

Capítulo 2

Diagnosis y validación del modelo

2.1. Introducción

En el Capítulo 1 se ha construido el modelo completamente aleatorizado y el contraste de hipótesis de igualdad de medias mediante el test F . En el Capítulo 2 se han estudiado diversos procedimientos para comparar entre sí las medias de los efectos de los tratamientos. Todo ello se ha hecho en el supuesto de que las hipótesis básicas del modelo, dadas en la sección ?? del Capítulo 1, sean ciertas. Por ello, surge de inmediato la cuestión de cómo comprobar que dichas hipótesis se verifican y en caso de no verificarse cómo actuar ante esa situación.

Cuando se selecciona un modelo para un conjunto de datos, frecuentemente no se puede estar seguro a priori de que ese modelo sea adecuado. En el caso del ANOVA, puede suceder que alguna o varias características del modelo tales como la normalidad de los términos de error o la independencia en los datos investigados no se verifiquen. Por lo tanto, es importante examinar la adecuación del modelo a los datos antes de realizar un análisis de los mismos basado en dicho modelo; sin embargo, por razones metodológicas, se han estudiado antes los procedimientos de inferencia en los capítulos 1 y 2.

En este capítulo se trata la diagnosis y la validación del modelo, que consiste en estudiar si las hipótesis básicas del modelo están o no en contradicción con los datos observados. Para ello, se discuten algunos métodos gráficos sencillos y procedimientos estadísticos. También se estudian los efectos que produce el incumplimiento de las suposiciones fundamentales.

Como sabemos, las principales hipótesis sobre el modelo se expresan en términos de las perturbaciones aleatorias u_{ij} , $i = 1, \dots, I$; $j = 1, \dots, n_i$, que según dichas hipótesis deben ser variables aleatorias independientes con distribución Normal de media 0 y varianza constante σ^2 . Para verificar tales hipótesis tendremos que hacer uso de estimadores de las

perturbaciones. Recordemos que tomábamos como estimadores de las perturbaciones los residuos, definidos como

$$e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_i .$$

Los métodos gráficos y los contrastes estadísticos que se describen en este capítulo para diagnosticar la adecuación del modelo de análisis de la varianza son, respectivamente, *las gráficas de residuos y los contrastes de igualdad de varianzas*. También se aborda el uso de transformaciones como método para mejorar la validez del modelo de análisis de la varianza y se considera la influencia que tienen las desviaciones de las hipótesis del modelo ANOVA en la inferencia. Podemos, por tanto, considerar que en el estudio de un experimento se debe seguir un proceso secuencial formado por los pasos siguientes:

- 1^o) Plantear un modelo que explique los datos.
- 2^o) Examinar la adecuación del modelo planteado. Si el modelo no es adecuado, tomar las medidas correctoras, tales como empleo de transformaciones de los datos, o modificar el modelo.
- 3^o) Si el modelo es adecuado se realiza el análisis estadístico de los datos. A continuación, se evalúa su efectividad o el grado de ajuste del modelo y se estudia su sensibilidad ante diversas situaciones.

Como se dirá posteriormente, los modelos ANOVA son razonablemente *robustos* frente a ciertos tipos de desviaciones del modelo, tales como que los términos de error no estén exactamente normalmente distribuidos. Por lo tanto, podemos decir que el principal objetivo del examen de la adecuación del modelo es detectar graves desviaciones de las condiciones supuestas en el modelo.

2.2. Comprobación de la idoneidad del modelo

Hemos dicho en el Capítulo 1 que el análisis de la varianza de una vía se basa en las siguientes suposiciones:

- * Las observaciones vienen descritas por el modelo

$$y_{ij} = \mu + \tau_i + u_{ij} \quad , \quad i = 1, 2, \dots, I \quad ; \quad j = 1, 2, \dots, n_i \quad . \quad (2.1)$$

- * Los errores u_{ij} son variables aleatorias independientes y normalmente distribuidas con media cero y varianza constante σ^2 .

Mediante el análisis de los residuos se puede comprobar, entre otras cosas, la presencia de observaciones anómalas. Para poder identificar dichas observaciones, como veremos más adelante, se pueden emplear los *residuos estandarizados* que pasamos a definir a continuación.

Residuos estandarizados

Se llaman *residuos estandarizados* y se denotan por d_{ij} , al cociente entre los residuos e_{ij} y la desviación típica poblacional, es decir

$$d_{ij} = \frac{e_{ij}}{\sigma} . \quad (2.2)$$

Puesto que σ suele ser desconocida, se modifica la definición de d_{ij} sustituyendo σ por la raíz cuadrada de la varianza residual,

$$d_{ij} = \frac{e_{ij}}{\sqrt{\hat{S}_R^2}} . \quad (2.3)$$

El análisis de los residuos tiene por objeto contrastar a posteriori las hipótesis del modelo. Así pues, este análisis va encaminado a comprobar:

- a) Si existe suficiente evidencia para no aceptar la incorrelación de los errores.
- b) Si la distribución de los errores es aproximadamente normal.
- c) Si existe alguna observación anómala que altere gravemente la estimación del modelo y que, tal vez, debería ser eliminada del estudio.
- d) Si la varianza de los errores es constante (homocedasticidad).

A continuación, iniciamos el estudio del análisis de los residuos. Para ello estudiamos, en primer lugar, su independencia estadística.

2.2.1. Independencia entre los residuos

Uno de los problemas que puede surgir en el estudio de un modelo de análisis de la varianza es la existencia de autocorrelación entre los residuos. Para el estudio de dicha autocorrelación se utiliza el gráfico de residuos frente al número de orden o de experiencia. Asimismo, se puede utilizar el test de Durbin-Watson.

Hay que tener en cuenta que los residuos e_{ij} no son variables aleatorias independientes, ya que, como sabemos, la suma de los residuos dentro de cada tratamiento es siempre igual a cero. Por lo tanto, únicamente $N - I$ de los N residuos son independientes. Sin embargo, si el tamaño muestral es grande en comparación con el número de tratamientos en el modelo, el efecto de dependencia entre los residuos es relativamente poco importante y se puede ignorar. En consecuencia, se pueden considerar los residuos como prácticamente independientes. La verificación del cumplimiento de esta hipótesis la abordaremos sólo gráficamente.

Gráfico de residuos en función del tiempo.

Un procedimiento gráfico para analizar la existencia de autocorrelación entre datos secuenciales es su representación a lo largo del tiempo. En el caso de los residuos, dicho procedimiento consiste en la representación de los residuos frente al orden en que se recopilaban los datos y buscar en dicho gráfico rachas de residuos de igual signo, así como cualquier tendencia creciente o decreciente en los mismos, lo cual sería un claro indicio de correlación entre los términos de error y el tiempo. Esto implica que la suposición de independencia de los residuos ha sido violada. En las figuras 3-1a y 3-1b se muestran ejemplos de esta situación

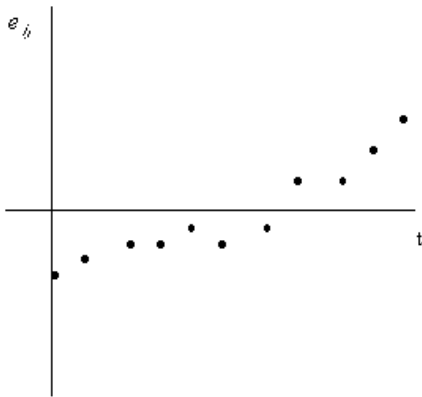


Figura 3-1a

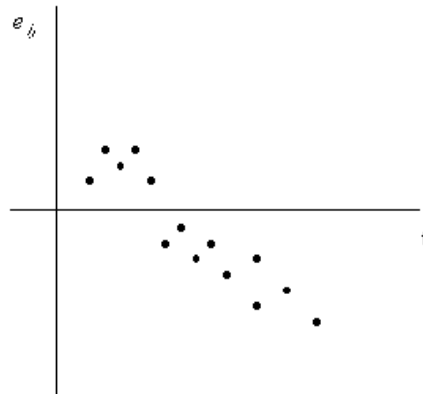


Figura 3-1b

El efecto de la dependencia puede ser un problema muy grave y difícil de corregir, por ello es importante prevenirlo cuando los datos se están recopilando. El método más eficaz para prevenir la dependencia es realizar un procedimiento apropiado de aleatorización.

