A possibilistic interpretation of the expectation of a fuzzy random variable

Inés Couso[1], Enrique Miranda[2], and Gert de Cooman[3]

[1] University of Oviedo (Spain), Dep. of Statistics and O.R. couso@pinon.ccu.uniovi.es
[2] Rey Juan Carlos University (Spain), Dep. of Informatics, Statistics and Telematics emiranda@escet.urjc.es
[3] Universiteit Gent (Belgium), Onderzoeksroep SYSTeMS. gert.decooman@ugent.be

1 Introduction

Since their first introduction, fuzzy random variables have been given a number of different definitions. In [10], Krätschmer gives an unified approach of all of them. All of these authors try to model situations where both randomness and imprecision are present, and they define a fuzzy random variable as a map assigning to any element of the initial space a fuzzy subset of the final space; however, they differ in the measurability condition imposed on this map and in the characteristics of the final space. On the other hand, the study of statistical parameters, such as expectation and variance, of a fuzzy random variable has followed two different approaches: some authors define them as fuzzy sets ([1, 11, 14, 17]); others as (crisp) numerical values, as in [9, 12, 13]. One of the reasons for the existence of these different approaches is that a fuzzy set can be given many different interpretations, as the survey conducted in [7] testifies; and any of these interpretations can be carried over to fuzzy random variables and their parameters. In the present paper, we intend to give fuzzy random variables a *possibilistic* interpretation. The value of a fuzzy set in a point will represent a degree of possibility, which is a specific type of upper probability. We shall see that this interpretation fits nicely into the framework of the theory of imprecise probabilities ([18]), and we shall be able to associate with any statistical parameter of interest an interval of possible values. This is a compromise between the two approaches considered above (precise numerical values and fuzzy sets): as sets of possible values these intervals have a straightforward interpretation in the context of the theory of imprecise probabilities, and the fact that these intervals do not generally reduce to a precise single value allows us to take into account the imprecision that a fuzzy random variable represents.

We shall focus here on the expectation of a fuzzy random variable, for which our approach yields an interval of possible values that is, in general, sharper than the support of the fuzzy expectation defined by Puri and Ralescu in [17]. Indeed, we shall prove that it coincides with the mean value ([6]) of this fuzzy expectation. We also present some additional discussion and give a number of interesting additional properties for our expectation.

2 Preliminary concepts

Let us first introduce some notation we will use along the paper. We denote the power set of Ω by $\wp(\Omega)$. The notations $\beta_{[0,1]}$ and $\beta_{\mathbb{R}}$ respectively stand for the (usual) Borel σ-algebras on the unit interval and on the space of real numbers. We will denote by $\lambda_{[0,1]}$ the uniform distribution on the unit interval. $(\mathrm{C})\int_A Y\,d\mu$ is defined to be the asymmetric Choquet integral of Y with respect to μ on the set A. id will denote the identity function.

As we stated in the introduction, a fuzzy random variable is a map that assigns to any element of the initial space a fuzzy subset of the final space. The definitions proposed in the literature differ with respect to the assumptions made about the final space, as well as to the measurability condition imposed on this map. We shall work with the definition given by Puri and Ralescu in [17]: a fuzzy random variable assumes values in the class $\mathcal{F}_0(\mathbb{R}^n)$ of fuzzy sets $u\colon \mathbb{R}^n \to [0,1]$ satisfying:

(a) $u_\alpha = \{x \in \mathbb{R}^n \colon u(x) \geq \alpha\}$ (the weak α-cut of u) is compact for all $\alpha \in (0,1]$.
(b) $\{x \in \mathbb{R}^n \colon u(x) = 1\} \neq \emptyset$.

The measurability condition is based upon the notion of strong measurability of a multi-valued map.

Definition 1 ([16]). *Let $(\Omega,\mathcal{A})$ and $(\Omega',\mathcal{A}')$ be two measurable spaces. A multi-valued map $\Gamma\colon \Omega \to \wp(\Omega')$ is called* **strongly measurable** *when*

$$\Gamma^*(A) = \{\omega \in \Omega \colon \Gamma(\omega) \cap A \neq \emptyset\} \in \mathcal{A}, \quad \forall A \in \mathcal{A}'.$$

We refer to [8] for a thorough study of this condition. We now come to the notion of a fuzzy random variable.

