Criterio de selección de reglas de decisión con observaciones continuas

Por PILAR IBARROLA MUÑOZ Y

RICARDO VELEZ IBARROLA

Recibido: 2 noviembre 1983

Presentado por el Académico Numerario D. Sixto Rios.

Abstract

In this work we deal with a decision model using time continuous observations; we define an appropiate class (D_1) of decision rules for which the risk function is well defined. It is considered a subclass D_m of D_1, whose rules mean intuitively that when time increases, and therefore the available information, we are to hope that the final decision improves in the sense of decreasing the associate loss function value. Topological properties of the class D_m are studied, as well as its relations with other decision criteria.

Resumen

Se considera un modelo de decisión con observaciones continuas en el tiempo; se definen las reglas de decisión que se consideran como aptas (clase D_1) y su función de riesgo. Dentro de la clase D_1 se considera una subclase D_m, teniendo las reglas de esta clase el significado intuitivo de que al aumentar el tiempo t, y por tanto la información, es de esperar que la decisión terminal mejore en el sentido de disminuir la pérdida a ella asociada. Se estudian las propiedades topológicas de la clase D_m y su relación con otros criterios de decisión.

1. DESCRIPCION DEL MODELO

De acuerdo con Irle–Schmitz (1974) consideramos un modelo de decisión con observaciones continuas definido por

$$[(\chi,\mathcal{F}), (\mathcal{F}_t)_{t\in R_+}, \Theta, \mathfrak{T} = (P_\theta)_{\theta\in\Theta}, (\mathfrak{A}, \beta_a), (L_\theta)_{\theta\in\Theta}, (c_\theta)_{\theta\in\Theta}]$$

siendo

1) $(\chi,\mathcal{F})$ un espacio medible que representa el espacio muestral.

2) $(\mathcal{F}_t)_{t\in R_+}$ una familia no decreciente de sub–σ–álgebras de $\mathcal{F}$, donde $\mathcal{F}_t$ representa la familia de sucesos observables hasta el instante t inclusive.

3) Θ es el espacio paramétrico y P_θ es una medida de probabilidad en $(\chi, \mathcal{F})$ para cada $\theta \in \Theta$.

4) $(\mathfrak{A}, \beta_a)$ es un espacio medible que representa el espacio de acciones del decisor.

5) Para cada $\theta \in \Theta$, L_θ es una función β_a–medible de $\mathfrak{A}$ en R_+ que representa la pérdida asociada a la acción a si el estado de la naturaleza es θ.

6) Para cada $\theta \in \Theta$, c_θ es una función no decreciente y continua por la derecha de R_+ en R_+, de forma que $c_\theta(t)$ representa el coste de las obsevaciones hasta el tiempo t inclusive cuando el estado de la naturaleza es θ.

Acerca de la filtración $(\mathcal{F}_t)_{t\in R_+}$ supondremos en adelante las "condiciones habituales": $(\mathcal{F}_t)_{t\in R+}$, es continua por la derecha y $\mathcal{F}_0$ contiene los conjuntos de $\mathcal{F}$ con medida P_θ –nula $\forall\theta \in \Theta$.

Introduciremos además las siguientes hipótesis

a) $\mathfrak{A}$ es un espacio métrico separable, β_a la σ–algebra de Borel y, para cada $\theta \in \Theta$, L_θ es continua y acotada.

b) $\forall t \in R_+$, $(P_\theta)_{\theta\in\Theta}$ está dominada, sobre $\mathcal{F}_t$, por una medida de probabilidad equivalente P_t respecto a la cual $\mathcal{F}_t$ es completa.

Representaremos por $\mathfrak{A}^*$ el conjunto de medidas de probabilidad sobre $(\mathfrak{A}, \beta_a)$

– Una aplicación δ_t de χ en $\mathfrak{A}^*$ se denominará una regla de decisión (apta) para el instante t si $\forall B \in \beta_a$, $\delta_t(x, B)$ es $\mathcal{F}_t$–medible cómo función de χ en R_+

De ello se deduce obviamente que, para cualquier función f de $\mathfrak{A}$ en R_+, β_a–medible, la aplicación $f * \delta_t$ de χ en R_+ definida por

$$f * \delta_t(x) = \int_a f(a)\, \delta_t(x, da) \quad \text{es } \mathcal{F}_t\text{–medible}$$

El conjunto de reglas de decisión (aptas) para el instante t será denotado por M_t. En el caso en que $\mathfrak{A}$ sea un espacio métrico y β_a la σ–álgebra de Borel en $\mathfrak{A}$, es fácil ver que $\delta_t \in M_t$ sí y sólo si $f * \delta_t$ es $\mathcal{F}_t$–medible para cualquier función continua y acotada f de $\mathfrak{A}$ en R.

