

ESTIMACION BAYESIANA MULTIPLE DE UN PARAMETRO

Ricardo Vélez Ibarrola
Departamento de Estadística e I.O.
Universidad Complutense

ABSTRACT

The problem to be analyzed in this paper deals with the finding of n values $x_1, x_2, \dots, x_n \in \mathbb{R}$ which minimizes the function

$$E[\min_{i=1 \dots n} c(\xi - x_i)]$$

where ξ is a one-dimensional random variable with known distribution function ϕ and c is a measurable and positive function.

First, conditions on c in order to ensure the existence of a solution to this problem are determined. Next, necessary conditions to be satisfied by the point (x_1, x_2, \dots, x_n) in which the function attains the minimum are looked for. Finally, conditions on ϕ are given to guarantee the uniqueness of the point which verifies the necessary conditions for minimum, when the function c is $c(y) = |y|$ and $c(y) = y^2$.

1. Introducción

La estimación bayesiana de un parámetro unidimensional sólo consiste, una vez determinada su distribución a posteriori, en encontrar un único punto que minimice el riesgo a posteriori, esperanza de la función de pérdida respecto a dicha distribución.

En determinadas situaciones el interés no estará centrado en un único punto, sino en un conjunto de n puntos que se adapte a la distribución a posteriori de manera óptima, en el sentido precisado por la función de pérdida.

Analizaremos esta situación en el caso de que la función de pérdida dependa de la diferencia entre el verdadero valor del parámetro y su estimación; de forma que los resultados pueden considerarse como una generalización de las medidas de localización clásicas, que podría presentar interés por ejemplo en el caso en que el soporte de la distribución a posteriori no fuese conexo.

La solución serviría así mismo para ciertos problemas de Investigación Operativa de enunciado sencillo: Si un fenómeno se produce en un punto aleatorio, con distribución conocida, a lo largo de una recta, ¿dónde han de situarse n observadores de manera que se haga mínima la esperanza de la distancia al fenómeno del más próximo de ellos?

Concretando el planteamiento, consideremos una variable aleatoria unidimensional ξ , con función de distribución $\phi(x)$, y una función medible $c : \mathbb{R} \rightarrow \mathbb{R}_+$; trataremos de determinar n valores $x_1, x_2, \dots, x_n \in \mathbb{R}$ que minimicen $E[\min_{i=1 \dots n} c(\xi - x_i)]$.

2. Existencia de solución óptima

Nos ocuparemos en primer lugar de buscar condiciones para que el problema tenga solución.

Teorema 1: Si c es continua, decreciente en $(-\infty, 0)$, creciente en $(0, \infty)$, con $\lim_{x \rightarrow \pm\infty} c(x) = \infty$, y existe algún valor x_0 tal que $E[c(\xi - x_0)] < \infty$, entonces la función $E_n(x_1, \dots, x_n) = E[\min_{i=1 \dots n} c(\xi - x_i)]$ alcanza su mínimo para cada $n \in \mathbb{N}$.

Demostración: Empezaremos considerando el caso $n = 1$. Es claro que el conjunto $A = \{x \in \mathbb{R} \mid E_1(x) < \infty\}$ es un intervalo no vacío, puesto que si $x_1 < x_2 \in A$ es $\forall x \in (x_1, x_2)$

$$\begin{aligned}
E_1(x) &= \int_{-\infty}^x c(y-x) \phi(dy) + \int_x^{\infty} c(y-x) \phi(dy) \leq \\
&\leq \int_{-\infty}^x c(y-x_2) \phi(dy) + \int_x^{\infty} c(y-x_1) \phi(dy) < \infty
\end{aligned}$$

Si A no se reduce a un punto, E es continua en el interior de A , puesto que $c(y-z) \xrightarrow{z \rightarrow x} c(y-x)$ y $0 \leq c(y-z) \leq c(y-x+\epsilon) + c(y-x-\epsilon)$ si $z \in (x-\epsilon, x+\epsilon)$, de donde si $(x-\epsilon, x+\epsilon) \subset A$, según el teorema de la convergencia dominada, $E_1(z) \xrightarrow{z \rightarrow x} E_1(x)$.

Sobre el comportamiento de la función E_1 en los extremos a_1, a_2 del intervalo A , el lema de Fatou permite asegurar que $\liminf_{z \rightarrow a_i, z \in A} E_1(z) \geq E_1(a_i)$ (siendo $E_1(a_i) = \int_{-\infty}^{\infty} \lim_{x \rightarrow a_i} c(y-x) \phi(dy)$, que es infinito en el caso en que $a_i = \pm \infty$). Si $E_1(a_i) = \infty$ ello asegura que $\lim_{x \rightarrow a_i, x \in A} E_1(x) = \infty$; mientras que si $E_1(a_i) < \infty$, como $0 \leq c(y-z) \leq c(y-a_i) + c(y-a_i + (-1)^i \epsilon)$ para todo z del intervalo de extremos $a_i, a_i + (-1)^i \epsilon$, será $\lim_{x \rightarrow a_i, x \in A} E_1(x) = E_1(a_i)$.

Según las consideraciones anteriores $\{x \in A \mid E_1(x) \leq K\}$ es un compacto no vacío, si K es suficientemente grande, y E_1 es continua en él, de manera que alcanza su mínimo.

Pasando ahora al caso $n > 1$, tendremos:

$$E_n(x_1, x_2, \dots, x_n) = \int_{-\infty}^{\infty} \min_{i=1 \dots n} c(y-x_i) \phi(dy)$$

donde el integrando es desde luego una función continua de $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

El conjunto donde la función E_n es finita es

$$A_n = \{x \in \mathbb{R}^n \mid \exists a \in A \text{ con } \min_{i=1 \dots n} x_i \leq a \leq \max_{i=1 \dots n} x_i\}$$

puesto que si $\min_{i=1 \dots n} x_i \leq a \leq \max_{i=1 \dots n} x_i$ es $\min_{i=1 \dots n} c(y - x_i) \leq c(y - a) \forall y \notin [\min x_i, \max x_i]$; mientras que si por ejemplo es $\max x_i < a \forall a \in A$ será:

$$\int_{\max x_i}^{\infty} \min c(y - x_i) \phi(dy) = \int_{\max x_i}^{\infty} c(y - \max x_i) \phi(dy)$$

siendo esta última integral divergente, ya que

$$\int_{-\infty}^{\infty} c(y - \max x_i) \phi(dy) = \infty$$

y sin embargo:

$$\int_{-\infty}^{\max x_i} c(y - \max x_i) \phi(dy) < \int_{-\infty}^{\max x_i} c(y - a) \phi(dy) < \infty \quad \forall a \in A$$

La frontera de A_n es consecuentemente $\partial A_n = \{x \in \mathbb{R}^n \mid \max x_i = a_1 \text{ ó } \min x_i = a_2\}$.

