat{\gamma}_{k(i)} &= \bar{y}_{i\cdot k\cdot} - \bar{y}_{i\cdots} \end{aligned}$$

Del mismo modo, para $\forall i, j, k$ fijados:

$$\begin{aligned} \frac{\partial}{\partial(\gamma\beta)_{k(i)j}} \sum_{i,j,k,l} \left(y_{ijkl} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\widehat{\alpha\beta})_{ij} - \hat{\gamma}_{k(i)} - (\gamma\beta)_{k(i)j} \right)^2 &= 0 \Longrightarrow \\ \frac{\partial}{\partial(\gamma\beta)_{k(i)j}} \sum_{i,j,k,l} \left(y_{ijkl} - \bar{y}_{ij\cdot\cdot} - \bar{y}_{i\cdot k\cdot} + \bar{y}_{i\cdots} - (\gamma\beta)_{k(i)j} \right)^2 &= 0 \\ -2 \sum_{l=1}^{n} \left(y_{ijkl} - \bar{y}_{ij\cdot\cdot} - \bar{y}_{i\cdot k\cdot} + \bar{y}_{i\cdots} - (\gamma\beta)_{k(i)j} \right) &= 0 \Longrightarrow \end{aligned}$$

y queda

$$\widehat{(\gamma\beta)}_{k(i)j} = \bar{y}_{ijk\cdot} - \bar{y}_{ij\cdot\cdot} - \bar{y}_{i\cdot k\cdot} + \bar{y}_{i\cdots}$$

Las respectivas sumas de cuadrados y grados de libertad son:

Sumas de Cuadrados		**grados de libertad**
$SCT = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}(y_{ijkl} - \bar{y}_{\cdots\cdot})^2$	$\Longrightarrow$	$abcn - 1$
$SCA = bcn\sum_{i=1}^{a}(\bar{y}_{i\cdots} - \bar{y}_{\cdots\cdot})^2$	$\Longrightarrow$	$a - 1$
$SCB = acn\sum_{j=1}^{b}(\bar{y}_{\cdot j\cdot\cdot} - \bar{y}_{\cdots\cdot})^2$	$\Longrightarrow$	$b - 1$
$SCAB = cn\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij\cdot\cdot} - \bar{y}_{i\cdots} - \bar{y}_{\cdot j\cdot\cdot} + \bar{y}_{\cdots\cdot})^2$	$\Longrightarrow$	$ab - a - b + 1 =$ $= (a-1)(b-1)$
$SCC(A) = bn\sum_{i=1}^{a}\sum_{k=1}^{c}(\bar{y}_{i\cdot k\cdot} - \bar{y}_{i\cdots})^2$	$\Longrightarrow$	$ac - a = a(c-1)$
$SCBC(A) = n\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}(\bar{y}_{ijk\cdot} - \bar{y}_{ij\cdot\cdot} - \bar{y}_{i\cdot k\cdot} + \bar{y}_{i\cdots})^2$	$\Longrightarrow$	$abc - ab - ac + a =$ $a(b-1)(c-1)$
$SCE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}(y_{ijkl} - \bar{y}_{ijk\cdot})^2$	$\Longrightarrow$	$abcn - abc =$ $= abc(n-1)$

Modelos de diseños cruzado-anidados con efectos mixtos

Se supone que hay 3 factores A, B y C tales que

$C \subset B * A$
A y B son factores fijos
C es un factor aleatorio

El modelo es

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{k(ij)} + \varepsilon_{ijkl}$$

donde
$i = 1, \ldots, a$
$j = 1, \ldots, b$
$k = 1, \ldots, c$
$l = 1, \ldots, n$

sujeto a las restricciones

$$\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = \sum_{i=1}^{a} (\alpha\beta)_{ij} = \sum_{j=1}^{b} (\alpha\beta)_{ij} = 0$$

$\forall (i, j)$ fijo

$$\begin{aligned} \gamma_{k(ij)} &\sim N(0, \sigma_\gamma^2) \\ \varepsilon_{ijkl} &\sim N(0, \sigma^2) \end{aligned}$$

independientes entre sí.

