Available online at www.sciencedirect.com
ScienceDirect
International Journal of Approximate Reasoning
45 (2007) 82–105
INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

Joint propagation of probability and possibility in risk analysis: Towards a formal framework

Cédric Baudrit [a], Inés Couso [b,*], Didier Dubois [a]

[a] *Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, France*
[b] *Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, Spain*

Received 15 April 2005; received in revised form 29 June 2006; accepted 7 July 2006
Available online 11 August 2006

Abstract

This paper discusses some models of Imprecise Probability Theory obtained by propagating uncertainty in risk analysis when some input parameters are stochastic and perfectly observable, while others are either random or deterministic, but the information about them is partial and is represented by possibility distributions. Our knowledge about the probability of events pertaining to the output of some function of interest from the risk analysis model can be either represented by a fuzzy probability or by a probability interval. It is shown that this interval is the average cut of the fuzzy probability of the event, thus legitimating the propagation method. Besides, several independence assumptions underlying the joint probability–possibility propagation methods are discussed and illustrated by a motivating example.

Keywords: Random sets; Possibility distributions; Fuzzy random variables; Independence

1. Introduction

It is now more and more widely acknowledged that all facets of uncertainty cannot be captured by a single probability distribution. In risk analysis, the basic task consists of

* Corresponding author. Address: Universidad de Oviedo, Dep. de Estadística e I.O. y D.M. E.U. Ingeniería Técnica Industrial de Gijón, Módulo 1-Planta 4 33071 Gijón (Asturias), Spain. Tel.: +34 985191806; fax: +34 985191802.
E-mail addresses: baudrit@irit.fr (C. Baudrit), couso@uniovi.es (I. Couso), dubois@irit.fr (D. Dubois).

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2006.07.001

exploiting the mathematical model of some phenomenon, so as to predict its output when some inputs or parameters of the model are ill-known.

There are two basic reasons why such parameters or inputs cannot be assigned precise values. First, some quantities are subject to intrinsic variability. For instance, when predicting the effect of radioactivity pollution on the health of people, it is clear that this effect depends on the particulars of individuals (their weight, for instance), and such characteristics differ from one individual to another. Another reason for uncertainty is the plain lack of knowledge about relevant parameters. This lack of knowledge may stem from a partial lack of data, either because this data is impossible to collect, or too expensive to collect, or because the measurement devices have limited precision, or yet because only human experts can provide some imprecise information.

Under such a situation, the traditional attitude was to represent each and every ill-known parameter or input by means of a probability distribution, for instance resorting to Laplace principle of Insufficient Reason that prescribes the use of uniform distributions in the absence of information. A more refined approach is the appeal to subjective probability distributions, whereby additive degrees of belief are supplied by experts via an exchangeable betting procedure. However, several scholars have complained that predictions obtained under such assumptions were generally not conservative and can only be attached a limited trust. Indeed, this purely probabilistic view has the major defect of not taking the idea of lack of knowledge for granted, and confuses it with the variability phenomenon. It seems that part of the controversies between subjective and objective probabilities are due to this confusion. Indeed, a variable quantity, if precisely observable, can faithfully be modeled by a single probability distribution built from the observation of frequencies. However a fixed but ill-known quantity is more naturally modeled in terms of partial lack of knowledge, for instance by means of confidence intervals, where confidence levels are subjectively assessed.

More often than not, the 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. These theories have proposed non-additive set-functions which most of the time combine set-valued and probabilistic representations. The most general setting is the one of imprecise probabilities developed at length by Peter Walley [35]. In this theory, sets of probability functions capture the notion of partial lack of probabilistic information. Slightly more restrictive is the theory of evidence, initiated by Dempster [9], an approach relying on the notion of random set, each set-valued realization representing a plainly incomplete information item. The set-functions generated in this mathematical framework were further exploited by Shafer [32] and Smets [33], within a purely subjectivist, non-statistical approach to uncertain evidence. Even more restrictive is the framework of possibility theory, where pieces of information take the form of fuzzy sets of possible values [39], which can be interpreted as consonant (nested) random sets. The merit of this framework lies it its great simplicity, which enables incomplete probabilistic information on the real line to be encoded in the form of fuzzy intervals [15,13]. Possibility distributions can also straightforwardly accommodate linguistic information on quantitative scales. 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. Different types of uncertain quantities can be considered:

1. Random variables observed with total precision. This is the standard case in probabilistic modelling, where only variability is present.
2. Deterministic parameters whose value is imprecisely known. Our information about it can be modeled by
   - 2a. A random variable: when there is a random error in the measurement of this deterministic value.
   - 2b. A set: when the information about the parameter is the fact that it lies in a given range.
   - 2c. A fuzzy set, interpreted as a possibility distribution, when the information about the parameter is linguistic (like "the temperature is high").
   - 2d. A nested family of confidence intervals supplied by an expert along with the corresponding levels of confidence. This case like the previous one can be modeled by a possibility distribution.
   - 2e. A random set: when the random error in the measurement goes along with a systematic error taking the form of a perception interval.
3. Imprecisely observed random variables: in this case variability and incompleteness come together because each observation in a statistic is set-valued by lack of precision, due to a limitation of the observation device.

