Probabilities of Events Induced by Fuzzy Random Variables

Cédric Baudrit
Université Paul Sabatier
baudrit@irit.fr

Inés Couso
Universidad de Oviedo
couso@pinon.ccu.uniovi.es

Didier Dubois
Université Paul Sabatier
dubois@irit.fr

Abstract

Propagating possibilistic and probabilistic variables through a mapping yields a fuzzy random variable. We propose a method to attach probability intervals to events pertaining to the output variable. We show that this method is consistent with classical approaches to fuzzy random variables and that the obtained probability interval is the mean value of the fuzzy probability defined by viewing a fuzzy random variable as higher order possibilistic uncertainty.

Keywords: Possibility measure, random set, fuzzy random variable.

1 Introduction

More often than not, uncertainty pervading parameters and inputs to a mathematical model is not of a single nature. Namely, randomness, as objective variability, and incomplete information may coexist, especially due to the presence of several, heterogeneous sources of knowledge, as for instance statistical data and expert opinions. In the last thirty years, a number of uncertainty theories have emerged that explicitly recognized incompleteness as a feature distinct from randomness. All such theories are coherent with each other, in the sense that they all represent upper and lower probability bounds, thus proposing a common framework for randomness and incomplete information. In this paper we are especially interested in the joint uncertainty propagation through mathematical models involving quantities respectively modeled by probability and possibility distributions. We will consider three different types of uncertain quantities: random variables observed with total precision, deterministic parameters whose value is imprecisely known, and imprecisely observed random variables.

Observe that the second type of uncertain quantities can be modeled in a natural way by possibility distributions, while the first type can be represented by a classical probability measure. The joint uncertainty propagation of the two first types of uncertain quantities yields a fuzzy random variable for the output. The aim of this paper is to explain how to describe the probability of events pertaining to such a fuzzy random variable. We discuss several approaches: one that generates a probability interval using random sets, a more classical view based on random fuzzy sets and an approach based on the idea of higher order uncertainty, yielding fuzzy probabilities. We show that these various approaches are consistent with one another. Lastly, the case when the three kinds of uncertain quantities are involved is discussed.

2 Preliminaries and notation

A fuzzy set is interpreted as a possibility distribution π from a finite set S to the unit interval associated to some unknown quantity x. Then $\pi(s)$ is interpreted as the possibility that $x = s$.

A random set on S is defined by a mass assignment m which is a probability distribution on the power set of S. We assume that m assigns a positive mass to a family of subsets of S called the set $\mathcal{F}$ of focal subsets. Generally $m(\emptyset) = 0$ and $\sum_{E \subseteq S} m(E) = 1$. A random set induces set functions called plausibility and belief measures, respectively denoted by Pl and Bel, and defined by Shafer [14] as follows.

$$\text{Pl}(A) = \sum_{E \cap A \neq \emptyset} m(E);\ \text{Bel}(A) = \sum_{E \subseteq A} m(E). \quad (1)$$

These functions are dual to each other in the sense that $\text{Pl}(A) = 1 - \text{Bel}(A^c)$, where A^c denotes the complement of A in S. The possibility distribution induced by a mass assignment m is defined as $\pi_m(s) = \sum_{E:s \in E} m(E)$. It is the one-point coverage function of the random set. Generally m cannot be recovered from π_m. However if the set of focal sets $\mathcal{F}$ is nested, then the information conveyed by m and π_m is the same. In this case the plausibility measure is called a possibility measure and is denoted Π, while the belief function is called a necessity measure and is denoted N. It can be checked that

$$\Pi(A) = \max_{s \in A} \pi_m(s);\ N(A) = \min_{s \notin A} 1 - \pi_m(s) \quad (2)$$

Dempster ([8]) introduced belief and plausibility functions as lower and upper probabilities by using a set-valued mapping Γ from a probability space $(\Omega, \mathcal{A}, P)$ to S (yielding a random set), where $\mathcal{A}$ is an algebra of measurable subsets of Ω. For simplicity assume $\forall \omega \in \Omega, \Gamma(\omega) \neq \emptyset$.