Las **esperanzas de los cuadrados medios** son:

$$\begin{cases} E[MCA] = \sigma^2 + n\sigma_\gamma^2 + \dfrac{bcn \sum_{i=1}^{a} \alpha_i^2}{a-1} \\ E[MCB] = \sigma^2 + n\sigma_\gamma^2 + \dfrac{acn \sum_{j=1}^{b} \beta_j^2}{b-1} \\ E[MCAB] = \sigma^2 + n\sigma_\gamma^2 + \dfrac{cn \sum_{i=1}^{a} \sum_{j=1}^{b} (\alpha\beta)_{ij}^2}{(a-1)(b-1)} \\ E[MCC(AB)] = \sigma^2 + n\sigma_\gamma^2 \\ E[MCE] = \sigma^2 \end{cases}$$

Las **estimas de los componentes de la varianza** son:

$$\begin{cases} \hat{\sigma}^2 = MCE \\ \hat{\sigma}_\gamma^2 = \dfrac{MCC(AB) - MCE}{n} \end{cases}$$

Los **contrastes de hipótesis** son:

Si $F_A = \dfrac{MCA}{MCC(AB)} > F_{a-1,ab(c-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \alpha_i = 0, \forall i$

Si $F_B = \dfrac{MCB}{MCC(AB)} > F_{b-1,ab(c-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \beta_j = 0, \forall j$

Si $F_{AB} = \dfrac{MCAB}{MCC(AB)} > F_{(a-1)(b-1),ab(c-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv (\alpha\beta)i_j = 0, \forall ij$

Si $F_{C(AB)} = \dfrac{MCC(AB)}{MCE} > F_{ab(c-1),abc(n-1);\alpha}$ se rechaza la hipótesis nula, $H_0 \equiv \sigma_\gamma^2 = 0$

Regla para determinar las medias de cuadrados esperadas (de Hesse)

En principio hay que diferenciar entre los factores fijos y los aleatorios. Las interacciones entre los factores fijos y los aleatorios se consideran aleatorias.

A cada factor aleatorio se le asigna su componente de la varianza $\sigma_\bullet^2$ y a cada factor fijo se le asigna su efecto fijo representado por la suma de cuadrados de los parámetros asociados a ese factor entre sus grados de libertad, e.g. si A es un factor fijo se le asocia $\frac{\sum_{i=1}^{a}\alpha_i^2}{a-1}$.

Se construye la siguiente tabla:

	Tipo Factor	**F**	**A**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	$\sum_{i=1}^{a}\alpha_i^2/(a-1)$	0	b	n
β_j	σ_β^2	a	1	n
$(\alpha\beta)_{ij}$	$\sigma_{\alpha\beta}^2$	1	1	n
$\varepsilon_{(ij)k}$	σ^2	1	1	1

1. El término del error del modelo $\varepsilon_{ij\ldots m}$ se representa como $\varepsilon_{(ij\ldots)m}$ donde m es el índice de las replicaciones, es decir, fijados $ij\ldots$ se consideran m réplicas aleatorias.

2. Los subíndices de cada término se subdividen en tres clases:

   – *activos*: están en el término y no se encuentran entre paréntesis.

   – *pasivos*: están en el término pero se encuentran entre paréntesis.

   – *ausentes*: No están en el término aunque pertenecen al modelo.

3. En cada fila se escribe un 1 si uno de los subíndices pasivos (también están entre paréntesis los correspondientes a los anidados) del componente de la fila coincide con el subíndice de la columna.

	Tipo Factor	**F**	**A**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	$\sum_{i=1}^{a} \alpha_i^2 / (a-1)$			
β_j	σ_β^2			
$(\alpha\beta)_{ij}$	$\sigma_{\alpha\beta}^2$			
$\varepsilon_{(ij)k}$	σ^2	1	1	

4. Si algún subíndice de la fila coincide con el subíndice de la columna, se escribe:

   – Un 1 si es un factor aleatorio

   – Un 0 si es un factor fijo

	Tipo Factor	**F**	**A**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	$\sum_{i=1}^{a} \alpha_i^2 / (a-1)$	0		
β_j	σ_β^2		1	
$(\alpha\beta)_{ij}$	$\sigma_{\alpha\beta}^2$	1	1	
$\varepsilon_{(ij)k}$	σ^2	1	1	1

   En los huecos restantes, se escribe el número de niveles que tiene la columna correspondiente. Queda así la siguiente tabla:

	Tipo Factor	**F**	**A**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	$\sum_{i=1}^{a} \alpha_i^2 / (a-1)$	0	b	n
β_j	σ_β^2	a	1	n
$(\alpha\beta)_{ij}$	$\sigma_{\alpha\beta}^2$	1	1	n
$\varepsilon_{(ij)k}$	σ^2	1	1	1

5. Para obtener el valor esperado de la media de cuadrados de cualquier componente del modelo, se hace:

   a) Se tapan todas las columnas encabezadas por los subíndices activos de ese componente.

   b) Se multiplican los números de las filas que tienen al menos los mismos subíndices que el componente, multiplicándolos a su vez por el factor fijo (suma de cuadrados) o el factor aleatorio (varianza) obtenidos.

c) La suma de estos productos es el valor esperado de la media de cuadrados.

