SOBRE LA PROBABILIDAD DIFUSA

F.J. Montero de Juan

Departamento de Estadística e I.O.

Facultad de Ciencias Matemáticas

Universidad Complutense de Madrid

RESUMEN

El concepto de "probabilidad difusa" propuesto por Zadeh resulta pobre para algunas aplicaciones prácticas. En este trabajo se generaliza el tratamiento de Yager a tales problemas y se introduce en el contexto de la difusidad comprehensiva.

Palabras clave: Probabilidad difusa, Difusidad Comprehensiva.

Introducción

El primer intento de formalizar el concepto de "probabilidad difusa" es el elaborado por Zadeh [7] . Un conjunto difuso A queda definido por el propio Zadeh a través de una "función de pertenencia"

$$\mu_A : X \dashrightarrow [0,1]$$

sobre un conjunto - en el sentido clásico - de objetos, de modo que cuanto más próximo esté $\mu_A(x)$ a 1, más pertenece el elemento x al conjunto difuso A. Si a un σ-álgebra $\mathcal{Q}(X)$ de subconjuntos de X se le asocia una medida de probabilidad

$$P : \mathcal{Q}(X) \longrightarrow [0,1]$$

y se supone que la función de pertenencia es medible respecto de este σ-álgebra y el σ-álgebra de Borel en el intervalo $[0 , 1]$, entonces tiene sentido la definición de Zadeh, que propone como "probabilidad difusa" de A la

La definición de Zadeh corresponde al caso $L = [0,1]$ con la ordenación natural.

Pero el caso con más analogía con la definición de Zadeh es suponer, en vez de $L = [0,1]$, un conjunto L totalmente ordenado. Es decir, un conjunto dotado de una relación de orden "$\leq$" de modo que $a \leq b$ o bien $b \geq a$ para cualesquiera $a, b \in L$. Es evidente entonces que para todo par de elementos de L existen el supremo y el ínfimo, y por tanto L es un retículo, aunque no se pueda asegurar en general que sea completo. Pero desde luego todo conjunto totalmente ordenado es retículo distributivo.

Cuando L sea un conjunto ordenado, entonces podremos asignar a cada elemento $x \in X$ un grado de verificación, que será mayor cuanto más alta sea la posición de $\mu_A(x) \in L$ dentro de la cadena que determina el orden en L. Si L es completo, están definidos un mínimo $0 = \bigwedge_{a \in L} a \in L$ y un máximo $1 = \bigvee_{a \in L} a \in L$ que corresponden a los conjuntos nítidos asociados. Parece natural aceptar que, dado un conjunto difuso A, se puede decir cuándo un elemento está totalmente en A o totalmente fuera de A.

Definición .- Sea $(X, \mathcal{Q}(X), P)$ un espacio de probabilidad y sea $\mu : X \rightarrow L$ una función de pertenencia en un retículo L totalmente ordenado, con μ medible en el sentido de que

$$\mu^{-1}\{b \in L \,/\, a \leq b\} \in \mathcal{Q}(X) \qquad \forall a \in L$$

Entonces llamaremos "función de distribución difusa" del conjunto caracterizado por μ a la función

$$S_\mu : L \dashrightarrow [0,1]$$

$$S_\mu(a) \quad P(\mu^{-1}\{b \in L \,/\, a \leq b\}) = \int_{\{x \in X/\mu(x) \geq a\}} dP$$

El significado de esta valoración de L en el intervalo $[0,1]$ es claro: S(a) representa la probabilidad de que, tomando al azar un elemento de la población X, éste pertenezca al conjunto difuso al menos al nivel a. En con-

que sería la esperanza de la función de pertenencia.

Tal definición se puede justificar básicamente por su analogía con el caso clásico: por una parte, la probabilidad de un conjunto clásico B resulta ser la esperanza de su función característica, que es en este caso la función indicatriz I_B, que toma el valor $I_B(x) = 0$ si $x \notin B$ y el valor $I_B(x)=1$ si $x \in B$. De este modo, si A es un conjunto nítido (con función de pertenencia $\mu_A(x) \in \{0,1\}$, $\forall x \in X$) entonces la definición de Zadeh coincide con la definición clásica.