Los efectos relacionados con el tiempo también pueden conducir a crecimientos o decrecimientos de la varianza del error. Por ejemplo, la destreza del investigador puede cambiar a medida que el experimento se desarrolla, así el experimentador puede obtener medidas más exactas dando lugar a una disminución de la varianza del error al transcurrir

el tiempo. Las tendencias de este tipo se manifiestan en un gráfico de los residuos frente al tiempo en el que hay mayor dispersión en un extremo que en el otro. Un ejemplo de esta situación se muestra en la siguiente figura

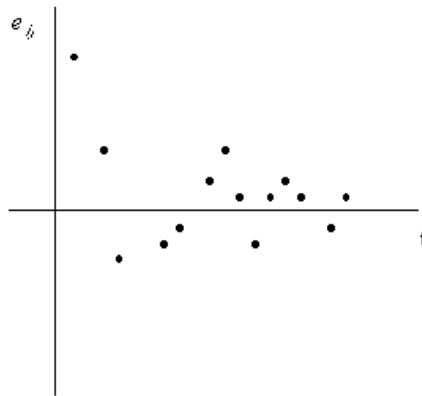


Figura 3-2

A continuación se muestra el gráfico de los residuos en función del tiempo para el ejemplo de referencia.

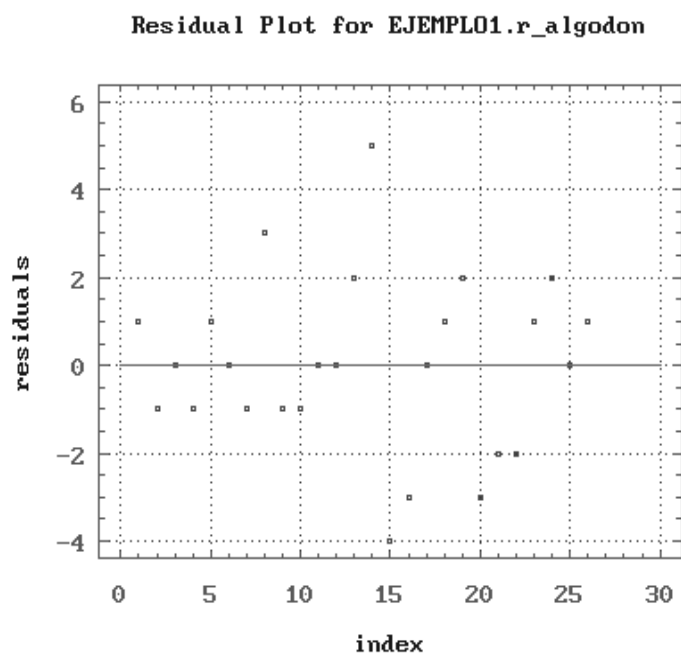


Figura 3-3

En esta gráfica no observamos ninguna tendencia sistemática que haga sospechar del incumplimiento de la suposición de independencia.

Después de analizar la independencia de los residuos, pasamos a estudiar la distribución de los mismos.

2.2.2. Normalidad de los residuos

Continuaremos el análisis de los residuos estudiando su distribución. Para ello, en primer lugar se debe hacer una inspección general dibujando los residuos en un diagrama de puntos, como el realizado para el ejemplo de referencia, en la Figura 3-4

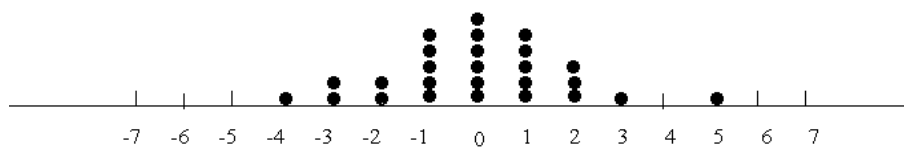


Figura 3-4: Diagrama de puntos de los residuos

Cuando el número de residuos es grande (≥ 20) se suelen agrupar los datos y construir un histograma, como el que muestra la siguiente figura, para el ejemplo de referencia.

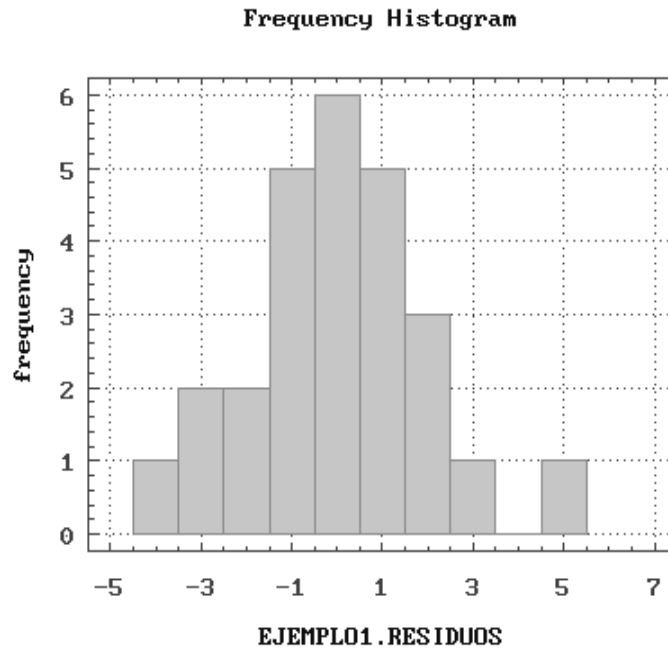


Figura 3-5

Si las hipótesis básicas del modelo son ciertas, ambas gráficas, figuras 3-4 y 3-5, tendrán en general la apariencia de una distribución normal centrada en cero. Desafortunadamente, cuando las muestras son pequeñas con frecuencia aparecen fluctuaciones considerables, por lo que la apariencia de no normalidad moderada no indica necesariamente la violación de la hipótesis de normalidad. Cuando hay grandes desviaciones se debe hacer un análisis más profundo y si es posible, realizar algunas transformaciones para corregir dichas desviaciones.

Las formas de discrepancias más frecuentes que se observan en estos gráficos ocurren cuando algunos residuos tienen un valor muy distanciado de los demás. Estos valores suelen corresponder a *datos anómalos* (outliers en terminología inglesa) que estudiaremos en la subsección 2.2.3.

También se puede representar un diagrama de puntos para cada tratamiento. En dicha gráfica podemos comprobar si hay un comportamiento anormal de los residuos en alguno de los tratamientos, así como examinar si hay muchas discrepancias entre las varianzas del

error. En las figuras 3-6a ; 3-6e se muestran dichos gráficos para el ejemplo de referencia.

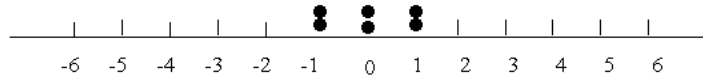


Figura 3-6a: Diagrama de residuos para el tratamiento 1



Figura 3-6b: Diagrama de residuos para el tratamiento 2



Figura 3-6c: Diagrama de residuos para el tratamiento 3

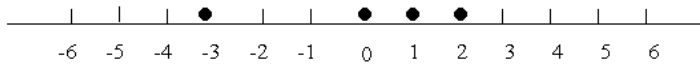


Figura 3-6d: Diagrama de residuos para el tratamiento 4

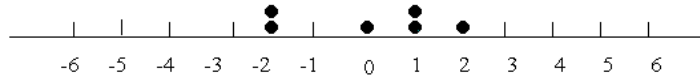


Figura 3-6e: Diagrama de residuos para el tratamiento 5

Otro procedimiento gráfico muy utilizado en la práctica para detectar el posible incumplimiento de la hipótesis de normalidad es el *gráfico probabilístico normal*, también denominado *gráfico gaussiano o representación en papel probabilístico normal*¹. La idea en que se basa este gráfico consiste en representar la función de distribución de una variable aleatoria X , $N(\mu, \sigma)$, en una escala transformada apropiada de forma que la gráfica quede linealizada. De esta manera, si se disponen de datos extraídos aleatoriamente de esa distribución, la representación en la escala transformada de la función de distribución empírica o muestral no debe separarse gráficamente de la línea recta teórica.