Definition 2. *Let $(\Omega,\mathcal{A},P)$ be a probability space. A map $\tilde{X}\colon \Omega \to \mathcal{F}_0(\mathbb{R}^n)$ is a* **fuzzy random variable** *if, for any $\alpha \in (0,1]$, the multi-valued map $\tilde{X}_\alpha\colon \Omega \to \wp(\mathbb{R}^n)$, defined by $\tilde{X}_\alpha(\omega) = \left\{x \in \mathbb{R}^n \colon \tilde{X}(\omega)(x) \geq \alpha\right\}$ is strongly measurable with respect to $\mathcal{A}$ and $\beta_{\mathbb{R}^n}$.*

We shall consider the possibilistic interpretation of fuzzy sets. On this interpretation, the image $\tilde{X}(\omega)$ of any $\omega \in \Omega$, is a map from $\mathbb{R}^n$ to $[0,1]$ that represents a possibility distribution on $\mathbb{R}^n$. More specifically, we assume that we have two experiments: one taking values on a set Ω, which is determined by a probability measure P defined on a σ-field of sets $\mathcal{A} \subseteq \wp(\Omega)$; and one taking values on $\mathbb{R}^n$ for which we have imprecise information. The relationship between the two experiments is given by the fuzzy random variable $\tilde{X}$: if the outcome of the first experiment is $\omega \in \Omega$, then our information about the outcome of the second experiment is given by the fuzzy set $\tilde{X}(\omega)$, which determines a (conditional) possibility measure on $\mathbb{R}^n$ that we shall denote by $\Pi(\cdot|\omega)$. In other words, $\Pi(A|\omega)$ represents the degree of possibility (or the upper probability) that the outcome of the second experiment belongs to A, if we know that the outcome of the first has been ω. This representation of our knowledge is related to the work developed in [15].

We shall also assume that there is a (precise but unknown) probability defining the relationship between the two experiments, i.e., for any ω in the initial space we assume the existence of a probability measure on the final space that represents the law governing the second experiment when the outcome of the first is ω. Therefore, the relationship between the two experiments is given by a **transition probability** $Q(\cdot|\cdot)$ on $\beta_{\mathbb{R}^n} \times \Omega$, i.e., a function such that:

(a) $Q(\cdot|\omega)$ is a probability measure for all $\omega \in \Omega$.
(b) $Q(A|\cdot)$ is $\mathcal{A} - \beta_{[0,1]}$-measurable for all $A \in \beta_{\mathbb{R}^n}$,

and the available knowledge about this transition probability is modelled by the conditional possibility measures $\{\Pi(\cdot|\omega)\}_{\omega \in \Omega}$, in the sense that $Q(\cdot|\omega) \leq \Pi(\cdot|\omega)$ for all $\omega \in \Omega$.

This model is related to the interpretation considered by Kruse and Meyer in [11], but it is nevertheless more general: these authors assume that the fuzzy random variable is a model for the imprecise observation of a random variable. Hence, for any ω of the initial space $\tilde{X}(\omega)(x)$ is the possibility that we give to x being the 'true' image of ω, and therefore, following the previous notation, they would only consider degenerate probability measures $Q(\cdot|\omega)$. Nevertheless, the reasoning made in [11] leads naturally to a second-order possibility distribution (such as those considered in [3, 5]), that is, to a possibility measure defined on the class of probability measures; we shall see that the combination of the information present in our model provides us with a first-order model.

