RELATIONSHIPS BETWEEN UNCERTAINTY MEASURES, FUZZINESS MEASURES AND DIVERGENCE MEASURES

Susana Montes
Dept. of Statistics and O.R.
University of Oviedo
smr@pinon.ccu.uniovi.es

Inés Couso
Dept. of Statistics and O.R.
University of Oviedo
couso@pinon.ccu.uniovi.es

Pedro Gil
Dept. of Statistics and O.R.
University of Oviedo
pedro@pinon.ccu.uniovi.es

Summary

The aim of this communication is to establish some relationships between uncertainty measures or entropies, fuzziness measures and divergence measures, that is, between different ways to measure the uncertainty, which can be produced by a random experiment and/or the inaccuracy in the definition of the subsets in the experiment. Different relationships will be studied until we establish a one-to-one function between some important families of uncertainty, fuzziness and divergence measures.

Keywords: uncertainty measure, fuzziness measure and divergence measure.

1 INTRODUCTION

Our work regards the study of uncertainty associated to systems in a fuzzy environment. The starting point of our research has been the axiomatic Information Theory of B.Forte and J.Kampé de Fèriet, where uncertainty is directly associated with a collection of (crisp) subsets of a space Ω (see for instance [1]). In the frame of this theory it is possible to guess that there exists a fairly strong relationship between uncertainty and fuzziness, and between uncertainty and classical divergence. In this respect, a fundamental work has been developed by De Luca-Termini, who introduced a kind of fuzziness measure (the non-probabilistic entropy of a fuzzy set) based on a probabilistic uncertainty measure.

In our opinion, the link between measures of uncertainty and imprecision in fuzzy environments, may lie in what we refer to as divergence measure, because of the analogy with the classical meaning of the term used by various authors ([9], ...) in comparing two probability distributions.

We will begin this work by introducing definitions and characterizations of these concepts in Section 2. Afterwards, in Section 3, we will study some relationships between uncertainty measures and fuzziness measures, and in Section 4, we will introduce these last ones and divergence measures, to end establishing an isomorphism between uncertainty, fuzziness and divergence measures.

2 PRELIMINARY

Shannon ([10]) defined the first probabilistic uncertainty measure (or entropy) in the context of Communication Theory. According to the author, the uncertainty underlying a random experiment can be measured by means of the quantity $H(P) = -\sum_{i=1}^{n} p_i log_2(p_i)$, where the values p_i are the probabilities of the possible results of that experiment. Some authors have generalized the previous measure and given other families of uncertainty measures. Salicru et al. (see [9]) proved that all these measures of entropy are part of a wider family and named them $h-\phi-$entropies. This family is slightly more general than Ben Bassat's family of f-entropies that were defined as those functions that can be expressed like $H(P) = \sum_{i=1}^{n} f(p_i)$, where f is a concave function. In another communication of us ([3]), we have introduced and characterized a family which is a bit more general than Ben Bassat's one (but different from the family of $h-\phi-$entropies) in the case of discrete distributions. We named them quasi-ϕ-entropies, and they are expressed by $H(P) = \sum_{i=1}^{n} \phi(p_i)$, where ϕ is a function such that $\phi(\lambda x+(1-\lambda)y) \geq \lambda\phi(x)+(1-\lambda)\phi(y), \forall x, y \in [0,1], x+y \leq 1$.

On the other hand, an uncertainty measure is said to verify Pigou-Dalton's condition (or Principle of Transfer) if given two probability distributions P and P' with parameters $(p_1, p_2, \ldots, p_n)$ and $(p_1', p_2', \ldots, p_n')$ respectively, then $H(P) \leq H(P')$, where, $p_k = p_k', \forall k \notin \{i, j\}$, $p_i, p_j, \delta \leq (p_i - p_j)/2, p_i' = p_i + \delta, p_j' = p_j - \delta$. In other words, the more similar the probabilities of two outcomes of the experiment are, the higher the entropy is.

