UNA FAMILIA DE MEDIDAS DIFUSAS ASOCIADA A CADA EVIDENCIA *

José A. Herencia
Dpto. Informática y Análisis Numérico
Campus de Rabanales (Edif. C-2)
Universidad de Córdoba
jaherencia@uco.es

M. Teresa Lamata
Dpto. Ciencias de la Computación e I.A.
E.T.S. Ingeniería Informática
Universidad de Granada
mtl@decsai.ugr.es

Resumen

A partir de cualquier asignación básica de probabilidad definida sobre un conjunto finito, se genera una familia de medidas difusas a la que pertenecen la creencia y la plausibilidad propias de tal evidencia. Tanto la definición como el estudio de ciertas propiedades de dicha familia se obtiene mediante funciones crecientes del intervalo unidad en sí mismo.

Palabras clave: Asignación Básica de Probabilidad, Evidencia, Posibilidad, Probabilidad, Medida Difusa.

1 PRELIMINARES Y NOTACIÓN

En este trabajo consideramos conjuntos finitos $X = \{x_1, x_2, \ldots, x_n\}$. Denotamos por $\mathcal{P}(X)$ al conjunto de partes de X. Para cualquier subconjunto $A \in \mathcal{P}(X)$, consideramos su cardinal $|A|$ y su complementario $A' := X \setminus A$. También usaremos el conjunto $\mathbb{N}$ de números enteros positivos.

Se denomina **medida difusa** sobre X a cualquier aplicación $g : \mathcal{P}(X) \to [0,1]$ que sea creciente ($A \subseteq B \Rightarrow g(A) \leq g(B)$) y que tome los dos valores siguientes: $g(\emptyset) = 0$, $g(X) = 1$. Dada cualquier medida difusa g sobre X, su **medida dual** es la medida difusa g', definida como sigue: $g'(A) := 1 - g(A'),\ \forall A \subseteq X$. Es inmediato comprobar que $g'' = g$, por lo que toda medida difusa puede estudiarse junto a su dual; denominándose g **autodual** en caso de ser $g' = g$. La mayoría de las medidas difusas que se manejan prácticamente determinan **parejas ordenadas**, en el sentido de que g y g' son comparables: siempre es una de ellas menor o igual que la otra. Para la clasificación de los distintos tipos de medidas difusas nos remitimos a [5] y [6].

Se denomina **asignación básica de probabilidad (A.B.P.)** sobre X a cualquier aplicación $m : \mathcal{P}(X) \to [0,1]$ que verifique las dos condiciones siguientes:

$$m(\emptyset) = 0, \qquad \textstyle\sum_{A \subseteq X} m(A) = 1.$$

Denotamos por $\mathcal{M}_X$ al conjunto de asignaciones básicas de probabilidad definidas sobre el conjunto X. Dada cualquier $m \in \mathcal{M}_X$, se denominan **elementos focales** de m a los subconjuntos B de X que cumplen la condición $m(B) > 0$. Asociada a tal m se considera la pareja de medidas difusas ordenadas formada por la **creencia** y la **plausibilidad**, definidas respectivamente como sigue ($\forall A \subseteq X$):

$$Bel(A) := \sum_{B \subseteq A} m(B), \qquad Pl(A) := \sum_{B \cap A \neq \emptyset} m(B)$$

A partir de la función Bel, se puede calcular la única A.B.P. m de la que es medida de creencia, mediante la fórmula siguiente ($\forall A \subseteq X$):

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} Bel(B)$$

Por tanto, podemos obtener las tres funciones m, Bel y Pl una vez conocida cualquiera de ellas (ya que Bel y Pl son medidas duales). La información que aporta cualquiera de estas funciones es denominada una **evidencia** sobre X. Hay dos tipos de evidencia especialmente importantes:

(1). Cuando los elementos focales forman una cadena (respecto a la inclusión de conjuntos), tenemos una evidencia **consonante**, que está determinada por la **distribución de posibilidad** $\pi : X \to [0,1]$ siguiente: $\pi(x) := \sum_{x \in B} m(B)$, $\forall x \in X$. En este caso, las funciones Bel y Pl se denominan, respectivamente, medidas de **necesidad** y **posibilidad**, siendo ($\forall A \subseteq X$):

$$Bel(A) = \min_{x \notin A} (1 - \pi(x)), \qquad Pl(A) = \max_{x \in A} \pi(x)$$

La **ignorancia** [**total**] corresponde al caso particular en que toda la evidencia es asignada al conjunto X (es decir: $m(X) = 1$), no aportando ninguna información respecto a sus subconjuntos estrictos.

