Revista Colombiana de Estadística
N° 3 - 1981

MODELO II DE EISENHART

Leonardo Bautista

Estudiante de Posgrado
Universidad de Dormund
Alemania Federal.

Motivación.

Uno de los problemas al que con más frecuencia debe enfrentarse un estadístico que trabaje en la práctica es el recibir datos obtenidos con bastante esfuerzo y a altos costos con la misión de extraer de ellos una cierta o la mayor información estadística posible. El trabajo de encontrar "buenos" estimadores o de realizar pruebas de hipótesis "óptimas" es bastante elemental cuando estos datos provienen de un modelo de diseño de experimentos conocido y "limpiamente" realizado, pero lamentablemente no es ésta la generalidad. En la mayoría de los casos los datos provienen de modelos no conocidos, no completamente ortodoxos y el estadístico es el encargado de aplicar métodos, como el aquí expuesto, para obtener lo que de él se exige.

A nivel teórico representa este artículo un avance importante en la teoría de los modelos lineales y el diseño de experimentos, aunque lo que aquí se expone tiene una edad mayor de 20 años y científicamente tan solo un caso especial de métodos más sofisticados y modernos (1977-1978) constituye una buena manera de comenzar a familiarizarse con los modelos lineales.

Por último quiero aclarar que este no constituun trabajo original (a nivel científico) sino un resumen sencillo y sin demostraciones del trabajo realizado en 1961 por Graybill y Hultquist [9]. Este artículo fué el manuscrito en mi charla del 25 de Sep./80 en la U. Nal.

Modelo.

$$Y = X_0\beta_0+X_1\beta_1+\ldots+X_{k+1}\beta_{k+1}$$

donde

Y = $(n\times 1)$-vector de observaciones

β_0 = constante fija desconocida (normalmente μ)

$$X_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \text{Notación } X_0 = 1_n$$

X_i = $(n\times p_i)$-matrices de constantes conocidas.

$X_{k+1} = I_{n\times n}$ (Matriz de identidad).

B_{k+1} = $(n\times 1)$-vector aleatorio (normalmente llamado vector aleatorio de error).

$$B_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{ip_i} \end{bmatrix} \quad (i = 1,\dots,k) \text{ vector aleatorio}$$

Supuestos.

1. Independencia estocástica de los vectores aleatorios B_i y entre las variables aleatorias dentro de los vectores.

2. B_i $(i = 1,\dots,k+1)$ tiene una distribución multinormal con media 0 (vector cero) y varianza $\sigma_i^2 \, I_{p_i}$.

3. Sea $X_i \cdot X_i' = A_i$ (B' es la notación para B traspuesto) se exige entonces:

a) A_i $(i = 0,1,\dots,k+1)$ son linealmente independientes.

b) El producto $A_i \cdot A_j = A_j \cdot A_i$ para todo $i,j = 0,1,\dots,k+1$.

Nota: Si $X_0 = 1_n$ entonces $A_0 = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \cdot & \dots & \vdots \\ 1 & 1 & & 1 \end{bmatrix}_{n\times n}$

emplearemos para esta matriz la notación $\mathbf{1}_n$.

4. La matriz X_i es tal que:

a) $1_n' \cdot X_i = r_i \, 1_{p_i}' \; ; \; r_i \in Z^+$

b) $X_i \cdot 1_{p_i} = 1_n$

es decir

a) $\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix} \cdot \begin{bmatrix} X_i \end{bmatrix} = r_i \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}$

b) $\begin{bmatrix} X_i \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

5. Si

$$B_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{ip_i} \end{bmatrix} \quad \text{se exige} \quad \begin{matrix} E[v_{ij}^3] = E[v_{ik}^3] \\ E[v_{ij}^4] = E[v_{ik}^4] \end{matrix} < \infty \begin{cases} \text{para todo} \\ j, k = 1, \ldots, p_i \end{cases}$$

para todo $i = 1, \ldots, k+1$

6. $B_0, \sigma_1^2, \ldots, \sigma_{k+1}^2$ son funcionalmente independientes.