In cases 2b–d, the natural tool for representing uncertainty is a possibility distribution – a binary one in case 2b. The fact that cases 1 and 2a can be modeled by a probability distribution should not hide the fact that in case 2a randomness does not pertain to the observable, but to the measurement device. So, this case could be captured more naturally by confidence intervals, enclosing the ill-known fixed value, and derived from the statistical analysis of the observations. Case 2e extends this situation to when the measurement device produces set-valued (imprecise) observations of a fixed ill-known quantity. It is formally similar to case 3, but in the latter, the observed quantity is subject to intrinsic variability.

If a mathematical model involves both random variables and imprecisely known parameters, the predictions are likely to take the form of a fuzzy random variable, which is a random variable taking fuzzy sets as values. Fuzzy random variables have been introduced in slightly different settings (see [20] for an overview). The original motivation stemmed from putting together random variables and linguistic variables introduced by Zadeh [38]. However, a special case of a fuzzy random variable is a random set, and not all fuzzy sets come from linguistic data, since for instance possibility distributions are fuzzy sets that can encode nested families of confidence intervals. Fuzzy random variables can be generated by computing a function whose arguments involve random variables and possibilistic variables without referring to linguistic information [21].

The aim of this paper is to describe basic principles underlying the combination of these three sources of information for the purpose of uncertainty propagation. An important aspect of the discussion concerns the way to represent knowledge about heterogeneous variables. We provide three different approaches to jointly propagate probability distributions and possibility distributions, and compute upper and lower probabilities of output events. Each one of these propagation models reflects a particular situation. We analyze the relations among these models, and relate them with other models in the literature. Section 2 provides the necessary technical background. Section 3 presents an example that will illustrate how the same possibility distribution can be interpreted differently in different situations. Its combination with the same probability measure leads us to different propagation models. In particular, the case when a random quantity is described by a possibility measure is distinguished from the case when a deterministic (constant) value is ill-known. In Section 4, we consider the joint propagation of subjective possibilistic information and well-observed random variables. We show that it generates a fuzzy random variable in Section 5, where we prove that the imprecise probability intervals of events obtained in the previous section are average intervals of the fuzzy probabilities generated by a higher order approach. In Section 6, we consider the joint propagation of random and deterministic but ill-known quantities. Section 7 provides a more general propagation framework. We end the paper with some concluding remarks and open problems.

2. Preliminaries and notation

In this section, some definitions needed in the rest of the paper are recalled. A fuzzy set F is identified with a membership function from a finite set S to the unit interval. The value $F(s)$ is the membership grade of element s in the fuzzy set. A fuzzy set is not empty if its membership function is normalized, that is, $F(s) = 1$ for some element s. In this paper, a fuzzy set is interpreted as a possibility distribution π associated to some unknown quantity x. Then $\pi(s)$ is interpreted as the possibility that $x = s$. Throughout the paper, we will use the notation π to denote a possibility distribution and – the membership function of – its associated fuzzy set. 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 only to a finite 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 [32] as follows:

$$\mathrm{Pl}(A) = \sum_{E\cap A\neq\emptyset} m(E); \tag{1}$$

$$\mathrm{Bel}(A) = \sum_{E\subseteq A} m(E). \tag{2}$$

These functions are dual to each other in the sense that $\mathrm{Pl}(A) = 1 - \mathrm{Bel}(A^c)$, 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

$$\Pi(A\cup B) = \max(\Pi(A), \Pi(B)); \quad N(A\cap B) = \min(N(A), N(B)), \tag{3}$$

$$\Pi(A) = \max_{s\in A} \pi_m(s); \quad N(A) = \min_{s\notin A}(1 - \pi_m(s)). \tag{4}$$

Suppose that $\pi(s_1) = 1 \geqslant \pi(s_2) \geqslant \cdots \geqslant \pi(s_n) \geqslant \pi(s_{n+1}) = 0$, and $E_i = \{s_1, \ldots, s_i\}$, then

$$m(E_i) = \pi(s_i) - \pi(s_{i+1}), \quad \forall i = 1, \ldots, n. \tag{5}$$

These set-functions can be interpreted as families of probability measures, even if this view does not match the original motivations of Shafer [32] and Smets [33] for belief functions nor the ones of Zadeh [39] for possibility theory. Let $\mathscr{P}$ be a set of probability measures on S. They induce upper and lower probability functions respectively defined by

$$P^*(A) = \sup_{Q \in \mathscr{P}} Q(A); \quad P_*(A) = \inf_{Q \in \mathscr{P}} Q(A). \tag{6}$$

