Distancias e Indices Parciales de Medidas Difusas

Luis Daniel Hernández Molinero
Dpto. Informática y Sistemas
Universidad de Murcia
e-mail: ldaniel@dif.um.es

Antonio Salmerón Cerdán
Dpto. Estadística y Matemática Aplicada
Universidad de Almería
e-mail: asc@stat.ualm.es

Resumen

Un elemento fundamental en la teoría de las medidas difusas es el estudio de índices que determinen la calidad de ésta (i.e. medidas de disonancia, confusión o discordancia). Dichos estudios se centran principalmente en al teoría de la posibilidad y de la evidencia, y todos ellos consideran toda la información que codifica la medida. En este trabajo proponemos dos técnicas para determinar la probabildad más cercana a una medida difusa general así como el establecimiento de un conjunto de índices de (in)certidumbre asocidos a estas medidas. En su desarrollo establecemos el concepto de distancia entre medidas difusas considerando sólo una parte de la información que éstas suministran.

1 Introducción

La pérdida de aditividad, al ampliar el campo de las medidas de probabilidad a las medidas difusas en general, dota de una especial relevancia al análisis de la medida de subconjuntos no unitarios del referencial [11, 7, 10]. En efecto, una medida de probabilidad queda caracterizada completamente por la medida de los conjuntos de cardinal uno, lo que no siempre ocurre con las medidas no aditivas, en las que, para su determinación, generalmente es necesario conocer la medida de todos los subconjuntos. Precisamente el estudio de la relación entre la medida de subconjuntos de distinto cardinal se convierte en elemento clave para la comprensión de las propiedades de una medida difusa.

Un elemento fundamental en la teoría de las medidas difusas es la búsqueda de índices que caractericen determinadas propiedades de la(s) medida(s) en estudio y nos permitan comparar su grado de cumplimiento; entre tales índices destacan los de entropía, especificidad, no aditividad o similaridad entre medidas [12, 5, 6, 8, 9]. Algunas de estas propiedades pueden estudiarse mediante extensiones directas o conceptuales de los conocidos para medidas de probabilidad; sin embargo otras características de las medidas difusas requieren enfoques nuevos, entre los que cabe destacar el propuesto por De Campos [2, 3] que se basa en el establecimiento de una distancia entre dos medidas difusas, y del que se derivan conceptos como el de probabilidad mas próxima, índice de no aditividad, índice de certidumbre e incertidumbre, etc...

Todos los conceptos e índices propuestos hasta el momento atienden a la globalidad de la medida difusa (ver p.e. [1]). Nos proponemos aquí la definición de índices que consideren de forma diferenciada la medida de los subconjuntos de distinto cardinal, en un intento de captar mejor las características de una medida difusa sobre un referencial finito. En este sentido, y utilizando el enfoque de las distancias antes mencionado definiremos la probabilidad mas próxima a la familia de subconjuntos de un cardinal determinado, así como índices de certidumbre e incertidumbre de la medida difusa en esas familias.

2 Definiciones Generales

Consideraremos siempre un conjunto finito $X = \{x_1, \ldots, x_n\}$ y el álgebra $\mathcal{A} = \mathcal{P}(X)$.

Definición 1 Sea $g : \mathcal{A} \longrightarrow [0,1]$. Se dice que g es una medida difusa sobre $(X, \mathcal{A})$ sii

1. $g(\emptyset) = 0$, $g(X) = 1$
2. Si $A, B \in \mathcal{A}$ y $A \subseteq B$, entonces $g(A) \leq g(B)$

Definición 2 Sea g una medida difusa. Se define la medida dual de g, y la notaremos por g^*, a una función $g^* : \mathcal{A} \to [0,1]$ definida por:

$$g^*(A) = 1 - g(\bar{A}) \quad \forall A \subseteq X$$

Una clase especialmente inportante de medidas difusas son las generadas por una evidencia. El concepto fundamental es el de asignación básica de probabilidad (a.b.p.).

Definición 3 Una aplicación $m : \mathcal{P}(X) \longrightarrow [0,1]$ es una a.b.p. sii $m(\emptyset) = 0$ y $\sum_{A\subseteq X} m(A) = 1$

Sin embargo las a.b.p. no son medidas difusas ya que no son funciones de conjunto monótonas sino tan sólo funciones que asignan "masas de creencia" sobre determinados conjuntos. Sin embargo, dada una a.b.p. podemos construir dos medidas difusas duales.

