Diseños factoriales

José Gabriel Palomo Sánchez
gabriel.palomo@upm.es

E.U.A.T.
U.P.M.

Julio de 2011

ÍNDICE

Diseños factoriales con dos factores I

En ocasiones, el experimentador está interesado en estudiar el efecto sobre la variable respuesta de varios factores.

- En estos casos la alternativa a la experimentación clásica, en la que se estudia el efecto de cada factor en experimentos independientes, es el **diseño factorial.**

Definición

Un diseño factorial con dos factores consiste en experimentar con todos los tratamientos que se obtienen al combinar cada nivel de un factor con los niveles del otro.

Ejemplo I

- Se desea analizar si el rendimiento de un determinado cultivo depende del tipo de semilla y de fertilizante empleados. Se dispone de dos semillas (A y B) y de tres fertilizantes 1, 2 y 3.

- El diseño factorial consta de $2 \times 3 = 6$ tratamientos, como se muestra en la tabla:

		SEMILLA A	B
FERT	1	y_{A1}	y_{B1}
	2	y_{A2}	y_{B2}
	3	y_{A3}	y_{B3}

Diseños factoriales con dos factores II

- En general, si existen dos factores de interés, el primero con K niveles y el segundo con J niveles distintos, el conjunto de datos del diseño factorial se resume en una tabla del tipo:

FACTOR 1

FACTOR 2		1	2	$\cdots$	K
	1	y_{11}	y_{21}	$\cdots$	y_{K1}
	2	y_{12}	y_{22}	$\cdots$	y_{K2}
	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
	J	y_{1J}	y_{2J}	$\cdots$	y_{KJ}

- El valor y_{ij} representa la observación realizada de la variable respuesta, en el nivel i del primer factor, y en el valor j del segundo.

Diseños factoriales con dos factores III

Las ventajas fundamentales del diseño factorial frente a la experimentación clásica son las siguientes:

- Eficiencia: con menos experimentos se estiman los efectos con la misma precisión.
- Mayor información, pues con la experimentación clásica no se exploran todas las combinaciones de los niveles de los factores.
- Mayor rango de validez de las conclusiones.

Diseños factoriales con dos factores IV. Objetivos

- Los objetivos de un diseño factorial con dos factores son los siguientes:

1. Contrastar si existen diferencias entre las medias de la variable respuesta en cada uno de los niveles del factor 1.
2. Contrastar si existen diferencias entre las medias de la variable respuesta en cada uno de los niveles del factor 2.
3. Contrastar si los dos factores **interaccionan**.

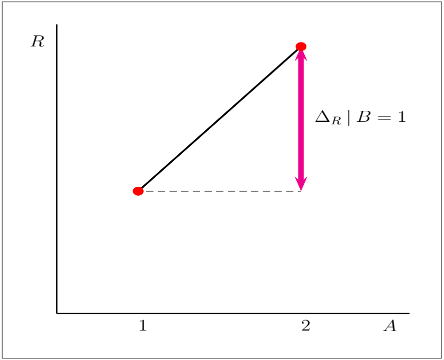
El concepto de interacción

- Empíricamente se comprueba que, en ocasiones, el efecto que, sobre la variable respuesta, se produce por el cambio de nivel en un determinado factor **depende** del nivel al que se encuentre el otro.

  Cuando esto sucede se dice que ambos factores **interaccionan**.

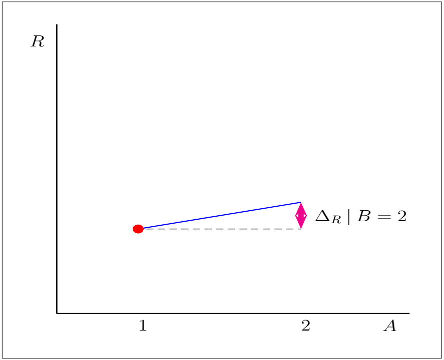
Ejemplo II

En el gráfico adjunto se observa el cambio en la respuesta producido al pasar del nivel 1 al 2 del factor A, cuando el factor B se encuentra en su nivel 1.



