Pub. Mat. UAB
Nº 22 Nov. 1980
Actes VII JMHL

UNA MEDIDA DE CENTRALIZACION n-DIMENSIONAL PARA VARIABLES ALEATORIAS

Juan-Antonio Cuesta Albertos

Colegio Universitario de Burgos

Abstract

Let X be a real random variable of finite variance, μ the probability which induces and d the usual distance in $\mathbb{R}$. We proof that exists a n-uple ($n>2$) in wich is the minimum of $\int d^2[x:\{P_1,\dots,P_n\}]d\mu$.

1. Introducción.-

Sea X una variable aleatoria real de varianza finita (v. a. v. f.) μ la probabilidad que induce y d la distancia usual en $\mathbb{R}$. Existen problemas en los que una medida de centralización más adecuada que la media, sería la formada por la n-upla en que se alcanzase el mínimo de:

$$(1) \qquad D[X;(P_1,\dots,P_n)] = \int_{-\infty}^{\infty} d^2[x;\{P_1,\dots,P_n\}]d\mu(x)$$

Es claro que si $(P_1,\dots,P_n)$ es una n-upla en que se alcanza el mínimo de (1), la medida de dispersión apropiada para esta de centralización es el valor de (1) en este punto.

Es bien conocido que si n=1, $P_1=E(X)$ es el mínimo de (1). En [1] se demuestra que si n=2, este mínimo también existe, si bien puede no ser único.

En este trabajo, se obtiene que el mínimo de (1) se alcanza para cualquier natural n.

2. Teorema de existencia.-

Sea X una v. a. v. f.; μ la probabilidad que induce y A un conjunto de la σ-álgebra de Borel, $\mathcal{B}$, usual en $\mathbb{R}$. Definimos:

$$a[A] = \begin{cases} \dfrac{\int_A x d\mu}{\mu(A)} & ; \text{ si } \mu(A) \neq 0 \\ 0 & ; \text{ si } \mu(A) = 0 \end{cases}$$

Sea $h = (h^1, \ldots, h^{n-1}) \in \mathbb{R}^{n-1}$; colocamos sus coordenadas en orden creciente, es decir, $h^{j_1} \leq h^{j_2} \leq \ldots \leq h^{j_{n-1}}$. Donde $\{j_1, \ldots, j_{n-1}\}$ es una permutación de $\{1, \ldots, n-1\}$. Definimos los intervalos:

$$A_h^1 = (-\infty, h^{j_1}); \; A_h^i = [h^{j_{i-1}}, h^{j_i}), \; i=2,\ldots,(n-1) \text{ y } A_h^n = [h^{j_{n-1}}, \infty).$$

Consideremos el espacio probabilístico $(\mathbb{R}, \mathcal{B}, \mu)$ y para cada $h \in \mathbb{R}^{n-1}$ la variable definida en él por la expresión:

$$g_h = \sum a[A_h^i] \cdot \chi_{A_h^i}$$

χ es la variable indicador de conjunto.

De forma análoga, B_h^i, $i = 1,2,\ldots,n$; representa un intervalo cualquiera de extremos los mismos que A_h^i. Con estos conjuntos, podemos definir en $(\mathbb{R}, \mathcal{B}, \mu)$, para cada $h \in \mathbb{R}^{n-1}$, la variable aleatoria:

$$g'_h = \sum a[B_h^i] \cdot \chi_{B_h^i}$$

Sea Y una v. a. v. f. definida en $(\mathbb{R}, \mathcal{B}, \mu)$; sea I la variable identidad definida en este mismo espacio. Utilizaremos la notación $D(X;Y)$ para representar la cantidad $E[(I-Y)^2]$.

<u>Proposición 2.1</u> Sea X una v. a. v. f. Para todo $h \in \mathbb{R}^{n-1}$, se cumple que:

$$D(X, g_h) = E[X^2] - E[g_h^2]$$

Además:

$$\sup_h E[g_h'^2] = \sup_h E[g_h^2] \tag{2}$$

Alcanzándose el superior de la derecha de (2) si y solo si se alcanza el de la izquierda.

<u>Proposición 2.2</u> Sea X una v. a. v. f. Se cumple que:

$$\underset{(P_1,\ldots,P_n)}{\text{inferior}} D[X;(P_1,\ldots,P_n)] = \underset{h \in \mathbb{R}^{n-1}}{\text{inferior}} D(X, g_h)$$

Además, si en h se alcanza el $\sup_h E[g_h^2]$; en $(a[A_h^1], \ldots, a[A_h^n])$, se alcanza el inferior de (1).

<u>Proposición 2.3</u> Sea X una v. a. v. f. no igualmente distribuida que una variable que toma a lo sumo n-1 valores. Sea $h \in \mathbb{R}^{n-2}$, entonces, existe $h' \in \mathbb{R}^{n-1}$ tal que $D(X, g_h) > D(X, g_{h'})$.

