"El término de error experimental (ε) en los modelos estadísticos de análisis de varianza. Condiciones subyacentes en el ANVA referidas a la variable aleatoria ε"

Por Francisco Javier TEJEDOR TEJEDOR

Los resultados de un experimento pueden estar afectados por dos causas fundamentales o fuentes de variación: la diversidad de los tratamientos experimentales, es decir, los distintos valores atribuidos a la variable independiente (variación intergrupo) y el conjunto de factores extraños al experimento (variación intragrupo o variación debida al error experimental).

Desde hace no pocos años, la metodología experimental está estudiando procedimientos para lograr una mayor exactitud en los resultados de los trabajos experimentales, intentando, por una parte, reducir al máximo el error experimental y, por otra, estudiarlo desde el punto de vista estadístico, puesto que su inclusión en los modelos así lo requiere.

Esquemáticamente, los procedimientos para reducir el error experimental serían:

a) Aumento del tamaño de las muestras utilizadas en el experimento (recordar la relación existente entre el error típico de un estadístico y el tamaño de la muestra).
b) Utilización de técnicas más perfeccionadas que aseguren en el mayor grado posible la uniformidad de aplicación de los tratamientos experimentales, la obtención de medidas precisas, etc.
c) Utilización de un diseño experimental adecuado. Uno de los objetivos básicos del diseño experimental es la minimización del error. Se consigue, por ejemplo, incluyendo dos o más variables independientes (diseños factoriales), controlando otras, lo que reducirá el número de variables extrañas influyentes en el experimento.

¿Por qué tanta preocupación por el término de error? Sencillamente, porque el contraste de la hipótesis nula formulada se efectúa comparando la varianza estimada entre las medias de los grupos (varianza intergrupo o cuadrado medio intergrupo, que recoge la acción de los tratamientos experimentales) y la varianza estimada dentro de los grupos (varianza intragrupo, cuadrado medio intragrupo o varianza del error experimental, que recoge únicamente la variación entre los sujetos, debida al azar, sin reflejar el efecto de la variable independiente o tratamientos; por esta razón, con la debida justificación matemática, sabemos que la varianza intragrupo es una estimación insesgada de la varianza de la población).

Recordemos la fórmula de F, estadístico de contraste, $F = \dfrac{\sigma^2_{inter}}{\sigma^2_{intra}}$.

De la misma forma, en los experimentos factoriales la hipótesis nula formulada para cada uno de los factores y para las interacciones entre los mismos se contrasta en todos los casos comparando separadamente la variación debida a cada uno de los factores o a las interacciones con la variación debida al error experimental.

Así, pues, conviene intentar minimizar el error experimental al plantearse la realización de un experimento. Pero no será posible eliminarlo totalmente, por lo que se hace necesario, si queremos ser realistas, incluirlo en el modelo, siendo, por tanto, necesario su estudio desde el punto de vista estadístico. Este último aspecto es el objetivo de nuestro trabajo.

CONDICIONES SUBYACENTES EN EL ANVA REFERIDAS A LA VARIABLE ALEATORIA ε

El modelo estadístico supuesto con más frecuencia en aplicaciones de análisis de varianza (ANVA) es un modelo lineal al cual se añaden algunas restricciones. El modelo para el caso más simple —un factor— es:

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

siendo:

μ = Efecto medio verdadero.

α_i = Efecto medio verdadero del iésimo tratamiento.

ε_{ij} = Efecto medio verdadero de la jotaésima unidad experimental sometida al iésimo tratamiento. Incluye también los efectos de todos los factores extraños al experimento.

¿Por qué aparece ε en el modelo? Las razones por las que aparece en el modelo ε_i, definido como término de perturbación, como término de error, podrían ser:

a) No conocemos todos los factores o variables independientes que pueden influir en la variable dependiente estudiada. Por esta razón dijimos que ε incluye el efecto de los factores extraños no controlados en un determinado experimento.

b) Puede entenderse como el elemento que refleja la aleatoriedad de las respuestas humanas.

c) Puede entenderse como el elemento que recoge los errores de observación o medida.

d) Puede entenderse como el elemento que refleja las diferencias individuales (sujetos que con valores α iguales no presentan iguales valores X; por ejemplo, sujetos con igual inteligencia no tienen el mismo rendimiento).

