HOMENAJE AL PROFESOR SIXTO RIOS
TRABAJOS DE ESTADISTICA Y DE INVESTIGACION OPERATIVA
Vol. 36, Núm. 3, 1985, pp. 153 a 164

DISEÑO SECUENCIAL PARA DISCRIMINAR ENTRE MODELOS,BASADO EN LA INFORMACION CUADRATICA

Pilar García-Carrasco A.

Departamento de Estadística e Investigación Operativa.
Facultad de Ciencias Matemáticas. Universidad Complutense.
Ciudad Universitaria, 28040-Madrid.

La información cuadrática es una buena alternativa a la información de Shannon para todos aquellos problemas que, por su naturaleza, interesa tratarlos con una utilidad no local. El objetivo de este trabajo es dar, para estas situaciones, un método secuencial de construción de diseños para discriminación entre modelos, basado en la maximización de la información cuadrática.

Después de una introducción, donde se resumen los conceptos y resultados principales sobre información cuadrática, se plantea el problema, reduciéndolo a un caso particular de comparación de experimentos, con un espacio paramétrico de tipo mixto, y donde la cantidad de interés es alternativamente un subconjunto de sus componentes o todo él. Se dan los diseños óptimos en ambos casos, su relación y posibles alternativas y aplicaciones.

Palabras Clave: Estadística Bayesiana; Diseño Secuencial; Información cuadrática; Utilidad propia; Utilidad no local; Ganancia de Energía Informacional; Discriminación entre modelos.

Clasificación AMS (1980): Primaria, 62L05; Secundaria, 62B15.

Sequential Design to Discriminate between Models, based on the Quadratic Information

Quadratic information is a good alternative to Shannon information for all those problems, that ought to their nature, we are interested to treat with a nonlocal utility. The purpose of this paper is to give, for all those situations, a sequential method for building designs to discriminate between models. This is based on the maximization of the quadratic information.

After an introduction, where the main results and concepts about the quadratic information are summarized, the problem is presented and reduced to a particular case of experiment comparison, with a mixed type parameter space and where the interest quantity is alternatively a subset of its components or the whole set. In both cases, optimum designs are given, as well as its relations, possible alternatives and applications.

Key Words: Bayesian Statistics. Sequential Design. Quadratic Information. Proper Utility. Nonlocal Utility. Informational Energy Gain. Discrimination between Models.

AMS Classification (1980): Primary, 62L05; Secundary, 62B15.

1. INTRODUCCION

El objetivo de este trabajo es dar un método secuencial de construcción de diseños para discriminar entre modelos, basado en la maximización de la información cuadrática.

El problema se plantea en la linea de Box y Hill (1967), Fedorov (1971) y del Río (1977), los cuales proponen una solución basada en la maximización de la información de Shannon. Esto no significa que el estudio aquí realizado sea totalmente paralelo al allí contenido, ya que en el actual trabajo se considera una clase más amplia de modelos, que comprende a los de regresión, y se logra un tratamiento unificado de los criterios marginal y conjunto al considerar la información cuadrática sobre transformaciones no biyectivas del parámetro. En los trabajos anteriores, al no considerar la información de Shannon generalizada para transformaciones del parámetro, era necesaria una mayor casuística en el tratamiento de ambos criterios y la demostración de diversos resultados parciales que podrían haberse obtenido directamente de las propiedades de la información generalizada.

Habiendo demostrado (G-Carrasco (1983) y (1984)) que la información cuadrática es una buena alternativa a la de Shannon, útil en aquellos problemas que por su naturaleza deban ser tratados con una utilidad no local, proporcionando en particular un buen criterio para comparar experimentos, vamos a aplicarlo en este trabajo al caso de discriminación entre modelos.