En el ejemplo, quedaría:

$E[MCA] = bn\frac{\sum_i \alpha_i^2}{a-1} + n\sigma^2_{\alpha\beta} + \sigma^2$
$E[MCB] = an\sigma^2_\beta + n\sigma^2_{\alpha\beta} + \sigma^2$
$E[MCAB] = n\sigma^2_{\alpha\beta} + \sigma^2$
$E[MCE] = \sigma^2$

Por tanto, los contrastes de hipótesis son:

$$F_A = \frac{MCA}{MCAB} \text{ para } H_0 \equiv \alpha_i = 0, \forall i$$

$$F_B = \frac{MCB}{MCAB} \text{ para } H_0 \equiv \sigma^2_\beta = 0$$

$$F_{AB} = \frac{MCAB}{MCE} \text{ para } H_0 \equiv \sigma^2_{\alpha\beta} = 0$$

De todas formas, no siempre se pueden construir contrastes de hipótesis para cualquier modelo. Una posible solución es suponer que algunas interacciones son nulas.

Ejemplo. Supongamos un modelo bifactorial con efectos aleatorios:

	Tipo Factor	**A**	**A**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	σ^2_α	1	b	n
β_j	σ^2_β	a	1	n
$(\alpha\beta)_{ij}$	$\sigma^2_{\alpha\beta}$	1	1	n
$\varepsilon_{(ij)k}$	σ^2	1	1	1

Las medias de cuadrados son:

$E[MCA] = bn\sigma^2_\alpha + n\sigma^2_{\alpha\beta} + \sigma^2$
$E[MCB] = an\sigma^2_\beta + n\sigma^2_{\alpha\beta} + \sigma^2$
$E[MCAB] = n\sigma^2_{\alpha\beta} + \sigma^2$
$E[MCE] = \sigma^2$

Por tanto, los contrastes de hipótesis son:

$$F_A = \frac{MCA}{MCAB} \text{ para } H_0 \equiv \sigma_\alpha^2 = 0, \forall i$$

$$F_B = \frac{MCB}{MCAB} \text{ para } H_0 \equiv \sigma_\beta^2 = 0$$

$$F_{AB} = \frac{MCAB}{MCE} \text{ para } H_0 \equiv \sigma_{\alpha\beta}^2 = 0$$

Ejemplo. Supongamos un modelo bifactorial con efectos fijos:

	Tipo Factor	**F**	**F**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	$\sum_{i=1}^a \alpha_i^2 / (a-1)$	0	b	n
β_j	$\sum_{j=1}^b \beta_j^2 / (b-1)$	a	0	n
$(\alpha\beta)_{ij}$	$\sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2 / (a-1)(b-1)$	0	0	n
$\varepsilon_{(ij)k}$	σ^2	1	1	1

Las medias de cuadrados son:

$$E[MCA] = bn\frac{\sum_{i=1}^a \alpha_i^2}{a-1} + \sigma^2$$
$$E[MCB] = an\frac{\sum_{j=1}^b \beta_j^2}{b-1} + \sigma^2$$
$$E[MCAB] = n\frac{\sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2}{(a-1)(b-1)} + \sigma^2$$
$$E[MCE] = \sigma^2$$

Por tanto, los contrastes de hipótesis son:

$$F_A = \frac{MCA}{MCE} \text{ para } H_0 \equiv \alpha_i = 0, \forall i$$

$$F_B = \frac{MCB}{MCE} \text{ para } H_0 \equiv \beta_j = 0, \forall j$$

$$F_{AB} = \frac{MCAB}{MCE} \text{ para } H_0 \equiv (\alpha\beta)_{ij} = 0, \forall i, j$$

Ejemplo. Supongamos un modelo anidado con dos factores: $B \subset A$ donde B es aleatorio y A es fijo:

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

siendo

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

donde

$$\sum_{i=1}^{a} \alpha_i = 0$$

para todo i,

$$\beta_{j(i)} \sim N\left(0, \sigma_\beta^2\right)$$

$$\varepsilon_{ijk} \sim N\left(0, \sigma^2\right)$$

donde todas las $v.a.$ son independientes.