No es difícil concluir, usando el teorema de la convergencia dominada, que si $x \in \overset{\circ}{A}_n$ o $x \in \partial A_n$ y $E_n(x) < \infty$, es $\lim_{z \rightarrow x, z \in A_n} E_n(z) = E_n(x)$.

También, en virtud del lema de Fatou, si $x \in \partial A_n$ y $E_n(x) = \infty$, es $\lim_{z \rightarrow x, z \in A_n} E_n(z) = \infty$. En definitiva $E: \bar{A}_n \rightarrow \bar{\mathbb{R}}_+$ es una función continua.

Consideremos ahora $e_n = \inf_{x \in A_n} E_n(x)$ y supongamos que e_{n-1} es accesible en $(x_1^*, \dots, x_{n-1}^*)$. En la situación trivial en que $e_n = e_{n-1}$, sería obviamente $e_n \leq E_n(x_1^*, \dots, x_{n-1}^*, x_n) \leq E_{n-1}(x_1^*, \dots, x_{n-1}^*) = e_{n-1} \forall x_n \in \mathbb{R}$ y e_n resultaría accesible. Podemos por tanto suponer

$e_n < e_{n-1}$ y considerar el conjunto no vacío $\{x \in A_n \mid E_n(x) \leq (e_n + e_{n-1})/2\}$. Este conjunto es desde luego cerrado y podremos concluir que es acotado si observamos que:

$$\begin{aligned} \lim_{x_j \rightarrow \pm\infty} E_n(x_1, x_2, \dots, x_n) &= \lim_{x_j \rightarrow \pm\infty} \int_{-\infty}^{\infty} \min_{i=1 \dots n} c(y - x_i) \phi(dy) = \\ &= \int_{-\infty}^{\infty} \min_{\substack{i=1 \dots n \\ i \neq j}} c(y - x_i) \phi(dy) = E_{n-1}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \geq e_{n-1} \end{aligned}$$

puesto que $\min_{i=1 \dots n} c(y - x_i) \uparrow \min_{\substack{i=1 \dots n \\ i \neq j}} c(y - x_i)$ cuando $x_j \rightarrow \pm\infty$.

La función E_n continua en el compacto anterior alcanza en él su mínimo e_n .

Una vez completada la demostración señalemos que la función

$$c(y) = \begin{cases} 1 - e^{2y} & \text{si } y < 0 \\ 10(1 - e^{-y}) & \text{si } y \geq 0 \end{cases}$$

junto con la distribución exponencial de parámetro 1, muestran que el requerimiento de que $c(x) \rightarrow \infty$ no es inútil. Sin embargo en el caso

en que c sea simétrica, aún cuando sea acotada es sencillo concluir que la función $E_n(x_1, \dots, x_n)$ alcanza su mínimo.

Nos mantendremos a partir de ahora en las condiciones del teorema anterior.

Una observación inmediata, que necesitaremos más adelante, es que, si S es el menor intervalo cerrado que contiene al soporte de ϕ , el punto en que se alcanza el mínimo de $E_n(x_1, \dots, x_n)$ tiene sus n coordenadas contenidas en S ; efectivamente si $x \notin S$ y s es el extremo de S más próximo a x se tiene $c(y - x) > c(y - s) \forall y \in S$, así que si se substituyen las coordenadas x_i exteriores a S por el extremo de S más próximo a x_i la función $\min_{i=1 \dots n} c(y - x_i)$ disminuirá en todo punto de S .

3. Condiciones necesarias de óptimo

Abordaremos ahora el problema de la determinación del punto (x_1, \dots, x_n) en que se alcanza el mínimo de E_n . Suponiendo que la distribución ϕ no está concentrada en un número de puntos menor o igual que n , podemos suponer, sin pérdida de generalidad, la función E definida para $x_1 < x_2 < \dots < x_n \in S$ y designaremos por $\tilde{x}_i, i = 1, \dots, n-1$, la raíz de la ecuación $c(y - x_i) = c(y - x_{i+1})$ que evidentemente existe, es única y está contenida en el intervalo (x_i, x_{i+1}) ; llamaremos además $\tilde{x}_0 = -\infty$ y $\tilde{x}_n = \infty$.

Será entonces:

$$\begin{aligned} E_n(x_1, \dots, x_n) &= \int_{-\infty}^{\infty} \min_{i=1 \dots n} c(y - x_i) \phi(dy) = \\ &= \sum_{i=1}^n \int_{(\tilde{x}_{i-1}, \tilde{x}_i]} c(y - x_i) \phi(dy) \quad (*) \end{aligned}$$

Sea \hat{X}_i el conjunto de puntos en que se alcanza el mínimo de la función $E_1(x)$ correspondiente a la distribución ϕ truncada por \tilde{x}_{i-1} y \tilde{x}_i , es decir:

$$\frac{\phi(x) - \phi(\tilde{x}_{i-1})}{\phi(\tilde{x}_i) - \phi(\tilde{x}_{i-1})}$$

Según la observación del párrafo anterior $\hat{X}_i \subset [\tilde{x}_{i-1}, \tilde{x}_i] \forall i \leq n$ y será:

$$\int_{(\tilde{x}_{i-1}, \tilde{x}_i]} c(y - \hat{x}_i) \phi(dy) < \int_{(\tilde{x}_{i-1}, \tilde{x}_i]} c(y - x) \phi(dy) \quad \forall \hat{x}_i \in \hat{X}_i \quad \forall x \notin \hat{X}_i$$

(*) Por comodidad de notación, convengamos que $(a, \infty]$ significa (a, ∞) y $[-\infty, a]$ significa $(-\infty, a]$.

