Revista Colombiana de Estadística
Nº 14, Diciembre de 1986

FRONTERA INFERIOR DEL RIESGO CUADRADO EN EL DISEÑO SECUENCIAL DE EXPERIMENTOS

Ramón A. Matos M.

Resumen. Este artículo tiene como objetivo establecer la frontera inferior del riesgo cuadrado de la estimación de parámetros de regresión en el diseño secuencial de experimentos. Mediante un ejemplo se ilustra el cálculo de esta frontera. Se describe formalmente el esquema secuencial de experimentos.

Summary. This paper has the ojective to establish the lower bound for the quadratic risk of the parameters' estimates of the regression model of the sequentially designed. An example ilustrates the calculation of this bound. The scheme of sequential designs formally is de-

scribed.

La teoría matemática del diseño secuencial de experimentos iniciada por Chernoff (1959) ha sido desarrollada para el caso de la estimación de parámetros en los modelos de regresión. Maljutov (1983) encontró dos tipos de fronteras minimáximas asintóticas locales para el riesgo cuadrado de los parámetros estimados en los modelos de regresión F.

La frontera asintótica propuesta en este artículo mejora las anteriormente obtenidas para casos similares. Se analiza el problema de estimación de parámetros desconocidos en el esquema del diseño secuencial de experimentos para los modelos F.

1. Conceptos Preliminares.

1.1 Esquema Secuencial de Experimentos.

La idea en la que se fundamenta el procedimiento del diseño secuencial podemos resumirla así: En cada paso de una experimentación cualquiera podemos decidir sobre la continuación o no de las observaciones. Si las observaciones se concluyen, es escogido el parámetro de la distri bución de las observaciones en base a los experimentos realizados. Si por el contrario se decide continuar el experimento, se escoge el control para las siguientes observaciones (el punto de realización del nuevo experimento) dependiendo de los experimentos anteriores. Procedemos de esta manera hasta llegar al final de las observaciones.

La formalización de este esquema es la siguiente: Sean $(X,\beta x)$, $(Y,\beta y)$, (Θ,β_θ), (Λ,β_Λ)- espacios medibles, donde β_x, β_y, β_θ, β_Λ- son las σ-álgebras correspondentes a los espacios métricos X, Y, Θ, Λ. P_θ^x es una familia regular de medidas en Y. Denotamos por $(Z,\beta_Z) = (X\times Y, \beta_X \otimes \beta_Y)$ el producto de los espacios (X,β_X), (Y,β_Y); $Z_n = Z\times..\times Z$; $\beta Z_n = \beta_Z \otimes ... \otimes \beta_Z$; $Z_\infty = \underbrace{Z\times...}_{\infty}$, $\beta Z_\infty = \beta_Z \otimes ...$, $\theta \in \Theta$ es el parámetro desconocido, $x \in X$ es el control que el experimentar tiene la posibilidad de escoger.

La terna de reglas (U,N,Q) control, de tensión y toma de decisión recibe el nombre de estrategia secuencial (S). Las reglas (U,N,Q) son definidas por las sucesiones de funciones medibles x_n, τ_n, d_n respectimamente, donde

$$x_n : C_n \to X; \quad \tau_n : Y_n \to \{0,1\}; \; d_n : Y_n | C_n \to \Lambda;$$

$$C_n = \{(y_1, \ldots y_n): \sum_{i=1}^{N} \tau_i(y_1, \ldots, y_n) = 0\};$$

$$N = \min\{n \mid \tau_n(y_1, \ldots, y_n) = 1\}$$

(el subíndice n está referido al orden de las observaciones). La pareja (U,N) recibe el nombre de diseño secuencial (ξ).

En virtud del teorema Tulcea (Neveu J. 1964 cap. VI), a cada estrategia S con distribución inicial g para x_1 le corresponde una única medida P_θ^x en el conjunto $(Z_\infty, \beta_{Z_\infty})$ tal que

$$P_\theta^S(y_1 \in B) = \int g(dx) P_\theta^x(y \in B) \tag{1.1}$$

$$P_\theta^S(y_{n+1} \; B | y_1, \ldots, y_n) = P_\theta^{x_{n+1}}(B) \quad \text{si} \quad (y_1, \ldots, y_n) \in C_n$$

$$x_{n+1} = x_n \, , \; y_{n+1} = y_n \quad \text{si} \quad (y_1, \ldots, y_n) \notin C_n$$

Por E_θ^S denotamos la integral con medida P_θ^S.