Para la construcción práctica del gráfico probabilístico normal se procede de la siguiente forma

- 1^a) Se ordenan las observaciones muestrales en orden creciente

$$x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[n]}$$

- 2^a) Se obtienen las proporciones acumuladas corregidas ξ_i que se pueden determinar de diversas formas, como por ejemplo

$$\xi_i = \frac{i}{n+1} \quad , \quad \xi_i = \frac{i - \frac{1}{2}}{n} \quad \text{y} \quad \xi_i = \frac{i}{n} \quad (2.4)$$

También, se suelen utilizar las expresiones sugeridas por Bernard & Bos-Levenbach (1953) y Blom (1958), respectivamente

$$\xi_i = \frac{i - 0,3}{n + 0,4} \quad , \quad (2.5)$$

$$\xi_i = \frac{i - \frac{3}{8}}{n + \frac{1}{4}} \quad , \quad (2.6)$$

¹Este gráfico es un caso particular de los gráficos de cuantiles teóricos frente a observados que, en terminología inglesa, se conocen como *Q-Q-Plot*.

siendo esta última, la expresión que utiliza STATGRAPHICS para elaborar *el gráfico probabilístico normal*.

3^a) A partir de los valores ξ_i se determinan los cuantiles teóricos, q_i , de la distribución normal mediante la fórmula

$$q_i = \Phi^{-1}(\xi_i) \quad ,$$

donde Φ es la función de distribución de la normal tipificada.

4^a) Se representa q_i frente a $x_{[i]}$

Cuanto mejor se ajuste la nube de puntos representada a una recta menos evidencia tendremos para suponer la violación de la hipótesis de normalidad de los residuos. Este gráfico, obviamente, se puede realizar para los valores de los residuos o para sus valores estandarizados.

El gráfico de q_i frente a los valores $x_{[i]}$ será una recta de ecuación,

$$x_{[i]} = \mu + \sigma q_i \quad ,$$

cuya ordenada en el origen estimará el valor de μ y la pendiente estimará el valor de σ .

Este procedimiento está automatizado utilizando el papel probabilístico normal en el que se representa $[i]$ o ξ_i frente a $x_{[i]}$.

En la tabla 3-2, se muestran los residuos organizados en orden ascendente, correspondientes a los datos del ejemplo de referencia, y los valores ξ_i dados por la expresión (2.6).

Tabla 3-2. Residuos y valores ξ_i para el Ejemplo 1-1

Orden	Resid.	ξ_i	Orden	Resid.	ξ_i	Orden	Resid.	ξ_i
i	e_{ij}	$\frac{i-3/8}{n+1/4}$	i	e_{ij}	$\frac{i-3/8}{n+1/4}$	i	e_{ij}	$\frac{i-3/8}{n+1/4}$
1	-4	0,023	10	-1	0,366	19	1	0,709
2	-3	0,061	11	0	0,404	20	1	0,747
3	-3	0,085	12	0	0,442	21	-1	0,785
4	-2	0,138	13	0	0,480	22	2	0,823
5	-2	0,176	14	0	0,519	23	2	0,861
6	-1	0,214	15	0	0,557	24	2	0,900
7	-1	0,252	16	0	0,595	25	3	0,938
8	-1	0,290	17	1	0,633	26	5	0,976
9	-1	0,328	18	1	0,671			

En la Figura 3-7 se muestra la gráfica en papel probabilístico normal para estos residuos. En esta gráfica los residuos están representados frente a $\xi_i \times 100$ en el eje de ordenadas.

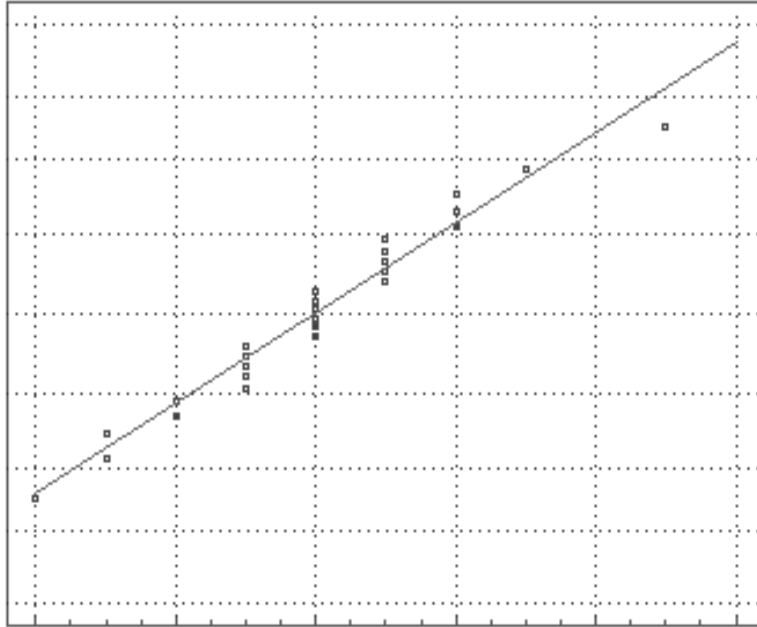


Figura 3-7

Los puntos en la Figura 3-7 están razonablemente próximos a la línea recta, excepto una observación que parece apartada de dicha línea y que estudiaremos en la subsección siguiente como posible valor anómalo.

El análisis de las desviaciones del modelo con respecto a la normalidad es, en muchos aspectos, más dificultoso que el correspondiente a los otros tipos de desviaciones. Por ejemplo, los residuos pueden parecer que no están normalmente distribuidos porque el modelo no sea el adecuado o porque la varianza del error no sea constante.

2.2.3. Residuos y datos anómalos

Los residuos extremos pueden identificarse en las gráficas de residuos (preferiblemente basadas en residuos estandarizados) y en la gráfica en papel probabilístico normal que también detecta observaciones atípicas. Una observación de tal naturaleza no es detectable durante el proceso de estimación del modelo y su presencia se puede hacer relevante cuando se examinan los residuos.

La presencia de uno o más residuos anómalos puede afectar “gravemente” el análisis de la varianza, en tales circunstancias es recomendable realizar una investigación minuciosa.

Lo primero que debe hacerse es revisar detenidamente los datos originales por si dicha discrepancia se debe a un error en la codificación de los datos, en los cálculos realizados. Si se detecta cualquiera de estas causas se debe corregir el dato y se seguirá con el experimento. Si no se puede corregir, entonces debe eliminarse del análisis.

Por otra parte, puede suceder que el residuo anómalo corresponda a un valor particular deseable en el experimento; por ejemplo, dicho dato puede contener información importante, como puede ser: alto rendimiento de la semilla de algodón, baja calificación del alumno, etc. En este caso, este residuo anómalo puede proporcionar más información que el resto de los datos. Una norma segura, frecuentemente sugerida, es descartar un residuo atípico sólo si hay suficiente evidencia de que representa un error de cálculo, de mal funcionamiento del equipo o una circunstancia de tipo similar.

A continuación, pasamos a comprobar si hay algún residuo atípico, por ejemplo, en la gráfica en papel probabilístico normal correspondiente al ejemplo de referencia. Para ello, recordemos que los residuos son las estimaciones de los errores aleatorios, los cuales son variables aleatorias $N(0, \sigma)$ y que, por lo tanto, los residuos estandarizados deben ser aproximadamente normales con media cero y varianza igual a uno. Así, debe cumplirse que aproximadamente el 68 % de los residuos estandarizados se encuentran entre los límites ± 1 , alrededor del 95 % entre ± 2 , y más del 99 % entre ± 3 . Entonces, podemos considerar como residuo potencialmente anómalo aquel cuya distancia del origen sea superior, en valor absoluto, a 3 o 4 desviaciones estándar.

Comprobamos que el residuo estandarizado máximo del ejemplo de referencia es

$$d_{3,3} = \frac{e_{3,3}}{\sqrt{\widehat{S}_R^2}} = \frac{5}{\sqrt{4,66}} = 2,316 \quad ,$$

que no se puede considerar de ningún modo como un posible residuo anómalo.

Algunos autores como Barnett & Lewis (1978), John & Prescott (1975) y Stefansky (1972) sugieren procedimientos estadísticos más complejos para detectar residuos anómalos.