Consecuentemente, dado un punto $x_1 < x_2 < \dots < x_n$, tendremos, si $z_i \in \hat{X}_i \forall i$,

$$E_n(x_1, \dots, x_n) \geq \sum_{i=1}^n \int_{(\bar{x}_{i-1}, \bar{x}_i]} c(y - z_i) \phi(dy) \geq \int_{-\infty}^{\infty} \min_{i=1 \dots n} c(y - z_i) \phi(dy) = E_n(z_1, \dots, z_n) \quad (1)$$

Para que la primera desigualdad no sea estricta es necesario que $x_i \in \hat{X}_i \forall i \leq n$ y llegamos por tanto a la conclusión siguiente:

Teorema 2: Si se cumplen las hipótesis del teorema 1 y en el punto (x_1, \dots, x_n) con $x_1 < \dots < x_n$ se alcanza el mínimo de la función E_n entonces $x_i \in \hat{X}_i \forall i = 1, \dots, n$.

La utilidad del resultado anterior quedará más clara aplicándolo a las dos funciones c más típicas en este tipo de problemas: $c_1(y) = |y|$ y $c_2(y) = y^2$; suponiendo respectivamente existencia de primer o segundo momento de la distribución ϕ para estar en las condiciones del teorema de existencia.

Puesto que ambas son funciones simétricas será en ambos casos:

$$\bar{x}_i = \frac{1}{2} (x_i + x_{i+1}) \quad \forall i = 1 \dots n$$

Es bien sabido que en el primer caso el mínimo de E_1 se alcanza en la mediana de la distribución, así que:

$$\hat{X}_i = \{x \in \mathbb{R} \mid \phi(x) \geq \frac{1}{2} (\phi(\bar{x}_i) + \phi(\bar{x}_{i-1})) \geq \phi(x^-)\}$$

Tenemos pues:

Corolario 1: Si ξ tiene momento de primer orden finito, y en el punto $(x_1 < \dots < x_n)$ se alcanza el mínimo de $E[\min_{i=1 \dots n} |\xi - x_i|]$ entonces:

$$\phi(x_i) \geq \frac{1}{2} (\phi(\bar{x}_i) + \phi(\bar{x}_{i-1})) \geq \phi(x_{i-}) \quad \forall i = 1 \dots n,$$

siendo $\bar{x}_i = \frac{1}{2} (x_i + x_{i+1})$.

En particular si es continua ha de verificarse:

$$\phi(x_i) = \frac{1}{2} (\phi(\bar{x}_i) + \phi(\bar{x}_{i-1})) \quad \forall i \leq n$$

Para la segunda función propuesta, el mínimo de E_1 se alcanza en la media de la distribución, así que \hat{X}_i se reduce al valor:

$$\int_{(\bar{x}_{i-1}, \bar{x}_i]} y \phi(dy) / (\phi(\bar{x}_i) - \phi(\bar{x}_{i-1}))$$

Corolario 2: Si ξ tiene momento de segundo orden finito, y en el punto $(x_1 < \dots < x_n)$ se alcanza el mínimo de la función $E[\text{mín}_{i=1 \dots n} (\xi - x_i)^2]$ entonces:

$$x_i [\phi(\bar{x}_i) - \phi(\bar{x}_{i-1})] = \int_{(\bar{x}_{i-1}, \bar{x}_i]} y \phi(dy) \quad \forall i = 1 \dots n$$

siendo $\bar{x}_i = \frac{1}{2} (x_i + x_{i+1})$

Merece la pena señalar que los sistemas de ecuaciones contenidos en los dos últimos corolarios, son los que se obtienen, suponiendo buenas condiciones de diferenciabilidad e imponiendo $\partial E_n / \partial x_i = 0 \quad \forall i = 1 \dots n$.

El razonamiento de donde se deduce el teorema anterior tiene alguna consecuencia más. Concretamente para que la segunda desigualdad en (1) no sea estricta es necesario que $\phi(\bar{z}_i) = \phi(\bar{x}_i) \quad \forall i = 1 \dots n - 1$. Así que:

Teorema 2 (continuación): Si $(x_1 < \dots < x_n)$ es un punto en que se alcanza el mínimo de E_n y existe $(z_1 < \dots < z_n) \neq (x_1 < \dots < x_n)$ con $z_i \in \tilde{X}_i \ \forall i \leq n$ entonces $\phi(\tilde{z}_i) = \phi(\tilde{x}_i) \ \forall i \leq n - 1$ y $E_n(z_1, \dots, z_n) = E_n(x_1, \dots, x_n)$.

4. Unicidad de la solución de las condiciones de óptimo

Interesa naturalmente localizar situaciones en las cuales exista un único punto que satisfaga las condiciones necesarias expresadas en el teorema 2. Trataremos dicha cuestión en el contexto de los corolarios del párrafo anterior, es decir, limitándonos a las funciones $c_1(y) = |y|$ y $c_2(y) = y^2$.

En la primera situación la posibilidad de existencia de diversas soluciones de las condiciones necesarias, aún en “buenas” condiciones, es clara a partir de una sencilla interpretación geométrica de las ecuaciones contenidas en el corolario 1: Un punto $(x_1, \tilde{x}_1, x_2, \dots, \tilde{x}_{n-1}, x_n)$ satisface dichas ecuaciones si sus componentes son abcisas de los puntos medios de los lados de una poligonal, de lados paralelos a los ejes, creciente de 0 a 1, y dichos puntos medios están sobre la gráfica de $\phi(y)$.