Definición 4 La medida de creencia asociada a una a.b.p. m, es una aplicación $Bel : \mathcal{P}(X) \longrightarrow [0,1]$ definida por: $Bel(A) = \sum_{B\subseteq A} m(B)$

Su medida dual recibe el nombre de medida de plausibilidad y viene dada por la expresión: $Pl(A) = \sum_{B\cap A\neq\emptyset} m(B) = 1 - Bel(\bar{A})$

Definición 5 Llamaremos probabilidades asociadas a una medida difusa g, a las $n!$ probabilidades P_σ dadas por:

$P_\sigma(x_{\sigma(1)}) = g(\{x_{\sigma(1)}\})$
$\ldots$
$P_\sigma(x_{\sigma(i)}) = g(\{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\}) - g(\{x_{\sigma(1)}, \ldots, x_{\sigma(i-1)}\})$
$\ldots$
$P_\sigma(x_{\sigma(n)}) = 1 - g(\{x_{\sigma(1)}, \ldots, x_{\sigma(n-1)}\})$

Para cada $\sigma = (\sigma(1), \ldots, \sigma(n)) \in S_n$ con S_n el conjunto de las permutaciones de n elementos.

Es claro que dada una medida difusa g sobre un referencial X, ésta define un único conjunto de medidas de probabilidad. Sin embargo, el conocimiento de las probabilidades asociadas a g, no siempre permite reconstruir la medida difusa g ya que distintas medidas pueden generar el mismo conjunto de probabilidades asociadas. Sin embargo, si conocemos exactamente de qué permutación procede cada probabilidad asociada entonces sí es posible reconstruir la medida difusa [2, 3]. Así, una medida difusa puede identificarse con $n!$ puntos en el espacio n-dimensional:

$$g = \{(P_\sigma(x_1), \ldots, P_\sigma(x_n));\ \sigma \in S_n\}$$

Así, al existir una relación biunívoca entre los puntos correspondientes a dos medidas difusas, dada en términos de la permutación correspondiente, De Campos propone definir la distancia entre dos medidas difusas mediante el establecemiento de métricas que vengan determinadas por las distancias que se definan entre las probabilidades asociadas de las dos medidas. Para definir una distancia entre probabilidades, se consideran que éstas son puntos de $I\!R^n$ y por tanto se puede emplear cualquier distancia usual de este espacio. Más concretamente:

Definición 6 Se define la distancia entre dos medidas difusas g y g' sobre X, con probabilidades asociadas P_σ y P'_σ ($\sigma \in \mathcal{R}^n$) respectivamente, como:

$$S_q(g,g') = \sqrt[q]{\frac{1}{2\times n!}\sum_{\sigma\in S_n}\sum_{i=1}^{n}|p_{\sigma_i} - p'_{\sigma_i}|^q} \quad (1)$$

$$S_v(g,g') = \frac{1}{2\times n!}\sum_{\sigma\in S_n}\sum_{i=1}^{n}|p_{\sigma_i} - p'_{\sigma_i}| \quad (2)$$

$$S_m(g,g') = \max_{\sigma\in S_n}\max_{1\leq i\leq n}|p_{\sigma_i} - p'_{\sigma_i}| \quad (3)$$

donde $p_{\sigma_i} = P_\sigma(x_i)$ y $p'_{\sigma_i} = P'_\sigma(x_i)$ $\forall\sigma \in S_n$ con $i = 1, 2, \ldots, n$

Obsérvese cómo las expresiones (1), (2) y (3) responden a las distancias de Minkowski, valor absoluto y del máximo respectivamente salvo constantes; constantes que se introducen con objeto de que $S_.(g,g') \leq 1$.

Nótese también que el cálculo de las distancias está basado en las diferencias $|p_i - p'_i|$. Puesto que trabajamos con $I\!R^n$ y estamos considerando todas las posibles permutaciones de S_n, en general serán necesarias calcular $n \times n!$ diferencias del tipo $|p_i - p'_i|$, cálculo excesivo con n no necesariamente muy grande (p.e. si $n = 10$, $n \times n! = 36.288.000$). Además, para el cálculo de distancias, se utiliza un tipo muy particular de medidas difusas, las probabilidades, independientemente del tipo o "status" que tengan las medidas que se consideran (que pueden ser probabilidades, evidencias, medidas representables, capacidades, etc).

