Trabajos de Estadística y
de Investigación Operativa
Vol. 35, Núm. 2, 1984, pp. 139 a 153

COMPATIBILIDAD DEL MÉTODO DE DE GROOT PARA LLEGAR A UN CONSENSO CON LA FÓRMULA DE BAYES

E. Caro, J. I. Domínguez, F. J. Girón
Departamento de Estadística
e Investigación Operativa.
Facultad de Ciencias.
Universidad de Málaga.

SUMMARY

De Groot's simple method of reaching a consensus when several decision makers have different prior opinions expressed in terms of p.m.'s is shown to be compatible with Bayes rule when sample information is included as an aid to decision making. It is proven that uptdating the priors of the different d.m.'s by means of Bayes theorem and then applying De Groot's method for reaching a consensus (whenever possible) yields the same as by first reaching a consensus and the applying Bayes theorem to the unanimons p.m.

The key point of the proof is that the transition matrix is also updated in every iteration when sample information is considered by means of the predictive distribution.

RESUMEN

En este artículo se prueba que el sencillo método propuesto por De Groot para llegar a un consenso cuando los varios decisores tienen opiniones diferentes

(*) Recibido, Marzo, 1982

expreadas en términos de distribuciones de probabilidad es compatible con la regla de Bayes cuando se tiene en cuenta la información muestral. Se demuestra que si se calcula primero las distribuciones a posteriori y después se aplica el método de De Groot para alcanzar un consenso (cuando esto sea posible), es lo mismo que realizar primero el consenso y después aplicar el teorema de Bayes a esta distribución consensuada.

La clave de la demostración es que tambien la matriz de transición en cada etapa se transforma, en presencia de la información muestral, mediante la utilización de la distribución predictiva.

1. INTRODUCCIÓN

El problema de amalgamar las opiniones de un grupo de decisores aun en el caso de que estos estén de acuerdo en una función de utilidad común no ha encontrado solución satisfactoria. Savage (1954, cap. 10) propone un modelo basado en el principio del minimax que utiliza, en principio, estrategias mixtas. Esta idea se ha utilizado recientemente, basándose en el modelo de juegos cooperativos de Nash, por Weerahandi y Zidek (1981).

Uno de los métodos más simples para amalgamar opiniones, expresadas en términos de medidas de probabilidad subjetiva fue propuesto por Stone (1961) denominado «amalgamación lineal de opiniones». Este método de amalgamación ha sido caracterizado axiomáticamente por Bacharach (1975) y Madansky (1978).

Otros métodos de amalgamar opiniones, en particular el método de «amalgamación logarítmica de opiniones» se han considerado (véase, p.e. Bacharach (1979), Mandansky (loc. cit.) y Weerahandi y Zidek (loc. cit.)). En esta última referencia se dan razones por las que este tipo de amalgamación presenta ventajas sobre el método propuesto por Stone.

En este artículo estudiamos uno de los métodos más simples para obtener un consenso, el dado por De Groot (1974) —y que está emparentado con el método de «amalgamación lineal de opiniones»— en relación con la incorporación de información muestral proporcionada por un experimento.

Este problema fue tratado por Raiffa (1968, cap. 8, secs. 10 a 13) quien argumentaba que era mejor combinar (linealmente) las opiniones para formar una única distribución a priori antes de incorporar la información muestral mediante la aplicación del teorema de Bayes, que aplicar individualmente el teorema de Bayes y después llegar a un consenso. Esta falta de conmutatividad es debida a que el método de amalgamación lineal no es «externamente bayesiano», según la terminología de Madansky (1978). Esto significa que la información muestral no sólamente afecta a las distribuciones a priori sino que los pesos de los coeficientes de la combinación lineal también se alteran según la fórmula de Bayes tomando para estos coeficientes como función de verosimilitud la densidad predictiva, tal como hace Rios (1981) en relación con los modelos jerárquicos, de los cuales el método de amalgamación lineal de opiniones puede considerarse como un caso particular (véase p.e., la sugerencia de De Groot en la discusión del artículo de Girón y Rios (1980)).