–Una regla de decisión terminal, $\delta = (\delta_t)_{t\in R_+}$ es una familia de reglas (aptas) para cada instante, de forma que el conjunto de reglas de decisión terminales será

$$D = \prod_{t\in R_+} M_t$$

Se denomina regla de decisión al par (δ, τ) constituído por una regla de decisión terminal $\delta \in D$ y un tiempo de parada τ respecto a $(\mathcal{F}_t)_{t \in R}$ (es decir una función $\tau : \chi \to R$ tal que $\{\tau \leq t\} \in \mathcal{F}_t, \forall t \in R_+$).

El conjunto de tiempos de parada respecto a $(\mathcal{F}_t)_{t \in R_+}$ se designará por $\mathcal{T}$, mientras que $\mathcal{T}_e$, $\mathcal{T}_d$ denotarán aquellos tiempos de parada tales que $\mathcal{T}(x)$ es $\mathcal{T}-c\ s$, un conjunto finito o numerable respectivamente.

Para dar una definición de riesgo adecuada a nuestros propósitos destacaremos la clase de reglas de decisión terminales

$$D = \Big\{\delta \in D \mid \forall \theta \in \Theta, (L_\theta * \delta_t)_{t \in R_+} \text{ es un proceso tal que,}$$

$$P_\theta - c\ s \text{ existe } \lim_{s \in Q_+, s \downarrow t} L_\theta * \delta_s = L_\theta * \delta_t \quad \forall t \in R_+\Big\}$$

Evidentemente $\forall \delta \in D$, $(L_\theta * \delta_t)_{t \in R_+}$ es un proceso adaptado a $(\mathcal{F}_t)_{t \in R}$ pero además, cuando $\delta \in D_I$, $(L_\theta * \delta_t)_{t \in R_+}$ es un proceso progresivamente medible respecto a $(\mathcal{F}_t)_{t \in R_+}$ (cf Meyer cap. IV –17)

Dado un tiempo de parada τ, sea $\{t_i^\theta\}_{i=1}^{\infty}$ el conjunto numerable de puntos de R_+ para los cuales $P_\theta(A_i^\theta) > 0$ siendo $A_i^\theta = \{x \in \chi \mid \tau(x) = t_i^\theta\}$. Desde luego $A_i^\theta \in \mathcal{F}_{t_i^\theta}$ y por tanto $A_i^\theta \in \mathcal{F}_\tau$ (σ–álgebra de los sucesos anteriores a τ) con ello $A^\theta = \bigcup_{i=1}^{\infty} A_i^\theta$ y $\chi - A^\theta$ pertenecerán también a $\mathcal{F}_\tau$.

Si (δ, τ) es una regla de decisión con $\delta \in D_I$ ó $\tau \in \mathcal{T}_d$ definiremos su riesgo mediante

$$R(\theta, (\delta, \tau)) = E_\theta [L_\theta * \delta_\tau + c_\theta \cdot \tau] = \int_\chi \Big\{L_\theta * \delta_\tau(x) + c_\theta(\tau(x))\Big\} P_\theta(dx)$$

siendo

$$L_\theta * \delta_\tau(x) = \begin{cases} L_\theta * \delta_{t_i^\theta}(x) \text{ si } x \in A_i^\theta \text{ para } i = 1, 2, \ldots \\ L_\theta * \delta_{\tau(x)}(x) \text{ si } x \in \chi - A^\theta \end{cases}$$

Obsérvese que $c_\theta \cdot \tau$ y $L_\theta * \delta_\tau$ son $\mathcal{F}_\tau$–medibles (debido en el segundo caso, a que $(L_\theta * \delta_t)_{t \in R_+}$ es progresivamente medible).