The set of probability measures dominated by an upper probability P^* is denoted by $\mathscr{P}(P^*) = \{Q, P^*(A) \geqslant Q(A), \forall A \subseteq S\}$. If the upper probability measure P^* is generated by the family $\mathscr{P}$, then $\mathscr{P}(P^*)$ is generally a proper superset of $\mathscr{P}$. In the case of a plausibility function Pl, the set $\mathscr{P}(\mathrm{Pl})$ of probability functions dominated by Pl is not empty and it generates lower and upper probability functions that coincide with the belief and plausibility functions, i.e.

$$\mathrm{Pl}(A) = \sup_{P \in \mathscr{P}(\mathrm{Pl})} P(A); \quad \mathrm{Bel}(A) = \inf_{P \in \mathscr{P}(\mathrm{Pl})} P(A). \tag{7}$$

This view of belief and plausibility functions as lower and upper probabilities was actually originally put forward by Dempster [9] using a set-valued mapping Γ from a probability space $(\Omega, \mathscr{A}, P)$ to S (yielding a random set), where $\mathscr{A}$ 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:

$$P^*(A) = \sup_{f \in S(\Gamma)} P_f(A); \quad P_*(A) = \inf_{f \in S(\Gamma)} P_f(A). \tag{8}$$

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:

$$P^*(A) = P(A^*) = \mathrm{Pl}_\Gamma(A) = \sum_{E \cap A \neq \emptyset} m_\Gamma(E); \tag{9}$$

$$P_*(A) = P(A_*) = \mathrm{Bel}_\Gamma(A) = \sum_{E \subseteq A} m_\Gamma(E). \tag{10}$$

So, the approach by mass assignments, the one using selection functions, and the one using multiple valued mappings are equivalent, as to the probability bounds they induce on events.

A fuzzy random variable is a generalization of the Dempster setting [9] 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. Joint probability–possibility propagation: A motivating example

As indicated in the introduction of this work, our final goal is to propose a framework for the propagation of three types of information: precise information about a random variable, incomplete or linguistic information about a fixed parameter and incomplete information about a random variable. A possibility distribution can model imprecise information about a fixed unknown parameter and it can also serve as an approximate representation of incomplete observation of a random variable. Although the same possibility distribution can describe two different types of information, the way in which it must be combined with a probability distribution representing a random variable will be different in each case. Let us provide a simple example that illustrates these different situations.

A game consists of two steps. In the first step, we can choose one of two keys. One of them opens a safe containing a reward worth 1000 euros. The other key does not open anything. The second step is partially unknown. It leads either to a win of at least 700 euros, else a possibly smaller reward of at least 50 euros. A more exact evaluation cannot be given.

A random variable, X, represents the reward obtained in the first step. It takes the values 0 or 1000, each of them with probability 0.5. More precisely, the results of the key experiment are $\omega_1 =$ "the chosen key opens the safe" and $\omega_2 =$ "the chosen key does not open the safe", and the variable is $X : \{\omega_1, \omega_2\} \to \mathbb{R}$ defined as $X(\omega_1) = 1000$ and $X(\omega_2) = 0$.

Let the quantity $Y : \Omega \to \mathbb{R}$ denote the reward in the second step. We can model our knowledge about $Y(\omega_1)$ and $Y(\omega_2)$ by means of the necessity measure:

$$N((50,\infty)) = 1, \quad N((700,\infty)) = 0.5.$$

It contains the same information as the possibility distribution:

$$\pi(x) = \begin{cases} 1 & \text{if } x > 700, \\ 0.5 & \text{if } 50 < x \leqslant 700, \\ 0 & \text{if } x \leqslant 50. \end{cases}$$

Alternatively, it can be described by the random set with mass assignment $m((50,\infty)) = 0.5$, $m((700,\infty)) = 0.5$.

Three possible scenarios articulate the second step with respect to the first one.

(a) Whether the additional reward depends on the chosen key or not is unknown.
(b) The additional reward is an (ill-known) constant y_0, independent from the chosen key.
(c) If the chosen key opens the safe, then an additional reward (in addition to the 1000 initial euros) worth more than 700 euros is received. If the chosen key does not open the safe, there will be a consolation prize worth more than 50 euros.

In this paper, there is a given underlying probability space $(\Omega, \mathscr{F}, P)$. The random variable $X : \Omega \to \mathbb{R}$ is known. There is also some information about the second reward Y. In each of the three cases, the incomplete information about Y is determined by the same possibility distribution, π, but a different type of information is given in each case. Yet, we do not know how $Y : \Omega \to \mathbb{R}$ is defined. For each element ω in the initial space, we cannot determine each image, $Y(\omega)$. Thus, we can neither determine the probability distribution $P_{(X,Y)}$, nor the probability measure induced by T, P_T. But we shall represent the available information about them by means of upper probability measures.