Por otra parte, la definición de Zadeh trata de evaluar el grado en el que una muestra posee la propiedad a la que se refiere el conjunto difuso A, desde un punto de vista frecuentista. El experimento aleatorio asociado daría como resultado una familia de observaciones $x_i \in X$, más o menos pertenecientes a A. Para una muestra de tamaño n, la "frecuencia" de A sería la proporción de elementos en A, y entonces

$$P(A) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mu_A(x_i)$$

Pero tal definición tiene un significado exacto de "promedio" que no va a permitir responder muchas preguntas de interés. Por ejemplo, ni siquiera podremos responder cuál es la probabilidad del conjunto de elementos que está totalmente en A, cuando éste sea un conjunto difuso. Y difícilmente se puede decir que la probabilidad difusa P(A) representa la probabilidad de obtener al azar un elemento de A, como ocurre en el caso nítido, máxime cuando cada "elemento de A" no viene descrito sino por su grado de pertenencia $\mu_A(x)$.

Distribución Difusa

En lo que sigue vamos a utilizar la generalización del concepto de conjunto difuso debida a Goguen [2]. Si L es un conjunto parcialmente ordenado, un conjunto difuso en L será una aplicación

$$\mu : X \longrightarrow L$$

Lo usual es considerar a L un retículo distributivo completo.

$$P^*_\mu \{ b \in L \,/\, a \geq b \} = S_\mu(a) \qquad \forall a \in L$$

Relación con la Difusidad comprehensiva

Ahora bien: el retículo L suele ser un conjunto discreto, y toda la información de tal estructura puede entonces asimilarse como una característica difusa, en el sentido de [4] . Por ejemplo, si a un elemento $x \in X$ se le asigna un valor $\mu(x) \in L$, y sólo uno, podemos considerar la familia de conjuntos difusos

$$\{ \mu_b : X \longrightarrow [0,1] \}_{b \in L}$$

de modo que $\mu_b(x) = 1$ cuando $b = \mu(x)$ y cero en caso contrario. Es trivial comprobar que de este modo hemos obtenido una característica difusa, aunque ésta estructura es desde luego más general, pues admite para cada $b \in L$ grados de verificación en el intervalo $[0,1]$. Su relación con la lógica $[0,1]$ - valorada es también inmediata (ver [1] , por ejemplo).

Sean entonces $(X, \mathcal{A}(X), P)$ un espacio de probabilidad y $(X, \mathcal{L}, \mu)$ un "espacio difuso" en el sentido de [4] , es decir:

- $\mathcal{A}$ es un σ-álgebra de etiquetas con I características:

$$\mathcal{A} = \sigma(\bigcup_{i \in I} J_i)$$

donde

$$L_i = \{ b_{ij} \}_{j \in N_i} \subset \mathbb{N}$$

es el conjunto de etiquetas básicas de la característica J_i que es cerrada bajo una operación "adición": $b \in J_i$ si y sólo si existe una familia de etiquetas $N(b) \subset N_i$ tal que

$$b = \sum_{j \in N(b)} b_{ij}$$

El σ-álgebra $\mathcal{A}$ es cerrado bajo operaciones numerables del tipo "disjunción" $(\bigvee_n b_n)$, "negación" (b') o "conjunción" $(\bigwedge_n b_n) = (\bigvee_n b'_n)'$.

reto, S(1) será la probabilidad de que tal elemento pertenezca totalmente al conjunto difuso.

Si el retículo fuese $L \subset \mathbb{R}$, entonces tendría sentido el cálculo de la esperanza matemática de la función de pertenencia de modo que ésta coincide con la "probabilidad difusa" de Zadeh cuando $L = [0,1]$. En este último caso, si además el espacio muestral X es finito, observamos que S(a) coincide con la probabilidad del conjunto de nivel A_a (ver Yager [5]), a priori los conjuntos de mayor interés en el estudio del conjunto difuso, con idéntico significado: probabilidad de obtener al menos un nivel a - en este caso en $[0,1]$ - de verificación del conjunto difuso A.

Evidentemente S es una función monótona no creciente, de modo que, caso de existir un mínimo $o \in L$ y un máximo $1 \in L$, entonces $S(o) = 1$ y $S(1) = P\{ x \in X, \ \mu(x) = 1 \}$.

