TRABAJOS DE ESTADISTICA Y
DE INVESTIGACION OPERATIVA
Vol. 34. Núm. 2, 1983, pp. 40 a 51

CARACTERIZACION AXIOMATICA PARA LA VARIANZA (*)

María Angeles Gil Alvarez
Departamento de Matemáticas
Facultad de Ciencias (Oviedo)

ABSTRACT

In the article ([7]), M. Martín propounds two axiomatic characterizations for the variance, suggesting the possibility of characterizing it in a more intuitive way as an uncertainty measure which takes into account the support of probability, in addition to its value.

The present paper is dedicated to the establishment of a characterization in the preceeding sense, following the path already covered by D.K. Faddejew when he characterizes Shannon's entropy and the one propounded in ([3]) for the measure defined in ([2]).

We want to outline that a study of variance as a "measure of unquietness" has been done in ([4]) where we have analysed its basic properties along with the ones relative to the mentioned measure at the end of the former paragraph.

RESUMEN

En el artículo ([7]) M. Martín propone dos caracterizaciones axiomáticas para la varianza sugiriendo la posibilidad de caracterizarla de forma más intuitiva como una medida de incertidumbre que tenga en cuenta el soporte de la probabilidad, además del valor de ésta.

El presente trabajo está dedicado a establecer una caracterización en tal sentido, siguiendo la linea de la axiomática de D.K. Faddejew para la entropía de Shannon y de la axiomática propuesta en ([3]) para la medida definida en ([2]).

Queremos reseñar que ya se ha realizado un estudio sobre la varianza como "medida de inquietud" en ([4]), donde se analizan sus propiedades básicas paralelamente a las relativas a la medida mencionada al final del párrafo anterior.

(*) Recibido Marzo, 1981

Palabras y frases clave: separation between the utilities, uncertainty measure, unquietness measure.

Separación entre las utilidades, medida de incertidumbre, medida de inquietud.

1. Introducción

Supongamos una experiencia constituída por un conjunto de alternativas con probabilidades de ocurrir conocidas y que para cierto decisor representan determinados intereses que él es capaz de evaluar subjetivamente de forma numérica.

En tal situación la inquietud surge para el decisor cuando, existiendo duda o incertidumbre sobre la alternativa que va a presentarse antes de realizar la experiencia, los intereses que las diversas alternativas representan para el decisor, y que numéricamente se traducen en valores que denominaremos "utilidades", son distintos.

Esta distinción de intereses se lleva a cabo en la varianza mediante la consideración de las diferencias entre los diversos valores otorgados a las alternativas, por lo que la varianza será más efectiva como medida de inquietud si el decisor establece su evaluación de intereses atendiendo a tales diferencias, sin que ello obste para ser aplicable en cualquier otro caso.

2. Caracterización axiomática

Sea A una experiencia dependiente del azar y sean $A_1, A_2, ..., A_n$ sus posibles alternativas, cuyas probabilidades respectivas son $p_1, p_2, ..., p_n$ ($p_i \geqslant 0$ para $i = 1, 2, ..., n$ y $\sum_{i=1}^{n} p_i = 1$).

Supongamos que cierto individuo (decisor), que ha asignado a $A_1, A_2, ..., A_n$ unas valoraciones reales $u_1, u_2, ..., u_n$ para su comparación respecto a la importancia que para él revisten, quiere evaluar la inquietud que le afecta antes de la realización de A, o inquietud que A reporta al decisor, con el fin de poder confrontar posteriormente dos experiencias distintas a través de tal característica.

Formulamos entonces el siguiente teorema:

TEOREMA. – Sea $HU^*(1;u)$, $HU^*(p_1,p_2;\ u_1,u_2)$, ..., $HU^*(p_1, ..., p_n;u_1, ..., u_n)$, ... una sucesión de funciones asociadas a experiencias dotadas de una distribución de probabilidad y un esquema finito de utilidades, formadas por 1, 2, ..., n, ... sucesos elementales (o alternativas), respectivamente. Si se satisface que:

I) *(Axioma de continuidad)*

$HU^*(p,\ 1-p;u_1,u_2)$ es una función continua respecto a p en el intervalo $[0,\ 1]$, cualesquiera que sean u_1 y u_2 reales.