Nota: En la práctica los modelos comúnmente más utilizados satisfacen el modelo y las suposiciones planteadas, por ejemplo: los modelos de diseño de experimentos con igual número de subclases.

- Modelos cruzados de tamaño n con o sin interacción.
- Modelos anidados de tamaño n.
- Modelos de parcelas divididas,

además si olvidamos los supuestos 3b) y 4. el modelo de regresión es un modelo II de Eisenhart.

Problema: El problema a resolver es encontrar a) estadísticos que estimen la constante B_0 y σ_i^2, $i = 1,\ldots,k+1$; b) las distribuciones de estos estadísticos y c) las restricciones que se deben imponer al modelo para que los estimadores sean "óptimos" es decir insesgados y únicos de mínima varianza.

Desarrollo. Dado que existe independencia estocástica (sup.1) y distribución normal de los B_i ($i = 1,\ldots,k+1$) (sup.2) sabemos que la densidad de Y es

$$f_Y(y) \approx exp\{(Y-1_nB_0)'V(y)^{-1}(Y-1_nB_0)\}$$

nuestro propósito es expresar esta forma cuadrática en forma tal que podamos distinguir la pertenencia de esta familia a la clase exponencial.

Llamemos $V(y) = V$ entonces

$$V = V[\sum_{i=0}^{k+1} X_iB_i] = \sum_{i=1}^{k+1} V[X_iB_i] = \sum_{i=1}^{k+1} X_iX_i' \; V(B_i) = \sum_{i=1}^{k+1} \sigma_i^2 A_i$$

Llamemos ahora $W = E(YY')$ entonces

$$W = E[(\sum_{i=0}^{k+1} X_iB_i)(\sum_{i=0}^{k+1} X_iB_i)'] = E[(\sum_{i=0}^{k+1} X_iB_i)(\sum_{i=0}^{k+1} B_i'X_i')]$$

$$= \sum_{i=0}^{k+1} X_iE(B_iB_i')X_i' + \underbrace{\sum_{i=0}^{k+1} X_i\, E(B_iB_j')X_j'}_{0} = \sum_{i=0}^{k+1} \sigma_i^2\, A_i.$$

Teorema 1. (Alg.Lin.) *Sea s la cantidad de valores propios diferentes entre sí de la matriz W, entonces $s \geq k+2$.*

Teorema 2. (Alg. Lin.) *Si $A_0, A_1, \ldots, A_{k+1}$ son matrices simétricas que conmutan dos a dos entonces existe una matriz ortogonal P tal que $PA_iP' = D_i$ donde D_i es la matriz diagonal conformada por los valores propios de A_i.* [1] pag.35 Teor.54, [2]pag.189.

Entonces podemos diagonalizar la matriz W

$$PWP' = P(\sum_{i=0}^{k+1} \sigma_i^2 A_i)P' = \sum_{i=0}^{k+1} \sigma_i^2 D_i$$

Veamos ahora que según nota del supuesto 3. $A_0 = \mathbb{1}_n$; esta matriz tiene rango uno y por lo tanto un solo valor propio diferente de cero y este es n.

Si escogemos apropiadamente P podemos obtener que D_0 sea $\begin{bmatrix} n & & & \\ & 0 & & \\ & & 0 & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}$ y como ya establecimos que s sea la cantidad de valores propios diferentes entre si de la matriz W entonces

$$PWP' = \begin{bmatrix} n_0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} B_0^2 + \sum_{i=0}^{k+1} \sigma_i^2 D_i = \begin{array}{c|l} \begin{bmatrix} d_1 & & & & & & \\ & d_2 & & & & & \\ & & \ddots & & & & \\ & & & d_2 & & & \\ & & & & \ddots & & \\ & & & & & d_s & \\ & & & & & & \ddots \; d_s \end{bmatrix} & \begin{array}{l} 1 \text{ vez} \\ n_2 \text{ veces} \\ \\ n_s \text{ veces} \end{array} \end{array}$$