Con tal concepto de riesgo, dos reglas de decisión terminales δ y δ', equivalentes en el sentido de que $\forall t \in R_+$, $\delta_t = \delta_t'$ $\mathcal{T} c.s$ verifican

$$R(\theta, (\delta, \tau)) = R(\theta, (\delta', \tau)) \ \ \forall \theta \in \Theta \quad \text{si } \tau \in \mathcal{T}_d \text{ ó } \tau \in \mathcal{T} \text{ y } \delta, \delta' \in D_I$$

De manera que el decisor puede identificar todas las reglas de decisión terminales equivalentes pertenecientes a D_I. (1)

(1) En [4], supuesto que δ es tal que el proceso $(L_\theta * \delta_t)_{t \in R_+}$ sea medible $\forall \theta \in \Theta$, Irle–Schmitz definen el riesgo como $\int_\chi (L_\theta * \delta_{\tau(x)}(x) + c_\theta(\tau(x))\ P_\theta(dx)$. Sin embargo se restringen inmediatamente a la situación en que $(L_\theta * \delta_t)_{t \in R_+}$ tenga trayectorias continuas por la derecha, en cuyo caso tal definición coincide con la propuesta aquí.

2. CLASES COMPACTAS DE REGLAS DE DECISION TERMINALES

Vamos a considerar una clase de reglas de decisión terminales D_m definida por $\delta \in D_m$ si $\forall \theta \in \Theta$, $(L_\theta * \delta_t)_{t \in R_+}$ es una supermartingala, es decir,

$$L_\theta * \delta_t \geqslant E_\theta [L_\theta * \delta_s / \mathcal{F}_t] \quad \forall s > t \quad \forall \theta \in \Theta$$

Es sabido entonces (cf. [6] Meyer Cap. VI – 2) que $D_m \subset D_I$, de manera que $R(\theta, (\delta, \tau))$ está definido $\forall \tau \in \mathcal{T}$ si $\delta \in D_m$

Intuitivamente que $\delta \in D_m$ significa que al aumentar t y por tanto la información del estadístico, es de esperar que su decisión final mejore en el sentido de disminuir la pérdida a ella asociada.

Sea $C(\mathfrak{A})$ el conjunto de funciones continuas y acotadas de $\mathfrak{A}$ en R.

Resulta natural introducir en M_t, o mejor dicho en el conjunto cociente que se obtiene identificado δ_t y δ'_t si $\delta_t = \delta'_t$, P_t–cs, la topología de Haussdorf, η_t para la cual la familia de conjuntos

$$\left\{ \delta'_t \in M_t \,/\, \left| \int (f^* \delta'_t) g \, dP_t - \int (f^* \delta_t) g \, dP_t \right| < \epsilon \right\}$$

con

$$f \in C(\mathfrak{A}),\ g \in L^1(\chi, \mathcal{F}_t, P_t) \text{ y } \epsilon > 0,$$

constituye una subbase de entornos de δ_t

Con ello la aplicación $\delta_t \to \int (f^* \delta_t) g \, dP_t$ es continua $\forall f \in C(\mathfrak{A})$, $\forall g \in L^1 (\chi, \mathcal{F}_t, P_t)$ y en particular $\delta_t \to \int (L_\theta * \delta_t) \, dP_\theta = E_\theta [L_\theta * \delta_t]$ es continua $\forall \theta \in \Theta$

Irle prueba el siguiente resultado:

Si H es un subconjunto de $\mathfrak{A}^*$ convexo, débilmente cerrado y tal que $\forall \epsilon > 0$ existe un compacto K_ϵ de $\mathfrak{A}$ con $\mu(K_\epsilon) > 1 - \epsilon$, $\forall \mu \in H$ (apretado) entonces

$$M_t (H) = \left\{ \delta_t \in M_t \,/\, \delta_t (x) \in H,\ P_t\text{–cs} \right\}$$

es compacto y secuencialmente compacto en M_t.

Representaremos por:

$$D(H) = \prod_{t \in R_+} M_t(H) \quad \text{y} \quad D_m(H) = D_m \cap D(H)$$

Teorema 2.1.

Si H es convexo, débilmente cerrado y "apretado" en $\mathfrak{A}^*$, $D_m(H)$ es compacto en $D(H)$ con la topología producto.