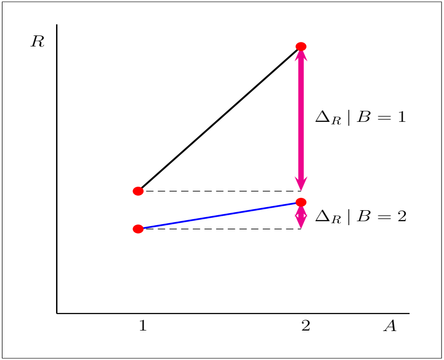
Ejemplo II

En este nuevo gráfico se observa el cambio en la respuesta producido al pasar del nivel 1 al 2 del factor A, cuando el factor B se encuentra en su nivel 2.



EJEMPLO II

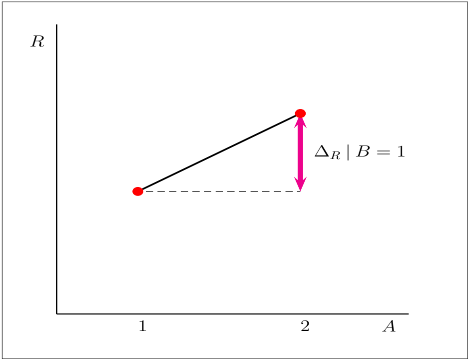
Al comparar ambos gráficos, se observa cómo el cambio producido en la respuesta al pasar del nivel 1 al 2 el factor A, cuando el factor B se encuentra en su nivel 1, es significativamente distinto del cambio producido cuando el factor B se encuentra a nivel 2.



- En este caso ambos factores interaccionan.

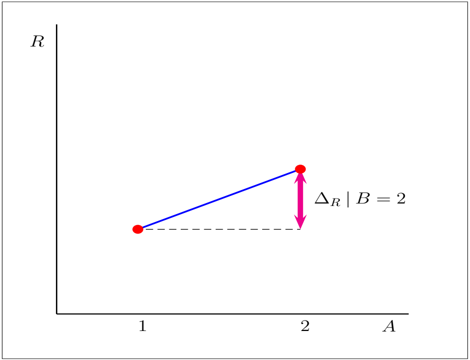
Ejemplo III

En el nuevo gráfico adjunto se observa, en otro caso, el cambio en la respuesta producido al pasar del nivel 1 al 2 del factor A, cuando el factor B se encuentra en su nivel 1.



Ejemplo III

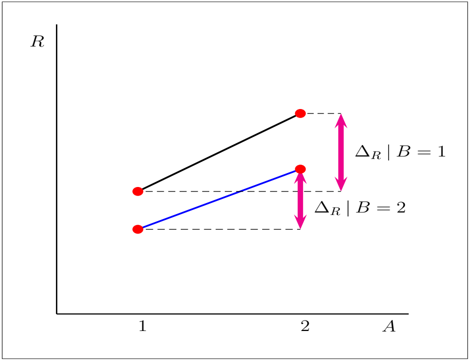
En este gráfico se observa el cambio en la respuesta producido, en este caso, al pasar del nivel 1 al 2 del factor A, cuando el factor B se encuentra en su nivel 2.



Ejemplo III

Al comparar ambos gráficos, se observa, en esta ocasión, cómo el cambio producido en la respuesta al pasar del nivel 1 al 2 el factor A, cuando el factor B se encuentra en su nivel 1, no es significativamente distinto del cambio producido cuando el factor B se encuentra a nivel 2.

- En este caso ambos factores **no** interaccionan.



Hipótesis del modelo I

La significatividad de los factores y de la interacción entre ellos se analiza bajo las siguientes hipótesis:

- La variable respuesta se puede descomponer en la forma:

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij}, \quad \text{donde:}$$

  - μ representa la media general.
  - α_i es la desviación de la media general debida a que la observación se realiza en el nivel i del primer factor.
  - β_j es la desviación de la media general debida a que la observación se realiza en el valor j del segundo factor.
  - $(\alpha\beta)_{ij}$ es la desviación de la media debida a la confluencia del nivel i del primer factor con el nivel j del segundo. Este término representa la interacción entre los dos factores.
  - e_{ij} representa el error aleatorio.