Nuestro método alternativo para calcular la distancia entre dos medidas difusas usa tan sólo los $2^n - 2$ valores que definen a cualquier medida sobre $\mathcal{P}(X)$, sin necesidad de recurrir a ningún otro tipo de "medida difusa auxiliar", lo que permite obviar los dos incovenientes citados.

3 Distancias Basadas en Subconjuntos

Para nuestro desarrollo usaremos la siguiente notación. Dado un referencial finito $X = \{x_1, x_2, \ldots, x_n\}$, el conjunto de todos los subconjuntos de X con cardinal m será denotado por A_m. Ya que existen $k = \begin{pmatrix} n \\ m \end{pmatrix}$ subonjuntos de cardinal m, A_m viene dado por la familia $\{A_m^p\}_{p=1}^k$ donde A_m^p denota al p-ésimo subconjunto de X con cardinal m. En lo que sigue, supondremos que, si $A_m^p = \{x_{p_1}, x_{p_2}, \ldots, x_{p_m}\}$ entonces $p_i < p_j$ si $i < j$, lo que nos permite construir el siguiente orden entre los subcojuntos del referencial X.

Dado un conjunto A_m^p consideremos el vector que definen los subíndices de los elementos que lo componen, esto es, si $A_m^p = \{x_{p_1}, x_{p_2}, \ldots, x_{p_m}\}$ denotaremos por

a_m^p al vector $(p_1, p_2, \ldots, p_m) \in I\!R^m$. Con esta notación definimos las siguientes relaciones:

Definición 7
Dados dos vectores $a_m^p = (p_1, p_2, \ldots, p_m) \in I\!R^m$ y $a_l^{p'} = (p'_1, p'_2, \ldots, p'_l) \in I\!R^l$ con $p_i < p_j$ y $p'_i < p'_j$ para todo $i < j$; definimos la siguiente relación:

$$a_m^p \prec^* a_l^{p'} \iff \begin{cases} m < l \\ \\ si\ m = l \implies a_m^p < a_m^{p'} \\ \text{Según el orden lexicográfico} \end{cases}$$

Definición 8 Diremos que $A_m^p \prec A_m^{p'} \iff a_m^p \prec^* a_m^{p'}$

Ejemplo 9 Sea $X = \{x_1, x_2, x_3\}$, entonces:

$$\begin{array}{llll} A_0^1 = \emptyset & A_1^1 = \{x_1\} & A_2^1 = \{x_1, x_2\} & A_3^1 = \{x_1, x_2, x_3\} \\ & A_1^2 = \{x_2\} & A_2^2 = \{x_1, x_3\} & \\ & A_1^3 = \{x_3\} & A_2^3 = \{x_2, x_3\} & \end{array}$$

y según la relación anterior, se tiene que:

$$A_0^1 \prec A_1^1 \prec A_1^2 \prec A_1^3 \prec A_2^1 \prec A_2^2 \prec A_2^3 \prec A_3^1$$

□

3.1 Distancias basadas en la medida de todos los subonjuntos

Una vez establecido un orden en los elementos de $\mathcal{P}(X)$ y teniendo en cuenta que siempre $g(\emptyset) = 0$ y $g(X) = 1$, una medida difusa podemos identificarla como un punto de $I\!R^{2^n-2}$. Así, si dos medidas difusas vienen determinadas por:

$$g \equiv (g(A_1^1), \ldots, g(A_i^1), \ldots, g(A_i^{k_i}), \ldots, g(A_{n-1}^{k_{n-1}}))$$

$$g' \equiv (g'(A_1^1), \ldots, g'(A_i^1), \ldots, g'(A_i^{k_i}), \ldots, g'(A_{n-1}^{k_{n-1}}))$$

con $k_i = \left(\begin{array}{c} n \\ i \end{array} \right)$ y $\sum_{i=1}^{n-1} k_i = 2^n - 2 = s$; el criterio más sencillo para establecer una distancia entre g y g' sería considerar simplemente la distancia entre los dos puntos que definen en $I\!R^s$, es decir, usar la comparación de los valores de g y g' en cada subconjunto A_m^p de X.

