TRABAJOS DE ESTADISTICA
Vol. 3, núm. 1, 1988 pp. 15 a 31

MEDIDAS DE ENTROPIA EN LA TEORIA DE LA EVIDENCIA

LAMATA, M., y MORAL, S.
Dpto. Ciencias Computación e I.A.
Fac. de Ciencias
Univ. de Granada

ABSTRACT

The aim of this paper is to define global measures of uncertainty in the framework of Dempster-Shafer's Theory of Evidence. Starting from the concepts of entropy and specificity introduced by Yager, two measures are considered; the lower entropy and the upper entropy.

Key words: Theory of Evidence, entropy.

Title: Entropy Measures in the Theory of Evidence.

RESUMEN

El objetivo de este trabajo es definir medidas globales de incertidumbre en el marco de la Teoría de la Evidencia de Dempster-Shafer. Partiendo de los conceptos de entropía y especificidad introducidos por Yager, se consideran dos medidas; la entropía inferior y la entropía superior.

Palabras clave: Teoría de la Evidencia, entropía.

Clasificación AMS: 94A17, 94D05.

1. INTRODUCCION

Sea X una variable que toma valores en un conjunto finito U y $X(\omega)$ un valor particular, desconocido, de esta variable. Si X es una variable aleatoria podremos disponer de una información sobre dicho valor,

Revisado, diciembre 1987

representable mediante una medida de probabilidad. Pero, en general, existen otros tipos de informaciones no representables de «forma probabilística». Entre ellas se encuentran, por ejemplo, las llamadas informaciones difusas (ver Zadeh (1978)). Según Kampé de Feriet (1974): «... La información es dada por un informador a un receptor bajo la forma de una proposición, en el sentido de la lógica formal, es decir de una frase que asocia por medio del verbo ser un nombre (objeto) y un predicado (propiedad, atributo; generalmente expresado por un adjetivo)». Una información difusa sobre $X(\omega)$ será una asociación entre dicho valor y una cualidad de tipo impreciso (Zadeh (1965)). Por ejemplo, si U es un conjunto de números, una proposición de la forma «$X(\omega)$ es pequeño será una información difusa.

La Teoría de la Evidencia de Dempster (1967) y Shafer (1976) proporciona un modelo para la representación de nuestro conocimiento sobre $X(\omega)$ (una evidencia) a través de una aplicación

$$m: \mathscr{P}(U) \to [0,1]$$

que verifique

$$m(\phi) = 0 \quad \text{y} \quad \sum_{A \subset U} m(A) = 1$$

Estas aplicaciones, llamadas asignaciones básicas de probabilidad, no son medidas de probabilidad, ni tan siquiera medidas difusas en el sentido de Sugeno (1974). Sin embargo, cada una de ellas lleva asociadas dos medidas difusas; una medida de plausibilidad, dada por

$$\mathrm{Pl}(A) = \sum_{B \cap A \neq \phi} m(B)$$

y otra de credibilidad

$$\mathrm{Bel}(A) = \sum_{B \subset A} m(B)$$

Para las propiedades de estas medidas puede consultarse Shafer (1976), Dubois-Prade (1982) o Lamata (1986).

Un aspecto importante de esta estructura es que proporciona un marco común donde se pueden representar, entre otras, las informaciones difusas y las de tipo probabilístico.

Una distribución de probabilidad p en U se representa mediante una asignación básica m dada por

$$\begin{aligned} m(\{a\}) &= p(a) \quad , \quad a \in U \\ m(A) &= 0 \qquad \text{en otro caso.} \end{aligned}$$

Obsérvese que tanto Pl como Bel coinciden con la medida de probabilidad P asociada a p. Las asignaciones básicas que, como ésta, se concentren en conjuntos unitarios diremos que son de tipo probabilístico y a p le llamaremos distribución de probabilidad asociada a m.

