

## **DISTANCES BETWEEN FUZZY MEASURES THROUGH ASSOCIATED PROBABILITIES: SOME APPLICATIONS**

Luis M. de CAMPOS

*Departamento de Ciencias de la Computación e I.A., Colegio Universitario de Jaén,  
Universidad de Granada, 23071 Jaén, Spain*

Maria T. LAMATA and Serafin MORAL

*Departamento de Ciencias de la Computación e I.A., Facultad de Ciencias, Universidad de  
Granada, 18071 Granada, Spain*

Received March 1988

Revised July 1988

*Abstract:* In a recent paper a method to study fuzzy measures by means of certain sets of associated probabilities has been developed. In this paper, distances in the set of fuzzy measures are defined through distances between the associated probabilities. Finally, some applications of these distances to measure the uncertainty and specificity of fuzzy measures as well as to approximate fuzzy measures, are considered.

*Keywords:* Fuzzy measure; associated probabilities; distances; specificity and uncertainty indices.

### **1. Introduction**

Fuzzy measures are appropriate tools to represent information or opinion states.

Recently, a method for associating a set of probabilities to each fuzzy measure has been developed (see Campos and Bolaños [4]). The purpose of this paper is to use this method to define distances between fuzzy measures, and briefly to show how distances could be used for studying some characteristics of fuzzy measures.

Section 2 is devoted to show the necessary elements we need for further development. Specifically the associated probabilities to a fuzzy measure and some properties of them are presented.

In Section 3 we build distances between fuzzy measures through distances between their associated probabilities. It is appropriate to underline the flexibility of this method, which permits us to define many different distances.

In Section 4 we study some applications of the distance concept: the probability nearest to a fuzzy measure, which permits us to evaluate the lack of specificity of fuzzy measures; the problem of approximating fuzzy measures, and the investigation of the quantity of uncertainty contained in a fuzzy measure.

## 2. Preliminary concepts

Let us consider a finite reference set  $X = \{x_1, x_2, \dots, x_n\}$ . A fuzzy measure on  $X$  in Sugeno's sense [15] is a set function

$$g: \mathfrak{P}(X) \rightarrow [0, 1],$$

where  $\mathfrak{P}(X)$  is the power set of  $X$ , satisfying

- (i)  $g(\emptyset) = 0, g(X) = 1$ ;
- (ii)  $\forall A, B \subseteq X, A \subseteq B \Rightarrow g(A) \leq g(B)$ ;

that is,  $g$  is a normalized and monotone set function. We denote by  $\mathcal{M}(X)$  the set of fuzzy measures on  $X$ , and by  $\mathcal{P}(X)$  the set of probability measures on  $X$ .

Two fuzzy measures  $g$  and  $g^*$  are called dual fuzzy measures if and only if the following relation is satisfied:

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

where  $\bar{A}$  is the complement of  $A$ .

Duality is an important concept, since it permits us to obtain alternative representations of a piece of information: we will consider the situation that two dual fuzzy measures contain the same information, but codified in a different way.

The following property of the Choquet integral [5] (also called monotone expectation): the integral of a function  $h$  with respect to a fuzzy measure  $g$  coincides with the mathematical expectation of  $h$  with respect to a probability measure which depends only on  $g$  and the ordering of the values of  $h$ , was the starting point for Campos and Bolaños in [4] to associate a set of  $n!$  probabilities to each fuzzy measure, in the following way:

**Definition 2.1.** Let  $g$  be a fuzzy measure on  $X$ . The probability functions  $P_\sigma$  defined by

$$\begin{aligned} P_\sigma(x_{\sigma(1)}) &= g(\{x_{\sigma(1)}\}), \dots, \\ P_\sigma(x_{\sigma(i)}) &= g(\{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(i)}\}) - g(\{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(i-1)}\}), \dots, \\ P_\sigma(x_{\sigma(n)}) &= 1 - g(\{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n-1)}\}), \end{aligned}$$

for each  $\sigma = (\sigma(1), \sigma(2), \dots, \sigma(n)) \in S_n$ , are called the associated probabilities to the fuzzy measure  $g$ , where  $S_n$  is the group of permutations of a set with  $n$  elements.