Definición 10 Definimos las distancia entre dos medidas difusas g y g' como:

$$\tilde{d}_q(g, g') = \sqrt[q]{\frac{1}{s} \sum_{A \subset X} |g(A) - g'(A)|^q} \quad [q \geq 2] \quad (4)$$

$$\tilde{d}_v(g, g') = \frac{1}{s} \sum_{A \subset X} |g(A) - g'(A)| \quad (5)$$

$$\tilde{d}_m(g, g') = \max_{A \subset X} |g(A) - g'(A)| \quad (6)$$

Obsérvese que dichas definiciones no son más que la distancias de Minkowski, del valor absoulto y del máximo, respectivamente, en el espacio s-dimensional. Las tres distancias oscilan entre 0 y 1. Obsérvese también que tan sólo se realizan $2^n - 2$ diferencias del tipo $|g - g'|$ frente a las $n \times n!$ diferencias necesarias si se utilizan las probabilidades asociadas.

3.2 Distancias basadas en la medida de familias de subonjuntos

Si bien las distancias que acabamos de definir usan directamente la información que suministran las medidas, pensamos que no son "buenas definiciones" en el sentido de que no establecen distinción alguna entre cada una de las coordenadas de g y g' en $I\!R^s$. Más concretamente, si las medidas difusas fuesen en particular medidas de probabilidad, éstas vienen determinadas por la medida sobre los conjuntos unitarios ya que, por la propiedad de aditividad, se puede conocer la medida sobre los conjuntos de cardinal mayor; es decir, estos últimos **no** suministran nueva información. Esto repercute en el cálculo de las diferencias $|g - g'|$. En efecto, es fácil comprobar que

$$|P(A) - P'(A)| = |P(\bar{A}) - P'(\bar{A})|$$

Es decir, la diferencia de los conjuntos de cardinal 1 coincide con el de los conjuntos de cardinal $n-1$, la diferencia de los conjuntos de cardinal 2 coincide con los de cardinal $n-2$, etc... y puesto que ambas diferencias (las de los conjuntos de cardinal m y las de los conjuntos de cardinal $n - m$) proceden de las difencias de las de cardinal 1, por la aditividad, tan sólo tienen interés las diferencias en los elementos atómicos. En consecuencia cabría plantearse que sentido tiene el calcular todas las diferencias $|g - g'|$ en el caso probabilístico.

Por otro lado, si consideramos dos medidas difusas g y g' cualesquiera, es necesario distinguir los conjuntos de distinto cardinal, pues la información sobre singletones no predice la información de los de cardinal superior (lo único que se verifica es la monotonía). Por ejemplo, si decimos que una medida g la interpretamos como el grado subjetivo de ocurrencia de un suceso y es tal que $g(x) = .9$ y que $g(x, y) = .901$, estamos reflejando el hecho de que tenemos un gran optimismo en que se dé $\{x\}$ y $\{x, y\}$, pero ello no quiere decir que necesariamente seamos pesimistas en la ocurrencia de $\{y\}$ (una probabilidad establecería que $g(y) = .001$). Es por ello, que tampoco nos basta considerar que una medida g es tan sólo un vector de $I\!R^s$ en el que todas sus componentes tienen la misma importancia, sino como un vector formado por distintas componentes donde cada una trabaja sobre conjuntos del mismo cardinal (i.e. son de la misma categoría), que pueden considerarse "comparables". El siguiente ejemplo nos permite aclarar esta idea.

Ejemplo 11 Consideremos las medidas g y g' definidas por:

B_i	g	g'	$\lvert g-g'\rvert$
$\{x_1\}$	0	.9	.9
$\{x_2\}$	.1	.6	.5
$\{x_3\}$	.2	.7	.5
$\{x_1, x_2\}$	.9	.9	0
$\{x_1, x_3\}$	.9	.9	0
$\{x_2, x_3\}$	.7	.7	0

De un modo intuitivo, cabría establecer un criterio en el que se reflejara que las medidas son muy distintas en conjuntos de cardinal 1, ya que se diferencian casi en el máximo valor posible (i.e. uno) en la coordenada del x_1 y en la mitad de este valor en el resto; es decir, g se diferencia de g' en "la dirección" de $\{x_1\}$ en 0.9, y en las "direcciones" de $\{x_2\}$ y $\{x_3\}$ en 0.5. Sin embargo, ambas medidas verifican que $g = g'$ para conjuntos de cardinal 2, por lo que considerando conjuntos de este tamaño ambas son indistinguibles. □

