Aleatoriedad y vaguedad en sucesos difusos

POR MARIA TERESA LOPEZ

Recibido: 10 de abril de 1985

Presentado por el académico numerario D. Francisco Azorín

Resumen

Dentro del contexto difuso tienen gran relevancia las medidas de incertidumbre no probabilística (o medidas de difusión); pero es interesante considerar medidas de incertidumbre probabilística para sucesos difusos, medidas que de alguna forma generalicen la conocida entropía de Shannon, para poder utilizar la mayor información disponible en los posibles criterios de comparación de experimentos.

1. INTRODUCCION

Un concepto básico de la estadística clásica es el de incertidumbre, que está asociado a la noción de experimento aleatorio, de ahí la gran importancia que tienen las medidas de incertidumbre para saber la información que se gana (o la incertidumbre que se pierde) al realizar un experimento aleatorio.

Dentro de estas medidas la más conocida, por sus propiedades y fácil manejo, es la medida de entropía de Shannon que fue punto de partida de numerosos trabajos. Posteriormente se pensó que además de las probabilidades asociadas al experimento aleatorio se debería considerar, en algunos casos, la "importancia" de los posibles resultados del experimento aleatorio; surgen entonces las medidas de incertidumbre útil (Belis y Guiasu 1968, P. Gil 1975, A. Gil 1979).

Por otro lado, a partir de 1965, se desarrolla la teoría de conjuntos difusos para tratar problemas en los que aparece una imprecisión o vaguedad debido a la falta de fronteras definidas en los conjuntos que se utilizan, y que en muchos casos se acercan más a los fenómenos reales que los conjuntos clásicos. La vaguedad o indeterminación de estos conjuntos se refleja en la función de pertenencia, que indica el "grado" con que un elemento verifica la propiedad que caracteriza al conjunto difuso; numerosos autores (De Luca y Termini 1972, 1978, Knofmacher 1975, Kaufmann 1975, Loo 1977, Emptoz 1981, etc.) estudian este tipo de indeterminación y proponen medirla a través de lo que llaman medidas de entropía no probabilística o medidas de difusión (o borrosidad, como diría el Prof. Azorín).

Estas medidas tienen un parecido formal con la entropía de Shannon y otras entropías, al menos en un principio, y por eso se llaman así, pero su significado es completamente distinto.

Nosotros pensamos en enlazar estos dos conceptos: *aleatoriedad* y *vaguedad* al considerar conjuntos difusos, es decir trabajaremos con sucesos difusos (conjuntos difusos definidos en un conjunto de posibles resultados de un experimento aleatorio) porque creemos que no es conveniente tratar este problema sólo como difuso o sólo como probabilístico.

2. INCERTIDUMBRE PROBABILISTICA DE UN SUCESO DIFUSO

En este apartado se busca y estudia una medida de incertidumbre probabilística para sucesos difusos, es decir como Zadeh (1968) supondremos que consideramos un conjunto difuso A definido en un referencial X que es el conjunto de posibles resultados de un experimento aleatorio, y por tanto en X hay definida una distribución de probabilidad que nos permite definir probabilidades de sucesos difusos según Zadeh:

Definición 1.– La probabilidad de un suceso difuso es la esperanza de su función de pertenencia

$$P(A) = \sum_{x_i \in X} p(x_i)\, \mu_A(x_i)$$

(esta definición será la que se utilice en todo el trabajo).

Por otro lado, como la medida de incertidumbre probabilística más conocida es la de Shannon trataremos de buscar una generalización de ella; en esta línea Zadeh propone como medida de entropía de un suceso difuso A el valor:

$$H^P(A) = - \sum_{x_i \in X} p(x_i)\, \mu_A(x_i) \log p(x_i)$$

aunque él mismo observa que esta medida no coincide con la de Shannon cuando A es un conjunto clásico, por otro lado otro inconveniente que tiene es que, si consideramos conjuntos completamente difusos (aquellos que tienen función de pertenencia constante $\mu_A(x) = \mu$) se verificaría que $H^P(A) = \mu\, \boldsymbol{H}(X)$ (siendo $\boldsymbol{H}(X)$ la entropía de Shannon del referencial) lo que significaría que $H^P(A)$ disminuye al disminuir la función de pertenencia, aunque al ser esta constante la incertidumbre relativa a las probabilidades debe ser la misma.

