ESTADISTICA ESPAÑOLA
Núm. 95, 1982, págs. 73 a 83

Contrastes de hipótesis en el modelo de análisis de la varianza multivariante de un factor con efectos aleatorios

por ANTONIO J. BAIGORRI MATAMALA
Departamento de Estadística.
Facultad de Ciencias Económicas.
Universidad de Bilbao

RESUMEN

En este trabajo se presentan los contrastes de hipótesis acerca de los parámetros de un modelo del análisis de la varianza. Tal modelo resulta de extender a vectores de variables el de un factor de clasificación con efectos aleatorios. La deducción de las distribuciones en el muestreo de los estadísticos en los que se fundamentan los test propuestos es estudiada en detalle. En general, los resultados obtenidos guardan paralelismo con los del modelo univariante, si bien basados en una teoría distribucional distinta.

Palabras clave: Análisis de varianza multivariante, modelos de efectos aleatorios, criterios multivariantes.

1. INTRODUCCION

En las técnicas de análisis de la varianza, los modelos de efectos aleatorios fueron por primera vez estudiados rigurosamente por Eisenhart (1947) y Henderson (1953). En este tipo de modelos el tema que ha recibido mayor atención ha sido la estimación de parámetros. Un resumen de los métodos propuestos, así como una solucción computacionalmente factible del método de máxima verosimilitud son dados en Harville (1977).

La consideración de los modelos de efectos aleatorios en vectores de variables no plantea problemas adicionales en lo referente a la estimación de parámetros. Este no es el caso en relación a los contrastes de hipótesis, ya que las distribuciones a utilizar son diferentes en los modelos univariantes y multivariantes. Este trabajo se centra en este tema; en particular, en los contrastes de hipótesis del modelo de análisis de la varianza multivariante de un factor con efectos aleatorios. Tal modelo ha sido estudiado con anterioridad por Ahrens (1977), en cuestiones relacionadas con la estimación de sus parámetros. El tema relativo a contrastes de hipótesis y que aquí se expone es por tanto nuevo.

Dedicaremos el resto de este apartado a la exposición de una serie de resultados de la estadística multivariante que serán utilizados posteriormente. Las pruebas correspondientes pueden consultarse, por ejemplo, en Anderson (1958). Introduciremos el siguiente convenio notacional:

$$A \sim W_p[\Sigma, n]$$

indicará que la matriz aleatoria A de orde $p * p$ sigue la distribución de Wishart de parámetros Σ, n.

1.1. Sea $X = (x_1, x_2, \ldots, x_p)^T$ una v. a. con distribución $\vec{X} \sim N(\vec{m}, \Sigma)$ y sea X_1, X_2, ..., X^N una muestra aleatoria de $\vec{X}$: $N \geq p$.

La matriz

$$A = \sum_{i=1}^{N} (\vec{X}_i - \overline{\vec{X}})(\vec{X}_i - \overline{\vec{X}})^T \; : \; \overline{\vec{X}} = 1/N \sum_{i=1}^{N} \vec{X}_i$$

sigue la distribución

$$A \sim W_p[\Sigma, n] \; : \; n = N - 1$$

independiente de $\overline{\vec{X}}$.

1.2. Si A_1, A_2, ..., A_r son matrices de v. a. con distribuciones independientes:

$$A_i \sim W_p[\Sigma, n_i] \; : \; i = 1, 2, \ldots, r$$

la matriz suma se distribuye

$$\sum_{i=1}^{r} A_i \sim W_p[\Sigma, s] \; : \; s = \sum_{i=1}^{r} n_i$$

1.3. Si A $\sim W_p[\Sigma, n]$ y $\vec{X} = (x_1, x_2, ..., x_p)^T$ está distribuido independientemente de A.

$$\vec{X}^T A \vec{X} / \vec{X}^T \Sigma \vec{X} \sim \chi^2(n)$$

1.4. Sean $\vec{X}_i = (x_{i1}, x_{i2}, ..., x_{ip})^T : i = 1, 2, ..., N$ vectores de números reales y consideremos la matriz

$$\chi = [\vec{X}_1 | \vec{X}_2 | ... | \vec{X}^N]^T$$

Sean A, B_1, B_2, ..., B_r matrices de orden N $*$ N y de rango a b_1, b_2, ..., br, respectivamente.