No es difícil imaginar dos de tales poligonales y hacer pasar una curva estrictamente creciente y derivable por los puntos medios de sus lados. Sin embargo

Teorema 3: Sea ϕ una función de distribución continua, con momento de primer orden finito, cuyo soporte sea un intervalo I de \mathbb{R} (finito o infinito) en cuyo interior ϕ tiene derivada φ continua y estrictamente positiva. Si se verifica $\forall a < b \in I$

$$\left\{ \begin{array}{l} \varphi(a) + \varphi(b) \leq 2\varphi(c) \text{ siendo } c \text{ tal que } \phi(a) + \phi(b) = 2\phi(c) \\ \varphi(b) \leq 2\varphi(c) \text{ siendo } c \text{ tal que } \phi(b) = 2\phi(c) \\ \varphi(a) \leq 2\varphi(c) \text{ siendo } c \text{ tal que } 1 + \phi(a) = 2\phi(c) \end{array} \right.$$

con una al menos de las tres desigualdades estricta, entonces el sistema de ecuaciones $2\phi(x_i) = \phi(\tilde{x}_i) + \phi(\tilde{x}_{i-1}) \quad i = 1 \dots n$, $2\tilde{x}_i = x_i + x_{i+1} \quad i = 1 \dots n-1$ tiene una única solución con $x_1 < x_2 < \dots < x_n \in I$.

Demostración: Los resultados de los apartados anteriores aseguran, puesto que ϕ tiene momento de primer orden finito, la existencia de al menos una solución. Establezcamos un hecho elemental.

Lema: Sea $F(x, y)$ una función diferenciable con continuidad, definida para $x < y$, con $F_y > 0$, $F_x + F_y \leq 0$ y $\lim_{x \uparrow y} F(x, y) < 0$. Si existen $x_0 < y_0$ tales que $F(x_0, y_0) = 0$, entonces $\forall y \in \mathbb{R}$ existe un único $f(y) \in \mathbb{R}$ tal que $F(f(y), y) = 0$. Además f es derivable y $f'(y) \leq 1$, siendo la desigualdad estricta si $F_x + F_y < 0$.

Demostración: Puesto que en la dirección del vector $(1, 1)$ la función F no crece, sobre la recta $x - y = x_0 - y_0$ la función es positiva si $y < y_0$ y negativa si $y > y_0$. Análogamente sobre la recta $x = x_0$, la función F crece y es por tanto estrictamente negativa para $y < y_0$ y estrictamente positiva para $y > y_0$.

Fijado $y_1 > y_0$, la función continua $F(x, y_1)$ es estrictamente positiva en x_0 y negativa en $x_0 - y_0 + y_1$, existe pues $x_1 \in (x_0, x_0 - y_0 + y_1]$ en el cual $F(x_1, y_1) = 0$ y dicho valor es único puesto que $F_x < 0$. Análogamente para $y_1 \in (x_0, y_0)$ existe un único $x_1 \in [x_0 - y_0 + y_1, x_0)$ en el cual $F(x_1, y_1) = 0$.

Para $y_1 \leq x_0$, puesto que $F(x, y_1)$ es positiva para $x = x_0 - y_0 + y_1$ y negativa estrictamente para x suficientemente próximo a y_1 , existe un único $x_1 \in [x_0 - y_0 + y_1, y_1)$ con $F(x_1, y_1) = 0$.

Establecida la existencia de f , el teorema de funciones implícitas garantiza su derivabilidad y es $f' = -F_y/F_x \leq 1$.

Volviendo a la demostración del teorema, como ϕ es estrictamente creciente y continua en I , las n primeras ecuaciones del sistema:

$$\left. \begin{aligned} 2\phi(x_1) &= \phi(\tilde{x}_1) \\ 2\phi(x_i) &= \phi(\tilde{x}_{i-1}) + \phi(\tilde{x}_i) \quad i = 2 \dots n-1 \\ 2\phi(x_n) &= 1 + \phi(\tilde{x}_{n-1}) \end{aligned} \right\}$$

definen cada x_i como función de \tilde{x}_{i-1} y \tilde{x}_i siendo además:

$$\left\{ \begin{aligned} \frac{\partial x_1}{\partial \tilde{x}_1} &= \frac{\varphi(\tilde{x}_1)}{2\varphi(x_1)} \\ \frac{\partial x_i}{\partial \tilde{x}_{i-1}} &= \frac{\varphi(\tilde{x}_{i-1})}{2\varphi(x_i)}, \quad \frac{\partial x_i}{\partial \tilde{x}_i} = \frac{\varphi(\tilde{x}_i)}{2\varphi(x_i)} \quad i = 2, \dots, n-1 \\ \frac{\partial x_n}{\partial \tilde{x}_{n-1}} &= \frac{\varphi(\tilde{x}_{n-1})}{2\varphi(x_n)} \end{aligned} \right.$$

Consideremos ahora las $n-1$ funciones:

$$\left\{ \begin{aligned} F_1(\tilde{x}_1, \tilde{x}_2) &= x_1 + x_2 - 2\tilde{x}_1 \\ F_k(\tilde{x}_{k-1}, \tilde{x}_k, \tilde{x}_{k+1}) &= x_k + x_{k+1} - 2\tilde{x}_k \quad k = 2, \dots, n-2 \\ F_{n-1}(\tilde{x}_{n-2}, \tilde{x}_{n-1}) &= x_{n-1} + x_n - 2\tilde{x}_{n-1} \end{aligned} \right.$$

y será:

$$\frac{\partial F_1}{\partial \tilde{x}_1} = \frac{\varphi(\tilde{x}_1)}{2\varphi(x_1)} + \frac{\varphi(\tilde{x}_1)}{2\varphi(x_2)} - 2, \quad \frac{\partial F_1}{\partial \tilde{x}_2} = \frac{\varphi(\tilde{x}_2)}{2\varphi(x_2)}$$

$$\frac{\partial F_k}{\partial \tilde{x}_{k-1}} = \frac{\varphi(\tilde{x}_{k-1})}{2\varphi(x_k)}, \quad \frac{\partial F_k}{\partial \tilde{x}_k} = \frac{\varphi(\tilde{x}_k)}{2\varphi(x_k)} + \frac{\varphi(\tilde{x}_k)}{2\varphi(x_{k+1})} - 2,$$

$$\frac{\partial F_k}{\partial \tilde{x}_{k+1}} = \frac{\varphi(\tilde{x}_{k+1})}{2 \varphi(x_{k+1})} \quad k = 2 \dots n-2$$

$$\frac{\partial F_{n-1}}{\partial \tilde{x}_{n-2}} = \frac{\varphi(\tilde{x}_{n-2})}{2 \varphi(x_{n-1})}, \quad \frac{\partial F_{n-1}}{\partial \tilde{x}_{n-1}} = \frac{\varphi(\tilde{x}_{n-1})}{2 \varphi(x_{n-1})} + \frac{\varphi(\tilde{x}_{n-1})}{2 \varphi(x_n)} - 2$$