Con la notación introducida, es claro que $\mathcal{P}(X) = \{A_m^p /\ p = 1, \ldots, k\ ;\ \ m = 0, \ldots, n\}$ y que para cada posible cardinal m, una medida difusa define un único punto en $I\!R^k$, con $k = \begin{pmatrix} n \\ m \end{pmatrix}$, dado por:

$$g(A_m) = (g(A_m^1), g(A_m^2), \ldots, g(A_m^p), \ldots, g(A_m^k)) \tag{7}$$

Notar que con el orden establecido anteriormente estamos imponiendo el modo en que deben venir ordenados los valores de las medidas difusas al calcular las distancias.

En estas condiciones, vamos a definir distancias entre dos medidas difusas, en primer lugar, en términos de cada subconjunto de un cardinal determinado, para, posteriormente, definir distancias globales entre ellas considerando los distinos cardinales.

Definición 12 Consideremos dos medidas difusas g y g' definidas sobre el mismo referencial finito X. Definimos la distancia de g a g' en la dirección de A_m^p como

$$d_{A_m^p}(g, g') = |g(A_m^p) - g'(A_m^p)|$$

donde $|.|$ denota el valor absoluto en $I\!R$.

Si consideramos ahora todos los conjuntos de un mismo cardinal m,

$$A_m = (A_m^1, A_m^2, \ldots, A_m^p, \ldots, A_m^k)$$

podemos extender la definición anterior a **distancias entre vectores de conjuntos** del mismo cardinal:

Definición 13 Consideremos dos medidas difusas g y g' definidas sobre el mismo referencial finito X de cardinal n. Definimos la distancia entre g y g' para el cardinal m (o a nivel m) como:

$$d_{A_m}(g, g') = \|g(A_m) - g'(A_m)\|$$

donde $g(A_m)$ está definida según (7) y $\|.\|$ denota cualquiera de las normas de $I\!R^k$.

Por último, como $\mathcal{P}(X) = \{A_0 = \emptyset, A_1, \ldots, A_{n-1}, A_n = X\}$, se puede calcular la distancia entre

$$\begin{aligned} g &\equiv (g(A_0), g(A_1), \ldots, g(A_{n-1}), g(A_n)) \\ g' &\equiv (g'(A_0), g'(A_1), \ldots, g'(A_{n-1}), g'(A_n)) \end{aligned}$$

donde $g(A_m)$ está definida según (7); es decir, calcular la distancia global de g y g'.

En este punto las definiciones de esta última distancia son múltiples, atendiendo a la importancia o preferencia que uno estime entre $g(A_i)$ y $g'(A_i)$. Por ejemplo, podemos identificar cada vector $g(A_i)$ con algún valor promedio α_i y establecer la distancia entre g y g' como la distancia en $I\!R^n$ entre los puntos $(\alpha_1, \ldots, \alpha_n)$ y $(\alpha'_1, \ldots, \alpha'_n)$. Particularmente, creemos que lo mejor sería establecer algún criterio de media ponderada sobre cada una de las componentes, es decir, considerar las distancias respecto a cada uno de los posibles cardinales y establecer una media. Así, por ejemplo, se pueden establecer las distancias:

$$\begin{aligned} d_1(g, g') &= \frac{1}{n} \sum_{m=1}^{n} d_{A_m}(g, g') \\ d_2(g, g') &= \sum_{m=1}^{n} \frac{1}{m} d_{A_m}(g, g') \end{aligned}$$

4 Probabilidad Más Próxima a una Medida Difusa

Puesto que disponemos ya de una forma de medir el parecido entre medidas difusas (por medio de las distancias d_{A_m}), y una probabilidad es un caso particular de ellas, podemos plantearnos hallar la probabilidad mas próxima a una medida difusa dada, lo que permite:

- Si la aproximación de una medida difusa a su probabilidad más cercana es suficientemente buena, se podría trabajar con dicha probabilidad en lugar de con la medida difusa, lo que en general será más eficiente.
- Establecer índices de aditividad. Si evaluamos la distancia entre una medida difusa y la probabilidad más cercana, disponemos de un valor que cuantifica la desviación de esa medida respecto de un comportamiento aditivo.