Dickey y Gunel (1978) obtienen un resultado análogo (la dependencia de los pesos de la información muestral) para el caso de los «factores de Bayes» del grupo en relación con los de los individuos.

Consideramos un grupo de k individuos, cada uno de los cuales tiene una distribución de probabilidad subjetiva sobre un parámetro desconocido $\theta \in \Omega$ (espacio paramétrico).

Sea:

$$\mathbf{P}^{(0)} = (P_1^{(0)}, \ldots, P_k^{(0)})'$$

el vector formado por las distribuciones de probabilidad a priori que cada individuo asigna al parámetro θ.

De igual modo que en De Groot (*loc. cit.*) y en Berger (1981), suponemos que cada indivíduo, después de ser informado de las distribuciones de los restantes, revisa su distribución y la sustituye por una combinación lineal convexa de las distribuciones $P_1^{(0)}, \ldots, P_k^{(0)}$.

Sea p_{ij} el peso que el individuo «i» asigna a la distribución del individuo «j» (p_{ij} denota la importancia relativa que para el individuo «i» tiene la opinión del individuo «j»). Suponemos además que las p_{ij} son no negativas y que

$$\sum_{j=1}^{k} p_{ij} = 1 \quad (1 \leqslant i \leqslant k).$$

Por tanto, después de la primera revisión, la nueva distribución para el individuo «i» será:

$$P_i^{(1)} = \sum_{j=1}^{k} p_{ij} P_j^{(0)} \tag{1.1}$$

con lo que si denotamos por $P = (p_{ij})$ $(1 \leqslant i, j \leqslant k)$, la matriz estocástica asociada al problema de describir la importancia relativa de las oponiones, entonces:

$$\mathbf{P}^{(1)} = P\mathbf{P}^{(0)}$$

y en general, para sucesivas revisiones, se obtendría,

$$\mathbf{P}^{(n)} = P^n \mathbf{P}^{(0)} \quad (n = 1, 2, 3, \ldots, \text{ad inf.})$$

expresión que nos da el vector de distribuciones después de n revisiones.

De Groot (*loc. cit.*) establece una condición necesaria y suficiente para llegar a una distribución de consenso, distribución en la cual todos los indivíduos estarían de acuerdo, y que llamaremos simplemente «consenso».

Posteriormente Berger (*loc. cit.*) demostró, mediante un contraejemplo, que dicha condición sólo es suficiente, y establece una condición bajo la cual el consenso siempre existe.

De hecho si cada uno de los k decisores aplica la regla de Cromwell (véase Lindley (1980)), es decir, tiene en cuenta las opiniones de todos los decisores, incluidos él mismo, entonces la matriz de pesos $P = (p_{ij})$ es estrictamente positiva y, es bien sabido, que la matriz $P^{(n)}$ siempre converge hacia una matriz de filas iguales, y por lo tanto el consenso siempre se obtiene*.

La expresión analítica del consenso viene dada como una combinación lineal convexa de los elementos del vector $\mathbf{P}^{(0)}$ y cuyas ponderaciones vienen caracterizadas por la forma que toma el límite de la matriz P^n de importancia relativas.

Nuestro principal objetivo consistirá en comprobar la conmutatividad de la regla de Bayes con el método del consenso. Esta comprobación la realizaremos en dos fases:

---

* Esto concuerda bastante bien con la idea intuitiva de que si (algún $p_{ij} = 0$) no se tienen en cuenta las opiniones de los demás, entonces un consenso es improbable.

i) Hacer una observación, aplicar la regla de Bayes y obtener el consenso.

Como se comprobará, la clave de la demostración está en que en cada etapa la matriz de transición también se transforma en presencia de la información muestral.

ii) Efectuar el consenso y a continuación hacer una observación para aplicar la regla de Bayes.