La inclusión de ε en el modelo es, ante todo, una postura realista, y más en el contexto de la experimentación en las ciencias humanas. Conviene recordar que un modelo, además de permitir un adecuado manejo estadístico, debe suponer ante todo credibilidad. No incluir ε_i equivaldría a aceptar relaciones funcionales exactas entre las variables, lo que en el contexto de las ciencias humanas supondría alejarnos de la realidad.

Las suposiciones fundamentales que conlleva el modelo son:

1.ª $X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$ (es decir, la idoneidad del propio modelo; se analiza en términos estadísticos de linealidad y aditividad).

2.ª Referida a α_i. Dependerá que sigamos el modelo I (efectos fijos), el modelo II (efectos aleatorios) o el modelo III (modelo mixto). Las suposiciones son:

Modelo I: $\sum^{t} \alpha_i = 0$.
Modelo II: α_i DNI $(0, \sigma^2_\alpha)$.
Modelo III: Una combinación de ambos (se entiende que para tener lugar este modelo el diseño ha de incluir, como mínimo, dos factores).

3.ª ε_{ij} es una variable aleatoria distribuida normal e independientemente con media 0 y varianza σ^2. Se expresa: DNI $(0, \sigma^2)$.

Pretendemos comentar el cumplimiento de estas tres condiciones formuladas para ε_i. Es decir, comprobar que las muestras que se incluyen en el ANVA son normales, tienen varianzas homogéneas y son independientes.

El esquema que seguiremos en la exposición será:

1. Contrastes para verificar el cumplimiento de esas condiciones.
2. Efectos del no cumplimiento de alguna de las condiciones.
3. Pruebas alternativas no paramétricas.

1. CONTRASTES PARA VERIFICAR EL CUMPLIMIENTO DE LAS CONDICIONES DE LA VARIABLE ALEATORIA ε

1.1. NORMALIDAD

1.1.1. Representación gráfica

La representación gráfica en papel probabilístico de las frecuencias acumuladas puede permitirnos obtener una cierta idea del «grado de normalidad» de la variable. Si la variable es normal, el gráfico sería una línea recta.

Se trata de un procedimiento aproximativo. Por requerir un elevado número de frecuencias en cada muestra, lo que no es frecuente en aplicaciones de ANVA, no estimamos el procedimiento demasiado útil.

1.1.2. Asimetría y curtosis

Se podrían calcular los coeficientes de asimetría y curtosis, que deberían resultar próximos a 0 y 3, respectivamente.

No parece tampoco un procedimiento recomendable.

1.1.3. Contraste W de Shapiro y Wilk (1)

Extraordinariamente útil en muestras menores de 50. El procedimiento de aplicación consiste en:

— Ordenar las puntuaciones de la muestra de mayor a menor.
— Calcular la suma de cuadrados en cada muestra respecto a la media de esa muestra:

$$S^2 = \sum_{1}^{n_i} X^2 - \frac{(\sum_{1}^{n_i} X)^2}{n_i}$$

(no confundir S^2 con la varianza)

(1) Más detalles sobre la prueba, tablas y ejemplos de aplicación pueden encontrarse en Ruiz-Maya (página 234).

— Calcular b:

$$b = \sum_{1}^{k} a_{n-i+1} (x_{n-i+1} - x_i)$$

Los coeficientes a están tabulados.

Si n es par, $k = n/2$.

Si n es impar, $k = (n - 1)/2$.

— Calcular b^2.

— Obtener el estadístico $W = \dfrac{b^2}{S^2}$.

— El estadístico de contraste W está tabulado. La interpretación del W obtenido se realiza en los siguientes términos: Si el valor W obtenido es mayor que el valor W tabular para un determinado tamaño de muestra y un determinado nivel de significación, aceptamos la hipótesis de que la muestra es normal. Nótese que, en contra de lo que es usual, la hipótesis de normalidad se rechaza para valores más pequeños de los $W_{(n,\alpha)}$.

1.1.4. **Contraste χ^2**

Hamdam opina que el número mínimo de intervalos necesarios para aplicar la prueba χ^2 en el contraste de normalidad ha de ser 10. Ostle acepta la opinión de Hamdam y estima que las frecuencias mínimas en los intervalos extremos no puede ser inferior a 5, lo que supondría aproximadamente una muestra de 600 sujetos (2).

No creemos que sea útil esta prueba para contrastar la normalidad en las aplicaciones de ANVA, ya que las muestras suelen ser de tamaño reducido. La prueba de «bondad de ajuste» nos parece útil para una muestra de tamaño grande.