In order to make the results of this paper clearer to the reader, let us recall some ideas from (Classical) Probability Theory. Let us consider a probability measure, P, on $(\Omega, \mathcal{A})$, and a transition probability, $Q(\cdot|\cdot)$, between Ω and $\mathbb{R}^n$, as described above. We can combine them to construct a probability measure on the product space $(\Omega \times \mathbb{R}^n, \mathcal{A} \otimes \beta_{\mathbb{R}^n})$. Its marginal on $\beta_{\mathbb{R}^n}$ is given by the formula:

$$Q_2(A) = \int_{\Omega} Q(A \mid \omega) \, \mathrm{d}P(\omega), \ \forall A \in \beta_{\mathbb{R}^n}. \quad (1)$$

On the other hand, the expected value of an integrable random variable, $Y : \mathbb{R}^n \to \mathbb{R}$ with respect to Q_2 will be given by the formula:

$$\int_\Omega Y \, \mathrm{d}Q_2 = \int_\Omega \left(\int_{\mathbb{R}^n} Y \, \mathrm{d}Q(\cdot \mid \omega) \right) \, \mathrm{d}P(\omega).$$

Let us now suppose that each probability measure $Q(\cdot \mid \omega)$ is degenerate on some $X(\omega) \in \mathbb{R}$. It is easy to prove that the mapping $X : \Omega \to \mathbb{R}$ is measurable and its induced probability measure coincides with Q_2. On the other hand, its expectation with respect to P can be calculated as follows:

$$\int_\Omega X \, \mathrm{d}P = \int_{\mathbb{R}} \mathrm{id} \, \mathrm{d}P_X = \int_{\mathbb{R}} \mathrm{id} \, \mathrm{d}Q_2 = \int_\Omega \left(\int_{\mathbb{R}} \mathrm{id} \, \mathrm{d}Q(\cdot \mid \omega) \right) \, \mathrm{d}P(\omega).$$

3 Fuzzy random variables as conditional possibility distributions

As we pointed out in the preceding section, our model considers either the probability measure P over $\mathcal{A}$, that governs the first sub-experiment, and the family of conditional possibility measures $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega}$ defined as follows:

$$\Pi(A|\omega) = \sup_{x\in A} \tilde{X}(\omega)(x),$$

for all $A \in \beta_{\mathbb{R}^n}$ and $\omega \in \Omega$. In words, the value $\Pi(A|\omega)$ is the upper bound of the probability of the final result being in A, provided that the outcome of the initial experiment is ω.

Since we have assumed that $\tilde{X}(\omega)$ is in the class $\mathcal{F}_0(\mathbb{R}^n)$, the possibility measure $\Pi(\cdot|\omega)$ is normal for every ω. Therefore, recalling the results in [2], we know that $\Pi(\cdot|\omega)$ is the upper envelope of the family of (σ-additive) probability measures dominated by it. If we combine every transition probability measure compatible with $\Pi(\cdot|\cdot)$ with the probability measure P, as described in Eq. (1), we get a marginal probability measure over $\beta_{\mathbb{R}^n}$. Indeed, the supremum of all of these probability measures is the upper probability describing the available information about the second sub-experiment. On the other hand, the supremum of the expectations of the identity function with respect to such probability measures (when defined over $\beta_{\mathbb{R}}$) will represent an upper bound of the expectation of the imprecisely known probability distribution.

The expressions for both suprema, as a function of the probability measure P and the family of (conditional) possibility measures $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega}$, will be developed near the end of this section. We begin here with a general result that allows us to express them as particular cases.

Theorem 1. *Consider a probability space $(\Omega, \mathcal{A}, P)$ and a fuzzy random variable $\tilde{X} \colon \Omega \to \mathcal{F}_0(\mathbb{R}^n)$, and denote its induced family of possibility measures by $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega}$. Let us define*

$$\mathcal{H} = \{Q(\cdot|\cdot) \textit{ transition prob.}\colon Q(A|\omega) \leq \Pi(A|\omega), \quad \forall\omega \in \Omega, \quad \forall A \in \beta_{\mathbb{R}^n}\}.$$

Given a random variable $Y\colon \mathbb{R}^n \to \mathbb{R}$ that is bounded below, the following equality holds:

$$\int_\Omega \left((C) \int_{\mathbb{R}^n} Y \,\mathrm{d}\Pi(\cdot|\omega) \right) \mathrm{d}P(\omega) = \sup_{Q(\cdot|\cdot)\in\mathcal{H}} \int_\Omega \left(\int_{\mathbb{R}^n} Y \,\mathrm{d}Q(\cdot|\omega) \right) \mathrm{d}P(\omega). \tag{2}$$

