

# Capítulo 7

## Modelos de Análisis de Varianza

En los capítulos anteriores se ha estudiado un tipo particular de modelo lineal, el cual incluye variables explicativas de tipo numérico. En el presente capítulo se presentará otro tipo de modelos lineales, en los cuales las variables explicativas son *categorías*, es decir, indican la pertenencia de un sujeto a alguna categoría. Ejemplos de este tipo de variables son: sexo, condición socioeconómica, nivel educativo, tipo de alimentación, etc.

### 7.1 Análisis de Varianza con un criterio de clasificación

Consideremos la siguiente situación:

**Ejemplo 7.1** *La tabla 7.1 corresponde a los tiempos de coagulación (en segundos) para muestras de sangre tomadas de 24 ratones de laboratorio, los cuales han recibido 4 dietas diferentes que llamaremos A, B, C y D. ¿Existe evidencia para pensar que la dieta a la cual ha sido sometido el animal afecta el tiempo de coagulación de su sangre?*

En la figura 7.1 se presentan los diagramas de caja correspondientes a cada una de las dietas. A partir de este gráfico, pareciera que el tiempo más largo de coagulación corresponde a animales que han estado sometidos a la dieta C, mientras que los tiempos más cortos parecen corresponder a las dietas A y D. Ahora bien, ¿son estas diferencias suficientes para afirmar que existen diferencias reales entre las distintas dietas?

Dieta			
A	B	C	D
62	63	68	56
60	67	66	62
63	71	71	60
59	64	67	61
	65	68	63
	66	68	64
			63
			59

Tabla 7.1: Tiempos de coagulación para muestras de sangre tomadas de ratones sometidos a 4 dietas diferentes (Ejemplo 7.1)

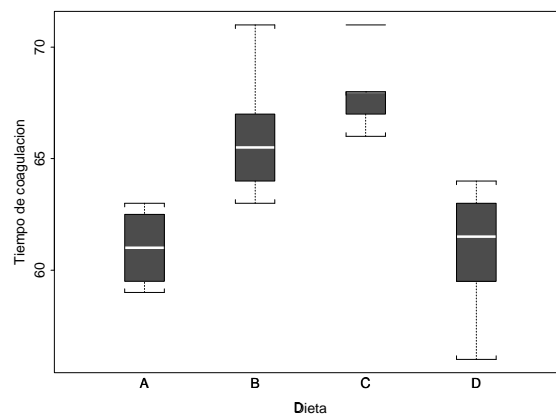


Figura 7.1: Diagramas de caja para los tiempos de coagulación.

El ejemplo anterior es un caso particular del siguiente tipo de problema: se tienen  $k$  grupos de datos, cada uno de los cuales tiene  $n_i$  elementos,  $i = 1, \dots, k$ .

$y_{11}$	$y_{21}$	$\dots$	$y_{k1}$
$y_{12}$	$y_{22}$	$\dots$	$y_{k2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y_{1n_1}$	$y_{2n_2}$	$\dots$	$y_{kn_k}$

Se desea saber si existen diferencias entre las medias de los distintos grupos. Un modelo razonable para esta situación es:

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad , \quad i = 1, \dots, k \\ j = 1, \dots, n_i$$

Este modelo puede escribirse en la forma  $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ , donde:

$$\mathbf{Y} = \begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1,n_1} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2,n_2} \\ \vdots \\ y_{k1} \\ y_{k2} \\ \vdots \\ y_{k,n_k} \end{pmatrix}, \quad X = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2,n_2} \\ \varepsilon_{1n_1} \\ \vdots \\ \varepsilon_{k1} \\ \varepsilon_{k2} \\ \vdots \\ \varepsilon_{k,n_k} \end{pmatrix}$$

Si se supone que los errores son independientes, tienen distribución normal media cero, varianza común  $\sigma^2$  ( $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 I)$ ), estamos nuevamente ante un modelo lineal como los estudiados en los capítulos anteriores. Es decir, podemos aplicar todos los métodos antes desarrollados (estimación por mínimos cuadrados, pruebas  $F$ , etc.) a la inferencia con este modelo.

No es difícil ver que los estimadores de mínimos cuadrados para los parámetros del modelo,  $\mu_1, \dots, \mu_k$ , son los promedios correspondientes a cada uno de los grupos,

$$\hat{\mu}_i = \bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$$

Se trata, sin duda, de un resultado que resulta atractivo desde el punto de vista intuitivo.

Para el ejemplo 7.1, esto significa tomar

$$\begin{aligned}\hat{\mu}_A &= 61 \\ \hat{\mu}_B &= 66 \\ \hat{\mu}_C &= 68 \\ \hat{\mu}_D &= 61\end{aligned}$$

Ahora bien, deseamos saber si estas medias son realmente diferentes. Para ello, probaremos las hipótesis:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \text{ vs } H_1 : \text{algún } \mu_i \text{ distinto}$$

Probar las hipótesis anteriores equivale a comparar los modelos:

$$\begin{aligned}(1) \quad y_{ij} &= \mu + \varepsilon_{ij} \text{ vs} \\ (2) \quad y_{ij} &= \mu_i + \varepsilon_{ij}\end{aligned}$$

No es difícil ver que los modelos anteriores son anidados (al sumar las columnas de la matriz de diseño del modelo (2) obtenemos un vector de unos, el cual corresponde a la matriz de diseño del modelo (1)). Por consiguiente, esta prueba puede hacerse por medio de la tabla de análisis de varianza correspondiente al modelo (2). Llamando  $N = \sum_{i=1}^k n_i$ , dicha tabla de análisis de varianza puede escribirse de la siguiente manera

Fuente	gl	SS	MS	F
Tratamientos	$k - 1$	$SSTr = \sum_{i=1}^k n_i (\bar{y}_{i.} - \bar{y}_{..})^2$	$MSTr = \frac{SSTr}{k-1}$	$\frac{MSTr}{MSE}$
Error	$N - k$	$SSE = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$	$MSE = \frac{SSE}{N-k}$	
Total	$N - 1$	$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$		

Igual que antes, rechazamos  $H_0$  si  $F > F_{k-1, N-k}^{\alpha}$ .

En el ejemplo 7.1, la tabla de análisis de varianza es la siguiente:

Fuente	gl	SS	MS	F
Tratamientos	3	228	76.0	13.57143
Error	20	112	5.6	
Total	23	340		

Al comparar el valor calculado de la F con el valor tabulado para  $\alpha = 0.05$ ,  $F_{3,20}^{0.05} = 3.098$ , llegamos a la conclusión de que los datos proporcionan evidencia suficiente para rechazar la hipótesis de igualdad de las medias de los cuatro grupos. Por lo tanto, podemos considerar que el tipo de dieta afecta el tiempo de coagulación de la sangre.

Observemos que el procedimiento anterior sólo nos permite decidir si *existe* alguna media distinta, pero no permite identificar *cuáles* de ellas son distintas. En la siguiente sección, estudiaremos el problema de comparar las medias de cada grupo.

Nuevamente, debe tomarse en cuenta que cualquier conclusión obtenida a partir de un modelo de este tipo se basa en las suposiciones de que los residuos tienen distribución normal con media cero y varianza común. Es decir, es necesario hacer un análisis cuidadoso de los residuos del modelo con el fin de determinar si estas suposiciones son ciertas, y si los resultados obtenidos son válidos.