Una información difusa de la forma «$X(\omega)$ es A» donde A es un subconjunto difuso de U con función de pertenencia μ, normalizada ($\operatorname{Max}_{a \in A} \mu(a) = 1$), puede representarse mediante la asignación básica de probabilidad

$$\begin{aligned} m(A_i) &= \mu(a_i) - \mu(a_{i-1}) \quad , \quad \text{para} \quad 1 \leqslant i \leqslant n-1 \\ m(U) &= \mu(a_n) \\ m(A) &= 0 \quad , \quad \text{en otro caso} \end{aligned}$$

donde $U = \{a_1, ..., a_n\}$ y $A_1 = \{a_1, ..., a_i\}$ con $1 = \mu(a_1) \geqslant \cdots \geqslant \mu(a_n)$.

Para la medida de plausibilidad asociada, que anotaremos π, se tiene

$$\pi(B) = \operatorname*{Max}_{a \in B} \{\mu(a)\}.$$

La medida de credibilidad, N, tiene la forma

$$N(B) = \operatorname*{Min}_{a \notin B} \{1 - \mu(a)\}$$

Observese que π coincide con la medida de posibilidad asociada a la misma información difusa en Zadeh (1978). Por este motivo, este tipo de asignaciones básicas de probabilidad serán llamadas también, en lo que sigue, de tipo posibilístico.

Yager (1983) se ha planteado el problema de determinar la calidad de una información representada por una asignación básica de probabilidad, m, midiendo la cantidad de incertidumbre asociada a la misma.

Yager observa que existen dos fuentes de incertidumbre para una asignación básica de probabilidad particular. Una primera que se origina cuando m se concentra en subconjuntos de U disjuntos entre sí (discordancia de la información) y una segunda que se produce cuando existen subconjuntos de U de cardinal mayor que uno con asignación positiva (imprecisión de la información).

Nuestro objetivo en este trabajo es medir la cantidad total de incertidumbre asociada a una evidencia sea cual sea su origen. Para este fin, propondremos varias medidas, que llamaremos de entropía, y de las que estudiaremos sus propiedades más importantes. Previamente necesitamos las siguientes definiciones.

Definición (Moral (1985)). Sean m y m^* dos asignacones básicas de probabilidad en U. Diremos que m está incluida en m^* ($m \subset m^*$) si y sólo si para todo $A \subset U$, existe una aplicación t_A: $\mathscr{P}(A) \to [0, 1]$ verificando

$$m(A) = \sum_{B \subset A} t_A(B) \quad \text{y} \quad m^*(C) = \sum_{B \supset C} t_B(C)$$

Definición. Sea m una asignación básica de probabilidad en el producto cartesiano $U_1 \times U_2$. Se llama evidencia marginal en U_1 ($i = 1, 2$) a la dada por la asignación

$$m_i(A) = \sum_{p_i(B) = A} m(B)$$

donde

$$p_1(B) = \{x \in U_1 / \exists \ y \in U_2 \quad \text{con} \quad (x, y) \in B\}$$
$$p_2(B) = \{x \in U_2 / \exists \ y \in U_1 \quad \text{con} \quad (x, y) \in B\}$$

Definición (Lamata (1986)). Sea m una asignación básica de probabilidad en $U_1 \times U_2$ con marginales m_1 y m_2. Se dice que existe independencia fuerte entre las marginales si y sólo si

$$m(A \times B) = m_1(A) \cdot m_2(B) \quad , \quad \forall\, A \subset U_1, B \subset U_2.$$

Es inmediato comprobar que si m_1 y m_2 son fuertemente independientes entonces se verifican las siguientes propiedades:

— $Pl(A \times B) = Pl_1(A) \cdot Pl_2(B)$.
— $Bel(A \times B) = Bel_1(A) \cdot Bel_2(B)$.
— $m(C) > 0 \Rightarrow C = A \times B \quad \text{con} \quad A \subset U_1 \text{ y } B \subset U_2$.

Ejemplo 1. Supongamos que tenemos que determinar la posición de un objeto en una zona determinada que tenemos dividida en tres bandas horizontales ($\{h_1, h_2, h_3\}$) y tres verticales ($\{v_1, v_2, v_3\}$), dando lugar a 9 casillas. Supongamos que tenemos dos informaciones sobre esta posición,

I_1: «La banda horizontal es h_2».
I_2: «La banda vertical es v_1».