Si definimos una distancia única, podemos plantearnos el conocer la probabilidad más cercana a cierta medida g y en consecuencia medir en cierto modo la aditividad de la medida [5, 2]. Sin embargo, con el planteamiento expuesto, no tiene sentido el cálculo de la probabilidad más cercana a la medida difusa considerada, pero sí el de las probabilidades más cercanas *en términos de conjuntos de distinto cardinal.* Presentamos para esta tarea dos planteamientos:

• **Primer Planteamiento: Exigencia de aditividad según un cardinal considerado.**

Consideremos una medida difusa g, y supongamos que la suma de la medida en los conjuntos de cardinal m es t_m, es decir:

$$g(A_m^1) + \cdots + g(A_m^{p_m}) + \cdots + g(A_m^{k_m}) = t_m \qquad (8)$$

donde k_m es el número de subconjuntos de cardinal m.

Si la medida fuese una probabilidad, entonces, por su comportamiento aditivo se verificaría:

$$P(A_m^1) + \cdots + P(A_m^{p_m}) + \cdots + P(A_m^{k_m}) = \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} = r_m \qquad (9)$$

Por tanto, si consideramos los subconjuntos de cardinal m, y el vector asociado:

$$g(A_m) = (g(A_m^1), g(A_m^2), \ldots, g(A_m^{p_m}), \ldots, g(A_m^{k_m}))$$

y si g fuese una probabilidad debería de verificarse que $t_m = r_m = \frac{(n-1)!}{(m-1)!\ (n-m)!}$ [ver (8) y (9)]. Puesto que en general dicha igualdad no se verifica, si "exigimos" un comportamiento aditivo en la medida, nos motiva la siguiente definición:

Definición 14 Definimos la probabilidad más cercana a la medidad difusa g, y la notamos por P_g, como aquella que viene determinada por el sistema de ecuaciones:

$$\left\{ \begin{array}{lll} \sum_{x \in A_m^1} p(x) = & P_g(A_m^1) & = g(A_m^1)\dfrac{r_m}{t_m} \\ \vdots & \vdots & \vdots \\ \sum_{x \in A_m^{k_m}} p(x) = & P_g(A_m^{k_m}) & = g(A_m^{k_m})\dfrac{r_m}{t_m} \end{array} \right.$$

$$con \quad \sum_{x \in X} p(x) = 1$$

Cabe llamar la atención al caso particular en que $m = 1$ (i.e. conjuntos unitarios), ya que si $t_1 = 1$ la probabilidad más cercana a g es ella misma pues en este caso g es una probabilidad (en los conjuntos unitarios) y por tanto:

$$P_g(x_i) = g(\{x_i\}) = g(A_1^i)$$

y en el caso de que $t_1 \neq 1$ la probabilidad más cercana es aquella que obtendríamos al realizar una normalización usual:

$$P_g(x_i) = g(\{x_i\}) \frac{1}{\sum_i g(\{x_i\})}$$

• **Segundo Planteamiento: Determinación de la probabilidad más próxima a partir del convexo asociado a la medida** (según una distancia considerada).

Puesto que en general, para un cierto cardinal m, $\sum_{l=1}^{k_m} g(A_m^l) = t_m \neq r_m$, podemos conseguir que esta igualdad se cumpla añadiendo o restando la cantidad necesaria en cada una de las componentes, $g(A_m^l)$, de $g(A_m)$; con ello conseguimos k_m probabilidades asociadas a la medida g, $\mathcal{P} = \{P_1, \ldots, P_{k_m}\}$, y definidas por:

$$P_l \equiv (g(A_m^1), \ldots, p_l, \ldots, g(A_m^{k_m}))$$

con $p_l \in \mathcal{R}$ tal que

$$g(A_m^1) + \cdots + p_l + \cdots + g(A_m^{k_m}) = r_m \qquad (10)$$

Como en el planteamiento anterior, la idea responde a considerar un cardinal m, para el cual si g fuese aditiva verificaría que el punto $g(A_m) \in I\!R^{k_m}$ se encontraría en el hiperplano $\sum g(A_m^p) = r_m$. Si no lo está entonces fijamos todas las coordenadas menos una (p.e. $g(A_m^l)$) y ésta la "prolongamos" hasta llegar al hiperplano dado por (10). Al repetir este proceso para todas las coordendas de $g(A_m)$, se obtiene el conjunto de probabilidades $\mathcal{P}$ que constiutyen los puntos extremos del convexo que definen.