La figura 7.2 muestra los gráficos de residuos para el ejemplo 7.1. Según estos gráficos, no existen razones para dudar de la validez de las suposiciones del modelo.

Para facilitar la observación de las diferencias entre los grupos, frecuentemente se emplea la siguiente reparametrización del modelo anterior:

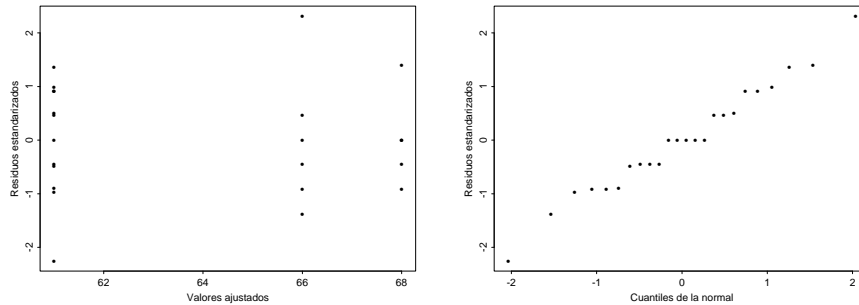


Figura 7.2: Gráficos de residuos para el modelo de análisis de varianza con un criterio de clasificación ajustado a los datos del Ejemplo 7.1

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad , \quad \begin{array}{l} i = 1, \dots, k \\ j = 1, \dots, n_i \end{array}$$

Nótese que, si bien este modelo es equivalente al anterior ( $\mu_i = \mu + \alpha_i$ , con lo cual estamos proyectando sobre el mismo subespacio), tiene un parámetro adicional. Ésto implica que las ecuaciones normales correspondientes a esta parametrización van a tener infinitas soluciones. Cuando este tipo de situación se produce, diremos que los parámetros del modelo *no son identificables*.

Para resolver este problema, es necesario imponer una restricción sobre los  $\alpha_i$  con el fin de obtener una única solución. Muchas restricciones son posibles. Una de las más usadas es

$$\sum_{i=1}^k n_i \alpha_i = 0$$

Cuando el experimento es balanceado (es decir, cuando todos los grupos tienen el mismo número de datos), esta restricción se transforma en  $\sum_{i=1}^k \alpha_i = 0$ .

Con esta restricción,  $\mu$  puede interpretarse como una media común a todos los grupos (obsérvese que  $\mu = \sum_{i=1}^k \left(\frac{n_i}{N}\right) \mu_i / k$ , promedio de las  $\mu_i$  ponderado por el tamaño del grupo), y  $\alpha_i$  representa la diferencia entre la media general  $\mu$  y la media del grupo  $\mu_i$  ( $\alpha_i = \mu_i - \mu$ ).

Usando esta restricción, se obtienen los siguientes estimadores de mínimos cuadrados:

$$\begin{aligned}\hat{\mu} &= \bar{y}_{..} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}}{N} \\ \hat{\alpha}_i &= \bar{y}_{i.} - \bar{y}_{..}\end{aligned}$$

En nuestro ejemplo, los parámetros estimados son:

$$\begin{aligned}\hat{\mu} &= 64 \\ \hat{\alpha}_1 &= -3 \\ \hat{\alpha}_2 &= 2 \\ \hat{\alpha}_3 &= 4 \\ \hat{\alpha}_4 &= -3\end{aligned}$$

Obsérvese, además, que esta parametrización *no altera la tabla de análisis de varianza*. En efecto, independientemente de la parametrización usada estamos proyectando sobre el mismo subespacio, de modo que los vectores que se comparan en la prueba  $F$  son iguales.

## 7.2 Comparación de las medias de los grupos

Una vez que se ha rechazado la hipótesis de que todos los grupos son iguales, es natural preguntarse cuáles de los grupos son iguales, y cuales son distintos entre sí. Es decir, deseamos probar las hipótesis:

$$H_0 : \mu_i = \mu_j \text{ vs } H_1 : \mu_i \neq \mu_j \quad (7.1)$$

Para cada par de medias  $\mu_i, \mu_j$ .

Estudiaremos a continuación cómo realizar esta prueba.

### 7.2.1 Mínima Diferencia Significativa

Una manera de realizar la prueba de hipótesis indicada en 7.1 es la siguiente: si  $H_0$  es cierta, puede demostrarse que

$$t = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{MSE \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}} \sim t_{N-k}$$

Por lo tanto, podemos usar la distribución  $t$  con  $n - k$  grados de libertad como referencia para esta prueba. Si se selecciona un nivel  $\alpha$  para la prueba, rechazaremos  $H_0$  cuando  $|t| > t_{N-k}^{\alpha/2}$ . Esto equivale a rechazar  $H_0$  cuando

$$|\bar{y}_i - \bar{y}_j| > t_{N-k}^{\alpha/2} \sqrt{MSE \left( \frac{1}{n_i} + \frac{1}{n_j} \right)},$$

es decir, cuando la diferencia entre ambas medias sea “grande” con respecto a la referencia escogida. Nótese que si todos los grupos tienen el mismo tamaño, es decir, si  $n_i = n$ , todas las diferencias entre medias serán comparadas contra el mismo número, el cual será la “mínima diferencia significativa”.

Para el ejemplo 7.1, obtenemos los siguientes resultados al aplicar el método de mínima diferencia significativa con  $\alpha = 0.05$  ( $t_{20}^{0.025} = 2.086$ ):

$$\begin{aligned} |\bar{y}_A - \bar{y}_B| &= 5 > 2.086 \times \sqrt{5.6 \left( \frac{1}{4} + \frac{1}{6} \right)} = 3.2 & (*) \\ |\bar{y}_A - \bar{y}_C| &= 7 > 2.086 \times \sqrt{5.6 \left( \frac{1}{4} + \frac{1}{6} \right)} = 3.2 & (*) \\ |\bar{y}_A - \bar{y}_D| &= 0 < 2.086 \times \sqrt{5.6 \left( \frac{1}{4} + \frac{1}{8} \right)} = 3.02 \\ |\bar{y}_B - \bar{y}_C| &= 2 < 2.086 \times \sqrt{5.6 \left( \frac{1}{6} + \frac{1}{6} \right)} = 2.85 \\ |\bar{y}_B - \bar{y}_D| &= 5 > 2.086 \times \sqrt{5.6 \left( \frac{1}{6} + \frac{1}{8} \right)} = 2.67 & (*) \\ |\bar{y}_C - \bar{y}_D| &= 7 > 2.086 \times \sqrt{5.6 \left( \frac{1}{6} + \frac{1}{8} \right)} = 2.67 & (*) \end{aligned}$$

Se han marcado con un asterisco aquellos pares de medias para los cuales puede rechazarse la hipótesis de igualdad.

Podemos resumir los resultados anteriores colocando las medias ordenadas de menor a mayor, y subrayando aquellas para las cuales no puede rechazarse la hipótesis de igualdad, de la siguiente manera:

$$\underline{\bar{y}_A. \quad \bar{y}_D.} \quad \underline{\bar{y}_B. \quad \bar{y}_C.}$$

Según estos resultados, las dietas A y D se comportan en forma similar, así como las dietas B y C. Adicionalmente, las dietas B y C parecen retardar el tiempo de coagulación de la sangre en los ratones.



### 7.2.2 Intervalos de Confianza

Otra manera en la cual pueden probarse las hipótesis de (7.1) es calculando *intervalos de confianza* para cada una de las medias.