Observemos que de acuerdo con las hipótesis es $\frac{\partial F_1}{\partial \tilde{x}_2} > 0$, $\frac{\partial F_1}{\partial \tilde{x}_1} + \frac{\partial F_1}{\partial \tilde{x}_2} \leq 0$ y $\lim_{\tilde{x}_1 \uparrow \tilde{x}_2} F_1(\tilde{x}_1, \tilde{x}_2) < 0$, luego según el lema anterior $F_1 = 0$ define \tilde{x}_1 como una cierta función f_1 de \tilde{x}_2 , siendo $f_1'(\tilde{x}_2) \leq 1$.

Sustituyendo dicha relación en F_2 será:

$$F_2(f_1(\tilde{x}_2), \tilde{x}_2, \tilde{x}_3) = G_2(\tilde{x}_2, \tilde{x}_3)$$

$$\text{con} \quad \frac{\partial G_2}{\partial \tilde{x}_2} = \frac{\partial F_2}{\partial \tilde{x}_1} f_1' + \frac{\partial F_2}{\partial \tilde{x}_2} \leq \frac{\varphi(\tilde{x}_1)}{2 \varphi(x_2)} + \frac{\varphi(\tilde{x}_2)}{2 \varphi(x_2)} + \frac{\varphi(\tilde{x}_2)}{2 \varphi(x_3)} - 2$$

$$\text{y} \quad \frac{\partial G_2}{\partial \tilde{x}_3} = \frac{\varphi(\tilde{x}_3)}{2 \varphi(x_3)}$$

de manera que $\frac{\partial G_2}{\partial \tilde{x}_3} > 0$, $\frac{\partial G_2}{\partial \tilde{x}_2} + \frac{\partial G_2}{\partial \tilde{x}_3} \leq 0$ y además $\lim_{\tilde{x}_2 \uparrow \tilde{x}_3} G_2(\tilde{x}_2, \tilde{x}_3) < 0$ con lo cual $G_2 = 0$ define \tilde{x}_2 como una cierta función f_2 de \tilde{x}_3 con $f_2'(\tilde{x}_3) \leq 1$.

Procediendo reiteradamente hasta la función F_{n-2} , obtendríamos $\tilde{x}_{n-2} = f_{n-2}(\tilde{x}_{n-1})$ con $f_{n-2}'(\tilde{x}_{n-1}) \leq 1$.

La curva del espacio \mathbb{R}^{n-1} , $C = \{(\tilde{x}_i)_{i=1}^{n-1} \mid \tilde{x}_i = f_i(\tilde{x}_{i+1}) \ i = 1 \dots n-2\}$ es el conjunto de puntos en que se anulan simultáneamente F_1, F_2, \dots, F_{n-2} . Por último $F_{n-1}(f_{n-2}(\tilde{x}_{n-1}), \tilde{x}_{n-1}) = G_{n-1}(\tilde{x}_{n-1})$ cumple:

$$\frac{dG_{n-1}}{d\tilde{x}_{n-1}} = \frac{\partial F_{n-1}}{\partial \tilde{x}_{n-2}} f_{n-2}' + \frac{\partial F_{n-1}}{\partial \tilde{x}_{n-1}} \leq \frac{\varphi(\tilde{x}_{n-2})}{2 \varphi(x_{n-1})} + \frac{\varphi(\tilde{x}_{n-1})}{2 \varphi(x_{n-1})} +$$

$$+ \frac{\varphi(\tilde{x}_{n-1})}{2\varphi(x_n)} - 2 \leq 0$$

siendo la última desigualdad estricta, puesto que al menos uno de los menores o iguales empleados en los pasos anteriores será estricto. Ello garantiza que existe un único punto sobre la curva C que anule también a F_{n-1} .

Complementariamente podemos probar que para que se verifiquen las hipótesis del resultado anterior es suficiente que ϕ sea una distribución con momento de primer orden finito y función de densidad

$$\varphi(x) = \begin{cases} K e^{g(x)} & \text{si } x \in \dot{I} \\ 0 & \text{si } x \notin \dot{I} \end{cases}$$

siendo $g: \dot{I} \rightarrow \mathbb{R}$ con derivada segunda $g''(x) < 0 \quad \forall x \in \dot{I}$

En efecto, si para $a < b \in \dot{I}$, c es tal que $2\phi(c) = \phi(a) + \phi(b)$ será $\frac{\partial c}{\partial a} = \frac{\varphi(a)}{2\varphi(c)}$ y $\frac{\partial c}{\partial b} = \frac{\varphi(b)}{2\varphi(c)}$; la función $L(a, b) = \varphi(a) + \varphi(b) - 2\varphi(c)$ definida para $a < b \in \dot{I}$ cumple:

$$\frac{\partial L}{\partial a} = \varphi'(a) - 2\varphi'(c) \frac{\varphi(a)}{2\varphi(c)} = K e^{g(a)} [g'(a) - g'(c)] > 0$$

$$\frac{\partial L}{\partial b} = \varphi'(b) - 2\varphi'(c) \frac{\varphi(b)}{2\varphi(c)} = K e^{g(b)} [g'(b) - g'(c)] < 0$$

y como $\lim_{a \uparrow b} L(a, b) = 0$, será pues $L(a, b) < 0 \quad \forall a < b \in \dot{I}$

Por otra parte si c es tal que $2\phi(c) = \phi(b)$ será $\frac{dc}{db} = \frac{\varphi(b)}{2\varphi(c)}$; la función $M(b) = \varphi(b) - 2\varphi(c)$ para $b \in \dot{I}$ cumple:

$$\frac{dM}{db} = \varphi'(b) - 2\varphi'(c) \frac{\varphi(b)}{2\varphi(c)} = K e^{g(b)} [g'(b) - g'(c)] < 0$$

de manera que $M(b)$ es constantemente menor que su límite en el extremo inferior de I , el cual existe y es negativo.