Demostración: Puesto que $M_t(H)$ es compacto $\forall t \in R_+$, de acuerdo con el teorema de Tychonoff $D(H)$ es compacto con la topología producto y basta ver que $D_m(H)$ es cerrado en $D(H)$ o bien que D_m es cerrado en D.

Sea $\delta^{(a)}$ una red en D_m con $\delta^{(a)} \to \delta_t$ con lo cual $\delta^{(a)} \to \delta_t$, $\forall t \in R_+$ por tanto

$$\int (f * \delta_t^{(a)}) g \, dP_t \to \int (f * \delta_t) g \, dP_t, \ \forall f \in C(\mathfrak{A}), \forall g \in L^1(\chi, \mathfrak{F}_t, P_t), \forall t \in R_+$$

y en particular, tomando $f = L_\theta$ y $g = I_A \dfrac{dP_\theta}{dP_t}$

resulta

$$\int_A (L_\theta * \delta_t^{(a)}) \, dP_\theta \to \int_A (L_\theta * \delta_t) \, dP_\theta, \ \forall \theta \in \Theta, \forall A \in \mathfrak{F}_t, \forall t \in R_+$$

Ahora bien, cómo $\delta^{(a)} \in D_m$ será

$$E_\theta [L_\theta * \delta_s^{(a)} / \mathfrak{F}_t] \leqslant L_\theta * \delta_t^{(a)} \quad \forall s > t, \forall a, \forall \theta \in \Theta$$

o dicho de otro modo

$$\int_A (L_\theta * \delta_s^{(a)}) \, dP_\theta \leqslant \int_A (L_\theta * \delta_t^{(a)}) dP_\theta, \ \forall A \in \mathfrak{F}_t, \forall s > t, \forall a, \forall \theta \in \Theta$$

Por consiguiente

$$\int_A (L_\theta * \delta_s) dP_\theta \leqslant \int_A (L_\theta * \delta_t) \, dP_\theta, \ \forall A \in \mathfrak{F}_t, \forall s > t, \ \forall \theta \in \Theta$$

es decir que $\delta \in D_m$.

3. SEMICONTINUIDAD INFERIOR DE LA FUNCION DE RIESGO

De acuerdo con lo anteriormente expuesto nos limitaremos a la consideración de reglas de decisión (δ, τ) con $\tau \in \mathfrak{T}_d$ ó $\delta \in D_m$

Teorema 3.1.

$\forall \tau \in \mathfrak{T}_d$, $\forall \theta \in \Theta$, $R(\theta, (\delta, \tau))$ es función semicontinua inferiormente de δ en D con la topología producto.

Demostración: Puesto que el sumando del riesgo correspondiente a coste es independiente de δ, bastará analizar la pérdida esperada $E_\theta [L_\theta * \delta_\tau]$

si

$$\tau \in \mathfrak{T}_d \ \text{ es } \ \tau(x) = \sum_{i=1}^{\infty} t_i I_{A_i}(x)$$

siendo:

$$A_i \in \mathcal{F}_{t_i}\ \forall i, \text{ con } \bigcup_{i=1}^{\infty} A_i = \chi \text{ y } A_i \cap A_j = \phi \text{ si } i \neq j;$$

por tanto

$$E_\theta [L_\theta * \delta_\tau] = \sum_{i=1}^{\infty} \int_{A_i} (L_\theta * \delta_{t_i}) \, dP_\theta$$

o como la función de pérdida es no negativa

$$E_\theta [L_\theta * \delta_\tau] = \sup_n \sum_{i=1}^{n} \int_{A_i} (L_\theta * \delta_{t_i}) \, dP_\theta$$

Ahora bien, según hemos visto en el apartado 2, la aplicación

$$\delta_{t_i} \to \int_{A_i} (L_\theta * \delta_{t_i}) \, dP_\theta$$

es continua en M_{t_i} con la topología η_{t_i} de forma que, considerando en $\prod_{t \epsilon R_+} M_t$ la topología producto, para todo n, la aplicación $\delta \to \sum_{i=1}^{n} \int_{A_i} (L_\theta * \delta_{t_i}) \, dP_\theta$ será continua. En definitiva $E_\theta [L_\theta * \delta_\tau]$ es función semicontinua inferiormente de δ, por ser supremo de funciones continuas.