Una condición necesaria y suficiente para que la descomposición

$$\chi^T A \chi = \chi^T B_1 \chi + \chi^T B_2 \chi + ... + \chi^T B_r \chi$$

pueda expresarse en términos de una matriz transformada Y = X D, de tal forma que los sumandos del segundo miembro dependan de elementos distintos de Y es que se verifique

$$a = b_1 + b_2 + ... + b_r$$

Este resultado algebraico, cuando es aplicado a matrices X de v. a., establece que si se satisface la condición anterior, las formas cuadráticas multivariantes $\chi^T B_1 \chi$, $\chi^T B_2 \chi$, ..., $\chi^T B_r \chi$ son independientes.

2. MODELO MULTIVARIANTE DE UN SOLO FACTOR DE CLASIFICACION ALEATORIO

Se considera un conjunto de N observaciones de una característica p-dimensional, que se supone han sido clasificadas de acuerdo a un factor R de tratamientos o condiciones experimentales de interés para el investigador y que recogemos en la siguiente tabla.

$$\begin{array}{lccccc}
\text{Trat } 1 & \vec{X}_{11} & \vec{X}_{12} & \cdots & \vec{X}_{1T} & \vec{X}_{10} \\
\text{Trat } 2 & \vec{X}_{21} & \vec{X}_{22} & \cdots & \vec{X}_{2T} & \vec{X}_{20} \\
\vdots \quad \vdots & \vdots & \vdots & & \vdots & \\
\text{Trat } R & \vec{X}_{R1} & \vec{X}_{R2} & \cdots & \vec{X}_{RT} & \vec{X}_{R0} \\
 & & & & & \vec{X}_{00}
\end{array} \qquad [1]$$

en donde $T = N/R$ y X_{00} y X_{r_0} son los vectores:

$$\vec{X}_{00} = 1/N \sum_{r=1}^{R} \sum_{t=1}^{T} \vec{X}_{rt} \qquad [2]$$

$$X_{r_0} = 1/T \sum_{t=1}^{T} \vec{X}_{rt}; \qquad r = 1, 2, \ldots, R \qquad [3]$$

Se admite que cada una de las N observaciones satisface el siguiente modelo:

$$\vec{X}_{rt} = \vec{\mu} + \vec{a}_r + \vec{\varepsilon}_{rt} \qquad \begin{array}{l} r = 1, 2, \ldots, R \\ t = 1, 2, \ldots, T \end{array} \qquad [4]$$

en donde se suponen las hipótesis:

$\vec{\mu}$ es un vector de constantes [5]

$\vec{a} \sim N(\vec{0}, \Sigma a)$; $r = 1, 2, \ldots, R$ [6]

$\vec{\varepsilon}_{rt} \sim N(\vec{0}, \Sigma)$; $r = 1, 2, \ldots, R$; $t = 1, 2, \ldots, T$ [7]

y las condiciones de incorrelación:

$E\vec{a}_r\vec{a}_{r'}^T = [0]_{p \times p}$ si $r \neq r'$; $r, r' = 1, 2, \ldots, R$ [8]

$E\vec{\varepsilon}_{rt}\vec{\varepsilon}_{r't'}^T = [0]_{p \times p}$ si $r \neq r'$ o $t \neq t'$; $r, r' = 1, 2, \ldots, R$; $t, t' = 1, 2, \ldots, T$ [9]

$E\vec{a}_r\vec{\varepsilon}_{rt}^T = [0]_{p \times p}$ $\forall r, t$; $r = 1, 2, \ldots, R$; $t = 1, 2, \ldots, T$ [10]

Nos referiremos al modelo anterior como «Análisis de la varianza multivariante de un factor de clasificación con efectos aleatorios».

Cuando en la condición [6] se consideran los vectores $\vec{a}_r$; $r = 1, 2, \ldots, R$, constantes, se tiene el modelo multivariante de un factor fijo. Tal modelo tiene por objeto el inferir en base a las N observaciones, si dichos parámetros son nulos, es decir, que no existe efecto debido a los tratamientos.