1.2 Resultado Auxiliar.

Formularemos un lema que puede interpretarse como una generalización de la identidad de Wald. Su demostración se logra con ayuda de al-

gunos resultados de la teoría de Martingalas.

A cada diseño secuencial ξ le hacemos corresponder una medida aleatoria Π^{ξ} en (X,β_X), definida para todo $A \in B_X$ por la fórmula.

$$\Pi^{\xi}(A) = (E_{\theta}^{S}N)^{-1} \sum_{i=1}^{N} \mathbb{1}(X_i \in A)$$

($\mathbb{1}(\ )$ es el indicador).

Esta construcción recibe el nombre de proyección predecible (predictable) del diseño ξ. Su importancia queda fundamentada en el siguiente hecho.

Lema 1.1 *Si $g(x,y)$ es una función integrable en Z y el diseño ξ es tal que $E_{\theta}^{S}N < \infty$, entonces*

$$\sum_{i=1}^{N} E_{\theta}^{S}(g(x_i,y_i)|\beta_{Z_{i-1}}) = E_{\theta}^{S}N\int_X \phi(x)\Pi^{\xi}(dx),$$

donde

$$\phi(x) = \int_Y g(x,y)P^{x}(dy)$$

El lema 1.1 afirma que la medida de una función dependiente de las observaciones secuencialmente diseñadas (con ayuda del diseño ξ) se calcula integrando la función con la medida aleatoria Π^{ξ}.

2. Condiciones.

En este parágrafo enunciaremos las condiciones para el cumplimiento de la afirmación central del presente artículo.

2.1 Diseños.

Dada la sucesión de diseños $\xi^t = (U^t, N^t)$, donde a cada diseño ξ^t le corresponde un espacio medible $(Z_\infty^t, \beta_{Z_\infty}^t)$ y un flujo de σ-álgebras $\beta_{Z_n}^t$. A cada t y $\theta \in \Theta$ le corresponde la medida $P_\theta^{S^t}$ en $(Z_\infty^t, \beta_{Z_\infty}^t)$ definida por la fórmula (1.1). Esta medida la denotanos por P_θ^t. Consideraremos que los diseños ξ cumplen la condición $E_\theta^{S^t} N < \infty$, donde E^{S^t} es la integral con medida P_θ^t. En lugar de S^t escribiremos sólo el subíndice t referido al orden de la sucesión de los diseños.

La sucesión de diseños ξ^t converge débilmente, o sea, existe un espacio probabilístico (Ω, F, P) y una medida en X Π regular, condicional respecto a (Ω, F, P) tal que para toda función continua $f(x)$ en X: $\int f(x)\Pi^t(dx) \xrightarrow{d(P^t)} \int f(x)\Pi(dx)$ (convergencia en distribución.

2.2 El Modelo.

El diseño secuencial de experimentos en el modelo de regresión F se describe con el siguiente esquema:

$$E_\theta^t(y_i|\beta_{Z_{i-1}}^t) = \eta(x_i,\theta)$$
$$D_\theta^t(y_i|\beta_{Z_{i-1}}^t) = V(x_i,\theta)$$

La igualdad tiene lugar en el conjunto $\{N^t \geqslant i\}$ con probabilidad P_θ^t de 1. Θ- es un conjunto abierto en $\mathbb{R}^P$. La familia de medidas P_θ^x y las funciones de control, de tensión, decisión se consideran Borelianas.

Las funciones η, V cumplen las condiciones.

M1. η, $\phi_i = \frac{\partial\eta}{\partial\theta_i}$, $i = 1,..,P$, V-

toman valores en $\mathbb{R}^1$ y son continuas y acotadas.

M2. $\inf\limits_{X\times\Theta} V(x,\theta) > 0$

M3. $\frac{\partial^2\eta}{\partial\theta_i\partial\theta_j}$, $\frac{\partial V}{\partial\theta_i}$, $i,j = 1,..,P$-

toman valores en $\mathbb{R}^1$ y son continuas y acotadas.