Se completa el trabajo comprobando la validez de estos resultados para el caso en que cada individuo tenga incertidumbre parcial sobre su distribución de probabilidad subjetiva, y para el caso en que los individuos expresen su incertidumbre mediante un modelo jerárquico.

2. CONMUTATIVIDAD DE LA REGLA DE BAYES CON EL MÉTODO DEL CONSENSO

2.1. 1.ª Fase

Observamos una v.a. X, que da información sobre θ, con lo que el vector $\mathbf{P}^{(0)}$ de las distribuciones a priori para el valor $X = x$ de la observación se transformará en el nuevo vector:

$$\mathbf{P}_x^{(0)} = (P_1^{(0)}(\theta/x), \ldots, P_k^{(0)}(\theta/x))'$$

de distribuciones a posteriori, en donde la distribución a posteriori para el individuo «i» viene dada por la fórmula de Bayes:

$$\frac{dP_i^{(0)}(\theta/x)}{dP_i^{(0)}(\theta)} = \frac{f(x/\theta)}{\int_\Omega f(x/\theta)\, dP_i^{(0)}(\theta)} \qquad (1 \leqslant i \leqslant k)$$

siendo $f(x/\theta)$ la verosimilitud asociada.

Por tanto, la primera revisión (1.1) se transforma, después de aplicar el Lema 1.2.1 (Ríos, M. J.), en:

$$P_i^{(1)}(\theta/x) = \sum_{j=1}^{k} {}_1p_{ij}(x) P_j^{(0)}(\theta/x)$$

donde

$$ {}_1p_{ij}(x) = \frac{p_{ij} f(x/P_j^{(0)})}{\sum_{j=1}^{k} p_{ij} f(x/P_j^{(0)})} \quad (i \text{ fijo}) $$

y

$$ f(x/P_j^{(0)}) = \int_\Omega f(x/\theta)\, dP_j^{(0)}(\theta) $$

Es decir, con esto, la matriz inicial de pesos o importancia de las opiniones P se transforma en una nueva matriz, también estocástica, como fácilmente puede comprobarse, que representa los pesos a posteriori que cada indivíduo asigna en una primera etapa después de la observación muestral. Llamemos:

$$ P_1(x) = P(x) = ({}_1p_{ij}(x))_{i,j} \quad (1 \leqslant i, j \leqslant k) $$

y con la idea de simplificar cálculos, expresaremos esta matriz en función de las correspondientes verosimilitudes asociadas.

Denotando por $l_i^{(h)}$ la verosimilitud asociada al individuo «i» en la etapa «h», tendremos:

(2.1) $\quad l_i^{(0)} = f(x/P_i^{(0)}) \quad (1 \leqslant j \leqslant k) \quad$ y

(2.2) $\quad l_i^{(1)} = f(x/P_i^{(1)}) = \int_\Omega f(x/\theta)\, dP_i^{(1)}(\theta) =$

$$ = \int_\Omega f(x/\theta)\, d\left(\sum_{j=1}^{k} p_{ij} P_j^{(0)}(\theta) \right) = \sum_{j=1}^{k} p_{ij} \int_\Omega f(x/\theta)\, dP_j^{(0)}(\theta) = $$

$$ = \sum_{j=1}^{k} p_{ij} f(x/P_j^{(0)}) = \sum_{j=1}^{k} p_{ij} l_j^{(0)}. $$

Con esto, obtenemos:

i) La matriz de pesos a posteriori en la primera etapa toma la forma

$$ P_1(x) = \left(\frac{p_{ij} l_j^{(0)}}{l_i^{(1)}} \right)_{i,j} \quad (1 \leqslant i, j \leqslant k) $$

con lo que el vector formado por las distribuciones a posteriori de cada individuo después de hacer la primera revisión, viene dado por:

$$ \mathbf{P}_x^{(1)} = P_1(x) \mathbf{P}_x^{(0)} $$

ii) si denotamos por

$$\mathbf{L}^1 = (l_1^{(1)}, \ldots, l_k^{(1)})'$$

tenemos que:

$$\mathbf{L}^1 = P\mathbf{L}^0$$

y, en general, si recurrimos al desarrollo realizado en (2.2), es fácil comprobar que:

$$\mathbf{L}^n = P\mathbf{L}^{n-1} = \ldots = P^n\mathbf{L}^0$$

resultado que precisamente coincide con la verosimilitud que se obtendría si comenzamos a aplicar Bayes en la n-ésima etapa, es decir

$$l_j'^{(n)} = f(x/P_j^{(n)}) = \int_\Omega f(x/\theta)\, dP_j^{(n)}(\theta) = \int_\Omega f(x/\theta)\, d\left(\sum_{i=1}^{k} p_{ji}^{(n)} P_i^{(0)}(\theta)\right) =$$

$$= \sum_{i=1}^{k} p_{ji}^{(n)} \int_\Omega f(x/\theta)\, dP_i^{(0)}(\theta) = \sum_{i=1}^{k} p_{ji}^{(n)} f(x/P_i^{(0)})$$

donde $p_{ji}^{(n)}$ es el elemento (j, i) de la matriz P^n. Por lo tanto:

$$\mathbf{L}'^n = P^n\mathbf{L}^0 = \mathbf{L}^n \qquad \forall n.$$

En base a los resultados obtenidos en la primera revisión después de aplicar Bayes, si se efectúa una segunda revisión, cada indivíduo vuelve a asignar una nueva distribución al parámetro, que viene dada por la fórmula

$$P_i^{(2)}(\theta/x) = \sum_{j=1}^{k} {}_2p_{ij}(x) P_j^{(1)}(\theta/x),$$

en donde

$${}_2p_{ij}(x) = \frac{p_{ij} f(x/P_j^{(1)})}{\sum_{j=1}^{k} p_{ij} f(x/P_j^{(1)})} = \frac{p_{ij} l_j^{(1)}}{l_i^{(2)}},$$

y que en forma matricial será:

$$\mathbf{P}_x^{(2)} = P_2(x)\mathbf{P}_x^{(1)}, \quad \text{con} \quad P_2(x) = ({}_2p_{ij}(x))_{i,j} = \left(\frac{p_{ij} l_j^{(1)}}{l_i^{(2)}}\right)_{i,j} \quad (1 \leqslant i, j \leqslant k)$$

Con esto:

$$\mathbf{P}_x^{(2)} = P_2(x)\mathbf{P}_x^{(1)} = P_2(x)P_1(x)\mathbf{P}_x^{(0)} = Q_2(x)\mathbf{P}_x^{(0)}$$

y donde,

$$Q_2(x) = P_2(x)P_1(x) = \left(\frac{p_{ij}l_j^{(1)}}{l^{(2)}}\right)_{i,j}\left(\frac{p_{ij}l_j^{(0)}}{l_i^{(1)}}\right)_{i,j} = \left(p_{ij}^{(2)}\frac{l_j^{(0)}}{l_i^{(2)}}\right)_{ij}(1 \leqslant i,j \leqslant k).$$

Siguiendo un procedimiento análogo para las sucesivas revisiones, obtenemos en general que:

$$P_n(x) = \left(p_{ij}\frac{l_j^{(n-1)}}{l_i^{(n)}}\right)_{i,j} \quad (1 \leqslant i,j \leqslant k);\ (n = 1, 2, \ldots, \text{ad inf.})$$

y en consecuencia, $\mathbf{P}_x^{(n)} = Q_n(x)\mathbf{P}_x^{(0)}$ donde

$$Q_n(x) = \prod_{i=0}^{n-1} P_{n-i}(x) = \left(p_{ij}^{(n)}\frac{l_j^{(0)}}{l_i^{(n)}}\right)_{i,j}, \quad \begin{array}{l}(1 \leqslant i,j \leqslant k);\\ (n = 1, 2, \ldots, \text{ad inf.})\end{array}$$

Con esto, estamos en condiciones de aplicar el método del consenso a estas distribuciones a posteriori.