Determinado este convexo podemos considerar como probabilidad más cercana P aquella que minimice la distancia d_{A_m} con g (para los conjuntos de cardinal considerado). Es decir, P viene dada por la resolución del siguiente programa:

$$\min_{P \in \mathcal{P}} d_{A_m}(g, P)$$

s.a.

$$\min_{P_l \in \mathcal{P}} P_l(A_m^l) \leq P(A_m^l) \leq \max_{P_l \in \mathcal{P}} P_l(A_m^l)$$

$$\sum_{l=1}^{k_m} P(A_m^l) = r_m$$

5 Indices de Certidumbre e Incertidumbre

La determinación del grado de (in)certidumbre de una medida puede ser de gran ayuda en determinados problemas. Por ejemplo, en sistemas de razonamiento bajo incertidumbre, es frecuente la imposibilidad de manejar toda la información disponible sobre ciertas variables en el proceso de inferencia. En estos casos, habrá que recurrir a métodos aproximados donde debe elegirse qué información es realmente informativa o no. En el caso de probabilidades, el problema se trató en [4].

Existen dos tipos de índices clásicos que permite determinar la calidad de la medida: la entropía y la especificidad [12, 10]. La combinación de ambos índices proporciona una idea de la calidad de la medida. La principal limitación es que sólo sirve para evidencias y no se puede extender a medidas difusas en general. De Campos [2] presenta otras alternativas totalmente diferentes para definir índices de incertidumbre para medidas difusas, que se basan en la distancia entre medidas. Ello es posible si se tiene en cuenta que las medidas que proporcionan más incertidumbre son las de ignorancia:

$$Bel_0(A) = \begin{cases} 0 & si\ A \neq X \\ 1 & si\ A = X \end{cases} \qquad Pl_0(A) = \begin{cases} 1 & si\ A \neq \emptyset \\ 0 & si\ A = \emptyset \end{cases} \tag{11}$$

y las que proporcionan mayor certidumbre son las medidas de Dirac.

De este modo, considerando las distancias definidas en este trabajo, podemos establecer índices atendiendo al cardinal del tipo de subconjuntos en el que estemos interesados. Así se puede definir un índice de certidumbre en términos de conjuntos de distinto cardinal midiendo la distancia entre la medida g y una medida de ignorancia: a mayor distancia, menos incertidumbre genera g.

Puesto que tenemos dos codificaciones de la ignorancia, Bel_0 y Pl_0, cabría considerar, en principio, las distancias $d(g, Bel_0)$ y $d(g, Pl_0)$, puesto que dada cualquier medida g existe una dual g^* y, en términos muy generales, podemos decir que podemos dividir las medidas en optimistas y pesimistas, consideraremos la medida de ignorancia que más se aproxime a g (i.e. la mínima de las distancias).

Definición 15 Sea g una medida difusa sobre el referencial $X = \{x_1, \ldots, x_n\}$.Definimos el índice de certidumbre de g a nivel m como:

$$C_m(g) = \min(d_{A_m}(g, Bel_0), d_{A_m}(g, Pl_0))$$

con d_{A_m} según la definición 13.

De forma análoga podemos definir un índice de incertidumbre en términos de conjuntos de distinto cardinal midiendo la distancia entre la medida g y las probabilidades degeneradas $P_i(x_j) = \delta_{i,j},\ i = 1, \ldots, n$ ya que éstas son las que menos incertidumbre provocan: a mayor distancia, más incertidumbre proporciona g. Como disponemos de n medidas de Dirac sobre un referencial de cardinal n, parace lógico establecer como índice de incertidumbre la menor de las distancias entre g y P_j.

Definición 16 Sea g una medida difusa sobre el referencial X. Definimos el índice de incertidumbre a nivel m como:

$$I_m(g) = \min_{1 \leq j \leq n} d_{A_m}(g, P_j)$$

Teniendo en cuenta que es equivalente codificar una información por una medida difusa g que por su dual g^*, cabe esperar que el índice de (in)certidumbre definido mida el mismo grado independientemente de la medida utilizada. En efecto, es fácil demostrar que:

Proposición 17 *Dada una pareja de medidas difusas (g, g^*) sobre un referencial de cardinal n se verfica:*

$$\begin{aligned} d_{A_m}(g, Pl_0) &= d_{A_{n-m}}(g^*, Bel_0) \\ d_{A_m}(g, P_j) &= d_{A_{n-m}}(g^*, P_j) \end{aligned}$$

donde Pl_0y Bel_0 vienen dadas por (11) y P_j es la distribución de probabilidad que vale 1 en x_j y cero en el resto.

