REVISTA

DE LA

REAL ACADEMIA DE CIENCIAS

EXACTAS, FISICAS Y NATURALES

DE

MADRID

TOMO LXX

CUADERNO SEGUNDO

MADRID
DOMICILIO DE LA ACADEMIA
VALVERDE, 22.—TELEFONO 221-25-29
1 9 7 6
RCFNAT 70 (2) 233-474 (1976)

Artículo 39 de los Estatutos de la Academia:

«La Academia no se hace solidaria de las opiniones cuestionables, en materia científica, de sus individuos. Cada autor es responsable de las proposiciones y asertos que contengan los escritos del mismo que aquélla publique.»

Depósito Legal: M. 894.-1958

TALLERES GRÁFICOS VDA. DE C. BERMEJO - J. GARCÍA MORATO, 122 - MADRID

NUEVOS CRITERIOS DE ORDENACION DE REGLAS DE DECISION

Sixto Ríos

Recibido 4-II-76

Decision problems in which one can admit that *a priori* probability distributions are vectors belonging to an ***uncertainty convex cone*** K, are considered. A ***dominance-K*** order and the relation between this order (a quasibayesien order) and the bayesien order are established. Properties of the admissible-K classes and complete-K classes are given.

1. El problema de dar criterios plausibles para la ordenación de reglas de decisión parece fundamental. Aceptado un criterio, el problema matemático es la determinación de las reglas óptimas u optimales de decisión (*).

Los tres criterios más usados son los de dominancia, minimax y Bayes. Los dos primeros se aplican cuando es *totalmente desconocida* la distribución *a priori* de los estados, y el de Bayes se aplica cuando es *conocida* dicha distribución *a priori* o, al menos, es posible asignarla subjetivamente, tras alguna información relativa a la misma.

El punto de vista que aquí vamos a adoptar, es suponer *parcialmente conocida* la distribución de probabilidad *a priori* y dar un criterio de orden parcial, generalización del orden parcial de dominancia y un criterio minimax, generalización del clásico. Al final establecemos relaciones entre los nuevos criterios y los tradicionales.

2. Sea un problema de decisión cuyos datos son: a) el espacio de estados de la Naturaleza $\Omega\ \{\omega_1, \ldots, \omega_n\}$, que suponemos finito;

(*) Un resumen de este trabajo se encuentra en las Actas del Congreso Internacional de Estadística de Varsovia (1975). 40th Session of the International Statistical Institute, Sixto Ríos, Ordre quasi bayesien des régles de décision, página 594).

b) una familia de medidas de probabilidad $P_\omega(z)$ definidas sobre un espacio Z; c) un conjunto Δ de decisiones δ; d) una clase Φ de reglas de decisión aleatorias ϕ; e) una función de pérdida

$$L(\omega, \phi) = -U[r(\omega, \phi)],$$

que representa la utilidad negativa que viene asociada al estado ω y a la decisión ϕ.

Supondremos que la función de riesgo $R(\omega, \phi)$, correspondiente a una regla de decisión $\phi \in \Phi$, existe y es finita. Como Ω se ha supuesto finito, la función de riesgo viene dada por n números

$$R(\omega_i, \phi) \quad (i = 1, 2, \dots, n).$$

Si definimos la aplicación

$$T: \Phi \longrightarrow R^n,$$

es decir,

$$T(\phi) = [R(\omega_1, \phi), \dots R(\omega_n, \phi)]$$

el conjunto imagen de Φ es un subconjunto $S \subset R^n$. Este S es el llamado *conjunto de riesgo* (o de *puntos de riesgo*).

Diremos que si para dos reglas ϕ_1, ϕ_2 es

$$R(\omega_i, \phi_1) = R(\omega_i, \phi_2) \quad (i = 1, \dots, n),$$

tales reglas son *equivalentes*. De este modo los elementos de S son las clases de equivalencia del conjunto Φ.

Si damos un criterio para seleccionar un $s \in S$, basta luego tomar una $\phi = T^{-1}(\bar{s})$.

Se observa también, que si consideramos un preorden parcial (completo) en las reglas de decisión, al pasar a los puntos de riesgo correspondientes se obtiene un orden parcial (completo), suponiendo Ω finito.

Comencemos por dar ideas intuitivas en el caso de dos estados $\Omega = \{\omega_1, \omega_2\}$. Puede ocurrir que no sea posible conocer con precisión las probabilidades p_1, p_2 de los mismos y sí únicamente saber que se verifica la acotación

$$\alpha_1 \leqslant p_1 \leqslant \alpha'_1, \tag{1}$$

la cual implica

$$1 - \alpha'_1 \leqslant p_2 \leqslant 1 - \alpha_1. \tag{2}$$

Este conjunto de vectores $P = (p_1, p_2)$ forman ángulo determinado por los vectores $(\alpha_1, 1 - \alpha_1)$, $(\alpha'_1, 1 - \alpha'_1)$. Si consideramos el riesgo total