Teorema 3.2.

$\forall \tau \in \mathcal{T}$, $\forall \theta \in \Theta$, $R(\theta, (\delta, \tau))$ es función semicontinua inferiormente de δ en D_m.

Demostración: Fijados $\theta \in \Theta$ y $\tau \in \mathcal{T}$, definamos $\{t_i^\theta\}_{i=1}^{\infty}$, $\{A_i^\theta\}_{i=1}^{\infty}$ y A^θ como en la definición de la función de riesgo del apartado 1, y sean

$$\tau_n^\theta(x) = \begin{cases} \inf \left\{ \frac{k}{2^n} \; / \; \tau(x) \leqslant \frac{k}{2^n} \right\} & \text{si } x \in \chi - A^\theta \\ \infty & \text{si } x \in A^\theta \end{cases}$$

de manera que $\tau_n^\theta \in \mathcal{T}_d$ pues $\left\{ \tau_n^\theta \leqslant t \right\} = (\chi - A^\theta) \cap \left\{ \tau \leqslant \frac{[2^n t]}{2^n} \right\} \in \mathcal{F}_t$

por ser

$$A^\theta \in \mathcal{F}_\tau \text{ y } \frac{[2^n t]}{2^n} \leqslant t.$$

Además

$$\tau_n^\theta(x) \downarrow \tau(x), \forall x \in \chi - A^\theta,$$

con lo cual

$$L_\theta * \delta_{\tau_n^\theta}(x) \xrightarrow[n\to\infty]{} L_\theta * \delta_\tau(x) \quad P_\theta - c.s \text{ en } \chi - A^\theta$$

y por consiguiente

$$\int_{\chi - A^\theta} (L_\theta * \delta_\tau(x))\, dP_\theta = \lim_{n\to\infty} \int_{\chi - A^\theta} (L_\theta * \delta_{\tau_n^\theta}(x)) dP_\theta$$

Así pues

$$E_\theta [L_\theta * \delta_\tau] = \sum_{i=1}^{\infty} \int_{A_i^\theta} (L_\theta * \delta_{t_i^\theta})\, dP_\theta + \lim_{n\to\infty} \int_{\chi - A^\theta} (L_\theta * \delta_{\tau_n^\theta})\, dP_\theta$$

y tanto el primer sumando como los términos de la sucesión del segundo sumando son funciones semicontinuas inferiormente de δ en D, en virtud del razonamiento empleado en el teorema 3.1. y en particular en D_m.

Ahora bien, para $\delta \in D_m$ si T_n es el conjunto de racionales de la forma $\frac{k}{2^n}$ con $k \in N$, como $(L_\theta * \delta_t)_{t \in T_{n+1}}$ es una supermartingala positiva

$$E [L_\theta * \delta_{\tau_n^\theta} \mid \mathcal{F}_{\tau_{n+1}^\theta}] \leqslant L_\theta * \delta_{\tau_{n+1}^\theta} \qquad \text{(cf [6] Meyer Cap. VI – 10)}$$

y en particular, puesto que $\chi - A^\theta \in \mathcal{F}_\tau \subset \mathcal{F}_{\tau_{n+1}^\theta}$

$$\int_{\chi - A^\theta} (L_\theta * \delta_{\tau_n^\theta})\, dP_\theta \leqslant \int_{\chi - A^\theta} (L_\theta * \delta_{\tau_{n+1}^\theta})\, dP_\theta \qquad \forall n \in N$$

de forma que

$$\lim_{n\to\infty} \int_{\chi - A^\theta} (L_\theta * \delta_{\tau_n^\theta})\, dP_\theta = \sup_{n \in N} \int_{\chi - A^\theta} (L_\theta * \delta_{\tau_n^\theta})\, dP_\theta$$

Por tanto $\lim_{n\to\infty} \int_{\chi - A^\theta} (L_\theta * \delta_{\tau_n^\theta})\, dP_\theta$ es semicontinua inferiormente en D_m.

4. RELACION CON OTROS CRITERIOS DE DECISION

I) *Criterio minimax*

Según los resultados de los apartados anteriores para $\tau \in \mathcal{T}_d$ ó $\delta \in D_m$, tenemos que la clase de reglas de decisión terminales consideradas es un conjunto compacto y que la función de riesgo R $(\theta, (\delta, \tau))$ es función semicontinua inferiormente en δ.