En general, supongamos que  $X_1, \dots, X_n$  son una muestra de una distribución  $f(x|\theta)$ . Un *intervalo de confianza* del  $(1 - \alpha) \times 100\%$  para  $\theta$  basado en la muestra  $X_1, \dots, X_n$  se define como una pareja de estadísticos (funciones de los datos)  $L(\mathbf{x}) < U(\mathbf{x})$  tales que

$$P(L(\mathbf{X}) < \theta < U(\mathbf{X})) = (1 - \alpha)$$

La expresión anterior nos tienta a pensar que “la probabilidad de que  $\theta$  esté entre  $L(\mathbf{X})$  y  $U(\mathbf{X})$  es del  $(1 - \alpha)\%$ . Sin embargo, al pensarlo un poco mejor nos damos cuenta de que  $\theta$  *no es una variable aleatoria* y por lo tanto carece de sentido hablar de la probabilidad de que pertenezca a algún intervalo. La probabilidad puede usarse porque  $X_1, \dots, X_n$ , antes de ser observadas, son variables aleatorias.

La interpretación correcta de un intervalo de confianza es, por tanto, la siguiente: *para el  $(1 - \alpha)\%$  de las posibles muestras de tamaño  $n$  de la distribución  $f(x|\theta)$ , el intervalo así calculado contendrá al verdadero valor del parámetro  $\theta$ .*

Consideremos, por ejemplo, el caso en el cual tenemos un conjunto de datos  $X_1, \dots, X_n$  provenientes de una distribución  $N(\mu, \sigma^2)$ , con  $\sigma^2$  desconocido. En este caso, se probó en el capítulo 3 que

$$\frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}} \sim t_{n-1}$$

Podemos afirmar, entonces, que

$$P\left(-t_{n-1}^{\alpha/2} < \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}} < t_{n-1}^{\alpha/2}\right) = P\left(\bar{X} - t_{n-1}^{\alpha/2} \sqrt{\frac{S^2}{n}} < \mu < \bar{X} + t_{n-1}^{\alpha/2} \sqrt{\frac{S^2}{n}}\right) = 1 - \alpha$$

y por tanto,  $\left(\bar{X} - t_{n-1}^{\alpha/2} \sqrt{\frac{S^2}{n}}, \bar{X} + t_{n-1}^{\alpha/2} \sqrt{\frac{S^2}{n}}\right)$  es un intervalo de confianza del  $(1 - \alpha) \times 100\%$  para  $\mu$ .

Para ilustrar el comportamiento de este intervalo de confianza, se generaron 10 muestras de 30 datos cada una a partir de una distribución Normal(0,1). En la figura 7.3 se muestran la densidad real de los datos y los resultados de calcular el intervalo de confianza del 90% para la media correspondiente a cada una de las 10 muestras. Obsérvese que el intervalo de confianza correspondiente a la muestra 7 *no contiene al verdadero valor de la media*.

En general, no podemos saber si el intervalo de confianza que hemos calculado contiene al verdadero valor del parámetro. Confiamos en que así sea, pues ser ía “mala suerte”

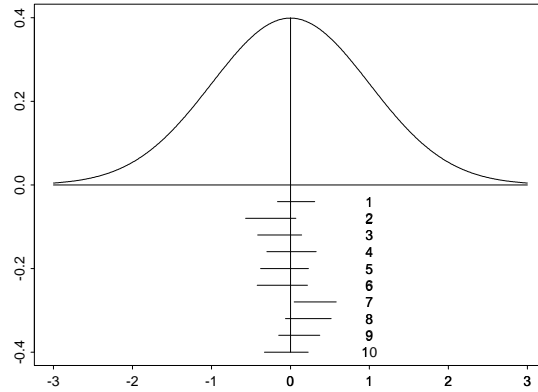


Figura 7.3: Intervalos de confianza para la media de una normal(0,1), calculados a partir de 10 muestras simuladas de tamaño 30.

que nuestra muestra perteneciera al pequeño porcentaje de las muestras que generan un intervalo de confianza que no contiene al verdadero valor del parámetro, pero es sin embargo necesario estar conscientes de que ésto puede suceder.

Para aplicar las ideas de intervalos de confianza al problema de comparar las medias de los distintos grupos en un modelo de análisis de varianza con un criterio de clasificación, comencemos por calcular el intervalo de confianza para cada una de estas medias. Usando las propiedades de los estimadores de mínimos cuadrados presentadas en el capítulo 4, podemos afirmar que

$$t = \frac{\bar{y}_i - \mu_i}{\sqrt{\frac{MSE}{n_i}}} \sim t_{N-k}$$

y por tanto

$$P \left( -t_{N-k}^\alpha < \frac{\bar{y}_i - \mu_i}{\sqrt{\frac{MSE}{n_i}}} < t_{N-k}^\alpha \right) = 1 - \alpha$$

De allí que los intervalos de confianza para los  $\mu_i$  estén dados por:

$$\bar{y}_i - t_{N-k}^{\alpha/2} \sqrt{\frac{MSE}{n_i}} < \mu_i < \bar{y}_i + t_{N-k}^{\alpha/2} \sqrt{\frac{MSE}{n_i}}$$

Para probar la hipótesis  $H_0 : \mu_i = \mu_j$  observamos si los intervalos para  $\mu_i$  y  $\mu_j$  se superponen. Cuando ésto no sucede, rechazamos la hipótesis de igualdad de ambas medias.

Los intervalos de confianza para los datos del ejemplo 7.1 son:

$$\begin{aligned}\mu_A: & (58.53, 63.47) \\ \mu_B: & (63.98, 68.02) \\ \mu_C: & (65.98, 70.02) \\ \mu_D: & (59.25, 62.75)\end{aligned}$$

La figura 7.4 muestra estos cuatro intervalos de confianza superpuestos. Es fácil ver, entonces que los intervalos de confianza para las dietas B y C se superponen, y producen tiempos de coagulación mayores que que las dietas A y D, cuyos intervalos de confianza también se superponen. Obtenemos, por tanto, las mismas conclusiones que cuando se usaba el método de mínima diferencia significativa.

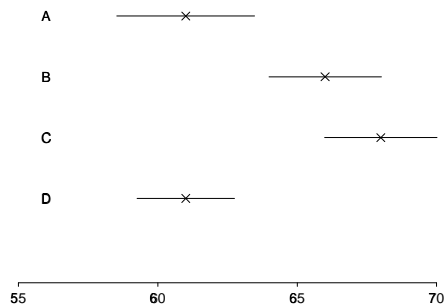


Figura 7.4: Intervalos de confianza para los tiempos de coagulación medios correspondientes a las cuatro dietas, ejemplo 7.1.

Nótese que cuando se rechaza la hipótesis de igualdad de las medias porque no se superponen los intervalos de confianza, también deberá rechazarse usando mínima diferencia significativa. En efecto, supongamos que se ha rechazado la hipótesis  $H_0 : \mu_i = \mu_j$  usando intervalos de confianza (es decir, los intervalos no se superponen). Suponiendo que  $\bar{y}_i < \bar{y}_j$ , ésto quiere decir que

$$\bar{y}_i + t_{N-k}^{\alpha/2} \sqrt{\frac{MSE}{n_i}} < \bar{y}_j + t_{N-k}^{\alpha/2} \sqrt{\frac{MSE}{n_j}}$$

y por tanto

$$|\bar{y}_i - \bar{y}_j| > t_{N-k}^{\alpha/2} \left( \sqrt{\frac{MSE}{n_i}} + \sqrt{\frac{MSE}{n_j}} \right) > t_{N-k}^{\alpha/2} \left( \sqrt{MSE \left( \frac{1}{n_i} + \frac{1}{n_j} \right)} \right)$$

Es decir, si se aplica el método de mínima diferencia significativa la hipótesis  $H_0 : \mu_i = \mu_j$  también deberá ser rechazada.

