LOCAL DIVERGENCE MEASURES ON INFINITE REFERENTIALS

Jorge Jiménez
Dept. of Mathematics
University of Oviedo
meana@etsiig.uniovi.es

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

Carlo Bertoluzza
Dept. of Computer Eng. and Systems
University of Pavia
carlo@unipv.it

Summary

The aim of this work is to define and to characterize the local divergence measures when working with infinite referentials , that is, those measures which allow us to study the degree of similarity between two fuzzy sets which are decomposables. Once this concept is characterized with several different definitions, some of their properties are studied and some examples are presented.

Keywords: measurable space, measure σ-additive, divergence measure.

1 INTRODUCTION

In several previous papers ([6], [7], ...), it emerged a new form to quantify the degree of difference between two fuzzy subsets within any referential Ω from a function, which will be defined below, and it will be called divergence measure, which has as particular cases the customary distances between fuzzy sets already known. This new concept emerged in a sense as an extension of the classic probabilistic divergence measures ([8], ...), taking advantage of the strong relationship existing between some concepts of the Information Theory and the Fuzzy sets Theory as, for example, the probabilistic entropy of Shannon ([9]) and the non-probabilistic entropy of De Luca & Termini ([2]).

Once the definition of divergence measure was given, special emphasis was made on the study of the defined divergences on finite spaces which could be decomposed as the sum of the value of a function in each point of the referential . The extension of all this theory to numerable referentials was a simple exercise, but not thus the extension to no numerable infinite spaces, which will go through the measure spaces' study, and where the proofs, though they could seem a simple generalization, need an analytical rigor that will make them considerably more complex. This is precisely the study made in this paper, in which we find a Section 2 that presents the measure definition of local divergence for a non-numerable infinite referential and are studied several related results; in Section 3 we study some interesting properties of the measures defined in the previous section, these properties, as well as all the results of the Section 2, do not include the proofs which are sometimes very long, but we simply make some short commentaries on the most relevant parts. Finally, in Section 4, possible applications and future projects related to this one are commented.

2 LOCAL DIVERGENCE MEASURES

The measure of the difference of two fuzzy subsets is defined axiomatically on the basis of the following natural properties:
• It is a non-negative and symmetric function of the two fuzzy subsets to be compared.
• It becomes zero when the two sets coincide.
• It decreases as the two subsets become "more similar" in some sense.
Whereas it is easy to formulate analytically the first and the second condition, the third one depends on the formalization of the concept of "more similar". We base our approach on the fact that if we add (in the sense of union) a subset $\widetilde{C}$ to both fuzzy subsets $\widetilde{A}$ and $\widetilde{B}$, we obtain two subsets which are closer to each other; the same happens with the intersection.

Before giving the divergence measure definition, we need to include a previous lemma, for this definition to make sense.

Lemma 2.1 *Let $(\Omega, \mathcal{A})$ be a measure space, where Ω is any set and $\mathcal{A}$ is a σ-field of crisp sets of Ω, and let $\widetilde{P}(\Omega)$ be the set of all the fuzzy subsets of Ω, then the* *set $\mathcal{A}^*$ defined by*

$$\mathcal{A}^* = \{\widetilde{A} \in \widetilde{P}(\Omega) / \text{ the membership function of } \widetilde{A}, \\ \widetilde{A} : (\Omega, \mathcal{A}) \longrightarrow ([0,1], \beta) \text{ is a measurable function}\}$$

is a σ-field over Ω.

Proof. The proof is a consequence of the fact that the Borel's σ-field (β) is generated by the real intervals, $\mathcal{A}$ is a σ-field and that the definitions of union and intersection which we use, are the maximum and the minimum considered by Zadeh. ∎

So we propose the following

Definition 2.2 *Let Ω the universe under study, and let $\mathcal{A}$ be a σ-field defined on Ω (we can consider in particular the family of the fuzzy subset of Ω, that is, $P(\Omega)$). A map $D : \mathcal{A}^* \times \mathcal{A}^* \longrightarrow I\!R$ is a divergence measure if and only if $\forall \widetilde{A}, \widetilde{B} \in \mathcal{A}^*$, satisfies the following conditions:*

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

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 space is local, if and only if, $\forall \widetilde{A}, \widetilde{B} \in P(\Omega)$ and $\forall x_i \in \Omega$ we have that $D(\widetilde{A}, \widetilde{B}) - D(\widetilde{A} \cup \{x_i\}, \widetilde{B} \cup \{x_i\}) = h(\widetilde{A}(x_i), \widetilde{B}(x_i))$. To extend this definition to the general case we need the following result:

Lemma 2.3 *Let $(\Omega, \mathcal{A})$ be a measurable space, and $P(\Omega)$ the set formed by all the crisp subsets of Ω, then we have that $P(\Omega) \cap \mathcal{A}^* = \mathcal{A}$, that is, all the elements of the σ-field belong to $\mathcal{A}^*$, all the crisp elements of $\mathcal{A}^*$ belong to $\mathcal{A}$.*

Proof. The proof is followed from the fact that, $\mathcal{A}$ is a σ-field and that the inverse image of the intervals $(-\infty, z]$ by the membership function of a crisp subset, A is either the empty set, or Ω, or A^C. ∎

From now on, we will consider the measure space $(\Omega, \mathcal{A}, \mu)$ ([3]), where Ω is a infinite referential such that $\mu(\Omega) < \infty$.The extension of the local divergence measure to this kind of referentials can't be made point by point, so we have to extend this concept subset by subset.Thus,

Definition 2.4 *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. A divergence measure D, is said to be local if and only if there exists a function $h : [0,1] \times [0,1] \longrightarrow I\!R$ measurable respect to the Borel's σ-fields such that*

$$D(\widetilde{A}, \widetilde{B}) - D(\widetilde{A} \cup Z, \widetilde{B} \cup Z) = \int_Z h(\widetilde{A}(w), \widetilde{B}(w)) d\mu(w),$$

for all $Z \in \mathcal{A}$ and for any $\widetilde{A}$ y $\widetilde{B} \in \mathcal{A}^*$, where $\int$ denote the Lebesgue-Stielges integral.

All the integrals which we calculate in the previous definition are well defined by applying the lemmas 2.1 and 2.3.

Before going to the next result, we must see a previous lemma which will be necessary for its proof.

Lemma 2.5 *Let $\widetilde{A}$ be a fuzzy subset of Ω such that $\widetilde{A}(w)$ is constant for all $w \in \Omega$, then $\widetilde{A}$ is measurable with regard to the σ-field $\mathcal{A}^*$.*

Proof. The proof is followed from the fact that the inverse image of the intervals $(-\infty, z]$ by the membership function of a crisp subset, A is either the empty set, or Ω. ∎

The following statement characterizes the local divergence measures.

Proposition 2.6 *Let $(\Omega, \mathcal{A}, \mu)$ be a measurable space. If a map $D : \mathcal{A}^* \times \mathcal{A}^* \longrightarrow I\!R$ is a local divergence then there exists a function $h : [0,1] \times [0,1] \longrightarrow I\!R$ such that*

$$D(\widetilde{A}, \widetilde{B}) = \int_\Omega h(\widetilde{A}(w), \widetilde{B}(w)) d\mu(w)$$

where
i) $h(x,y) = h(y,x), \forall x, y \in [0,1]$;
ii) $h(x,x) = 0, \forall x \in [0,1]$;
iii) $h(x,z) \geq \max\{h(x,y), h(y,z)\}, \forall x, y, z \in [0,1]$ with $x \leq y \leq z$.
iv) h is measurable.
Conversely, If D is a function defined by $D(\widetilde{A}, \widetilde{B}) = \int_\Omega h(\widetilde{A}(w), \widetilde{B}(w)) d\mu(w)$ where h verify the four before properties then h is a local divergence measure.

Proof. To prove the fist implication is necessary to consider in Definition 2.4 $Z = \Omega$, and apply Lemma 2.5 together with the fact that $\mu(\Omega) < \infty$. we also have to define the constant sets $\widetilde{A}(w) = x, \widetilde{B}(w) = y$ y $\widetilde{C}(w) = z$ for all $w \in \Omega$ and to consider that the Lebesgue-Stielges integral is only defined for measurable functions, with the aim to prove the four properties of h.

In the converse proof we have to consider the fact that the Lebesgue-Stielges integral of a positive measurable function always exists, the properties of h, and the decomposition of Ω in the five following disjoined subsets: $\Omega_1 = \{w \in \Omega / \widetilde{A}(w) \leq \widetilde{B}(w) \leq \widetilde{C}(w) \;\; \widetilde{B}(w) < \widetilde{A}(w) \leq \widetilde{C}(w)\}$, $\Omega_2 = \{w \in \Omega / \widetilde{A}(w) \leq \widetilde{C}(w) < \widetilde{B}(w)\}$, $\Omega_3 = \{w \in \Omega / \widetilde{B}(w) \leq \widetilde{C}(w) < \widetilde{A}(w)\}$, $\Omega_4 = \{w \in \Omega / \widetilde{C}(w) < \widetilde{A}(w) \leq \widetilde{B}(w)\}$ y $\Omega_5 = \{w \in \Omega / \widetilde{C}(w) < \widetilde{B}(w) < \widetilde{A}(w)\}$, which are measurable subsets, because of being measurable a continuous function of measurable functions. ■

We can use the previous proposition to generate divergence measures easily in infinite referentials. The importance of this fact leads us to the searching of others possible conditions for the function h, which will be equivalent. So we have the following result

Corollary 2.7 *Condition iii) in Proposition 2.6 for h can be replaced by $h(\cdot, y)$ is a function decreasing in $[0, y]$ and increasing in $[y, 1]$.*

Up to now, We have seen, a local divergence measure based on the union with a crisp set. In the following result, we are going to give an equivalent condition based on the intersection with a crisp set.