Supongamos, además, que estas informaciones no son del todo fiables en el sentido de que si nos han dicho que está en una banda, entonces,

— El 70 % de las ocasiones es correcta seguro.
— El 10 % de las veces puede haber una equivocación y estar en la banda contigua.
— El 20 % de las ocasiones la información no significa nada, pudiendo estar en cualquier banda.

Estas informaciones se pueden representar mediante sendas evidencias, m_1 y m_2, en los conjuntos $U_1 = \{h_1, h_2, h_3\}$, $U_2 = \{v_1, v_2, v_3\}$, dadas por

$$m_1(\{h_2\}) = 0{,}7 \quad , \quad m_1(U_1) = 0{,}3$$

$$m_2(\{v_1\}) = 0.7 \quad , \quad m_2(\{v_1, v_2\}) = 0{,}1 \quad , \quad m_2(U_2) = 0{,}2$$

Si suponemos que estas informaciones son fuertemente independientes, tendremos una información sobre la localización del objeto en el mapa dada por

$$m(\{(h_2, v_1)\} = 0{,}49 \quad , \quad m(U_1 \times \{v_1\}) = 0{,}21$$

$$m(\{h_2\} \times \{v_1, v_2\}) = 0{,}07 \quad , \quad m(U_1 \times \{v_1, v_2\}) = 0{,}03$$

$$m(\{h_2\} \times U_2) = 0{,}014 \quad , \quad m(U_1 \times U_2) = 0{,}06$$

2. MEDIDAS DE CALIDAD DE YAGER

Sea m una asignación básica de probabilidad en U, las siguientes definiciones pretenden determinar medidas de la calidad de m y se deben a Yager (1983).

Definición. Se llama entropía de m a la cantidad

$$E(m) = \sum_{A \subset U} m(A)L(Pl(A))$$

donde $L(.)$ es la función logaritmo neperiano.

Definición. Se llama especificidad de m a

$$S(m) = \sum_{A \subset U} m(A)/|A|$$

donde $|.|$ representa el cardinal y se supone $m(\phi)/|\phi| = 0$.

Entre las propiedades demostradas en Yager (1983) para estos valores se encuentran:

1) $E(m)$ y $S(m)$ existen y son finitas cualquiera que sea m.

2) Si m es de tipo probabilístico entonces
 — $E(m) = -\sum_{a \in U} m(\{a\})\, L(m(\{a\}))$
 — $S(m) = 1$
 Nótese que $E(m)$ coincide con la entropía de Shannon de la distribución de probabilidad asociada.

3) — $E(m) = 0 \Leftrightarrow (\forall\, A, B \subset U,\ m(A) > 0,$
 $m(B) > 0 \Rightarrow A \cap B \neq \phi)$
 — $S(m)$ alcanza su mínimo valor si $m(U) = 1$ y entonces $m(U) = 1/|U|$.

4) Si m y m^* son de tipo posibilístico con medidas asociadas π y π' y se verifica $\pi(A) \geqslant \pi'(A)$, $\forall\, A \subset U$, entonces $S(m^*) \geqslant S(m)$.

De las propiedades anteriores se deduce que, mientras la entropía mide la «discordancia» de una evidencia, la especificidad mide la «precisión» de la misma: si m se concentra en conjuntos de cardinal pequeño la especificidad será alta y si lo hace en conjuntos de cardinal grande ésta será baja. La combinación de ambos valores proporciona una medida de la calidad de la evidencia: cuanto menor sea la entropía y mayor la especificidad, tendremos mejores informaciones.

En lo que sigue vamos a relacionar los anteriores conceptos con los de independencia fuerte e inclusión de asignaciones básicas de probabilidad.

Teorema. Si m es una asignación básica de probabilidad en $U_1 \times U_2$ con marginales m_1 y m_2 fuertemente independientes, entonces

a) $E(m) = E(m_1) + E(m_2)$.
b) $S(m) = S(m_1) \cdot S(m_2)$.