### 7.2.3 Observaciones

Los métodos antes mencionados (mínima diferencia significativa e intervalos de confianza) tienen el inconveniente de ser muy “liberales”, es decir, de señalar como diferentes muchos pares de medias, sobre todo cuando es necesario realizar un número grande de comparaciones. Se han propuesto distintos métodos para solucionar este inconveniente; no hablaremos de ellos en esta guía, pero pueden ser encontrados en un gran número de los textos que cubren el tópico de análisis de varianza.

## 7.3 Transformaciones para estabilizar la varianza

Hemos visto anteriormente que al analizar los residuos obtenidos al ajustar un modelo lineal pueden evidenciarse problemas con los residuos. Uno de estos problemas puede ser la heterocedasticidad de los residuos, es decir, la evidencia de que los errores no tienen una varianza común. Este problema puede resolverse en general realizando una transformación adecuada de los datos, y en el caso del modelo de análisis de varianza es posible estimar empíricamente cuál es esa transformación.

**Ejemplo 7.2** *Se mide el contenido de plata (en porcentaje) en diferentes monedas encontradas en Chipre y provenientes del reinado del rey Manuel I Comnenus (1143-1180). Se han encontrado nueve monedas de una primera acuñación, siete de una segunda acuñación posterior, cuatro de una tercera y finalmente siete de una cuarta acuñación. Se desea saber si existen diferencias significativas entre el contenido de plata de*

*las monedas acuñadas al principio y al final del reinado del rey Manuel. Los datos se muestran en la tabla 7.2*

Acuñaación			
1	2	3	4
5.9	6.9	4.9	5.3
6.8	9.0	5.5	5.6
6.4	6.6	4.6	5.5
7.0	8.1	4.5	5.1
6.6	9.3		6.2
7.7	9.2		5.8
7.2	8.6		5.8
6.9			
6.2			

Tabla 7.2: Porcentaje de plata en la elaboración de monedas encontradas en Chipre, correspondientes al reinado de Manuel I (Ejemplo 7.2)

Nuevamente, parece adecuado emplear un modelo de análisis de varianza con un criterio de clasificación. Al ajustar dicho modelo obtenemos los siguientes estimadores de las medias:

$$\mu_1 = 6.74$$

$$\mu_2 = 8.24$$

$$\mu_3 = 4.88$$

$$\mu_4 = 5.61$$

y la tabla de análisis de varianza toma la forma:

Fuente	gl	SS	MS	F
Acuñaación	3	37.74753	12.58251	26.27201
Error	23	11.01544	0.47893	
Total	26	48.76297		

Al observar el valor del estadístico  $F$ , parece posible afirmar que las cuatro acuñaaciones contienen diferentes porcentajes de plata; sin embargo, cuando se analizan los residuos es clara la existencia de heterocedasticidad, tal y como se observa en la figura 7.5.

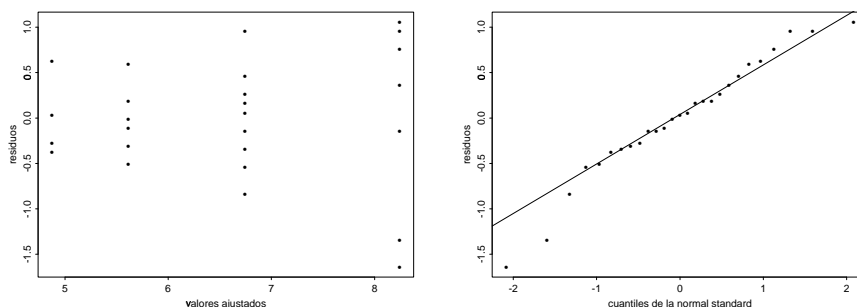


Figura 7.5: Gráficos de residuos para el ejemplo 7.2.

Por lo tanto, es necesario llevar a cabo una transformación de los datos con el fin de eliminar ese problema. Veamos entonces un método empírico para estimar dicha transformación.

Sea  $E(y_{ij}) = \mu_i$  la esperanza de  $y_{ij}$ , y supongamos que la desviación standard de  $y_{ij}$ ,  $\sigma_i$ , es proporcional a alguna potencia de  $\mu_i$

$$\sigma_i \propto \mu_i^\alpha$$

Por supuesto, éste no es el único tipo posible de relación entre la media y la varianza. Sin embargo, es lo suficientemente flexible como para obtener resultados prácticos en la mayoría de los casos.

Se desea, entonces, encontrar una transformación que establezca la varianza.

Para ello, definamos  $y_{ij}^* = y_{ij}^\lambda$ , una transformación potencial de los datos originales. Si llamamos  $\sigma_i^*$  a la desviación standard de  $y_{ij}^*$ , puede probarse que:

$$\sigma_i^* \propto \mu_i^{\lambda+\alpha-1}$$

Claramente, si tomamos  $\lambda = 1 - \alpha$  obtendremos  $\sigma_i^* = \sigma^*$  constante, con lo cual resolveremos el problema de heterocedasticidad. Es necesario, entonces, buscar una manera de estimar  $\alpha$ . Para ello, observemos que  $\sigma_i \propto \mu_i^\alpha$  equivale a decir  $\sigma_i = \theta \mu_i^\alpha$ , donde  $\theta$  es una constante de proporcionalidad. Tomando logaritmos, la relación multiplicativa se convierte en una relación aditiva, la cual puede escribirse como:

$$\log(\sigma_i) = \log \theta + \alpha \log \mu_i$$

Lo anterior sugiere el siguiente procedimiento para la estimación de la transformación:

1. Para cada grupo, calcular la media muestral  $\bar{y}_i$  y la desviación standard muestral  $s_i$
2. Graficar  $\log s_i$  vs  $\log \bar{y}_i$
3. Determinar la pendiente de la recta. Ese será el valor estimado de  $\alpha$ .
4. Hacer la transformación  $y_{ij}^* = y_{ij}^\lambda$ , donde  $\lambda = 1 - \alpha$

Si se obtiene  $\lambda = 0$ , se toma como transformación el logaritmo de los datos.

La tabla 7.3 presenta algunas de las transformaciones más usuales obtenidas por este método. En general, es recomendable usar valores simples de  $\lambda$ , pues la diferencia práctica entre 0.5 y 0.55, por ejemplo, es casi imperceptible y la interpretación de valores sencillos es mucho más clara.

Relación	$\alpha$	$\lambda = 1 - \alpha$	Transformación
$\sigma \propto \text{cte}$	0	1	Ninguna
$\sigma \propto \mu^{\frac{1}{2}}$	$\frac{1}{2}$	$\frac{1}{2}$	Raíz cuadrada
$\sigma \propto \mu$	1	0	Logaritmo
$\sigma \propto \mu^{\frac{3}{2}}$	$\frac{3}{2}$	$-\frac{1}{2}$	Raíz cuadrada inversa
$\sigma \propto \mu^2$	2	-1	Inverso

Tabla 7.3: Transformaciones usuales para estabilizar la varianza en caso de heterocedasticidad.