Once we have established some result about uncertainty measure, we are going to introduce another very important concept, studied by many authors ([6], [11], ...), which was named fuzziness measure.

Definition 2.1
A fuzziness measure is a real function f defined on $\widetilde{P}(\Omega)$, where $\widetilde{P}(\Omega) = \{$fuzzy subsets of $\Omega\}$, such that

1. $f(\widetilde{A}) = 0 \Longleftrightarrow \widetilde{A}$ *is a crisp set.*
2. *If $\widetilde{A}, \widetilde{B} \in \widetilde{P}(\Omega)$ and $\widetilde{A}$ is "sharper" than $\widetilde{B}$, then $f(\widetilde{A}) \leq f(\widetilde{B})$.*
3. *$f(\widetilde{A})$ assumes the maximum value if and only if $\widetilde{A}$ is "maximally fuzzy".*

Depending on the fuzziness degree meaning and measurement, the terms "sharper than" and "maximally fuzzy" will be considered in different ways. We will work with these two criteria, $\widetilde{A}$ is said to be sharper than $\widetilde{B}$ if and only if

1. either $A(x) \leq B(x) \leq 1/2$ or $A(x) \geq B(x) \geq 1/2, \forall x \in \Omega$ (see [5]),
   or more generally,
2. $\widetilde{A}$ is said to be sharper than $\widetilde{B}$ if and only if $|\widetilde{A}(x) - 1/2| \leq |\widetilde{B}(x) - 1/2|, \forall x \in \Omega$ (see [4]).

A very important family of fuzziness measure is the Knopfmacher's class ([6]), that is, the family formed by the functions f such that

$$f(\widetilde{A}) = F\left(\sum_{w \in \Omega} c_w \cdot g_w(\widetilde{A}(w))\right), \forall \widetilde{A} \in \widetilde{P}(\Omega),$$

where $c_w \in I\!R^+$; g_i is a real-valued function such that $g_w(0) = g_w(1) = 0, g_w(t) = g_w(1-t), \forall t \in [0,1]$ and g_w is strictly increasing on $[0, 1/2]$; F is a positive strictly increasing function with $F(0) = 0$.

In other works ([7], [8]) we have considered the particular Knopfmacher fuzziness measure such that F is the identity, g_w is the same for all $w \in \Omega$ (we have denoted g_w by u_f or simply u) and u is concave. Any function in this family was named local fuzziness measure.

Finally, we will introduce the last concept that we are going to relate to the two previous ones, which was studied firstly in [8].

Definition 2.2 *([7]) Let Ω be the referential. A map $D : \widetilde{P}(\Omega) \times \widetilde{P}(\Omega) \longrightarrow I\!R$ is a divergence measure if and only if $\forall \widetilde{A}, \widetilde{B} \in \widetilde{P}(\Omega)$, D satisfies the following conditions:*

1. $D(\widetilde{A}, \widetilde{B}) = D(\widetilde{B}, \widetilde{A})$;
2. $D(\widetilde{A}, \widetilde{A}) = 0$;
3. $\max\{D(\widetilde{A} \cup \widetilde{C}, \widetilde{B} \cup \widetilde{C}), D(\widetilde{A} \cap \widetilde{C}, \widetilde{B} \cap \widetilde{C})\} \leq D(\widetilde{A}, \widetilde{B}), \forall \widetilde{C} \in \widetilde{P}(\Omega)$.

According to this definition, if we consider only the case where Ω is finite, since the couples $(\widetilde{A}, \widetilde{B}), (\widetilde{A} \cup \{x_i\}, \widetilde{B} \cup \{x_i\})$ only differ in the i^{th} element, it seems natural to suppose that the variation of divergence only depends on what has been changed, and according to this, it is said that a divergence measure on a finite referential is local if and only if for all $\widetilde{A}, \widetilde{B} \in \widetilde{P}(\Omega)$ we have that $D(\widetilde{A}, \widetilde{B}) - D(\widetilde{A} \cup \{x_i\}, \widetilde{B} \cup \{x_i\}) = h(\widetilde{A}(x_i), \widetilde{B}(x_i))$.