La idea fundamental de ambos trabajos era considerar el problema de inferencia como uno particular de decisión, con conjunto de acciones las posibles distribuciones sobre el espacio paramétrico y función de utilidad la cudrática, por ser la más sencilla entre las propias, no locales e invariantes frente a transformaciones biyectivas. En el primero se estudiaba la información esperada de un experimento cuando la cantidad de interés era el parámetro de partida θ. En el segundo se extendía esta medida a situaciones en las cuales se puede estar interesado solo en parte de la información dada por el experimento; es decir, la cantidad de interés es $\psi = \psi(\theta)$, función en general no biyectiva de θ, comprendiendo como caso particular el anterior al tomar $\psi(\theta) = \theta$. Entre los diversos conceptos estudiados vamos a destacar y resumir algunos a los que posteriormente haremos referencia.

Sea Ω el espacio paramétrico, $p(\theta)$ $\theta \in \Omega$ una distribución a priori sobre él, $X = \{X;\ p(x / \theta),\ \theta \in \Omega\}$ un experimento y $\psi = \psi(\theta) \in \Psi$ la cantidad de interés, en general no biyectiva de θ.

Si Ψ es discreto y finito, concretamente $\Psi = \{\psi_1,\ldots,\psi_N\}$, tomando para cualquier distribución $p^+(.)$ sobre ψ cuando el valor del parámetro es ψ_i la utilidad

$$u\,(p^+(.), \psi_i) = 2\,p^+(\psi_i) - \Sigma_j\, p^+(\psi_j)^2 \ ; \quad j=1,\ldots,N \tag{1.1}$$

se obtiene, siguiendo el principio general (bayesiano) de maximización de la utilidad esperada (Raiffa y Schlaifer (1961)), la información cuadrática proporcionada por X sobre ψ cuando la distribución a priori es $p(\theta)$.

$$G(X;\psi;p(\theta)) = E_x(\,\Sigma_i\, p(\psi_i /x)^2) - \Sigma_i\, p(\psi_i)^2 = GAN(X;\psi;p(\theta)) \ ; \quad i=1,\ldots N \tag{1.2}$$

donde GAN es la ganancia de enegía informacional estudiada en éstos y otros artículos anteriores (Pardo (1977), G- Carrasco (1982)).

Para el caso de ψ continuo, con una densidad $p(\psi)$ implicada por $p(\theta)$, tomando para cualquier distribución $p^+(.)$ sobre ψ, cuando el verdadero valor del parámetro es ψ_0, la utilidad

$$u_\pi(p^+(.), \psi_0) = 2\{\,p^+(\psi_0) / \pi(\psi_0)\,\} - \int_\Psi (\,p^+(\psi)^2 / \pi(\psi)\,)\, d\psi - 1 \tag{1.3}$$

donde $\pi(.)$ es una densidad de referencia con $u(\pi(.), \psi_0) = 0$, se obtiene, maximizando la utilidad esperada, la información cuadrática proporcionada por X sobre ψ cuando la distribución a priori es $p(\theta)$

$$G_\pi(X;\psi;p(\theta)) = \int_X \int_\Psi p(x)\, \{p(\psi/x)^2/\pi(\psi_0)\}\, d\psi dx - \int_\Psi \{p(\psi)^2/\pi(\psi)\}\, d\psi \tag{1.4}$$

siendo de particular interés tomar como distribución de referencia la a priori, resultando

$$G_p(X;\psi;p(\theta)) = \int_X \int_\Psi\, p(\psi/x)\, p(x/\psi)\, d\psi\, dx - 1 \tag{1.5}$$

Es importante observar que, cuando se requiere uniformidad para los casos discreto y continuo, puede introducirse en (1.1) una distribución $\pi(.)$ de referencia.

Una vez centrados en estas medidas de información cuadrática, se plantea en la sección 2 el problema de construcción de diseños secuenciales para discriminar entre modelos, reduciéndolo en cada etapa a uno de comparación de experimentos con espacio paramétrico mixto. En la sección 3 se aplica el criterio de maximización de la información cuadrática distinguiendo dos casos: el primero, en el que se considera solo la información debida a los modelos, ψ es un subconjunto de las componentes de θ, discreto y finito, y queda por tanto la información cuadrática resumida a la expresión (1.2); el segundo, en el que se considera conjuntamente la información debida a los modelos y parámetros, ψ es igual a θ, y al ser el espacio mixto es necesario definir y estudiar la información cuadrática adecuada, generalización de (1.2) y (1.5), así como posibles alternativas. En la sección 4 se estudia la relación entre los dos criterios anteriores, marginal y conjunto. Por último, en la sección 5, se considera el caso particular de modelos de regresión.