entonces

$$PVP' = \begin{bmatrix} d_1 & & & & & \\ & d_2 & & & & \\ & & \ddots d_2 & & & \\ & & & \ddots d_s & & \\ & & & & \ddots d_s \end{bmatrix} - \begin{bmatrix} n & & & \\ & 0 & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \end{bmatrix} B_0^2 = \begin{bmatrix} d_1^* & & & \\ & d_2 \cdot I_{n_2} & & \\ & & \ddots & \\ & & & d_s \cdot I_{n_s} \end{bmatrix}$$

donde $d_1^* = d_1 - nB_0^2$.

Ahora volviendo al problema inicial llamemos

$Q = (Y-1_nB_0)' \; V^{-1}(Y-1_nB_0)$ entonces

$Q = (Y-1_nB_0)' \; P'P \; V^{-1}P'P(V-1_nB_0)$

$= (PY-P1_nB_0)' \; (PVP')^{-1}(PY-P1_nB_0)$ y así

$$Q = (PY-P1_nB_0)' \begin{bmatrix} 1/d_1^* & & & \\ & (1/d_2)I_{n_2} & & \\ & & \ddots & \\ & & & \left(\frac{1}{d_s}\right) I_{n_s} \end{bmatrix} (PY-P1_nB_0)$$

Ahora observamos que $PA_0P' = \text{diag.}(n00\ldots0)$ es decir $PX_0X_0'P' = P1_n1_n'P' = \text{diag.}(n00\ldots0)$ y si particionamos P

$\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_s \end{bmatrix}$ donde P_u es una $(n_u \times n)$-matriz; obtenemos

$(P_1 1_n)^2 = n \Rightarrow P_1 1_n = n^{\frac{1}{2}}$ y

$(P_u 1_n)^2 = 0$ para todo $u = 2,\ldots,s$ y así podemos expresar Q:

$$Q=\begin{bmatrix} P_1 y - B_0 n^{\frac{1}{2}} \\ P_2 y \\ \vdots \\ P_s y \end{bmatrix}' \begin{bmatrix} 1/d_1^* & & & \\ & 1/d_2 \cdot I_n & & \\ & & \ddots & \\ & & & 1/d_s I_{n_s} \end{bmatrix} \begin{bmatrix} P_1 y - B_0 n^{\frac{1}{2}} \\ P_2 y \\ \vdots \\ P_s y \end{bmatrix}$$

es decir $Q = 1/d_1^*(P_1 y - B_0 n^{\frac{1}{2}})^2 + \sum_{u=2}^{s} \frac{1}{d_u} y' P_u' P_u P$ y así la densidad

$$f_Y(y) \approx \exp\{1/d_1^*(P_1 y - B_0 n^{\frac{1}{2}})^2 + \sum_{u=2}^{s} \frac{1}{d_u} y' P_u' P_u y\}$$

Así obtenemos s estadísticas suficientes según el criterio de Neymann [3] pág.150 teor.3.22.

$P_1 y,\ y' P_2' P_2 y, \ldots, y' P_s' P_s y$ y según teorema de Lehmann-Scheffé podemos comprobar que éstas conforman un conjunto minimal suficiente [4].

Teorema 3. *Si $s = k+2$ entonces el conjunto de estadísticas P_1, $y' P_u' P_u y$ $(u = 2,\ldots,s)$ generan una familia completa,* [5].

Veamos por último la distribución de cada una de ellas.

1. $P_1 y \sim n(P_{1+n} B_0, P_1 V P_1') = n(B_0 n^{\frac{1}{2}}, d_1^*)$

2. Según teorema de Graybill [6] se cumple (también

en este caso)

$$\frac{y'P_u'P_u y}{d_u} \sim \chi^2_{n_u} \quad \text{y debido a que} \quad P_u'P_u V P_v' P_v = 0 \ (u \neq v)$$

se puede demostrar la independencia estocástica de todas estas variables aleatorias; es decir tenemos estimadores insesgados para B_0 y $d_2, \ldots, d_s$.