Los valores de  $s_i$  para el ejemplo 7.2 son, respectivamente, 0.543, 1.1, 0.45 y 0.363. Al graficar  $\log s_i$  vs  $\log \bar{y}_i$  se obtiene el resultado que se muestra en la figura 7.6. Si bien el primer punto no parece estar sobre la recta, los otros tres presentan un crecimiento claro. Si elegimos los grupos 2 y 3 para determinar la pendiente de la recta, obtenemos un valor de  $-1.2$ , lo cual sugiere llevar usar el inverso como transformación para estabilizar la varianza.

La tabla de análisis de varianza para los datos transformados es

Fuente	gl	SS	MS	F
Acuñaación	3	0.02147491	0.007158305	33.91868
Error	23	0.00485399	0.000211043	
Total	26	0.0263289		

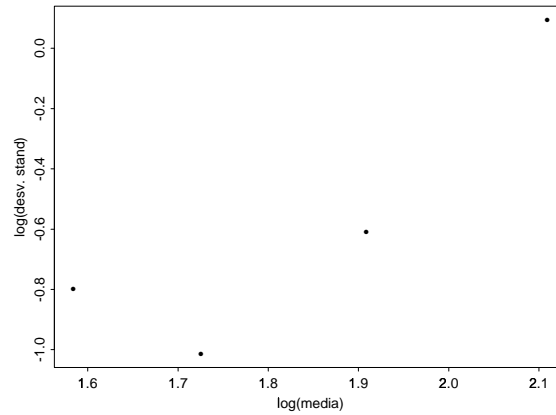


Figura 7.6: Gráficos de  $\log s_i$  vs  $\log \bar{y}_i$ , ejemplo 7.2.

Nuevamente, se concluye que existen diferencias entre las cuatro acuñaciones.

Los gráficos de residuos contra valores ajustados y de residuos en papel normal se muestran en la figura 7.7. Puede verse que ha desaparecido la apertura, y que el gráfico en papel normal está mucho más cerca de ser una recta. Por tanto, parece más adecuado proseguir el análisis a partir de estos datos transformados.

## 7.4 Análisis de varianza con dos criterios de clasificación

**Ejemplo 7.3** *Un ingeniero diseña una batería para ser usada en cierto equipo, el cual se sabe estará sometido a variaciones grandes de temperatura. El único parámetro de diseño que puede ser seleccionado es el material usado en las placas, para el cual tiene tres posibles selecciones. Una vez que la batería es enviada a su lugar de destino, el ingeniero no tiene control sobre la temperatura de operación de la misma, y su experiencia le dice que la temperatura de operación puede afectar el tiempo de vida de la batería. Sin embargo, es posible controlar la temperatura de operación en el laboratorio con la finalidad de realizar pruebas.*

*El ingeniero decide probar los tres materiales a tres niveles de temperatura () pues considera que esos tres valores de temperatura son consistentes con el ambiente en el cual*



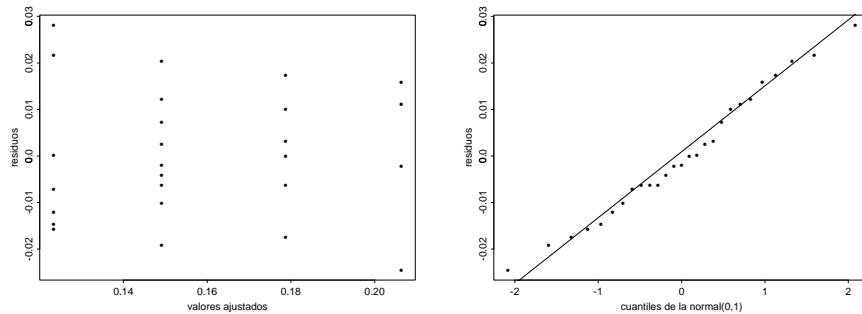


Figura 7.7: Gráficos de residuos para los datos transformados, ejemplo 7.2.

será usado el equipo. Se prueban cuatro baterías para cada combinación de material y temperatura, y se mide su tiempo de vida (en horas). Los resultados del experimento se encuentran en la tabla 7.4. El ingeniero desea responder dos preguntas

1. ¿Cómo es afectado el tiempo de vida de la batería por el tipo de material y la temperatura?
2. ¿Existe una selección de material que produzca un tiempo de vida uniforme independientemente de la temperatura elegida? Esto sería, sin duda, importante, pues permitiría diseñar una batería robusta ante los cambios de temperatura.

En este ejemplo, existen dos posibles criterios de clasificación para las observaciones (material usado en las placas y temperatura), y deseamos determinar cómo se diferencia el voltaje de salida para cada combinación de estos factores.

En general, si tenemos un criterio A con  $a$  posibles clasificaciones, o equivalentemente un factor A con  $a$  niveles, y un factor B con  $b$  niveles, la situación correspondiente al ejemplo 7.3 puede describirse de la siguiente manera:

Material	Temperatura (C)					
	8		15		22	
1	130	155	34	40	20	70
	74	180	80	75	82	58
2	150	188	136	122	25	70
	159	126	106	115	58	45
3	138	110	174	120	96	104
	168	160	150	139	82	60

Tabla 7.4: Tiempos de vida para baterías con tres diferentes materiales en las placas, a tres temperaturas diferentes de funcionamiento (ejemplo 7.3)

		Factor B			
		$B_1$	$B_2$	$\dots$	$B_b$
Factor A	$A_1$	$y_{111}$	$y_{121}$	$\dots$	$y_{1b1}$
		$\vdots$	$\vdots$	$\dots$	$\vdots$
		$y_{11r}$	$y_{12r}$	$\dots$	$y_{1br}$
	$A_2$	$y_{211}$	$y_{221}$	$\dots$	$y_{2b1}$
		$\vdots$	$\vdots$	$\dots$	$\vdots$
	$y_{21r}$	$y_{22r}$	$\dots$	$y_{2br}$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	$A_a$	$y_{a11}$	$y_{a21}$	$\dots$	$y_{ab1}$
		$\vdots$	$\vdots$	$\dots$	$\vdots$
		$y_{a1r}$	$y_{a2r}$	$\dots$	$y_{abr}$

Cuando tenemos un conjunto de datos con esta estructura, diremos que estamos ante un problema de *análisis de varianza con dos criterios de clasificación*. En un caso como éste, es de interés determinar si el factor A y el factor B influyen sobre la variable de respuesta. Sin embargo, cabe preguntarse si los diferentes niveles del factor A tendrán la misma influencia sobre cada nivel del factor B. Por ejemplo, ¿tendrá la temperatura el mismo efecto sobre el tiempo de vida para los tres materiales, o existirán diferencias en este aspecto?

Este comentario hace necesario, antes de proseguir, el estudio del concepto de *interacción*.

### 7.4.1 Interacción

En muchos experimentos, observamos que los cambios en la variable de respuesta originados, digamos, por el factor B no son los mismos para diferentes niveles del factor A. En estos casos, diremos que existe *interacción* entre los factores A y B.