Si las hipótesis [4] a [10] son asumidas en su totalidad, las observaciones $\vec{X}_{rt}$, $r = 1, 2, \ldots, R$; $t = 1, 2, \ldots, T$ siguen distribuciones independientes $N \sim (\vec{\mu}, \Sigma + \Sigma a)$. Esta situación modeliza el caso en que los R tratamientos son una muestra de un colectivo más amplio cuyo estudio es de interés para el investigador. En el modelo de efectos aleatorios también se centrará la atención en inferir la existencia o no de efectos debidos a los tratamientos. La diferencia, en este caso, es que tal inferencia concierne no sólo los R tratamientos seleccionados, sino todos los elementos de la población de la cual proceden.

En términos de la tabla 1 diremos que mientras el modelo de un factor fijo supone inferencia «en horizontal», en el modelo aleatorio es en ambos sentidos: «vertical» y «horizontal».

3. CONTRASTES DE HIPOTESIS

3.1. INTRODUCCIÓN

En el modelo presentado son de especial interes los contrastes de hipótesis en relacción Σa, matriz de covarianza de los tratamientos. Dos test de hipótesis concernientes a $\vec{\mu}$ y $\Sigma + \Sigma_a$ caen dentro de la inferencia acerca de los parámetros de la distribución normal multivariante, por lo que no serán tratados aquí. En los contrastes acerca de Σ_a se distinguirán dos casos: Σ conocida y Σ desconocida. Estudiaremos algunas cuestiones introductorias que facilitarán la obtención de las distribuciones de los estadísticos en que se basan los contrastes propuestos.

Consideremos las matrices:

$$S = \sum_{r=1}^{R} \sum_{t=1}^{T} (\vec{X}_{rt} - \vec{X}_{00})(\vec{X}_{rt} - \vec{X}_{00})^T$$

$$D = \sum_{r=1}^{R} \sum_{t=1}^{T} (\vec{X}_{rt} - \vec{X}_{r_0})(\vec{X}_{rt} - \vec{X}_{r_0})^T$$

$$F = T \sum_{r=1}^{R} (\vec{X}_{r_0} - \vec{X}_{00})(\vec{X}_{r_0} - \vec{X}_{00})^T$$

que satisfacen la relación algebraica $S = D + F$.

Si se tienen en cuenta las siguientes expresiones deducidas de [4]:

$$\vec{X}_{r_0} = \vec{\mu} + \vec{a}_r + \varepsilon_{r_0}; \qquad r = 1, 2, \ldots, R$$
$$\vec{X}_{00} = \vec{\mu} + \vec{a}_0 + \varepsilon_{00}$$

en donde $\vec{a}_0$, ε_{r_0}; $r = 1, 2, \ldots, R$, y ε_{00} designan los vectores:

$$\vec{a}_0 = 1/R \sum_{r=1}^{R} a_r \qquad [11]$$

$$\vec{\varepsilon}_{r_0} = 1/T \sum_{t=1}^{T} \vec{\varepsilon}_{rt} \qquad [12]$$

$$\vec{\varepsilon}_{00} = 1/N \sum_{r=1}^{R} \sum_{t=1}^{T} \varepsilon_{rt} \qquad [13]$$

la matriz F puede escribirse:

$$F = T \sum_{r=1}^{R} (\vec{\varepsilon}_{r_0} - \vec{\varepsilon}_{00})(\varepsilon_{r_0} - \varepsilon_{00})^T + T \sum_{r=1}^{R} (\vec{a}_r - \vec{a}_0)(\vec{a}_r - \vec{a}_0)^T \qquad [14]$$

Similarmente D:

$$D = \sum_{r=1}^{R} D_r \qquad [14]$$

$$D_r = \sum_{t=1}^{T} (\vec{\varepsilon}_{rt} - \vec{\varepsilon}_{r_0})(\vec{\varepsilon}_{rt} - \varepsilon_{r_0})T \qquad [15]$$

3.2. CONTRASTES ACERCA DE Σa CUANDO Σ ES CONOCIDA

Comenzaremos estudiando el contraste de la hipótesis:

$\mathcal{H}_0$: $\Sigma a = [0]_{p\times p}$

$\mathcal{H}a$: $\Sigma a \neq [0]_{p\times p} \cap \Sigma a$ definida positiva y simétrica,

es decir, la no existencia de efecto debido a los tratamientos.