* Financiado por la DGICYT, Proyecto PB95-1181

(2). La evidencia es **probabilística** cuando todos los elementos focales de m son unitarios. En este caso $p(x) := m(\{x\})$, $\forall x \in X$, es una **distribución de probabilidad** y se cumple que $(\forall A \subseteq X)$:

$$Bel(A) = Pl(A) = \sum_{x \in A} p(x)$$

Se tiene **certeza** en el caso particular de que toda la evidencia esté concentrada en un solo elemento: $\exists x \in X$, $m(\{x\}) = 1$.

Cuando hay algún elemento focal B no unitario, la *"masa de probabilidad"* $m(B)$ se asigna a todo el conjunto B sin que se atribuya una cantidad determinada de la misma a cada uno de sus elementos. Esto hace que la creencia no sea aditiva (al ser para B mayor que la suma de creencias de sus elementos), por lo que no se trata de una probabilidad. No obstante, se considera a veces una probabilidad asociada a cualquier A.B.P., obtenida de la siguiente forma: la masa de probabilidad $m(B)$ de cada elemento focal se reparte uniformemente entre todos sus elementos. Así, en la distribución de probabilidad obtenida, se suma a cada elemento la probabilidad $m(B)/|B|$. En definitiva, resulta la siguiente medida de probabilidad:

$$Pw(A) := \sum_{m(B)>0} m(B) \frac{|B \cap A|}{|B|}, \quad \forall A \subseteq X. \qquad (1)$$

Partiendo de esta expresión, consideramos una generalización de la misma en la Sección 2, resultando una familia de medidas difusas que contiene (como casos extremos) la creencia y la plausibilidad. Dicha familia se obtiene mediante funciones crecientes del intervalo unidad en sí mismo; de forma que algunas propiedades de estas funciones inducen propiedades en las medidas difusas que originan, como vemos en la Sección 3. Ahí también consideramos algunos casos particulares y ejemplos, concluyendo con algunas cuestiones abiertas.

2 DEFINICIÓN DE LA FAMILIA DE MEDIDAS DIFUSAS $\{f_\rho\}_{\rho \in \mathcal{I}_0}$

Al estudiar medidas de entropía asociadas a evidencias [1, 2] aparecen bajo un mismo tratamiento las medidas difusas Bel, Pl y Pw, ya que la *aleatoriedad* de m puede determinarse por la expresión:

$$H(m) = -\sum_{m(B)>0} m(B) \log f(B). \qquad (2)$$

La cual corresponde a:

- La *medida de confusión* de Höhle [3] si $f = Bel$.
- La *medida de disonancia* de Yager [7] si $f = Pl$.
- La *medida de discordancia (o conflicto)* de Klir y Ramer [4] cuando $f = Pw$.

Esto nos ha motivado a explicitar una formulación que permita obtener una familia de medidas difusas que englobe a las tres citadas. Para ello nos basamos en los hechos siguientes. Consideremos cualquier A.B.P. m sobre X y cualquier $A \subset X$. Entonces, los elementos focales B que no cortan a A (para los cuales $|B \cap A| = 0$) se descartan en el cálculo de $Bel(A)$ y de $Pl(A)$. Por el contrario, aparecen en ambos cálculos los elementos focales B que están incluidos en A (para los cuales $|B \cap A| = |B|$). La diferencia entre $Bel(A)$ y $Pl(A)$ se debe a los elementos focales B "intermedios", que cortan a A pero no están incluidos en A (es decir, los elementos focales B para los que $0 < |B \cap A| < |B|$). En la medida de probabilidad Pw, considerada en (1), se asigna el peso $\frac{|B \cap A|}{|B|}$ a cada elemento focal B. De aquí surge la generalización consistente en transformar el peso $\frac{|B \cap A|}{|B|}$, cambiándolo por el factor $\rho\left(\frac{|B \cap A|}{|B|}\right)$, siendo $\rho : [0,1] \to [0,1]$ cualquier función creciente. Concretando, tenemos:

DEFINICIÓN 2.1 *Denominemos $\mathcal{I}$ a la familia de funciones crecientes de $[0,1]$ en $[0,1]$. Para cualquier A.B.P. $m \in \mathcal{M}_X$ y cualquier $\rho \in \mathcal{I}$, consideramos la función $f_\rho : \mathcal{P}(X) \to [0,1]$ definida por:*

$$f_\rho(A) := \sum_{B \neq \emptyset} m(B) \rho\left(\frac{|B \cap A|}{|B|}\right), \quad \forall A \subseteq X. \qquad (3)$$

PROPOSICIÓN 2.2 *Para cualquier A.B.P. $m \in \mathcal{M}_X$ y cualquier $\rho \in \mathcal{I}$, se cumple:*

1. *Las condiciones necesarias y suficientes para que f_ρ sea una medida difusa sobre X son las siguientes: $\rho(0) = 0$ y $\rho(1) = 1$. Por eso consideramos el conjunto $\mathcal{I}_0 := \{\rho \in \mathcal{I} : \rho(0) = 0,\ \rho(1) = 1\}$.*
2. *Si $\rho \in \mathcal{I}_0$, entonces la medida dual de f_ρ es:*

   $f'_\rho = f_{\rho'}$ *donde* $\rho'(x) = 1 - \rho(1-x)$, $\forall x \in [0,1]$.

DEMOSTRACIÓN:

1. Si en la ecuación (3) tomamos $A = \emptyset$ (resp. $A = X$) se obtiene la suma, extendida a todos los elementos focales B, de $\rho(0)m(B)$ (resp. $\rho(1)m(B)$). Por tanto: $f_\rho(\emptyset) = \rho(0)$ y $f_\rho(X) = \rho(1)$. Teniendo en cuenta estas igualdades, sólo falta comprobar que el crecimiento de ρ implica el crecimiento de f_ρ. En efecto, consideremos subconjuntos cualesquiera tales que $A_1 \subseteq A_2 \subseteq X$. Entonces, para cualquier $B \neq \emptyset$ se cumple que $\frac{|B \cap A_1|}{|B|} \leq \frac{|B \cap A_2|}{|B|}$. De donde, por ser ρ creciente, resulta efectivamente que $f_\rho(A_1) \leq f_\rho(A_2)$.
2. Partiendo de que f_ρ sea una medida difusa, basta

calcular su dual: $f'_\rho(A) = 1 - f_\rho(A') =$

$$= \sum_{B\neq\emptyset} m(B) - \sum_{B\neq\emptyset} m(B)\rho\left(\frac{|B\cap A'|}{|B|}\right).$$

Usando la igualdad $|B\cap A'| = |B| - |B\cap A|$ y asociando términos obtenemos: $f'_\rho(A) =$

$$= \sum_{B\neq\emptyset} m(B)\left[1-\rho\left(1-\frac{|B\cap A|}{|B|}\right)\right] = f_{\rho'}(A). \ \Box$$

COROLARIO 2.3 *Asociadas con cualquier A.B.P. $m \in \mathcal{M}_X$, podemos considerar las siguientes medidas difusas:*

1. *Para cada $\alpha \in [0,1]$, tenemos la medida difusa P_α definida por ($\forall A \subseteq X$):*

$$P_\alpha(A) := \sum_{B\subseteq A} m(B) + \alpha \sum_{\substack{B\not\subseteq A\\ B\cap A\neq\emptyset}} m(B) \qquad (4)$$

   Su medida dual es: $P'_\alpha = P_{1-\alpha}$. En particular, $P_0 = Bel$ y $P_1 = Pl$.