Hipótesis del modelo II

Además, se supone que:

- Para todos los valores de i y j, se tiene que:

$$e_{ij} \approx N(0, \sigma).$$

- Todos los e_{ij} son independientes entre sí.
- Por último, para evitar problemas de indeterminación en la estimación del modelo, se impondrá que:

$$\sum_{i=1}^{K} \alpha_i = \sum_{j=1}^{J} \beta_j = \sum_{i=1}^{K} (\alpha\beta)_{ij} = \sum_{j=1}^{J} (\alpha\beta)_{ij} = 0.$$

Consecuencias de las hipótesis del modelo I

Como consecuencia de las hipótesis se cumple que:

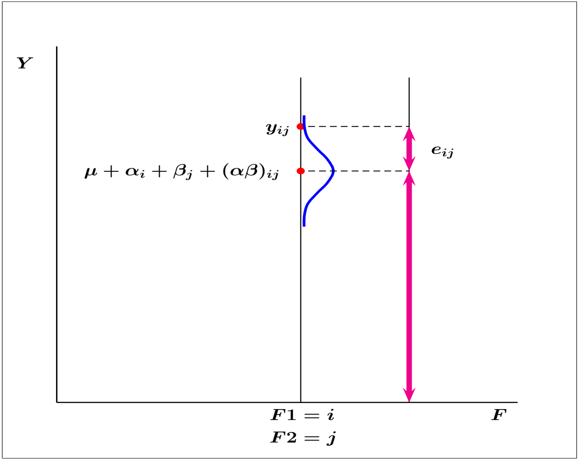
- La variable respuesta en los individuos sometidos al nivel i del primer factor, y al valor j del segundo, sigue una distribución:

$$(Y|F1 = i \wedge F2 = j) \approx N(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}, \sigma).$$

- Todos los y_{ij} son independientes entre sí.

Consecuencias de las hipótesis del modelo II

Gráficamente,



Observaciones I

1. El objetivo fundamental del estudio es analizar si existen diferencias significativas entre los valores de los α_i, de los de β_j, y de los de $(\alpha\beta)_{ij}$.

2. El número de parámetros del modelo que hay que estimar es:

$$1+(I-1)+(J-1)+(I-1)\times(J-1)+1 = I\times J+1 = n+1,$$

por lo que no hay datos suficientes para la estimación de todos los parámetros del modelo.

Observaciones II

- El problema de la falta de un número suficiente de observaciones para la estimación de todos los parámetros del modelo tiene distintas soluciones en la literatura.

- La solución más empleada para soslayar este problema, cuando es económicamente viable, es **replicar** el experimento varias veces.

- Una réplica de un experimento consiste en la realización completa, de nuevo, de dicho experimento. Lo que incluye la aleatorización del orden de los nuevos ensayos, etc.

Observaciones III

- En lo sucesivo, se considerará que el experimento realizado consta de un diseño completo replicado R veces:

FACTOR 1

		1	2	$\cdots$	K
FACTOR 2	1	y_{111}	y_{211}	$\cdots$	y_{K11}
	1	y_{112}	y_{212}	$\cdots$	y_{K12}
	1	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	1	y_{11R}	y_{21R}	$\cdots$	y_{K1R}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	J	y_{1J1}	y_{2J1}	$\cdots$	y_{KJ1}
	J	y_{1J2}	y_{2J2}	$\cdots$	y_{KJ2}
	J	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	J	y_{1JR}	y_{2JR}	$\cdots$	y_{KJR}

Estimación del modelo I

- Para estimar la media global μ, se utiliza la media de todas las observaciones:

$$\hat{\mu} = \bar{y}_{\bullet\bullet\bullet} = \frac{\sum_{i=1}^{K} \sum_{j=1}^{J} \sum_{r=1}^{R} y_{ijr}}{n}$$

- El valor de α_i, desviación de la media global atribuible a que la observación se realiza en el nivel i del factor 1, se estima por:

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

donde:

$$\bar{y}_{i\bullet\bullet} = \frac{\sum_{j=1}^{J} \sum_{r=1}^{R} y_{ijr}}{J \times R}$$

Estimación del modelo II. Observación

Debe observarse que $\hat{\alpha}_i$ se calcula como la diferencia entre la media de todas las observaciones realizadas en el nivel i del primer factor y la media global, lo que es coherente con el significado de α_i en el modelo.