$$\rho (p_1) = p_1 x_1 + p_2 x_2,$$

las rectas de Bayes correspondientes a los valores extremos de p_1 que pasan por un punto $M (\bar{x}_1, \bar{y}_1)$, son:

$$MA: \alpha_1 \; x_1 + (1 - \alpha_1) \; x_2 = \alpha_1 \bar{x}_1 + (1 - \alpha_1) \bar{x}_2$$
$$MB: \alpha'_1 x_1 + (1 - \alpha'_1) x_2 = \alpha'_1 \bar{x}_1 + (1 - \alpha'_1) \bar{x}_2$$

y todas las rectas de Bayes intermedias (*):

$$p_1 x_1 + (1 - p_1) x_2 = p_1 \bar{x}_1 + (1 - p_1) \bar{x}_2 \; (\alpha_1 \leqslant p_1 < \alpha'_1)$$

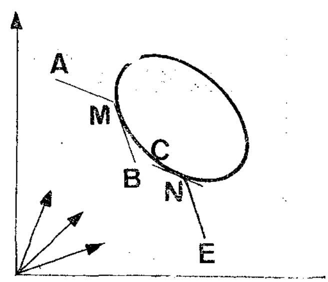

Fig. 1.

nos dan un cierto conjunto de soluciones Bayes cuyos correspondientes puntos de riesgo son los del arco M N.

Se observa que el riesgo de Bayes $\rho (\alpha_1)$, correspondiente a puntos situados en el semiplano inferior limitado por M A, no es superior al riesgo de Bayes $\rho (\alpha_1)$ en M y el riesgo de Bayes $\rho (1 - \alpha'_1)$ de los puntos situados en el semiplano inferior limitado por M B no es superior al riesgo $\rho (1 - \alpha'_1)$ en M; luego en los puntos del ángulo A M B, convexo, intersección de ambos, el riesgo no es superior al riesgo en M, tanto si se evalúa por $\rho (\alpha_1)$ como por $\rho (1 - \alpha'_1)$.

Esto nos da la idea de definir un orden parcial vectorial a partir

(*) Parece natural admitir que el conjunto de vectores de probabilidad que representan la indeterminación inherente al problema sea convexo.

del ángulo A M B como cono positivo y considerar como soluciones del problema de decisión los puntos minimales (decisiones admisibles) correspondientes a tal orden parcial. Se ve intuitivamente que se obtiene así un cierto conjunto M N que coincide con el de las soluciones Bayes correspondientes a los vectores definidos por el ángulo [1] (con análogas excepciones a las que se presentan al comparar las soluciones Bayes con las admisibles con la definición tradicional).

Esto nos sugiere la idea de generalizar la noción corriente de orden natural y relacionarlo con la admisibilidad y las soluciones Bayes.

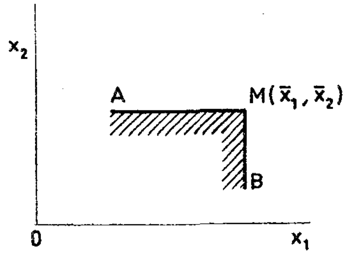

Fig. 2.

Se observa que en el caso del orden natural los dos semiplanos son

$$x_1 \leqslant \bar{x}_1$$
$$x_2 \leqslant \bar{x}_2,$$

y esto corresponde a considerar como posibles todas las probabilidades tales que $0 \leqslant p_1 \leqslant 1$, $p_2 = 1 - p_1$, es decir, $\alpha_1 = 0$, $\alpha'_1 = 1$.

Aparecen así la recta M A como la que corresponde a limitar inferiormente la probabilidad p_1 y la M B a limitarla superiormente.

En este mismo orden de ideas consideraremos también el orden minimax generalizado, asociado al orden parcial de dominancia que hemos introducido.

En definitiva, el objeto de este trabajo es estudiar los métodos de decisión a que conducen las ideas intuitivas indicadas y las relaciones entre los mismos y con los clásicos Bayes, minimax y dominancia.

3. Sea $K \subset K_0$ el *cono de incertidumbre*, cono convexo y cerrado de R^n que contiene el conjunto de los vectores

$$P = (p_1, \ldots, p_n),\ p_i \geqslant 0,\ \sum_1^n p_i = 1, \qquad [1]$$

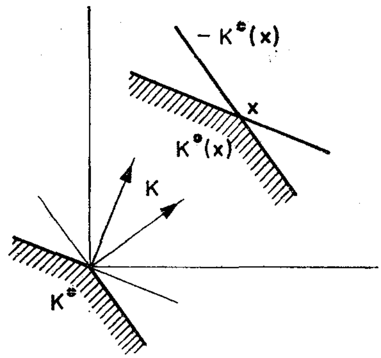

Fig. 3.

que pueden representar distribuciones de probabilidad posibles. (La condición de convexidad no se introduce sólo por facilidad matemática, sino que parece intuitivamente condición natural.)