2. *Para cada $k \in \{2i-1 : i \in \mathbb{N}\} \cup \{\frac{1}{2i-1} : i \in \mathbb{N}\}$ tenemos la medida difusa autodual $P_{[k]}$ definida por ($\forall A \subseteq X$):*

$$P_{[k]}(A) := \sum_{B\neq\emptyset} \frac{m(B)}{2}\left[1+\left(2\frac{|A\cap B|}{|B|}-1\right)^k\right]$$

   En particular, tenemos la probabilidad $Pw = P_{[1]}$ considerada en (1).

DEMOSTRACIÓN:

1. Basta considerar la medida $P_\alpha = f_{\rho_\alpha}$, donde $\rho_\alpha \in \mathcal{I}_0$ está definida por:

$$\rho_\alpha(x) := \begin{cases} 0, & x=0 \\ \alpha, & 0<x<1 \\ 1, & x=1 \end{cases} \qquad (5)$$

   Además es inmediato comprobar que $\rho'_\alpha = \rho_{1-\alpha}$, $P_0 = Bel$ y $P_1 = Pl$.

2. En este caso se trata de la medida $P_{[k]} = f_{\rho_{[k]}}$, siendo $\rho_{[k]}(x) = \dfrac{1+(2x-1)^k}{2}$. También se comprueba fácilmente que $\rho_{[k]} \in \mathcal{I}_0$ y que $\rho'_{[k]} = \rho_{[k]}$. En el caso particular $k = 1$, se tiene simplemente que $\rho_{[1]} = Id$ es la función identidad en $[0,1]$, que origina la probabilidad Pw. $\Box$

NOTA: Es obvio que podemos considerar muchas otras familias de medidas difusas, sin más que tomar las funciones $\rho \in \mathcal{I}_0$ adecuadas. Algunos ejemplos son:

- Para cada $r \in (0,+\infty)$ podemos tomar la función $\rho(x) = x^r$, obteniendo la siguiente medida difusa:

$$P_{(r)}(A) := \sum_{B\neq\emptyset} m(B)\left(\frac{|A\cap B|}{|B|}\right)^r$$

  De nuevo tenemos el caso particular $Pw = P_{(1)}$.
- Para cada $b \in (1,+\infty)$ tenemos la medida difusa correspondiente a la función $\rho(x) = \dfrac{b^x-1}{b-1}$.
- También con $b \in (1,+\infty)$ podemos tomar $\rho(x) = \log_b[1+(b-1)x]$.

Al contrario de lo que ocurría en las familias de funciones consideradas en el corolario 2.3, estas nuevas familias de funciones no contienen a sus respectivas duales. Observemos también que siempre que $\rho \in \mathcal{I}_0$ sea estrictamente creciente, podemos considerar la función $\rho^{-1} \in \mathcal{I}_0$ (siendo éste el procedimiento que relaciona los dos últimos ejemplos). $\triangleright$

3 ALGUNAS PROPIEDADES DE LAS MEDIDAS DIFUSAS f_ρ

Vemos ahora cómo algunas propiedades de las funciones $\rho \in \mathcal{I}_0$ se transfieren a las correspondientes medidas difusas.

PROPOSICIÓN 3.1 *Para cualesquiera $\rho, \hat\rho \in \mathcal{I}_0$ se cumple:*

1. *$\forall\beta \in [0,1]$, $\beta\rho + (1-\beta)\hat\rho \in \mathcal{I}_0$, siendo $f_{\beta\rho+(1-\beta)\hat\rho} = \beta f_\rho + (1-\beta) f_{\hat\rho}$.*
2. *Si $\rho \le \hat\rho$ entonces $f_\rho \le f_{\hat\rho}$.*
3. *Si la gráfica de ρ es simétrica respecto al punto $\left(\frac{1}{2},\frac{1}{2}\right)$, entonces f_ρ es autodual.*