II) *(Axioma de simetría)*

$HU^*(p_1,p_2, ..., p_n;\ u_1,u_2, ..., u_n)$ es una función simétrica respecto a los pares (p_i, u_i), para $i = 1, 2, ..., n$.

III) *(Axioma de ramificación)*

$$HU^*(p_1,p_2,p_3, ..., p_n;u_1,u_2,u_3, ..., u_n)=$$

$$=HU^*\left(p_1+p_2,p_3, ..., p_n;\frac{p_1u_1+p_2u_2}{p_1+p_2},u_3, ..., u_n\right)+$$

$$+(p_1+p_2)HU^*\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2};u_1,u_2\right)$$

IV) *(Axioma de separación entre las utilidades)*

$$HU^*(1/2,\ 1/2;u_1,u_2)=K\,(u_1-u_2)^2$$

siendo K una constante positiva.

V) *(Axioma de normalización)*

$$HU^*(1/2,\ 1/2;\ 1,-1)=1$$

Entonces:

$$HU^*(p_1,p_2, ..., p_n;u_1,u_2, ..., u_n)=\sum_{i=1}^{n}p_i\,[u_i-E(u)]^2$$

siendo:

$$E(u) = \sum_{i=1}^{n} p_i u_i$$

La interpretación de los dos primeros axiomas es evidente e inmediata. Vamos ahora, antes de introducirnos en la demostración del resultado que acabamos de enunciar, a realizar una traducción intuitiva del tercer y cuarto axioma:

Axioma III: Supongamos la experiencia $A = \{A_1, A_2, ..., A_n\}$. La inquietud que A reporta al decisor vendrá dada por:

$$HU^*(p_1, p_2, p_3, ..., p_n; u_1, u_2, u_3, ..., u_n)$$

Si se reunen A_1 y A_2 en una única alternativa A_0, el valor $HU^*(p_1 + p_2, p_3, ..., p_n; \frac{p_1 u_1 + p_2 u_2}{p_1 + p_2}, u_3, ..., u_n)$ corresponde a la inquietud que la experiencia $B = \{A_0, A_3, ..., A_n\}$ reporta al decisor siempre que éste atribuya a que ocurra A_0 la utilidad esperada de la experiencia $C = \{A_1, A_2\}$.

Es obvio que la inquietud $HU^*(A)$ debe superar a la $HU^*(B)$, puesto que la realización de B hace desaparecer toda la inquietud sobre B pero no toda la existente sobre A puesto que en dicha realización ha podido presentarse la alternativa A_0. Como la experiencia C reporta una inquietud igual a $HU^*\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}; u_1, u_2\right)$ y A_0 aparece con probabilidad $(p_1 + p_2)$ es natural suponer que $HU^*(A)$ excede a $HU^*(B)$ en $(p_1 + p_2) HU^*(C)$.

El hecho de que la contribución que $HU^*(C)$ tiene en $HU^*(A)$ esté ponderada por $(p_1 + p_2)$ indica que al medir la inquietud de una experiencia realmente se está considerando la inquietud esperada.

Axioma IV: Si una experiencia consta de dos alternativas equiprobables parece plausible admitir que la inquietud únicamente debe venir motivada por la "separación" de intereses o utilidades entre esas dos alternativas y ser una función creciente con esa separación entre intereses. A este fin, y buscando la función que simultáneamente resulte más operativa y significativa hemos tomado la expresión $K\,(u_1 - u_2)^2$, donde u_1 y u_2 son las utilidades de las alternativas y K es una constante positiva.

Una vez efectuadas las interpretaciones que justifican de manera intuitiva la elección de los axiomas considerados, vamos a comprobar que efectivamente en su conjunto caracterizan a la varianza de la variable aleatoria que toma los valores $u_1, u_2, ..., u_n$ con probabilidades $p_1, p_2, ..., p_n$, respectivamente.