Demostración

a) Por definición

$$E(m) = - \sum_{C \subset U_1 \times U_2} m(C) L(Pl(C))$$

Ahora bien, si existe independencia fuerte se verifica

$$m(C) > 0 \Rightarrow C = A \times B \quad \text{con} \quad A \subset U_1 \text{ y } B \subset U_2$$

$$m(A \times B) = m_1(A) \cdot m_2(B) \text{ y}$$

$$Pl(A \times B) = Pl_1(A) \cdot Pl_2(B)$$

Por tanto, podemos poner

$$\begin{aligned} E(m) &= - \sum_{A \times B \subset U_1 \times U_2} m(A \times B) L(Pl(A \times B)) = \\ &= - \sum_{A \times B \subset U_1 \times U_2} m_1(A) m_2(B) [L(Pl_1(A)) + L(Pl_2(B))] = \\ &= \left[\sum_{B \subset U_2} m_2(B) \right] \left[- \sum_{A \subset U_1} m_1(A) L(Pl_1(A)) \right] + \\ &+ \left[\sum_{A \subset U_1} m_1(A) \right] \left[- \sum_{B \subset U_2} m_2(B) L(Pl_2(B)) \right] = \end{aligned}$$

$E(m_1) + E(m_2)$, como queríamos probar.

b) Teniendo en cuenta las propiedades de la independencia fuerte, tenemos

$$S(m) = \sum_{C \subset U_1 \times U_2} \frac{m(C)}{|C|} = \sum_{A \times B \subset U_1 \times U_2} \frac{m(A \times B)}{|A \times B|} =$$

$$= \sum_{A \times B \subset U_1 \times U_2} \frac{m_1(A) \cdot m_2(B)}{|A| \cdot |B|} =$$

$$= \left(\sum_{A \subset U} \frac{m_1(A)}{|A|} \right) \cdot \left(\sum_{B \subset U_2} \frac{m_2(B)}{|B|} \right) = S(m_1) \cdot S(m_2)$$

Con lo que concluye la demostración.

Teorema. Si m y m^* son dos asignaciones básicas de probabilidad en U de tal forma que $m \subset m^*$ entonces $S(m) \leqslant S(m^*)$.

Demostración. Por definición, si $m \subset m^*$ entonces $\forall A \subset U$ existe una aplicación t_A: $\mathscr{P}(A) \to [0, 1]$ tal que se verifica

$$m(C) = \sum_{D \subset C} t_C(D) \quad ; \quad m^*(E) = \sum_{F \supset E} t_F(E).$$

Se tiene entonces que

$$S(m) = \sum_{A \subset U} (m(A)/|A|) = \sum_{A \subset U} \left(\sum_{B \subset A} (t_A(B)/|A|) \right) \leqslant$$

$$\leqslant \sum_{A \subset U} \left(\sum_{B \subset A} (t_A(B)/|B|) \right) = \sum_{B \subset U} (1/|B|) \left(\sum_{A \supset B} t_A(B) \right) =$$

$$= \sum_{B \subset U} (m^*(B)/|B|) = S(m^*) \quad , \text{ como queríamos probar.}$$

3. ENTROPIAS INFERIOR Y SUPERIOR

En este apartado vamos a estudiar medidas de la incertidumbre total asociada a una evidencia. Dicha incertidumbre proviene de la discordancia y de la imprecisión de la información. Por tal motivo proponemos, en primer lugar, el funcional entropía inferior que damos en la siguiente definición.

Definición. Sea m una asignación básica de probabilidad en U con medida de plausibilidad asociada Pl. Llamaremos entropía inferior de m a

$$H_*(m) = - \sum_{A \subset U} m(A)L(Pl(A)/|A|)$$

donde se supone $m(\phi)L(Pl(\phi)/|\phi|) = 0$.