For example, the associated probability for the permutation  $(1, 2, \dots, n)$  is

$$\begin{aligned} P_{(1,2,\dots,n)}(x_1) &= g(\{x_1\}), \\ P_{(1,2,\dots,n)}(x_2) &= g(\{x_1, x_2\}) - g(\{x_1\}), \dots, \\ P_{(1,2,\dots,n)}(x_i) &= g(\{x_1, \dots, x_i\}) - g(\{x_1, \dots, x_{i-1}\}), \dots, \\ P_{(1,2,\dots,n)}(x_n) &= 1 - g(\{x_1, x_2, \dots, x_{n-1}\}). \end{aligned}$$

**Remark.** It is obvious that each  $p_\sigma$  is a probability function because  $p_\sigma(x_{\sigma(i)}) \geq 0$   $\forall i$  (by virtue of monotonicity of  $g$ ) and  $\sum_{i=1}^n p_\sigma(x_{\sigma(i)}) = 1$ .

If  $P_\sigma$  and  $P'_\sigma$ ,  $\sigma \in S_n$ , are the associated probabilities to different fuzzy measures  $g$  and  $g'$  respectively, the sets  $\{P_\sigma \mid \sigma \in S_n\}$  and  $\{P'_\sigma \mid \sigma \in S_n\}$  could be equal, but in such a case, probabilities cannot be generated by the same permutations, that is, it is not possible that  $P_\sigma = P'_\sigma \forall \sigma \in S_n$ . Moreover if  $P_\sigma = P'_\sigma \forall \sigma \in S_n$  then we can assure that  $g$  and  $g'$  are equal. So, we can interpret any fuzzy measure in terms of its associated probabilities.

The most remarkable case where different fuzzy measures provide the same set of  $n!$  probabilities, but ordered in a different way, is the case of dual fuzzy measures:

Two fuzzy measures  $g$  and  $g'$  are dual if and only if they have the same  $n!$  associated probabilities corresponding to dual permutations ( $\sigma$  and  $\sigma^*$  are dual permutations if  $\sigma^*(i) = \sigma(n - i + 1)$ ,  $i = 1, 2, \dots, n$ ), that is,  $P_\sigma = P'_{\sigma^*}$  if  $\sigma$  and  $\sigma^*$  are dual.

Another interesting case is when the fuzzy measure is a probability:

A fuzzy measure is a probability if and only if its  $n!$  associated probabilities coincide.

Proofs of these results can be found in [4].

### 3. Distances between fuzzy measures

Several applications of the concept of associated probabilities are possible. Now we focus our attention on the definition of distances between fuzzy measures. This may be interesting because we could measure differences between information or opinion states. Moreover, in the next section we will study other possible applications of distances.

The procedure that we use is to extend distances between probabilities to fuzzy measures by means of the associated probabilities.

There are many distances for probability measures (see, for example, [2] or [10]). Any of them could be used by our method, depending on the particular problem that we consider.

Given two fuzzy measures  $g$  and  $g'$ , we need a function that starting from the distance between their associated probabilities, gives a value representing the distance between  $g$  and  $g'$ .

We rank the  $n!$  permutations of  $S_n$  with some criterion (for example by means of the lexicographical order) to number them, and thus to represent the set  $\{P_\sigma \mid \sigma \in S_n\}$  as an  $n!$ -tuple  $(P_1, P_2, \dots, P_{n!})$ .

For convenience, we will call  $m = n!$  and

$$T = \{(y_1, y_2, \dots, y_m) \in \mathbb{R}^m \mid y_i \geq 0 \forall i\}.$$

**Proposition 3.1.** Consider a function  $f: T \rightarrow \mathbb{R}$  satisfying the following properties:

- (1)  $f(y_1, y_2, \dots, y_m) = 0 \Leftrightarrow y_1 = y_2 = \dots = y_m = 0$ ;
- (2)  $f$  is monotone non-decreasing:

$$y_i \leq z_i \forall i \Rightarrow f(y_1, y_2, \dots, y_m) \leq f(z_1, z_2, \dots, z_m);$$

(3)  $f$  is subadditive:

$$f(y_1 + z_1, y_2 + z_2, \dots, y_m + z_m) \leq f(y_1, y_2, \dots, y_m) + f(z_1, z_2, \dots, z_m).$$

Then, for any distances  $d_i$ ,  $i = 1, \dots, m$ , defined on the set  $\mathcal{P}(X)$ , the function  $D: \mathcal{M}(X) \times \mathcal{M}(X) \rightarrow \mathbb{R}$  defined by

$$D(g, g') = f(d_1(P_1, P'_1), d_2(P_2, P'_2), \dots, d_m(P_m, P'_m))$$

is a distance on the set  $\mathcal{M}(X)$ .