1.1.5. **Contraste de Kolmogorov-Smirnov**

Su aplicación no necesita número mínimo de intervalos ni número mínimo de frecuencias en cada intervalo (3):

— Se calculan las frecuencias empíricas relativas acumuladas (F_e).

— Se calculan las frecuencias teóricas relativas acumuladas (F_t).

— Se obtiene la diferencia, en valor absoluto, $D = |F_t - F_e|$.

— Seleccionamos el valor máximo de D.

— Se compara con el valor tabular. Si el D seleccionado es mayor que el valor tabular $D_{(n,\alpha)}$, se rechaza la hipótesis de normalidad de la muestra.

1.2. HOMOGENEIDAD DE LAS VARIANZAS (HOMOCEDASTICIDAD)

1.2.1. **Contraste de Bartlett**

Este contraste es sensible a la no-normalidad, por lo que no es recomendable su utilización cuando ésta es muy notoria. Box opina que el contraste encubre diferencias entre las varianzas si la curtosis es positiva e indica diferencias inexistentes si la curtosis es negativa (4). Para su aplicación:

---

(2) Más detalles de la prueba pueden encontrarse en numerosos textos de estadística. Ejemplos de su aplicación a las ciencias humanas pueden verse en Siegel (p. 64) y en Tejedor (p. 174).

(3) Más detalles de la prueba, ejemplos y tablas en Siegel (p. 69).

(4) Más detalles sobre la prueba y ejemplos de aplicación pueden encontrarse en Ostle (p. 161) y en Ruiz-Maya (p. 238).

— Se obtienen las cuasivarianzas de las muestras:

$$s_i^2 = \frac{\sum^{n_i} x^2}{n_i - 1}$$

($n_i - 1 = v_i$, son los grados de libertad de cada muestra).

— Se obtiene la cuasivarianza ponderada:

$$s^2 = \frac{\sum^{k} v_i s_i^2}{\sum^{k} v_i} = \frac{\sum^{k} v_i s_i^2}{v}$$

($\sum^{k} v_i = v$, grados de libertad del total; k es el número de muestras o grupos).

— Se calcula el estadístico de contraste B:

$$B = v \cdot \ln s^2 - \sum^{k} v_i \ln s_i^2$$

(ln = logaritmo neperiano).

— El estadístico B sigue aproximadamente la distribución χ^2 con $k-1$ grados de libertad, para valores $v_i \geqq 5$.
— Si $v_i \geqq 5$ se compara el valor B obtenido con el valor tabular de $\chi^2_{(k-1)}$, y se interpreta en la forma conocida: Si $B > \chi^2$ se rechaza la hipótesis de homogeneidad de varianzas.
— Si $v_i \leqq 5$ existen tablas que proporcionan mayor precisión que las de χ^2: tablas de Merrigton-Thompson, que pueden encontrarse en Bennett-Franklin, página 198.
— Si B resultara sólo un poco mayor que χ^2, habría que corregir el valor de B y obtener el valor B/C, que también se distribuye aproximadamente como $\chi^2_{(k-1)}$. El valor C es:

$$C = 1 + \frac{\sum^{k} (1/v_i) - 1/v}{3(k-1)}$$

1.2.2. **Contraste de Cochran**

Es también sensible a la no-normalidad. Su aplicación (5) exige que las n de las muestras sean iguales:

— Se obtienen las varianzas de las muestras.
— Se obtiene el estadístico R:

$$R = \frac{s^2 \text{ mayor}}{\text{Suma de las varianzas}}$$

— Los valores $R_{(n,r)}$ están tabulados. Si $R > R_{(n,r)}$ se rechaza la hipótesis de homogeneidad de las varianzas.

(5) Más detalles sobre la prueba y tablas pueden encontrarse en Tejedor (p. 216).

1.2.3. Contraste de Levene

No es sensible a la no-normalidad. Utiliza la técnica ordinaria de ANVA. Toma como datos del ANVA los valores absolutos de las diferencias entre cada puntuación y su media (6).

1.2.4. ANVA con los logaritmos de las cuasivarianzas

Aplicable cuando es notoria la no-normalidad. Es necesario establecer una etapa de submuestreo, pudiendo entonces ocurrir que las submuestras tengan o no igual tamaño (debiendo, por tanto, seguir el modelo adecuado, es decir, el modelo equilibrado o el no equilibrado) (7).