A selection from Γ is a function f from Ω to S such that $\forall \omega \in \Omega, f(\omega) \in \Gamma(\omega)$. The set of measurable selections from Γ is denoted $S(\Gamma)$, and we write $f \in S(\Gamma)$ for short. Each selection f yields a probability measure P_f on S such that $P_f(A) = P(f^{-1}(A))$. Now define the following upper and lower probabilities:

$$P^*(A) = \sup_{f \in S(\Gamma)} P_f(A);\ P_*(A) = \inf_{f \in S(\Gamma)} P_f(A).$$

Let the upper and lower inverse images of subsets $A \subseteq S$ be measurable subsets A^* and A_* of Ω defined by $A^* = \{\omega, \Gamma(\omega) \cap A \neq \emptyset\}$, $A_* = \{\omega, \Gamma(\omega) \subseteq A\}$. Define the mass assignment m_Γ on S by $m_\Gamma(E) = P(\{\omega, \Gamma(\omega) = E\})$. Then belief and plausibility functions are retrieved as follows:

$$\text{Pl}_\Gamma(A) = P(A^*); \text{Bel}_\Gamma(A) = P(A_*). \quad (3)$$

A fuzzy random variable [13] is a generalization of the Dempster setting to when the set-valued mapping Γ is changed into a fuzzy set valued mapping Φ. It is supposed that $\forall \omega \in \Omega, \Phi(\omega)$ is a normalized fuzzy set of S. To each fuzzy subset F of S with membership function π_F is attached a probability mass $m_\Phi(F) = P(\{\omega, \Phi(\omega) = F\})$.

3 Probability-possibility propagation

In this section we study how to combine precise random information about a (monodimensional) variable with incomplete, possibilistic information about a (monodimensional) fixed parameter.

3.1 The propagation model

Let us now consider a random variable X, that takes values $x_1, \ldots, x_m$ with respective probabilities $p_1, \ldots, p_m$. Let us assume that we know these values and probabilities. Let us consider, on the other hand, a fixed (constant) parameter, y_0, imprecisely known: let us suppose that our information about it is given by "confidence levels". Thus, we will assume that there is a family of nested sets, $A_1 \supseteq \ldots \supseteq A_q$, containing y_0, with their respective confidence levels, $1 - \alpha_1 \geq \ldots \geq 1 - \alpha_q$. The available information about y_0 takes the form of lower probability bounds:

$$P(A_j) \geq 1 - \alpha_j,\ j = 1, \ldots, q.$$

These inequalities reflect information given by an expert: "the parameter belongs to the set A_j with a confidence degree $1 - \alpha_j$". Notice that we have "pure probabilistic" information about X, which may reflect a phenomenon of variability and "possibilistic" information about y_0 because of the nested structure of confidence sets: Following [9] (finite universes) and [3, 4] (general setting), the set of probability measures $\{P \mid P(A_j) \geq 1 - \alpha_j,\ \forall j = 1, \ldots, q\}$ coincides with the set of probability measures $\mathcal{P}(\Pi)$ dominated by the possibility measure Π with distribution π:

$$\pi(y) = \begin{cases} 1 & \text{if } y \in A_q \\ \alpha_j & \text{if } y \in A_{j-1} \setminus A_j, \forall j < q \end{cases}$$

Therefore, in the first case (variable X) there is randomness and total precision. In the second case, there exists imprecision, but no randomness (y_0 is a constant). Let us now consider the random variable T that, for each possible value of X, x_i, takes value $f(x_i, y_0)$, where $f : I\!R^2 \to I\!R$ is a known mapping. Now we need to represent the available information about the probability measure induced by function f. We easily observe that, when X takes value x_i ($i \in \{1, \ldots, m\}$), T is in the set $T_{ij} = f(x_i, A_j) = \{f(x_i, y) \mid y \in A_j\}$ with a confidence degree $1 - \alpha_j$. Recalling again the results from [4] and [9], we observe that, for each $i \in \{1, \ldots, m\}$, the set of probability measures $\{P \mid P(T_{ij}) \geq 1 - \alpha_j,\ \forall j = 1, \ldots, q\}$ coincides with the set of probability measures dominated by the possibility measure with distribution π^i given by:

$$\pi^i(t) = \begin{cases} 1 & \text{if } t \in T_{iq} \\ \alpha_j & \text{if } t \in T_{i(j-1)} \setminus T_{ij}, \forall j < q \end{cases}$$

induced by the mass assignment m_i: $m_i(T_{i1}) = \nu_1 = \alpha_1$, $m_i(T_{ij}) = \nu_j = \alpha_j - \alpha_{j-1},\ j = 2, \ldots, q$. This possibility distribution is related to π by the extension principle of fuzzy set theory:

$$\pi^i(t) = \sup_{y \mid f(x_i, y) = t} \pi(y),\ \forall t \in I\!R. \qquad (4)$$