2. PLANTEAMIENTO DEL PROBLEMA

Consideremos la siguiente situación: para cada valor t de una variable controlable dentro de una región T se realiza un determinado experimento cuyo resultado es una variable aleatoria Y_t. Se supone que entre esta variable aleatoria medida y las variables de control existe una cierta relación.

Concretamente se tienen m modelos M_i, $i=1,...,m$, candidatos a explicar el tipo de dependencia de Y_t de t, la cual ocurre para cada i $(i=1,...,m)$ a través de un cierto vector de parámetros desconocidos $w_i = (w_{i1},...,w_{ip_i})' \in W_i \subset R^{p_i}$.

Supongamos que se ha realizado el diseño $D_n = \{t_1,...,t_n; y_1,...,y_n\}$, es decir, se han realizado n observaciones $(y_1,...,y_n)$ obtenidas respectivamente en los puntos $t_1,...,t_n$ de T, no necesariamente distintos. La densidad correspondiente a la observación n+1-ésima, y , realizada en $t \in T$, bajo el modelo M_i, cuando el valor del parámetro es $w_i \in W_i$, condicionada al diseño D_n, se supone conocida y la denotamos $f_{n+1}(y / M_i, w_i; t)$. La dependencia en t de la v.a. Y_t viene explicitada por tanto a través de esa familia de densidades; concretamente, para cada $i=1,...,m$, identificamos M_i con el conjunto de densidades.

$$\{ f_{n+1}(y / M_i, w_i; t);\ \forall\ D_n,\ \forall\ n \geq 0\ ,\ \forall\ w_i \in W_i,\ t \in T\} \tag{2.1}$$

Supongamos además, dado nuestro enfoque bayesiano, que se conocen las probabilidades a priori de los modelos $p_0(M_i)$ i=1,...,m, y las distribuciones a priori de los parámetros en cada modelo, que suponemos continuas, (siempre que no se especifique lo contrario) y con densidades $f_0(w_i/M_i)$ $w_i \in W_i$, i=1,...,m.

A partir de estos datos, $p_0(M_i)$, $f_0(w_i/M_i)$ y $f_{n+1}(y/M_i, w_i; t)$, es posible conocer las nuevas probabilidades de los modelos $p_n(M_i)$ y las densidades de los parámetros dentro de cada modelo $f_n(w_i/M_i)$, una vez realizado el diseño D_n.

Concretamente, la densidad de la primera observación, realizada en $t \in T$ bajo M_i será

$$f_1(y / M_i; t) = \int_{W_i} (f_1 (y / M_i, w_i; t) f_0(w_i / M_i) dw_i$$

y la distribución marginal de y, realizada en $t \in T$ será

$$f_1(y; t) = \Sigma_i\, p_0(M_i)\, f_1(y / M_i; t) \;; \qquad i=1,...,m$$

Con estas notaciones, la probabilidad del modelo M_i y la densidad del párametro w_i bajo M_i, cuando se ha realizado el diseño $D_1= \{t; y\}$, se obtiene respectivamente por las expresiones

$$p_1(M_i) = \{ p_0(M_i)\, f_1(y / M_i; t) \} / f_1(y; t)$$

$$f_1(w_i / M_i) = \{ f_0(w_i / M_i)\, f_1(y / M_i, w_i; t) \} / f_1(y / M_i; t)$$

De modo análogo, supuestas obtenidas estas cuatro distribuciones para n, se obtienen para n+1:

1) La densidad de la observación n+1-ésima, y, realizada en $t \in T$ bajo M_i y condicionada a $D_n= \{t_1,...,t_n;y_1,...,y_n\}$,

$$f_{n+1}(y / M_i; t)= \int_{W_i} f_{n+1}(y / M_i,w_i; t)\, f_n(w_i / M_i)\, dw_i \tag{2.2}$$