1.2.5. Contraste de Lehman

Descrito en Ghosh (1972). Requiere el cálculo de tres coeficientes para cada combinación del número de varianzas y del tamaño de la muestra. Ghosh ha tabulado estos coeficientes para $k < 10$ y $n \leqslant 5$, siendo todas las n iguales.

1.3. INDEPENDENCIA

1.3.1. Coeficiente de correlación serial de separación 1

Se forman los pares:

X	Y
X_i	X_{i+1}
X_{i+1}	X_{i+2}
...	...

Se obtiene el valor Rs (fórmula de Pearson entre X e Y). Para valores $n \leqslant 30$ deben consultarse las tablas de Anderson (Bennet-Franklin, página 686).

Para valores de $n > 30$, Rs se distribuye de forma aproximadamente normal con:

$$\mu = -\frac{1}{n-1}$$

$$\sigma^2 = \frac{n-2}{(n-1)^2}$$

Por lo que $\dfrac{(n-1)\,Rs + 1}{\sqrt{n-2}}$ se distribuye aproximadamente N (0, 1).

1.3.2. Contraste de rachas

Se convierte la variable en dicotómica (+ ,—), según que una puntuación sea mayor o menor que la anterior.

Otro criterio para dicotomizar la variable es asignar el signo + ó — según que la puntuación sea mayor o menor que la mediana.

Se determina: n_1 = número de signos +; n_2 = número de signos —; r = número de

(6) Más detalles sobre la prueba pueden encontrarse en Glass y Stanley (p. 374).
(7) Más detalles en Ruiz-Maya (p. 240).

rachas (las rachas son tantas como bloques de signos iguales resulten al dicotomizar la variable).

Si n_1 ó n_2 son menores que 20, existen tablas especiales que dan los valores críticos (8).

Si n_1 y n_2 son mayores de 20, la r se distribuye de forma aproximadamente normal con:

$$\mu = \frac{2\, n_1 n_2}{n_1 + n_2} + 1$$

$$\sigma^2 = \frac{2\, n_1 n_2 (2\, n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}$$

Por lo que $\dfrac{r - \mu}{\sigma}$ se distribuye de forma aproximadamente normal.

2. EFECTOS DEL NO CUMPLIMIENTO DE LAS CONDICIONES

2.1. NORMALIDAD

El efecto de la no-normalidad se traduce en una modificación del nivel nominal de significación, α, que se convierte en cada circunstancia en lo que podemos llamar «nivel real de significación».

Pearson y Box y Andersen han estudiado las probabilidades reales (nivel real de significación) asociadas al contraste de hipótesis para diferentes grados de desviación respecto a la simetría y a la curtosis.

Las conclusiones son que las probabilidades reales y las nominales son prácticamente iguales. Por ejemplo, para una probabilidad nominal $\alpha = 0{,}05$, la probabilidad real, para asimetría ± 1 y curtosis 4, resulta ser 0,049.

Por tanto, juzgamos irrelevante el cumplimiento de la condición de normalidad para el contraste de medias en el ANVA.

Mayor influencia puede tener la no-normalidad en el contraste de la homogeneidad de las varianzas, por lo que seguimos recomendando el conocimiento del grado de aproximación de las muestras a la normalidad.

De las pruebas comentadas juzgamos como más adecuada la prueba de Shapiro y Wilk.

Por último, advertir que podría ser más grave la no-normalidad en los modelos de efectos aleatorios que en los modelos de efectos fijos.

2.2. HOMOGENEIDAD DE VARIANZAS

Su no cumplimiento tiene igualmente un efecto de modificación del nivel de significación nominal, pero mucho más importante que el comentado para casos de no-normalidad.

El tema ha sido extensamente tratado: Scheffé, Lindquist, Cochran, Box, Box y Andersen, Boneau...

(8) Más detalles sobre la prueba, tablas y ejemplos en Siegel (p. 74).

Las conclusiones, prácticamente coincidentes en todos, pueden resumirse así:

— Con n iguales, los efectos de la heterogeneidad de varianzas son mínimos, aunque la diferencia entre probabilidades α nominales y reales puede variar según el número de grupos y según el valor de la razón de heterogeneidad.

— Si las muestras menores (mayores) son extraídas de las poblaciones con mayor (menor) varianza, el valor α real aumenta.

— Si las muestras menores (mayores) son extraídas de las poblaciones con menor (mayor) varianza, el valor α real disminuye.