Obsérvese que,

$$H_*(m) = - \sum_{A \subset U} m(A)L(Pl(A)) - \sum_{A \subset U} m(A)L(1/|A|) =$$
$$= E(m) + - \sum_{A \subset U} m(A)L(1/|A|).$$

El valor $I(m) = - \sum_{A \subset U} m(A)L(1/|A|)$ ha sido considerado por Dubois-Prade (1984) como una medida de imprecisión para la evidencia dada por la asignación básica de probabilidad m. Así H_* es la suma de una medida de discordancia y una de imprecisión pudiéndose considerar, por tanto, como una medida de la incertidumbre total.

Las siguientes propiedades se deducen inmediatamente de la definición.

1) $H_*(m)$ existe y es finita cualquiera que sea m.

2) Si m es de tipo probabilístico, entonces

$$H_*(m) = - \sum_{a \in U} m(\{a\})L(m(\{a\}))$$

Es decir, coincide con la entropía de Shannon.

3) $H_*(m) = 0$ si y sólo si m es de tipo probabilístico degenerada en un elemento de U ($\exists\, a \in U$ con $m(\{a\}) = 1$).

4) Si m es una asignacón básica de probabilidad en $U_1 \times U_2$ con marginales m_1 y m_2 fuertemente independientes, entonces

$$H_*(m) = H_*(m_1) + H_*(m_z).$$

La siguiente proposición relaciona H_* con la entropía de Shannon de la distribución de probabilidad asociada a una evidencia por Dubois-Prade (1982).

Proposición. Si m es una asignación básica de probabilidad en U, entonces $H_*(m)$ es menor o igual que la entropía de Shannon de la distribución de probabilidad dada por

$$a \in U,\ p(a) = \sum_{A \ni a} m(A)/|A|$$

Demostración

La entropía de Shannon de $\{p(a)\}_{a \in U}$ vale

$$H(p) = -\sum_{a \in U} p(a)L(p(a)) = -\sum_{a \in U}\left(\sum_{A \ni a} m(A)/|A|\right)L(p(a)) =$$

$$= -\sum_{A \subset U} (m(A)/|A|)\left(\sum_{a \in A} L(p(a)).\right.$$

Teniendo en cuenta que $-L(x)$ es una función convexa, se verifica:

$$\frac{-\sum_{a \in A} L(p(a))}{|A|} \{ -L\left(\sum_{a \in A} \frac{p(a)}{|A|}\right)$$

y, por tanto, se tiene

$$H(p) \geqslant -\sum_{A \subset U} m(A)L\left((\sum_{a \in A} p(a))/|a|\right).$$

Como además se verifica

$$\sum_{a \in A} p(a) = \sum_{a \in A}\left(\sum_{B \ni a} (m(B)/|B|)\right) = \sum_{B \cap A \neq \phi} (m(B)/|B|)|B \cap A| \leqslant$$

$$\leqslant \sum_{B \cap A \neq \phi} m(B) = Pl(A)$$

obtenemos, finalmente,

$$H(p) \geqslant - \sum_{A \subset U} m(A) L(Pl(A)/|A|) = H_*(m)$$

que es lo que queríamos probar.

Proposición. Si m es una asignación básica de probabilidad en U, entonces $H_*(m) \leqslant L(|U|)$, alcanzándose esta cota superior para cualquier asignación m tal que exista una partición $\{A_i\}$ de U con

$$m(A) = (|A_i|/|U|) \qquad \text{si } A = A_i$$
$$m(A) = = 0 \qquad\qquad \text{en otro caso}$$

Demostración

De acuerdo con la proposición anterior $H_*(m) \leqslant H(p)$ siendo p una distribución de probabilidad en U. Por otra parte, por las propiedades de la entropia de Shannon, $H(p) \leqslant L(|U|)$. Así pues, $H_*(m) \leqslant L(|U|)$ para toda m.