The following statement characterizes the local divergence measures.

Proposition 2.3 *A map D is a local divergence if and only if there exists a function $h_D : [0,1] \times [0,1] \longrightarrow I\!R$ such that*

$$D(\widetilde{A}, \widetilde{B}) = \sum_{w \in \Omega} h_D(\widetilde{A}(w), \widetilde{B}(w))$$

where

i) $h_D(x, y) = h_D(y, x), \forall x, y \in [0,1]$;

ii) $h_D(x, x) = 0, \forall x \in [0,1]$;

iii) $h_D(x, z) \geq \max\{h_D(x, y), h_D(y, z)\}$, $\forall x, y, z \in [0,1]$ *with* $x \leq y \leq z$.

3 UNCERTAINTY MEASURES AND FUZZINESS MEASURES

From now on, we are going to consider that the referential Ω is finite.

Proposition 3.1 *Let H be an uncertainty measure that verifies the Pigou-Dalton's condition and such that*

$H(P) = 0 \iff P$ is degenerate. Let us consider a function $u_H : [0,1] \longrightarrow I\!R^+$ such that $u_H(x) = H(x, 1-x), \forall x \in [0,1]$. Then, the map $f : \widetilde{P}(\Omega) \longrightarrow I\!R^+$ defined by $f(\widetilde{A}) = h\left(\sum_{w\in\Omega} u_H(\widetilde{A}(w))\right)$, where $h : I\!R^+ \longrightarrow I\!R^+$ is a non-decreasing function such that $h(x) = 0 \iff x = 0$, is a fuzziness measure according to the first definition of "fuzzy subset sharper than" and it belongs to Knopfmacher's class.

Proof. It is a direct consequence of Theorem 2.1 in [2]. ■

The uncertainty measures more widely used are symmetrical, and so, if they have this property, the fuzziness measures they generate are according to the second definition of "fuzzy subset sharper than", which is stronger than the first one and we also think that is more natural.

Proposition 3.2 *In the preceding proposition, if H has also a symmetry property, that is, if $H(x, 1-x) = H(1-x, x), \forall x \in [0,1]$, then f is also a fuzziness measure according to the second definition "fuzzy subset sharper than".*

Proof. This is again a direct consequence of Theorem 2.1 in [2]. ■

Thus, if we have a uncertainty measure H satisfying adequate conditions, we can construct a fuzziness measure with this one.

Lemma 3.3 *A uncertainty measure H is a quasi-ϕ-entropy if and only if H is symmetrical, expansible, branching, continue and it fulfils Pigou-Dalton's condition.*

Proof. To prove this result, it is necessary to use Proposition 2.3 in [3]. ■

Once we have characterized the quasi-ϕ-entropies, we can use this result to establish a one-to-one function between uncertainty and fuzziness measures.

Proposition 3.4 *Let $F_1 : \mathcal{H} \longrightarrow \mathcal{F}$ be a function such that $F_1(H)(\widetilde{A}) = \sum_{w\in\Omega} u_H(\widetilde{A}(w))$, where u_H is the function defined in Proposition 3.1, $\mathcal{H}$ is a family of symmetrical and continue uncertainty measures satisfying Pigou-Dalton's condition and such that $H(P) = 0 \iff P$ is degenerate, and $\mathcal{F}$ is a family of fuzziness measures. The function F_1 is injective if and only if any uncertainty measure in $\mathcal{H}$ is a quasi-ϕ entropy with ϕ continue, $\phi(x) = \phi(1-x), \forall x \in [0, \frac{1}{2}$ and $\phi(x) = 0 \iff x = 0$.*

Proof. In this proof we have to consider Proposition 3.2 to see that F_1 is actually a function. Besides, $F_1(H)$ is a local fuzziness measure and then it is easy to prove that F_1 is injective. In order to prove the converse, we need to use Lemma 3.3. ■