Veamos unos ejemplos (9):

Razón entre las varianzas de las muestras	**Tamaño de la muestra**	α **nominal**	α **real**
1, 1, 1, 1, 3	5, 5, 5, 5, 5	0,05	0,074
1, 1, 3	9, 5, 1	0,05	0,17
1, 1, 1, 1, 3	1, 5, 5, 5, 9	0,05	0,025

En resumen, aceptamos Bartlett como el mejor contraste. Si la no-normalidad fuera muy acusada, comprobar la homogeneidad de varianzas mediante la prueba de ANVA con los logaritmos neperianos de las cuasivarianzas o el contraste de Levene, que no es sensible a la no-normalidad (similitud entre las α nominales y reales), y que presenta un grado de eficiencia respecto al de Bartlett del 90 por 100 aproximadamente.

2.3. INDEPENDENCIA

Quizá su violación suponga los efectos más graves, pero es la condición más fácil de conseguir con un proceso adecuado de aleatorización.

Pueden presentarse problemas de dependencia con «muestras repetidas» (diferentes medidas a un mismo grupo de sujetos), siendo lo más recomendable en este caso la utilización de alguna prueba alternativa no paramétrica (prueba Q de Cochran).

2.4. VALORES F MENORES QUE UNO

Es relativamente frecuente obtener valores F menores que uno. Ostle opina que puede deberse a:

— incumplimiento de alguna condición respecto a ε_i,

— inadecuación del modelo.

Debemos, pues, en primer lugar, verificar, si no se ha hecho, el cumplimiento de las condiciones referidas a ε_i. Si todas ellas se cumplen, aconseja:

a) Como respuesta más inmediata, más simple, interpretar el inverso del valor F obtenido invirtiendo igualmente los grados de libertad; es decir, interpretar 1/F con v_2, v_1 grados de libertad. Si 1/F es mayor que $F_{(v_2, v_1\,;\,\alpha)}$, se rechaza la hipótesis de igualdad de medias.

b) Como respuesta más exigente, analizar la inadecuación del modelo que pudiera atribuirse a la falta de aditividad, la cual, a su vez, pudiera deberse a:

(9) Una tabla de información más completa puede encontrarse en Glass y Stanley (p. 372).

— observaciones equivocadas,
— existencia de interacciones,
— verdaderos efectos multiplicativos.

Deberemos entonces revisar el proceso de obtención de puntuaciones; estudiar las interacciones, si el modelo lo permite, y, por último, recurrir a las transformaciones de la variable, que sobre todo corrigen la no aditividad, teniendo efectos secundarios positivos en cuanto que contribuyen a mejorar la normalidad y a favorecer la homogeneidad de las varianzas.

2.5. TRANSFORMACIONES EN LA VARIABLE

Las transformaciones más recomendables (10) serían:

a) **Transformación logarítmica.**—Si la verdadera relación entre las variables es multiplicativa: $X = \mu \alpha \varepsilon$, la transformación logarítmica convertiría la relación en aditiva:

$$X' = \log X = \log \mu + \log \alpha + \log \varepsilon$$

Una relación es multiplicativa cuando el incremento de α en una cantidad fija no supone el incremento de X en esa misma cantidad, sino en una determinada proporción.

b) **Transformación raíz cuadrada.**—Si las observaciones son frecuencias o siguen la distribución de Poisson, puede utilizarse la transformación:

$$X' = \sqrt{X}$$

Si $n < 10$, Bartlett aconseja que se haga mejor $X' = \sqrt{X + 0{,}5}$, y Anscombe opina que da mejores resultados la transformación $X' = \sqrt{X + 3/8}$.

c) **Transformación arco seno.**—Si las observaciones son tantos por uno (distribución binomial) puede utilizarse la transformación $X' = \mathrm{arc} \cdot \mathrm{sen} \sqrt{p}$, que puede expresarse también como $X' = \mathrm{sen}^{-1} \sqrt{p}$.

Bartlett opina que los valores $p = 0$ sean sustituidos por $1/2n$ y los valores $p = 1$ por $1 - (1/2n)$.

Anscombe opina que obtendremos mejores resultados tomando $p = \dfrac{x + (3/8)}{n + (3/4)}$.

3. PRUEBAS ALTERNATIVAS NO PARAMETRICAS

Si alguna de las suposiciones que conlleva el modelo no se cumplen, habrá que tener mucho cuidado con la interpretación de la hipótesis y con las conclusiones a obtener, sobre todo si el valor F obtenido está cercano al valor F tabular.