Thus, according to the probability distribution of X and our information about y_0, the probability measure of T is imprecisely determined by means of the basic assignment m that assigns probability mass $\nu_{ij} = p_i\, \nu_j$ to each focal T_{ij}. This view comes down to considering a random fuzzy set as a standard random set, using a two-stepped procedure: first select a fuzzy set with membership function π^i with probability p_i and then select the α-cut A_j of π^i with probability ν_j.

Besides, we can observe that the plausibility measure coincides with the arithmetic mean of the possibility measures Π^i (weighted by the probabilities of the different values of X), i.e., $\forall\, A$:

$$\mathrm{Pl}(A) = \sum_{i=1}^{q} \sum_{j \mid A \cap T_{ij} \neq \emptyset} p_i\, \nu_j = \sum_{i=1}^{q} p_i \Pi^i(A). \qquad (5)$$

Similarly the belief function coincides with the arithmetic mean of the necessity measures N^i (similarly weighted).

Taking into account the properties of possibility measures as upper envelopes of sets of probability measures (see [9], for finite universes and [3, 4], for the general case), we get the equality:

$$\mathrm{Pl}(A) = \sup \left\{ \sum_{i=1}^{m} p_i\, P_i(A) \mid P_i \in \mathcal{P}(\Pi^i), \forall i \right\} \qquad (6)$$

and similarly with the degree of belief using inf instead of sup. These equations suggest another probabilistic interpretation of these plausibility and belief functions: let us consider an arbitrary event A. According to our information about y_0, if we observe the value x_i for the random variable X, then the probability $P(T \in A | X = x_i) = P_i(A)$ that T takes a value in A is less than or equal to $\Pi^i(A)$, and at least equal to $N^i(A)$. On the other hand, the probability that X takes each value x_i is p_i. Thus, according to the Theorem of Total Probability, all we know about the probability $P_T(A) = P(T \in A)$ is that it can be expressed as $\sum_{i=1}^{m} p_i\, P_i(A)$, where P_i is a probability measure dominated by Π^i, for each i. Hence, according to equation (6), we can interpret the value $\mathrm{Pl}(A)$ as the tightest upper bound for the "true" probability of A, according to the available information. The same holds for the dual belief function, Bel, as the tightest lower bound.

4 Related approaches

In the literature different notions of fuzzy random variables exist. Older approaches consider a fuzzy random variable "classically", that is as a random variable with a special range. These approaches differ by the involved measurability assumption; see [12] for a more general approach subsuming them. One may also view a fuzzy random variable as expressing higher-order uncertainty. Here we show the relations between the plausibility and belief measures here defined and each of those interpretations.

4.1 Relationship with the "classical" model

A fuzzy random variable is a random variable whose values are fuzzy subsets. This "classical" vision of a fuzzy random variable as a measurable function agrees with the interpretation given by Puri and Ralescu ([13]). In that paper, the authors consider that the outcomes of some random experiments are not numerical ones, but they can be vague linguistic terms. In this context, the information provided by the fuzzy random variable can be summarized by means of the probability measure it induces in the final space. When the fuzzy random variable takes a finite number of different "values", its induced probability is determined by the mass function. Therefore, it is enough to specify the different images of the fuzzy random variable and the probability of occurrence of each one of them. Thus, different probability values will be assigned to different linguistic labels (the probability that the result is "high" is 0.5, etc.). In our particular problem, the fuzzy random variable is constructed specifically as follows.