Ahora si se cumple que $s = k+2$, es decir si hay completez, podemos aplicar el teorema de Lehmann-Scheffé [7] que dice: Si $U(T)$ es insesgado para Θ.

T conforma un conjunto de estadísticas suficientes y la familia de densidades $Z(T)$ es completa entonces $U(T)$ es el único estimador insesgado de mínima varianza para Θ.

Y este era precisamente nuestro propósito, veamos ahora un ejemplo práctico.

<u>Ejemplo</u>:

$$y_{ij} = \mu+\beta_i+\tau_j+e_{ij} \qquad (i = 1,2 \quad j = 1,2)$$
$(i = 3,4 \quad j = 3,4)$ es decir

		B			
		1	2	3	4
τ	1	x	x		
	2	x	x		
	3			x	x
	4			x	x

$$B_i \sim n(0,\ \sigma_b^2)$$
$$\tau_j \sim n(0,\ \sigma_t^2) \qquad \text{i.i.d.}$$
$$e_{ij} \sim n(0,\ \sigma_e^2) \qquad \text{para todo } i, j\ ;$$

en forma matricial tenemos

$$y=\begin{bmatrix} y_{11}\\ y_{12}\\ y_{21}\\ y_{22}\\ y_{33}\\ y_{34}\\ y_{43}\\ y_{44}\end{bmatrix} = 1_8\cdot\mu+\begin{bmatrix}1&&&\\1&&&\\&1&&\\&1&&\\&&1&\\&&1&\\&&&1\\&&&1\end{bmatrix}\begin{bmatrix}B_1\\B_2\\B_3\\B_4\end{bmatrix}+\begin{bmatrix}1&0&&\\0&1&&\\1&0&&\\0&1&&\\&&1&0\\&&0&1\\&&1&0\\&&0&1\end{bmatrix}\begin{bmatrix}\tau_1\\\tau_2\\\tau_3\\\tau_4\end{bmatrix}+I_8\begin{bmatrix}e_{11}\\e_{12}\\e_{21}\\e_{22}\\e_{33}\\e_{34}\\e_{43}\\e_{44}\end{bmatrix}$$

$k+1 = 3$; $A_0 = \mathbb{1}_8$ $\quad A_1 = \text{diag}(\mathbb{1}_2,\ \mathbb{1}_2,\ \mathbb{1}_2,\ \mathbb{1}_2)$

$$A_2 = \text{diag}\left(\begin{bmatrix}1&0&1&0\\0&1&0&1\\1&0&1&0\\0&1&0&1\end{bmatrix},\ \begin{bmatrix}1&0&1&0\\0&1&0&1\\1&0&1&0\\0&1&0&1\end{bmatrix}\right)$$

y por supuesto $A_3 = I_8$.

$$P = \frac{1}{8^{\frac{1}{2}}}\begin{bmatrix}1&1&1&1&1&1&1&1\\ -1&-1&-1&-1&1&1&1&1\\ -1&-1&1&-1&-1&1&-1&1\\ 1&-1&1&-1&-1&1&-1&1\\ -1&-1&1&1&-1&-1&1&1\\ 1&1&-1&-1&-1&-1&1&1\\ 1&-1&-1&1&1&-1&-1&1\\ -1&1&1&-1&1&-1&-1&1\end{bmatrix}\begin{matrix}P_1\\P_2\\P_3\\ \\P_4\\ \\P_5\\ \end{matrix} \qquad D_i = PA_iP'$$

$$\mathcal{D}_0 = \text{diag}(8,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0)$$

$$\mathcal{D}_1 = \text{diag}(2,\ 2,\ 0,\ 0,\ 2,\ 2,\ 0,\ 0)$$

$$\mathcal{D}_2 = \text{diag}(2,\ 2,\ 2,\ 2,\ 0,\ 0,\ 0,\ 0)$$

$$\mathcal{D}_3 = I_8$$

$$PWP' = \sum_{i=1}^{3} \sigma_i^2 \mathcal{D}_i = \text{diag}(8^2+2\sigma_b^2+2\sigma_t^2+\sigma_e^2,\ 2\sigma_b^2+2\sigma_t^2+\sigma_e^2,$$