**Proof.** As  $d_i(P_i, P'_i) \geq 0 \forall i$  and  $f$  is non-negative (because of properties (1) and (2)), it follows that  $D(g, g') \geq 0 \forall g, g' \in \mathcal{M}(X)$ .

$$\begin{aligned} D(g, g') = 0 &\Leftrightarrow f(d_1(P_1, P'_1), \dots, d_m(P_m, P'_m)) = 0 \\ &\Leftrightarrow d_i(P_i, P'_i) = 0 \forall i \Leftrightarrow P_i = P'_i \forall i \Leftrightarrow g = g'. \end{aligned}$$

It is evident that  $D(g, g') = D(g', g)$ , because  $d_i$ ,  $i = 1, \dots, n$ , are distances.

Finally,  $\forall g, g', g'' \in \mathcal{M}(X)$ , by properties (2) and (3), and taking into account that  $d_i$  are distances,

$$\begin{aligned} D(g, g'') &= f(d_1(P_1, P''_1), \dots, d_m(P_m, P''_m)) \\ &\leq f(d_1(P_1, P'_1) + d_1(P'_1, P''_1), \dots, d_m(P_m, P'_m) + d_m(P'_m, P''_m)) \\ &\leq f(d_1(P_1, P'_1), \dots, d_m(P_m, P'_m)) \\ &\quad + f(d_1(P'_1, P''_1), \dots, d_m(P'_m, P''_m)) \\ &= D(g, g') + D(g', g''), \end{aligned}$$

and the triangular inequality holds. Therefore  $D$  is a distance on  $\mathcal{M}(X)$ .  $\square$

**Remark.** Although we can use a distance  $d_i$  for each variable, it seems reasonable to use always the same distance:  $d_i = d \forall i$ .

If we want the distance  $D$  to be an extension of the distance  $d$  (used to define it), we must also demand the property of idempotency to  $f$ .

On the other hand, since the ordering of permutations (which has permitted us to change the set of associated probabilities by an  $m$ -tuple) is very arbitrary, it is desirable that the distance does not depend on this ordering. We can obtain it by requiring  $f$  to be a function symmetric with respect to its arguments.

Thus, we define a function generatrix of distances as follows:

**Definition 3.1.** Let  $f: T \rightarrow \mathbb{R}$  be a function satisfying the following conditions:

- (1)  $f(y_1, y_2, \dots, y_m) = 0 \Leftrightarrow y_1 = y_2 = \dots = y_m = 0$ ;
- (2)  $f$  is monotone non-decreasing:

$$y_i \leq z_i \forall i \Rightarrow f(y_1, y_2, \dots, y_m) \leq f(z_1, z_2, \dots, z_m);$$

- (3)  $f$  is subadditive:

$$f(y_1 + z_1, y_2 + z_2, \dots, y_m + z_m) \leq f(y_1, y_2, \dots, y_m) + f(z_1, z_2, \dots, z_m);$$

(4)  $f$  is idempotent:

$$f(y, y, \dots, y) = y \quad \forall y \geq 0;$$

(5)  $f$  is symmetric with respect to its arguments:

$$f(y_{\sigma(1)}, y_{\sigma(2)}, \dots, y_{\sigma(m)}) = f(y_1, y_2, \dots, y_m) \quad \forall \sigma \in S_m.$$

Then  $f$  is called a function generatrix of distances.

**Proposition 3.2.** *Let  $f: T \rightarrow \mathbb{R}$  be a function generatrix of distances, and let  $d$  be a distance on  $\mathcal{P}(X)$ . Then the function*

$$D: \mathcal{M}(X) \times \mathcal{M}(X) \rightarrow \mathbb{R}, \quad D(g, g') = f(d(P_1, P'_1), \dots, d(P_m, P'_m))$$

is a distance on  $\mathcal{M}(X)$  and

(a)  $D$  is an extension of  $d$ , that is,  $D(P, P') = d(P, P') \quad \forall P, P' \in \mathcal{P}(X)$ .

(b)  $D(g, g') = D(g^*, g'^*)$ , where  $g^*$  and  $g'^*$  are the dual fuzzy measures of  $g$  and  $g'$  respectively.

**Proof.**  $D$  is a distance because of properties (1)–(3) and Proposition 3.1.

(a) As  $P$  and  $P'$  are probabilities, their associated probabilities satisfy  $P_i = P, P'_i = P' \quad \forall i = 1, \dots, m$ . Therefore

$$\begin{aligned} D(P, P') &= f(d(P_1, P'_1), \dots, d(P_m, P'_m)) \\ &= f(d(P, P'), \dots, d(P, P')) = d(P, P'), \end{aligned}$$

because of condition (4).