Let $(\Omega, \mathcal{A}, P)$ be a probability space, and $X : \Omega \rightarrow I\!R$ be a random variable, the observation of a certain characteristic of each element of Ω. Assume that X, takes a finite number of different values, $x_1, \ldots, x_m$ with respective probabilities $p_1, \ldots, p_m$. On the other hand, the available (imprecise) information about a fixed parameter, y_0 is given by the possibility distribution $\pi : I\!R \rightarrow [0,1]$, viewed as a constant fuzzy mapping, $\tilde{Y}$ from Ω to the set of measurable fuzzy subsets of $I\!R$, that assigns to every element of Ω, the same fuzzy set, π. This scheme indicates that the parameter y_0 does not depend on each particular individual, $\omega \in \Omega$. Let us now consider a map $f : I\!R^2 \rightarrow I\!R$ and the random variable given as $T = f \circ (X, y_0)$. T takes values $t_1 = f(x_1, y_0), \ldots, t_m = f(x_m, y_0)$ with respective probabilities $p_1, \ldots, p_m$. The available information about each t_i is given by the fuzzy set π_i defined as $\pi_i(t) = \sup_{\{y \in I\!R | f(x_i, y) = t\}} \pi(y), \ \forall\, t \in I\!R$. These m fuzzy sets and their respective probabilities uniquely determine the probability distribution induced by $\tilde{T}$, considered as a classical measurable function. The plausibility function defined in the previous section can be expressed as: $\mathrm{Pl}(A) = \sum_{i=1}^{m} p_i \sup_{t \in A} \pi_i(t), \ \forall\, A$, in agreement with (5).

4.2 Relationship with the imprecise higher order uncertainty setting

In [5], a fuzzy r.v. $\tilde{T} : \Omega \rightarrow \tilde{\mathcal{P}}(I\!R)$ represents imprecise information about the random variable $T_0 : \Omega \rightarrow I\!R$: for each $\alpha > 0$, the probability of the event "$T_0(\omega) \in [\tilde{T}(\omega)]_\alpha, \ \forall\, \omega \in \Omega$" is greater than or equal to $1 - \alpha$. Under this interpretation we can say that, for each confidence level $1 - \alpha$, the probability distribution associated to T_0 belongs to the set $\mathcal{P}_{\tilde{T}_\alpha} = \{P_T \mid T \in S(\tilde{T}_\alpha)\}$, where $S(\tilde{T}_\alpha)$ is the set of selections from the random set $\tilde{T}_\alpha$.

Thus, given an arbitrary event A on the final space, the probability $P_{T_0}(A)$ belongs to the set

$$\mathcal{P}_{\tilde{T}_\alpha}(A) = \{P_T(A) \mid T \in S(\tilde{T}_\alpha)\}$$

with confidence level $1 - \alpha$. This family of nested confidence sets determines the fuzzy number $\tilde{P}_{\tilde{T}}(A)$, defined as

$$\tilde{P}_{\tilde{T}}(A)(p) = \sup\{\alpha \in [0,1] \mid p \in \mathcal{P}_{\tilde{T}_\alpha}(A)\}, \ \forall\, p,$$

that represents our imprecise information about the quantity $P_{T_0}(A) = P(T_0 \in A)$. Thus, the value $\tilde{P}_{\tilde{T}}(A)(p)$ represents the degree of possibility that the "true" degree of probability $P_{T_0}(A)$ is p. The possibility measure $\tilde{P}_{\tilde{T}}$ is a "second order possibility measure". We use this term because it is a possibility distribution defined over a set of probability measures ([2, 15]).

There exists a strong relationship between the plausibility measure defined in section 3 and the fuzzy set $\tilde{P}_{\tilde{T}}(A)$ defined in [5]. Define the "mean value" of a fuzzy number π, as the interval:

$$M(\pi) = \{E(P) \mid P \leq \Pi\},$$

where $E(P)$ represents the expected value associated to the probability measure P [10]. Then, we can state the following new result:

Theorem. *Given an arbitrary event A, the interval* $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$ *coincides with the "mean value" of the fuzzy set* $\tilde{P}_{\tilde{T}}(A)$.

