Expected pair-wise comparison of the outcomes of a fuzzy random variable

Inés Couso[1], Laura Garrido[1], Susana Montes[1], and Luciano Sánchez[2]

[1] Dep. of Statistics and O.R, University of Oviedo, Spain
couso,garridolaura,montes@uniovi.es
[2] Dep. of Computer Sciences, University of Oviedo, Spain
luciano@uniovi.es

Summary. We introduce the notion of expected pair-wise comparison of a fuzzy random variable. It inludes some well-known parameters such as the quadratic entropy of a random variable, the upper probability induced by a random set or the scalar variance of a fuzzy random variable as particular cases. The special case of expected dissimilitude is highlighted and shown as a useful alternative to the scalar variance when the images of the fuzzy random variable are not necessarily convex, nor in a numerical scale.

Key words: Comparison measure, divergence measure, similarity measure, semantics of fuzzy sets, fuzzy random variable

1 Introduction

Fuzzy random variables (frv for short) were first introduced by Féron in 1976, as functions that assign a fuzzy subset to each possible output of a random experiment, extending the notions of random variable and random set. Later on, several variants were proposed. The different definitions in the literature vary on the measurability conditions imposed to this mapping, and in the properties of the output space, but all of them intend to model situations that combine fuzziness and randomness. Apart from the differences among the formal definitions, fuzzy random variables have been also given different interpretations. Thus, a frv can be viewed ([3]) as a random object, an ill-known random variable or as a conditional upper probability. Each of those interpretations leads to a different way of extending parameters as the expectation, the variance, etc. The case of the variance is discussed in detail in [3]. In this paper, we will treat fuzzy random variables as random objects. We will introduce the notion of *expected pair-wise comparison* and we will discuss in some detail the specific notion of *expected dissimilitude measure*. The expected dissimilitude will average the "degrees of difference" or "divergence" between pairs of outcomes of the frv. In order to find a suitable quantification the differences between two outcomes of the frv, we will provide a brief discussion about some previous notions in the literature such as divergence measures ([8]) or distance measures ([6]), and we will introduce the notion of *dissimilitude*. Once being able to quantify such differences, the expected divergence will average them into a single quantity. We will show how the expected dissimilitude encompasses some different kinds of existing measures: on the one side, it extend the notion of scalar variance of a frv ([3, 7]). On the other hand, it also extends the quadratic entropy of a random variable. So, depending on the specific dissimilitude measure we use, we can extend the notion of variance, entropy, or a mixture of them. Furthermore, it allows us to quantify the expected difference between the different outcomes of the frv when the universe is not a numerical scale. The existing definitions of scalar variance are not easily adaptable to this kind of universes, because they involve the notion of expectation. In this paper, we do not average the dissimilitude degree between each possible outcome and the expectation, but between pairs of outcomes, by taking into account a pair of independent copies of the frv.

The rest of the paper is organized as follows: in Section 2, we briefly discuss the state of art about comparison measures, we provide some new results relating the notions of dissimilarity [2], divergence [8], distance [6] and metric, and we introduce the new notion of *dissimilitude measure*. In Section 3, we propose the concepts of expected pair-wise comparison and expected pair-wise dissimilitude of a frv, studying some interesting properties, and illustrating them with examples. We end the paper with some concluding remarks.

2 Dissimilitude measures for pairs of fuzzy sets

As we pointed out in the last section, an initial step in the construction of an expected dissimilitude measure will be the study of different options to compare pairs of outcomes of a fuzzy random variable. Let us denote by $\mathcal{F}(U)$ the family of fuzzy subsets of a universe U. A *comparison measure* [2] is a mapping $S : \mathcal{F}(U) \times \mathcal{F}(U) \to [0, 1]$ expressed as:

$$S(A, B) = G_S(A \cap B, A - B, B - A),\ \forall\, A, B \in \mathcal{F}(U),$$

for some $G_S : \mathcal{F}(U) \times \mathcal{F}(U) \times \mathcal{F}(U) \to [0, 1]$. This notion includes the idea of similarity and dissimilarity, and other kinds of comparison of pairs of fuzzy sets in a common framework. In this paper, we will slightly relax the assumptions for a comparison measure, and we will not force them to take values in the unit interval. We will pay attention to the quantification of the the degree of "difference" between two fuzzy outcomes. To this purpose, we will survey some previous proposals in the literature and we will check some properties and relations between them.