En lo que sigue, supondremos que en nuestro problema se dan las condiciones para la existencia del consenso dadas por Berger (loc. cit.). Sea:

$$Q_*(x) = \lim_{n\to\infty} Q_n(x) = \left(\Pi_{ij}\frac{l_j^{(0)}}{l_i^*}\right)_{i,j}$$

donde,

$$(\Pi_{ij})_{i,j} = P^* = \lim_{n\to\infty} P^n \quad \text{y} \quad \mathbf{L}^* = \lim_{n\to\infty} L^n = P^*L^0,$$

cuyos elementos son de la forma:

$$l_i^* = \sum_{j=1}^{k} \Pi_{ij}l_j^{(0)} = \sum_{j=1}^{k} \Pi_j l_j^{(0)} = l^* \quad \forall i$$

Es decir,: $\mathbf{L}^* = (l^*, \ldots, l^*)$, y por tanto

$$Q_*(x) = \left(\Pi_j \frac{l_j^{(0)}}{l^*}\right) = l^{*-1} \quad (\Pi_j l_j^{(0)})_j \quad (1 \leqslant j \leqslant k)$$

En definitiva, se obtiene:

$$\mathbf{P}_x^* = \lim_{n \to \infty} \mathbf{P}_x^n = \lim_{n \to \infty} Q_n(x)\mathbf{P}_x^{(0)} = Q_*(x)\mathbf{P}_x^{(0)} = (P_x^*, \ldots, P_x^*),$$

donde

$$P_x^* = \sum_{j=1}^{k} \Pi_j \frac{l_j^{(0)}}{l^*} P_j^{(0)}(\theta/x) = \sum_{j=1}^{k} \alpha_j P_j^{(0)}(\theta/x), \text{ con}$$

$$\alpha_j = \frac{\Pi_j l_j^{(0)}}{l^*} \geqslant 0 \quad \forall j \quad \text{y} \quad \sum_{j=1}^{k} \alpha_j = 1.$$

Es decir, el consenso viene dado por una combinación lineal convexa cuya interpretación es la «media ponderada de las distribuciones a posteriori en la etapa inicial».

En el hipotético caso en que la matriz estocástica de pesos P sea doblemente estocástica, los pesos α_j toman la forma simplificada:

$$\alpha_j = l_j^{(0)} \Big/ \sum_{j=1}^{k} l_j^{(0)} \qquad (1 \leqslant j \leqslant k).$$

2.2. 2.ª Fase

Planteado el problema en su forma estándar y obtenido el consenso, si existe según Berger, llamemos:

$$\Pi = \lim_{n \to \infty} P^n$$

y P^* la distribución del consenso, que será la distribución a priori consensual de los indivíduos.

Si se toma una observación x y se calcula la distribución a posteriori del consenso, la verosimilitud tomará la forma:

$$l_j^{**} = f(x/P^*) = \int_\Omega f(x/\theta)\, dP^*(\theta) = \int_\Omega f(x/\theta)\, d\left(\sum_{j=1}^{k} \Pi_j P_j^{(0)}\right) =$$

$$= \sum_{j=1}^{k} \Pi_j \int_{\Omega} f(x/\theta)\, dP_j^{(0)}(\theta) = \sum_{j=1}^{k} \Pi_j f(x/P_j^{(0)}) =$$

$$= \sum_{j=1}^{k} \Pi_j l_j^{(0)} = l_j^* = l^* \qquad (1 \leqslant j \leqslant k),$$

con lo que la verosimilitud es la misma se haga el consenso antes o después de tomar la observación.