M4. $M(\theta) = \int\phi(x,\theta)V^{-1}(x,\theta)\phi^T(x,\theta)\Pi(dx) > 0$

2.3 Los Estimadores.

Sea L la clase de estimadores para el valor real (θ^*) en el modelo F del tipo:

$$\tau^t(y,\theta) = \theta + A_\theta^t(y) \sum_{i=1}^{N} B_\theta^t(x_i)(y_i - \eta(x_i,\theta))$$

donde

$$y = (y_1,\ldots,y_{N^t})^T$$

Las funciones A_θ^t, B_θ^t cumplen las siguientes condiciones

T1. $A_\theta^t(y)\overset{*}{N}{}^t$ converge en probabilidad P_θ^t a la matriz $a(\theta)$ de dimensiones $(p\times p)$, además $\sup_{\mathbb{Z}\in W(\tau)} \|A_\theta^t(y)\overset{*}{N}{}^t\| < \infty$ donde $\overset{*}{N}{}^t = E_{\overset{*}{\theta}}^t N^t$, $\mathbb{Z}$- es el conjunto de los enteros, $W(\tau) = \{\theta \mid \|\theta-\overset{*}{\theta}\| < \tau\}$

T2. B_θ^t es una sucesión que toma valores en $\mathbb{R}^p$ y converge uniformemente a $B_\theta(x)$ en el conjunto $X\times W(\tau)$

3.

Para juzgar sobre la optimización de los estimadores de la clase L establecemos la frontera inferior del riesgo

$$R_\ell(\tau^t) = \lim_{c\to\infty} \lim_{t\to\infty} \sup_{\theta\in W(c\mid \overset{*}{N}{}^t)} \overset{*}{N}{}^t E_{\overset{*}{\theta}}^t [\ell^t(\tau^t(y,\theta)-\overset{*}{\theta})]^2$$

donde $\ell \in \mathbb{R}^p$, $\ell \neq 0$.

Teorema 3.1. *Si se cumplen las condiciones del parágrafo 2, para todo* $\ell \in \mathbb{R}^p$

$$R_\ell(\tau^t) \geqslant \ell^T EM^{-1}(\overset{*}{\theta})\ell$$

Demostración. Usando la igualdad (2) en "Matos R. 1985" obtenemos que

$$R_\ell(\tau^t) = \lim_{c\to\infty} \lim_{\to\infty} \sup_{W(c|\overset{*}{N}^t)} [\sqrt{\overset{*}{N}^t}\, \ell^T (\overset{*}{E}^t \tau^t - \overset{*}{\theta})]^2 +$$

$$+\lim_{t\to\infty} \overset{*}{N}^t \overset{*}{D}^t \ell^T \tau(y,\overset{*}{\theta}) \tag{3.1}$$

Si existe una sucesión t_i tal que

$$\lim_{c\to\infty} \lim_{t\to\infty} \sup_{W(c|\overset{*}{N}^t)} \sqrt{\overset{*}{N}^{t_i}}\ell^T (\overset{*}{E}^{t_i}\tau^{t_i} - \overset{*}{\theta}) = \infty$$

entonces automáticamente se cumple que

$$R_\ell(\tau^t) \geqslant E(\ell^t \overset{*}{M}^{-1} \ell)$$

Por lo tanto, examinaremos el caso cuando para todo t

$$\lim_{c\to\infty} \lim_{t\to\infty} \sup_{W(c|\overset{*}{N}^t)} \sqrt{\overset{*}{N}^t}\ell^T (E^t \tau^t - \overset{*}{\theta}) > \infty \tag{3.2}$$