Montes el al. introduced in [8] an axiomatic definition for the *divergence* between pairs of fuzzy subsets based on the following natural properties:

- It is nonnegative and symmetric.
- It becomes zero when the two fuzzy sets coincide.
- It decreases when two fuzzy sets become "more similar".

Different formalizations of the third idea lead to different axiomatic definitions. Two of them are:

Definition 1. *(Bouchon-Meunier et al., [2]) Consider a universe* U *and let* $\mathcal{F}(U)$ *the family of fuzzy subsets of* U*. A comparison measure* $S : \mathcal{F}(U) \times \mathcal{F}(U) \to \mathbb{R}$ *is a* dissimilarity measure *when, for any pair* $A, B \in \mathcal{F}(U)$ *the following conditions hold:*

Diss1.- G_S *does not depend on its first argument (intersection) and it is increasing in the other two (differences) w.r.t. the fuzzy inclusion.*
Diss2.- $S(A, A) = 0$

Definition 2. *(Montes et al., [8]) Consider a universe* U *and let* $\mathcal{F}(U)$ *the family of fuzzy subsets of* U*. A mapping* $D : \mathcal{F}(U) \times \mathcal{F}(U) \to \mathbb{R}$ *is a* divergence measure *when, for any pair* $A, B \in \mathcal{F}(U)$ *the following conditions hold:*

Div1.- $D(B, A) = D(A, B)$.
Diss2.- $D(A, A) = 0$.
Div3.- $D(A \cup C, B \cup C) \leq D(A, B)$.
Div4.- $D(A \cap C, B \cap C) \leq D(A, B)$.

There is a strong relationship between dissimilarities and divergences: according to the following result, the divergence between two crisp sets A and B does not depend on their intersection, and it increases with their difference. (We omit the proof.)

Proposition 1. *Consider the function* $D : \mathcal{F}(U) \times \mathcal{F}(U) \to \mathbb{R}$

- *If* D *satisfies axiom Div3, then* $D(A, B) \leq D(A-B, B-A)$, $\forall A, B \in \wp(U)$ *(the class of crisp subsets of* U*).*
- *If* D *satisfies axiom Div4, then* $D(A, B) \geq D(A - B, B - A)$, $\forall A, B \in \mathcal{F}(U)$.
- *If* D *satisfies axiom Div4, then* $D(A, B) \leq D(C, B)$*, for all* $A, B, C \in \wp(U)$ *such that* $C \cap B = \emptyset$ *and* $A \subseteq C$.

Thus, the above measures (divergence and dissimilarity measures) focus on the differences between two fuzzy sets, but they do not care about their similarities. Sometimes, we need to take into account the similarities between sets, other times, we do not. Let us illustrate this with an easy example.

Example 1. Consider the set of languages:

$U = \{$English (e), Spanish (s), French (f), Italian (i), Dutch (d), Russian (r)$\}$

and let the crisp subsets $E = \{e, s, f, i, d\}$, $G = \{e, s, f, d, r\}$, $A = \{i\}$ and $V = \{r\}$ denote the respective communication skills of four persons called Enrique, Gert, Angelo and Vladimir. Enrique and Gert share much more language skills than Angelo and Vladimir, but those commonalities cannot be detected by means of the above measures (divergences and dissimilarities). If we just wanted to focus on the differences, those measures would be useful. But, if also take into account their common skills, we should use different comparison measures.

We can find in the literature some other measures that detect the differences between two fuzzy subsets, but they are not necessarily independent on the commonalities. Let us show some of them:

Definition 3. *(Fan J, Xie W, [6]) Consider a universe U and let $\mathcal{F}(U)$ the family of fuzzy subsets of U. A mapping $d : \mathcal{F}(U) \times \mathcal{F}(U) \to \mathbb{R}$ is a* distance measure *when:*

Div1.- $d(B, A) = d(A, B),\ \forall A, B \in \mathcal{F}(U)$.
Diss2.- $d(A, A) = 0,\ \forall A \in \mathcal{F}(U)$.
DM3.- $d(D, D^c) = \max_{A,B\in\mathcal{F}(U)} d(A, B)$, *for any crisp set* $D \in \wp(U)$.
DM4.- If $A \subseteq B \subseteq C$, *then* $\max\{d(A, B), d(B, C)\} \leq d(A, C)$.