Sea ahora m como en la proposición. Obviamente

$$H_*(m) = -\sum_i m(A_i) L(Pl(A_i)/|A_i|).$$

Puesto que $A_i \cap A_j = \phi$ si $i \neq j$, se tiene

$$H_*(m) = -\sum_i \frac{|A_i|}{|U|} L\left(\frac{|A_i|}{|U|} \cdot \frac{1}{|A_i|}\right) =$$

$$= -\left((\sum_i |A_i|)/|U|\right) \cdot L(1/|U|) = L(|U|),$$

que es lo que queríamos probar.

Ejemplo 2. Continuando con el ejemplo 1, la entropía inferior en U_1, después de la información I_1 es

$$E_*(m_1) = -0{,}7L(1) - 0{,}3L(1/3) = 0{,}3L(3),$$

mientras que la entropía inferior de m_2 es

$$E_*(m_2) = -0{,}7L(1) - 0{,}1L(1/2) - 0{,}2L(1/3) = 0{,}2L(3) + 0{,}1L(1/2)$$

En ambos casos, la entropía de Yager es cero. La entropía inferior no proviene en ninguna de las dos situaciones de la contradicción interna de las informaciones, sino de la falta de precisión de las mismas. $E_*(m_2)$ es un poco menor que $E_*(m_1)$ ya que I_2 proporciona un poco más de precisión que I_1.

La entropía del problema global de localizar el objeto será

$$E_*(m) = E_*(m_1) + E_*(m_2) = 0{,}5L(3) + 0{,}1L(2).$$

El nombre de entropía inferior para H_* proviene de que en numerosas ocasiones, se obtiene una medida de la incertidumbre menor que la intuitivamente esperada, como muestra el siguiente ejemplo:

Ejemplo 3. Partiendo del problema del ejemplo 1, supongamos que tenemos la siguiente información sobre la banda horizontal en la que está el objeto buscado:

— La tercera parte de las veces está en la primera o segunda banda ($m(\{h_1, h_2\}) = 1/3$).
— La tercera parte de las veces está en la primera o tercera banda ($m(\{h_1, h_3\}) = 1/3$).
— La tercera parte de las veces está en la segunda o tercera banda ($m(\{h_2, h_3\}) = 1/3$).

Como m es simétrica en todos sus elementos (la información no apunta hacia ninguna banda en particular) cabría esperar que la incertidumbre fuese máxima (igual a $L(3)$). Sin embargo, se puede comprobar que $H_*(m) = -L(1/2) = L(2)$.

Como complemento a H_*, vamos a definir una entropía que constituya una cota superior para la cantidad de incertidumbre asociada a una evidencia.

Definición. Sea m una asignación básica de probabilidad en U, llamaremos entropía superior de m al valor

$$H^*(m) = -\sum_{A \subset U} m(A)L(m(A)/|A|)$$

donde se supone $m(\phi)L(m(\phi)/(\phi) = 0$.

Las siguientes propiedades de la entropía superior son de comprobación inmediata.

1) $H^*(m)$ existe y es finita cualquiera que sea m.

2) Si m es de tipo probabilístico $H^*(m)$ coincide con la entropía de Shannon:

$$H^*(m) = -\sum_{a\in U} (\{a\})L(m\{a\}))$$

3) $H^*(m) = 0$ si y sólo si m degenera en un subconjunto unitario de U ($\exists\, a \in U$ con $m(\{a\}) = 1$).

4) Si m es una asignación básica de probabilidad en $U_1 \times U_2$ con marginales m_1 y m_2 fuertemente independientes, entonces

$$H^*(m) = H^*(m_1) + H^*(m_2)$$

5) Para toda m se verifica $0 \leqslant H_*(m) \leqslant H^*(m)$.

La siguiente proposición justifica el nombre de entropía superior dado a H^*.

Proposición. Si m es una asignación básica de probabilidad en U y p una distribución de probabilidad asociada a una evidencia de tipo probabilístico que incluya a m; entonces $H^*(m)$ es mayor o igual que la entropía de Shannon de p.