Basándonos en esta "medida" de entropía y en la de Shannon para subconjuntos de X podría pensarse en asociar a cada elemento $x_i \in X$ un peso absoluto $p(x_i)\, \mu_A(x_i)$ y calcular la entropía de A como el valor

$$H(A) = - \sum_{x_i \in X} \frac{p(x_i)\, \mu_A(x_i)}{P(A)} \log p(x_i)\, \mu_A(x_i)$$

con lo que se tendría en el caso de que A fuese un conjunto clásico un valor igual a la entropía de Shannon de A, pero que también tiene inconvenientes como aumentar al disminuir la función de pertenencia, por lo que todo conjunto difuso tendría una entropía mayor o igual que la entropía de Shannon del referencial, resultado que no parece convincente.

Como paso siguiente se piensa en definir la incertidumbre probabilística de un suceso difuso A como una función de las incertidumbres puntuales de los elementos del referencial $(-\log p(x_i))$, como ya ocurría con la entropía de Shannon para conjuntos ordinarios, y donde la función de pertenencia aparece como un elemento de ponderación, llegándose al siguiente resultado:

Definición 2.– Si A es un suceso difuso de un referencial X se define la *incertidumbre probabilística* de A respecto a la distribución de probabilidades $\{p(x_i)\}_{x_i \in X}$ al valor:

$$HP(A) = -\sum_{x_i \in X} \frac{p(x_i)\, \mu_A(x_i)}{P(A)} \log p(x_i)$$

Si $P(A) = 0$ se definirá $HP(A) = 0$, lo que no supone restricción alguna ya que si $P(A) = 0$ entonces $p(x_i)\, \mu_A(x_i) = 0$, $x_i \in X$, y por tanto los elementos posibles $(p(x_i) > 0)$ no pertenecen al conjunto A con lo que no habría incertidumbre.

Así, la incertidumbre correspondiente a la probabilidad de un suceso difuso puede interpretarse como un valor medio de las incertidumbres puntuales de los elementos que pertenecen al conjunto difuso.

A continuación se estudian las propiedades de la medida que se acaba de proponer:

Teorema 2.1.– $HP(A)$ coincide con $H(A)$ si A es un subconjunto clásico del referencial.

Demostración.– La demostración de este teorema es inmediata si se tiene en cuenta que la función de pertenencia para un conjunto clásico sólo puede tomar los valores 0 y 1.

Teorema 2.2.– La incertidumbre probabilística de un suceso difuso A es no negativa:

$$HP(A) \geqslant 0$$

Además $HP(A)$ se anula si y sólo si hay un resultado seguro que "pertenece" al conjunto A, ó $P(A) = 0$.

Demostración.– En efecto, $HP(A)$ es un valor medio de autoinformaciones que son valores no negativos.

Por otro lado, si $P(A) \neq 0$ entonces $\exists\, j \,/\, p(x_j)\, \mu_A(x_j) \neq 0$, pero como $HP(A) = 0$ entonces $p(x_i)\, \mu_A(x_i) \log p(x_i) = 0$, $\forall\, i$, y en consecuencia $\log p(x_j) = 0$, por tanto $p(x_j) = 1$.

Teorema 2.3.– Si A es un conjunto completamente difuso, en el sentido que su función de pertenencia es constante, la incertidumbre probabilística asociada a A coincide con la entropía de Shannon del referencial.

Es decir, si la función de pertenencia es constante, no influye en la incertidumbre probabilística, que en este caso se debe sólo a las probabilidades.

Teorema 2.4.– Si X_1, X_2 son dos referenciales probabilísticos que se diferencian en un suceso de probabilidad nula, y A y B son dos conjuntos difusos sobre X_1 y X_2, respectivamente, tales que sus funciones de pertenencia sólo se diferencian en el elemento de probabilidad nula; entonces A y B tienen la misma incertidumbre correspondiente a las probabilidades.

Por tanto, en virtud de este teorema, y desde el punto de vista de la incertidumbre probabilística puedo considerar que los conjuntos difusos están definidos en un referencial cuyos elementos tienen todos probabilidad positiva.