There are some relationships between divergence and distance measures. In fact, it is checked in [8] that any function satisfying Div3 and Div4 fulfills DM4. Furthermore, any *local*[3] divergence satisfies DM3. Thus, any local divergence measure satisfies Definition 3. Let us mention that the term *distance measure* is used in [6] without referring to the mathematical notion of *metric*. Nevertheless, both notions are somehow related, as they quantify the degree of difference between fuzzy subsets. We can find in the recent literature some metrics and pseudo-metrics defined on classes of fuzzy sets, such as the well known Hamming distance, the Puri-Ralescu [9] pseudo-metric[4] and other families of metrics proposed in [1, 7, 10] on some specific classes of convex fuzzy sets, for instance.

We can find some relationships between the above metrics and the notions of divergence, distance and comparison measure. In this short paper, we will only list them, without referring to formal details about the domain of definition of each measure, and without detailing the proofs:

Proposition 2.

- *All the metrics and pseudo-metrics cited above can be expressed as comparison measures on their respective domains of definition and they satisfy axioms Div1, Diss2, Div3 and DM4.*

[3] A divergence measure is called *local* ([8]) when there exists a function $h : [0, 1] \times [0, 1] \to \mathbb{R}$ such that: $D(A, B) - D(A \cup \{x\}, B \cup \{x\}) = h(A(x), B(x)),\ \forall x \in U$.

[4] Puri and Ralescu introduce a metric in the class of fuzzy subsets of $\mathbb{R}^n$ with compact and non-empty level cuts. It can be easily extended to more general families of fuzzy subsets as a pseudo-metric.

- *Only the Hamming distance satisfies axioms Diss1, Div4 and DM3.*

According to the above proposition, axioms Diss1, Div4 and DM3 exclude most of the mentioned (pseudo)-metrics. In the rest of the paper, we will use the term *dissimilitude measure* for those comparison measures safisfying Div1, Diss2, Div3 and DM4.

3 Expected pair-wise comparison of a fuzzy random variable

Consider a probability space $(\Omega, \mathcal{A}, P)$ and a frv defined on it, i.e., an $\mathcal{A}-\sigma$ measurable mapping $\tilde{X} : \Omega \to \mathcal{F}$, where σ is a σ-field defined on a certain class of fuzzy subsets $\mathcal{F} \subseteq \mathcal{F}(U)$. (This general definition encompasses several specific proposals in the literature). Any fuzzy random variable induces a probability measure on σ by means of the formula:

$$P_{\tilde{X}}(\mathcal{C}) = P(\{\omega \in \Omega \,:\, \tilde{X}(\omega) \in \mathcal{C}\}),\ \forall \mathcal{C} \in \sigma.$$

Now consider the product probability $P \otimes P : \mathcal{A} \otimes \mathcal{A} \to [0,1]$ as the only probability measure satisfying the restriction

$$(P \otimes P)(A \times B) = P(A) \cdot P(B)\ \forall A, B \in \mathcal{A}.$$

Let $\tilde{X}_1$ and $\tilde{X}_2$ two (identically distributed) copies of $\tilde{X}$, and consider a comparison measure on $\mathcal{F}$, $S : \mathcal{F} \times \mathcal{F} \to [0,1]$.

Definition 4. *We define the* expected pair-wise comparison *of* $\tilde{X}$ *as the quantity*

$$E_S(\tilde{X}) = \int_{\Omega \times \Omega} S(\tilde{X}_1(\omega), \tilde{X}_2(\omega'))\, d(P \otimes P)(\omega, \omega'),$$

provided that the mapping $g(\omega, \omega') = S(\tilde{X}_1(\omega), \tilde{X}_2(\omega')),\ \forall (\omega, \omega') \in \Omega \times \Omega$ *is* $\mathcal{A} \otimes \mathcal{A} - \beta_{\mathbb{R}}$ *measurable.*

The above definition generalizes some well-known quantities, as we show in the following examples.

Example 2. Consider a finite population Ω and the set of languages $U = \{e, s, f, i, d, r\}$ of Example 1. Consider the multi-valued mapping $\Gamma : \Omega \to \wp(U)$ that assigns to each person $\omega \in \Omega$ the subset of languages in U (s)he can speak[5]. The following expected pair-wise comparison measures provide interesting information about such attribute:

[5]Let us assume that $\Gamma(\omega) \neq \emptyset,\ \forall \omega \in \Omega$, so everybody is assumed to be able to speak some language in the set U. Multi-valued mappings represent special cases of fuzzy-valued mappings. Furthermore, if we consider the power set as the initial σ-field, they are measurable with respect to any σ-field on the final space.