Let us try to evaluate the resulting knowledge on the total gain $T = X + Y$ for each scenario. We shall see that the available information about the probability distribution of T is different in each situation.

(a) Here the link between X and Y is unknown. In particular, we do not know whether the second reward Y is a constant (independence) or not (total dependence). However, we can model our *knowledge* about the value of this reward by means of a (constant) fuzzy random variable $\tilde{Y} : \{\omega_1, \omega_2\} \to \tilde{\mathscr{P}}(\mathbb{R})$ that associates to both results about the key the same fuzzy set $\tilde{Y}_0$ with distribution π (the possibility distribution above), as we have got the same information about $Y(\omega_1)$ and $Y(\omega_2)$. Here, we must combine the precise information about a random variable (the value of the first reward) with an independent incomplete information about another one (the value of the second reward). If the chosen key opens the safe, the total reward is, surely, greater than 1050 euros. In addition, with probability at least 0.5, it is greater than 1700 euros. If the key does not open the safe, the total reward is, surely, greater than 50 euros. In addition it is greater than 700 euros with probability at least 0.5. In other words, the possibility that $T(\omega_1) = t$ is

$$\pi^1(t) = \begin{cases} 0.5 & \text{if } t \in (1050, 1700], \\ 1 & \text{if } t \in (1700, \infty), \\ 0 & \text{otherwise.} \end{cases}$$

And the possibility that $T(\omega_2)$ coincides with t is

$$\pi^2(t) = \begin{cases} 0.5 & \text{if } t \in (50, 700], \\ 1 & \text{if } t \in (700, \infty), \\ 0 & \text{otherwise.} \end{cases}$$

For an arbitrary event A, the probability $P_T(A) = P(T \in A)$ is

$$P(T \in A) = \sum_{i=1}^{2} P(T \in A | X = x_i) \cdot P(X = x_i),$$

where $P(T \in A | X = x_i)$ is dominated by $\Pi^i(A)$ (the possibility measure associated to π^i). Indeed, since the process driving the choice of the additional reward is unknown, all that is known about T is of the form $X + \tilde{Y}_0 = x_i + \tilde{Y}_0$, a fuzzy set with distribution π^i. Hence, $P(T \in A)$ is dominated by $P_a^*(A) = \sum_{i=1}^{2} P(X = x_i)\Pi^i(A)$. In the example, P_a^* is the possibility measure associated to the mass assignment

$$m_a((1700, \infty)) = m_a((1050, \infty)) = m_a((700, \infty)) = m_a((50, \infty)) = 1/4.$$

It can be obtained via Dempster's rule applied to the probability P_X attached to X and the possibility π describing Y, followed by a projection on $T = X + Y$. This fact points out that even if the relation between X and Y is unknown, the pieces of information about them are independent. P_a^* corresponds to the possibility distribution

$$\pi_a(t) = \begin{cases} 0 & \text{if } t \leqslant 50, \\ 0.25 & \text{if } t \in (50, 700], \\ 0.5 & \text{if } t \in (700, 1050], \\ 0.75 & \text{if } t \in (1050, 1700], \\ 1 & \text{if } t > 1700. \end{cases}$$

(b) In the second scenario, we have additional information about Y: we know that it is a fixed number, i.e. $Y(\omega_1) = Y(\omega_2) = y_0$. So, we have more information about T than in the previous case: we know that the reward in the second step is fixed (it does not depend on the chosen key.) Hence we have the following information about the conditional probability values $P(T = t | Y = y)$:

$$P(T = t|Y = y) = P(X = t - y) = \begin{cases} 0.5 & \text{if } t = 1000 + y, \\ 0.5 & \text{if } t = y. \end{cases}$$

On the other hand, we know that $P(T \in A) = \int_{y \geqslant 50} P(T \in A|y)\,\mathrm{d}Q_Y(y)$ where $Q_Y \leqslant \Pi$, that is

$$P(T \in A) \leqslant P^*_b(A) = \sup_{Q_Y \leqslant \Pi} \left\{ \int_{y \geqslant 50} P(T \in A|Y = y)\,\mathrm{d}Q_Y \right\}.$$

This supremum coincides with the Choquet integral of $g_A(y) = P(T \in A | Y = y)$ with respect to Π (where Π represents the possibility measure associated to the possibility distribution π.) Hence, using the mass function induced by π:

$$P^*_b(\{t\}) = 0.5 \sup_{y \geqslant 50} P_X(t - y) + 0.5 \sup_{y \geqslant 700} P_X(t - y) = \begin{cases} 0.25 & \text{if } t \in [50, 700), \\ 0.5 & \text{if } t \geqslant 700, \\ 0 & \text{otherwise.} \end{cases}$$

The set-function P^*_b that is obtained is not necessarily a possibility measure (since $\sup_t P^*_b(T = t) < 1$), nor a belief function. In Section 5, we will show that P^*_b is dominated by P^*_a (it represents more precise information), even if, as shown in the table below, $\pi_a(t) = P^*_a([0, t]) = P^*_b([0, t])$.