Sea K^* el *cono polar* correspondiente, es decir:

$$K^* \{ Y \in R^n / P^T \cdot Y \leqslant 0; \quad \forall P \in K \}.$$

Si $P_0 \in K$, K^* está contenido en el semiespacio $P_0^T \cdot Y \leqslant 0$.

Sea K_0 el *ortante fundamental* definido por los n vectores unitarios $[(1, 0, \ldots, 0); (0, 1, \ldots, 0); \ldots (0, \ldots, 0, 1)]$.

Definimos el *cuantante inferior* relativo a X, $K^*(X)$, como el conjunto $K^* + X$.

Si $X \in R^n$, $Y \in R^n$, decimos que Y *K-domina a* X y escribimos $Y \underset{k}{\gtrsim} X$, si $Y \in K^*(X)$.

Decimos que $Y \underset{k}{>} X$ (*K-domina estrictamente*) si $Y \underset{k}{\gtrsim} X$ y no es $X \underset{k}{\gtrsim} Y$.

Decimos que el orden definido por $\gtrsim$ *está incluido en el orden* definido por $\gtrsim^*$, si y sólo si la relación $Y > X$ implica la relación $Y >^* X$, y la relación $Y \sim X$ implica la relación $Y \overset{*}{\sim} X$. Escribiremos $(\gtrsim)$ Inc $(\overset{*}{\gtrsim})$. (Es inmediato ver que esta relación «Inc» es un preorden parcial).

Si $K_1 \supset K_2$, *el orden* $Y \underset{\widetilde{k_1}}{>} X$ *está incluido en el orden* $Y \underset{\widetilde{k_2}}{>} X$, *y recíprocamente.*

Pues si $Y \underset{k_1}{>} X$, es $Y \in K_1^*(X)$, y como $K_1^*(X) \subset K_2^*(X)$, será $Y \in K_2^*(X)$, y como además no es $Y = X$, resulta $Y \underset{k_2}{>} X$.

Teniendo en cuenta la relación evidente

$$K_0 \supset K \supset \{P_0\}$$

resulta *que el orden natural (u orden de dominancia* (*) *correspondiente al ortante* K_0) *está incluido en el orden K y éste en el orden bayesiano correspondiente al vector* P_0. Escribiremos

$$(D_{k_0}) \text{ Inc } (D_k) \text{ Inc } (B_{p_0}).$$

Diremos que X es un punto de riesgo *admisible* (*K*) si no existe otro punto $Y \in S$, tal que $Y \underset{k}{>} X$.

El conjunto de puntos admisibles-K se llama la *clase admisible-K* y la designamos por $A_{(K)}$.

Diremos que C es una *clase completa-*(*K*) de puntos de riesgo si dado un punto de riesgo $X \notin C$ existe otro punto de riesgo $Y \in C$, tal que $Y \underset{k}{>} X$.

Una clase se dice minimal completa-(K) si es completa (K) y no contiene una subclase propia que sea completa (K).

Diremos que X es *Bayes-K* si es Bayes respecto a algún vector $P \in K$.

Es inmediato que en una clase minimal completa (K), si existe $X \in C$, no puede existir $X' \in C$ tal que $X' \underset{k}{>} X$.

TEOREMA 1.—*Si existe una clase C que es minimal completa* (*K*) *coincide con la clase* $A_{(k)}$ *de las reglas admisibles* (*K*).

Supongamos que C es completa (K) y probemos que $A_{(k)} \subset C$. Si $X \in A_{(k)}$, necesariamente $X \in C$, pues si $X \notin C$, existe $Y \in C$, tal que $Y \underset{k}{>} X$, luego no podría ser $X \in A_{(k)}$.

Probemos que si C es minimal completa (K) es $C \subset A_{(k)}$. Supongamos que existe $X \in C$ y $X \notin A_{(k)}$ y probemos que existe $X' \underset{k}{>} X$, $X' \in C$. Esto sería una contradicción con que C es mini-

(*) Véase T. Ferguson. 1967. *Mathematical Statistics* (Academic Press), p. 54.

mal completa (K). Como X $\notin A_{(K)}$, existe $Y \underset{k}{>} X$. Luego si Y $\in$ C estaría demostrado. Si Y $\notin$ C, existirá X' $\in$ C, tal que $X' \underset{k}{>} Y$, y como $Y \underset{k}{>} X$, será $X' \underset{k}{>} X$, luego tenemos una contradicción. Luego $C \subset A_{(k)}$ y por tanto $C = A_{(k)}$.

Es interesante observar que la demostración anterior vale si damos como definiciones de dominancia, admisibilidad, clase completa, minimal completa las que resultan de cualquier preorden parcial introducido en el conjunto S.

TEOREMA 2. — *Si $K_2 \subset K_1 \subset K_0$, la clase de reglas admisibles $A_{(k_2)}$ es un subconjunto de la clase de reglas admisibles $A_{(k_1)}$.*

Hay que probar que si X $\in A_{(k_2)}$, X $\in A_{(k_1)}$. Pues si X $\notin A_{(k_1)}$, existe un $X' \underset{k_1}{>} X$, luego como el orden $\underset{k_1}{>}$ está incluido en el orden $\underset{k_2}{\gtrsim}$, sería $X' \underset{k_2}{>} X$, es decir, X $\notin A_{(k_2)}$, contra la hipótesis.