2) La distribución marginal de y, realizada en $t \in T$ y condicionada a D_n,

$$f_{n+1}(y; t) = \Sigma_i\, p_n(M_i)\, f_{n+1}(y / M_i; t) \;; \quad i=1,...,m \tag{2.3}$$

3) La probabilidad del modelo M_i cuando se ha realizado el diseño $D_{n+1}= \{t_1,\ldots,t_n,t;\ y_1,\ldots,y_n,\ y\}$,

$$p_{n+1}(M_i) = \{\ p_n(M_i)\ f_{n+1}(y\ /\ M_i;\ t)\ \}/\ f_{n+1}(y;\ t) \tag{2.4}$$

4) La densidad del parámetro w_i en el modelo M_i cuando se ha realizado el diseño D_{n+1},

$$f_{n+1}(w_i\ /\ M_i)= \{\ f_{n+1}(y\ /\ M_i,\ w_i;\ t)\ f_n(w_i\ /\ M_i)\ \}\ /\ f_{n+1}(y\ /\ M_i;\ t) \tag{2.5}$$

Observamos que la dependencia del diseño D_n solo se hace explícita en la notación a través de n, por simplificación y por no dar en general lugar a error. En cualquier caso, en alguno de los desarrollos posteriores, serán utilizadas alternativamente notaciones que ponen de manifiesto parte de esta dependencia; concretamente,

$$p_{n+1}(M_i)= p_{n+1}(M_i\ /\ y)\ \text{ y }\ f_{n+1}(w_i\ /\ M_i)= f_{n+1}(w_i\ /\ M_i,\ y) \tag{2.6}$$

En el siguiente apartado vamos a dar un método, basado en la maximización de la información cuadrática, para la construcción del diseño secuencial. Suponiendo que estamos en la etapa n+1, concretamente, que se ha utilizado un diseño D_n, se trata de elegir el punto $t \in T$ donde se efectuará la observación n+1.

3. METODO SECUENCIAL BASADO EN LA INFORMACION CUADRATICA

En el problema de decisión planteado para la etapa n+1, el espacio paramétrico es $\Omega = \{\ (M_i,\ w_i),\ w_i \in W_i,\ \ i=1,\ldots,m\ \}$, con una distribución a priori $p_n(M_i)\ f_n(w_i\ /\ M_i)$, $w_i \in W_i$, $i=1,\ldots,m$. El experimento correspondiente a realizar la n+1 observación en $t \in T$ es $Y_{t,n+1}= \{R;\ f_{n+1}(y\ /\ M_i,w_i;\ t),\ (M_i,w_i) \in \Omega\}$. El espacio de acciones es el conjunto de punto $t \in T$, ó lo que es análogo, el conjunto de experimentos $Y_{t,n+1}$ con $t \in T$.

Por otro lado, en el contexto de la discriminación entre modelos, estamos interesados en dos tipos de construcción de diseños secuenciales. Uno, que denominamos marginal, considera unicamente la información debida a los modelos, lo que equivale a decir que la cantidad de interés es $\psi(M_i,\ w_i) = M_i$, para todo $w_i \in W_i$, $i=1,\ldots,m$. Otro que denominamos conjunto, considera la información conjunta de modelos y parámetros, en cuyo caso la cantidad de interés es el parámetro original, es decir, $\psi(M_i,w_i)= (M_i,w_i), w_i\in W_i$, $i=1,...,m$. Dada la particular estrutura de Ω, vamos a construir la información cuadrática esperada del experimento $Y_{t,n+1}$ sobre ψ, adecuada a cada uno de estos dos casos.