Although in this theorem we have focused on the upper bound, it is possible to establish a similar result with respect to the lower bound. That is, if we consider the family of necessity measures $\{N(\cdot|\omega)\}_{\omega\in\Omega}$ that can be derived from $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega}$ using conjugacy, it is easily checked that $\int_\Omega((C)\int_{\mathbb{R}^n} Y \,\mathrm{d}N(\cdot|\omega))\,\mathrm{d}P(\omega)$ coincides with the infimum of the set of values considered on the right hand side of Eq. (2).

The result has two other interesting consequences. Since in particular $\Pi(A|\omega) = (C)\int_{\mathbb{R}^n} I_A \,\mathrm{d}\Pi(\cdot|\omega)$, we find for $Y = I_A$ that

$$\int_\Omega \Pi(A|\omega)\,\mathrm{d}P(\omega) = \sup\left\{ \int_\Omega Q(A|\omega)\,\mathrm{d}P(\omega)\colon Q(\cdot|\cdot) \in \mathcal{H} \right\}$$

for any $A \in \beta_{\mathbb{R}^n}$. This means that $\int_\Omega \Pi(A|\omega)\,\mathrm{d}P(\omega)$ is the smallest upper bound that we can give to the probability of A, taking into account the information provided by P and $\tilde{X}$. We shall call this number the **upper probability** $\overline{P}_{\tilde{X}}(A)$ that $\tilde{X}$ assumes a value in the set A (and similarly for the **lower probability** $\underline{P}_{\tilde{X}}(A)$).

Secondly, if $n = 1$ and the support of $\tilde{X}(\omega)$ is a compact set for each $\omega \in \Omega$, recalling the results in [4] it is easily checked that

$$\int_\Omega \left((C) \int_{\mathbb{R}} \mathrm{id}\,\mathrm{d}\Pi(\cdot|\omega) \right) \mathrm{d}P(\omega) = \sup_{Q(\cdot|\cdot)\in\mathcal{H}} \left\{ \int_\Omega \left(\int_{\mathbb{R}} \mathrm{id}\,\mathrm{d}Q(\cdot|\omega) \right) \mathrm{d}P(\omega) \right\}.$$

Thus, the first term of the equality is then the smallest upper bound we can give to the expectation of a random variable whose imprecise observation is modelled by the fuzzy random variable $\tilde{X}$. We shall call this number the **upper expectation** $\overline{E}(\tilde{X})$ of the fuzzy random variable $\tilde{X}$. A similar comment can be made with respect to the greatest lower bound, called the **lower expectation** $\underline{E}(\tilde{X})$. In the next section, we shall concentrate on the study of these lower and upper expectations, and relate them to some existing definitions for the *expectation* of a fuzzy random variable in the literature.

To conclude our discussion of Theorem 1, let us comment on how this result relates to Walley's theory of imprecise probabilities ([18]). It turns out that the left hand side of Eq. (2) is an expression for what Walley would call the *marginal extension* of the marginal precise probability P and the conditional upper probability $\Pi(\cdot|\cdot)$ (see [18, Section 6.7]). The fact that this marginal extension is equal to the expression on the right hand side, namely an upper envelope of the marginal extensions of P and the (σ-additive) conditional probabilities compatible with $\Pi(\cdot|\cdot)$, reminds us of his Lower Envelope Theorem 6.7.4 in [18], where he proves a similar result but for envelopes involving *finitely* additive conditional probabilities.