En la demostración que sigue se enuncian en primer lugar siete lemas que constituyen una traslación de los que aparecen en la demostración dada por A. Renyi para caracterizar la entropía de C.E. Shannon con los axiomas dados por D.K. Faddejew. Su verificación no se desarrolla en el presente trabajo aunque los mismos lemas se prueban detalladamente en ([3]).

Lema 1.– Se verifica que:

$$HU^*(1, 0; u_1, u_2) = 0 \qquad (2.1)$$

cualesquiera que sean u_1 y u_2 reales.

Lema 2.– Se verifica que:

$$HU^*(p_1, p_2, ..., p_{n-1}, 0; u_1, u_2, ..., u_{n-1}, u_n) =$$
$$= HU^*(p_1, p_2, ..., p_{n-1}; u_1, u_2, ..., u_{n-1}) \qquad (2.2)$$

cualesquiera que sean $u_1, u_2, ..., u_{n-1}, u_n$ reales y $p_1, p_2, ..., p_{n-1}$ no negativos con $\sum_{i=1}^{n-1} p_i = 1$.

Lema 3.– Se verifica que:

$$HU^*(p_1, ..., p_k, p_{k+1}, ..., p_n; u_1, ..., u_k, u_{k+1}, ..., u_n) =$$

$$= HU^*(p_1 + \cdots + p_k, p_{k+1}, ..., p_n; \frac{p_1u_1 + \cdots + p_ku_k}{p_1 + \cdots + p_k}, u_{k+1}, ..., u_n) +$$

$$+(p_1+\cdots+p_k)\,HU^*\left(\frac{p_1}{p_1+\cdots+p_k},\ldots\frac{p_2}{p_1+\cdots+p_k};u_1,\ldots u_k\right)\quad(2.3)$$

para $k=1,2,\ldots,n$ con $n\in\mathbb{N}$ y cualesquiera que sean $u_1,u_2,\ldots,u_n$ reales y $p_1,p_2,\ldots,p_n$ no negativos con $\sum_{i=1}^{n} p_i=1$.

Lema 4.— Se verifica que:

$$HU^*(p_{11},\ldots p_{1m_1},\ldots p_{n1},\ldots p_{nm_n};u_{11},\ldots u_{1m_1},\ldots u_{n1},\ldots u_{nm_n})=$$

$$=HU^*(p'_1,\ldots p'_n;u'_1,\ldots u'_n)+$$

$$+\sum_{i=1}^{n} p'_i\,HU^*\left(\frac{p_{i_1}}{p'_i},\ldots\frac{p_{i m_i}}{p'_i};u_{i_1},\ldots u_{i m_i}\right)\quad(2.4)$$

siendo:

$$p'_i=\sum_{j=1}^{m_i} p_{ij}\quad\text{y}\quad u'_i=\sum_{j=1}^{m_i}\frac{p_{ij}\,u_{ij}}{p'_i}$$

cualesquiera que sean u_{ij} reales y p_{ij} no negativos con $\sum_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}=1$.

Lema 5.— Si denotamos por $L^*(n;u)=HU^*(1/n,\overset{n)}{\ldots}1/n;u,\ldots u)$, se verifica que:

$$L^*(n\cdot m;u)=L^*(n;u)+L^*(m;u)\quad(2.5)$$

$\forall\, m,n\in\mathbb{N}$ y cualquiera que sea u real.

Lema 6.— Se verifica que:

$$\lim_{n\to\infty}\frac{L^*(n;u)}{n}=0\quad(2.6.1)$$

y:

$$\lim_{n \to \infty} [L^*(n;u) - L^*(n-1;u)] = 0 \qquad (2.6.2)$$

cualquiera que sea u real.

Lema 7.– Si $L^*(n;u)$ es una función real que satisface las condiciones:

$$L^*(n \cdot m;u) = L^*(n;u) + L^*(m;u)$$

siendo n, $m \in \mathbb{N}$ tales que n y m son primos entre sí, y:

$$\lim_{n \to \infty} [L^*(n;u) - L^*(n-1;u)] = 0$$

se verifica que:

$$L^*(n;u) = \lambda \log n \qquad (2.7)$$

cualquiera que sea u real y λ constante real.