3.1. CRITERIO MARGINAL

En esto caso $\psi(M_i,w_i) = M_i$, $w_i \in W_i$, $i=1,...,m$; es decir, ψ toma un número discreto y finito de valores; partiendo por tanto, como caso particular de (1.1), de la utilidad cuadrática

$$u(p^+(.), M_i) = 2\, p^+(M_i) - \Sigma_j\, p^+(M_j)^2 \quad ; \quad j=1,...,m \tag{3.1}$$

y obteniendo de (1.2) para el experimento $Y_{t,n+1}$ la información cuadrática

$$\begin{aligned} &GAN(Y_{t,n+1}; \{M_i\}; p_n(M_i)\, f_n(w_i / M_i)) = \\ &\int_R \Sigma_i\, p_{n+1} (M_i / y)^2 f_{n+1} (y; t)\, dy \ - \ \Sigma_i\, p_n(M_i)^2 \ ; \quad i=1,...,m \end{aligned} \tag{3.2}$$

Por tanto, cuando uno está interesado en la información debida solo a los modelos, debe realizar la observación n+1 en el punto $t \in T$ que maximice la cantidad marginal

$$G_M(n+1; t) = \int_R \Sigma_i\, p_{n+1} (M_i / y)^2 f_{n+1} (y; t)\, dy \quad ; \quad i=1,...,m \tag{3.3}$$

3.2 CRITERIO CONJUNTO.

Dada la estructura de $\psi(M_i,w_i)= (M_i,w_i)$, $w_i \in W_i$, $i=1,...,m$, para una construcción rigurosa de la información cuadrática en este caso, parece lógico definir para cualquier distribución $p^+(.)\, f^+(. / .)$ sobre ψ, de modo análogo a (1.1) y (1.3), la utilidad cuadrática

$$\begin{aligned} &u_\pi(p^+(.)\, f^+(. / .), (M_i,w_i)) = 2 \{ p^+(M_i)\, f^+(w_i / M_i) / \pi(w_i / M_i) \} - \\ &- \Sigma_j \int_{W_j}\{ p^+(M_j)^2 f^+(w_j / M_j)^2 / \pi(w_j / M_j) \}\, dw_j \quad j=1,...,m \end{aligned} \tag{3.4}$$

Propiedad. La utilidad definida en (3.4) es propia, invariante frente a transformaciones biyectivas del parámetro, y no local, para cualquier distribución de referencia $\pi(w_j / M_j)$ sobre W_j, $j=1,...,m$; verificando además que

$$u_{\pi}(p^+(.)\ \pi(.\ /\ .),\ (M_i,\ w_i)) = 2p^+(M_i) - \Sigma_j\, p^+(M_j)^2 = u(p^+(.),\ M_i)$$

considerada en (3.1). En efecto: Es propia, puesto que para cualquier par de distribuciones $p^+(.)$ $f^+(.\ /\ .)$ y $p(.)\ f(\ .\ /\ .)$ sobre ψ, se verifica que

$$\Sigma_i \int_{Wi} u_{\pi'}(p^+(.)\ f^+(.\ /\ .),\ (M_i,\ w_i))\ p(M_i)\ f(w_i\ /\ M_i)\ dw_i \leq$$
$$\leq \Sigma_i \int_{Wi} u_{\pi}(\ p\ (.)\ f\ \ (.\ /\ .),\ (M_i,\ w_i))\ p(M_i)\ f(w_i\ /\ M_i)\ dw_i \ \ ; \ \ i=1,...,m$$

por ser ésta, después de algunas operaciones, equivalente a la desigualdad trivial

$$\Sigma_i \int_{Wi} \{\ (\ p^+(M_i)\ f^+(w_i\ /\ M_i)\ -\ p(M_i)\ f(w_i\ /\ M_i\ ))^2\ \}\ /\ \{\ \pi(w_i\ /\ M_i\ )\ \}\ dw_i \geq 0$$

Es no local, por la definición. Es invariante frente a transformaciones biyectivas; concretamente, si $g(M_i,w_i) = (M'_i,w'_i)$ es biyectiva, para las distribuciones implicadas $p^*(M'_i)$, $f^*(w'_i\ /\ M'_i)$ y $\pi^*(w'_i\ /\ M'_i)$ por $p^+(M_i)$, $f^+(w_i\ /\ M_i)$ y $\pi(w_i\ /\ M_i)$ respectivamente, es immediato comprobar que

$$u_{\pi^*}(p^*(.)\ f^*(.\ /\ .),\ (M'_i,w'_i)) = u_{\pi}(p^+(.)\ f^+(.\ /\ .),\ (M_i,\ w_i))\ .$$