4. Se dice que un punto X del conjunto convexo S pertenece a la *frontera inferior-K*, que designamos por λ_k (S), si

$$K^* (X) \cap \bar{S} = \{ X \}.$$

Se dice que el conjunto S es *cerrado K-inferiormente* si

$$\lambda_k (S) \subset S.$$

Se dice que *S es acotado-K inferiormente* si existe un X tal que $S \subset -K^* (X)$.

Nuestro objetivo es ahora: 1.º) demostrar teoremas de existencia para soluciones Bayes-K y soluciones admisibles-K; 2.º) relacionar los conjuntos formados por unas y otras y con los λ_k (S); 3.º) relacionarlas con las clases completas-K y minimal completas-K.

Veamos algunas propiedades que relacionan las decisiones Bayes-K con las decisiones admisibles — K. Algunos teoremas incluyen la condición de acotado-K inferiormente, que es superflua en el caso de número finito de estados y decisiones.

TEOREMA 3.—*Supongamos que el conjunto S sea acotado-K inferiormente y cerrado-K inferiormente. A todo vector P $\in$ Int K le corresponde un punto Y $\in$ S que es Bayes-K y admisible-K.*

Consideremos el conjunto B de los números $b = P^T \cdot Y$ en que $Y \in S$. Como $S \subset - K^* (A)$ para un A conveniente, será:

$$b = P^T \cdot Y = P^T \cdot A + P^T \cdot (Y - A) \geqslant m,$$

ya que por ser $Y - A \in - K^*$ y $P \in K$, es $P^T \cdot (Y - A) \geqslant 0$.

Sea $b_0 = \inf_{b \in k} b$ y sea $b^{(h} \in B$ una sucesión de números tales que $\lim b^{(h} = b_0$. A cada número $b^{(h}$ le corresponde un punto

$$Y^{(h} = (y_1^{(h}, \ldots, y_n^{(h})$$

tal que

$$b^{(h} = P^T \cdot Y^{(h} = \sum_1^k p_j \, y_j^{(h} \longrightarrow b_0.$$

Como es $P \in \operatorname{Int} K$, es cada $p_j > 0$ y cada sucesión $y_j^{(h}$ es acotada superiormente, luego existe $\lim y_j^{(n} = y_j^0$.

Por la continuidad

$$\Sigma \, p_j \, y_j^{(h} \longrightarrow \Sigma \, p_j \, y_j^0,$$

luego $\Sigma \, p_j \, y_j^0 = b_0$. Probemos que

$$Y^0 = (y_1^0, \ldots, y_n \,) \in \lambda_k (S).$$

Como Y^0 es un punto límite de puntos de S, $Y^0 \in \overline{S}$, luego

$$\{ Y^0 \} \subset K^* (Y^0) \cap \overline{S}.$$

Veamos que

$$K^* (Y^0) \cap \overline{S} \subset \{ Y^0 \}.$$

Si $Y' \neq Y^0$ y fuera $Y' \in K^* (Y^0)$ e $Y' \in S$, sería $P^T \cdot Y' < b_0$ y habría puntos $Y' \in S$ tales que $P^T \cdot Y' < b_0$ y esto contradiría la hipótesis de que $b_0 = \inf B$, luego

$$K^* (Y^0) \cap \overline{S} = \{ Y_0 \},$$

es decir, $Y^0 \in \lambda_k (S)$.

Como S es cerrado-K inferiormente $\lambda_k (S) \subset S$ y el mínimo valor de $P^T \cdot Y$ se obtiene para un punto Bayes-K, ya que siempre hemos supuesto $P \in \operatorname{Int} K$.

Supongamos que Y_0 no fuera admisible-K. Entonces existe $Y' \underset{k}{>} Y_0$, es decir,

$$Y' \in K^* (Y_0) \tag{1}$$

y

$$Y_0 \notin K^* (Y'). \tag{2}$$

De [1] resulta $Y' - Y_0 \in K^*$, y teniendo en cuenta [2], es $Y' - Y_0 \neq 0$. Como P es interior a K, será $P^T (Y' - Y_0) < 0$. Resulta, pues, $P^T Y' < P^T Y_0$, con lo que Y_0 no puede ser un punto de Bayes, y tenemos una contradicción.

La necesidad de la condición de que P_0 sea interior a K resulta de que si, por ejemplo, $K = K_0$, sin tal condición podría ser P_0 un vector no esencialmente positivo.

TEOREMA 4.—*Si un punto $X \in S$ es tal que $X \in \lambda_K (S)$ es X admisible-K.*

La hipótesis es que $K^* (X) \cap \overline{S} = \{X\}$.

Si X no fuera admisible existiría $X' \underset{k}{>} X$; luego $X' \in K^* (X)$ y $X' \neq X$, luego sería

$$K^* (X) \cap \overline{S} \supset \{ X, X' \}$$

contra la hipótesis.