Análogamente para c tal que $2\phi(c) = 1 + \phi(a)$ será $\frac{dc}{da} = \frac{\phi(a)}{2\phi(c)}$ y $N(a) = \phi(a) - 2\phi(c)$ para $a \in \dot{I}$ tendrá $\frac{dN}{da} = K e^{g(a)} [g'(a) - g'(c)] > 0$ con lo cual $N(a)$ es siempre menor que su límite en el extremo superior de I , que es negativo.

Entre otras, estas últimas condiciones suficientes son satisfechas por las distribuciones normal, gamma con $p > 1$ y beta con $p, q \geq 1$ (p ó $q \neq 1$). Comprobaciones inmediatas permiten añadir la distribución uniforme y la exponencial entre las distribuciones que cumplen las hipótesis del teorema anterior, cubriendo con ello la mayor parte de las distribuciones a posteriori usuales.

Pasando ahora al estudio correspondiente a la función $c_2(y) = y^2$, podemos enunciar:

Teorema 4: Sea ϕ una función de distribución continua, con momento de segundo orden finito, y que tenga derivada continua φ en el interior del menor intervalo cerrado I que contenga al soporte de ϕ . Si se verifica:

$$(b-c)\varphi(b) + (c-a)\varphi(a) \leq \phi(b) - \phi(a) \text{ siendo } c = \int_a^b y \phi(dy) / \phi(b) - \phi(a)$$

$$(b-c)\varphi(b) \leq \phi(b) \quad \text{siendo } c = \int_{-\infty}^b y \phi(dy) / \phi(b)$$

$$(c-a)\varphi(a) \leq 1 - \phi(a) \quad \text{siendo } c = \int_a^{\infty} y \phi(dy) / 1 - \phi(a)$$

con al menos una de las tres desigualdades estricta, entonces el sistema de ecuaciones:

$$x_i [\phi(\tilde{x}_i) - \phi(\tilde{x}_{i-1})] = \int_{\tilde{x}_{i-1}}^{\tilde{x}_i} y \phi(dy) \quad i = 1 \dots n, \quad 2\tilde{x}_i = x_i + x_{i+1}, \quad i = 1 \dots n-1$$

tiene una única solución con $x_1 < x_2 < \dots < x_n \in I$.

Demostración: Como en el caso anterior la existencia de solución queda garantizada por tener ϕ momento de segundo orden finito.

Los valores de x_i vienen dados explícitamente en función de \tilde{x}_{i-1} , \tilde{x}_i por las n primeras ecuaciones del sistema y se tiene:

$$\left\{ \begin{array}{l} \frac{\partial x_1}{\partial \tilde{x}_1} = (\tilde{x}_1 - x_1) \frac{\varphi(\tilde{x}_1)}{\phi(\tilde{x}_1)} \\ \frac{\partial x_i}{\partial \tilde{x}_{i-1}} = (x_i - \tilde{x}_{i-1}) \frac{\varphi(\tilde{x}_{i-1})}{\phi(\tilde{x}_i) - \phi(\tilde{x}_{i-1})}, \\ \frac{\partial x_i}{\partial \tilde{x}_i} = (\tilde{x}_i - x_i) \frac{\varphi(\tilde{x}_i)}{\phi(\tilde{x}_i) - \phi(\tilde{x}_{i-1})} \quad i = 2 \dots n-1 \\ \frac{\partial x_n}{\partial \tilde{x}_{n-1}} = (x_n - \tilde{x}_{n-1}) \frac{\varphi(\tilde{x}_{n-1})}{1 - \phi(\tilde{x}_{n-1})} \end{array} \right.$$

Para las $n-1$ funciones:

$$\left\{ \begin{array}{l} F_1(\tilde{x}_1, \tilde{x}_2) = x_1 + x_2 - 2\tilde{x}_1 \\ F_k(\tilde{x}_{k-1}, \tilde{x}_k, \tilde{x}_{k+1}) = x_k + x_{k+1} - 2\tilde{x}_k \quad k = 2 \dots n-2 \\ F_{n-1}(\tilde{x}_{n-2}, \tilde{x}_{n-1}) = x_{n-1} + x_n - 2\tilde{x}_n \end{array} \right.$$

será entonces:

$$\left\{ \begin{array}{l} \frac{\partial F_1}{\partial \tilde{x}_1} = (\tilde{x}_1 - x_1) \frac{\varphi(\tilde{x}_1)}{\phi(\tilde{x}_1)} + (x_2 - \tilde{x}_1) \frac{\varphi(\tilde{x}_1)}{\phi(\tilde{x}_2) - \phi(\tilde{x}_1)} - 2, \\ \frac{\partial F_1}{\partial \tilde{x}_2} = (\tilde{x}_2 - x_2) \frac{\varphi(\tilde{x}_2)}{\phi(\tilde{x}_2) - \phi(\tilde{x}_1)} \\ \frac{\partial F_k}{\partial \tilde{x}_{k-1}} = (x_k - \tilde{x}_{k-1}) \frac{\varphi(\tilde{x}_{k-1})}{\phi(\tilde{x}_k) - \phi(\tilde{x}_{k-1})}, \end{array} \right.$$

$$\left\{ \begin{aligned} \frac{\partial F_k}{\partial \bar{x}_k} &= (\bar{x}_k - x_k) \frac{\varphi(\bar{x}_k)}{\phi(\bar{x}_k) - \phi(\bar{x}_{k-1})} + (x_{k+1} - \bar{x}_k) \frac{\varphi(\bar{x}_k)}{\phi(\bar{x}_{k+1}) - \phi(\bar{x}_k)} - 2, \\ \frac{\partial F_k}{\partial \bar{x}_{k+1}} &= (\bar{x}_{k+1} - x_{k+1}) \frac{\varphi(\bar{x}_{k+1})}{\phi(\bar{x}_{k+1}) - \phi(\bar{x}_k)}, \\ \frac{\partial F_{n-1}}{\partial \bar{x}_{n-2}} &= (x_{n-1} - \bar{x}_{n-2}) \frac{\varphi(\bar{x}_{n-2})}{\phi(\bar{x}_{n-1}) - \phi(\bar{x}_{n-2})}, \\ \frac{\partial F_{n-1}}{\partial \bar{x}_{n-1}} &= (\bar{x}_{n-1} - x_{n-1}) \frac{\varphi(\bar{x}_{n-1})}{\phi(\bar{x}_{n-1}) - \phi(\bar{x}_{n-2})} + (x_n - \bar{x}_{n-1}) \frac{\varphi(\bar{x}_{n-1})}{1 - \phi(\bar{x}_{n-1})} - 2 \end{aligned} \right.$$

El razonamiento se completa entonces de manera idéntica al del teorema 3.