Una vez definida la utilidad (3.4) y visto que verifica las propiedades deseadas, consideramos de modo análogo a (1.5) como distribución de referencia la a priori $f_n(w_i\ /\ M_i)$. Maximizando la correspondiente utilidad esperada, con la notación (2.6) y realizando algunas operaciones, se obtiene la información cuadrática para el experimento $Y_{t,n+1}$

$$G_{f_n}(\ Y_{t,n+1};\ (M_i,w_i);\ p_n(M_i)\ f_n(w_i\ /\ M_i)\ ) =$$

$$= \int_R \Sigma_i \int_{Wi} u_{f_n}(\ p_{n+1}(.\ /\ y)\ f_{n+1}(.\ \ /.\ ,\ y),\ (M_i,w_i)\ )$$

$$p_{n+1}(M_i\ /\ y)\ f_{n+1}(w_i\ /\ M_i,\ y)\ f_{n+1}(y;\ t)\ dw_i\ dy\ -$$

$$-\ \ \Sigma_i \int_{Wi} u_{f_n}(\ p_n(.)\ f_n(.\ \ /.\ ),\ (M_i,w_i)\ )\ p_n(M_i)\ f_n(w_i\ /\ M_i)\ dw_i =$$

$$= \int_R \Sigma_i \int_{Wi} \{\ p_{n+1}(M_i\ /\ y)^2 f_{n+1}(w_i\ /\ M_i,\ y)^2 /\ f_n(w_i\ /\ M_i)\ \}\ f_{n+1}(y;\ t)\ dw_i\ dy\ -$$

$$-\ \ \Sigma_i \int_{Wi} p_n(M_i)^2 f_n(w_i\ /\ M_i)\ dw_i =$$

$$= \int_R \Sigma_i \int_{Wi} p_{\ n+1}(M_i\ /\ y)\ f_{n+1}(w_i\ /\ M_i,\ y)\ f_{n+1}(y\ /\ M_i,w_i;\ t)\ p_n(M_i)\ dw_i\ dy$$

$- \ \Sigma_i \, p_n(M_i)^2 \qquad\qquad i=1,...,m \qquad\qquad (3.5)$

Por tanto, cuando uno está interesado en la información debida a los modelos y parámetros, debe realizar la observación n+1 en el punto $t \in T$ que maximice la cantidad conjunta

$$G_C(n+1;\,t) = \int_R \Sigma_i \int_{W_i} p_{\,n+1}(M_i\,/\,y)\; f_{n+1}(w_i\,/\,M_i,\,y)\; f_{n+1}(y\,/\,M_i,w_i;\,t)\; p_n(M_i)\; dw_i\, dy$$
$$i=1,...,m \qquad\qquad (3.6)$$

Observación. De modo alternativo a (3.4) podrían haberse considerado otras utilidades $u(p^+(.)\, f^+(.\,/\,.),\,(M_i,w_i))$ cuadráticas (propias, invariantes y no locales); por ejemplo, tomando distribuciones de referencia $\pi(M_i)$ $\pi(w_i\,/\,M_i)$ sobre (M_i,w_i), ó una referencia $k_i\,\pi(w_i\,/\,M_i)$ con $k_i > 0$ constante para cada $i=1,...,m$.

También podría definirse la utilidad cuadrática no invariante (propia y no local)

$$u(p^+(.)\, f^+(.\,/\,.),\,(M_i,w_i))= \qquad\qquad (3.7)$$

$$= 2\, p^+(M_i)\, f^+(w_i\,/\,M_i) - \ \Sigma_j \int_{W_j} p^+(M_j)^2\, f^+(w_j\,/\,M_j)^2\, dw_j \quad ; \ j=1,...,m$$

que llevaría a tomar la observación n+1 en el punto $t \in T$ que maximizara la ganancia de energía informacional.

Es interesante, por último, observar que dada una utilidad propia u (en particular todas las cuadráticas aquí consideradas), para cualquier transformación lineal $u' = Au + B(M_i,w_i)$ con $A > 0$, los diseños secuenciales basados en la maximización de la utilidad esperada coinciden para ambas, u y u'.