La matriz F de [14] puede escribirse

$$F = \sum_{r=1}^{R} [\sqrt{T}(\vec{\varepsilon}_{r_0} + \vec{a}_r - \varepsilon_{00} - \dot{a}_0)][\sqrt{T}(\vec{\varepsilon}_{r_0} + \vec{a}_r - \vec{\varepsilon}_{00} - \vec{a}_0)^T]$$

Puesto que $\sqrt{T}(\varepsilon_{r_0} + a_r)$ están distribuidas independientemente, $N(0, \Sigma + T\Sigma a)$ y $\sqrt{T}(\vec{\varepsilon}_{00} + \vec{a}_0) = 1/R \sum_{r=1}^{R} \sqrt{T}(\vec{\varepsilon}_{r_0} + \vec{a}_r)$ por 1.1 se tiene que

$$F \sim W_p[\Sigma + T\Sigma a, R - 1]$$

y F es independiente de $\sqrt{T}(\varepsilon_{00} + a_0)$.

El resultado anterior permite fundamentar el contraste de la hipótesis que aquí nos ocupa. Puesto que $\vec{\mu}$ y $\sqrt{T}$ son constantes F y $\vec{X}_{00}$ son independientes, y por 1.3

$$\phi = \vec{X}_{00}^T F \vec{X}_{00}/\vec{X}_{00}\Sigma\vec{X}_{00} \sim \chi^2(R - 1)$$

cuando H_0 es cierta.

Si $\chi^2\varphi$ designa la abscisa que en la distribución χ^2 deja hacia la izquierda un área de φ, el contraste se establecerá en los siguientes términos. Rechazar H_0 si evaluado ϕ a partir de una muestra $\phi < \chi^2\varphi$. Esta prueba, por construcción, es de nivel de significación φ.

En general, para hallar la curva característica de la prueba es necesario considerar hipótesis alternativas más restringidas que la propuesta como $\mathcal{H}a)$ $\Sigma a = k\Sigma$; $k > 0$. Bajo $\mathcal{H}a$ el estadístico

$$\phi = \vec{X}_{00} F\vec{X}_{00}/\vec{X}_{00}^{T}\Sigma\vec{X}_{00} \sim \Gamma[1/2(k+1), (R-1)/2]$$

por tanto, para cada valor de k puede evaluarse la probabilidad de que la variable ϕ pertenezca a la región $(\chi^2\varphi, \infty)$, es decir, la probabilidad de error de tipo II.

El contraste de la hipótesis anterior no es el único que puede ser de interés en relacción a Σa. Ocurre que otro tipo de hipótesis sufre de las limitaciones de la teoría distribucional de la que estamos haciendo uso. Sólo aquellas hipótesis que suponen proporcionalidad entre las matrices Σ y Σa pueden tratarse adecuadamente. Supongamos que la matriz Σa es de la forma $\Sigma a = k\Sigma$, siendo k un parámetro desconocido, y se desea contrastar la hipótesis:

$$\mathcal{H}_0: \Sigma a \leq \bar{k}\Sigma$$
$$\mathcal{H}a: \Sigma a > \bar{k}\Sigma$$

Una prueba de nivel de significación φ es, en este caso, rechazar H_0 si $\phi < \Gamma\varphi$, siendo $\Gamma\varphi$ la abscisa que en la distribución $\Gamma[1/2(\bar{k}+1), (R-1)/2]$ deja hacia la izquierda un área de α.

Para valores de $k > \bar{k}$ el estadístico ϕ sigue la distribución $\Gamma[1/2(k+1)\,(R-1)/2]$, por tanto, en base a esta distribución, pueden obtenerse las probabilidades de aceptación de las hipótesis alternativas.