Lemma 2.8 *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, if D is a local divergence measure, then for all crisp subset Z of Ω it is verified that*

$$D(\widetilde{A} \cup Z, \widetilde{B} \cup Z) = D(\widetilde{A} \cap Z^c, \widetilde{B} \cap Z^c), \forall \widetilde{A}, \widetilde{B} \in \mathcal{A}^*.$$

Proof. To prove this lemma is necessary to apply the definition of local divergence measure to the sets $\widetilde{A} \cap Z^c$ and $\widetilde{B} \cap Z^c$ with respect to Z. ■

Proposition 2.9 *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, if D is a divergence measure, then D is local if and only if there exists a measurable function $h : [0,1] \times [0,1] \longrightarrow I\!R$ such that*

$$D(\widetilde{A}, \widetilde{B}) - D(\widetilde{A} \cap Z, \widetilde{B} \cap Z) = \int_{Z^c} h(\widetilde{A}, \widetilde{B}) d\mu,$$

$\forall \widetilde{A}, \widetilde{B} \in \mathcal{A}^*$ and $\forall Z \in P(\Omega)$.

Proof. In the proof we have to use the previous lemma and the definition of local divergence measure applied to Z^c. ■

To end this section, we will enunciate a result which will allow us to generate divergence measures from a local divergence measure.

Proposition 2.10 *Let D be a local divergence measure associated to a measurable function h, if $\phi : I\!R \longrightarrow I\!R$ is a non-decreasing function and $\phi(0) = 0$, then the maps D_ϕ and D^ϕ are divergence measures, and, besides, D_ϕ is local, where*

$$D_\phi(\widetilde{A}, \widetilde{B}) = \int_\Omega \phi(h(\widetilde{A}(w), \widetilde{B}(w)) d\mu(w),$$
$$D^\phi(\widetilde{A}, \widetilde{B}) = \phi\left(\int_\Omega h(\widetilde{A}(w), \widetilde{B}(w) d\mu(w)\right).$$

Proof. To prove this proposition, we must take into account that $(\phi \circ h)^{-1}$ is equal to either$((-\infty, z]) = h^{-1}((-\infty, k_1])$ or $h^{-1}((-\infty, k_1))$ or $h^{-1}(I\!R)$, where $k_1 = \sup\{k \in I\!R / \phi(k) \leq z\}$, and that any of these three sets belong to the σ-field $\mathcal{A}$. ■

3 PROPERTIES OF THE LOCAL DIVERGENCE MEASURES

We will study several properties which are verified by the local divergence measures. This properties relate them to the fuzziness degree of the sets. Before stating the different properties we must explain the notation. $\widetilde{A} \lll \widetilde{B}$ will denote that couple of sets such that verify $|\widetilde{A}(w) - 1/2| \geq |\widetilde{B}(w) - 1/2|, \forall w \in \Omega$, and $\widetilde{A} \ll \widetilde{B}$ will denote that couple of sets such that verify $\widetilde{A}(w) \leq \widetilde{B}(w) \leq 1/2$ o $\widetilde{A}(w) \geq \widetilde{B}(w) \geq 1/2, \forall w \in \Omega$.

Proposition 3.1 *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, and let $\widetilde{A}$, $\widetilde{B}$ be in $\mathcal{A}^*$. Given D, a divergence measure with the local property, we have that*