Veamos ahora que las hipótesis del teorema anterior se verifican si ϕ es una distribución con momento de segundo orden finito y función de densidad

$$\varphi(x) = \begin{cases} K e^{g(x)} & \text{si } x \in \dot{I} \\ 0 & \text{si } x \notin \dot{I} \end{cases}$$

siendo $g : \dot{I} \rightarrow \mathbb{R}$ con $g''(x) < 0 \quad \forall x \in \dot{I}$

En primer lugar si $c = \int_a^b y \varphi(y) dy / \phi(b) - \phi(a)$ con $a < b \in \dot{I}$ será:

$$\frac{\partial c}{\partial a} = (c - a) \frac{\varphi(a)}{\phi(b) - \phi(a)} \quad \text{y} \quad \frac{\partial c}{\partial b} = (b - c) \frac{\varphi(b)}{\phi(b) - \phi(a)}$$

La función $L(a, b) = (b - c) \varphi(b) + (c - a) \varphi(a) - \phi(b) + \phi(a)$ en $a < b \in \dot{I}$ tendrá:

$$\frac{\partial L}{\partial a} = \frac{c - a}{\phi(b) - \phi(a)} \{ \varphi'(a)[\phi(b) - \phi(a)] - \varphi(a)[\varphi(b) - \varphi(a)] \} =$$

$$= \frac{c - a}{\phi(b) - \phi(a)} \varphi(a) \{g'(a) [\phi(b) - \phi(a)] - \varphi(b) + \varphi(a)\}$$

y

$$\begin{aligned} \frac{\partial L}{\partial b} &= \frac{b - c}{\phi(b) - \phi(a)} \{\varphi'(b) [\phi(b) - \phi(a)] - \varphi(b) [\varphi(b) - \varphi(a)]\} = \\ &= \frac{b - c}{\phi(b) - \phi(a)} \varphi(b) \{g'(b) [\phi(b) - \phi(a)] - \varphi(b) + \varphi(a)\} \end{aligned}$$

Llamando:

$$\ell_1(a, b) = g'(a) [\phi(b) - \phi(a)] - \varphi(b) + \varphi(a) \quad y$$

$$\ell_2(a, b) = g'(b) [\phi(b) - \phi(a)] - \varphi(b) + \varphi(a)$$

será:

$$\frac{\partial \ell_1}{\partial a} = g''(a) [\phi(b) - \phi(a)] < 0, \quad \frac{\partial \ell_1}{\partial b} = \varphi(b) [g'(a) - g'(b)] > 0$$

y

$$\frac{\partial \ell_2}{\partial a} = \varphi(a) [g'(a) - g'(b)] > 0 \quad \frac{\partial \ell_2}{\partial b} = g''(b) [\phi(b) - \phi(a)] < 0$$

con lo cual $\ell_1(a, b) > 0$ y $\ell_2(a, b) < 0 \quad \forall a < b \in \dot{I}$ puesto que $\lim_{a \uparrow b} \ell_1(a, b) = \lim_{a \uparrow b} \ell_2(a, b) = 0$. Por tanto $\partial L / \partial a > 0$ y $\partial L / \partial b < 0$, así que $L(a, b) < 0 \quad \forall a < b \in \dot{I}$ puesto que $\lim_{a \uparrow b} L(a, b) = 0$.

En segundo lugar, si

$$c = \int_{-\infty}^b y \varphi(y) dy / \phi(b) \quad \text{será} \quad \frac{dc}{db} = (b - c) \frac{\varphi(b)}{\phi(b)}$$

la función $M(b) = (b - c) \varphi(b) - \phi(b)$ para $b \in \dot{I}$ tendrá:

$$\frac{dM}{db} = \frac{b-c}{\phi(b)} [\varphi'(b) \phi(b) - \varphi^2(b)] = \frac{b-c}{\phi(b)} \varphi(b) [g'(b) \phi(b) - \varphi(b)]$$

siendo esta derivada negativa puesto que:

$$g'(b) \phi(b) - \varphi(b) \leq \int_{I \cap (-\infty, b]} g'(x) \varphi(x) dx - \varphi(b) = - \lim_{x \downarrow i} \varphi(x)$$

donde i representa el extremo inferior de I . Por tanto $M(b) \leq \lim_{b \downarrow i} M(b)$

y dicho límite es cero; ello es evidente en el caso de ser i finito y se deduce fácilmente, en el caso i infinito, del hecho de que:

$$M(b) = \int_{-\infty}^b \phi(x) dx \frac{\varphi(b)}{\phi(b)} - \phi(b)$$

Por último, si

$$c = \int_a^{\infty} y \varphi(y) dy / 1 - \phi(a) \quad \text{será} \quad \frac{dc}{da} = (c - a) \frac{\varphi(a)}{1 - \phi(a)}$$

la función $N(a) = (c - a) \varphi(a) - 1 + \phi(a)$ para $a \in \dot{I}$ tendrá:

$$\begin{aligned} \frac{dN}{da} &= \frac{c-a}{1-\phi(a)} \{ \varphi^2(a) + \varphi'(a) [1 - \phi(a)] \} = \\ &= \frac{c-a}{1-\phi(a)} \varphi(a) \{ \varphi(a) + g'(a) [1 - \phi(a)] \} \end{aligned}$$

siendo esta derivada positiva puesto que:

$$g'(a) [1 - \phi(a)] + \varphi(a) \geq \int_{I \cap [a, \infty)} g'(x) \varphi(x) dx + \varphi(a) = \lim_{x \uparrow i'} \varphi(x)$$

siendo i' el extremo superior de I . Por tanto $N(a) \leq \lim_{a \uparrow i'} N(a)$ y dicho límite es cero como en el caso anterior.