Proposición 2.4 Sea $[h_m]$ una sucesión de elementos de $\mathbb{R}^{n-1}$ tal que para cada natural m, $h_m = (h_m^1,\ldots,h_m^{n-1})$, donde $h_m^i \leq h_m^{i+1}$. Utilizaremos la notación A_m^i para representar al conjunto $A_{h_m}^i$.

Sea X una v. a. v. f. y μ la probabilidad que induce en $\mathbb{R}$. Supongamos que existe $i_0 \in \{1,\ldots,n\}$ tal que $\mu[A_m^{i_0}] \longrightarrow 0$.

Entonces existe $J \subset \{1,\ldots,n-1\} - \{i_0\}$ y una subsucesión (a la que representaremos con la misma notación que a la inicial) tales que si llamamos h'_m al vector $(h_m^j)_{j\in J}$; se cumple que:

$$D(X,g_{h_m}) - D(X,g_{h'_m}) \longrightarrow 0.$$

DEMOSTRACION.- Puesto que $\sum\mu(A_m^i) = 1$, existe al menos un i tal que $\mu(A_m^i)$ no converge a cero. Supongamos que existen $i>i_0$ y $j<i_0$ con ésta propiedad (en el resto de los casos la demostración es similar). Sea i_1 el menor de los números mayores que i_0 y j_1 el mayor de los menores que la cumplen.

Evidentemente, existen $r>0$ y una subsucesión (que representamos con la misma notación que a la sucesión original) tales que:

$$\mu[A_m^{i_1}] > r \text{ y } \mu[A_m^{j_1}] > r$$

Tomando $J = \{1,\ldots,j_1,i_1,\ldots,n\}$ y aplicando la desigualdad de Schwartz, se demuestra que esta subsucesión cumple la proposición.

Proposición 2.5 Sea X una v. a. v. f. y n un número natural cualquiera. Se cumple que existe $h \in \mathbb{R}^{n-1}$, tal que:

$$E[g_h^2] = \sup_h E[g_h^2]$$

DEMOSTRACION.- Puesto que si X es igual en distribución que una v. a. que toma a lo sumo n valores el problema es trivial, supondremos que no lo es.

En [1], se demuestra este mismo teorema para n=2. Supondremos por tanto que $n>2$ y razonaremos por inducción, es decir, le supondremos cierto para todo natural menor que n.

De la proposición 2.1 se deduce que el conjunto $[E(g_h^2)]_{h\in\mathbb{R}^{n-1}}$ está acotado. Sea M su extremos superior y $[h_m]$ una sucesión contenida en $\mathbb{R}^{n-1}$ tal que:

$$(3) \qquad E[g_{h_m}^2] \longrightarrow M.$$

Sea $h_m = (h_m^1,\ldots,h_m^{n-1})$. Podemos suponer sin pérdida de generalidad que para todo m y todo i: $h_m^i \leq h_m^{i+1}$.

Veamos que la sucesión $[h_m^1]$ está acotada.

Supongamos que no lo está inferiormente. Podemos extraer una sub sucesión (que también representamos como $[h_m^1]$) convergiendo a $-\infty$.

A la subsucesión de $[h_m]$ asociada con ésta, la podemos aplicar la proposición 2.4 ya que $\mu(A_m^1) \longrightarrow 0$. Ello unido a la hipótesis de inducción y a la proposición 2.3, implica la imposibilidad de que se cumpla (3) Por tanto, la sucesión original, $[h_m^1]$ está acotada inferiormente.

Análogamente, demostraríamos que también lo está superiormente.

Por consiguiente, podemos extraer una subsucesión monótona y convergente. Representaremos con la misma notación que a las originales, tanto a la subsucesión obtenida como a la de $[h_m]$ asociada con ella, ya que en lo sucesivo, únicamente nos referiremos a ellas.

Aplicando el mismo procedimiento a la sucesión obtenida, tendríamos una subsucesión de $[h_m^2]$ monótona y convergente.

Reiterando el procedimiento, obtendríamos, finalmente, una subsucesión $[h_m]$ tal que para todo i, la sucesión $[h_m^i]$ es monótona y convergente. En el resto de la demostración, trabajaremos únicamente con esta subsucesión.

Sea $B_0^i = \lim_m A_m^i$, que existe para todo i, por ser monótonas las sucesiones que forman los extremos de los intervalos A_m^i.

Definiendo:

$g' = \sum a(B_0^i)\cdot\chi_{B_0^i}$; utilizando el teorema de la convergencia mayorada, de (3) y la proposición 2.1, se deduce ésta.

<u>Teorema 2.6</u> Sea X una v. a. v. f. y n un número natural cualquiera. Existe $(P_1,\ldots,P_n)$ en el que se alcanza el mínimo de (1).

DEMOSTRACION.- Se deduce de las proposiciones 2.5 y 2.2

Para finalizar, queremos indicar que, en general, el mínimo de (1) no se alcanza en un único punto.

BIBLIOGRAFIA.-

[1] J. A. Cuesta - Una mediad de centralización para variables aleatorias

[2] M. Loève - Probability Theory - D. Van Nostrand Co., 1960