t	$[0, 50)$	$[50, 700)$	$[700, 1050)$	$[1050, 1700)$	$\geqslant 1700$
$\sup_{y \geqslant 50} P_X([0, t] - y)$	0	0.5	0.5	1	1
$\sup_{y \geqslant 700} P_X([0, t] - y)$	0	0	0.5	0.5	1
$P^*_b([0, t])$	0	0.25	0.50	0.75	1

(c) In this situation the same random process is at work in the result of the two steps. Now, the information about $(X, Y) : \Omega \to \mathbb{R}^2$ should be represented by the multi-valued map $\Gamma_{X,Y} : \{\omega_1, \omega_2\} \to \mathscr{P}(\mathbb{R})^2$ such that $\Gamma_{X,Y}(\omega_1) = \{1000\} \times (700, \infty)$ and $\Gamma_{X,Y}(\omega_2) = \{0\} \times (50, \infty)$. Each subset so obtained has probability 0.5. All we know about (X, Y) is that it is a selection of $\Gamma_{X,Y}$. We easily observe that the marginal upper probability of $\Gamma_{X,Y}$ on the Y axis is the possibility measure associated to the possibility distribution π previously considered. (All we know about the probability distribution of the second reward, P_Y, is that it is dominated by this possibility measure.) In this third scenario, we can thus represent the available information about $T = X + Y$ by means of a random set $\Gamma_T : \Omega \to \Omega(\mathbb{R})$ given by

$$\Gamma_T(\omega_1) = 1000 + (700, \infty) = (1700, \infty) \quad \text{and} \quad \Gamma_T(\omega_2) = (50, \infty),$$

each with probability 0.5. (If the chosen key opens the safe, the total reward is surely greater than 1700 euros. If it does not open it, we only know that the total reward is greater than 50 euros.) In this case, our information about the probability distribution of T can be represented by the basic assignment:

$$m_c((1700,\infty)) = 1/2 = m_c((50,\infty)).$$

The upper probability associated to m_c, P^*_c, is a possibility measure that neither dominates nor is dominated by the first upper probability measure, P^*_a.

As we have stated at the beginning of the paper, our aim is to show how we should combine a probability measure and a possibility measure in different settings. In the following sections, we will describe in detail the three different propagation models outlined in this example. We will show the relationships and the differences among them. To conclude this section, let us observe that, in the first case (case (a)), we can represent, in a natural way, the available information about the random variable T by a fuzzy random variable. It takes several different "values" (images), each one of them related to a value of X. (When X takes value x_i, the image of the fuzzy random variable is the fuzzy set π^i). We can find in the literature different models to represent the probabilistic information provided by a fuzzy random variable. In next section we will show the relations between the plausibility measure here defined, P^*_a, and each of those models.

4. When the joint propagation of probability and possibility yields a fuzzy random variable

This section systematizes the situation of scenario (a) in the motivating example. Let us now consider a random variable $X : \Omega \to \mathbb{R}$, that takes the 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, another variable, $Y : \Omega \to \mathbb{R}$ imprecisely known. Let us suppose we have the same information about all its images $Y(\omega)$, and that it is given by means of "confidence sets", which are cuts of a fuzzy set $\tilde{Y}_0$. Namely, we will assume that there is a family of nested sets, $A_1 \supseteq \cdots \supseteq A_q$, with their respective confidence levels, $1-\alpha_1 \geqslant \cdots \geqslant 1-\alpha_q$. The available information about $Y(\omega)$ takes the form of lower probability bounds:

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

These inequalities reflect information given by an expert: "the value $Y(\omega)$ belongs to the set A_j with a confidence degree $1-\alpha_j$". (For instance, if $q=1$ and $\alpha_1=0$ we should reflect that the expert only knows the range of Y, not anything else.) Notice that we have "pure probabilistic" information about X, which may reflect a phenomenon of variability and "possibilistic" information about Y because of the nested structure of confidence sets. However, even if the *knowledge* $\tilde{Y}_0$ about Y does not depend on a random phenomenon, the actual value of Y may fluctuate according to the value of X. There is no information about this possible (objective) dependence between X and Y, but the source of information about X (standard statistical data) is independent from the source of information about Y (a human expert).