Y, en consecuencia:

$$C_m(g) = \min(d_{A_m}(g, Bel_0), d_{A_m}(g, Pl_0)) =$$

$$\min(d_{A_{n-m}}(g^*, Pl_0), d_{A_{n-m}}(g^*, Bel_0)) = C_{n-m}(g^*)$$

$$I_m(g) = \min_{1 \leq j \leq n} d_{A_m(g,P_j)} =$$

$$\min_{1 \leq j \leq n} d_{A_{n-m}(g^*,P_j)} = I_{n-m}(g^*)$$

Es decir, fijado cierto nivel m, el grado de (in)certidumbre que establece la medida g coincide con con el grado de (in)certidumbre que establece la medida g^* a nivel $n - m$. De esta forma, utilizando los índices de certidumbre o incertidumbre puede medirse, de forma parcial y detallada, la información que sobre la verificación de un predicado nos dé la medida difusa.

6 Conclusiones y estudios futuros

En este trabajo hemos introducido un nuevo enfoque para la caracterización de medidas difusas. La idea principal consiste en considerar que el comportamiento de una medida difusa puede ser radicalmente distinto dependiendo de los subconjuntos del referencial sobre los que se esté evaluando. En ese sentido, hemos definido nuevas medidas de distancia entre medidas difusas atendiendo al comportamiento de las medidas en los distintos subconjuntos del referencial, para un cardinal dado. Usando estas distancias, definimos dos índices de certidumbre e incertidumbre, de los cuales se estudian algunas propiedades. Así mismo, actualizamos el concepto de probabilidad más cercana a una medida difusa. Pensamos que el presente trabajo puede servir para una mejor comprensión del comportamiento de las medidas difusas.

Como futuros trabajos, podría estudiarse el comportamiento de las distancias e índices definidos para clases concretas de medidas difusas. Este estudio puede ser tanto teórico como experimental, seleccionando una batería de medidas y calculando sobre ellas las distancias y los índices de (in)certidumbre.

Referencias

[1] Machine intelligence and pattern recognition. In B.M Ayyrub, M.M. Gupta, L.N. Kanan, eds. *Analysis and Management: Theory and Applications*, vol. 13. North Holland, 1992.

[2] L.M. de Campos. *Caracterización y estudio de medidas e integrales difusas a partir de probabilidades.* PhD thesis, Dpto. Ciencias de la Computación e I.A. U. de Granada, 1987.

[3] L.M. de Campos, M.T. Lamata, S. Moral. Distances between fuzzy measures through associated probabilities: some applications. *Fuzzy Sets and Systems*,(35):57-68, 1990.

[4] J.E. Cano, L.D. Hernández, S. Moral. Importance sampling algorithms for belief networks. To appear in *International J. of App. Reasoning*, 1996.

[5] D. Dubois, H. Prade. Unfair coins and necessity measures: towards a possibilistic interpretation of histograms. *Fuzzy Sets and Syst.*, 10:15-20, 1983.

[6] D. Dubois, H. Prade. A note on measures of specificity for fuzzy sets. *Busefal*, 19:83-89, 1984.

[7] G.J. Klir. Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like?. *Fuzzy Sets and Systems*, (24):141-160, 1987.

[8] G.J. Klir. The role of uncertainty principles in inductive systems modelling. *Kybernetes*, 17(2):23-34, 1988.

[9] G.J. Klir. Probabilistic versus possibilistic conceptualization of uncertainty.

In *Analysis and Management of uncertainty: Theory and Applications*, p.13-25. Elsevier Science Pub., 1992.

[10] M.T. Lamata, S. Moral. Classification of fuzzy measures. *Fuzzy Sets and Syst.*, 33:243-253, 1989.

[11] M.Sugeno. *Theory of fuzzy integrals and its applications.* PhD. Thesis, Tokio Ins. of Tech., 1974.

[12] R.R. Yager. Entropy and specificity in a mathematical theory of evidence. *Int. J. of general systems*, 9:249-260, 1983.