Para ilustrar este concepto, consideremos la siguiente tabla:

	$B_1$	$B_2$
$A_1$	20	30
$A_2$	40	52

Si dibujamos las respuestas con respecto a los distintos niveles del factor B, y unimos con una línea los puntos correspondientes al mismo nivel del factor A, obtenemos el gráfico que se presenta en la figura 7.8. En dicho gráfico puede verse que el cambio en el factor B afecta de la misma manera a ambos niveles del factor A. Supongamos, por ejemplo, que los datos corresponden a un experimento químico en el cual medimos la cantidad de cierto producto obtenida usando dos temperaturas de reacción diferentes (A) y añadiendo o no un catalizador (B). En ese caso, el gráfico nos diría que al añadir el catalizador obtenemos el mismo incremento en la producción para las dos temperaturas. Además, observamos que a la temperatura A2 siempre se obtiene mayor cantidad de producto que a la temperatura A1, y que siempre es mejor añadir el catalizador. En casos como este, diremos que *no hay interacción* entre los factores A y B (o usando la analogía anterior, entre la temperatura y el catalizador), o bien que los efectos de los factores A y B son *aditivos* (es decir, se suman).

Consideremos ahora una nueva tabla del mismo tipo:

	$B_1$	$B_2$
$A_1$	20	40
$A_2$	50	12

Al dibujar el gráfico equivalente, obtenemos el resultado presentado en la figura 7.9. En este nuevo gráfico puede observarse que para el nivel A1 del factor A, el cambio de B produce un incremento de la variable de respuesta, mientras que para el nivel A2 se produce un descenso en el valor de la variable de respuesta. Usando nuevamente la analogía con el experimento químico, observaríamos que al realizar la reacción a la temperatura A1, la adición del catalizador produce un aumento en la producción, mientras que a la temperatura A2 obtenemos el efecto contrario. Adicionalmente, observemos que para este

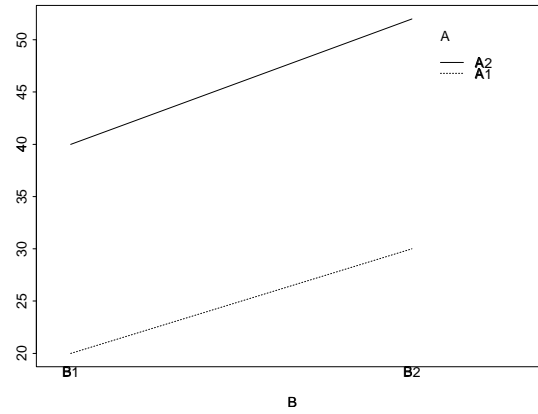


Figura 7.8: Gráfico correspondiente a un experimento sin interacción

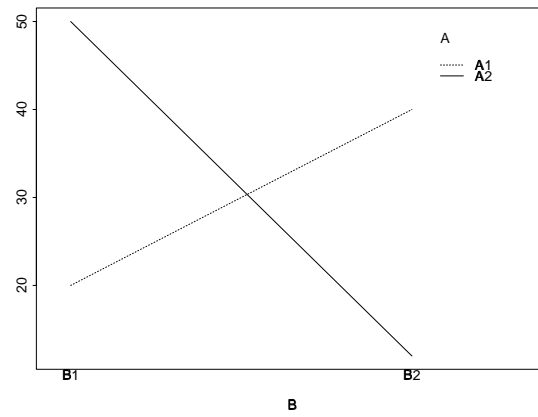


Figura 7.9: Gráfico correspondiente a un experimento con interacción

experimento, contrariamente al caso anterior, no puede hablarse de una “mejor temperatura”, o de que sea en general mejor añadir o no el catalizador. En casos como éste, diremos que *existe interacción* entre los factores A y B.

Los gráficos anteriores suelen llamarse (nada sorprendentemente) *gráficos de interacción*. En este tipo de gráficos la falta de paralelismo entre las líneas evidencia la existencia de interacción entre los factores estudiados.

El gráfico de interacción correspondiente a los datos del ejemplo 7.3 se presenta en la figura 7.10. Puede observarse en este gráfico que, para los materiales 1 y 2, el tiempo de vida promedio de la batería disminuye al aumentar la temperatura de 8 a 15 grados, mientras que para el material 3 hay un ligero incremento. Al aumentar la temperatura de 15 a 22 grados, el tiempo de vida promedio de las baterías fabricadas con el material 1 experimenta un ligero ascenso; en el caso de los materiales 2 y 3, el tiempo de vida promedio se reduce. Parece, entonces, que baterías elaboradas con diferentes materiales se comportan de manera distinta al cambiar las temperaturas; ésto indica la posible presencia de una interacción importante entre el material y la temperatura. Observemos además que, si bien las baterías elaboradas con el material 2 parecen tener un tiempo de vida mayor al operarse a 8 grados centígrados, el material 3 proporciona un rendimiento mucho más uniforme al cambiar la temperatura.

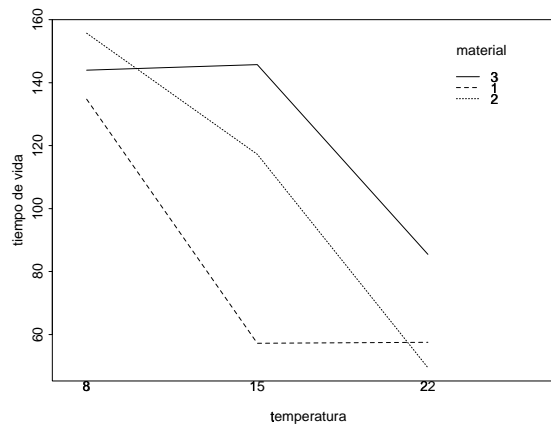


Figura 7.10: Gráfico de interacción para el tiempo de vida con respecto a la temperatura y al material usado en las placas de una batería (ejemplo 7.3)

### 7.4.2 Ajuste del modelo

Para los datos de la tabla 7.4 parece razonable ajustar el siguiente modelo:

$$\begin{aligned} y_{ijk} &= \mu_{ij} + \varepsilon, & i &= 1, \dots, a; \\ & & j &= 1, \dots, b; \\ & & k &= 1, \dots, r. \end{aligned} \quad (7.2)$$

Donde  $\mu_{ij}$  es la media de cada celda, es decir, el valor promedio de la variable de respuesta que se obtiene al combinar el nivel  $i$  del factor A con el nivel  $j$  del factor B. Este modelo tiene, entonces,  $a \times b$  parámetros, de modo que sólo puede ser ajustado cuando se dispone de más de una observación por celda, es decir, cuando se tienen dos o más *réplicas* del experimento ( $r \geq 2$ ). En efecto, si se dispone de una sola observación por celda, y hay que estimar  $a \times b$ , no quedarán grados de libertad disponibles para la estimación del error. Más adelante se estudiará el caso en el cual se dispone sólo de una observación por celda. Por ahora, supondremos que es posible estimar el modelo completo.

El modelo anterior presenta el inconveniente de que no permite separar la influencia de filas, columnas e interacción. Para poder separar dichas influencias, se emplea la siguiente reparametrización:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}, \quad (7.3)$$

donde los parámetros se interpretan como sigue:

- $\mu$ : Media general.
- $\alpha_i$ : Efecto de la fila.
- $\beta_j$ : Efecto de la columna.
- $(\alpha\beta)_{ij}$ : Efecto de la interacción.