2.2.4. Diagnósis de heterocedasticidad

La detección de la heterocedasticidad, es decir, de la violación de la hipótesis de igualdad de varianzas en el modelo, es bastante complicada y más aún su tratamiento. Existen múltiples pruebas conducentes a comprobar la presencia de heterocedasticidad. Entre los procedimientos gráficos mostraremos las representaciones de los residuos frente a los valores ajustados y frente a ciertas variables de interés. Con dichas gráficas se puede observar, por ejemplo en la última mencionada, si dichas variables están influyendo en algún sentido en la variabilidad de las observaciones, haciendo que ésta aumente o disminuya.

También estudiaremos, como hemos dicho anteriormente, algunos contrastes estadísticos como los tests de Barlett, Cochran y Hartley.

Gráfico de los residuos frente a los valores ajustados.

En esta gráfica se representan los residuos, $e_{ij} = y_{ij} - \hat{y}_{ij}$, frente a los valores previstos por el modelo para cada tratamiento, \hat{y}_{ij} . Dicha gráfica puede poner de manifiesto la violación de la hipótesis de homocedasticidad si, por ejemplo, se observa que la variabilidad de los residuos depende de la respuesta media en cada nivel del factor. Esto se manifiesta en que el gráfico de los residuos frente a \hat{y}_{ij} tiene forma de embudo. Una gráfica de esta forma pone de manifiesto un aumento o disminución de los errores en función de los niveles del factor y entonces podemos concluir que el modelo no es el adecuado.

Las figuras 3-8a y 3-8b representan tipos de gráficas en las que la varianza del error no es constante

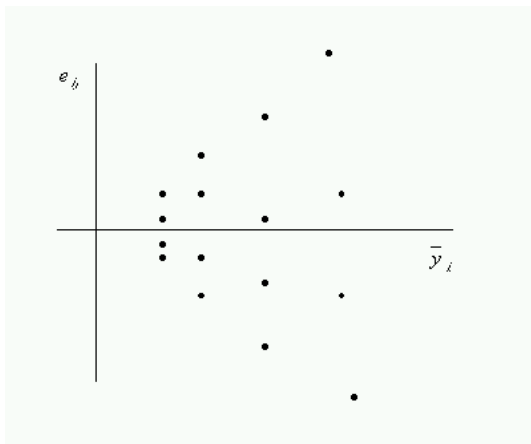


Figura 3-8a

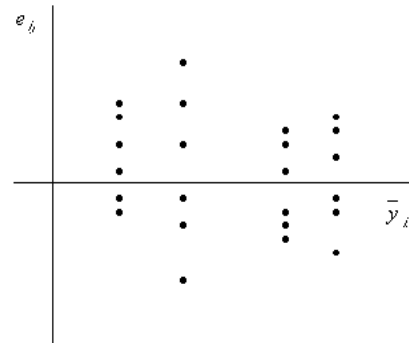


Figura 3-8b

En la Figura 3-8a se muestra un gráfico de residuos en forma de embudo, donde puede observarse que la variabilidad de los residuos aumenta con los niveles del factor.

En la Figura 3-8b se observa que la variabilidad de los residuos es mayor en el nivel 2.

A continuación se presenta la gráfica de los residuos frente a los valores estimados, realizada con STATGRAPHICS, para el ejemplo de referencia.

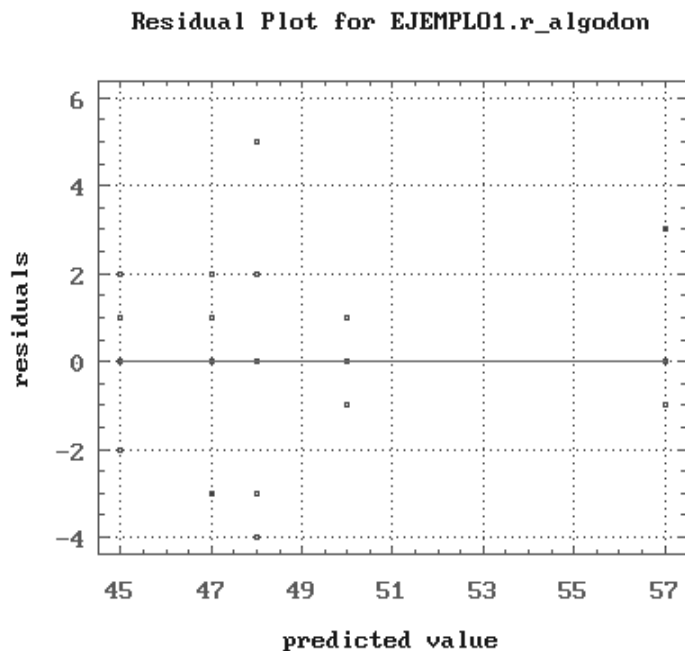


Figura 3-9

En dicha gráfica comprobamos que los residuos fluctúan en forma más o menos aleatoria alrededor de la línea base; es decir, no revelan ningún patrón obvio y ésto es indicativo de que los términos de error tienen varianza constante y los residuos son independientes de los valores ajustados.

Gráfico de los residuos en función de variables de interés

Otra representación importante es la gráfica de los residuos en función de otras variables que se crean de interés en cada experimento. Esta gráfica, como la anterior, no es sólo útil para comprobar si los residuos son independientes de estas variables de interés, sino también para examinar si la varianza de los términos de error es constante.

La siguiente gráfica muestra la representación de los residuos frente a los niveles del factor para el ejemplo de referencia.

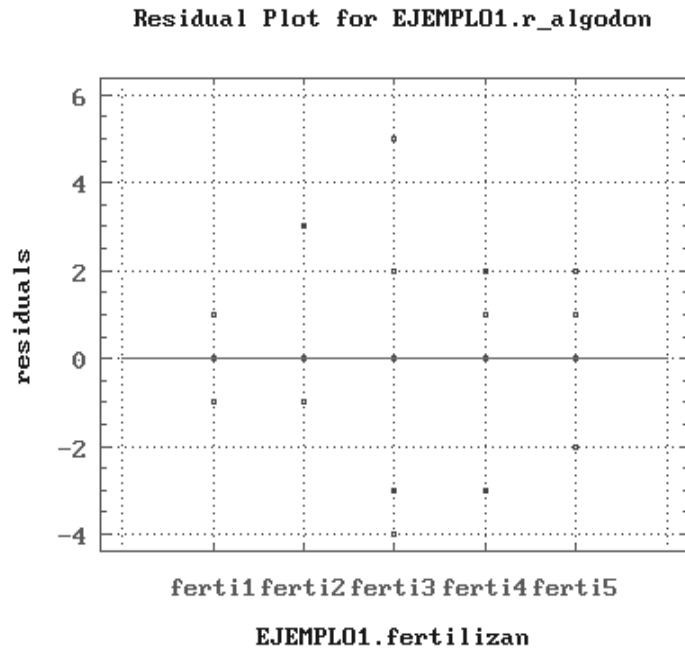


Figura 3-10

Las conclusiones que se obtienen de esta gráfica son las mismas que las obtenidas en la gráfica 3-9.

Además de las gráficas de los residuos, que frecuentemente se utilizan para diagnosticar la igualdad de las varianzas, existen contrastes para comprobar la homocedasticidad. Algunos de estos se estudian a continuación

Contrastes para la igualdad de las varianzas

Entre los contrastes que se utilizan para estudiar si las varianzas son iguales en los distintos tratamientos, como requiere el modelo ANOVA, es decir, para decidir entre las hipótesis

$$\begin{aligned}
 H_0 : \sigma_1^2 &= \sigma_2^2 = \dots = \sigma_I^2 \\
 H_1 : \sigma_i^2 &\neq \sigma_j^2 \text{ para algún par } (i, j) \text{ ,}
 \end{aligned}
 \tag{2.7}$$

consideramos los tests de Barlett, Cochran y Hartley. Estos contrastes suponen que y_{ij} son variables aleatorias independientes con distribución normal.

El test de Barlett puede usarse para tamaños de muestra iguales o desiguales mientras que los tests de Cochran y Hartley se aplican sólo cuando los tamaños de las muestras son

iguales.

En primer lugar estudiamos el test de Barlett.

Test de Barlett

La idea básica del test de Barlett es sencilla. Consiste en utilizar las medias aritmética y geométrica de las varianzas muestrales para construir un estadístico cuya distribución sea aproximadamente una χ^2 .