TEOREMA 5.—*Si S es cerrado y X es admisible-K, es $X \in \lambda_K (S)$.*

X admisible-K equivale a que no existe $X' \in S$ tal que

$$X' \in K^* (X) - \{ X \},$$

o sea que

$$S \cap K^* (X) = \{ X \}.$$

Como $S = \overline{S}$, será

$$\overline{S} \cap K^* (X) = \{ X \},$$

es decir, $X \in \lambda_K (S)$.

TEOREMA 6.—*Si S es acotado-K inferiormente, $\lambda_K (S)$ no es vacío.*

Es consecuencia inmediata del teorema 3.

TEOREMA 7.—*Si X es admisible-K y Ω es finito, X es Bayes-K.*

Si X es admisible-K

$$K^*(X) \cap S = \{X\}.$$

Como $K^*(X) - \{X\}$ y S son convexos y no tienen punto común, existe un vector P tal que $P^T Y \leqslant P^T Z$ para todo $Y \in K^*(X) - \{X\}$ y todo $Z \in S$ (por el teorema del hiperplano separador).

Demostremos que $P \in K$. Sea $X \in S$, es $P^T Y \leqslant P^T X$ para todo $Y \in K^*(X)$. Como $Y \in K^*(X)$, es $Y = X + V$, siendo $V \in K^*$, luego

$$P^T (X + V) = P^T X + P^T V \leqslant P^T X,$$

luego $P^T V \leqslant 0$, y como $V \in K^*$, $P \in K$ (salvo normalización).

TEOREMA 8.—*Si Ω es finito y S es acotado-K y cerrado-K inferiormente, la clase de Bayes-K es completa-K y la de reglas admisibles-K minimal completa-K.*

Probemos primero que $\lambda_k(S)$ es completa-K. Supongamos que $X \notin \lambda_K(S)$. Como $X \in S$, resulta que $S_1 = K^*(X) \cap \overline{S}$ es no vacío, convexo y acotado-K, luego $\lambda_K(S_1) \neq \phi$. Sea $Y \in \lambda_K(S_1)$, es decir, $\{Y\} = K^*(Y) \cap \overline{S}_1$. Como $Y \in K^*(X)$ porque

$$Y \in \overline{S}_1 = \overline{K^*(X) \cap \overline{S}} \subset \overline{K^*(X)} = K^*(X).$$

Resulta $Y \in \lambda_K(S)$, pues

$$\{Y\} = K^*(Y) \cap \overline{S}_1 = \overline{K^*(Y) \cap K^*(X)} \cap \overline{S} =$$

$$= K^*(Y) \cap K^*(X) \cap S = K^*(Y) \cap \overline{S},$$

y por otra parte, como $Y \neq X$, $Y \in K^*(X) - X$, luego $Y \underset{k}{>} X$.

Por el teorema 4, todo $X \in \lambda_K(S)$ es admisible-K; luego ninguna subclase de $\lambda_K(S)$ puede ser completa-K, ya que una clase completa-K debe contener todas las reglas admisibles-K. Luego $\lambda_K(S)$ es minimal completa-K.

Por el teorema 7, la clase de las reglas de Bayes contiene a $\lambda_K(S)$, luego es completa-K.

TEOREMA 9.—*Si el conjunto S es acotado-K_1 y cerrado-K_1 inferiormente y $K_1 \supset K_2$, se verifica $\lambda_{K_1}(S) \supset \lambda_{K_2}(S)$.*

Se observa que las condiciones acotado-K_1 y cerrado K_1 inferiormente, implican las condiciones acotado K_2 y cerrado K_2 inferiormente.

En efecto, $X \in \lambda_{K_2}(S)$ equivale a que

$$K_2^*(X) \cap \overline{S} = \{X\}$$

y vamos a probar que $X \in \lambda_{K_2}(S)$ implica $X \in \lambda_{K_1}(S)$.

Como

$$K_2^*(X) \supset K_1^*(X)$$

será

$$K_1^*(X) \cap \overline{S} \subset K_2^*(X) \cap \overline{S} = \{X\}.$$

Como $X \in K_1^*(X)$ y $X \in \overline{S}$, es

$$K_1^*(X) \cap \overline{S} \supset \{X\},$$

luego

$$K_1^*(X) \cap \overline{S} = \{X\},$$

luego $X \in \lambda_{K_1}(S)$.

En el caso particular de que el cono K se reduzca a un vector $\{P\}$, definiremos como $K^*_{\{P\}}$ el conjunto formado por el origen y el conjunto de puntos Y, tales que $P^T \cdot Y < 0$ (*).

El conjunto $K^*_{\{P\}}(X)$ será

$$K^*_{\{P\}}(X) = X + K^*_{\{P\}}$$

y diremos que $X \in \lambda_{\{P\}}(S)$ si

$$K^*_{\{P\}}(X) \cap \overline{S} = \{X\}.$$

(*) Se observa que esta definición es algo diferente de la que resultaría como generalización natural de K^*, que sería definir $K^*_{\{P\}}$ como el conjunto de puntos Y tales que $P^T \cdot Y \leqslant 0$.