Teorema 2.5.– La incertidumbre probabilística asociada a un conjunto difuso del referencial X es mayor que la entropía de Shannon del referencial respecto a la distribución de probabilidad $p(x_i)\, \mu_A(x_i) \,/\, P(A)$. Dándose la igualdad si la función de pertenencia es constante.

$$HP(A) \geqslant \boldsymbol{H}(X;\; p(x_i)\, \mu_A(x_i) \,/\, P(A))$$

Demostración.– Es inmediata aplicando el lema de Gibbs a las distribuciones de probabilidad sobre X $\{p(x_i)\}$ y $\{p(x_i)\, \mu_A(x_i) \,/\, P(A)\}$.

Teorema 2.6.– Si $\{A_j\}_{j=1,\ldots,k}$ es una k-partición difusa del conjunto difuso A. Se verifica que la incertidumbre probabilística asociada a A es el valor medio de las incertidumbres probabilísticas asociadas a cada elemento de la k-partición:

$$HP(A) = \sum_{j=1}^{k} \frac{P(A_j)}{P(A)}\, HP(A_j)$$

Demostración.– Para demostrar esta propiedad hay que tener en cuenta que por la definición de k-partición se verifica que:

$$\mu_A(x_i) = \sum_{j=1}^{k} \mu_{A_j}(x_i), \qquad \forall\, x_i \in X$$

Este resultado puede considerarse como una generalización del correspondiente a la entropía de Shannon suministrada por una partición π_A del conjunto A.

Teorema 2.7.– Si los sucesos elementales del referencial son equiprobables, entonces la incertidumbre probabilística asociada a cualquier conjunto difuso de ese referencial es la misma y coincide con la entropía de Shannon del referencial.

Este resultado pone de manifiesto que la medida considerada se debe principalmente a las probabilidades, y que la función de pertenencia sólo actúa como un elemento ponderador.

Como consecuencia de estas propiedades creemos que el valor que se ha definido es una buena medida de incertidumbre probabilística para conjuntos difusos.

3. INCERTIDUMBRE TOTAL DE UN SUCESO DIFUSO

Al considerar conjuntos difusos también se podría medir la indeterminación o vaguedad de estos conjuntos a través de las conocidas medidas de difusión y, junto con la incertidumbre probabilística, encontrar un valor que nos mida la incertidumbre total del conjunto difuso. Por otro lado, como el conjunto difuso está definido en un referencial que es el conjunto de posibles resultados de un experimento aleatorio (hay definida una distribución de probabilidades) parece lógico suponer que las probabilidades deban influir en la medida de difusión del conjunto, ya que tenemos más información que la dada por la función de pertenencia (no es lo mismo hablar de personas altas en general, que hablar de personas altas sabiendo "casi seguro" que miden más de 1,80).

Como consecuencia de todo esto se propone una medida de incertidumbre total para sucesos difusos que consta de dos partes: una que es la incertidumbre probabilística y otra que es una medida de difusión influida por la distribución de probabilidades del referencial, y que por comodidad de cálculo se elige como aquella que recuerda a la entropía cuadrática (Emptoz 1981).

Definición 1.– Se llama incertidumbre total de un suceso difuso A al valor:

$$HT(A) = -\sum_{x_i \in X} p(x_i)\, \mu_A(x_i) \left[\frac{\log p(x_i)}{P(A)} + \mu_A(x_i) - 1 \right]$$

que se puede descomponer en $HT(A) = HP(A) + D_P(A)$ siendo $HP(A)$ la incertidumbre probabilística y

$$D_P(A) = \sum_{x_i \in X} p(x_i)\, \mu_A(x_i)\, (1 - \mu_A(x_i))$$

es un valor que verifica las condiciones que generalmente se imponen a una medida de difusión, y que además tiene la ventaja de ser un valor que es menor o igual al que tendríamos si no considerásemos las probabilidades (por ser $p(x_i) \leqslant 1$), de aquí la importancia de considerar la información probabilística sobre el referencial.

El valor $HT(A)$ así definido verifica unas propiedades que ponen de manifiesto que se puede considerar como una buena medida de incertidumbre total, entre ellas destaca:

Teorema 3.1.– La incertidumbre total de un suceso difuso A es cero si y sólo si los sucesos posibles no pertenecen al conjunto o existe un suceso seguro que "pertenece totalmente al conjunto A" (su función de pertenencia vale 1)*.

$$HT(A) = 0 \Leftrightarrow P(A) = 0 \quad \text{ó} \quad P(A) = 1$$

Del resto de las propiedades se mencionará el hecho de coincidir con la entropía de Shannon para conjuntos clásicos.