(b) Since the associated probabilities to a fuzzy measure and its dual fuzzy measure are equal (although they are obtained from different permutations), and  $f$  is symmetric, it is obvious that  $D(g, g') = D(g^*, g'^*)$ .  $\square$

The last result is reasonable, taking into account that dual fuzzy measures codify the same information, and we are comparing the same codifications.

**Remark.** Because of monotonicity and idempotency of  $f$ , the relation

$$\min(y_1, y_2, \dots, y_m) \leq f(y_1, y_2, \dots, y_m) \leq \max(y_1, y_2, \dots, y_m)$$

holds, and we can consider the distance between two fuzzy measures as an average of distances between their associated probabilities.

This method for building distances between fuzzy measures is very flexible: we can obtain many different distances using different functions  $f$  and/or different distances between probabilities.

Next, we will see some simple examples:

The functions

$$f_m(y_1, \dots, y_m) = \max(y_1, \dots, y_m), \quad (1)$$

$$f_q(y_1, \dots, y_m) = \sqrt[q]{\frac{1}{m} \sum_{j=1}^m y_j^q}, \quad q \geq 1, \quad (2)$$

satisfy conditions (1)–(5). Therefore they are generatrix functions of distances.

The functions

$$d_m(P, P') = \max_{1 \leq i \leq n} |p(x_i) - p'(x_i)|, \quad (3)$$

$$d_q(P, P') = \sqrt[q]{\sum_{i=1}^n |p(x_i) - p'(x_i)|^q}, \quad q \geq 1, \quad (4)$$

$$d_s(P, P') = \max_{A \subseteq X} |P(A) - P'(A)| \quad (5)$$

are distances between probabilities.

#### 4. Some applications of distances

In this section we will outline some applications of distances between fuzzy measures. We intend to show how the concept of distance may be used to measure the non-specificity of a fuzzy measure, to measure the uncertainty and to solve the problems of approximation and convergence of fuzzy measures.

##### 4.1. Approximation of a fuzzy measure by a probability

The problem is the following: Is it possible to find the probability nearest to a given fuzzy measure? Since we have a tool (the distance) for measuring proximity, we can set out the problem in the following way:

Given  $g \in \mathcal{M}(X)$ , find  $P_g \in \mathcal{P}(X)$  such that

$$D(g, P_g) = \min_{P \in \mathcal{P}(X)} D(g, P).$$

It can be interesting to solve this problem for replacing a fuzzy measure by a probability and for other reasons specific for some types of fuzzy measures (see [6]).

We have solved the problem before for a particular distance, which has been generated by the function

$$f_2(y_1, \dots, y_m) = \sqrt{\frac{1}{m} \sum_{j=1}^m y_j^2} \quad (6)$$

and the distance on  $\mathcal{P}(X)$

$$d_2(P, P') = \sqrt{\sum_{i=1}^n (p(x_i) - p'(x_i))^2}. \quad (7)$$

So, the distance  $D_2$  is

$$\begin{aligned} D_2(g, g') &= \sqrt{\frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n (p_j(x_i) - p'_j(x_i))^2} \\ &= \sqrt{\frac{1}{n!} \sum_{\sigma \in S_n} \sum_{i=1}^n (p_{\sigma}(x_i) - p'_{\sigma}(x_i))^2}. \end{aligned}$$

The probability nearest to a fuzzy measure  $g$  is

$$p_g(x_i) = \frac{1}{n!} \sum_{\sigma \in S_n} p_{\sigma}(x_i), \quad i = 1, \dots, n, \quad (8)$$

that is, it is the uniform average of the associated probabilities to  $g$ .

To find this distribution we calculate the values that make zero the partial derivatives of the convex function

$$f(P) = \sum_{\sigma \in S_n} \sum_{i=1}^n (p_{\sigma}(x_i) - p(x_i))^2,$$

with respect to each  $p(x_i)$ . It can be easily verified that these values are a minimum probability distribution and equal to  $p_g(x_i)$ .

An interesting particular case are the evidence measures (Shafer [14]). Dubois and Prade in [6] suggest to associate the probability obtained by the distribution of the evidence mass of each focal set among its elements in each evidence measure. They also develop a method to associate a probability to each possibility measure.

It can be proved (see [3]) that our nearest probability coincides with the probabilities proposed by Dubois and Prade in both cases.