4 The expectation of a fuzzy random variable

Let us next introduce the concept of expectation established by Puri and Ralescu in [17]. (We refer the reader to this paper for a further explanation.) It is based on the notion of an Aumann integral of a random set. Consider a probability space $(\Omega, \mathcal{A}, P)$ and a multi-valued map $\Gamma\colon \Omega \to \wp(\mathbb{R}^n)$. Γ is said to be **integrably bounded** if there is an integrable map $h\colon \Omega \to \mathbb{R}$ such that $||x|| \leq h(\omega)$ for all $x \in \Gamma(\omega)$ and all $\omega \in \Omega$. The **Aumann integral** of an integrably bounded multi-valued map Γ is defined as the set

$$(A)\int_{\Omega} \Gamma \,\mathrm{d}P := \left\{ \int_{\Omega} f \,\mathrm{d}P\colon f(\omega) \in \Gamma(\omega),\, \forall \omega \in \Omega,\ f \text{ measurable} \right\}.$$

Now, a fuzzy random variable $\tilde{X}\colon \Omega \to \mathcal{F}_0(\mathbb{R}^n)$ is called integrably bounded if $\tilde{X}_\alpha$ is integrably bounded for any $\alpha \in (0,1]$, and the **expectation** [17] of an integrably bounded fuzzy random variable $\tilde{X}\colon \Omega \to \mathcal{F}_0(\mathbb{R}^n)$ is the unique fuzzy set whose α-cuts are given by

$$[E(\tilde{X})]_\alpha := (A)\int_{\Omega} \tilde{X}_\alpha \,\mathrm{d}P, \quad \forall \alpha \in (0,1].$$

This concept of expectation is compatible with a second order possibility model, as that proposed in [3]: if we assume the existence of some unknown random variable modelling the relationship between the two experiments, this fuzzy expectation could be considered as a possibility distribution on the set of possible values for the 'true' expectation.

On the other hand, given a fuzzy number $u\colon \mathbb{R} \to [0,1]$, Dubois and Prade ([6]) define its **mean value** as

$$M(u) := \left\{ \int_{\mathbb{R}} \mathrm{id}\,\mathrm{d}P\colon P\colon \beta_{\mathbb{R}} \to [0,1] \text{ probability},\ P \leq U \right\}.$$

This is the set of possible values for the expectation of a random experiment whose probability distribution is dominated by the possibility measure U induced by u. We now relate the mean value of the expectation of Puri and Ralescu to the lower and upper expectations of $\tilde{X}$.

Theorem 2. *Consider a probability space $(\Omega, \mathcal{A}, P)$ and an integrably bounded fuzzy random variable $\tilde{X}\colon \Omega \to \mathcal{F}_0(\mathbb{R})$, and let $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega}$ denote the induced family of possibility measures. Then*

$$\int_{\Omega} \left((C) \int_{\mathbb{R}} \mathrm{id}\,\mathrm{d}\Pi(\cdot|\omega) \right) \mathrm{d}P(\omega) = \sup M(E(\tilde{X})).$$

It is clear that we may derive an analogous result with respect to the infimum value of $M(E(\tilde{X}))$ and the integral with respect to the necessity measures $N(\cdot|\omega)$ conjugate to the $\Pi(\cdot|\omega)$, i.e., the lower expectation of $\tilde{X}$. Therefore, under the possibilistic interpretation, the mean value of the fuzzy expectation of a fuzzy random variable (with compact support and integrably bounded) is the set of possible values for the expectation associated with the original probability distribution, taking into account the available information.

Let us show next that this value generally speaking gives us more information than the expectation of Puri and Ralescu, in the sense that not all the values in the support of that expectation are a possible value for the expectation with respect to the 'true' probability distribution.

Example 1. Consider the probability space $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ and the fuzzy random variable $\tilde{X}\colon [0,1] \to \mathcal{F}_0(\mathbb{R})$ given, for all $\omega \in [0,1]$, by

$$\tilde{X}(\omega)(x) = \begin{cases} 2\,(x-\omega) & \text{if } \omega \leq x \leq \omega + 1/2 \\ 2\,(\omega+1-x) & \text{if } \omega + 1/2 < x \leq \omega + 1 \\ 0 & \text{otherwise} \end{cases}$$

The support of Puri and Ralescu's expectation is $\mathrm{supp}(E(\tilde{X})) = [1/2, 3/2]$. Its mean value, $M(E(\tilde{X})) = [3/4, 5/4]$, is strictly included in the support. ♦

5 Conclusions and open problems

The model proposed in this paper for the representation of the information provided by a fuzzy random variable is a compromise between the fuzzy models proposed in [3, 11, 14] and the precise models considered in [9, 12, 13]: ours is a first order model, and associates to any parameter of the (unknown) probability distribution an upper and a lower bound. We have shown that, in the case of the expectation with respect to this unknown distribution, these bounds are the most precise ones we can consider, taking into account the available information. Moreover, the interval they determine coincides with the mean value of the expectation defined by Puri and Ralescu.