De nuevo, según esto, el resultado del teorema 4 es válido para las distribuciones normal, gamma con $p > 1$ y beta con $p, q \geq 1$ (p ó $q \neq 1$): pudiéndose añadir a ellas, mediante comprobaciones directas, la uniforme y la exponencial.

En relación con los dos teoremas anteriores podemos hacer algunas observaciones.

1) Nótese que las demostraciones hacen uso en realidad, respectivamente, de las desigualdades:

$$\left. \begin{aligned} & \frac{\varphi(a) + \varphi(b)}{2 \varphi(m)} + \frac{\varphi(b) + \varphi(c)}{2 \varphi(n)} \leq 2 \quad \forall a < b < c \in \dot{I} \\ & \text{siendo } \varphi(a) + \varphi(b) = 2 \varphi(m) \text{ y } \varphi(b) + \varphi(c) = 2 \varphi(n) \\ & \frac{\varphi(b)}{2 \varphi(m)} + \frac{\varphi(b) + \varphi(c)}{2 \varphi(n)} \leq 2 \quad \forall b < c \in \dot{I} \\ & \text{siendo } \varphi(b) = 2 \varphi(m) \text{ y } \varphi(b) + \varphi(c) = 2 \varphi(n) \\ & \frac{\varphi(a) + \varphi(b)}{2 \varphi(m)} + \frac{\varphi(b)}{2 \varphi(n)} \leq 2 \quad \forall a < b \in \dot{I} \\ & \text{siendo } \varphi(a) + \varphi(b) = 2 \varphi(m) \text{ y } \varphi(b) = 2 \varphi(n) \end{aligned} \right\}$$

$$\left\{ \begin{aligned} & \frac{(m-a)\varphi(a) + (b-m)\varphi(b)}{\varphi(b) - \varphi(a)} + \frac{(n-b)\varphi(b) + (c-n)\varphi(c)}{\varphi(c) - \varphi(b)} \leq 2 \\ & \forall a < b < c \in \dot{I} \text{ siendo } m = \frac{\int_a^b y \varphi(dy)}{\varphi(b) - \varphi(a)} \text{ y } n = \frac{\int_b^c y \varphi(dy)}{\varphi(c) - \varphi(b)} \end{aligned} \right.$$

$$\frac{(b-m)\varphi(b)}{\phi(b)} + \frac{(n-b)\varphi(b) + (c-n)\varphi(c)}{\phi(c) - \phi(b)} \leq 2 \quad \forall b < c \in I$$

$$\text{siendo } m = \frac{\int_{-\infty}^b y \phi(dy)}{\phi(b)} \quad \text{y} \quad n = \frac{\int_b^c y \phi(dy)}{\phi(c) - \phi(b)}$$

$$\frac{(m-a)\varphi(a) + (b-m)\varphi(b)}{\phi(b) - \phi(a)} + \frac{(n-b)\varphi(a)}{1 - \phi(a)} \leq 2 \quad \forall a < b \in I$$

$$\text{siendo } m = \frac{\int_a^b y \phi(dy)}{\phi(b) - \phi(a)} \quad \text{y} \quad n = \frac{\int_a^{\infty} y \phi(dy)}{1 - \phi(a)}$$

Desigualdades más débiles que las contenidas en los enunciados, pero también más difíciles de comprobar si no se cumplen aquellas.

2) La misma técnica de demostración, puede aplicarse utilizando las relaciones $\bar{x}_i = \frac{x_i + x_{i+1}}{2}$ para expresar \bar{x}_i en función de x_i y x_{i+1} y anulando sucesivamente las funciones F_1, F_2, \dots, F_n , diferencia entre los miembros de las n primeras ecuaciones de los sistemas.

Habría que imponer para ello, en el contexto del teorema 3:

$$\left\{ \begin{array}{l} \varphi\left(\frac{a+b}{2}\right) \leq 2\varphi(a), \quad \varphi\left(\frac{a+b}{2}\right) \leq 2\varphi(b) \quad \forall a < b \in I \\ \varphi\left(\frac{a+b}{2}\right) + \varphi\left(\frac{b+c}{2}\right) \leq 2\varphi(b) \quad \forall a < b < c \in I \end{array} \right.$$

y en el contexto del teorema 4:

$$(b-a)\varphi\left(\frac{a+b}{2}\right) \leq 2\phi\left(\frac{a+b}{2}\right),$$

$$(b - a) \varphi\left(\frac{a + b}{2}\right) \leq 2 \left[1 - \phi\left(\frac{a + b}{2}\right) \right] \quad \forall a < b \in I$$

$$(b - a) \varphi\left(\frac{a + b}{2}\right) + (c - b) \varphi\left(\frac{b + c}{2}\right) \leq \\ \leq 2 \left[\phi\left(\frac{b + c}{2}\right) - \phi\left(\frac{a + b}{2}\right) \right] \quad \forall a < b < c \in I$$

sistemas de desigualdades menos frecuentes que los de los enunciados, pero que pueden ser útiles en algún caso.

En resumen, para las distribuciones a posteriori y funciones de pérdida usuales se puede garantizar la existencia de un único punto (x_1, x_2, \dots, x_n) que minimice $E[\min_{i=1 \dots n} c(\xi - x_i)]$.

La determinación de dicho punto a partir de las condiciones necesarias que ha de verificar, depende obviamente de la distribución considerada, de la función de pérdida concreta e incluso del número n de puntos en cuestión. Salvo en casos triviales sería necesario el uso de procedimientos numéricos.

REFERENCIAS

- 1.— Apostol, T. (1965): *Análisis matemático*. Reverté. Barcelona.
- 2.— Loève, M. (1963): *Probability Theory*. Van Nostrand. New York.
- 3.— Johnson, N.- Leone, F. (1977): *Statistics and experimental design*. Wiley. New York.