En estas circunstancias recomendamos la utilización de alguna prueba alternativa no paramétrica. A continuación citamos algunas, especificando el modelo de ANVA adecuado y las condiciones que requiere su aplicación:

(10) Tablas de estas transformaciones pueden encontrarse en Ruiz-Maya.

Prueba	Modelo	Requiere	No requiere
Kruskal-Wallis	Un factor	Independencia	Normalidad Homogeneidad de varianzas
Q de Cochran	Un factor	n iguales	Normalidad Homogeneidad de varianzas Independencia
Welch	Un factor	Normalidad Independencia	Homogeneidad de varianzas
Test Matricial	Dos factores	Normalidad Independencia	Homogeneidad de varianzas
Rangos de Friedman	Dos factores	Independencia	Normalidad Homogeneidad de varianzas

4. EJEMPLO

En una Escuela de Idiomas se intenta experimentar cuál de los cinco métodos audiovisuales (diapositivas, cassettes y textos) existentes en el mercado produce mejores resultados.

Se forman cinco grupos de cinco sujetos asignados al azar a cada grupo.

Todos los grupos tienen una hora diaria de clase durante un curso académico, impartida por el mismo profesor.

A fin de curso se somete a todos los alumnos a una misma prueba. Los resultados (imaginarios) aparecen a continuación:

1	2	3	4	5
6	4	2	3	4
8	7	4	8	2
3	6	1	6	5
5	8	5	2	1
9	7	2	7	5

Nuestro objetivo fundamental será contrastar las condiciones que conlleva el modelo de ANVA respecto a la variable aleatoria ε_i. Posteriormente veremos si se produce algún tipo de anomalía en el propio modelo que nos impida dudar de la veracidad de nuestras conclusiones al interpretar la hipótesis nula (todos los grupos tienen el mismo rendimiento medio).

Puede sernos útil a lo largo del problema la información auxiliar del siguiente cuadro de datos:

	1	2	3	4	5	Total
n	5	5	5	5	5	25
$\sum X$	31	32	14	26	17	120
$\sum X^2$	215	214	50	152	71	702

4.1. NORMALIDAD

Aplicamos el contraste de Shapiro y Wilk. Comenzamos ordenando en cada grupo las puntuaciones de mayor a menor:

1	2	3	4	5
9	8	5	8	5
8	7	4	7	5
6	7	2	6	4
5	6	2	3	2
3	4	1	2	1

— Calculamos la suma de cuadrados en cada grupo:

$$SC_i = \Sigma X^2 - \frac{(\Sigma X)^2}{n}$$

$$SC_1 = 215 - \frac{(31)^2}{5} = 22{,}80$$

$$SC_2 = 9{,}20 \quad ; \quad SC_3 = 10{,}80 \quad ; \quad SC_4 = 26{,}80 \quad ; \quad SC_5 = 13{,}20$$

— Establecemos las diferencias $(X_{n-i+1} - X)$ para cada grupo. Como las muestras tienen todas el mismo tamaño, $n = 5$, obtenemos en todos los casos $k = (n - 1)/2$ diferencias:

Grupo	1	2	3	4	5
Diferencias	6	4	4	6	4
	3	1	2	4	3

— Fijamos los valores a_{n-i+1}. Consultamos la tabla de valores a y tomamos los coeficientes correspondientes a cada diferencia; coeficientes que serán iguales para todas las muestras, ya que todas ellas tienen $n = 5$:

	a_{n-i+1}
1.ª diferencia	0,6646
2.ª diferencia	0,2413

— Obtenemos los valores b para cada grupo:

$$b = \sum^{k} a_{n-i+1} (X_{n-i+1} - X)$$

Para el primer grupo tendremos:

$$b = 6.\ 0{,}6646 + 3.\ 0{,}2413 = 4{,}7115$$

— Obtenemos los valores b^2; para el primer grupo, $b^2 = 22{,}1982$.

— Obtenemos para cada grupo $W = \frac{b^2}{SC}$; para el primer grupo:

$$W = \frac{22{,}1982}{22{,}80} = 0{,}9736$$

— Fijamos el valor $W_{(5\,;\,0{,}05)} = 0{,}762$.

— Interpretamos los valores W obtenidos. Para el primer grupo resulta: Como W obtenido es mayor que el W tabular, no rechazamos la hipótesis de normalidad de la muestra.