4. INCERTIDUMBRE UTIL PARA SUCESOS DIFUSOS

Como continuación del estudio anterior se puede considerar el caso en el que se asigna a los distintos elementos del referencial, o a los distintos resultados de un experimento aleatorio, un valor que represente la "utilidad" respecto al objetivo final ya que, en general, todos estos elementos serán conocidos para poder asignar el valor de la función de pertenencia. En este caso trataremos de buscar una generalización de una medida de incertidumbre útil a los sucesos difusos; siguiendo un camino análogo al utilizado en el apartado 2, y tomando como base la medida de incertidumbre de M. A. Gil (1981), por creerla la más conveniente, se llega a:

Definición 1.– Se llama incertidumbre probabilística útil de un suceso difuso A al valor:

$$HPU(A) = - \sum_{x_i \in X} \frac{p(x_i)\,\mu_A(x_i)}{P(A)} \log \frac{p(x_i)\,u_i}{E(u)}$$

donde u_i es la utilidad asignada al elemento $x_i \in X$ y

$$E(u) = \sum_{x_i \in X} p(x_i)\,\mu_A(x_i)\,u_i \,/\, P(A)$$

---

* En todo el artículo el término posible se refiere a probabilidad positiva.

El valor $HPU(A)$ puede considerarse como un "valor medio" de incertidumbres puntuales útiles de los elementos del referencial que "pertenecen" al suceso difuso; y al igual que ocurría en el caso no difuso HPU puede descomponerse en dos partes: una correspondiente a la distribución de probabilidad y otra correspondiente a las utilidades:

$$HPU(A) = -\sum_{x_i \in X} \frac{p(x_i)\,\mu_A(x_i)}{P(A)} \log p(x_i) -$$

$$-\sum_{x_i \in X} \frac{p(x_i)\,\mu_A(x_i)}{P(A)} \log \frac{u_i}{E(u)} = HP(A) + HU(A)$$

Obsérvese que no se necesita asignar utilidades a todos los resultados del experimento aleatorio sino sólo a los que pertenecen al conjunto difuso que se considera.

Esta medida verifica las siguientes propiedades, que nos permiten considerarla como una buena generalización de la incertidumbre útil en el caso no difuso.

Teorema 4.1.– La incertidumbre útil de un suceso difuso es no negativa $HPU(A) \geqslant 0$.

Demostración.– Al poder descomponer la incertidumbre útil en la parte correspondiente a la distribución de probabilidad del referencial y en la parte correspondiente a las utilidades se tiene:

$$HP(A) \geqslant 0 \quad \text{(por el teorema 2.2)}$$

$$HP(A) = \log E(u) - \sum_{x_i \in X} \frac{p(x_i)\,\mu_A(x_i)}{P(A)} \log u_i \geqslant 0$$

(por la concavidad hacia arriba de la función logarítmica).

Además, se obtiene la igualdad si y sólo si existe un elemento del referencial que tiene probabilidad 1 y que "pertenece" al conjunto difuso.

Teorema 4.2.– Si A es un conjunto completamente difuso, la incertidumbre útil del conjunto difuso A sólo depende de la función de utilidad y de la distribución de probabilidad.

Demostración.– Por ser la función de pertenencia constante $P(A) = \mu$, si

$$\mu(x_i) = \mu, \quad x_i \in X, \quad \text{y} \quad E(u) = \sum_{x_i} p(x_i)\, u_i,$$

por tanto:

$$HPU(A) = -\sum_{x_i \in X} p(x_i) \log \frac{P(x_i)\, u_i}{E(u)}$$

Obteniéndose en este caso la incertidumbre útil del referencial.

Teorema 4.3.– Si la función de utilidad es constante, entonces la incertidumbre útil asociada a un conjunto difuso coincide con la incertidumbre probabilística de este conjunto difuso.

Demostración.– Es inmediata aplicando la definición.

Este resultado pone de manifiesto que lo que nos interesa no es tanto la utilidad absoluta de los elementos del referencial como las utilidades relativas de los elementos que pertenecen al conjunto difuso que se considera.