Supongamos que $s_1^2, s_2^2, \dots, s_I^2$ son las varianzas² muestrales de I poblaciones normales y gl_i son los grados de libertad asociados a s_i^2 . Por lo tanto, las medias aritmética y geométrica de dichas varianzas, ponderadas por el número de grados de libertad, son respectivamente,

$$MA = \frac{1}{gl_T} \sum_{i=1}^I gl_i s_i^2 \quad , \quad (2.8)$$

$$MG = \left[(s_1^2)^{gl_1} \cdot (s_2^2)^{gl_2} \cdot \dots \cdot (s_I^2)^{gl_I} \right]^{1/gl_T} \quad , \quad (2.9)$$

donde $gl_T = \sum_i gl_i$.

Es bien conocido que, para cualquier conjunto de valores, la media geométrica es siempre menor que la aritmética; es decir, $MG \leq MA$, verificándose la igualdad únicamente cuando todos los valores son iguales entre sí. Aplicando dicha propiedad:

- a) Si el cociente MA/MG está próximo a 1, tenemos evidencia de que las varianzas poblacionales son iguales.
- b) Si el cociente MA/MG es grande, indica que las varianzas poblacionales son desiguales.

Se puede razonar de forma similar si aplicamos logaritmos en el cociente MA/MG

$$\ln(MA/MG) = \ln MA - \ln MG \quad .$$

Barlett demostró que si las I muestras aleatorias provienen de poblaciones normales independientes, el estadístico

$$B = \frac{gl_T}{C} (\ln(MA) - \ln(MG)) \quad , \quad (2.10)$$

se distribuye, bajo la hipótesis nula y para tamaños de muestras grandes, aproximadamente como una χ^2 con $I - 1$ grados de libertad, donde

²Aunque nos referimos a s_i^2 como varianzas muestrales, en realidad representan cuasivarianzas muestrales.

$$C = 1 + \frac{1}{3(I-1)} \left[\left(\sum_{i=1}^I \frac{1}{gl_i} \right) - \frac{1}{gl_T} \right] . \quad (2.11)$$

Como se comprueba fácilmente, el término C es siempre mayor que 1.

Para una mayor simplicidad en los cálculos, este estadístico de contraste también se puede expresar como

$$B = \frac{1}{C} \left[(gl_T) \ln(MA) - \sum_{i=1}^I (gl_i) \ln s_i^2 \right] . \quad (2.12)$$

Cuando el test de Barlett se utiliza en un modelo unifactorial, entonces

$$gl_i = n_i - 1 \quad y \quad gl_T = \sum_i (n_i - 1) = N - I .$$

Por lo tanto el estadístico de contraste del test de Barlett, en este caso tiene la siguiente expresión:

$$B = \frac{1}{C} \left[(N - I) \ln(MA) - \sum_{i=1}^I (n_i - 1) \ln s_i^2 \right] , \quad (2.13)$$

donde

$$C = 1 + \frac{1}{3(I-1)} \left(\sum_{i=1}^I (n_i - 1)^{-1} - (N - I)^{-1} \right) \quad (2.14)$$

y

$$MA = \frac{\sum_{i=1}^I (n_i - 1) s_i^2}{N - I} .$$

Puesto que B se distribuye aproximadamente como una χ_{I-1}^2 cuando H_0 es cierta, el criterio de decisión, fijado un nivel de significación α , es el siguiente

$$\begin{aligned} \text{Si } B_{exp} \leq \chi_{\alpha, I-1}^2, & \text{ se acepta } H_0 \\ \text{Si } B_{exp} > \chi_{\alpha, I-1}^2, & \text{ se rechaza } H_0 \end{aligned} , \quad (2.15)$$

donde $\chi_{\alpha, I-1}^2$ es el valor crítico de la distribución χ^2 con $I - 1$ grados de libertad. Los percentiles de la distribución χ^2 están dados en la Tabla IV del Apéndice C

Hemos dicho que el estadístico B se distribuye aproximadamente como una χ^2 , dicha aproximación puede considerarse apropiada cuando los grados de libertad, gl_i , de cada una de las varianzas muestrales sean mayores o iguales a 4.

Para ilustrar este contraste utilizamos el Ejemplo 1-1. Una vez que se ha comprobado que las observaciones proceden de una distribución normal, se estudia si las varianzas σ_i^2 correspondientes a los cinco fertilizantes son iguales. Para ello, aplicamos la prueba de Barlett. Concretamente, en este ejemplo los tratamientos incumplen la condición impuesta a los grados de libertad por dicho test, ya que el fertilizante 4 tiene 3 grados de libertad. Posteriormente, volveremos a realizar el test utilizando la transformación de Cox para solventar este problema.

Para aplicar el test de Barlett, lo primero que hay que hacer es calcular las varianzas muestrales de cada tratamiento. Dichas varianzas valen:

$$s_1^2 = 0,8 \quad , \quad s_2^2 = 3 \quad , \quad s_3^2 = 13,5 \quad , \quad s_4^2 = 4,67 \quad , \quad s_5^2 = 2,8 \quad .$$

A continuación, para determinar el estadístico de contraste se puede construir la Tabla 3-3, organizando los datos de la siguiente manera:

Tabla 3-3

Fertilizantes	n_i	\bar{y}_i	s_i^2	$(n_i - 1)s_i^2$	$(n_i - 1) \ln s_i^2$	$(n_i - 1)^{-1}$
1	6	50	0,80	4	-1,11	1/5
2	5	57	3,00	12	4,39	1/4
3	5	48	13,5	54	10,41	1/4
4	4	47	4,67	14	4,62	1/3
5	6	45	2,80	14	5,14	1/5
	26			98	23,45	37/30

A fin de determinar el valor de B se calculan la media aritmética de las varianzas muestrales y el término C , obteniéndose los siguientes resultados:

$$MA = \frac{\sum_{i=1}^5 (n_i - 1)s_i^2}{N - I} = \frac{98}{21} = 4,67$$

$$C = 1 + \frac{1}{3(I-1)} \left(\sum_{i=1}^5 (n_i - 1)^{-1} - (N - I)^{-1} \right) = 1 + \frac{1}{3(4)} \left(\frac{3}{7} 30 - \frac{1}{21} \right) = 1,102 \quad ,$$

sustituyendo los valores de MA y C en la expresión de B , se obtiene,

$$B_{exp} = \frac{1}{C} \left[(N - I) \ln(MA) - \sum_{i=1}^5 (n_i - 1) \ln s_i^2 \right] = \frac{1}{1,102} (21 \ln 4,67 - 23,45) = 8,089 \quad .$$

Si realizamos el contraste al 5% y comparamos $B_{exp} = 8,089$ con el valor de la χ^2 teórica ($\chi_{0,05,4}^2 = 9,49$), se concluye que no puede rechazarse la hipótesis nula. Por lo tanto, se acepta que las varianzas de los cinco fertilizantes son iguales. El P -valor del contraste es $P\{\chi_4^2 > 8,089\} = 0,088$.

Si el numerador de B (sin calcular C) es menor que $\chi_{\alpha, I-1}^2$, como el valor de la constante C es siempre mayor que 1, entonces al dividir por C el valor de B disminuirá, por lo tanto en este caso no hace falta calcular el valor de C . Así, lo primero que se debe calcular es el numerador de B y sólo si éste es mayor que $\chi_{\alpha, I-1}^2$ se determinará el valor de C .

En el Ejemplo 1-1 el numerador de B , que denotamos por B_1 , es

$$B_1 = (N - I) \ln(MA) - \sum_{i=1}^I (n_i - 1) \ln s_i^2 = 21 \ln 4,67 - 23,45 = 8,914 \quad .$$

Por lo tanto, en este caso no es necesario determinar la constante C .

Hemos dicho que la prueba de Barlett no se debe aplicar cuando los grados de libertad de algunas o de todas las varianzas son menores que 4. En este caso, para aplicar dicho contraste se debe utilizar una transformación del estadístico de contraste, que mostramos a continuación.

Transformación de Box

Una transformación del estadístico de contraste de Barlett fué desarrollada por Box con el objetivo de contrastar la igualdad de varianzas cuando los grados de libertad tienen un valor pequeño. Dicho estadístico tiene la siguiente expresión

$$B' = \frac{f_2 BC}{f_1(A - BC)} \quad , \quad (2.16)$$

donde

- B está definido en (2.13)

- C está definido en (2.14)
- $f_1 = I - 1$
- $f_2 = \frac{I + 1}{(C - 1)^2}$
- $A = \frac{f_2}{2 - C + \frac{2}{f_2}}$.