The intuitive meaning of this last result is as follows. As explained before, in the second order imprecise model we represent our imprecise information by a pair of order 2 plausibility-necessity measures. For each α, we assign the lower probability (degree of necessity) $1-\alpha$ to a set of probability measures induced by a random set obtained via α-cuts. This necessity function is equivalent to a set of second order probability measures. Pick a particular second order probability measure, $I\!P$, belonging to this set, and an arbitrary event A. In this setting, we can define a random variable that takes each (probability) value $Q(A)$ with (higher-order) probability $I\!P(\{Q\})$[1]. If $I\!P$ were the "correct" second order probabil-

[1]For the sake of clarity, we are assuming that the second order probability measure, $I\!P$, is "discrete".

ity measure that models the second order experiment, then we could state that the "true" probability of A should coincide with the expectation of this random variable. In the last theorem we have shown that $\mathrm{Bel}(A)$ and $\mathrm{Pl}(A)$ respectively coincide with the lower and upper bounds of the set comprising the possible values of the expectations associated to each second order probability measure dominated by the pair of (second order) possibility-necessity measures. As a consequence of this, $\mathrm{Bel}(A)$ and $\mathrm{Pl}(A)$ represent, in the average, the most precise bounds for the "true" probability of A, under the available information. This result is related to one by Couso et al. [7], who prove that the mean interval of the (fuzzy) expectation of a fuzzy random variable is the expectation of the random set obtained by computing the mean intervals of the fuzzy realizations of this random variable.

5 Propagating general heterogeneous information

In this section, we try to combine and propagate the three kinds of information: pure random variables, imprecisely known fixed quantities, and imprecise random variables. $\vec{X} : \Omega \to I\!R^k$ is a random vector that is observed with total precision; $\vec{Y} = (y_{01}, \ldots, y_{0l})$, is a constant vector and we have partial information about it, represented by a fuzzy set, π, a constant map $\tilde{Y}$ from Ω to the set of measurable fuzzy subsets of $I\!R$, that assigns the fuzzy set π to each element $\omega \in \Omega$. Finally, $\vec{Z} : \Omega \to I\!R^n$ is a random vector observed with imprecision. In order to represent this vague observation, we consider a multi-valued mapping, $\Gamma : \Omega \to \mathcal{P}(I\!R^n)$ inducing a random set. In our model we suppose that there exists a unidimensional random variable, $T : \Omega \to I\!R$, that can be expressed in the form $T = f(\vec{X}, \vec{Y}, \vec{Z})$, where the mathematical model described by the function $f : I\!R^{k+l+n} \to I\!R$ is totally well-known. We will try to represent the information about the probability distribution of T based on the information available, about $\vec{X}$, $\vec{Y}$ and $\vec{Z}$, respectively.

First observe that $\vec{X}$ is a random vector and, therefore, is a particular case of multidimensional random set (a singleton in $I\!R^k$). Thus, in our model, we can assume it as part of vector $\vec{Z}$.

To simplify again the notation, temporarily suppose that the variable Z and the parameter y_0 are unidimensional. Let m and m' be the mass assignments of probability associated to y_0 and Z, respectively. We shall argue that the suitable combination of such assignments in this case is the "product rule".

Given a particular element of Ω, $\omega \in \Omega$, all that we know about the value $Z(\omega)$ it is that it belongs to the set $\Gamma(\omega)$. The imprecise knowledge about y_0 is modelled by a constant fuzzy function, with value π. Thus, with a confidence level $1 - \alpha$, the parameter y_0 belongs to α–cut $\pi_\alpha = \{x \in I\!R \mid \pi(x) \geq \alpha\}$. If we combine both sources of information, given an arbitrary element ω of the initial space and any $\alpha \in [0, 1]$, we can say that the pair $(Z(\omega), y_0)$ belongs to the set $\Gamma(\omega) \times \pi_\alpha$ with confidence $1 - \alpha$. Since we are working on a finite referential, we will suppose that the multi-valued mapping Γ has r different set-valued images, $C_1, \ldots, C_r$ with respective masses $m_1, \ldots, m_r$. On the other hand, the fuzzy set π has q different α–cuts, $\pi_{\alpha_1} \supseteq \ldots \supseteq \pi_{\alpha_q}$.