Demostración

Si la evidencia asociada a p incluye a m, entonces

$$\forall\, A \subset U,\ \exists\, t_A\colon A \to [0,1] \text{ tal que}$$

$$\sum_{a\in A} t_A(a) = m(A) \quad \text{y} \quad \sum_{B\ni a} t_B(a) = p(a)$$

Sustituyendo las anteriores expresiones en la definición de entropía superior, obtenemos

$$H^*(m) = -\sum_{A\subset U} m(A)L(m(A)/|A|) =$$

$$= -\sum_{A\subset U}\left(\sum_{a\in A} t_A(a)L(m(A)/|A|)\right).$$

Sea ahora un $A \subset U$ fijo y consideremos el problema de programación matemática en las variables $t(a)$, $a \in U$, dado por

$$\begin{array}{ll} \text{Min} & -\sum_{a \in A} t_A(a) L(t(a)) \\ \text{s.a.} & 0 \leqslant t(a) \leqslant 1 \\ & \sum_{a \in A} t(a) = m(A) \end{array}$$

Como es bien sabido, la solución de este problema es

$$t(a) = t_A(a) \quad , \quad \forall\, a \in U.$$

Por tanto, para H^* se verifica

$$H^*(m) \geqslant -\sum_{A \subset U} \left(\sum_{a \in A} t_A(a) L(t_A(a)) \right).$$

Por otra parte, y puesto que $p(a) \geqslant t_A(a)$, $\forall\, a \in A$, se tiene que

$$\begin{aligned} H^*(m) &\geqslant -\sum_{A \subset U} \left(\sum_{a \in A} t_A(a) L(t_A(a)) \right) \geqslant \\ &\geqslant -\sum_{A \subset U} \left(\sum_{a \in A} t_A(a) L(p(a)) \right) = \\ &= -\sum_{a \in U} L(p(a)) \left(\sum_{A \ni a} t_A(a) \right) = \\ &= -\sum_{a \in U} L(p(a)) p(a) = H(p). \end{aligned}$$

Como queríamos probar.

Proposición. Si m es una asignación básica de probabilidad en U, entonces

$$H^*(m) \leqslant L(|U| 2^{|U|-1})$$

alcanzándose esta cota superior para la asignación

$$m'(A) = \frac{|A|}{|U|2^{|U|-1}} \quad , \quad \forall\, A \subset U$$

Demostración

Consideremos el siguiente problema de programación en las variables $m(A)$, $A \subset U$:

$$\text{Max} \qquad - \sum_{A \subset U} m(A) L(m(A)/|A|)$$

$$\text{s.a.} \qquad m(A) \geqslant 0$$

$$\sum_{A \subset U} m(A) = 1$$

$$m(\phi) = 0$$

Al ser el objetivo una función cóncava y las restricciones lineales, la condición necesaria y suficiente para que $\bar{m}$ sea solución es que existan $\alpha \in R$ y $\lambda_A \geqslant 0$, $\forall\, A \subset U$, tales que

1) $L(\bar{m}(A)/|A|) - 1 + \alpha + \lambda_A = 0$, $\forall\, A \neq \phi$.

2) $\bar{m}(A) \geqslant 0$ y $\lambda_A . m(A) = 0$, $\forall\, A \neq \phi$.

3) $\sum_{A \subset U} \bar{m}(A) = 1$.

4) $\bar{m}(\phi) = 0$.

Despejando $\bar{m}(A)$ en 1), obtenemos si $A \neq \phi$.

$$\bar{m}(A) = k|A| \operatorname{Exp}(-\lambda_A) \quad , \quad \text{donde } k = \exp(1 - \alpha) > 0$$

de modo que $\bar{m}(A) > 0$ y $\lambda_A = 0$.

Teniendo ahora en cuenta 3), 4) y la igualdad $\sum_{r=0}^{n} r \binom{n}{r} = n2^{n-1}$ se obtiene

$$k = \frac{1}{\sum_{A \subset U} |A|} = \frac{1}{\sum_{r=0}^{|U|} r \binom{|U|}{r}} = \frac{1}{|U|2^{|U|-1}}$$

y, por tanto,

$$\bar{m}(A) = \frac{|A|}{|U|2^{|U|-1}} \quad , \quad \forall A \subset U$$

que coincide con la asignación m' del enunciado.