En consecuencia, las distribuciones, a posteriori coinciden tanto si se observa el valor de $X = x$ antes como después de hacer el consenso. Esquemáticamente, el resultado obtenido en este apartado se puede expresar de la siguiente forma:

$$1.^{\text{a}} \text{ Fase:} \quad \mathbf{P}^{(0)} \xrightarrow{\text{Bayes}} B_x(\mathbf{P}^{(0)}) \xrightarrow{\text{Consenso}} (B_x(\mathbf{P}^{(0)}))^*$$

$$2.^{\text{a}} \text{ Fase:} \quad \mathbf{P}^{(0)} \xrightarrow{\text{Consenso}} P^* \xrightarrow{\text{Bayes}} B_x(P^*)$$

y en consecuencia: $(B_x(\mathbf{P}^{(0)}))^* = B_x(P^*)$.

3. EXTENSIONES: EL CASO DE INFORMACIÓN PARCIAL. HIPERDISTRIBUCIONES

En lo que sigue, supondremos que Ω es un espacio métrico separable, lo que implica que Ω^*, (conjunto de las distribuciones de probabilidad que se pueden definir en $(\Omega, \beta(\Omega))$), es también un espacio métrico separable, con la métrica de Prohorov. (Véase Parthasaraty, 1967, th. 62, pág. 43). Vamos a considerar las dos siguientes situaciones:

3.1

La información que cada indivíduo posee acerca del verdadero valor del parámetro viene dada por un subconjunto $K_i^*(1 \leqslant i \leqslant k)$, convexo y compacto, de Ω^*. Como señalan Girón y Rios (1980), el exigir que K_i^* sea convexo es introducido más bien por una condición natural que por conveniencia matemática, ya que si el indivíduo tiene incertidumbre entre dos distribuciones, también la tendrá sobre cualquier distribución que se obtenga como una combinación lineal convexa de ellas.

La forma de tratar el problema en este caso, sería la siguiente: aplicar la técnica de estudio realizada en el parágrafo precedente a cada uno de los casos que se pueden presentar cuando cada indivíduo elige una distribución de probabilidad que pertenezca a su conjunto de distribuciones subjetivas, teniendo así para cada uno de estos casos una distribución de «consenso», si existe, y al hacer el estudio completo obtendríamos un subconjunto convexo y compacto K^* de Ω^*, en donde si la matriz P es ergódica, entonces

$$K^* = \sum_{j=1}^{k} \Pi_j K_j^*,$$

y en el caso de que no lo sea

$$K^* = \bigcap_{i,j} \Pi(i,j) K_{ij}^*,$$

en donde $\Pi(i,j)$ son los propuestos por Berger (loc. cit.) y que, según el apartado anterior, sería independiente de que la observación se realice antes o después de aplicar la regla de Bayes.

Es de interés resaltar, que si bien este caso presenta el aparente inconveniente de que puedan existir más de una solución de consenso, presenta la ventaja de que al trabajar con subconjuntos de Ω^*, y no con elementos, es más factible que las ecuaciones de Berger, que determinan el consenso, tengan solución.

3.2

Supongamos ahora que la información que cada indivíduo posee acerca del verdadero valor del parámetro viene dada por una hiperdistribución $\mu_i \in \Omega^{**}$ $(1 \leqslant i \leqslant k)$ (conjunto de las distribuciones de probabilidad que se pueden definir sobre $(\Omega^*, \beta(\Omega^*))$).

La idea consiste ahora en demostrar la validez de los resultados obtenidos en el apartado 2 para este caso.