- Let us fix an arbitrary subset $D \subseteq U$. Consider the comparison measure S_1 such that G_{S_1} is defined as $G_{S_1}(A,B,C) = \max\{M(A \cap B), M(A \cap B^c)\}$, where
$$M(E) = \begin{cases} 1 & \text{if } E \cap D \neq \emptyset, \\ 0 & \textit{otherwise.} \end{cases}$$
The expected pair-wise comparison of Γ, $E_{S_1}(\Gamma)$ coincides with the *upper probability* ([4]) of D and it represents the proportion of persons in the population that speak some of the languages included in D. If, for instance, D is equal to $\{d\}$, then $E_{S_1}(\Gamma)$ represents the proportion of persons that can speak Dutch, at least.
- Let us fix again an arbitrary subset $D \subseteq U$. Consider the comparison measure S_2 such that G_{S_2} is defined as $G_{S_2}(A,B,C) = \min\{M(A \cap B), M(A \cap B^c)\}$, where
$$M(E) = \begin{cases} 1 & \text{if } E \subseteq D, \\ 0 & \textit{otherwise.} \end{cases}$$
The expected pair-wise comparison of Γ, $E_{S_2}(\Gamma)$, coincides with the *lower probability* ([4]) of D and it represents the proportion of persons in the population that do not speak any language outside D.
- Consider the comparison measure $S_3(A,B) = \#(A \cap B)$. The expected pair-wise comparison $E_{S_3}(\Gamma)$ averages the capacity of communication between pairs of people in the population.
- Consider the Hamming distance $S_4(A,B) = d_H(A,B) = \#(A \triangle B)$. The expected pair-wise comparison $E_{S_4}(\Gamma)$ represents a degree of divergence about the language skills of the people in the population.

Example 3. The above example can be modified if we have more refined information about the communication skills of the people. We can use a frv $\tilde{X} : \Omega \to \mathcal{F}(U)$ to represent those abilities. The membership value $\tilde{X}(\omega)(u)$ will represent a degree of preference ([5]) in a $[0,1]$ scale for the language $u \in U$. Thus $\tilde{X}(\omega)(u) > \tilde{X}(\omega)(u')$ will mean that the person ω prefers to speak u than u', because (s)he is more familiar with it. Those degrees of preference can be determined as a function of the CEFR levels, for instance. For a specific dissimilitude measure, the expected dissimilitude of $\tilde{X}$ reflects an expected degree of difference in the language skills between pairs of persons in the population.

Example 4. Consider a set of days, Ω, and consider the multi-valued mapping $\Gamma : \Omega \to \wp(\mathbb{R})$, where $\Gamma(\omega) = [L(\omega), U(\omega)]$ represents the interval of minimum and maximum temperatures attained in Mieres on a date ω. Several expected pair-wise comparison measures return different informative quantities such as: the variance of the min temperatures, the variance of the max temperatures, a mixture (linear combination) of both variances, the variance of the amplitudes of the min-max intervals, the proportion of days where the min temperature exceeds a certain threshold, the variance of the middle points of the intervals, etc.

The particular case where the comparison measure S is a dissimilitude is remarkable. It extends some key notions in the literature, as we show in the following remarks.

Remark 1. On the one hand, it extends the notion of quadratic entropy of a random variable: If $\tilde{X}$ represents a random variable X on a finite universe $U = \{u_1, \dots, u_n\}$ in the sense that $\tilde{X}(\omega) = \{X(\omega)\},\ \forall\, \omega \in \Omega$, and $S(A,B) = d_H(A,B) = \#A \triangle B$ is the Hamming distance, then the expected pair-wise comparison of $\tilde{X}$ is the quadratic entropy of X:

$$E_S(\tilde{X}) = \sum_{i=1}^{n}\sum_{j=1}^{n} d_H(\{x_i\},\{x_j\}) p_i \cdot p_j = \sum_{i=1}^{n}\sum_{j=1}^{n}(1-\delta_{ij}) p_i \cdot p_j = 1 - \sum_{i=1}^{n} p_i^2,$$

where p_i denotes the probability $P(X = u_i)$, $i = 1, \dots n$.