For each $i \in \{1, \ldots, r\}$, the family of probability measures $\{P \mid P(C_i \times \pi_{\alpha_j}) \geq \alpha_j,\ \forall j = 1, \ldots, q\}$ coincides with the family of probability measures dominated by the possibility measure induced by the consonant mass assignment m_i: $m_i(C_i \times \pi_{\alpha_1}) = \nu_1 = \alpha_1$, $m_i(C_i \times \pi_{\alpha_j}) = \nu_j = \alpha_j - \alpha_{j-1},\ j = 2, \ldots, q$. Thus, according to the information about Z and y_0, the probability measure of (Z, y_0) is imprecisely determined by means of the basic assignment m that assigns the probability mass $\nu_{ij} = m_i \nu_j$ to each focal $C_i \times \pi_{\alpha_j}$. In other words, it is obtained as the "product" of the mass assignment associated to the random set Γ and the mass assignment associated to the possibility measure Π. It is clear that this approach is just a variant of, slightly more general than, the model proposed in Section 3.

6 Conclusion and open problems

The main result of this short note is that the joint propagation of possibility and probability through a mathematical model yields a fuzzy random variable consistent with classical views thereof, as well as more recent second order uncertainty interpretations. The higher order model in section 4.2 is richer and does not aggregate imprecision with randomness, since it assigns a fuzzy-interval-valued probability to each event. The model in section 3 yields a belief function (or a random set) which averages the fuzzy random variable and delivers for each event the (interval) mean value of its fuzzy-interval-valued probability. The proposed approach assumes independence between the statistical information about the random quantity and the imprecise information about the deterministic attribute, as witnessed by equation (5). Independence is between observation processes, not the observed quantities. We restricted to a finite seting for the sake of simplicity and because it is the one to be used in practical implementations. Extensions to the infinite setting are of course of interest. Future works should address in more details the multidimensional case, which require notions of independence in the presence of variability and imprecision to be further formalized, following the path opened in [6, 11]. More details can be found in the long version of this paper.

Acknowledgments

This work is partially supported by French Institutes B.R.G.M, I.R.S.N and I.N.E.R.I.S, a travel grant from Université Paul Sabatier, and the Spanish project MTM2004-01269.

References

[1] C. Baudrit, D. Dubois, D. Guyonnet, H. Fargier. Joint treatment of imprecision and randomness in uncertainty propagation. IPMU-2004, July 4-9, Perugia, Italy, 873-880, 2004.

[2] G. De Cooman, A behavioural model for vague probability assessments, to appear in *Fuzzy Sets and Systems.*

[3] G. De Cooman, D. Aeyels. Supremum-preserving upper probabilities. *Information Sciences*, 118, 173 -212, 1999.

[4] I. Couso, S. Montes, and P. Gil. The necessity of the strong α-cuts of a fuzzy set, *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9, 249-262, 2001.

[5] I. Couso, S. Montes, and P. Gil. Second order possibility measure induced by a fuzzy random variable. In C. Bertoluzza et al. (Eds.), *Statistical Modeling, Analysis and Management of Fuzzy Data*, 127–144. Springer, Heidelberg, 2002.

[6] I. Couso, S. Moral, P. Walley. A survey of concepts of independence for imprecise probabilities, *Risk Decision and Policy*, 5, 165–181, 2000.

[7] I. Couso, E. Miranda, G. De Cooman A possbilistic interpretation of the expectation of a fuzzy random variable. In M. Lopez-Diaz, et al. Eds. *Soft Methodology and Random Information Systems*, 133-140, Springer, Berlin, 2004.

[8] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.*, 38, 325-339, 1967.

[9] D. Dubois, H. Prade. When upper probabilities are possibility measures, *Fuzzy Sets and Systems*, 49, 1992, 65-74.

[10] D. Dubois and H. Prade. The mean value of a fuzzy number, *Fuzzy Sets and Systems*, 24, 279–300, 1987.

[11] T. Fetz. Sets of joint probability measures generated by weighted marginal focal sets, Proceedings of ISIPTA'01, 171–178, Ithaca, NY (USA), 2001.

[12] V. Krätschmer (2001) A unified approach to fuzzy random variables, *Fuzzy Sets and Systems* **123** 1-9.

[13] M.L. Puri, D. Ralescu. Fuzzy Random Variables, *J. Math. Anal. Appl.,* 114, 409-422, 1986.

[14] G. Shafer. A mathematical theory of evidence. Princeton University Press, 1976.

[15] P. Walley. Statistical inference based on a second-order possibility distribution, *Int. J. of General Systems*, 26, 337-383, 1997.