3.3. Contrastes acerca de Σ_a cuando Σ es desconocida

Consideremos de nuevo el contraste de la hipótesis

$\mathcal{H}_0$: $\Sigma a = [0]_{p \times p}$

$\mathcal{H}a$: $\Sigma a \neq [0]_{p\times p}$ $\cap$ Σa definida positiva y simétrica, suponiendo en este caso que la matriz Σ es desconocida.

El contraste anterior puede fundamentarse en ciertos estadísticos, función de las matrices D y F, cuyas distribuciones en el muestreo, cuando $\mathcal{H}_0$ es cierta, se estudiarán en este apartado. Un paso previo lo constituye la obtención de las distribuciones en el muestreo de D y F, así como la prueba de su independencia.

3.3.1. *Distribuciones en el muestreo de D y F*

La distribución de F fue ya obtenida en 3.1, resultando ser:

$$F \sim W_p[\Sigma, R - 1]$$

cuando $\mathcal{H}_0$ es cierta. Nos ocuparemos, por tanto, de la distribución de D.

Consideremos la matriz D en términos de las matrices D_r de [16]. Las expresiones [7], [9] y [13] hacen satisfacer a D_r los requisitos de 1.1, por tanto

$$D_r \sim W_p[\Sigma, (T - 1)]$$

La matriz D, de acuerdo a [4], [7], [15] y [16] es la suma de R matrices independientes con distribución de Wishart de parámetro matricial común. Luego, por 1.2 concluimos

$$D \sim W_p[\Sigma, R(T - 1)]$$

3.3.2. *Independencia de las matrices F y D*

Para la prueba de la independencia de las matrices F y D se introducirá una notación más adecuada (véase Fz. Troconiz (1981), c. 12.1). Definiremos las siguientes matrices y vectores

$$\chi = [\vec{X}_{11} | \vec{X}_{12} | \dots | \vec{X}_{RT}]$$

I_N matriz unidad de orden N

I_R matriz unidad de orden R

$\vec{1}_N$ vector columna formado con N unos

$\vec{1}_T$ vector columna formado con T unos

Con esta notación, las matrices S, F y D pueden expresarse:

$$S = \chi^T[I_N - 1/N\, \vec{1}_N \vec{1}_N^T]\chi = \chi^T P_1 \chi$$

$$D = \chi_T[I_N - I_R \otimes {}^1/_T \vec{1}_T \vec{1}_T]\chi = \chi P_2 \chi$$

$$F = \chi^T[I_R \otimes {}^1/_T \vec{1}_T \vec{1}_T^T - {}^1/_N \vec{1}_N \vec{1}_N^T]\chi = \chi P_3 \chi$$

Las matrices P_1, P_2 y P_3, por ser idempotentes, tienen su rango igual a su traza. En base a tal propiedad se tiene:

$$\text{ran } P_1 = t_r[I_N - {}^1/_N \vec{1}_N \vec{1}_N^T]$$

$$\text{ran } P_2 = t_r[I_N - I_R \otimes {}^1/_T \vec{1}_T \vec{1}_T^T]$$

$$\text{ran } P_3 = t_r[I_R \otimes {}^1/_T \vec{1}_T \vec{1}_T^T - {}^1/_N \vec{1}_N \vec{1}_N^T]$$

Puesto que $S = D + F$ y ran P_1 = ran P_2 + ran P_3, los requisitos de 1.4 son satisfechos, con lo que concluimos que F y D son independientes.

Resumiendo los resultados hasta este punto, se tiene que D y F están distribuidas independientemente:

$$D \sim W_p[\Sigma, R(T - 1)]$$

$$F \sim W_p[\Sigma, R - 1] \quad \text{cuando } \mathcal{H}_0 \text{ es cierto.}$$

por tanto, el contraste de la hipótesis $\mathcal{H}_0$: $\Sigma a = [0]_{p \times p}$ no es otro que el contraste de la hipótesis de que dos matrices con distribuciones Wishart independientes tienen el mismo parámetro matricial. El contraste de esta hipótesis puede basarse en cualquiera de los criterios multivariantes que han sido propuestos con este y otros propósitos (una recopilación crítica es dada por Pillai, 1975). Tales criterios suponen diferentes funciones de las raíces de la ecuación determinante:

$$|F - r^2 D| = 0 \qquad [16]$$

Así, el criterio de Wilks se define

$$\lambda = \prod_{i=1}^{p} 1/(1 - r_i^2) \qquad [17]$$

el de Hotelling

$$T_0^2 = \sum_{i=1}^{p} r_i^2 \qquad [18]$$

y el de Roy

$$\mathcal{R} = r_1^2 : r_1^2 = \text{mayor raíz de} \qquad [16]\ [18]$$

La distribución conjunta de las raíces de [16] fue hallada simultáneamente por varios autores en 1939: en base a tal distribución se han hallado las distribuciones de las funciones [17], [18] y [19], y que en la actualidad están tabuladas. En resumen, se tienen las distribuciones en el muestreo cuando $\mathcal{H}_0$ es cierta de varios estadísticos, pudiéndose, en base a cualquiera de ellos, fundamentar la prueba que aquí nos ocupa.

Una vez seleccionado un criterio de cálculo de la curva de potencia de la prueba, requiere la obtención de su distribución en el muestreo bajo $\mathcal{H}a$. Para la obtención de tales distribuciones es necesario la distribución conjunta bajo $\mathcal{H}a$ de las raíces de [16].

Esta distribución obtenida por Roy en 1942 depende de adicionales parámetros, como son las raíces características de la ecuación determinante:

$$|(\Sigma + T\Sigma a) - \lambda^2\Sigma| = 0 \qquad [20]$$

Aunque puesto en estos términos, el problema se reduce a hallar la distribución de una cierta función de varias variables aleatorias de las cuales se conoce su distribución conjunta; por una parte, el tipo de función que suponen los criterios, y por otra, las características de la distribución conjunta implicada han dado lugar a que sólo para alguno de ellos haya podido derivarse su distribución bajo la hipótesis alternativa, dependiendo también de las raíces características de [20].

Esta característica de las pruebas de hipótesis del modelo multivariante supone la generalización de un resultado del modelo univariante. En dicho modelo (ver, por ejemplo, Scheffe, 1959, pág. 227), la curva de potencia de la prueba de no existencia de efecto, debido a los tratamientos, es función del cociente entre la varianza de los tratamientos y la varianza del error. Este resultado constituye un caso particular del aquí estudiado. En efecto, si se considera $p = 1$, la única raíz de la ecuación [20] es precisamente dicho cociente.

BIBLIOGRAFIA

AHRENS, H.: «On a Invariance Property for First and Second Moments of Estimate of Variance and Covariance Components». *Biometrical Journal*, págs. 479-496, 1977.

ANDERSON, T. W.: *And Introduction to Multivariate Statistical Analysis*. John Wiley & Sons. New York, 1958.

EISENHART, C.: «The Asumptions Underlying the Analisis of Variance». *Biometrics*, 3, págs. 1-21, 1947.

FZ. de TROCONIZ: «Introducción a las Teorías de las Probabilidades y de la Estadística Clásica y Bayesiana». *Grafor*. Bilbao, 1981.

HARVILLE, D. A.: «Maximum Likelihood aproaches to variance and Covariance Component Estimation and to related Problems». *Journal of The American Statistical Association*, 72, págs. 320-338, 1977.

HENDERSON, C.: «Estimation of Variance Covariance Components». *Biometrics*, 9, págs. 226-252, 1953.

PILLAI, S.: «Distributions of Characteristics Roots in Multivariate Analysis: Null Distributions». *The Canadian Journal of Statistics*, págs. 157-184, 1976.

SCHEFFE, H.: *The Analysis of Variance*. John Wiley & Sons. New York, 1959.

SUMMARY

This paper deals with some tests of hipotesis concerning the parameters of an analysis of variance model. Such a model results of extending to vectors of variables the One Way Random Effect Model. The sampling distributions of the statistics which the tests are based on are studied in detail. Broadly speaking, the results obtained are an extension of those that come out in the univariate model although based on a different distribution theory.

Key words: Multivariate Analisis of Variance. Random Effect Models. Multivariate Criteria.

AMS. 1970, Subject classification: 62H15.