Como consecuencia de la serie de lemas precedente llegamos a que $\forall\, n \in \mathbb{N}$, $n \geqslant 1$ y cualquiera que sea u real:

$$L^*(n;u) = \lambda \log n$$

Y haciendo intervenir el axioma IV para $u_1 = u_2 = u$:

$$L^*(2;u) = K(u-u)^2 = 0$$

de donde se obtiene que $\lambda = 0$ y, por tanto, $L^*(n;u) = 0$.

Antes de proseguir con la demostración del teorema efectuamos un inciso para apuntar un resultado que precisaremos en lo sucesivo:

Si denotamos por $A = \{y \in \mathrm{IR}$ / existen r, $k \in \mathbb{N}$ con $r \leqslant 2^k, y = r/2^k\}$, se comprueba fácilmente que este conjunto A es denso en el intervalo $[0, 1]$ pues fijado $\epsilon > 0$ arbitrario y cualquiera que sea $p \in (0, 1)$ (ya que $p = 0$ ó 1 pertenecen obviamente al conjunto A), tomando $k \in \mathbb{N}$ con $1/2^k < \epsilon$ (donde la existencia del valor k está garantizada por no estar IN acotado superiormente), entonces forzosamente debe existir un $r \in$ IN tal que $p - \epsilon < r/2^k \leqslant p$ lo que indica que $|p - \frac{r}{2^k}| < \epsilon$.

Para demostrar en una segunda etapa que el teorema es cierto para $n = 2$ bastará, por el resultado que acabamos de exponer, probarlo para $p = r/2^k$ con r, $k \in$ IN. Con tal motivo establecemos un lema previo:

Lema 8. – Cualquiera que sea $k \in$ IN, se verifica que:

$$HU^*(1/2^k, 1/2^k, \dots 1/2^k; u_1, u_2, \dots u_{2^k}) =$$

$$= \sum_{i=1}^{2^k} \frac{1}{2^k} [u_i - E(u)]^2 \tag{2.8}$$

siendo $E(u) = \sum_{i=1}^{2^k} \frac{u_i}{2^k}$.

En efecto:

Para $k = 1$, se tiene que:

$$HU^*(1/2, 1/2; u_1, u_2) = K (u_1 - u_2)^2$$

y, por el axioma V:

$$K [1 + (-1)]^2 = 1$$

de donde:

$$HU^*(1/2, 1/2; u_1, u_2) = \frac{(u_1 - u_2)^2}{4} = \frac{1}{2} \left[\frac{u_1 - u_2}{2} \right]^2 +$$

$$+ \frac{1}{2} \left[\frac{u_2 - u_1}{2} \right]^2 = \frac{1}{2} \left[u_1 - \frac{u_1 + u_2}{2} \right]^2 + \frac{1}{2} \left[u_2 - \frac{u_1 + u_2}{2} \right]^2$$

Razonando por inducción, supongamos que (2.8) es cierta para $k = m$ y vamos a ver que también lo es para $k = m + 1$. Por el lema 4 es:

$$HU^*\left(\frac{1}{2^{m+1}}, \dots \frac{1}{2^{m+1}}; u_1, \dots u_{2^{m+1}}\right) =$$

$$= HU^*\left(\frac{1}{2^m}, \frac{1}{2^m}, \dots \frac{1}{2^m}; \frac{u_1+u_2}{2}, \frac{u_3+u_4}{2}, \dots \frac{u_{2^{m+1}-1}+u_{2^{m+1}}}{2}\right) +$$

$$+ \frac{1}{2^m} HU^*(1/2, 1/2; u_1, u_2) + \frac{1}{2^m} HU^*(1/2, 1/2; u_3, u_4) + \cdots +$$