Theorem 3.5 *If $\mathcal{H}$ is the class of ϕ-entropies such that ϕ is symmetrical with respect to $1/2$ and $H(P) = 0 \iff P$ is degenerate and if $\mathcal{F}$ is the class of local fuzziness, then the function F_1 defined in Proposition 3.4 is a one-to-one function.*

Proof. If H is a ϕ-entropy, then H is a quasi-ϕ-entropy, and then we have only to prove, by using Proposition 3.4, that F_1 is surjective. In order to see this, we have to consider the ϕ-entropy H defined by $\phi(x) = 1/2u_f(x), \forall x \in [0,1]$, and then $F_1(H) = f$. ■

4 FUZZINESS MEASURES AND DIVERGENCE MEASURES

In order to relate fuzziness measures and divergence measures, to continue with the ideas of fuzziness measures proposed by Yager ([12]), we need to establish a previous lemma for non-differentiable functions.

Lemma 4.1 *If a function $u : [0,1] \longrightarrow I\!R^+$ is concave, then the function $h : [0,1] \times [0,1] \longrightarrow I\!R^+$ defined by $h(x,y) = u\left(\frac{x+y}{2}\right) - \frac{u(x)+u(y)}{2}$ satisfies that $h(\cdot, y)$ is a non-increasing function in $[0,y]$ and a non-decreasing function in $[y,1]$, for any y fixed in $[0,1]$.*

Proof. To prove this lemma it is necessary to consider that if u is concave in $(0,1)$, then u is continue, there exist $u'(x^+), u'(x^-), \forall x \in (0,1)$, and for all $x_1, x_2 \in (0,1)$ such that $x_1 < x_2$, then $u'(x_1^-) \geq u'(x_1^+) \geq u'(x_2^+) \geq u(x_2^-)$ and we also have to prove a result similar to Lagrange's Theorem for functions r such that there exists $r(x^+)$ and $r(x^-)$ for all x, but r is non-differentiable. ■

By considering this result and the properties of local fuzziness measure, we have that

Proposition 4.2 *If we consider the function F_2 defined in $\mathcal{F}$ (family of local fuzziness measures) by the map $F_2(f) : \widetilde{P}(\Omega) \times \widetilde{P}(\Omega) \longrightarrow I\!R$ with*

$$F_2(f)(\widetilde{A}, \widetilde{B}) = f\left(\frac{\widetilde{A}+\widetilde{B}}{2}\right) - \frac{f(\widetilde{A})+f(\widetilde{B})}{2},$$
$$\forall \widetilde{A}, \widetilde{B} \in \widetilde{P}(\Omega),$$

then $F_2(f)$ is a local divergence measure such that $h(x,1-x)$ is a convex function of x and

$$h_D(x,y) = \tfrac{1}{2}h_D(x,1-x) + \tfrac{1}{2}h_D(y,1-y) - h_D\left(\tfrac{x+y}{2}, 1-\tfrac{x+y}{2}\right), \forall x,y \in [0,1].$$

Proof. The function $h_{F_2(f)}$ which define the local divergence measure is $h_{F_2(f)}(x,y) = u_f\left(\frac{x+y}{2}\right) - \frac{u(x)+u(y)}{2}, \forall x,y \in [0,1]$. By applying Lemma 4.1, we can prove that $F_2(f)$ defined by this $h_{F_2(f)}$ satisfies the condition in Proposition 2.3. It is trivial that $h_{F_2(f)}$ verifies the condition in the proposition. ∎

Thus, the local divergence measure considered in the previous proposition is completely characterized if we know the value of $h(x,1-x), \forall x \in [0,1]$.