Nótese que el modelo así escrito es equivalente al modelo 7.2, haciendo  $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ . Sin embargo, este segundo modelo tiene  $1 + a + b + ab = (a + 1)(b + 1)$  parámetros, lo cual indica que es necesario imponer  $a + b + 1$  restricciones. Igual que en el caso del modelo de análisis de varianza con un criterio de clasificación, son muchas las restricciones posibles; sin embargo, las más usadas son:

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

$$\sum_{j=1}^b \beta_j = 0$$

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

Aunque parece haber  $a + b + 2$  restricciones, en la última igualdad hay una redundancia ( $\sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij} = 0$ ), con lo cual se obtiene justo el número de restricciones necesarias.

Nuevamente se trata de un modelo lineal, y se supone que el vector de errores sigue una distribución  $\epsilon \sim N(\mathbf{0}, \sigma^2 I)$ . Por tanto, pueden emplearse todos los métodos ya presentados para este tipo de modelos.

Los estimadores de mínimos cuadrados para los parámetros del modelo son:

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

Nótese que  $\mu$ , la media general, se estima como el promedio de todas las observaciones;  $\alpha_i$ , el efecto de la fila  $i$ , se estima como el promedio de la fila  $i$  menos el promedio general, es decir, como la parte del promedio de la fila que no es explicada por la media general. Un estimador similar se emplea para  $\beta_j$ : el promedio de la columna  $j$  menos el promedio general. Finalmente, la interacción se estima como aquella parte del promedio de la celda  $ij$  que no puede explicarse como  $\mu + \alpha_i + \beta_j$ , es decir, la parte de la celda que no puede ser representada por medio de la suma de los efectos de fila y columna. Observemos, además, que  $\hat{y}_{ijk} = \bar{y}_{ij.}$ , es decir, cada observación se estima como el promedio de las observaciones en la celda en la cual se encuentra.

Para los datos del ejemplo 7.3 obtenemos los siguientes estimados de los parámetros del modelo:

$$\begin{aligned} \hat{\mu} &= 105.25 \\ \hat{\alpha}_1 &= -22.08, & \hat{\alpha}_2 &= 2.25, & \hat{\alpha}_3 &= 19.83 \\ \hat{\beta}_1 &= 39.58, & \hat{\beta}_2 &= 1.5, & \hat{\beta}_3 &= -41.08 \\ \widehat{\alpha\beta}_{11} &= 12.00, & \widehat{\alpha\beta}_{12} &= -27.42, & \widehat{\alpha\beta}_{13} &= 15.42 \\ \widehat{\alpha\beta}_{21} &= 8.67, & \widehat{\alpha\beta}_{22} &= 8.25, & \widehat{\alpha\beta}_{23} &= -16.92 \\ \widehat{\alpha\beta}_{31} &= -20.67, & \widehat{\alpha\beta}_{32} &= 19.17, & \widehat{\alpha\beta}_{33} &= 1.5 \end{aligned}$$

### 7.4.3 Pruebas de Hipótesis

Para simplificar el modelo, observemos que el modelo  $y_{ijk} = \mu + \varepsilon_{ijk}$  está anidado con el modelo completo descrito en las ecuaciones 7.2 y 7.3. Por lo tanto, puede escribirse una tabla de análisis de varianza similar a las usadas para el caso de los modelos de regresión, o de análisis de varianza con un criterio de clasificación, tal y como se muestra en la siguiente tabla:

Fuente	gl	SS	MS	F
Modelo	$ab$	$SSM$	$MSM = \frac{SSM}{a-1}$	$F_M = \frac{MSM}{MSE}$
Error	$ab(r-1)$	$SSE$	$MSE = \frac{SSE}{k-1}$	
Total	$abr-1$	$SST$		

donde

$$SSM = r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{...})^2$$

$$SSE = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y}_{ij.})^2$$

$$SST = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y}_{...})^2$$

Sin embargo, ésto no es suficiente, pues estamos interesados en separar la influencia de los dos factores y de la interacción entre ellos. Es decir, deseamos probar las siguientes hipótesis

- Todas las filas son iguales:

$$H_0^A : \alpha_1 = \alpha_2 = \dots = \alpha_a \text{ vs } H_1^A : \text{algún } \alpha_i \neq 0$$

- Todas las columnas son iguales:

$$H_0^B : \beta_1 = \beta_2 = \dots = \beta_b \text{ vs } H_1^B : \text{algún } \beta_j \neq 0$$

- No hay interacción entre filas y columnas:

$$H_0^{AB} : \alpha\beta_{ij} = 0 \text{ para todo } i, j \text{ vs } H_1^{AB} : \text{algún } \alpha\beta_{ij} \neq 0$$



Es importante destacar que, en el caso de que haya interacción presente, no es cierto que todas las filas o todas las columnas sean iguales; por tanto, carece de sentido probar  $H_0^A$  y  $H_0^B$  cuando se ha rechazado  $H_0^{AB}$ . En otras palabras, *si en un modelo aparece el término de interacción, también deben estar presentes los efectos de cada factor.*

Para llevar a cabo las pruebas de hipótesis anteriores, observemos que la suma de cuadrados para el modelo  $SSM$  puede descomponerse como:

$$\begin{aligned}
 SSM &= r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_{...})^2 \\
 &= r \sum_{i=1}^a \sum_{j=1}^b [(\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{.j} - \bar{y}_{...}) + (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})]^2 \\
 &= br \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 + ar \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{...})^2 + r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})^2 \\
 &= br \sum_{i=1}^a \hat{\alpha}_i^2 + ar \sum_{j=1}^b \hat{\beta}_j^2 + r \sum_{i=1}^a \sum_{j=1}^b (\hat{\alpha}\hat{\beta})_{ij}^2 \\
 &= SSA + SSB + SSAB
 \end{aligned}$$

Puede probarse que  $SSA \sim \chi_{a-1}^2$ ,  $SSB \sim \chi_{b-1}^2$ ,  $SSAB \sim \chi_{(a-1)(b-1)}^2$ , y que estas distribuciones son independientes. Nótese, además, que  $SSA$  representa la variación explicada por los efectos de las filas,  $SSB$  representa la variación explicada por los efectos de las columnas y, finalmente,  $SSAB$  expresa la variación explicada por la interacción en el modelo. Estos resultados pueden usarse para construir la siguiente tabla de análisis de varianza:

Fuente	gl	SS	MS	F
A	$a - 1$	$SSA$	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MSA}{MSE}$
B	$b - 1$	$SSB$	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$
AB	$(a - 1)(b - 1)$	$SSAB$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$	$F_{AB} = \frac{MSAB}{MSE}$
Error	$ab(r - 1)$	$SSE$	$MSE = \frac{SSE}{k-1}$	
Total	$abr - 1$	$SST$		

Tenemos, entonces, tres valores de  $F$ . El valor  $F_{AB}$  nos permite probar la hipótesis  $H_0^{AB}$ , la cual será rechazada si  $F_{AB} > F_{(a-1)(b-1), ab(n-1)}^\alpha$ . En forma análoga, rechazaremos  $H_0^A$  si  $F_A > F_{a-1, ab(n-1)}^\alpha$ , y  $H_0^B$  cuando  $F_B > F_{b-1, ab(n-1)}^\alpha$ .

Para los datos del ejemplo 7.3, la tabla de análisis de varianza que se obtiene se presenta en la tabla 7.5. Eligiendo  $\alpha = 0.05$  y comparando el valor calculado de  $F_{AB} = 3.54291$  con el valor tabulado  $F_{2,27}^{0.05} = 3.35$ , rechazamos  $H_0^{AB}$ , con lo cual ya no es necesario probar las hipótesis  $H_0^A$  y  $H_0^B$ .