#### 4.2. Degree of non-specificity

The concept of non-specificity is well established in the case of evidence and possibility measures (see [7, 9, 16]); it is connected with the size of the subsets (focal elements) that are designated as prospective locations of an unknown element in the reference set. The larger the subsets, the less specific the characterization.

Non-specificity is related with the concept of imprecision in the sense that the range of variability between the lower measure (belief) and the upper measure (plausibility) increases as non-specificity increases.

In the case of more general fuzzy measures, where we do not have a basic probability assignment, we may define non-specificity or imprecision by falling back on this last idea. So, a non-specificity measure or imprecision measure  $N_{sp}$  must satisfy the following conditions:

- (a)  $N_{sp}$  is minimum only for the probability measures.
- (b)  $N_{sp}$  is maximum only for the complete ignorance measure.
- (c)  $N_{sp}$  is monotonic with respect to some inclusion relation between fuzzy measures.

Since the probability measures are the most specific measures (see [16]), if we evaluate the expression  $NA(g) = D_2(g, P_g)$ , we have a value that could represent the degree of non-specificity of the fuzzy measure  $g$ . It is easy to check the following properties of  $NA(\cdot)$ :

- (1)  $NA(g) = 0 \Leftrightarrow g$  is a probability measure.
- (2)  $NA(g) = NA(g^*)$ , if  $g$  and  $g^*$  are dual fuzzy measures.
- (3) The function  $NA: \mathcal{M}(X) \rightarrow \mathbb{R}$ ,  $g \rightarrow NA(g)$ , is continuous.

Next, we give an example to illustrate the behaviour of  $NA(\cdot)$ :

**Example 4.1.** Consider the fuzzy measure  $g_\varepsilon$  on  $X = \{x_1, \dots, x_n\}$  defined by

$$g_\varepsilon(A) = \begin{cases} (1 - \varepsilon) \frac{|A|}{n} & \text{if } A \neq X, \\ 1 & \text{if } A = X, \end{cases} \quad 0 \leq \varepsilon \leq 1,$$

where  $|A|$  represents the cardinality of  $A$ .

$g$  is an evidence measure (more precisely a belief measure) whose basic probability assignment is

$$m_\varepsilon(A) = \begin{cases} (1 - \varepsilon) \frac{1}{n} & \text{if } A = \{x_i\}, i = 1, \dots, n, \\ \varepsilon & \text{if } A = X, \\ 0 & \text{otherwise,} \end{cases}$$

that is  $g(A) = \sum_{B \subseteq A} m_\varepsilon(B)$ .  $g(A)$  measures the total belief mass that is constrained to  $A$ , whereas  $m_\varepsilon(A)$  measures the total portion of belief that is confined to  $A$  yet none of which is confined to any proper subset of  $A$ .

The associated probabilities to  $g$  are

$$P_\sigma(x_i) = \begin{cases} (1 - \varepsilon) \frac{1}{n} & \text{if } \sigma(i) \neq n, \\ (1 - \varepsilon) \frac{1}{n} + \varepsilon & \text{if } \sigma(i) = n, \end{cases} \quad \forall i = 1, \dots, n, \quad \forall \sigma \in S_n.$$

The probability nearest to  $g_\varepsilon$  is

$$p_\varepsilon(x_i) = \frac{1}{n}, \quad i = 1, \dots, n,$$

that is, the uniform probability on  $X$ .

The distance  $D_2(g_\varepsilon, P_\varepsilon)$  is

$$D_2(g_\varepsilon, P_\varepsilon) = \sqrt{\frac{1}{n!} \sum_{\sigma \in S_n} \sum_{i=1}^n \left( p_\sigma(x_i) - \frac{1}{n} \right)^2} = \varepsilon \sqrt{\frac{n-1}{n}}.$$

Therefore  $NA(g_\varepsilon) = \varepsilon \sqrt{(n-1)/n}$ . This result is reasonable:

For  $\varepsilon = 0$ ,  $g_\varepsilon$  is the uniform probability.

For  $\varepsilon = 1$ ,  $g_\varepsilon$  is the measure of complete ignorance.

For  $0 < \varepsilon < 1$  we obtain intermediate situations.

Thus  $NA(\cdot)$  depends only on  $\varepsilon$  and non-specificity increases as  $\varepsilon$  increases.