$$2\sigma_t^2+\sigma_e^2,\ 2\sigma_t^2+\sigma_e^2,\ 2\sigma_b^2+\sigma_e^2,\ 2\sigma_b^2+\sigma_e^2,\ \sigma_e^2,\ \sigma_e^2)$$

$$= \text{diag}(d_1, d_2, d_3, d_3, d_4, d_4, d_5, d_5)$$

y por consecuencia P será particionada como arriba se señala.

$s = 5 > 4 = k+2$ es decir no hay completez, además $n_1 = 1$ $n_2 = 1$ $n_3 = 2$ $n_4 = 2$ y $n_5 = 2$. Las estadísticas suficientes que se obtienen son:

$$P_1 y = \frac{\sum_{i,j=1,2}^{0,4} y_{ij}}{8^{\frac{1}{2}}} = T_1 \sim n(8^{\frac{1}{2}}\mu, d_1^*)\ ;\ E\left[\frac{\sum_{i,j} y_{ij}}{8}\right] = \mu$$

$$y' P_2' P_2 y = 1/8\left(\sum_{i,j=3}^{4} y_{ij} - \sum_{i,j=1}^{2} y_{ij}\right)^2 = T_2\ ;\ \frac{T_2}{d_2} \sim \chi_1^2\ ;$$

$$E\left[\frac{T_2}{1}\right] = d_2$$

$$y' P_3' P_3 y = 1/4((y_{11}+y_{21}-y_{12}-y_{22})^2+(y_{33}+y_{43}-y_{34}-y_{44})^2)$$

$$= T_3$$

$$y'P_4'P_4y = 1/4((y_{11}+y_{12}-y_{21}-y_{22})^2+(y_{33}+y_{34}-y_{43}-y_{44})^2)$$
$$= T_4$$

$$y'P_5'P_5y = 1/4((y_{11}+y_{22}-y_{12}-y_{21})^2+(y_{33}+y_{44}-y_{34}-y_{43})^2)$$
$$= T_5$$

Y aunque no podemos decir mucho acerca de la "calidad" de estos estimadores si podemos realizar pruebas de hipótesis en base a la distribución F dado que estas estadísticas tienen una distribución χ^2.

Un último ejemplo es el conocido diseño de experimentos de bloques aleatorios (b-bloques y t-tratamientos); el modelo es:

$y = \mu 1 + X_1 B + X_2\tau + e$ donde $X_1 = \mathrm{diag}(1_t, 1_t, .., 1_t)$

$X_1 =$ (bt×b)-matriz

$X_2 = \mathrm{diag}(I_t, I_t, ..., I_t)$, $X_2 =$ (bt×b)-matriz

$A_1 = \mathrm{diag}(\mathbb{1}_t, \mathbb{1}_t, ..., \mathbb{1}_t)_{bt\times bt}$

$A_2 = \mathrm{diag}(X_2, X_2, ..., X_2)_{bt\times bt}$; los valores propios de W son:

σ_3^2, $t\sigma_1^2+\sigma_3^2$, $b\sigma_2^2+\sigma_3^2$ y $tb\mu^2+t\sigma_1^2+b\sigma_2^2+\sigma_3^2$;

y en este caso $s = 4 = k+2$ y los estimadores obtenidos son los estimadores cuadráticos únicos de mínima varianza.

Conclusiones y recomendaciones.

Enfrentados a un problema de este tipo y luego de cerciorarnos completamente de la validez de los supuestos es relativamente fácil aplicar el método para encontrar estimadores de "alta calidad" o realizar pruebas de hipótesis uniformemente más poderosas.

Cuando la completez no existe habrá que probar los estimadores a la cota de Rao-Cramer o tratar de obtener estimadores por otros métodos -Mingue ó Mivque- [8], y así poder hacer comparaciones (Varianzas).