Luego cuando A sea un conjunto nítido, con función de pertenencia $\mu(x)$ tomando los valores $0, 1 \in \mathbb{R}$ según sea $x \in A^c$ ó $x \in A$, entonces $S(1) = P(A)$.

Desde luego la representación de Yager [6] , modificación a su vez de la variante de Klement [3] a la inicial de Yager [5] , se puede utilizar en este contexto sustituyendo el intervalo $[0,1]$ por un conjunto L totalmente ordenado y completo.

Hay que señalar además dos cuestiones: por una parte, si $\mathcal{A}(X) = \mathcal{P}(X)$ está formado por todos los subconjuntos de X - esto es lo usual cuando X es discreto - entonces sabemos que toda función de pertenencia considerada es medible. Por otra parte, dado que a L se le puede dotar naturalmente de una topología de orden, cabe preguntarse acerca de la posibilidad de obtener como extensión una medida de probabilidad sobre el σ-álgebra de Borel $\mathcal{B}_L$ asociada a tal topología, es decir, el mínimo σ-álgebra que contiene a los abiertos en la topología de orden. Así, si L es un conjunto totalmente ordenado y completo, y S_μ es una función de distribución difusa según la definición anterior, entonces tendremos inducida una medida de probabilidad P^*_μ definida en el σ-álgebra de Borel $\mathcal{B}_L$ engrendrada por la topología de orden, tal que

- la aplicación $\mu : X \times \mathcal{A} \longrightarrow [0\ ,\ 1]$ verifica las siguientes propiedades:

$$\mu(x, \sum_{j \in N_i} b_{ij}) = 1 \qquad \forall x \in X \quad \forall i \in I$$

$$\mu(x, \sum_{j \in M} b_{ij}) = \sum_{j \in M} \mu(x, b_{ij}) \qquad \forall x \in X \quad \forall M \subset N_i \quad \forall i \in I$$

$$\mu(x, \bigvee_n b_n) = \sup_n \mu(x, b_n) \qquad \forall x \in X \quad \forall (b_n)_{n \in \mathbb{N}} \subset \mathcal{A}$$

$$\mu(x, b') = 1 - \mu(x, b) \qquad \forall x \in X \quad \forall b \in \mathcal{A}$$

Si suponemos que $\mu(\cdot, b)$ es medible para cualquier etiqueta $b \in \mathcal{A}$, entonces cada una de estas etiquetas induce una probabilidad difusa

$$P_b^* : \mathcal{B}_{[0,1]} \longrightarrow [0\ ,\ 1]$$

corresponde a la propiedad a la que se refiere la etiqueta. De este modo tiene sentido definir la "probabilidad difusa" inducida como una aplicación

$$P^* : \mathcal{B}_{[0,1]} \times \mathcal{A} \longrightarrow [0,1]$$

$$P^*(B,b) = P_b^*(B)$$

es decir, que la probabilidad difusa es así entendida como un proceso en el que el espacio de subíndices es un σ-álgebra de etiquetas, verificando las condiciones de coherencia que se deducen de la definición de espacio difuso. Por ejemplo, es claro que

$$P_b^*(B) = P_{b'}^*(1-B) \qquad \forall B \in \mathcal{B}_{[0,1]}$$

donde $1 - B = \{ y = 1 - x \}_{x \in B}$

REFERENCIAS

Czogala, E. "An Introduction to Probabilistic L-Valued logic" Fuzzy sets and Systems 13 (1984), 179-185.

Goguen, J.A. "L - Fuzzy Sets" J. Mathematical Analysis and Applications 18 (1967) 145-174.

Klement, E.P. "Some Remarks on a Paper of R.R. Yager" Information Sciences 27 (1982), 211-220.

Montero, F.J. "Comprehensive Fuzzines" Fuzzy Sets and Systems (pendiente de publicación)

Yager, R.R. "A Note on Probabilities of Fuzzy Events" Information Sciences 18 (1979) 113-122.

Yager, R.R. "A Representation of the Probability of a Fuzzy Subset" Fuzzy Sets and Systems 13 (1984) 273-283.

Zadeh, L.A. "Probability Measures of Fuzzy Events" 23 (1968), 421-427.