Estimación del modelo III

La siguiente tabla resume la estimación de los efectos del factor 1:

FACTOR 1

		1	$\cdots$	K
FACTOR 2	1	y_{111}	$\cdots$	y_{K11}
		$\vdots$	$\vdots$	$\vdots$
		y_{11R}	$\cdots$	y_{K1R}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	J	y_{1J1}	$\cdots$	y_{KJ1}
		$\vdots$	$\vdots$	$\vdots$
		y_{1JR}	$\cdots$	y_{KJR}
	$\hat{\alpha}$	$\hat{\alpha}_1 = (\bar{y}_{1\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})$	$\cdots$	$\hat{\alpha}_K = (\bar{y}_{K\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})$

Estimación del modelo IV

- El valor de β_j, desviación de la media global atribuible a que la observación se realiza en el valor j del factor 2, se estima por:

$$\hat{\beta}_j = \bar{y}_{\bullet j \bullet} - \bar{y}_{\bullet\bullet\bullet},$$

donde:

$$\bar{y}_{\bullet j \bullet} = \frac{\sum_{i=1}^{K} \sum_{r=1}^{R} y_{ijr}}{K \times R}$$

Estimación del modelo V. Observación

De manera similar a como se hizo en el caso de $\hat{\alpha}_i$, debe observarse que $\hat{\beta}_j$ se calcula como la diferencia entre la media de todas las observaciones realizadas en el nivel j del segundo factor y la media global, lo que es coherente con el significado de β_j en el modelo.

Estimación del modelo VI

La siguiente tabla resume la estimación de los efectos del factor 2:

		FACTOR 1				
		1	2	$\cdots$	K	$\hat{\beta}$
FACTOR 2	1	y_{111} $\vdots$ y_{11R}	y_{211} $\vdots$ y_{21R}	$\cdots$ $\vdots$ $\cdots$	y_{K11} $\vdots$ y_{K1R}	$\hat{\beta}_1 = (\bar{y}_{\bullet 1 \bullet} - \bar{y}_{\bullet\bullet\bullet})$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	J	y_{1J1} $\vdots$ y_{1JR}	y_{2J1} $\vdots$ y_{2JR}	$\cdots$ $\vdots$ $\cdots$	y_{KJ1} $\vdots$ y_{KJR}	$\hat{\beta}_J = (\bar{y}_{\bullet J \bullet} - \bar{y}_{\bullet\bullet\bullet})$

Estimación del modelo VII

- La interacción $(\alpha\beta)_{ij}$, que es la desviación de la media general producida por la realización del ensayo experimental en los niveles i del factor 1 y j del factor 2, simultáneamente, viene dada por:

$$(\hat{\alpha\beta})_{ij} = \bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet},$$

donde,

$$\bar{y}_{ij\bullet} = \frac{\sum_{r=1}^{R} y_{ijr}}{R}.$$

Es decir, $\bar{y}_{ij\bullet}$ representa la media de todas las observaciones obtenidas en los niveles i del factor 1 y j del factor 2, simultáneamente.

Estimación del modelo VIII

- Por último, la varianza del error experimental, σ^2, se estima por la varianza residual:

$$\hat{s}_R^2 = \frac{\sum\sum\sum e_{ijr}^2}{K \times J \times (R-1)},$$

donde

$$e_{ijr} = y_{ijr} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + (\widehat{\alpha\beta})_{ij}) = y_{ijr} - \bar{y}_{ij\bullet}.$$

Análisis de la varianza I

Como en los casos anteriores, el análisis de la existencia de diferencias en la variable respuesta debidas a los niveles de los factores, o a la interacción, se puede realizar comparando la variabilidad explicada por cada uno de estos términos con la variabilidad total.

Así:

$$(y_{ijr} - \bar{y}_{\bullet\bullet\bullet}) = \hat{\alpha}_i + \hat{\beta}_j + (\widehat{\alpha\beta})_{ij} + e_{ijr}.$$

Análisis de la varianza II

Llamando variabilidad total y variabilidad explicada por el factor 1, respectivamente, a los términos:

$$VT = \sum_{i=1}^{K}\sum_{j=1}^{J}\sum_{r=1}^{R}(y_{ijr} - \bar{y}_{\bullet\bullet\bullet})^2$$

y

$$VE(\alpha) = \sum_{i=1}^{K}\sum_{j=1}^{J}\sum_{r=1}^{R}\hat{\alpha}_i^2 = JR\sum_{i=1}^{K}\hat{\alpha}_i^2.$$