Los datos completos del problema aparecen en la tabla 1:

TABLA 1

Grupo o muestra	SC	$X_{n-i+1} — X$	a_{n-i+1}	b	b^2	W	Interpretación
1	22,80	6	0,6646	4,7115	22,1982	0,97	Normal
		3	0,2413				
2	9,20	4	0,6646	2,8997	8,4082	0,91	Normal
		1	0,2413				
3	10,80	4	0,6646	3,141	9,8659	0,91	Normal
		2	0,2413				
4	26,80	6	0,6646	4,9528	24,5302	0,91	Normal
		4	0,2413				
5	13,20	4	0,6646	3,3823	11,4399	0,87	Normal
		3	0,2413				

4.2. HOMOGENEIDAD DE LAS VARIANZAS

4.2.1. Contraste de Bartlett

Puesto que es aceptable el presupuesto de normalidad en las muestras, aplicamos el contraste de Bartlett. Para ello:

— Obtenemos las cuasivarianzas de las muestras:

$$s_i^2 = \frac{\Sigma X^2}{n — 1} — \frac{(\Sigma X)^2}{n(n — 1)}$$

Para el primer grupo tendremos:

$$s_1^2 = \frac{215}{4} — \frac{31^2}{20} = 5,7$$

$$s_2^2 = 2,3 \quad ; \quad s_3^2 = 2,7 \quad ; \quad s_4^2 = 6,7 \quad ; \quad s_5^2 = 3,3$$

— Fijamos los grados de libertad de cada muestra: $v_i = 4$.

— Fijamos los grados de libertad del total: $\Sigma v_i = v = 20$.

— Obtenemos la cuasivarianza ponderada total:

$$s^2 = \frac{\Sigma v_i s_i^2}{\Sigma v_i} = \frac{v_i \Sigma s_i^2}{v} =$$

$$= \frac{4(5,7 + 2,3 + 2,7 + 6,7 + 3,3)}{20} = 4,14$$

— Obtenemos el estadístico $B = v \ln s^2 — v_i \Sigma \ln s_i^2$

$B = 20 \ln 4,14 — 4 (\ln 5,7 + \ln 2,3 + \ln 2,7 + \ln 6,7 + \ln 3,3) =$

$= 20 \cdot 1,4207 — 4 (1,7405 + 0,8329 + 0,9932 + 1,9021 + 1,1939) =$

$= 28,4140 — 26,6506 = 1,7633$

— Fijamos el valor tabular de $\chi^2_{(4\,;\,0,05)} = 9,49$.

— Interpretamos: Como el valor de B es menor que el valor tabular de χ^2, no rechazamos la hipótesis de homogeneidad de las varianzas.

4.2.2. Contraste de Cochran

Reconfirmaremos la conclusión obtenida tras la aplicación del contraste de Bartlett utilizando el contraste de Cochran, ya que podemos hacerlo puesto que todas las muestras tienen el mismo tamaño. Para ello:

— Calculamos las varianzas en cada muestra:

$$s_i^2 = \frac{\sum X^2}{n} - \frac{(\sum X)^2}{n^2}$$

Para el primer grupo tendremos:

$$s_1^2 = \frac{215}{5} - \frac{31^2}{25} = 4{,}56$$

$$s_2^2 = 1{,}84 \quad ; \quad s_3^2 = 2{,}16 \quad ; \quad s_4^2 = 5{,}36 \quad ; \quad s_5^2 = 2{,}64$$

— La varianza mayor resulta ser: $s_4^2 = 5{,}36$.

— La suma de las varianzas: 16,56.

— El estadístico $R = \frac{5{,}36}{16{,}56} = 0{,}3237$.

— Interpretamos: Como el valor R obtenido es menor que el valor R tabular, $R_{(5,3\,;\,0{,}05)} = 0{,}5441$, no rechazamos la hipótesis de igualdad de varianzas.

4.3. INDEPENDENCIA

4.3.1. Coeficiente de correlación serial de separación 1

Formamos los pares seriados:

X	Y
6	8
8	3
3	5
5	9
9	4
4	7
7	6
6	8
8	7
7	2
2	4
4	1
1	5
5	2
2	3
3	8
8	6
6	2
2	7
7	4
4	2
2	5
5	1
1	5

Donde:

$\sum X = 115$
$\sum Y = 114$
$\sum X^2 = 687$
$\sum Y^2 = 676$
$\sum XY = 543$
$n = 24$

— El cálculo de la r_{xy} nos proporciona el valor $r_{xy} = -0{,}024$.