Box demostró que este estadístico, bajo la hipótesis nula, sigue una distribución F de Snedecor con f_1 y f_2 grados de libertad.

En este caso, para decidir entre las hipótesis planteadas en (2.7), la regla de decisión apropiada al nivel de significación α , es:

$$\begin{aligned} \text{Si } B'_{exp} \leq F_{\alpha; f_1, f_2}, & \text{ se acepta } H_0 \\ \text{Si } B'_{exp} > F_{\alpha; f_1, f_2}, & \text{ se rechaza } H_0 \end{aligned} \quad (2.17)$$

donde F_{α, f_1, f_2} es el valor de la distribución F de Snedecor con f_1 y f_2 grados de libertad que deja a su derecha un área igual a α . El valor de f_2 generalmente no será un número entero, por lo que será necesario interpolar en la Tabla V del Apéndice C.

Para ilustrar este procedimiento, aplicamos la prueba de Barlett modificada a los datos del Ejemplo 1-1. Empezaremos calculando los valores de f_1 , f_2 y A .

$$1^{\text{a}}) f_1 = I - 1 = 4$$

$$2^{\text{a}}) f_2 = \frac{I + 1}{(C - 1)^2} = \frac{6}{0,010} = 576,70$$

$$3^{\text{a}}) A = \frac{f_2}{2 - C + \frac{2}{f_2}} = \frac{576,70}{2 - 1,102 + \frac{2}{576,70}} = \frac{576,70}{0,901} = 640,06$$

Por lo tanto, el valor del estadístico B' es:

$$B'_{exp} = \frac{f_2 B_{exp} C}{f_1 (A - B_{exp} C)} = \frac{(576,70)(8,089)(1,102)}{4(640,06 - (8,089)(1,102))} = \frac{5140,74}{2524,58} = 2,036 \quad .$$

Si se realiza el contraste a un nivel de significación $\alpha = 0,05$, el valor de la F teórica correspondiente a 4 y 576.70 grados de libertad es 2.371, ($F_{0,05;4,576,70} = 2,371$). Puesto que $B'_{exp} = 2,036 < 2,371$, se concluye que los rendimientos de los cinco fertilizantes tienen varianzas iguales. Esta es la misma conclusión que obtuvimos con el estadístico de contraste de Barlett y la aproximación a la distribución χ^2 .

Comentarios 2.1

C1) El test de Barlett es bastante sensible a las desviaciones de la normalidad. Si la población se aparta bastante de esta hipótesis, no es recomendable aplicar dicho test para contrastar la igualdad de varianzas.

C2) Como hemos dicho anteriormente, el test F para la igualdad de medias es robusto frente a las desviaciones de la hipótesis de homocedasticidad cuando los tamaños de las muestras son aproximadamente iguales. Por tanto, si las poblaciones son razonablemente normales de forma que el test de Barlett puede emplearse y los tamaños de las muestras no difieren mucho, sólomente cuando se detectan grandes diferencias entre las varianzas se considera que el modelo ANOVA no es el adecuado.

Test de Cochran

Cuando los tamaños muestrales son iguales, Cochran determinó la distribución, bajo la hipótesis nula, del siguiente estadístico de contraste

$$C = \frac{\max(s_i^2)}{\sum_{i=1}^I s_i^2} . \quad (2.18)$$

Los percentiles más importantes de esta distribución vienen recogidos en la Tabla XIX del Apéndice C.

La regla de decisión, para controlar el riesgo de cometer un error de tipo I igual a α , es:

$$\begin{aligned} \text{Si } C_{exp} \leq C_{1-\alpha; I, gl} , & \text{ se acepta } H_0 \\ \text{Si } C_{exp} > C_{1-\alpha; I, gl} , & \text{ se rechaza } H_0 \end{aligned} , \quad (2.19)$$

donde $C_{1-\alpha; I, gl}$ es el percentil $(1 - \alpha)100$ de la distribución de C para I poblaciones y gl son los grados de libertad de cada varianza muestral. En el modelo unifactorial equilibrado $gl = n - 1$.

Aplicamos el test de Cochran al Ejemplo 1-2. Recordemos que en dicho ejemplo las varianzas muestrales son:

$$s_1^2 = 5,31 \quad ; \quad s_2^2 = 4,16 \quad ; \quad s_3^2 = 12,47 \quad ; \quad s_4^2 = 4,44 \quad .$$

Una vez verificado que las observaciones se aproximan a una normal, utilizamos el test de Cochran para comprobar si los 4 tratamientos tienen o no la misma varianza; es decir,

$$\begin{aligned} H_0 : \sigma_1^2 &= \sigma_2^2 = \sigma_3^2 = \sigma_4^2 \\ H_1 : \sigma_i^2 &\neq \sigma_j^2 \text{ por lo menos para un par}(i, j) \text{ .} \end{aligned}$$

Para $I = 4$ y $gl = n - 1 = 11$, si realizamos el contraste, por ejemplo al 1%, y comparamos el cociente

$$C_{exp} = \frac{\max(s_i^2)}{\sum_{i=1}^I s_i^2} = \frac{12,47}{26,392} = 0,4724 \text{ ,}$$

con el valor correspondiente de la C teórica $C_{0,99;4,11} \approx 0,5$, se concluye que no hay suficiente evidencia para rechazar la hipótesis de igualdad de las varianzas de los cuatro tratamientos.

Test de Hartley

Si cada una de las I varianzas muestrales s_i^2 , tienen el mismo número de grados de libertad, $gl_i = gl$, un test sencillo para decidir entre

$$\begin{aligned} H_0 : \sigma_1^2 &= \sigma_2^2 = \dots = \sigma_I^2 \\ H_1 : \sigma_i^2 &\neq \sigma_j^2 \text{ para algún par}(i, j) \text{ ,} \end{aligned}$$

se debe a Hartley. Dicho contraste está basado únicamente en la varianza muestral más grande, denotada por $\max(s_i^2)$, y la varianza muestral más pequeña, denotada por $\min(s_i^2)$, siendo el estadístico de contraste correspondiente:

$$H = \frac{\max(s_i^2)}{\min(s_i^2)} \text{ .} \quad (2.20)$$

Claramente, valores de H próximos a 1 apoyan la hipótesis nula y valores grandes de H están a favor de la hipótesis alternativa.

Como en el test de Cochran, y al tratarse de un modelo unifactorial equilibrado, $gl = n - 1$.

La distribución de H depende del número de tratamientos I y de los grados de libertad, gl . Dicha distribución, cuando H_0 es cierta, ha sido tabulada por Hartley y los percentiles más usuales se incluyen en la Tabla X del Apéndice C.

La regla de decisión, como en los tests anteriores, es:

$$\begin{aligned} \text{Si } H_{exp} \leq H_{1-\alpha;I,gl} , & \text{ se acepta } H_0 \\ \text{Si } H_{exp} > H_{1-\alpha;I,gl} , & \text{ se rechaza } H_0 \end{aligned} \quad (2.21)$$

donde $H_{1-\alpha;I,gl}$ es el percentil $(1 - \alpha)100$ de la distribución de H para I poblaciones y gl grados de libertad.

A continuación, para ilustrar este contraste, vamos a aplicarlo a los datos del Ejemplo 1-2, en cuyo caso el estadístico de contraste H toma el siguiente valor

$$H_{exp} = \frac{\max(s_i^2)}{\min(s_i^2)} = \frac{12,47}{4,16} = 2,99 \quad .$$

Si realizamos el contraste al 5 %, el valor correspondiente de la H teórica es $H_{0,95;4,11} \approx 5$. Puesto que, $H_{1-\alpha} = 5 > H_{exp} = 2,99$, se decide H_0 ; en otras palabras, concluimos que, a un nivel de significación del 5 %, las varianzas de los cuatro tratamientos son iguales.

Comentarios de los contrastes de Cochran y Hartley

- C1) Los tests de Cochran y Hartley exigen que los tamaños de las muestras sean iguales, $n_1 = n_2 = \dots = n_I$. Si los tamaños de muestras no son iguales pero no difieren mucho, ambos tests pueden utilizarse como contrastes aproximados. En este caso, se toma como valor de los grados de libertad la media de los tamaños muestrales.*
- C2) Los tests de Cochran y Hartley como el de Barlett, son bastante sensibles a las desviaciones de la suposición de normalidad y no se utilizan cuando las desviaciones de esta suposición son importantes.*

Supongamos que se ha comprobado tanto gráficamente como mediante algún contraste que las varianzas de los distintos niveles no son iguales. En esta situación, se debe intentar estabilizar las varianzas aplicando transformaciones apropiadas a los datos y una vez aplicadas dichas transformaciones se debe hacer un análisis de la varianza a los datos transformados. En la sección siguiente trataremos el problema de seleccionar la forma de la transformación adecuada.