Análisis de la varianza III

Y, de forma análoga, llamando variabilidad explicada por el factor 2, variabilidad explicada por la interacción, y variabilidad no explicada, respectivamente, a los términos:

$$VE(\beta) = \sum_{i=1}^{K}\sum_{j=1}^{J}\sum_{r=1}^{R}\hat{\beta}_j^2 = KR\sum_{j=1}^{J}\hat{\beta}_j^2.$$

$$VE(\alpha\beta) = \sum_{i=1}^{K}\sum_{j=1}^{J}\sum_{r=1}^{R}(\hat{\alpha\beta})_{ij}^2 = R\sum_{i=1}^{K}\sum_{j=1}^{J}(\hat{\alpha\beta})_{ij}^2.$$

y

$$VNE = \sum_{i=1}^{K}\sum_{j=1}^{J}\sum_{r=1}^{R}e_{ijr}^2.$$

Análisis de la varianza IV

Se puede demostrar que:

$$VT = VE(\alpha) + VE(\beta) + VE(\alpha\beta) + VNE.$$

- El término $VE(\alpha)$ depende sólo de la variabilidad entre los distintos $\hat{\alpha}_i$.
- El término $VE(\beta)$ depende sólo de la variabilidad entre los distintos $\hat{\beta}_j$.
- El término $VE(\alpha\beta)$ depende sólo de la variabilidad entre los distintos $(\hat{\alpha\beta})_{ij}$.
- El término VNE es una medida de la variabilidad de los residuos.

Análisis de la varianza V

Observaciones:

- En la medida en que $VE(\alpha)$ sea grande en relación con VNE, habrá evidencia de diferencia entre los valores de los efectos sobre las respuesta de los distintos niveles del factor 1.
- En la medida en que $VE(\beta)$ sea grande en relación con VNE, habrá evidencia de diferencia entre los valores de los efectos sobre las respuesta de los distintos niveles del factor 2.
- En la medida en que $VE(\alpha\beta)$ sea grande en relación con VNE, habrá evidencia de la existencia de interacción entre los dos factores.

Análisis de la varianza VI

- Discutir la magnitud de $VE(\alpha)$, $VE(\beta)$ y de $VE(\alpha\beta)$, requiere analizar sus distribuciones de probabilidad.

Teorema I

1. Si se verifica la hipótesis:

$$\alpha_1 = \alpha_2 = \cdots = \alpha_K = 0,$$

la variable $VE(\alpha)/\sigma^2$ se distribuye como una χ^2_{K-1}.

2. Si se verifica la hipótesis:

$$\beta_1 = \beta_2 = \cdots = \beta_J = 0,$$

la variable $VE(\beta)/\sigma^2$ se distribuye como una χ^2_{J-1}.

Análisis de la varianza VII

Teorema II

- Si se verifica que

$$(\alpha\beta)_{ij} = 0, \text{para todos los valores de} \quad i \quad \text{y} \quad j,$$

la variable $VE(\alpha\beta)/\sigma^2$ se distribuye como una $\chi^2_{(K-1)(J-1)}$.

Teorema III

La variable VNE/σ^2 se distribuye, en cualquier caso, como una $\chi^2_{KJ(R-1)}$ y es independiente de las tres distribuciones anteriores.

Análisis de la varianza VIII

Consecuencia I

Si se verifica la hipótesis:

$$\alpha_1 = \alpha_2 = \cdots = \alpha_K = 0$$

la variable

$$\frac{\frac{VE(\alpha)}{\sigma^2(K-1)}}{\frac{VNE}{\sigma^2 KJ(R-1)}} \longrightarrow F_{(K-1;KJ(R-1))}$$

Análisis de la varianza IX

Consecuencia II

Si se verifica la hipótesis:

$$\beta_1 = \beta_2 = \cdots = \beta_J = 0$$

la variable

$$\frac{\frac{VE(\beta)}{\sigma^2 (J-1)}}{\frac{VNE}{\sigma^2 KJ(R-1)}} \longrightarrow F_{(J-1;KJ(R-1))}$$