— La interpretación puede hacerse utilizando las tablas de Anderson o como la r obtenida en una muestra pequeña, mediante la fórmula:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}024\sqrt{22}}{\sqrt{1-0{,}024^2}} = 0{,}1126$$

— El valor de $t_{(22\,;\,0{,}05)} = 2{,}074$, luego el valor t obtenido es menor que el valor tabular, por lo que la r obtenida no es significativa. No podemos, por tanto, rechazar la hipótesis de independencia entre las muestras.

4.3.2. **Prueba de rachas**

Trataremos de confirmar la interpretación de la hipótesis de independencia aplicando la prueba de rachas, bajo el criterio de utilizar la mediana para dicotomizar la variable. La mediana resulta ser igual a 4,875. Tenemos, pues, la siguiente serie de signos:

$$+ + - + + - + + + + - - - + - - + + - + - - + - +$$

Determinamos: $n_1 = 14$; $n_2 = 11$; $r = 15$.

En tablas vemos que para $n_1 = 14$, $n_2 = 15$, al 0,05, el intervalo de rachas esperado por azar es 8 — 17.

Interpretamos: Como el número de rachas obtenido se encuentra entre los límites de dicho intervalo, no podemos rechazar la hipótesis de independencia.

4.4. CONCLUSIONES

Las condiciones requeridas por la variable aleatoria ε se cumplen en su totalidad, por lo que estimamos, en principio, legítima la aplicación de las técnicas de ANVA para resolver la situación experimental planteada.

La tabla de ANVA resultante para nuestro ejemplo es la tabla 2:

TABLA 2

F. de V.	SC	G. de L.	c. m.	c. m. e.	F
Intergrupo o tratamientos	53,2	4	13,3	$+ 5 \sum^{5} \alpha_i^2/4$	3,21
Intragrupo o error experimental	82,8	20	4,14		
Total	136	24			

El valor $F_{(4,20\,;\,0{,}05)} = 2{,}87$. Luego rechazamos la hipótesis nula, admitiendo que existe diferencia entre los rendimientos medios de los diferentes grupos.

Como no ha ocurrido que $F < 1$, no tiene lugar dudar sobre la idoneidad del modelo. Así, pues, aceptamos definitivamente como adecuada la prueba de ANVA para dar respuesta a la situación experimental planteada y la conclusión obtenida tras el contraste de la hipótesis formulada.

En general, recomendamos la consulta de los escritos originales de cada una de las pruebas presentadas (ver bibliografía). Para su utilización es suficiente consultar las referencias que se citan en cada caso.

BIBLIOGRAFIA

Bartlett, M. S.: «The use of transformation». **Biometrics**, vol. 3, p. 39, 1947.

Bennett, C. A. y Franklin, N.: «Statistical analysis». N. Y., Wiley, 1959.

Box, G. E.: «Some theorems on quadratic forms applied the study of analysis of variance problems. I. Effect of inequality of analysis of variance in the one-way classification». **Annals of Mat. Sta.**, 25, 290-302, 1954.

Cochran, W. G.: «Some consequences when the assumptions for the analysis of variance are not satisfied». **Biometrics**, vol. 3, p. 22, 1947.

Einsenhart, C.: «The assumptions underlying the analysis of variance». **Biometrics**, vol. 3, p. 1, 1947.

Ghosh, B. K.: «On Lehman's test for homogeneity of variances». **J. Roy. Sta. Soc.**, 34, 221-234, 1972.

Glass, G. y Stanley, J.: «Métodos estadísticos aplicados a las ciencias sociales». Prentice-Hall-Internacional, Madrid, 1974.

Ostle, B.: «Estadística aplicada». Limusa-Wiley, México, 1965.

Ruiz-Maya, L.: «Métodos estadísticos de investigación. Introducción al análisis de varianza». Instituto Nacional de Estadística, Madrid, 1977.

Siegel, S.: «Estadística no paramétrica». Trillas, México, 1975.

Shapiro, S. S. y Wilk, M. B.: «An analysis of variance test for normality». **Biometrika**, vol. 52, pp. 591-612, 1965.

Tejedor, F. J.: «Inferencia estadística y teoría de test». Edición propia. U.A.M., Madrid, 1976.