De acuerdo con la definición de τ, sustituyendo $\eta(x,\theta)$ por $\overset{*}{\eta}(x)-(\overset{*}{\eta}(x)-\eta(x,\theta))$ e igualmente, usando la igualdad

$$\overset{*}{\eta}(x)-\eta(x,\theta) = \overset{*}{\phi}^t(x)(\overset{*}{\theta}-\theta) \tag{3.3}$$

$$+ \int_0^1 (\phi^T(x,\overset{*}{\theta}+\lambda(\theta-\overset{*}{\theta}))-\overset{*}{\phi}^T(x))d\lambda(\overset{*}{\theta}-\theta)$$

tenemos que

$$\sup_{W(c/\overset{*}{N}^{t})} \sqrt{\overset{*}{N}^{t}}\,\ell^{T}(\overset{*}{E}^{t}\tau^{t}-\overset{*}{\theta})$$

$$= \sup_{W(c/\overset{*}{N}^{t})} \Big| \sqrt{\overset{*}{N}^{t}}\,\ell^{T}\overset{*}{E}^{t}A_{\theta}^{t}(y) \sum_{i=1}^{N} B_{\theta}^{t}(x_{i})(y_{i}-\overset{*}{\eta}(x_{i}))$$

$$+ \sqrt{\overset{*}{N}^{t}}\,\ell^{T}\overset{*}{E}^{t}A_{\theta}^{t}(y) \sum_{i=1}^{N} B_{\theta}^{t}(x_{i})\cdot$$

$$\cdot \int_{0}^{1} (\phi^{T}(x_{i},\overset{*}{\theta}+\lambda(\theta-\overset{*}{\theta}))-\overset{*}{\phi}^{T}(x_{i}))\,d\lambda(\overset{*}{\theta}-\theta)$$

$$+ \sqrt{\overset{*}{N}^{t}}\,\ell^{T}(\overset{*}{E}^{t}A_{\theta}^{t}(y) \sum_{i=1}^{N} B_{\theta}^{t}(x_{i})\overset{*}{\phi}^{T}(x_{i})(\overset{*}{\theta}-\theta)-(\overset{*}{\theta}-\theta))\Big|$$

En virtud del cumplimiento de las condiciones del parágrafo 2 y aplicando el lema 1.1 es fácil comprobar que el límite de los dos primeros sumandos, cuando $t \to \infty$, es igual a cero.

Por otro lado, es fácil comprobar que

$$\lim_{t\to\infty} \sup_{W(c/\overset{*}{N}^{t})} \Big|\sqrt{\overset{*}{N}^{t}}\,\ell^{T}(\overset{*}{E}^{t}A_{\theta}^{t}(y) \sum_{i=1}^{N} B_{\theta}^{t}(x_{i})\phi^{*T}(x_{i})(\overset{*}{\theta}-\theta)-(\overset{*}{\theta}-\theta))\Big|$$

$$= \sqrt{c}\,\|\ell^{T}(Ea(\overset{*}{\theta})\int_{X} \overset{*}{B}(x)\overset{*}{\phi}^{T}(x)\Pi(dx)-J_{p})\| \qquad (3.4)$$

donde J_{p} es la matriz unitaria $(p\times p)$.

De la igualdad (3.4) se concluye que para el cumplimiento de (3.2) es condición necesaria que

$$\|\ell^T(Ea(\overset{*}{\theta})\int_X \overset{*}{B}(x)\overset{*}{\phi}^T(x)\Pi(dx)-J_p)\| = 0 \qquad (3.5)$$

Examinando el límite del segundo sumando en la expresión 3.1, obtenemos

$$\lim_{t\to\infty} \overset{*}{N}^t \overset{*}{D}^t \ell^T \tau^t(y,\overset{*}{\theta})$$

$$= E\ell^T a(\overset{*}{\theta})\int_X \overset{*}{B}(x)\overset{*}{B}^T(x)\overset{*}{V}(x)\Pi(dx)a^T(\overset{*}{\theta})\ell \qquad (3.6)$$

Encontrar el mínimo de la expresión (3.6) con la condición (3.5) representa un problema sencillo de la teoría de optimización. De los resultados de esta teoría concluimos que

$$\lim_{t\to\infty} \overset{*}{N}^t \overset{*}{D}^t \ell^T \tau(y,\overset{*}{\theta}) \geq E\ell^T \overset{*}{M}^{-1}\ell$$

(puede utilizarse la metodología empleada en el caso similar por Matos R. 1985).

De lo anterior se deduce que $R_\ell(\tau^t) \geq E\ell^T \overset{*}{M}^{-1}\ell$.