A more suitable option is to consider the situation that the fuzzy measures with less specificity are the pair dual of measures of complete ignorance (Shafer [14]):

$$\text{Bel}_0(A) = \begin{cases} 0 & \text{if } A \neq X, \\ 1 & \text{if } A = X, \end{cases} \quad \text{Pl}_0(A) = \begin{cases} 1 & \text{if } A \neq \emptyset, \\ 0 & \text{if } A = \emptyset. \end{cases}$$

Then, we can use the distance between any fuzzy measure  $g$  and one of those two fuzzy measures as a specificity index:

$$C(g) = \min(D(g, \text{Bel}_0), D(g, \text{Pl}_0)) \quad (9)$$



If  $C(g)$  is a value near zero, then  $g$  is near  $\text{Bel}_0$  or  $\text{Pl}_0$ , and therefore  $g$  has not much specificity. However, if  $C(g)$  has a high value,  $g$  is not like  $\text{Bel}_0$  or  $\text{Pl}_0$ , and thus  $g$  has a high degree of specificity.

It is not difficult to prove the following properties of  $C(\cdot)$ :

- (1)  $C : \mathcal{M}(X) \rightarrow \mathbb{R}$  is a continuous function.
- (2)  $C(g) = C(g^*)$  if  $g$  and  $g^*$  are dual fuzzy measures.
- (3)  $C(g) = 0 \Leftrightarrow g = \text{Bel}_0$  or  $g = \text{Pl}_0$ .

**Example 4.2.** Let us consider the set  $X = \{x_1, x_2, x_3\}$ , and a belief measure  $\text{Bel}$  on  $X$ , whose basic probability assignment is  $m$ , that is  $\text{Bel}(A) = \sum_{B \subseteq A} m(B)$ .

If we denote  $m_i = m(\{x_i\})$ ,  $m_{ij} = m(\{x_i, x_j\})$  and  $m_{123} = m(X)$ , the associated probabilities to  $\text{Bel}$  are shown in Table 1.

We choose the distance generated by

$$f_1(y_1, \dots, y_m) = \frac{1}{m} \sum_{j=1}^m y_j$$

and

$$d(P, P') = \frac{1}{2} d_1(P, P') = \frac{1}{2} \sum_{i=1}^n |p(x_i) - p'(x_i)|,$$

that is,

$$D_1(g, g') = \frac{1}{2n!} \sum_{\sigma \in S_n} \sum_{i=1}^n |p_\sigma(x_i) - p'_\sigma(x_i)|.$$

Under these conditions,

$$\begin{aligned} C(\text{Bel}) &= D_1(\text{Bel}, \text{Bel}_0) = \frac{1}{3}(2m_1 + 2m_2 + 2m_3 + m_{12} + m_{13} + m_{23}) \\ &= \frac{1}{3}\{(3-1)(m_1 + m_2 + m_3) + (3-2)(m_{12} + m_{13} + m_{23}) + (3-3)m_{123}\} \\ &= 1 - \frac{1}{3}(m_1 + m_2 + m_3 + 2m_{12} + 2m_{13} + 2m_{23} + 3m_{123}). \end{aligned}$$

This expression can be generalized to a  $\text{Bel}$  on any finite set  $X$ , and we obtain

$$C(\text{Bel}) = \frac{1}{n} \sum_{A \subseteq X} (n - |A|)m(A) = 1 - \frac{1}{n} \sum_{A \subseteq X} |A| m(A). \tag{10}$$

This index has the properties of a specificity measure:  $C(\text{Bel})$  is maximum and equal to  $(1 - (1/n))$  if  $\text{Bel}$  is a probability measure, and it is minimum and equal to zero if  $\text{Bel}$  is the measure of complete ignorance  $\text{Bel}_0$ .

Table 1

$\sigma$	$P_\sigma(x_1)$	$P_\sigma(x_2)$	$P_\sigma(x_3)$
(1, 2, 3)	$m_1$	$m_2 + m_{12}$	$m_3 + m_{13} + m_{23} + m_{123}$
(1, 3, 2)	$m_1$	$m_2 + m_{12} + m_{23} + m_{123}$	$m_3 + m_{13}$
(2, 1, 3)	$m_1 + m_{12}$	$m_2$	$m_3 + m_{13} + m_{23} + m_{123}$
(2, 3, 1)	$m_1 + m_{12} + m_{13} + m_{123}$	$m_2$	$m_3 + m_{23}$
(3, 1, 2)	$m_1 + m_{13}$	$m_2 + m_{12} + m_{23} + m_{123}$	$m_3$
(3, 2, 1)	$m_1 + m_{12} + m_{13} + m_{123}$	$m_2 + m_{23}$	$m_3$