Remark 2. On the other hand, it extends some notions of scalar variance of a fuzzy random variable [3] in the literature: all the (pseudo-)metrics considered at the end of Section 2 satisfy the properties of dissimilitude measures. Furthermore, any non-decreasing function of a similitude satisfying the boundary condition $g(0) = 0$ is also a dissimilitude. If we construct the similitude measure $S = \frac{d^2}{2}$ on the basis of any of those distances d, and we take into account the specific arithmetic used in each context, in order to avoid the explicit use of the expectation, we can extend the existing notions of scalar variances [3] in the literature. Furthermore, expected dissimilitude measures even apply when the images of the frv are not necessarily convex, and/or they do not lay in a numerical scale, as we have illustrated in Example 2.

Some general properties of expected dissimilitude measures are given in the following proposition.

Proposition 3. *Let $S : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ be a dissimilitude measure. Then:*

- $E_S(A) = 0$, $\forall\, A \in \mathcal{F}$.
- $E_S(\tilde{X} \cup A) \le E_S(\tilde{X})$, *for all frv* $\tilde{X}$ *and all* $A \in \mathcal{F}$.
- *If* $S(A,B) = \sum_{x \in U} g(A(x), B(x))$ *then* $E(\tilde{X} \cup \tilde{Y}) \le E_S(\tilde{X}) + E_S(\tilde{Y})$, *for all frv* $\tilde{X}$ *and* $\tilde{Y}$.

Example 5. We can illustrate the above properties by referring to the language skills of Example 2. The first property would mean that the expected dissimilitude is null when everybody in the population owns the same communication skills. For the second property, let us assume that all the people in that population that do not speak Spanish take a course on this language. Then, the expected dissimilitude measure should decrease. Finally, suppose that we consider two separate groups of languages, and we consider the communication skills of the people within each group ($\tilde{X}$ denotes the abilities within the first group of languages, and $\tilde{Y}$ denotes the abilities within the second group.) Then, the expected dissimilitude in the whole set of languages cannot be strictly greater than the sum of the expected dissimilitude values within each group.

4 Concluding remarks

The notion of expected comparison of a fuzzy random variable encompasses several well known parameters associated to random variables, random sets and fuzzy random variables. In particular, the expected dissimilitude quantifies the dispersion of the outcomes of a fuzzy random variable. It generalizes some entropies for random variables and also some scalar variances of fuzzy random variables. The existing definitions of scalar variances that we can find in the literature [3, 7] are restricted to those situations where the outcomes of the frv are convex fuzzy subsets of $\mathbb{R}^n$. The new definition applies in a variety of situations, even for the cases where there is not a numerical scale. We have illustrated the utility of the new notion with several examples. In future works we plan to study some additional properties of the expected dissimilitude, for some specific dissimilitude measures, trying to lay bare the connection with the general notions of entropy and dispersion.

Acknowledgements This work has been supported by the FEDER-MEC Grants TIN2007-67418-C03-03, TIN2008-06681-C06-04 and MTM2007-61193.

References

1. Bertoluzza C, Corral N, Salas A, On a new class of distances between fuzzy numbers, Mathware & Soft Computing 2 (1995) 71-84.
2. Bouchon-Meunier B, Rifqi M, Bothorel S, Towards general measures of comparison of objects, Fuzzy Sets and Systems 84 (1996) 143-153.
3. Couso I, Dubois D, On the variability of the concept of variance of a fuzzy random variable, IEEE Transactions on Fuzzy Systems 17 (2009) 1070 - 1080.
4. Dempster AP, Upper and Lower Probabilities Induced by a Multivalued Mapping, Annals of Mathematical Statistics 38 (1967) 325-339.
5. Dubois D, Prade H, The three semantics of fuzzy sets, Fuzzy Sets and Systems 90 (1997) 141-150.
6. Fan J, Xie W, Some notes on similarity measure and proximity measure, Fuzzy Sets and Systems 101 (1999) 403-412.
7. Körner R, On the variance of fuzzy random variables, Fuzzy Sets and Systems 92 , (1997) 83-93.
8. Montes S, Couso I, Gil P, Bertoluzza C, Divergence measure between fuzzy sets, International Journal of Approximate Reasoning 30 (2002) 91–105.
9. Puri ML, Ralescu DA, Fuzzy random variables, Journal of Mathematical Analysis and Applications 114 (1986) 409–422.
10. Trutschnig W, González-Rodríguez G, Colubi A, Gil MA, A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread 179 (2009) 3964-3972.