Fuente	gl	SS	MS	F
Material	2	10633.17	5316.58	7.98341
Temperatura	2	39083.17	19541.58	29.34376
Material:Temp	4	9437.67	2359.42	3.54291
Error	27	17980.75	665.95	
Total	35	77134.76		

Tabla 7.5: Tabla de análisis de varianza para los datos de tiempo de vida de baterías (ejemplo 7.3)

Nótese que el resultado obtenido confirma lo observado en la sección anterior con respecto a la interacción.

Se pueden emplear las mismas técnicas que en el caso de análisis de varianza con un criterio de clasificación para comparar las medias de dos celdas, es decir, para probar hipótesis de la forma  $H_0 : \mu_{ij} = \mu_{kl}$  vs  $H_1 : \mu_{ij} \neq \mu_{kl}$ . Puede usarse el método de mínima diferencia significativa, o bien intervalos de confianza.

Para los datos del ejemplo 7.3 nos interesa comparar el efecto de las diferentes temperaturas sobre cada uno de los materiales, con el fin de determinar si alguno de ellos es robusto a los cambios de temperatura. En este caso la mínima diferencia significativa es la misma para todas las pruebas, y se calcula (usando  $\alpha = 0.05$  como

$$t_{27}^{0.05} \sqrt{\frac{MSE}{2}} = 31.08$$

Para el material 1 las diferencias de los tiempos de vida entre baterías que operan a diferentes temperaturas son:

$$\begin{aligned} |\bar{y}_{11} - \bar{y}_{12}| &= |134.75 - 57.25| = 77.5 > 31.08 \\ |\bar{y}_{11} - \bar{y}_{13}| &= |134.75 - 57.5| = 77.25 > 31.08 \\ |\bar{y}_{12} - \bar{y}_{13}| &= |57.25 - 57.5| = 0.25 < 31.08 \end{aligned}$$

es decir, las baterías elaboradas con el material 1 tienen un tiempo de vida significativamente menor, comparado con el tiempo de vida a 8 grados centígrados, cuando operan a 17 o 22 grados cent. No hay diferencias significativas entre estas dos temperaturas.

En el caso del material 2, las diferencias son las siguientes:

$$\begin{aligned} |\bar{y}_{21.} - \bar{y}_{22.}| &= |155.75 - 117.25| = 38.5 > 31.08 \\ |\bar{y}_{21.} - \bar{y}_{23.}| &= |155.75 - 49.5| = 106.25 > 31.08 \\ |\bar{y}_{22.} - \bar{y}_{23.}| &= |117.25 - 49.5| = 67.75 > 31.08 \end{aligned}$$

En este caso, los tiempos de vida para las tres temperaturas son diferentes, y el tiempo de vida es menor para cada temperatura.

Para el material 3 obtenemos las siguientes diferencias:

$$\begin{aligned} |\bar{y}_{31.} - \bar{y}_{32.}| &= |144 - 145.75| = 1.75 < 31.08 \\ |\bar{y}_{31.} - \bar{y}_{33.}| &= |144 - 85.5| = 58.5 > 31.08 \\ |\bar{y}_{32.} - \bar{y}_{33.}| &= |145.75 - 85.5| = 60.25 > 31.08 \end{aligned}$$

El desempeño de las baterías construidas usando el material 3 no se deteriora al subir la temperatura de 8 a 17 grados centígrados, pero el tiempo de vida sufre un descenso considerable cuando la temperatura sube a 22 grados.

A partir de estas observaciones y del gráfico de interacción, elaborar las baterías con el material 3 parece la mejor decisión. Para corroborar esto, veamos que el material 3 tiene un desempeño igual o mejor al de los materiales 1 y 2 a cualquier temperatura.

A 8 grados:

$$\begin{aligned} |\bar{y}_{31.} - \bar{y}_{11.}| &= |144 - 134.75| = 9.25 < 31.08 \\ |\bar{y}_{31.} - \bar{y}_{21.}| &= |144 - 155.75| = 11.75 < 31.08 \end{aligned}$$

es decir, a 8 grados los tiempos de vida de baterías elaboradas con los tres materiales son iguales.

A 15 grados:

$$\begin{aligned} |\bar{y}_{32.} - \bar{y}_{12.}| &= |145.75 - 57.25| = 88.5 > 31.08 \\ |\bar{y}_{32.} - \bar{y}_{22.}| &= |145.75 - 117.25| = 28.55 < 31.08 \end{aligned}$$

A temperatura de 15 grados, las baterías elaboradas con el material 3 son mejores que las de material 1 y comparables a las de material 2.

Finalmente, a 22 grados:

$$\begin{aligned} |\bar{y}_{31.} - \bar{y}_{11.}| &= |85.5 - 57.5| = 28 < 31.08 \\ |\bar{y}_{31.} - \bar{y}_{21.}| &= |85.5 - 49.5| = 36 > 31.08 \end{aligned}$$

A 22 grados, las baterías de material 3 son mejores que las de material 2 y son comparables a las de material 1.

Por todo lo anterior, la mejor decisión es diseñar las baterías usando el material 3 en las placas.

Al tratarse de un modelo lineal, es necesario nuevamente comprobar las suposiciones del modelo; es decir, debe llevarse a cabo el análisis de los residuos del modelo. La figura 7.11 presenta los gráficos de residuos vs valores ajustados y de residuos en papel normal para los datos del ejemplo 7.3. En el gráfico de residuos contra valores ajustados la varianza parece aumentar ligeramente hacia la derecha del gráfico, si bien esta impresión visual es causada principalmente por dos residuos de valor absoluto grande, uno negativo y otro positivo, los cuales además pertenecen a la misma celda. En el gráfico de residuos en papel normal estos mismos residuos parecen ocasionar un alejamiento de la recta  $y = x$ . Será necesario, por lo tanto, identificar la celda a la cual pertenecen esos datos problemáticos y determinar si hubo algún problema con ellos. De todas maneras, los problemas que se presentan no parecen lo bastante marcados como para invalidar las conclusiones sobre este conjunto de datos.

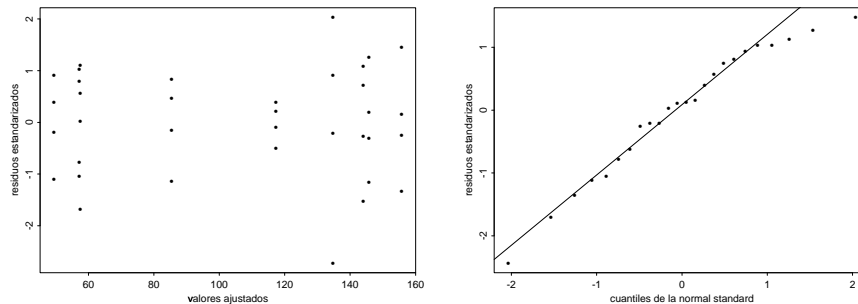


Figura 7.11: Gráficos de residuos para los datos de tiempo de vida de baterías (ejemplo 7.3)

En caso de que los residuos presenten heterocedasticidad pueden aplicarse las técnicas ya estudiadas para el modelo de análisis de varianza con un criterio de clasificación con el fin de estimar una transformación que permita estabilizar la varianza.