Following [15] (finite universes) and [3,4] (general setting), the set of probability measures $\{P : P(A_j) \geqslant 1-\alpha_j, \forall j=1,\ldots,q\}$ coincides with the set of probability measures that are dominated by the following possibility measure, Π:

$$\Pi(A) = \begin{cases} \alpha_1 & \text{if } A \cap A_1 = \emptyset, \\ \alpha_{j+1} & \text{if } A \cap A_j \neq \emptyset, \quad A \cap A_{j+1} = \emptyset, \;\; j=1,\ldots,q-1, \\ 1 & \text{if } A \cap A_q \neq \emptyset. \end{cases} \tag{11}$$

This possibility measure is determined by the basic assignment m:

$$m(A_j) = \nu_j = \alpha_{j+1} - \alpha_j, \quad j = 0, \ldots, q,$$

where $A_0 = \mathbb{R} \supseteq A_1 \supseteq \cdots \supseteq A_{q-1} \supseteq A_q$ and $0 = \alpha_0 \leqslant \alpha_1 \leqslant \cdots \leqslant \alpha_q \leqslant \alpha_{q+1} = 1$.

Therefore, for variable X, there is randomness but total precision, while the information regarding Y is incomplete: especially Y is possibly tainted with variability, but its relationship to X is just unknown.

Let us now consider the random variable $T = f(X, Y)$, where $f : \mathbb{R}^2 \to \mathbb{R}$ is a known mapping. Now we need to represent the available information about the probability measure induced by $T : \Omega \to \mathbb{R}$. We easily observe that, when X takes the value x_i ($i \in \{1, \ldots, m\}$), T is in the set $T_{ij} = f(x_i, A_j) = \{f(x_i, y) : y \in A_j\}$ with a confidence degree $1 - \alpha_j$. Recalling again the results from [4,15], we observe that, for each $i \in \{1, \ldots, m\}$, the set of probability measures $\{P : P(T_{ij}) \geqslant 1 - \alpha_j, \forall j = 1, \ldots, q\}$ coincides with the set of probability measures dominated by the possibility measure Π^i given by

$$\Pi^i(A) = \begin{cases} \alpha_1 & \text{if } A \cap T_{i1} = \emptyset, \\ \alpha_{j+1} & \text{if } A \cap T_{ij} \neq \emptyset, \quad A \cap T_{i(j+1)} = \emptyset, \;\; j = 1, \ldots, q-1, \\ 1 & \text{if } A \cap T_{iq} \neq \emptyset. \end{cases}$$

This possibility measure is related to Π as follows. Its possibility distribution, π^i, is obtained from π (the possibility distribution of Π) by the extension principle of fuzzy set theory [38]:

$$\pi^i(t) = \sup_{y : f(x_i, y) = t} \pi(y), \quad \forall t \in \mathbb{R}. \tag{12}$$

It is the membership function of the fuzzy set $f(x_i, \tilde{Y}_0)$. We also observe that Π^i is determined by the mass assignment m_i:

$$m_i(\mathbb{R}) = \nu_0 = \alpha_1, \quad m_i(T_{ij}) = \nu_j = \alpha_{j+1} - \alpha_j, \quad j = 1, \ldots, q-1,$$
$$m_i(T_{iq}) = \nu_q = 1 - \alpha_q.$$

Thus, according to the probability distribution of X and our information about y, the probability measure of T is imprecisely determined by means of the basic assignment m_T that assigns the probability mass $\nu_{ij} = p_i \nu_j$ to each focal T_{ij} ($T_{i0} = \mathbb{R}, \ \forall i$). The associated plausibility and belief functions are given by the expressions:

$$\mathrm{Pl}_T(A) = \sum_{(i,j) : A \cap T_{ij} \neq \emptyset} \nu_{ij}, \quad \forall A; \qquad \mathrm{Bel}_T(A) = \sum_{(i,j) : T_{ij} \subseteq A} \nu_{ij}, \quad \forall A.$$

This view comes down to considering the random fuzzy set, assigning to each realization x_i the fuzzy set π^i, 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 describing our information about T coincides with the arithmetic mean of the possibility measures Π^i (weighted by the probabilities of the different values of X,) i.e.

$$\mathrm{Pl}_T(A) = \sum_{(i,j) : A \cap T_{ij} \neq \emptyset} \nu_{ij} = \sum_{i=1}^{m} \sum_{j : A \cap T_{ij} \neq \emptyset} p_i \nu_j = \sum_{i=1}^{m} p_i \Pi^i(A), \quad \forall A. \tag{13}$$

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

$$\mathrm{Bel}_T(A) = \sum_{(i,j):T_{ij}\subseteq A} v_{ij} = \sum_{i=1}^{m} \sum_{j:T_{ij}\subseteq A} p_i v_j = \sum_{i=1}^{m} p_i N^i(A), \quad \forall A. \tag{14}$$

These expressions are special cases of definitions independently proposed by Dubois and Prade [11] in the mid-eighties, Yen [34] in the early nineties for fuzzy events. Taking into account the properties of possibility measures as upper envelopes of sets of probability measures (see [15], for finite universes and [3,4,24], for the general case), we get the equalities:

$$\mathrm{Pl}_T(A) = \sup\left\{\sum_{i=1}^{m} p_i P_i(A) : P_i \in \mathscr{P}(\Pi^i),\ i = 1,\ldots,m\right\}, \tag{15}$$

$$\mathrm{Bel}_T(A) = \inf\left\{\sum_{i=1}^{m} p_i P_i(A) : P_i \in \mathscr{P}(\Pi^i),\ i = 1,\ldots,m\right\}. \tag{16}$$