Para aquellos que se interesan desde el punto de vista teórico tendrán que recurrir al artículo original (donde se encuentran aproximadamente 10 errores tipográficos que hacen más difícil su comprensión) o al libro de Kendall y Stuart, The advanced theory of statistics, Griffin-London 1973, Vol. 2 y 3 cap. 23-36 donde el artículo ha sido trabajado en forma amplia. Otro caso diferente desarrollado en forma similar se encuentra en el artículo:

Hultquist,R., Atzinger,E. The mixed effects model and simultaneous diagonalization of symmetric matrices, the annals of math. Stat. 1973 Vol.43 pag. 2024-2030.

Para finalizar quiero de nuevo señalar que toda esta teoría se convierte en casos particulares del trabajo realizado por Seely en los últimos años, por eso son recomendables los siguientes artículos:

Seely,J. Minimal sufficient statistics and completeness for multivariate normal families, Sankhyá 1977 Vol. 39 pág. 170-185.

Seely,J.A. Compl. suff. statistics for the linear model under normality and a singular covariance matriz, Communic. Statistics 1978 Vol.15 pág. 1465-1473.

* * *

BIBLIOGRAFIA

[1] Ballman,R., *Introduction to matriz analysis*, Mc Graw-Hill, N.Y. (1960).

[2] Thrall,R., Tornheim,L., *Vector spaces and matrices*, John Wiley and sons, N.Y. (1957).

[3] Witting,H., *Mathematische Statistik*, B.g. Teubner Stuttgart (1978).

[4] Lehmann,L., Scheffé,H., *Completeness, similar region and unbiased estimation*, part. I, Sankhyá, Vol. 10 pág. 305-340, (1950).

[5] Gautschi,W., *Some remarks on Herbach's paper*, Ann. Math. Stat. Vol. 30 pág. 960-963, (1959).

[6] Graybill,F., Marsaglia,G., *Idempotent matrices and quadratic forms in the general linear hypothesis*, Ann. Math. Stat. Vol. 28 pag. 678-686, (1957).

[7] Mood,A.,Graybill,F., Boes,D., *Introduction of the theory of statistics*, Mc-Graw Hill, N.Y. (1950).

[8] Rao,C.R., *Minimum variance quadratic unbiased estimation of variance components*, Journal of multivariate analysis 1 N°4, pag. 445-456, (1971).

———, *Estimation of variance and covariance components MINQUE theory*, Journal of multivariate analysis 1 N°4, pag. 257-275, (1971).

Kleffe,J. *Optimal estimation of variance components, a survey*, Sankhyâ Vol. 39, (1977).

[9] Graybill,F., Hultquist,R., *Theorems concerning Eisenhart's modell II*, Ann. Math. Stat. Vol. 32 pág. 261-269, (1961).

*

SOCIEDAD COLOMBIANA DE MATEMATICAS

Simposios Especializados

TALLER
DE
MATEMATICA APLICADA

La Sociedad Colombiana de Matemáticas se propone realizar un seminario-taller de Matemática Aplicada sobre los siguientes temas específicos:

Optimización numérica

Métodos numéricos para sistemas lineales

Modelos matemáticos para recursos naturales

Sistemas de Ecuaciones Diferenciales de Sturm-Liouville

Programación de microcumputadores

Probabilidad y Estadística Matemática

Participantes invitados del exterior:

William R. Derrick (U.of Montana). Modelos de administración de recursos, optimización.
Julio C. Díaz (U.of Kentucky). Optimización numérica.
Bruce H. Edwards (U.of Florida). Métodos numéricos para ecuaciones diferenciales.
Alan Lazer (U. of Cincinnati). Sistemas de ecuaciones diferenciales. Análisis Funcional.
Shair Ahmad (U.of Florida). Sistemas de Sturm-Liouville.
John Horváth (U. of Maryland). Análisis Funcional.

Fechas: 10 al 22 de Agosto-1981

Informes: Sociedad Colombiana de Matemáticas
Apartado Aéreo 2 5 2 1
Bogotá. 1.