2.3. Transformaciones de los datos

En el caso de que las gráficas de los residuos u otros diagnósticos indiquen que el modelo ANOVA no es apropiado para los datos, se requieren posibles medidas correctoras.

Una de tales medidas es modificar el modelo. Esta opción puede tener el inconveniente de análisis más complejos. Otra medida es utilizar transformaciones en los datos. Y una tercera medida, cuando la dificultad básica es la falta de normalidad, es emplear algún test no paramétrico tal como el test de la mediana o el test de Kruskal-Wallis. Ambos tests se verán en el Capítulo 6.

Seguidamente, se estudian las transformaciones que se pueden realizar en los datos para intentar estabilizar las varianzas de los grupos.

2.3.1. Transformaciones para estabilizar la varianza

Se ha investigado bastante en el tema de la selección de la transformación apropiada de los datos en el caso de que las varianzas de los términos de error de cada nivel no sean iguales. Hay varios tipos de situaciones en las cuales dichas varianzas no son constantes y cada una de estas situaciones requiere diferente tipo de transformación de los datos. Así, por ejemplo, si el experimentador conoce la relación entre la varianza de las observaciones y la media de las mismas, podría usar esta información como guía para seleccionar la transformación adecuada de los datos, dando lugar a distintas situaciones que analizamos a continuación.

Sea y una variable aleatoria con media μ y varianza σ_y^2 . Supongamos que la varianza σ_y^2 está relacionada con la media μ mediante la función $\sigma_y^2 = g(\mu)$. Sea y^* una transformación cualquiera de y definida por $y^* = h(y)$. Entonces, utilizando el desarrollo de Taylor, se puede demostrar que la varianza $\sigma_{y^*}^2$ verifica

$$\sigma_{y^*}^2 \propto [h'(\mu)]^2 g(\mu) \quad (2.22)$$

donde $h'(\mu)$ es la primera derivada de $h(y)$ valuada en μ . Por tanto, si queremos hacer constante $\sigma_{y^*}^2$ podemos elegir $h(y)$ tal que (2.22) sea constante. Por conveniencia, elegimos esa constante igual a 1, $\sigma_{y^*}^2 = 1$. Por consiguiente,

$$[h'(\mu)]^2 g(\mu) = 1 \quad \text{o} \quad h'(\mu) = \frac{1}{\sqrt{g(\mu)}} \quad , \quad (2.23)$$

cuya solución, salvo la constante arbitraria, es

$$h(\mu) = \int \frac{d\mu}{\sqrt{g(\mu)}} \quad . \quad (2.24)$$

Vamos a enumerar algunos casos importantes y obtener la correspondiente transformación estabilizadora.

Supongamos que se cumplen las siguientes relaciones entre la varianza σ_y^2 y la media μ :

a) La varianza σ_y^2 es proporcional a μ .

Si $\sigma_y^2 = k\mu = g(\mu)$. Entonces de (2.24) se obtiene

$$h(\mu) = \int \frac{d\mu}{\sqrt{k\mu}} = \frac{2}{\sqrt{k}} \sqrt{\mu} \quad . \quad (2.25)$$

De esta forma, una transformación raíz cuadrada,

$$y^* = \sqrt{y} \quad ,$$

producirá una varianza constante en la variable transformada.

b) La desviación estándar σ_y es proporcional a μ . Sea $\sigma_y = k\mu$. Entonces $g(\mu) = k^2\mu^2$.

Por lo tanto, de (2.24) se obtiene

$$h(\mu) = \int \frac{d\mu}{k\mu} = \frac{1}{k} \ln(\mu) \quad . \quad (2.26)$$

De esta forma, una transformación logarítmica,

$$y^* = \ln(y) \quad ,$$

producirá una varianza constante en la variable transformada.

c) La desviación estándar σ_y es proporcional a μ^2 . Sea $\sigma_y = k\mu^2$. Entonces $g(\mu) = k^2\mu^4$.

Por lo tanto, de (2.24) se obtiene

$$h(\mu) = \int \frac{d\mu}{k\mu^2} = -\frac{1}{k} \left(\frac{1}{\mu} \right) \quad . \quad (2.27)$$

De esta forma, una transformación inversa,

$$y^* = \frac{1}{y} \quad ,$$

producirá una varianza constante en la variable transformada.

En general, si $\sigma_y = k\mu^\alpha$, entonces $g(\mu) = k^2\mu^{2\alpha}$. Por lo tanto, de (2.24) se obtiene

$$h(\mu) = \int \frac{d\mu}{k\mu^\alpha} = \frac{1}{k} \left(\frac{\mu^{1-\alpha}}{1-\alpha} \right) . \quad (2.28)$$

Por lo tanto, si la relación observada es del tipo $\sigma_y \propto \mu^\alpha$, transformando los datos

$$y^* = h(y) = y^{1-\alpha}$$

se obtienen nuevas variables con varianza constante.

Para todos estos casos mencionados eligiendo k adecuadamente se puede alcanzar la varianza unidad.

De la expresión (2.22) se deduce que la transformación obtenida por este procedimiento sólo estabiliza las varianzas aproximadamente. Por lo tanto, es importante inspeccionar los residuos de la variable transformada para comprobar si efectivamente las varianzas se han estabilizado.

Como las transformaciones resultantes $h(y)$ son del tipo potencial, en las subsecciones siguientes vamos a estudiar una de las familias de transformaciones potenciales más importantes, la familia de *transformaciones de Box-Cox*.

Box y Cox (1964) desarrollaron un procedimiento, que estudiaremos en la subsección 2.3.2, para elegir una transformación de la familia de transformaciones potenciales de la forma $y^* = y^{(\lambda)}$

$$y^{(\lambda)} = \begin{cases} \frac{(y+m)^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(y+m) & \text{si } \lambda = 0 \end{cases} \quad (2.29)$$

donde la constante m se elige de forma que $y+m$ sea positivo. Esta familia de transformaciones incluye como casos particulares, salvo términos constantes, las estudiadas en los apartados anteriores.

Seguidamente, se muestra en la Tabla 3-4 un resumen de algunos casos típicos para

conseguir varianza constante.

Tabla 3-4. Transformaciones para estabilizar las varianzas

Relación entre σ_i y μ	α	$\lambda = 1 - \alpha$	Transformación	Comentario
$\sigma_i = k\mu_i^2$	2	-1	Recíproca	
$\sigma_i = k\mu_i^{3/2}$	3/2	-1/2	Inversa de la raíz	
$\sigma_i = k\mu_i$	1	0	Logarítmica	Datos log-normal
$\sigma_i = k\mu_i^{1/2}$	1/2	1/2	Raíz cuadrada	Datos de Poisson
$\sigma_i = k$	0	1	No transformar	

2.3.2. Transformaciones para corregir la falta de normalidad

Ante la violación de la hipótesis de normalidad surge de inmediato la cuestión de qué hacer en tal situación. Evidentemente una de las *alternativas* sería ignorar este hecho y proceder con el estudio como si la hipótesis fuera cierta. Resolver la situación de esta forma no es recomendable puesto que puede conducir a conclusiones incorrectas. Una segunda alternativa sería intentar *transformar* dichos datos para conseguir la normalidad. Así, por ejemplo, en situaciones en las que se conoce la distribución de la variable respuesta, el análisis de la varianza se puede abordar utilizando las siguientes transformaciones normalizadoras

- * Para datos binomiales, $y_{ij} \rightsquigarrow B(n_i, p_i)$, la transformación a utilizar es el arcoseno cuya expresión óptima viene dada por

$$y_{ij}^* = \arcsen \frac{\sqrt{y_{ij} + 3/8}}{\sqrt{n_i + 3/4}} .$$

- * Para datos que siguen la distribución de Poisson, la transformación óptima es la raíz cuadrada

$$y_{ij}^* = \sqrt{y_{ij}} \quad \text{o bien} \quad y_{ij}^* = \sqrt{1 + y_{ij}}$$

- * Para datos que tienen una distribución log-normal, la transformación apropiada es la logarítmica

$$y_{ij}^* = \log y_{ij}$$

Los procedimientos que se han utilizado para conseguir la normalidad de los datos mediante transformaciones de los mismos son múltiples, siendo uno de los más empleados en la práctica el uso de la familia Box-Cox.