Moreover, it is easy to check that the index  $C(\cdot)$  can also be written as

$$C(\text{Bel}) = 1 - \frac{1}{n} \sum_{i=1}^n \text{Pl}(\{x_i\}), \quad (11)$$

where  $\text{Pl}(\cdot)$  is the dual plausibility measure of  $\text{Bel}(\cdot)$ . Thus, if two belief measures  $\text{Bel}$  and  $\text{Bel}'$  satisfies

$$\text{Bel}(A) \leq \text{Bel}'(A) \quad \forall A \subseteq X$$

(and therefore  $\text{Pl}'(A) \leq \text{Pl}(A) \quad \forall A \subseteq X$ ), then

$$C(\text{Bel}) \leq C(\text{Bel}'),$$

and the index  $C(\cdot)$  is monotonic with respect to disjunctive inclusion [8] and strong inclusion [13, 17].

### 4.3. Degree of uncertainty

Another topic that can be studied using distances between fuzzy measures is the problem of determining the quantity of uncertainty having a fuzzy measure. This problem has been studied by several authors from different points of view (see [7, 9, 12, 16]) for some classes of fuzzy measures (evidence and possibility measures).

The basic idea is to assume that the least uncertain fuzzy measures are the  $n$  Dirac measures:

$$\text{for } k = 1, \dots, n, \quad p^k(x_i) = \begin{cases} 0 & \text{if } i \neq k, \\ 1 & \text{if } i = k. \end{cases}$$

It seems reasonable to use some of the distances defined before as uncertainty index: the minimum distance between any fuzzy measure  $g$  and the Dirac measures could be an uncertainty index of  $g$ :

$$U(g) = \min_{1 \leq i \leq n} D(g, P^i). \quad (12)$$

If  $U(g)$  is near zero, then  $g$  is near some Dirac measure  $P^i$ , and therefore  $g$  has not much uncertainty. On the other hand, if  $U(g)$  is high,  $g$  is like no measure  $P^i$ , and we say that  $g$  generates great uncertainty.

It is not difficult to prove the following properties of  $U(\cdot)$ :

- (1)  $U: \mathcal{M}(X) \rightarrow \mathbb{R}$  is a continuous function.
- (2)  $U(g) = U(g^*)$  if  $g$  and  $g^*$  are dual fuzzy measures.
- (3)  $U(g) = 0 \Leftrightarrow g = P^k$  for some  $k = 1, \dots, n$ .

The following examples show the performance of index  $U(\cdot)$  for evidence and possibility measures.

**Example 4.3.** Let us consider again a belief measure  $\text{Bel}$  defined on  $X = \{x_1, x_2, x_3\}$ . If we make calculations, we obtain

$$\begin{aligned} D_1(\text{Bel}, P^k) &= \frac{1}{2} \frac{1}{n!} \sum_{\sigma \in S_n} 2(1 - p_\sigma(x_k)) \\ &= 1 - m_k - \frac{1}{2}m_{ki} - \frac{1}{2}m_{kj} - \frac{1}{3}m_{123}. \end{aligned}$$

This can be generalized to any finite set  $X$ , obtaining

$$D_1(\text{Bel}, P^k) = 1 - \sum_{\substack{A \subseteq X \\ x_k \in A}} \frac{m(A)}{|A|}.$$

Therefore,

$$U(\text{Bel}) = \min_{1 \leq i \leq n} \left( 1 - \sum_{\substack{A \subseteq X \\ x_i \in A}} \frac{m(A)}{|A|} \right) = 1 - \max_{1 \leq i \leq n} \sum_{\substack{A \subseteq X \\ x_i \in A}} \frac{m(A)}{|A|}. \quad (13)$$

$U(\cdot)$  seems a mixture of non-specificity and dissonance measures: it is minimum and equal to zero only in the case of Dirac measures, which have minimum dissonance and non-specificity; it is maximum for the complete ignorance measure and the uniform probability, which have respectively maximum non-specificity and dissonance. Moreover, Dubois and Prade [7] proved that the specificity measure defined by Yager [16],

$$S_p(\text{Bel}) = \sum_{A \subseteq X} \frac{m(A)}{|A|} \quad (14)$$

in the case of a possibility measure (which has not any dissonance), is equal to

$$S_p(\pi) = \max_{1 \leq i \leq n} \left( \sum_{\substack{A \subseteq X \\ x_i \in A}} \frac{m(A)}{|A|} \right). \quad (15)$$

Thus  $U(\pi) = 1 - S_p(\pi)$ . Therefore, in this case  $U(\cdot)$  measures only non-specificity.