La elección de la utilidad (3.4) considerada para el criterio conjunto se ha realizado por ser la más sencilla de las cuadráticas invariantes y por la buena relación que proporciona entre las informaciones conjunta y marginal.

4. RELACION ENTRE LOS CRITERIOS CONJUNTO Y MARGINAL

Propiedad. Si para todo $i=1,...,m$ existe un $w_{i0} \in W_i$ tal que $p(w_{i0}\,/\,M_i) = 1$, se verifica que $G_C(n+1;\,t) = G_M(n+1;\,t)$, para todo $n \geq 0$ y para todo $t \in T$.

En efeto, vamos a demostrar por inducción que para todo n se verfica que $f_{n+1}(y / M_i; t) = f_{n+1}(y / M_i, w_{i0}; t)$ y $p_n(w_{i0} / M_i) = 1$, para cada $i=1,...m$.

Por las definiciones (2.2) a (2.5), para $n=0$ y w_i discretas, es inmediato que en nuestras hipótesis $f_1(y / M_i; t) = f_1(y / M_i, w_{i0}; t)$ y $p_1(w_{i0} / M_i) = 1$; además, supuesto para n resulta que $f_{n+1}(y / M_i; t) = f_{n+1}(y / M_i, w_{i0}; t)$ y $p_{n+1}(w_{i0} / M_i, y) = p_{n+1}(w_{i0} / M_i) = 1$.

Por tanto,

$$G_C(n+1; t) = \int_R \Sigma_i \, p_{n+1}(M_i / y) \, f_{n+1}(y / M_i; t) \, p_n(M_i) \, dy =$$

$$= \int_R \Sigma_i \, p_{n+1}(M_i / y)^2 \, f_{n+1}(y ; t) \, dy = G_M(n+1; t) ; \qquad i=1,...,m$$

Observación 1. En las hipótesis de la propiedad no existe una densidad sobre W_i, sino que toda la masa se concentra en w_{i0}; aún así y en general para $W_i = \{w_{i1},...,w_{iki}\}$ es útil y conveniente considerar la utilidad cuadrática

$$u_{p_n}(p^+(.)\, p^+(. / .), (M_i, w_{i1})) = \{ 2\, p^+(M_i)(p^+(w_{i1} / M_i) \} / \{p_n(w_{i1} / M_i) \} -$$

$$- \Sigma_j \, \Sigma_k \, \{ p^+(M_j)^2 \, p^+(w_{jk} / M_j)^2 \} / \{ p_n(w_{jk} / M_j) \} ; j=1,...,m ; \; k=1,...,k_j \qquad (4.1)$$

definida de modo análogo a (3.4) y con distribución de referencia la a priori $p_n(w_{i1} / M_i)$, la cual se comprueba facilmente que es propia, no local e invariante, verficando además para todas las distribuciones $p^+(.)$ sobre M_i que $u_{p_n}(p^+(.)\, p_n(. / .), (M_i, w_{i1})) = u(p^+(.), M_i)$, definida en (3.1).

Para esta utilidad (4.1) se obtiene la informacion cuadrática conjunta análoga a (3.6) y utilizada en la demostración de la propiedad

$$G_C(n+1; t) = \int_R \Sigma_i \, \Sigma_l \, p_{n+1}(M_i / y) \, p_{n+1}(w_{i1} / M_i, y) \, f_{n+1}(y / M_i, w_{i1}; t) p_n(M_i) \, dy$$
$$i=1,...,m ; \; l=1,...,k_i \qquad (4.2)$$

Observación 2. También con la utilidad (1.1) para (M_i, w_{i1}) discreto y finito, y en las hipótesis de la propiedad se vertifica que

$$G_C(n+1; t) = \int_R \Sigma_i \, \Sigma_l \, p_{n+1}(M_i, w_{il} / y)^2 \, f_{n+1}(y; t) \, dy =$$

$$= \int_R \Sigma_i \, p_{n+1}(M_i / y)^2 \, f_{n+1}(y; t) \, dy = G_M(n+1; t) ; \quad i=1,...,m ; \; l=1,...,k_i$$

La duda entre utilizar como base del diseño, en las hipótesis de la propiedad, las utilidades (1.1) ó (4.1), por ser esta situación un caso particular de ambas, queda resuelta al coincidir las dos expresiones en este caso.