	Tipo Factor	**F**	**A**	**A**
	Num. Niveles	a	b	n
	Subíndices	i	j	k
α_i	$\sum_{i=1}^{a} \alpha_i^2 / (a-1)$	0	b	n
$\beta_{j(i)}$	σ_β^2	1	1	n
$\varepsilon_{(ij)k}$	σ^2	1	1	1

Las medias de cuadrados son:

$$E[MCA] = bn\frac{\sum_{i=1}^{a} \alpha_i^2}{a-1} + n\sigma_\beta^2 + \sigma^2$$
$$E[MCB(A)] = n\sigma_\beta^2 + \sigma^2$$
$$E[MCE] = \sigma^2$$

Por tanto, los contrastes de hipótesis son:

$$F_A = \frac{MCA}{MCB(A)} \text{ para } H_0 \equiv \alpha_i = 0, \forall i$$

$$F_B = \frac{MCB(A)}{MCE} \text{ para } H_0 \equiv \sigma_\beta^2 = 0$$

Ejemplo. Supongamos un modelo por bloques aleatorizados completos con un factor aleatorio y un bloque aleatorio:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

siendo

$$i = 1, \ldots, a$$
$$j = 1, \ldots, b$$

donde

$$\begin{aligned}\alpha_i &\sim N(0,\sigma_\alpha^2)\\ \beta_j &\sim N\left(0,\sigma_\beta^2\right)\\ \varepsilon_{ij} &\sim N\left(0,\sigma^2\right)\end{aligned}$$

donde todas las *v.a.* son independientes.

	Tipo Factor	**A**	**A**	**F***
	Num. Niveles	a	b	1
	Subíndices	i	j	k
α_i	σ_α^2	1	b	1
β_j	σ_β^2	a	1	1
ε_{ij}	σ^2	1	1	1

(*) En este caso no hay réplicas aleatorias y hay un sólo elemento fijo $(k=1)$. De este modo, k no es una réplica por lo que se pone ε_{ij}.

Las medias de cuadrados son:

$E[MCA] = b\sigma_\alpha^2 + \sigma^2$
$E[MCB] = a\sigma_\beta^2 + \sigma^2$
$E[MCE] = \sigma^2$

Por tanto, el contraste de hipótesis es:

$$F_A = \frac{MCA}{MCE} \text{ para } H_0 \equiv \sigma_\alpha^2 = 0$$

Ejemplo. Supongamos un modelo por bloques aleatorizados completos con un factor aleatorio y un bloque fijo:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

siendo

$i = 1, \ldots, a$
$j = 1, \ldots, b$

donde

$$\begin{aligned}\alpha_i &\sim N(0,\sigma_\alpha^2)\\ \sum_{j=1}^{b}\beta_j &= 0\\ \varepsilon_{ij} &\sim N\left(0,\sigma^2\right)\end{aligned}$$

donde todas las $v.a.$ son independientes.

	Tipo Factor	**A**	**F**	**F***
	Num. Niveles	a	b	1
	Subíndices	i	j	k
α_i	σ_α^2	1	b	1
β_j	$\sum_{j=1}^{b}\beta_j^2 \big/ (b-1)$	a	0	1
ε_{ij}	σ^2	1	1	1

(*) En este caso no hay réplicas aleatorias y hay un sólo elemento fijo $(k=1)$. De este modo, k no es una réplica por lo que se pone ε_{ij}.

Las medias de cuadrados son:

$E[MCA] = b\sigma_\alpha^2 + \sigma^2$
$E[MCB] = a\frac{\sum_{j=1}^{b}\beta_j^2}{b-1} + \sigma^2$
$E[MCE] = \sigma^2$

Por tanto, el contraste de hipótesis es:

$$F_A = \frac{MCA}{MCE} \text{ para } H_0 \equiv \sigma_\alpha^2 = 0$$

Aplicación con R

```
# Diseño anidado
datos <- read.table("datAnida.txt", header=T)
attach(datos)

terreno <- as.factor(terreno)
locali <- as.factor(locali)

# Dos formas de programarlo:

modelo <- aov(silice ~ terreno + terreno/locali)

# alternativamente:

modelo <- aov(silice ~ terreno + locali%in%terreno)

summary(modelo)
```