Debido a que ahora hay que utilizar un nuevo «operador» (baricentro Q) que viene definido por la fórmula:

$$Q(B) = \int_{\Omega^*} P(B) \mu(dP),$$

los casos a analizar, inicialmente, serán 3!,

1. $\mu_i \xrightarrow{Q} Q(\mu_i) \xrightarrow{C} [\mathrm{Q}(\mu_i)]^* \xrightarrow{B_x} B_x\{[Q(\mu_i)]^*\}$

2. $\mu_i \xrightarrow{C} \mu^* \xrightarrow{Q} Q(\mu^*) \xrightarrow{B_x} B_x[Q(\mu^*)]$

3. $\mu_i \xrightarrow{Q} Q(\mu_i) \xrightarrow{B_x} B_x[Q(\mu_i)] \xrightarrow{C} \{B_x[Q(\mu_i)]\}^*$

4. $\mu_i \xrightarrow{B_x} B_x(\mu_i) \xrightarrow{Q} Q[B_x(\mu_i)] \xrightarrow{C} \{Q[B_x(\mu_i)]\}^*$

5. $\mu_i \xrightarrow{C} \mu^* \xrightarrow{B_x} B_x(\mu^*) \xrightarrow{Q} Q[B_x(\mu^*)]$

6. $\mu_i \xrightarrow{B_x} B_x(\mu_i) \xrightarrow{C} [B_x(\mu_i)]^* \xrightarrow{Q} Q\{[B_x(\mu_i)]^*\}$

según el orden en que actúen los «operadores» Bayes (B_x), Consenso (C, *) y Baricentro (Q), y que abreviadamente denotamos por:

(1) QCB_x; (2) CQB_x; (3) QB_xC; (4) B_xQC; (5) CB_xQ y (6) B_xCQ.

Por el apartado 2, sabemos que los operadores B_x y C conmutan, con lo que si demostramos que los operadores Q y C, por un lado, y Q y B_x por otro, conmutan, se tendrá el resultado deseado.

La conmutatividad de Q y C es una consecuencia inmediata de la linealidad del operador Q, y la de Q y B_x es obvia si se tiene en cuenta que el baricentro de la hiperdistribución a posteriori es

$$Q(B_x(\mu_i)) = \int_{\Omega^*} P_x \mu_x(dP)$$

que coincide, precisamente, con la distribución a posteriori del baricentro.

Consideremos, por último, la estabilidad del modelo cuando la información viene dada por más de una observación. Los dos lemas que siguen nos garantizan que la incorporación de dicha información al modelo de De Groot puede realizarse en cualquier etapa, bien inicialmente, o bien durante el proceso de amalgamación, o bien cuando el consenso se haya alcanzado (si ha lugar), y en cualquier orden si la muestra es la v.a.i.i.d.

Por simplicidad supondremos una muestra de tamaño dos aunque los resultados son, evidentemente, válidos para cualquier tamaño muestral.

Representemos por x e y dos observaciones independientes de la v.a. X que depende del parámetro $\theta \in \Omega$ cuya función de densidad respecto de la medida dominante $\nu(x)$ es $f(x/\theta)$.

Lema 1. La función de verosimilitud generalizada para muestras de tamaño dos se factoriza como sigue:

$$f(x, y/P) = f(x/P)f(y/P_x) = f(y/P)f(x/P_y)$$

DEMOSTRACIÓN

$$f(x/P)f(y/P_x) = f(x/P)\int f(y/\theta)\, dP_x(\theta) = f(x/P)\int f(y/\theta)\frac{f(x/\theta)}{f(x/P)}\, dP(\theta) =$$

$$= \int f(y/\theta)f(x/\theta)\, dP(\theta) = \int f(x, y/\theta)\, dP(\theta) = f(x, y/P)$$

Análogamente se demuestra la otra igualdad.

El lema que sigue demuestra que el operador de Bayes es conmutativo y asociativo respecto de los pesos de una combinación lineal convexa.