TEOREMA 10.—*Si X es un punto de Bayes respecto a P, se verifica que $X \in \lambda_{\{P\}}(S)$ y recíprocamente.*

Es consecuencia fácil de las definiciones.

TEOREMA 11.—*Si el conjunto S es acotado-K_1 y cerrado-K_1 inferiormente y tenemos una sucesión de conos de incertidumbre*

$$K_1 \supset K_2 \supset \ldots \supset K_n \supset \ldots,$$

tal que $\lim K_n = \{P\}$ (siendo $P \in \operatorname{int} K_n$, para todo n), *se tiene*

$$\lim \lambda_{K_n}(S) = \lambda_{\{P\}}(S).$$

Por el teorema 9 la sucesión $\lambda_{K_n}(S)$ es monótona, es decir,

$$\lambda_{K_1}(S) \supset \lambda_{K_2}(S) \supset \ldots \supset \lambda_{K_n}(S) \supset \ldots$$

Vamos a probar que

$$X \in \lambda_{\{P\}}\{S\} \Rightarrow X \in \lambda_{K_n}(S)$$

para todo n.

Se tiene

$$K^*_{\{P\}}(X) \supset K_n^*(X)$$

para todo n, luego

$$K_n^*(X) \cap \overline{S} \subset K^*_{\{P\}}(X) \cap \overline{S} = \{X\}.$$

Como $X \in K_n^*(X)$ y $X \in \overline{S}$, resulta

$$K_n^*(X) \cap \overline{S} \supset \{X\}, \quad \text{luego} \quad X \in \lambda_{K_n}(X).$$

Como la sucesión $\lambda_{K_n}(S)$ es monótona decreciente y todo punto X de $\lambda_{\{P\}}(S)$ pertenece a $\lambda_{K_n}(S)$ para todo n, X pertenece a

$$\bigcap_{n=1}^{\infty} \lambda_{K_n}(S) = \lim_{n \to \infty} \lambda_{K_n}(S) = L(S),$$

es decir,

$$\lambda_{\{P\}}(S) \subset L(S).$$

Hay que demostrar que, recíprocamente, todo punto de L (S) pertenece a $\lambda_{\{P\}}$ (S). Si $X \in L(S)$, es $X \in \lambda_{K_n}(S)$ para todo *n*.

De aquí resulta:

$$\{X\} = K_n^*(X) \cap \overline{S}$$

para todo *n*, es decir,

$$X \in K_n^*(X)$$

para todo *n*, luego

$$X \in K^*_{\{P\}}(X),$$

luego

$$\{X\} \subset K^*_{\{P\}}(X) \cap \overline{S}.$$

Además, si hubiera otro $X' \neq X$, tal que

$$X' \in K^*_{\{P\}}(X) \cap \overline{S}$$

sería $X' \in K^*_{\{P\}}(X)$, luego $X' \in K_n^*(X)$ para algún *n*, luego

$$X' \in K_n^*(X) \cap \overline{S},$$

luego

$$K_n^*(X) \cap \overline{S} \supset \{X, X'\},$$

es decir, $X \notin \lambda_{K_n}(S)$ contra la hipótesis.

Este teorema sugiere llamar *casi-bayesianas* a las decisiones $X \in \lambda_K(S)$ para *n* suficientemente grande.

Se observa que en este teorema no se utiliza la propiedad de compacidad del conjunto de las soluciones Bayes, que es válida bajo ciertas hipótesis (véase Blackwell, pág. 126) (*). Utilizando dicha propiedad, se podría dar el teorema en forma algo distinta, por un lado más general, pero por otro más restringida.

5. Una ventaja importante de este método sobre otros, en que la información *a priori* es incompleta, reside en que permite com-

(*) *Theory of Games and Statistical Decision.* J. Wiley, 1954.

binarse con la regla de Bayes para incorporar la información dada por un experimento.

Sea K el cono de incertidumbre de $\mathbb{R}^n$, que es dato del problema, sean

$$P(\omega_1/x_i),\ P(\omega_2/x_i), \ldots,\ P(\omega_n/x_i) \quad [i = 1, 2, \ldots, m]$$

las probabilidades *a posteriori*, supuesto que el resultado del experimento haya sido x_i, calculadas a partir del vector

$$P = \{p_1, p_2, \ldots, p_n\} \in K.$$

Al variar P en el cono de incertidumbre *a priori* K, tendremos m conos $K_1, K_2, \ldots, K_m$ de las probabilidades *a posteriori*.

Es fácil ver que éstos son conos convexos si lo es el cono K de los vectores de probabilidad *a priori*.