Transformaciones de Box-Cox.

Consideremos la familia de transformaciones dada por (2.29) para $m = 0$

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln y & \text{si } \lambda = 0 \end{cases} \quad (2.30)$$

Box y Cox desarrollaron un procedimiento donde el parámetro λ se estima a partir de los datos. Dicho procedimiento consiste en suponer que las variables transformadas $y^{(\lambda)}$ siguen una distribución $N(\mu, \sigma)$ para algún λ adecuado.

A partir de la relación entre la función de densidad de los datos originales, y , y de los datos transformados, $y^{(\lambda)}$, se obtiene la log-verosimilitud $\mathbb{L}(\lambda, \mu, \sigma^2)$, dada por

$$\mathbb{L}(\lambda, \mu, \sigma^2) = -\frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln 2\pi + (\lambda - 1) \sum_i \ln y_i - \frac{1}{2\sigma^2} \sum_i \left(\frac{y_i^\lambda - 1}{\lambda} - \mu \right)^2. \quad (2.31)$$

Sustituyendo los valores de μ y σ^2 que maximizan la función (2.31) se obtiene la siguiente expresión para $\mathbb{L}(\lambda)$

$$\mathbb{L}(\lambda) = -\frac{N}{2} \ln \left[\frac{\frac{1}{N} \sum_i \left(y_i^{(\lambda)} - \bar{y}^{(\lambda)} \right)^2}{\bar{y}^{(\lambda-1)^2}} \right] \quad (2.32)$$

donde $\bar{y} = \left(\prod_{i=1}^N y_i \right)^{1/N}$ es la media geométrica de las observaciones.

Definiendo, por simplicidad, una nueva variable $z(\lambda)$ como

$$z(\lambda) = \frac{y^\lambda - 1}{\lambda \bar{y}^{\lambda-1}}, \quad (2.33)$$

la ecuación (2.32) se transforma en

$$\mathbb{L}(\lambda) = -\frac{N}{2} \ln \sum_i [z_i(\lambda) - \bar{z}(\lambda)]^2, \quad (2.34)$$

salvo términos constantes.

El valor de λ que maximice la ecuación (2.34) será el estimador de máxima verosimilitud. Dicho valor de λ también se puede obtener por procedimientos gráficos.³

Cuando los términos de error se distribuyen normalmente pero tienen las varianzas desiguales, una transformación de las observaciones para estabilizar las varianzas puede afectar a la normalidad. Afortunadamente, en la práctica la falta de normalidad y heterocedasticidad suelen ir unidos. Además, la transformación que ayuda a corregir la desigualdad de varianzas generalmente también es efectiva para conseguir que la distribución de los términos de error se aproxime más a la distribución Normal. Se recomienda, sin embargo, estudiar los residuos de los datos transformados para estar seguros de que la transformación aplicada ha sido efectiva estabilizando las varianzas y haciendo que la distribución de los términos de error sea razonablemente normal.

2.4. Efectos de la violación de las hipótesis del modelo

En las subsecciones precedentes, hemos mostrado como el análisis de los residuos y otras técnicas de diagnóstico pueden ser útiles para asegurar la adecuación del modelo ANOVA a los datos que se investigan. También discutimos el uso de transformaciones, principalmente para estabilizar las varianzas y obtener distribuciones del término error más próximas a la Normal. La cuestión que ahora surge es qué efectos produce la violación de las hipótesis establecidas en el modelo unifactorial, es decir la violación de las hipótesis:

- i) Independencia de los errores.
- ii) Normalidad de los errores.
- iii) Igualdad de varianzas de tratamientos.

Una revisión minuciosa de los estudios realizados sobre el efecto de la violación de las hipótesis del modelo se debe a Scheffé 1959. A continuación, exponemos las conclusiones más importantes de dicho trabajo.

Dependencia entre los términos de error

La falta de independencia entre los términos de error puede producir graves efectos en la inferencia, tanto en el modelo de efectos fijos como en el de efectos aleatorios. Puesto que este incumplimiento es a menudo difícil de corregir, es importante evitarlo cuando sea factible. Una forma de conseguirlo es mediante el uso de la aleatorización y otra forma es modificando el modelo.

³Para un desarrollo más detallado vease Peña (1989)

Desviación de los errores de la normalidad

Antes de tratar el efecto de la no-normalidad del término de error, vamos a establecer una distinción entre los dos tipos de inferencias que se realizan sobre un modelo de análisis de la varianza:

- a) *Inferencias sobre las medias*: presente en los modelos de efectos fijos y que concierne únicamente a los efectos del factor (estimación puntual o por intervalos para medias, contraste de igualdad de medias, contrastes múltiples de medias, etc.)
- b) *Inferencias sobre las varianzas*: presente tanto en el modelo de efectos fijos como en el de efectos aleatorios y que incluye, entre otras, la estimación puntual o por intervalos de la varianza o de componentes de la varianza e inferencia sobre cociente de varianzas.

El efecto de la desviación de la normalidad afecta de manera desigual a estos tipos de inferencias. En general, si la desviación de la normalidad no es muy grave, es poco importante en la inferencia sobre medias y más grave en la inferencia sobre varianzas.

Además, en ambos casos, las estimaciones puntuales continúan siendo insesgadas y en los contrastes de hipótesis se alteran el error de tipo I y su potencia. Generalmente, dicho error es ligeramente mayor que el nominal y la potencia menor que la teórica.

Heterocedasticidad del término error

Cuando las varianzas de los términos de error de cada nivel son desiguales, el test F para la igualdad de medias de los tratamientos en el modelo de efectos fijos unifactorial está poco afectado si todos los tamaños muestrales de los niveles del factor son iguales o difieren muy poco, pero no ocurre lo mismo cuando hay grandes diferencias entre dichos tamaños muestrales o cuando una varianza es mucho mayor que las otras.

En el caso de dos grupos de igual tamaño, la violación de la hipótesis de homocedasticidad para n grande es aún menos importante que en el caso general, así, por ejemplo, la potencia del test F será la teórica si y sólo si los tamaños de cada grupo son iguales. Sin embargo, cuando consideramos el caso de I grupos, con $I > 2$, la violación de la hipótesis de igualdad de varianzas tiene algún efecto incluso cuando los modelos son equilibrados, si bien los efectos parecen poco importantes.

Concretamente, cuando las varianzas son desiguales y los tamaños de las muestras son iguales o aproximadamente iguales, entonces:

- * El verdadero error de tipo I aumenta. Es un poco más grande que el nivel nominal α .

- * La potencia del test F es aproximadamente la teórica bajo la suposición de varianzas iguales.
- * El procedimiento de comparaciones múltiples de Scheffé, basado en la distribución F , está poco afectado.
- * El test F y los análisis relacionados son robustos frente a la desigualdad de las varianzas de los términos de error de cada nivel.

Para el modelo unifactorial de efectos aleatorios, aunque el modelo sea equilibrado, la no homogeneidad de las varianzas del error puede afectar a las inferencias sobre las componentes de la varianza.

Conclusiones

Lo expresado en esta subsección se pueden resumir brevemente como

- i) El efecto de la correlación en las observaciones puede ser grave en las inferencias sobre medias.*
- ii) La no-normalidad tiene efecto pequeño en las inferencias sobre las medias pero efectos graves en las inferencias sobre varianzas.*
- iii) La desigualdad de varianzas tiene efecto pequeño en las inferencias sobre medias si los grupos son del mismo tamaño y grave si los grupos difieren mucho de tamaño.*

Bibliografía utilizada

- * **García Leal, J. & Lara Porras, A.M.** (1998). *“Diseño Estadístico de Experimentos. Análisis de la Varianza.”* Grupo Editorial Universitario.
- * **Lara Porras, A.M.** (2000). *“Diseño Estadístico de Experimentos, Análisis de la Varianza y Temas Relacionados: Tratamiento Informático mediante SPSS”* Proyecto Sur de Ediciones.