Análisis de la varianza X

Consecuencia III

Si se verifica la hipótesis:

$$(\alpha\beta)_{ij} = 0, \text{ para todos los valores de } \ i \ \text{ y } \ j,$$

la variable

$$\frac{\frac{VE(\alpha\beta)}{\sigma^2(K-1)(J-1)}}{\frac{VNE}{\sigma^2 KJ(R-1)}} \longrightarrow F_{((K-1)(J-1);KJ(R-1))}$$

Análisis de la varianza XI

- Llamando $\hat{s}_e^2(\alpha)$ al valor de $VE(\alpha)$/(K-1),
- $\hat{s}_e^2(\beta)$ al valor de $VE(\beta)$/(J-1),
- y $\hat{s}_R^2$ al valor de VNE/(KJ(R-1)).

Se tiene que cuando $\alpha_1 = \alpha_2 = \cdots = \alpha_K = 0$:

$$\frac{\hat{s}_e^2(\alpha)}{\hat{s}_R^2} \longrightarrow F_{((K-1);KJ(R-1))}.$$

Y cuando $\beta_1 = \beta_2 = \cdots = \beta_J = 0$:

$$\frac{\hat{s}_e^2(\beta)}{\hat{s}_R^2} \longrightarrow F_{((J-1);KJ(R-1))}.$$

Análisis de la varianza XII

- Análogamente, llamando $\hat{s}_e^2(\alpha\beta)$ al valor de $VE(\alpha\beta)$/(K-1)(J-1):

Se tiene que cuando $(\alpha\beta)_{ij} = 0,$ para todos los valores de i y j,

$$\frac{\hat{s}_e^2(\alpha\beta)}{\hat{s}_R^2} \longrightarrow F_{((K-1)(J-1);KJ(R-1))}.$$

Análisis de la varianza XIII. El test de la F I

- Empleando los resultados anteriores, para discutir el contraste:

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_K = 0,$$

frente a

$$H_1 : \text{Existe al menos un } \alpha_i \text{ tal que } \alpha_i \neq 0,$$

basta con analizar el valor del estadístico

$$F = \frac{\hat{s}_e^2(\alpha)}{\hat{s}_R^2}$$

Análisis de la varianza XIV. El test de la F II

- De manera que, si se denomina F_α al valor tal que

$$P(F_{(K-1;KJ(R-1))} > F_\alpha) = \alpha,$$

cuando

$$F = \frac{\hat{s}_e^2(\alpha)}{\hat{s}_R^2} < F_\alpha$$

se aceptará la hipótesis nula, ($\alpha_i = 0$, para todo i), que se rechazará en caso contrario.

Observación: Nótese que el test de la F es un contraste unilateral, en coherencia con la hipótesis que se contrasta.

Análisis de la varianza XV. El test de la F III

- Análogamente, para discutir el contraste:

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_J = 0,$$

frente a

$$H_1 : \text{Existe al menos un } \beta_j \text{ tal que } \beta_j \neq 0,$$

basta con analizar el valor del estadístico

$$F = \frac{\hat{s}_e^2(\beta)}{\hat{s}_R^2}$$

Análisis de la varianza XVI. El test de la F IV

- De manera que, si se denomina F_α al valor tal que

$$P(F_{(J-1;KJ(R-1))} > F_\alpha) = \alpha,$$

cuando

$$F = \frac{\hat{s}_e^2(\beta)}{\hat{s}_R^2} < F_\alpha$$

se aceptará la hipótesis nula, ($\beta_j = 0$, para todo j), que se rechazará en caso contrario.

Observación: Nótese que el test de la F es un contraste unilateral, en coherencia con la hipótesis que se contrasta.

Análisis de la varianza XVII. El test de la F V

- De la misma manera, para discutir el contraste:

$$H_0 : (\alpha\beta)_{ij} = 0, \quad \text{para todos los valores de } i \text{ y } j,$$

frente a

$$H_1 : \text{Existe al menos un } (\alpha\beta)_{ij} \text{ tal que } (\alpha\beta)_{ij} \neq 0,$$

basta con analizar el valor del estadístico

$$F = \frac{\hat{s}_e^2(\alpha\beta)}{\hat{s}_R^2},$$

comparándolo con el F_α apropiado.