Aplicación con SAS

```
options ls=76 nodate nonumber;
title 'ANOVA ANIDADO';
data ano;
input terreno locali silice;
cards;
1 1 6
1 1 2
1 1 0
1 1 8
1 2 13
1 2 3
1 2 9
1 2 8
1 3 1
1 3 10
1 3 0
1 3 6
1 4 7
1 4 4
1 4 7
1 4 9
2 1 10
2 1 9
2 1 7
2 1 12
2 2 2
2 2 1
2 2 1
2 2 10
2 3 4
2 3 1
2 3 7
2 3 9
2 4 0
2 4 3
2 4 4
2 4 1
```

```
3 1 0
3 1 0
3 1 5
3 1 5
3 2 10
3 2 11
3 2 6
3 2 7
3 3 8
3 3 5
3 3 0
3 3 7
3 4 7
3 4 2
3 4 5
3 4 4
4 1 11
4 1 0
4 1 6
4 1 4
4 2 5
4 2 10
4 2 8
4 2 3
4 3 1
4 3 8
4 3 9
4 3 4
4 4 0
4 4 8
4 4 6
4 4 5
5 1 1
5 1 4
5 1 7
5 1 9
5 2 6
5 2 7
5 2 0
5 2 3
5 3 3
5 3 0
5 3 2
5 3 2
5 4 3
5 4 7
5 4 4
5 4 0
;
proc glm;
class terreno locali;
model silice=terreno locali(terreno);

contrast 'locali(terreno_A)'
locali(terreno) 1 -1  0   0 ,
locali(terreno) 1  0 -1   0 ,
locali(terreno) 1  0  0  -1 ;

contrast 'locali(terreno_B)'
locali(terreno) 0 0 0 0  1 -1  0   0 ,
locali(terreno) 0 0 0 0  1  0 -1   0 ,
locali(terreno) 0 0 0 0  1  0  0  -1 ;
```

```
contrast 'locali(terreno_C)'
locali(terreno) 0 0 0 0  0 0 0 0  1 -1  0   0 ,
locali(terreno) 0 0 0 0  0 0 0 0  1  0 -1   0 ,
locali(terreno) 0 0 0 0  0 0 0 0  1  0  0  -1 ;

contrast 'locali(terreno_D)'
locali(terreno) 0 0 0 0  0 0 0 0  0 0 0 0  1 -1  0   0 ,
locali(terreno) 0 0 0 0  0 0 0 0  0 0 0 0  1  0 -1   0 ,
locali(terreno) 0 0 0 0  0 0 0 0  0 0 0 0  1  0  0  -1 ;

contrast 'locali(terreno_E)'
locali(terreno) 0 0 0 0  0 0 0 0  0 0 0 0  0 0 0 0  1 -1  0   0 ,
locali(terreno) 0 0 0 0  0 0 0 0  0 0 0 0  0 0 0 0  1  0 -1   0 ,
locali(terreno) 0 0 0 0  0 0 0 0  0 0 0 0  0 0 0 0  1  0  0  -1 ;

means locali(terreno);
run;
```

ANOVA ANIDADO
The GLM Procedure

Class Level Information

Class	Levels	Values
terreno	5	1 2 3 4 5
locali	4	1 2 3 4

Number of observations 80

ANOVA ANIDADO
The GLM Procedure

Dependent Variable: silice

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	19	327.9500000	17.2605263	1.61	0.0823
Error	60	642.0000000	10.7000000		
Corrected Total	79	969.9500000			

R-Square	Coeff Var	Root MSE	silice Mean
0.338110	65.09623	3.271085	5.025000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
terreno	4	45.0750000	11.2687500	1.05	0.3876
locali(terreno)	15	282.8750000	18.8583333	1.76	0.0625

Source	DF	Type III SS	Mean Square	F Value	Pr > F
terreno	4	45.0750000	11.2687500	1.05	0.3876
locali(terreno)	15	282.8750000	18.8583333	1.76	0.0625

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
locali(terreno_A)	3	50.1875000	16.7291667	1.56	0.2076
locali(terreno_B)	3	126.1875000	42.0625000	3.93	0.0125
locali(terreno_C)	3	74.7500000	24.9166667	2.33	0.0835
locali(terreno_D)	3	6.5000000	2.1666667	0.20	0.8943
locali(terreno_E)	3	25.2500000	8.4166667	0.79	0.5061

Level of	Level of		------------silice-----------	
locali	terreno	N	Mean	Std Dev
1	1	4	4.00000000	3.65148372
2	1	4	8.25000000	4.11298756
3	1	4	4.25000000	4.64578662
4	1	4	6.75000000	2.06155281
1	2	4	9.50000000	2.08166600
2	2	4	3.50000000	4.35889894
3	2	4	5.25000000	3.50000000
4	2	4	2.00000000	1.82574186
1	3	4	2.50000000	2.88675135
2	3	4	8.50000000	2.38047614
3	3	4	5.00000000	3.55902608
4	3	4	4.50000000	2.08166600
1	4	4	5.25000000	4.57347424
2	4	4	6.50000000	3.10912635
3	4	4	5.50000000	3.69684550
4	4	4	4.75000000	3.40342964
1	5	4	5.25000000	3.50000000
2	5	4	4.00000000	3.16227766
3	5	4	1.75000000	1.25830574
4	5	4	3.50000000	2.88675135