Basta ver que si

$$[P(\omega_1/x_i),\ P(\omega_2/x_i) \ldots P(\omega_n/x_i)]$$

y

$$[Q(\omega_1/x_i),\ Q(\omega_2/x_i) \ldots Q(\omega_n/x_i)]$$

son vectores de probabilidad *a posteriori*, también lo es

$$[\lambda P(\omega_1/x_i) + \mu Q(\omega_1/x_i), \ldots] \quad (\lambda + \mu = 1,\ \lambda \geqslant 0,\ \mu \geqslant 0).$$

En efecto,

$$\lambda P(\omega_1/x_i) + \mu Q(\omega_1/x_i) = \lambda \frac{P(\omega_1) P(x_i/\omega_1)}{P(x_i)} + \mu \frac{Q(\omega_1) P(x_i/\omega_1)}{Q(x_i)}$$

Por otra parte, al vector *a priori*

$$\lambda' P(\omega_1) + \mu' Q(\omega_1)$$

le corresponde como probabilidad *a posteriori*:

$$\frac{(\lambda' P(\omega_1) + \mu' Q(\omega_1)) P(x_i/\omega_1)}{\lambda' P(x_i) + \mu' Q(x_i)}$$

luego basta determinar el (λ', μ') que hace

$$\frac{(\lambda' P(\omega_1) + \mu' Q(\omega_1)) P(x_i/\omega_1)}{\lambda' P(x_i) + \mu' Q(x_i)} = \lambda P(\omega_1/x_i) + \mu Q(\omega_1/x_i).$$

Considerando los conos polares correspondientes

$$K_1^*,\ K_2^*,\ \ldots,\ K_m^*,$$

podremos definir para cada uno K_i^* el conjunto Bayes-K_i, y entonces definiremos como regla de Bayes-K_1, K_2, ..., K_m), la que hace corresponder a x_i el conjunto Bayes-K_i.

Por ejemplo, si tomamos como distribución *a priori* P [0,8; 0,2] y como experimento $\begin{bmatrix} 0{,}4 & 0{,}9 \\ 0{,}6 & 0{,}1 \end{bmatrix}$ se obtienen:

Como distribuciones *a posteriori*

$$P(\omega_i/X_1):\ [0{,}64;\ 0{,}36]$$
$$P(\omega_i/X_2):\ [0{,}96;\ 0{,}04]$$

y análogamente se obtienen a partir del vector *a priori* P′ [0,5; 0,5] las distribuciones *a posteriori*

$$P'(\omega_i/X_1):\ \left(\frac{4}{13},\ \frac{9}{13}\right)$$

$$P'(\omega_i/X_2):\ \left(\frac{6}{7},\ \frac{1}{7}\right)$$

Es decir, que el cono *a priori* es el (P, P′) y el cono *a posteriori* correspondiente al resultado X_1 es el limitado por los vectores

$$\alpha = [0{,}64;\ 0{,}36] \quad \text{y} \quad \alpha' = \left[\frac{4}{13},\ \frac{9}{13}\right]$$

y el correspondiente al resultado X_2 es el limitado por los vectores

$$\beta = [0{,}96;\ 0{,}04] \quad \text{y} \quad \beta' = \left[\frac{6}{7},\ \frac{1}{7}\right].$$

En la figura están en gruesos los conos polares $K^*(X_1)$ de vértice J y $K^*(X_2)$ de vértice H. Se observa que si el experimento da X_1, las soluciones casi-Bayes son las del segmento MN y si el experimento da X_2, las soluciones casi-Bayes son las del segmento MT.

6. Según es bien sabido, en la ordenación minimax se define δ_1 como mejor que δ_2 si y sólo si

$$\sup_\theta R(\theta, \delta_1) \leqslant \sup_\theta R(\theta, \delta_2).$$

Se dice que δ_0 es una decisión minimax si es

$$\sup_{\theta} R(\theta, \delta_0) = \inf_{\delta} \sup_{\theta} R(\theta, \delta).$$

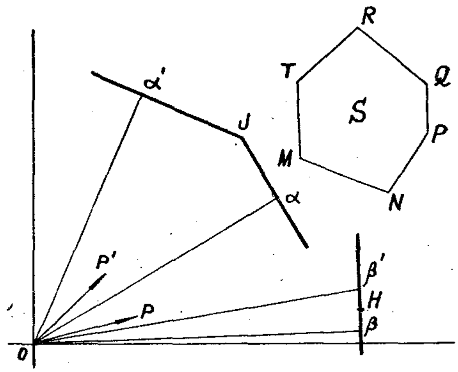

Fig. 4.

Esta definición conduce a la interpretación geométrica conocida. Todos los puntos de la frontera del recinto

$$Q(c) = K_0^*(X_c) = \{(y_1, \ldots, y_k);\ y_1 \leqslant c\}$$

se consideran equivalentes y se busca

$$c_0 = \inf\{c;\ Q(c) \cap S \neq \phi\}.$$

Si existen puntos X que verifican la condición de pertenecer a $Q(C_0) \cap S$, corresponden a reglas de decisión minimax.