Theorem 4.3 *If $\mathcal{D}$ is the class of local divergence measures such that $h_D(x,y) = \frac{1}{2}h_D(x,1-x) + \frac{1}{2}h_D(y,1-y) - h_D\left(\frac{x+y}{2}, 1-\frac{x+y}{2}\right), \forall x,y \in [0,1]$ and $h(x,1-x)$ is a convex function of x, then the map $F_2 : \mathcal{F} \longrightarrow \mathcal{D}$ is a one-to-one function.*

Proof. F_2 is a function by applying Proposition 4.2. It is injective since $h_{F_2(f)}$ is completely characterized by the values of $h_{F_2(f)}(x,1-x) = u_f(1/2) - u_f(x), \forall x \in [0,1]$. F_2 is surjective since for all $D \in \mathcal{D}$ we can consider de local fuzziness measure defined by $u_f(x) = h_D(1,0) - h_D(x,1-x), \forall x \in [0,1]$, and then $f_2(f) = D$. ∎

By joining Theorems 3.5 and 4.3 we have that

Corollary 4.4 *The set $\mathcal{H}$ of the class of quasi-ϕ entropy with ϕ continue, $\phi(x) = \phi(1-x), \forall x \in [0,1]$ and $\phi(x) = 0 \iff x = 0$, the set $\mathcal{F}$ of local fuzziness measures and the set $\mathcal{D}$ of the local divergence measures such that $h(x,y) = \frac{1}{2}h(x,1-x) + \frac{1}{2}h(y,1-y) - h\left(\frac{x+y}{2}, 1-\frac{x+y}{2}\right), \forall x,y \in [0,1]$ and $h(x,1-x)$ is a convex function of x, are isomorphic.*

5 CONCLUSIONS AND OPEN PROBLEMS

In this paper we have studied some relations between different ways to measure the uncertainty (probabilistic or non-probabilistic) and a measure of the difference between two fuzzy subsets. The main result is Corollary 4.4, where we have given a way to construct a divergence measure if we have an uncertainty measure or a fuzziness measure, and conversely.

With regard to the open problems, the most immediate is to find other relationships between this functions and other different ways to measure the uncertainty (fuzzy measure, aggregation operators, etc).

Acknowledgements

The research in this paper is supported in part by DGICYT Grant No. PB97-1286. Its financial support is gratefully acknowledged.

References

[1] C. Bertoluzza and M. Schneider (1974). Informations totalement composables. *Théories de l'Information, Lect. Notes Math.* **398**, pp 90–98.

[2] I. Couso, S. Montes and P. Gil (1996). Measures of fuzziness and Information Theory. *Proceedings of IPMU'96 Conference* pp 1030–1035.

[3] I. Couso and P. Gil (1998). Characterization of a family of entropy measures. *Proceedings of IPMU'98 Conference* pp 1053–1059.

[4] D. Dubois and H. Prade (1980). Fuzzy Sets and Systems. Theory and Applications. *Academic Press, Inc.*.

[5] G.J. Klir and T.A. Folger (1988). Fuzzy Sets, Uncertainty, and Information. *Prentice-Hall International (UK) Limited.*

[6] J. Knopfmacher (1975). On measures of fuzziness. *J. Math. Analysis and Applications* **49**, pp529–534.

[7] S. Montes, P. Gil and C. Bertoluzza. Divergence between fuzzy sets and fuzziness (1998). *Proceedings of IPMU'98 Conference* pp 1030–1035.

[8] S. Montes and P. Gil (1998). Some Classes of Divergence Measures Between Fuzzy Subsets and Between Fuzzy Partitions. *Mathware and Soft Computing* **5**, pp 253–265.

[9] L. Pardo (1997). Teoría de la Información estadística. Ed. Hesperides, Salamanca.

[10] C.E. Shannon (1948). A mathematical Theory of Communication. *Bell Syst. Tech. J.* **27**, pp 379–423, 623–656.

[11] E. Trillas and T. Riera (1978). Entropies in Finite Fuzzy Sets. *Inf. Sc.* **15**, 159–168.

[12] R.R. Yager (1995). Measures of Entropy and Fuzziness related to Agregation Operators. *Inf. Sc.* **82**, 147–166.