DEMOSTRACIÓN: Los dos primeros apartados se deducen inmediatamente de la ecuación (3). Para demostrar el tercero, al ser $f'_\rho = f_{\rho'}$ con $\rho'(x) = 1-\rho(1-x)$ (Proposición 2.2, apartado 2), basta comprobar que la simetría enunciada implica la igualdad $\rho' = \rho$. En efecto, tal simetría consiste en que $\forall u \in [-1/2, 1/2]$, $\rho(1/2+u) - 1/2 = 1/2 - \rho(1/2-u)$, lo que equivale a que $\forall u \in [-1/2,1/2]$, $1-\rho(1/2+u) = \rho(1/2-u)$. Entonces basta efectuar el cambio de variable $x = 1/2-u$ para obtener la condición buscada: $\forall x \in [0,1]$, $1-\rho(1-x) = \rho(x)$. $\Box$

Estos resultados indican la posibilidad de obtener nuevas funciones $\rho \in \mathcal{I}_0$, así como las correspondientes medidas difusas f_ρ, mediante combinaciones lineales convexas. Por ejemplo, podemos considerar las siguientes medidas ($\forall\alpha,\beta \in [0,1]$):

$$\alpha Pl + (1-\alpha)Bel = P_\alpha, \quad \beta Pw + (1-\beta)P_\alpha, \quad etc.$$

Asimismo, los resultados anteriores corroboran hechos conocidos como la igualdad $Pw' = Pw$ o las desigualdades siguientes:

$$m \leq Bel \leq Pw \leq Pl, \qquad Bel \leq f_\rho \leq Pl, \, \forall \rho \in \mathcal{I}_0.$$

De la última desigualdad se deduce el hecho obvio siguiente: cuando $m \in \mathcal{M}_X$ es una evidencia probabilística, entonces la familia $\{f_\rho\}_{\rho \in \mathcal{I}_0}$ se reduce a la medida de probabilidad determinada por m. Esto motiva la cuestión de buscar condiciones sobre m y sobre ρ que garanticen la obtención de determinados tipos de medidas f_ρ. En este sentido, tenemos:

Corolario 3.2 *Para cualquier $\rho \in \mathcal{I}_0$ se cumple:*

1. *Si $\rho(x)+\rho(1-x) \leq 1$, $\forall x \in [0,1]$, entonces resulta la pareja ordenada $f_\rho \leq f'_\rho$.*
2. *Si $\rho(x)+\rho(1-x) \geq 1$, $\forall x \in [0,1]$, entonces resulta la pareja ordenada $f_\rho \geq f'_\rho$.*

Demostración: Basta aplicar el segundo apartado de la Proposición 3.1 ya que la desigualdad $\rho(x)+\rho(1-x) \leq 1$ (resp. $\rho(x)+\rho(1-x) \geq 1$) implica que $\rho \leq \rho'$ (resp. $\rho \geq \rho'$). $\square$

Para ilustrar lo anterior podemos considerar cualquier $\alpha \in [0, 1/2]$, en cuyo caso la definición (5) muestra claramente que la función ρ_α cumple las condiciones del primer apartado, resultando efectivamente la pareja ordenada de medidas difusas siguiente: $P_\alpha \leq P_{1-\alpha}$. Vemos ahora un ejemplo donde se muestra que si ρ y ρ' no están ordenadas, entonces f_ρ y f'_ρ pueden no estarlo tampoco; lo que impide que f_ρ sea una medida que junto con su dual acoten una probabilidad (no pudiendo por tanto ser representable ni mucho menos una evidencia):

Ejemplo: Sea X un conjunto con cardinal $n = 4$, donde definimos la evidencia consonante dada por la siguiente A.B.P: $m(\{x_1\}) = m(\{x_1, x_2\}) = 1/4$, $m(X) = 1/2$. Tomamos además la función ρ definida como sigue:

$$\rho(x) := \begin{cases} 0, & x = 0 \\ 1/4, & 0 < x \leq 1/2 \\ 1, & 1/2 < x \leq 1 \end{cases}$$