In the future, we intend to compare our model with the second order model considered in [3]. In [19], Walley proposes a method to reduce second-order models into first-order ones, using techniques of natural extension. It would be interesting to see whether the reduction of the model considered in [3] leads to the model we propose in this paper, and, more generally, whether Walley's reduction method is equivalent to calculating the mean value of the possibility distributions involved in the second-order model. This would be essential if we want to develop a unified theory.

Acknowledgements We acknowledge the financial support of the project BFM-2001-3515, and research grant G.0139.01 of the Flemish Fund for Scientific Research (FWO).

References

1. I. Couso, S. Montes, and P. Gil. (1998). Función de distribución y mediana de variables aleatorias difusas. *Actas de ESTYLF'98*, pp. 279–284, Pamplona (Spain). In Spanish.
2. I. Couso, S. Montes, and P. Gil. (2001). The necessity of the strong α-cuts of a fuzzy set. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **9**, 249–262.
3. I. Couso, S. Montes, and P. Gil. (2002). Second order possibility measure induced by a fuzzy random variable. In C. Bertoluzza, M. A. Gil, and D. A. Ralescu (eds) *Statistical modeling, analysis and management of fuzzy data*. Springer, Heidelberg, pp. 127–144.
4. I. Couso, S. Montes, and P. Gil. (2002). Stochastic convergence, uniform integrability and convergence in mean in fuzzy measure spaces. *Fuzzy Sets and Systems* **129**, 95–104.
5. G. de Cooman and P. Walley. (2002). An imprecise hierarchical model for behaviour under uncertainty. *Theory and Decision* **52**, 327–374.
6. D. Dubois and H. Prade. (1987). The mean value of a fuzzy number. *Fuzzy Sets and Systems* **24**, 279–300.
7. D. Dubois and H. Prade. (1997). The three semantics of fuzzy sets. *Fuzzy Sets and Systems* **90**, 141–150.
8. C. J. Himmelberg. (1975). Measurable relations. *Fundamenta Mathematicae* **87**, 53–72.
9. R. Körner. (1997). On the variance of fuzzy random variables. *Fuzzy Sets and Systems* **92**, 83–93.
10. V. Krätschmer. (2001). A unified approach to fuzzy random variables. *Fuzzy Sets and Systems* **123**, 1–9.
11. R. Kruse and K. D. Meyer. (1987). *Statistics with vague data*. D. Reidel Publishing Company, Dordrecht.
12. Y. K. Liu and B. Liu. (2003). A class of fuzzy random optimization: expected value models. *Information Sciences* **155**, 89–102.
13. M. A. Lubiano, M. A. Gil, M. López-Díaz, and M. T. López-García. (2000). The $\overrightarrow{\lambda}$-mean squared dispersion associated with a fuzzy random variable. *Fuzzy Sets and Systems* **111**, 307–317.
14. K. D. Meyer and R. Kruse. (1990). On calculating the covariance in the presence of fuzzy data. In W. H. Janko, M. Roubens, and H. J. Zimmerman (eds), *Progress in fuzzy sets and systems*. Kluwer Academic Publishers, Dordrecht.
15. E. Miranda, G. de Cooman, and I. Couso. (2002). Imprecise probabilities induced by multi-valued mappings. In *Proceedings of the 9th IPMU Conference*, pp. 1061–1068, Annecy (France).
16. H. T. Nguyen. (1978). On random sets and belief functions. *Journal of Mathematical Analysis and Applications* **65**, 531–542.
17. M. Puri and D. Ralescu. (1986). Fuzzy random variables. *Journal of Mathematical Analysis and Applications* **114**, 409–422.
18. P. Walley. (1991). *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London.
19. P. Walley. (1997). Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems* **26**, 337–384.