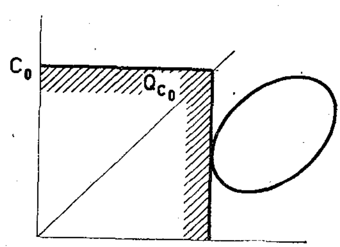

Fig. 5.

Si llamamos riesgo total de la decisión para una distribución

$$P = [p_1, p_2, \ldots, p_n]$$

a la esperanza matemática

$$\rho(P, \delta) = \sum_1^n p_i \, R(\theta_i, \delta)$$

se demuestra que

$$\sup_\theta R(\theta, \delta) = \sup_{P \in K_0} \rho(P, \delta).$$

De aquí resulta que

$$\inf_\delta \sup_\theta R(\theta, \delta) = \inf_\delta \sup_{P \in K_0} \rho(P, \delta);$$

luego δ_0 es tal que

$$\sup_{P \in K_0} \rho(P, \delta_0) = \inf_\delta \sup_{P \in K_0} \rho(P, \delta),$$

lo que puede interpretarse diciendo que el método minimax equivale a buscar para cada δ el $P(\delta) \in K_0$, que nos da el

$$\sup_{P \in K_0} \rho(P, \delta)$$

y buscar después la que nos da el mínimo de éstos.

Geométricamente esto se traduce en considerar los hiperplanos ortogonales a los vectores $P \in K_0$ que pasan por $(y_1^0, \ldots, y_n^0)$ y determinar aquel cuya distancia al origen es máxima. Se ve que esta distancia máxima corresponde a uno de los hiperplanos por $(y_1^0, \ldots, y_n^0)$ paralelo a uno de los coordenados, y de aquí la coincidencia con el método anterior.

Si en vez de considerar la total indeterminación que significa $P \in K_0$ tomamos un cono convexo de incertidumbre $K \subset K_0$, tendremos análogamente un método K-minimax en que se ha de determinar δ_0 tal que

$$\sup_{P \in K_0} \rho(P, \delta_0) = \inf_\delta \sup_{P \in K_0} \rho(P, \delta).$$

Geométricamente esto equivale a considerar equivalentes todos los puntos frontera del recinto

$$Q(c) = K^*(X_c),$$

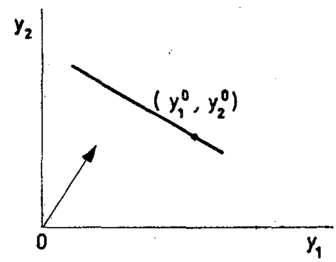

Fig. 6.

en que $X_c = (c, c, \ldots, c)$ y buscar

$$c_0 = \inf \{ c ; \ Q(c) \cap S \neq \phi \}.$$

Si existen puntos X que verifican la condición de pertenecer a $Q(c_0) \cap S$, corresponden a reglas minimax-K.

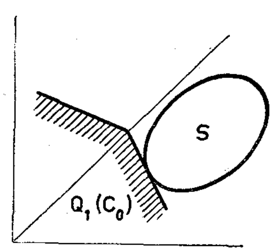

Fig. 7.

Si se toma $K = \{P\}$, es

$$\sup_{\{P\}} \rho(P, \delta) = \rho(P, \delta)$$

y el método se reduce a buscar δ_0 tal que

$$\rho(P, \delta_0) = \inf_{\delta} \rho(P, \delta),$$

es decir, la solución Bayes respecto a P.

8. Se puede combinar la idea de incorporar la información dada por un experimento con la aplicación del método minimax-K. Si el cono de incertidumbre *a priori* K se transforma mediante la información dada por el experiento en los conos *a posteriori*

$$K_1, K_2, \ldots, K_m,$$

como vimos en el § 5, se pueden considerar las soluciones minimax-K_i como más adecuadas que las minimax-K, cuando se sabe que el resultado del experimento ha sido x_i.

Es interesante observar que si el cono de incertidumbre *a priori* es el ortante fundamental K_0 (que conduce al método minimax convencional), todos los conos *a posteriori* coinciden con éste y no cabe incorporar por este procedimiento nueva información.

Tenemos, finalmente, las siguientes relaciones de inclusión entre los métodos aquí considerados:

$$\begin{tikzcd}
 & & (D_k) \arrow[drr, "\mathrm{Inc}"] & & \\
(D_{k0}) \arrow[urr, "\mathrm{Inc}"] \arrow[dr, "\mathrm{Inc}"'] & & & & (B) \\
 & (MM) \arrow[rr, "\mathrm{Inc}"'] & & (MM_k) \arrow[ur, "\mathrm{Inc}"'] &
\end{tikzcd}$$

En un próximo trabajo trataremos la fundamentación axiomática del criterio de dominancia-K en la línea del conocido trabajo de Milnor (*). Además, trataremos de extender estos resultados al caso de infinitos estados.

Agradezco a V. Horra su cuidadosa corrección de pruebas y algunas observaciones que mejoraron la exposición.

(*) *Games against Nature* (Thrall, Cooms, Davies, *Decision Processes*. J. Wiley, 1954, pág. 49).