Son por tanto aplicables los siguientes teoremas (Ferguson, [0]) (2.9)

Teorema 1.

Dado un juego (Θ, C, R) si existe una topología sobre C tal que *a*) C es compacto, *b*) $R(\theta, \delta)$ es semicontinua en $\delta \in C$, $\forall \theta \in \Theta$: el juego tiene un valor y el estadístico una estrategia minimax. Con lo cual para

$$(\delta, \tau) \in (D \times \mathcal{T}_d) \cup (D_m \times \mathcal{T})$$

Teorema del minimax 4.1.

Sea Z una σ–álgebra fija sobre Θ, $\Theta^* = \{\xi / \xi$ es una distribución probabilística sobre $(\Theta, Z)\}$ y $r(\xi, (\delta, \tau)) = \int_\Theta R(\theta,(\delta,\tau))\ \xi\ (d\theta)$ se cumplirá a τ prefijado:

a) Si $\tau \in \mathfrak{T}_d$:

$$\sup_{\xi\in\Theta^*} \inf_{\delta\in D} r(\xi,(\delta,\tau)) = \inf_{\delta\in D} \sup_{\xi\in\Theta^*} r(\xi,(\delta,\tau))$$

existirá $\delta_0 \in D$ tal que:

$$\sup_\theta R(\theta, (\delta_0,\tau)) \leqslant \sup_\theta R\ (\theta, (\delta,\tau)),\ \ \forall\ \delta \in D$$

b) Si $\tau \in \mathfrak{T}$:

$$\sup_{\xi\in\Theta^*} \inf_{\delta\in D_m} r(\xi, (\delta,\tau)) = \inf_{\delta\in D_m} \sup_{\xi\in\Theta^*} r\ (\xi, (\delta,\tau))$$

y existirá $\delta_0 \in D_m$ tal que $\sup_\theta R(\theta, (\delta_0,\tau)) \leqslant \sup_\theta R\ (\theta, (\delta,\tau)),\ \ \forall\ \delta \in D_m$

II) *Criterio Bayes*

Será igualmente aplicable el teorema (Ferguson [0] (2.10))

Teorema 2.

En las hipótesis del teorema 1 la clase de reglas extensión Bayes en C respecto a distribuciones finitas es esenciamente completa.

De aquí resulta para $(\delta,\tau) \in (D\times \mathfrak{T}_d) \cup (D_m \times\mathfrak{T})$ a τ prefijado:

Teorema de completitud 4.2.

a) Para $\tau \in \mathfrak{T}_d$ la clase de reglas extensión Bayes respecto a distribuciones finitas es esencialmente completa

b) Para $\tau \in \mathfrak{T}$ dentro de la clase D_m, la subclase de reglas extensión Bayes respecto a distribuciones finitas sobre Θ será esencialmente completa.

BIBLIOGRAFIA

[1] FERGUSON.: Mathematical Statistics A Decision Theoretic Approach. *Academic Press.* (1967)

[2] I. I. GIHMAN, A. V SKOROHOD.: The theory of stochastic Processes I. *Springer Verlag* (1974)

[3] I. I. GIHMAN, A V. SKOROHOD.: The theory of Stochastic Processes III. *Sprin Verlag.* (1979)

[4] A. IRLE.: Decision theory for continuous observations: minimax solutions. *Metrika t.* 24 n^{o} 3 p. 163–168. (1977).

[5] A. IRLE, N. SCHMITZ.: Decision theory for continuous observations. Bayes solutions. Trans seventh Prague conf. on Information theory, statis. Decision Functions. Random Processes and Eight European Meeting of statisticians.Tech, Univ. Prague, vol. B. (1974)

[6] E. LANERY.: Solutions Bayesiennes en theorie de la decision statistique. *Ann. Inst. Henri Poincaré. Section B. Vol.* XVIII, n^{o} 1 p. 55–79. (1982).

[7] P. A MEYER, C. DELLACHERIE.: Probabilités et potentiel. Hermann. (1978).

[8] A. WALD, Statistical Decision Functions. J. Wiley, New York. (1950). (reimprimido Chelsea Publishing co., Bronx N. y. 1971)