5. MODELOS DE REGRESION

Hemos definido en (2.1) un modelo M_i general como la familia de densidades $\{f_{n+1}(y \,/\, M_i, w_i;\, t);\ \ \forall\, D_n,\ \forall\, n \geq 0,\ \forall\, w_i \in W_i,\ \forall\, t \in T\}$.

Un caso particular de interés es el de observaciones independientes, es decir, $f_{n+1}(y \,/\, M_i,\, w_i;\, t) = f(y \,/\, M_i,\, w_i;\, t)$, no dependiendo de n ni de D_n, y estando por tanto M_i definido por la familia

$$\{f(y \,/\, M_i,\, w_i;\, t);\ \forall\, w_i \in W_i,\ \forall\, t \in T\} \tag{5.1}$$

Aitchison y Dunsmore (1975) denominan a éstos modelos de regresión.

En el caso mas frecuente de modelos de regresión, M_i viene especificado por una relación del tipo

$$Y_t = \eta_i(t,\, w_i) + \varepsilon_i, \tag{5.2}$$

donde η_i es una función conocida, $w_i \in W_i$ un vector de parámetros desconocidos y ε_i una variable aleatoria de distribución conocida y media cero; obteniendo a partir de ésto, al suponer observaciones independientes, la correspondiente familia de densidades del tipo (5.1).

Observamos además que en el caso (5.1), y en particular en los modelos de regresión (5.2), aunque las densidades $f(y \,/\, M_i, w_i;\, t)$ no dependen de n, las distribuciones $f_{n+1}(y \,/\, M_i;\, t)$, $f_{n+1}(y;\, t)$, $p_n(M_i)$ y $f_n(w_i \,/\, M_i)$ definidas en la sección 2, siguen dependiendo de D_n.

Un camino abierto a la investigación es la aplicación del diseño secuencial aquí obtenido, basado en la maximización de la información cuadrática, al caso particular de modelos de regresión lineales en hipótesis de normalidad, el cual es de esperar que proporcione, bajo ciertas hipótesis, buenas relaciones con otros criterios, como el de maximazación de la información de Shannon y el de D-optimalidad (Stone (1959), Fedorov (1972) y del Río (1977)).

REFERENCIAS

AITCHISON, J. AND DUNSMORE, I.R. (1975). Statistical Prediction Analysis. Cambridge: University Press.

BOX, G.E.P. and HILL, W.J. (1967). Discrimination among mechanistic models. *Technometrics*, **9**, 57 -71.

FEDOROV, V.V. (1971). Design of experiments for linear optimality criteria. *Theory Probab. Applic.* **16**, 189 -195.

GARCIA-CARRASCO, M.P. (1982). Criterio Bayesiano para la comparación de experimentos basado en la maximización de la ganancia de energía informacional. *Actas XIII Reunión Nacional de Estadística*, **2** , 65 -72.

GARCIA-CARRASCO, M.P. (1983). The quadratic scoring rule as a basis for experiment comparison. *Questió.* **8**, nº 3, 121-126.

GARCIA-CARRASCO, M.P. (1984). Algunas extensiones y consecuencias de la medida de información cuadrática. *Actas XIV Reunión Nacional de Estadística*, **1**, 444-454.

PARDO, L. (1977). La energía informacional como fundamento de una teoría de la información. *Escuela de Estadística de Madrid.*

RAIFFA, H. AND SCHLAIFER, R.(1961). *Applied Statistical Decision Theory.* Massachussets: The M.I.T. Press.

RIO BUENO, M. (1977). Algunas cuestiones sobre discriminación en modelos de regresión. Universidad de Madrid. Tesis doctoral.

STONE, M. (1959). Application of a measure of information to the design and comparison of regression experiments. *Ann. Math. Statist.* **30**, p.55 -70.