$$+ \frac{1}{2^m} HU^*(1/2, 1/2; u_{2^{m+1}-1}, u_{2^{m+1}}) =$$

$$= \frac{1}{2^m}\left[\frac{u_1+u_2}{2} - \frac{\sum_{i=1}^{2^{m+1}} u_i}{2^{m+1}}\right]^2_{2^{m+1}} + \frac{1}{2^m}\left[\frac{u_3+u_4}{2} - \frac{\sum_{i=1}^{2^{m+1}} u_i}{2^{m+1}}\right]^2 + \cdots +$$

$$+ \frac{1}{2^m}\left[\frac{u_{2^{m+1}-1}+u_{2^{m+1}}}{2} - \frac{\sum_{i=1}^{2^{m+1}} u_i}{2^{m+1}}\right] + \frac{1}{2^m}\frac{1}{2}\left[u_1 - \frac{u_1+u_2}{2}\right]^2 +$$

$$+ \frac{1}{2^m}\frac{1}{2}\left[u_2 - \frac{u_1+u_2}{2}\right]^2 + \frac{1}{2^m}\frac{1}{2}\left[u_3 - \frac{u_3+u_4}{2}\right]^2 +$$

$$+ \frac{1}{2^m}\frac{1}{2}\left[u_4 - \frac{u_3+u_4}{2}\right]^2 + \cdots + \frac{1}{2^m}\frac{1}{2}\left[u_{2^{m+1}-1} -\right.$$

$$\left. - \frac{u_{2^{m+1}-1}+u_{2^{m+1}}}{2}\right]^2 + \frac{1}{2^m}\frac{1}{2}\left[u_{2^{m+1}} - \frac{u_{2^{m+1}-1}+u_{2^{m+1}}}{2}\right]^2 =$$

$$= u_1^2/2^{m+1} + u_2^2/2^{m+1} + u_3^2/2^{m+1} + u_4^2/2^{m+1} + \cdots +$$

$$+ (u_{2^{m+1}-1})^2/2^{m+1} + (u_{2^{m+1}})^2/2^{m+1} - \left(\sum_{i=1}^{2^{m+1}} u_i/2^{m+1}\right)^2 =$$

$$= \sum_{j=1}^{2^{m+1}} \frac{1}{2^{m+1}}[u_j - \sum_{i=1}^{2^{m+1}} u_i/2^{m+1}]^2$$

con lo que el lema queda probado.

Nos hallamos en disposición de probar el teorema para experiencias con dos alternativas y en las que la probabilidad p de una de ellas puede expresarse como $r/2^k$ con $r,\ k \in \mathbb{N}$. Entonces:

$$HU^*(r/2^k,\ (2^k - r)/2^k;\ u_1, u_2) =$$

$$= HU^*(1/2^k, \overset{r)}{\ldots}\ 1/2^k, \overset{2^k-r)}{\ldots}\ 1/2^k;\ u_1, \overset{r)}{\ldots}\ u_1, u_2, \overset{2^k-r)}{\ldots}\ u_2) +$$

$$- \frac{r}{2^k} HU^*(1/r, \overset{r)}{\ldots}\ 1/r;\ u_1, \ldots u_1) -$$

$$- \frac{2^k - r}{2^k} HU^* \ \frac{1}{2^k - r}, \overset{2^k-r)}{\ldots}\ \frac{1}{2^k - r};\ u_2, \ldots u_2 \ =$$

$$= r \frac{1}{2^k} \left[u_1 - \frac{r u_1 + (2^k - r) u_2}{2^k} \right]^2 +$$

$$+ (2^k - r) \frac{1}{2^k} \left[u_2 - \frac{r u_1 + (2^k - r) u_2}{2^k} \right]^2$$

lo que por la continuidad postulada en el axioma I implica que:

$$HU^*(p_1, p_2;\ u_1, u_2) = p_1\, [u_1 - (p_1 u_1 + p_2 u_2)]^2 +$$

$$+ p_2\, [u_2 - (p_1 u_1 + p_2 u_2)]^2$$

Para ver finalmente que el conjunto de axiomas dado caracteriza a la varianza, para cualquier $n \in \mathbb{N}$, razonaremos por inducción. Supongamos que es cierto para $n = m$ y vamos a comprobar que lo es también para $n = m + 1$. Así:

$$HU^*(p_1, p_2, p_3, \ldots p_{m+1};\ u_1, u_2, u_3, \ldots u_{m+1}) =$$

$$= HU^*(p_1 + p_2, p_3, \ldots p_{m+1};\ \frac{p_1 u_1 + p_2 u_2}{p_1 + p_2}, u_3, \ldots u_{m+1}) +$$

$$+ (p_1 + p_2) HU^* \left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2} ; u_1, u_2 \right) =$$

$$= (p_1 + p_2) \left[\frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} - \sum_{i=1}^{m+1} p_i u_i \right]^2 + p_3 [u_3 - \sum_{i=1}^{m+1} p_i u_i]^2 +$$

$$+ \cdots + p_{m+1} [u_{m+1} - \sum_{i=1}^{m+1} p_i u_i]^2 + p_1 \left[u_1 - \frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} \right]^2 +$$

$$+ p_2 \left[u_2 - \frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} \right]^2 = (p_1 + p_2) \left[\frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} \right]^2 +$$

$$+ (p_1 + p_2) (\sum_{i=1}^{m+1} p_i u_i)^2 - 2 (p_1 u_1 + p_2 u_2) \sum_{i=1}^{m+1} p_i u_i + p_1 u_1^2 +$$

$$+ p_1 \left[\frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} \right]^2 - 2 p_1 u_1 \frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} + p_2 u_2^2 +$$

$$+ p_2 \left[\frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} \right]^2 - 2 p_2 u_2 \frac{p_1 u_1 + p_2 u_2}{p_1 + p_2} +$$

$$+ \sum_{j=3}^{m+1} p_j [u_j - \sum_{i=1}^{m+1} p_i u_i]^2 = \sum_{j=1}^{m+1} p_j [u_j - \sum_{i=1}^{m+1} p_i u_i]^2$$

3. Conclusiones finales

Es conveniente reseñar que la presente caracterización propone indirectamente un método para construir medidas de inquietud sin más que tomar como función creciente con la "separación" entre las utilidades cualquier otra a partir de la que sea posible la obtención de una medida.

Así mismo, puede modificarse también el axioma de ramificación considerando que la contribución de $HU^*(C)$ en $HU^*(A)$ viene afectada por un coeficiente dependiente de p_1, p_2, u_1, u_2, distinto de $(p_1 + p_2)$.

REFERENCIAS

[1] D.K. FADDEJEW, Zum Begriff der entropie eines endlichen Warhrschernlichkeits schema; Arbeiten zur Informations-theorie. V.E.B. Berlín, 1957.

[2] M.A. GIL, Incertidumbre y utilidad. Tesis. Oviedo, 1979.

[3] M.A. GIL, Caracterización axiomática de una medida para la incertidumbre correspondiente a las utilidades. *XII Reunión Nacional de la Sociedad de Estadística, Investigación Operativa e Informática*. Jaca, 1980.

[4] M.A. GIL, On two unquietness measures for the finite experiences. *(Pendiente de publicación).*

[5] S. GUIASU and R. THEODORESCU, La théorie matématique de l'information. *Dunod*. París, 1968.

[6] S. GUIASU and R. THEODORESCU, Incertitude et information. *Les presses de l'Université Laval*. Québec, 1971.

[7] M. MARTIN, Caracterización de la varianza. *Revista Trabajos de Estadística y de Investigación Operativa. Volumen XXVIII, Cuaderno 2 y 3*. Madrid, 1977.

[8] A. RENYI, Calcul des probabilités avec un appendice sur la théorie de l'information. *Dunod*. Paris, 1966.

[9] C.E. SHANNON, A mathematical theory of communication. *Bell. System. Techn. J-27*. 1948.

[10] A.M. YAGLOM and I.M. YAGLOM, Probabilité et information. *Dunod*. Paris, 1969.