El teorama queda demostrado.

De esta forma la exactitud asintótica de los diseños óptimos se caracteriza por la matriz de información $M(\overset{*}{\theta}) = M(\Pi,\overset{*}{\theta})$ y el problema de construir un diseño secuencial óptimo se reduce a encontrar un diseño ξ^* tal que la proyección predecible Π^* satisfaga la condición

$$E\Pi^* = arg\ \min_{\pi} \psi(EM(\Pi,\overset{*}{\theta})),$$

donde ψ_- es un funcional convexo definido en el conjunto de las matrices de información (criterio).

4. Ejemplo.

Sea $X = [-a\ b]$, $\theta \in \mathbb{R}^1$, $Y = \{y^{(1)},y^{(2)}\} \subset \mathbb{R}^1 \mathbb{Z}$; a,b son dos números enteros positivos, $y_n^{(2)}$ es (independientemente de $\theta x^1,\ldots,x^N$ y de la trayectoria $y^{(1)}$) la posición enésima de la caminata aleatoria unidimensional que desde la posición $j \neq 0$ es descrita por una partícula que se desplaza con probabilidad 2/3 un paso en dirección contraria al origen de las coordenadas y con probabilidad 1/3 se desplaza un paso hacia el origen de las coordenadas. Desde cero la partícula se desplaza con probabilidad 2/3 al punto +1 y con probilidad 1/3 a -1.

La distribución de la caminata $y_n^{(2)}$ es la siguiente

$$p_+ = \frac{1}{2} + \frac{1}{6}\, sign\ y_{n-2}^{(2)}; \quad p_- = 1 - p_+$$

donde P_+ es la probabilidad de desplazamiento de la partícula a la derecha del punto donde se encuentra, P_- la probabilidad de desplazamiento a la izquierda. Consideraremos que $y_0^{(2)} = 0$, o sea la segunda componente de nuestras observaciones está definida por la trayectoria descrita por la caminata $y_n^{(2)}$ que parte de cero.

$y_n^{(1)}$ es una variable aleatoria normal con parámetros $(X_n\theta, 1)$.

De las propiedades generales de los diseños continuos óptimos se deduce que el diseño ξ^* definido por la fórmula

$$x_n = \begin{cases} b, \text{ si } y_{n-1}^{(2)} \geqslant 0 & \text{con probabilidad } \Pi(b) \\ -a, \text{ si } y_{n-1}^{(2)} < 0 & \text{con probabilidad } \Pi(-a) \end{cases}$$

es óptimo para la estimación de $\overset{*}{\theta}$ con la condición

$$sign\ x_n = sign\ y_{n-1}^{(2)}$$

Es sabido, que la trayectoria $y_n^{(2)}$ toma valores $\pm\infty$ con probabilidad $\frac{1}{2}$ "Feller W. 1970".

En virtud del teorema 3.1 obtenemos que

$$\lim_{c\to\infty} \lim_{t\to\infty} \sup_{W(c|\overset{*}{N}^t)} (R^t(\overset{*}{\theta},\tau^t) | \{y_n^{(2)} \to \infty\}) \geq b^{-2} E\Pi^{-1}(C)$$

La respuesta obtenida depende de la proyección Π quien a su vez depende de la regla (aleatoria) de detención N.

*

BIBLIOGRAFIA

Chernoff., (1959). Sequential design of experiments. Ann. Math. Stat., V30, No.3, p.755-770.

Feller, W., (1970). An introduction to probability theory and its applications. John Wiley & sons. New York.

Maljutov, M.B., (1983). Izv. Vuzov. Matematika, No. 11, Moscú. p.19-41.

Matos, R., (1985). Sobre la optimización de estimadores de mínimo cuadrado. Revista Colombiana de Estadística Nº 11, Bogotá.

Neveu, J., (1964). Bases mathématiques du calcul des probabilités. Massonn et cie. Paris.

*

Nota. El resultado de este artículo forma parte del trabajo presentado por el Dr. Malutov M.B. y Matos R. al 17th. European meeting of statisticans en la Universidad de Aristóteles (Tesaló nica-Grecia 1987) organizada por la Sociedad Bernoulli.

*