Análisis de la varianza XVIII. La tabla ADEVA

Los resultados de los tests de la F se resumen en la

Tabla ADEVA

Fuentes de variac.	**Suma de cuadrados**	**Grados de libertad**	**Varian.**	F	p-v.
Factor 1	$JR\sum\hat{\alpha}_i^2$	$K-1$	$\hat{s}_e^2(\alpha)$	$\frac{\hat{s}_e^2(\alpha)}{\hat{s}_R^2}$	$p(\alpha)$
Factor 2	$KR\sum\hat{\beta}_j^2$	$J-1$	$\hat{s}_e^2(\beta)$	$\frac{\hat{s}_e^2(\beta)}{\hat{s}_R^2}$	$p(\beta)$
Interacción	$R\sum\sum(\hat{\alpha\beta})_{ij}^2$	(K-1)(J-1)	$\hat{s}_e^2(\alpha\beta)$	$\frac{\hat{s}_e^2(\alpha\beta)}{\hat{s}_R^2}$	$p(\alpha\beta)$
Residual	$\sum\sum\sum(e_{ijr})^2$	$KJ(R-1)$	$\hat{s}_R^2$		
Total	ns_Y^2	$n-1$	$\hat{s}_Y^2$		

Análisis de la varianza XIX. Diagnosis y validación del modelo

- Como en los casos anteriores, una vez realizado el análisis de la varianza, antes de emplear las conclusiones allí extraídas, es necesario verificar las hipótesis del modelo.
- Esta verificación se lleva a cabo por medio del análisis de los residuos.
  - La discusión de la normalidad se realiza a través del papel probabilístico normal.
  - La comprobación de la homocedasticidad requiere gráficos de los residuos frente a los distintos valores de los factores y frente a los valores previstos por el modelo.

Inferencia para los parámetros del modelo I

Una vez realizada la diagnosis del modelo, puede ser necesario hacer inferencia respecto de los parámetros del mismo.

- La inferencia respecto del valor de α_i se puede hacer teniendo en cuenta que:

$$\frac{\bar{y}_{i\bullet\bullet} - \alpha_i}{\hat{s}_R/\sqrt{JR}} \longrightarrow t_{KJ(R-1)}$$

- La comparación de dos desviaciones de la media general provocadas por dos valores distintos del factor 1, α_i y α_j, se puede realizar si se tiene en cuenta que:

$$\frac{(\bar{y}_{i\bullet\bullet} - \bar{y}_{j\bullet\bullet}) - (\alpha_i - \alpha_j)}{\hat{s}_R\sqrt{\frac{2}{JR}}} \longrightarrow t_{KJ(R-1)}$$

Inferencia para los parámetros del modelo II

Análogamente,

- La inferencia respecto del valor de β_j se puede hacer teniendo en cuenta que:

$$\frac{\bar{y}_{\bullet j \bullet} - \beta_j}{\hat{s}_R / \sqrt{KR}} \longrightarrow t_{KJ(R-1)}$$

- La comparación de dos desviaciones de la media general provocadas por dos valores distintos del bloque, β_i y β_j, se puede realizar si se tiene en cuenta que:

$$\frac{(\bar{y}_{\bullet i \bullet} - \bar{y}_{\bullet j \bullet}) - (\beta_i - \beta_j)}{\hat{s}_R \sqrt{\frac{2}{KR}}} \longrightarrow t_{KJ(R-1)}$$

Inferencia para los parámetros del modelo III

- La inferencia respecto de $(\alpha\beta)_{ij}$ se realiza considerando que

$$\frac{\bar{y}_{ij\bullet} - (\alpha\beta)_{ij}}{\hat{s}_R/\sqrt{R}} \longrightarrow t_{KJ(R-1)}$$

- La inferencia respecto de σ^2 se realiza teniendo en cuenta la siguiente distribución:

$$\frac{KJ(R-1)\hat{s}_R^2}{\sigma^2} \longrightarrow \chi^2_{KJ(R-1)}$$

Inferencia para los parámetros del modelo IV. Observaciones

- La extensión a modelos con más de dos factores se realiza de una manera natural.
- En el caso en que no se realicen réplicas, una técnica habitual es considerar nulas las interacciones de orden superior o igual a tres, con lo que se disminuye el número de parámetros del modelo a estimar.
- El lector interesado en los puntos anteriores puede consultar el libro de Daniel Peña **Regresión y diseño de experimentos**. Alianza editorial. (2002)