Lema 2. Si representamos por $\mu = (\mu_i, \ldots, \mu_k)$ la hiperdistribución asociada al baricentro $Q = \sum_{i=1}^{k} \mu_i P_i$ y por B_x, B_y, B_z los operadores de Bayes correspondientes a las muestras x, y y $z = (x, y)$, referidos a hiperdistribuciones, entonces se tiene que:

$$B_z(\mu) = B_x(B_y(\mu)) = B_y(B_x(\mu)).$$

DEMOSTRACIÓN. Si representamos por $\mu_i(z)$ los pesos a posteriori, tras haber observado z, estos vendrían dados por

$$\mu_i(z) = \frac{\mu_i f(z/P_i)}{f(z/Q)}$$

Por otro lado $\mu_i(x)$ vendrá dado por:

$$\mu_i(x) = \frac{\mu_i f(x/P_i)}{f(x/Q)}$$

Si ahora observamos y, los nuevos pesos, $\mu_i'(x, y)$ vendrán dados por

$$\mu_i'(x, y) = \frac{\mu_i(x) f(y/P_{ix})}{f(y/Q_x)}$$

donde

$$Q_x = B_x(Q) = \sum_{i=1}^{k} \mu_i(x) B_x(P_i) = \sum_{i=1}^{k} \mu_i(x) P_{ix}.$$

Demostrar que $B_z(\mu) = B_y(B_x(\mu))$ es pues equivalente a demostrar que $\mu_i(z) = \mu_i'(x, y)$. Sustituyendo $\mu_i(x)$ por su valor, en la última expresión, resulta

$$\mu_i'(x, y) = \frac{\mu_i f(x/P_i) f(y/P_{ix})}{f(x/Q) f(y/Q_x)}, \quad \text{y por el lema 1}$$

$$\mu_i'(x, y) = \frac{\mu_i f(x, y/P_i)}{f(x, y/Q)} = \frac{\mu_i f(z/P_i)}{f(z/Q)}.$$

Análogamente se demuestra la otra igualdad.

Corolario. Si

$$Q = \sum_{i=1}^{k} \mu_i P_i \quad \text{y} \quad z = (x, y), \text{ entonces}$$

$$B_z\left(\sum_{i=1}^{k} \mu_i P_i\right) = \sum_{i=1}^{k} B_y(B_x(\mu_i)) B_y(B_x(P_i)) = \sum_{i=1}^{k} B_x(B_y(\mu_i)) B_x(B_y(P_i)).$$

REFERENCIAS

BACHARACH, M. (1975). «Group decisions in the face of differences of opinion». *Manag. Sci.* **22**, 182-191.

BACHARACH, M. (1979). «Normal Bayesian dialogues». *Jour. Amer. Statist. Assoc.* **74**, 837-846.

BERGER, R. L. (1981). «A necessary and sufficient condition for reaching a consensus using De Groot's method». *Jour. Amer. Statist. Assoc.* **76**, 415-418.

De Groot, M. H. (1974). «Reaching a consensus». *Jour. Amer. Statist. Assoc.* **69**, 118-121.

Dickey, J. M. y Gunel, E. (1978). «Bayes factor from mixed probabilities». *Jour. Roy. Statist. Soc. B,* **40**, 43-46.

Giron, F. J. y Rios, S. (1980). «Quasi-Bayesian behaviour: a more realistic approach to decisionmaking?. T.E.I.O. **31**, 17-38.

Lindley, D. V. (1980). «The Bayesian approach to Statistics». Unpublished tecnnical report.

Madansky, A. (1978). «Externally Bayesian groups». Unpublished manuscript. University of Chicago.

Raiffa, H. (1968). «Decision Analysis: Introductory Lectures on Choices Under Uncertainty». Addison-Wesley, Reading, Mass.

Rios, M. J. (1981). «Problemas de decisón con información parcial a priori». Tesis Doctoral. Madrid.

Savage, L. J. (1954). «The Foundations of Statistics». Wiley. New York.

Stone, M. (1961). «The opinion pool». *Ann. Math. Statist.* **32**, 1339-1342.

Weerahandi, S. y Zidek, J. V. (1981). «Multi-Bayesian statistical decission theory». *Jour. Roy. Statist. Soc. A,* **144**, 85-93.