On the other hand, in the case of probability measures (which only have dissonance as source of uncertainty), the index  $U(\cdot)$  is

$$U(p) = 1 - \max_{1 \leq i \leq n} p(x_i). \quad (16)$$

That is, in absence of non-specificity  $U(\cdot)$  measures only dissonance. So,  $U(\cdot)$  may be considered as an index of global uncertainty.

**Remark.** In the general case the uncertainty index  $U(\cdot)$  can also be written as

$$U(g) = 1 - \max_{1 \leq i \leq n} p_g(x_i),$$

where  $P_g$  is the probability nearest to  $g$  studied above (Eq. (8)).

#### 4.4. Approximation and convergence of fuzzy measures

The problem considered here is a generalization of the above problems. If we have a family of fuzzy measures  $\mathcal{F}$ , a distance  $D$  and an arbitrary fuzzy measure  $g$ , the problem of approximating  $g$  by an element of  $\mathcal{F}$  may be stated in the following way:

Find the measure  $g' \in \mathcal{F}$  such that

$$D(g, g') = \min_{h \in \mathcal{F}} D(g, h).$$

The value  $D(g, g')$  may be considered as a degree of resemblance between  $g$  and the elements from  $\mathcal{F}$ .

Another problem that we may solve using the distance is to determine if a succession of fuzzy measures  $\{g_n\}_{n \in \mathbb{N}}$  converges to a fuzzy measure  $g$  ( $\{D(g_n, g)\}_{n \in \mathbb{N}} \rightarrow 0$ ). For this problem all the distances we have considered are equivalent.

## 5. Concluding remarks

Starting from the concept of associated probabilities to a fuzzy measure, we have defined distances between fuzzy measures in a very flexible way. Some of these distances have allowed us to study several characteristics of fuzzy measures.

Perhaps the use of other distances would be interesting to improve our indices.

## References

- [1] G. Banon, Distinction between several subsets of fuzzy measures, *Fuzzy Sets and Systems* **5** (1981) 325–339.
- [2] J. Burbea and C.R. Rao, Entropy differential metric, distance and divergence measures in probability spaces: an unified approach, *J. Multivariate Anal.* **12** (1983) 575–596.
- [3] L.M. de Campos, Caracterización y estudio de medidas e integrales difusas a partir de probabilidades, Ph.D. Thesis, Universidad de Granada (1987).
- [4] L.M. de Campos and M.J. Bolaños, Representation of fuzzy measures through probabilities, *Fuzzy Sets and Systems* **31** (1989) 23–36.
- [5] G. Choquet, Theory of capacities, *Ann. Inst. Fourier* **5** (1953) 131–295.
- [6] D. Dubois and H. Prade, Unfair coins and necessity measures: towards a possibilistic interpretation of histograms, *Fuzzy Sets and Systems* **10** (1983) 15–20.
- [7] D. Dubois and H. Prade, A note on measures of specificity for fuzzy sets, *Busefal* **19** (1984) 83–89.
- [8] D. Dubois and H. Prade, A set-theoretic view of belief functions, *Internat. J. General Systems* **12** (1986) 193–226.
- [9] M. Higashi and G.J. Klir, Measures of uncertainty and information based on possibility distributions, *Internat. J. General Systems* **9** (1983) 43–58.
- [10] P.J. Huber, *Robust Statistics* (Wiley, New York, 1981).
- [11] M.T. Lamata, Modelos de decisión con información general, Ph.D. Thesis, Universidad de Granada (1985).
- [12] M.T. Lamata and S. Moral, Measures of entropy in the theory of evidence, *Internat. J. General Systems* **14** (1987) 297–305.
- [13] S. Moral, Información difusa, relaciones entre probabilidad y posibilidad, Ph.D. Thesis, Universidad de Granada (1983).
- [14] G. Shafer, *A Mathematical Theory of Evidence* (Princeton University Press, Princeton, NJ, 1976).
- [15] M. Sugeno, Theory of fuzzy integrals and its applications, Ph.D. Thesis, Tokio Inst. of Technology (1974).
- [16] R.R. Yager, Entropy and specificity in a mathematical theory of evidence, *Internat. J. General Systems* **9** (1983) 249–260.
- [17] R.R. Yager, The entailment principle for Dempster–Shafer granule, Tech. Report MII-512, Iona College, New Rochelle, NY (1985).