1. *If $\widetilde{A} \lll \widetilde{B}$ then $D(\widetilde{A}, \widetilde{A}^c) \geq D(\widetilde{B}, \widetilde{B}^c)$.*
2. *$D(Z, Z^c) = D(V, V^c), \forall Z, V \in P(\Omega)$.*
3. *$D(\widetilde{A}, \widetilde{A}) \leq D(\widetilde{A}, \widetilde{B}) \leq D(Z, Z^c), \forall Z \in P(\Omega)$.*
4. *If the function h which generates D satisfies that $h(x, e) = h(c(x), e), \forall x \in [0, 1]$, where e is the fixed point of complementary function ^{c}, then $D(\widetilde{A}, \widetilde{E}_c) = D(\widetilde{A}^c, \widetilde{E}_c)$, where $\widetilde{E}_c(w) = e, \forall w \in \Omega$.*
5. *Given $\widetilde{A}, \widetilde{B}, \widetilde{C} \in \mathcal{A}^*$ such that for all $w \in \Omega$ we have that either $\widetilde{A}(w) \leq \widetilde{B}(w) \leq \widetilde{C}(w)$ or $\widetilde{A}(w) \geq \widetilde{B}(w) \geq \widetilde{C}(w)$, then it is got that $D(\widetilde{A}, \widetilde{C}) \geq \max\{D(\widetilde{A}, \widetilde{B}), D(\widetilde{B}, \widetilde{C})\}$.*
6. *If $\widetilde{A} \ll \widetilde{B}$ then $D(\widetilde{A}, \widetilde{E}) \geq D(\widetilde{B}, \widetilde{E})$ where $\widetilde{E}$ is the fuzzy set defined in property 4 when $c(x) = 1 - x$.*
7. *The function h which generates D, satisfies that $h(\cdot, 1/2)$ is symmetrical with regard to $1/2$ if and only if when $\widetilde{A} \lll \widetilde{B}$ then $D(\widetilde{A}, \widetilde{E}) \geq D(\widetilde{B}, \widetilde{E})$.*

8. *The function h which generates D satisfies that* $h(x,0)=h(1-x,1), \forall x \in [0,\frac{1}{2}]$, *if and only if* $\widetilde{A} < \ll \widetilde{B}$ *implies that* $D(\widetilde{A}, N_{\widetilde{A}}) \leq D(\widetilde{B}, N_{\widetilde{B}})$, *where* $N_{\widetilde{A}}, N_{\widetilde{B}}$ *represent the nearest crisp set to* $\widetilde{A}$ *and* $\widetilde{B}$, *respectively.*

9. *If* $h(\cdot, 1/2)$ *is symmetric with regard to* $1/2$, *then* $D(Z, \widetilde{E}) = D(V, \widetilde{E}), \forall Z, V \in P(\Omega)$.

10. *If* $\widetilde{A} \lll \widetilde{B}$, *we have that* $D(\widetilde{A}, \widetilde{E}) \geq D(\widetilde{B}, \widetilde{E})$, *if and only if the function h associated with D satisfies that* $h(\cdot, 1/2)$ *is symmetric with regard to* $1/2$.

Proof. To prove the first property is necessary to decompose Ω into three disjoined sets in the adequate way. These sets will be measurable by applying again the following property "a continuous function composed with measurable functions is measurable.
In the second property, we get that the divergence between a crisp set and its complementary is always constant and equal to $h(1,0) \cdot \mu(\Omega)$.
The third and fourth property are trivial by using the monotony of h and the hypothesis of the respective statements.
In the fifth one it is necessary to decompose again Ω in three measurable disjoined sets and to use the monotony of h.
The sixth one is consequence of the previous one.
In the seventh one, we have to consider the set $\widetilde{A}$ defined by $\widetilde{A}(w) = x, \forall w \in \Omega$ which verifies that $\widetilde{A} \lll \widetilde{A}^c$ and $\widetilde{A}^c \lll \widetilde{A}$, and we have to decompose Ω again in the four adequate measurable and disjoined sets.
The eighth one is proved using the previous decomposition of Ω and the monotony of the integral.
The ninth one is a particular case of the tenth one which we can prove using the fuzzy set $\widetilde{A}(w) = x \ \forall x \in \Omega$ and its complementary, the monotony and symmetry of h. ■

4 CONCLUSIONS AND OPEN PROBLEMS

In this paper we have introduced and studied the concept of local divergence measures defined on a infinite referentials. Its importance lies on the fact that all the applications of the local divergence measures can now be used in the infinite case, which presents some existence and integrability problems which are erased when defining this kind of measures in a coherent way.

We have the following open problems: First, to study the fuzziness measures on infinite referentials which generate these divergence measures. This first problem, has already been analysed and we obtain very satisfactory results which allow us to generate some of the already known fuzziness measures in the infinite case (Minkowski ([5]), Knopfmacher ([4]), etc), and this will be the target of a future paper.

Other immediate open problems will be the extension of all the studies made in the finite referential case to the infinite case.

As further purposes we have the composition of a parallel theory in the finite and infinite case of the studies related to the divergence measures which we are making.

Acknowledgements

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

References

[1] P. Billingsley (1979). Probability and Measure. *John Wiley & Sons, Inc.*.

[2] A. De Luca, and S. Termini (1972). A definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Sets Theory. *Inf. and Control* **20**, pp 301-312.

[3] M. Loevè. Probability theory (1963). *Van Nostrand Reinhold company.*

[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] 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.

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

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

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