Observemos que ρ verifica la condición del segundo apartado del Corolario 3.2 en todos los puntos excepto en $x = 1/2$. Pero eso es suficiente para que f_ρ y f'_ρ no estén ordenadas, ya que basta calcular los siguientes valores: $f_\rho(\{x_3\}) = 1/8 > 0 = f'_\rho(\{x_3\})$, mientras que $f_\rho(\{x_1, x_2\}) = 5/8 < 7/8 = f'_\rho(\{x_1, x_2\})$. $\triangleright$

Finalmente consideramos otro ejemplo para mostrar que ni la probabilidad Pw ni la medida difusa $P_{1/2}$ determinan unívocamente la A.B.P m de la cual provienen (hecho que sí ocurre con Bel y Pl como es sabido y con P_α, si $\alpha \neq 1/2$, como puede comprobarse).

Ejemplo: Sea $n = 2$, $r \in [0, 1/2]$ y consideremos la A.B.P. definida por: $m(\{x_1\}) = m(\{x_2\}) = r$, $m(X) = 1 - 2r$. Entonces basta efectuar simples cálculos para comprobar que: $\forall r \in [0, 1/2]$, $Pw(\{x_1\}) = Pw(\{x_2\}) = 1/2 = P_{1/2}(\{x_1\}) = P_{1/2}(\{x_2\})$. $\triangleright$

Las propiedades y ejemplos anteriores sugieren nuevas cuestiones (que indican una posible continuación de este trabajo). Por ejemplo:

- ¿Es inyectiva la aplicación que asigna a cada ρ la medida difusa f_ρ?
- Además de Bel y Pl, ¿cuáles son todas las medidas de la familia $\{f_\rho\}_{\rho \in \mathcal{I}_0}$ que permiten reconstruir la A.B.P. m de la cual provienen? ¿A qué tipo de funciones ρ corresponden tales medidas?

Nota: Una vez fijada la A.B.P. m, en la determinación de f_ρ intervienen solamente los valores $\rho\left(\frac{|B \cap A|}{|B|}\right)$ correspondientes a los distintos elementos focales B. Por tanto, al responder las cuestiones anteriores, deberemos considerar una relación de equivalencia en $\mathcal{I}_0$ de forma que cada ρ venga determinada exclusivamente por los valores que toma en el conjunto de puntos $\{\frac{p}{|B|} : 0 \leq p \leq |B|, B$ es elemento focal de $m\}$ (agradecemos a un referee la determinación de este conjunto). $\triangleright$

Referencias

[1] J.A. Herencia y M.T. Lamata, Una generalización de la entropía a través de la Teoría de Dempster–Shafer, *VI Congreso Español Sobre Tecnologías y Lógica Fuzzy* (Oviedo, 1996). Actas, pp. 143-147.

[2] J.A. Herencia and M.T. Lamata, Entropy Measures Associated with a Fuzzy Basic Probability Assignment, *Proc. of the FUZZ-IEEE'97* (I.I.I.A., Barcelona, Spain, 1997) 863-868.

[3] U. Höhle, Fuzzy plausibility measures. In: E.P. Klement (ed.), *Proc. Third Intern. Seminar on Fuzzy Set Theory* (Johannes Kepler Univ., Linz, 1981) 7-30.

[4] G.J. Klir and A. Ramer, Uncertainty in Dempster-Shafer theory: A critical re-examination, *Int. J. Gen. Syst.* 18 (1990) 155-166.

[5] M. Lamata and S. Moral, Classification of fuzzy measures, *Fuzzy Sets and Systems* 33 (1989) 243-253.

[6] J.F. Verdegay-López, Tratamiento de la incertidumbre mediante medidas difusas. En *Algunos aspectos del tratamiento de la Información en Inteligencia Artificial* (Publicaciones de la Universidad de Granada, 1991) 155-172.

[7] R.R. Yager, Entropy and specificity in a Mathematical Theory of Evidence, *Int. J. Gen. Syst.* 9 (1983) 249-260.