Teorema 4.4.– Fijada la distribución de probabilidades del referencial X la incertidumbre útil de un suceso difuso es mínima si todos los elementos que "pertenecen" al conjunto difuso y son posibles tienen la misma utilidad.

5. EXTENSIONES AL CASO CONTINUO

Todos los resultados de los apartados anteriores pueden extenderse al caso en el que el conjunto de resultados del experimento aleatorio es infinito. Es decir, cuando se considera como referencial un espacio de probabilidad con distribución de probabilidad absolutamente continua y como consecuencia se puede trabajar con su función de densidad $f(x)$.

En este caso A se considera un suceso difuso si su función de pertenencia es medible respecto al álgebra del espacio probabilístico; y aplicando la definición de Zadeh se tiene:

Definición 1.– Si A es un suceso difuso de un espacio probabilístico X cuya distribución de probabilidad está caracterizada por la función de densidad $f(x)$, se define su probabilidad como el valor

$$P(A) = \int_X \mu_A(x) f(x)\, dx$$

con lo que la extensión de la incertidumbre probabilística sería:

Definición 2.– La incertidumbre probabilística de un suceso difuso A es el valor

$$HP(A) = -\frac{1}{P(A)} \int_X \mu_A(x) f(x)\, dx$$

siempre que la integral exista y $P(A) \neq 0$; si $P(A) = 0$ se definiría $HP(A) = 0$.

Este valor verifica unas propiedades que son análogas al caso no difuso y entre las que merece destacarse el hecho de que $HP(A)$ puede ser un valor negativo, para comprobarlo basta el mismo ejemplo que se suele considerar en el caso no difuso, que es tomar como referencial un intervalo con la distribución uniforme. Así, si $X = (0, a)$ y $f(x) = U(0, a)$ se tendría:

$$HP(A) = -\frac{1}{P(A)} \int_0^a \mu_A(x) \frac{1}{a} \log \frac{1}{a} \, dx =$$

$$= \log a \frac{1}{P(A)} \int_0^a \mu_A(x) \frac{1}{a} \, dx = \log a = \boldsymbol{H}(X)$$

independientemente del conjunto difuso que se considere.

Análogamente a como ocurría en el caso no continuo, también se propone una medida de incertidumbre total que considere la aleatoriedad y la vaguedad.

Definición 3.– La entropía total de un suceso difuso A es el valor

$$HT(A) = -\int_X f(x) \mu_A(x) \left[\frac{\log f(x)}{P(A)} - 1 + \mu_A(x) \right]$$

cuando esta integral exista y $P(A) \neq 0$.

Este valor se puede descomponer en dos partes, una que es la incertidumbre probabilística y otra que corresponde a una medida de difusión

$$HT(A) = HP(A) + D_P(A)$$

siendo

$$D_P(A) = \int_X f(x) \mu_A(x) (1 - \mu_A(x)) \, dx$$

6. APLICACIONES

Como posible aplicación de la incertidumbre probabilística se está estudiando la medida de la información suministrada por un experimento aleatorio sobre un suceso difuso.

BIBLIOGRAFIA

[1] DE LUCA, A., TERMINI, S. (1972): *A definition a non probabilistic entropy in the setting of fuzzy sets theory*. Inform. Contr. 20, pp. 301-312.

[2] DE LUCA, A., TERMINI, S. (1979): *Entropy and energy measures of a fuzzy sets*. Advances in fuzzy sets theory and applications. North-Holland.

[3] EMPTOZ, H. (1981): *Nonprobabilistic entropies and indetermination measures in the setting of fuzzy sets theory*. Fuzzy sets and systems, 5, pp. 309-318.

[4] GIL, M. A. (1981): *Estudio de una medida de incertidumbre correspondiente a las probabilidades*. Trabajos Estadísticos Investigación Oper., 32, pp. 45-66.

[5] KNOFMACHER, J. (1975): *On measures of fuzziness*. J. Math. Anal. Appl. 49, pp. 529-534.

[6] LOO (1976): *Measures of fuzziness*. Cibernetica, 20, pp. 201-210.

[7] ZADEH, L. (1965): *Fuzzy sets*. Inform. Contr. 8, pp. 338-353.

[8] ZADEH, L. (1968): *Probability measures of fuzzy events*. J. Math. Anal. Appl., 23, pp. 421-427.

Departamento de Matemáticas
Universidad de Oviedo