These equations suggest another probabilistic interpretation of these plausibility and belief functions, laid bare in case (a) of the example: 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 $P(T \in A|X = x_i)$ that T takes a value in A is less than or equal to $\Pi^i(A)$, and at least equal to $N^i(A)$. In other words,

$$N^i(A) \leqslant P(T \in A|X = x_i) = P_i(A) \leqslant \Pi^i(A).$$

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 quantity $P_T(A) = P(T \in A)$ 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 Eq. (15), we can interpret the values $\mathrm{Pl}_T(A)$ and $\mathrm{Bel}_T(A)$ as the most precise bounds (the smallest for $\mathrm{Pl}_T(A)$, the largest for $\mathrm{Bel}_T(A)$) for the "true" probability of A, according to the available information.

5. Relationship with existing approaches to fuzzy random variables

At the end of Section 3, we have remarked that a fuzzy random variable could be used to represent our imprecise knowledge about the random variable T. It was patent in the previous section since the resulting knowledge about $T = f(X, Y)$ could be obtained by generating the fuzzy sets $f(x_i, \tilde{Y})$ by picking the x_i's randomly, while $\tilde{Y}$ is a constant fuzzy set, as done by Guyonnet et al. [21]. A fuzzy random variable admits as many interpretations as there are interpretations of fuzzy sets in the literature (see [16] for a detailed description). Next, we are going to briefly review two of these interpretations. According to each one of them, the information provided by the fuzzy random variable will be summarized in a specific way. Thus, we will see how two different interpretations can lead us to a classical model or an order 2 imprecise model, respectively. Our intention in this subsection is to compare the plausibility–belief model constructed in Section 4 with these two interpretations. As indicated in the introduction, a fuzzy random variable associates a fuzzy set, to each possible result of a random experiment. Different definitions of fuzzy random variables (see for instance [10,22,26] or [30]) differ in the way the classical measurability condition of random variables is transferred to this context.

5.1. The imprecise higher order uncertainty setting

In our particular problem, the fuzzy random variable is constructed specifically as follows. Let $(\Omega, \mathscr{A}, P)$ be a probability space. On the one hand, let us consider a random variable, $X : \Omega \to \mathbb{R}$ (X represents the observation of a certain characteristic of each element of Ω). Let us consider, on the other hand, another random variable $Y : \Omega \to \mathbb{R}$. Let us suppose that the available (imprecise) information about Y is given by the fuzzy set associated to the membership function $\pi : \mathbb{R} \to [0,1]$. Hence, it defines a constant mapping, $\tilde{Y} : \Omega \to \tilde{\mathscr{P}}(\mathbb{R})$ that assigns, to every element of Ω, the same fuzzy set, π. Thus, for each ω and each y, $\tilde{Y}(\omega)(y) = \pi(y)$ represents the possibility grade that $Y(\omega)$ coincides with y. This scheme illustrates that our knowledge about the image of Y does not depend on each particular individual, $\omega \in \Omega$. Let us now consider a map $f : \mathbb{R}^2 \to \mathbb{R}$ and the random variable given as $T = f \circ (X, Y)$. The available information about T is given by the fuzzy random variable $\tilde{T} : \Omega \to \tilde{\mathscr{P}}(\mathbb{R})$ defined as follows:

$$\tilde{T}(\omega)(t) = \sup_{\{y \in \mathbb{R} : f(X(\omega), y) = t\}} \pi(y).$$

Let us assume that the element ω of the population is selected. Then, the degree of possibility that the true image $T(\omega)$ is t coincides with the degree of possibility that $Y(\omega)$ belongs to the set $\{y | f(X(\omega), y) = t\}$. Thus, the fuzzy random variable $\tilde{T} : \Omega \to \tilde{\mathscr{P}}(\mathbb{R})$ represents the vague information available about the random variable $T : \Omega \to \mathbb{R}$. Following this approach, we can assign, to each event of the final space, a fuzzy subset of the unit interval. This fuzzy quantity reflects the (vague) information available about the true probability of the event.