Consideremos ahora una asignación cualquiera m; se verifica

$$H^*(m) \leqslant H^*(m') = - \sum_{A \subset U} \frac{|A|}{|U|2^{|U|-1}} L\left(\frac{|A}{|U|2^{|U|-1}} \cdot \frac{1}{|A|}\right) =$$
$$= L(|U|2^{|U|-1}),$$

con lo que concluye la demostración.

En el caso del ejemplo 3 la entropía superior vale

$$H^*(m) = -(1/3 + 1/3 + 1/3)L(1/6) = L(6) > L(3) > L(2) = H_*(m)$$

En el caso del ejemplo 2 tenemos

$$H^*(m) = H^*(m_1) + H^*(m_2) =$$
$$= -0{,}7L(0{,}7) - 0{,}3L(0{,}3) - 0{,}1L(0{,}1/2) - 0{,}2L(0{,}2/3) > H_*(m).$$

4. CONCLUSION

En este trabajo hemos estudiado distintas medidas de la cantidad de incertidumbre para la Teoría de la Evidencia. En primer lugar, hemos considerado las medidas de especificidad y entropía introducidas por Yager (1983) para pasar a las que hemos denominado entropía inferior (H_*) y entropía superior (H^*). Con ellas intentamos resumir en un único valor la incertidumbre total asociada a una evidencia.

Pensamos, sin embargo, que aún quedan numerosas investigaciones por realizar en este campo. En particular, se podrían considerar otras medidas como

$$H_{\mathrm{Bel}}(m) = - \sum_{A \subset U} m(A)L(\mathrm{Bel}\,(A)/|A|)$$
$$H_Q(m) = - \sum_{A \subset U} m(A)L(Q(A)/|A|)$$

donde Q viene dada por

$$Q(A) = \sum_{B \supset A} \cdot\ m(B).$$

Sería entonces necesario un estudio comparativo entre ellas y, en general, la determinación de un sistema de axiomas para caracterizar la familia de medidas de incertidumbre en la Teoría de la Evidencia.

BIBLIOGRAFIA

DEMPSTER, A. P. (1967): «Upper and lower probabilities induced by a multivalued mapping», *Ann. Math. Statis.*, 38,325-339.

DUBOIS, D., y H. PRADE (1982): «On several representations of an uncertain body of evidence», en *Fuzzy Information and Decision Processes* (M. M. Gupta, E. Sánchez, eds.), North-Holland, 161-181.

DUBOIS, D., y H. PRADE (1984): «A note on measures of specificity for fuzzy sets», *BUSEFAL*, 19,83-89.

GIL, P. (1981): *Teoría Matemática de la Información*, ICE, Madrid.

KAMPE DE FERIET, J. (1974): «La theorie generalisee de l'information et la measure subjetive de l'information», en *Theories de l'information* (J. Kampe de Feriet, F. Picard, eds.), Springer-Verlag, 1-28.

HIGASHI, M., y G. J. KLIR (1983): «Measures of uncertainty and information based on possibility distributions», *Int. J. General Systems*, 9,43-58.

LAMATA, M. T. (1986): *Modelos de Decisión con Información General.* Tesis Doctoral, Universidad de Granada.

MORAL, S. (1985): *Informaciones Difusas, relaciones entre Posibilidad y Probabilidad.* Tesis Doctoral, Universidad de Granada.

SHAFER, G. (1976): *A Mathematical Theory of Evidence*, Princeton University Press.

SUGENO, M. (1974): *Theory of Fuzzy Integrals and its Applications.* Ph. Thesis, *Tokyo Institute of Tecnology.*

YAGER, R. R. (1983): «Entropy and specificity in a Mathematical Theory of Evidence», *Int J. General Systems*, 9,249-260.

ZADEH, L. A. (1965): «*Fuzzy Sets*», *Information and Control*, 8,338-353.

ZADEH, L. A. (1978): «Fuzzy Sets as a basis for a Theory of Possibility», *Fuzzy Sets and Systems*, 1,3-28.