In [5] the second author proposed a model of this type. Here, we briefly recall how this model is constructed. Then, we will show the relationship between this model and the plausibility–belief model defined in Section 4. Let us first assume that the fuzzy random variable $\tilde{T} : \Omega \to \tilde{\mathscr{P}}(\mathbb{R})$ represents the following imprecise information about the random variable $T : \Omega \to \mathbb{R}$: for each $\alpha > 0$, the probability of the event "$T(\omega) \in [\tilde{T}(\omega)]_\alpha$, $\forall \omega \in \Omega$" is greater than or equal to $1 - \alpha$. This is in agreement with the fact that the possibility distribution associated with $\tilde{T}$ is equivalent to stating that for each cut $[\tilde{T}(\omega)]_\alpha$, the degree of necessity $N([\tilde{T}(\omega)]_\alpha) \geqslant 1 - \alpha$. Under this interpretation we can say that, for each confidence level $1 - \alpha$, the probability distribution associated to T belongs to the set $\mathscr{P}_{\tilde{T}_\alpha} = \{P_T : T \in S(\tilde{T}_\alpha)\}$, where $S(\tilde{T}_\alpha)$ is the set of selections from the random set $\tilde{T}_\alpha$. Thus, given an arbitrary event A of the final space, the probability $P_T(A)$ belongs to the set

$$\mathscr{P}_{\tilde{T}_\alpha}(A) = \{P_T(A) : T \in S(\tilde{T}_\alpha)\} \tag{17}$$

with confidence level $1 - \alpha$. In [5] the fuzzy set $\tilde{P}_{\tilde{T}}$ of probability functions, with membership function given by the equation:

$$\tilde{P}_{\tilde{T}}(Q) = \sup\{\alpha \in [0,1] : Q \in \mathscr{P}_{\tilde{T}_\alpha}\}, \quad \forall Q$$

is viewed as an imprecise representation of the probability measure P_T. In fact, $\tilde{P}_{\tilde{T}}$ is a possibility distribution on the space of probability functions. According to the available information, the quantity $\tilde{P}_{\tilde{T}}(Q)$ represents the possibility degree that Q coincides with the true probability measure associated to T, P_T. On the other hand, for each event A, the fuzzy number $\tilde{P}_{\tilde{T}}(A)$, defined as

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

represents our imprecise information about the quantity $P_T(A) = P(T \in A)$. Thus, the value $\tilde{P}_{\tilde{T}}(A)(p)$ represents the degree of possibility that the "true" degree of probability $P_T(A)$ is p. De Cooman recently proposed a behavioral interpretation of such fuzzy probabilities [2].

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 [36]. It is recalled in the second section that a possibility measure encodes a set of probability measures (the set of probability measures it dominates). Hence, a second order possibility measure is associated to a set of (meta-)probability measures, each of them defined, as well, over a set of probability measures. Thus, a second order possibility measure allows us to state sentences like "the probability that the true probability of the event A is 0.5 ranges between 0.4 and 0.7". On the other hand, it is easily checked that the set of values considered in Eq. (17) is upper bounded by

$$\mathrm{Pl}_\alpha(A) = P(\{\omega \in \Omega : [\tilde{T}(\omega)]_\alpha \cap A \neq \emptyset\})$$

and lower bounded by

$$\mathrm{Bel}_\alpha(A) = P(\{\omega \in \Omega : [\tilde{T}(\omega)]_\alpha \subseteq A\}).$$

Remark 1. In particular, if A is an interval of the form $(-\infty, x]$, and the final space is finite, then $\mathrm{Pl}_\alpha(A)$ and $\mathrm{Bel}_\alpha(A)$ respectively coincide with the values:

$$F^*_\alpha(x) = \mathrm{Pl}_\alpha((-\infty, x]) = P(\{\omega \in \Omega : \min[\tilde{T}(\omega)]_\alpha \leqslant x\})$$

and

$$F_{*\alpha}(x) = \mathrm{Bel}_\alpha((-\infty, x]) = P(\{\omega \in \Omega : \max[\tilde{T}(\omega)]_\alpha \leqslant x\}).$$

In [18], Ferson and Ginzburg represent the imprecise information about P_T by the nested family of sets of probability measures $\{\mathscr{P}(\{F_{*\alpha}, F^*_\alpha\})\}_{\alpha\in[0,1]}$. Here $\mathscr{P}(\{F_{*\alpha}, F^*_\alpha\})$ represents the set of probability measures obtained from $F_{*\alpha}$ and F^*_α, i.e., the set:

$$\{Q : F_{*\alpha}(x) \leqslant Q(-\infty, x] \leqslant F^*_\alpha(x),\ \forall x \in \mathbb{R}\}.$$

On the other hand, it is important to notice that the set of probability measures defined by the pair (Bel_α,Pl_α) is, in general, more precise than the set of probability measures induced by $(F_{*\alpha}, F^*_\alpha)$. In [6] the relationships between both sets of probabilities are studied.

Remark 2. In contrast, for each $\alpha \in (0,1]$, the set of probability measures $\mathscr{P}_{\tilde{T}_\alpha}$ is generally more precise than the set of probability measures associated to Bel_α and Pl_α. Hence, for an arbitrary event A, the set of values $\mathscr{P}_{\tilde{T}_\alpha}(A)$ is included in the interval $[\mathrm{Bel}_\alpha(A), \mathrm{Pl}_\alpha(A)]$. For finite universes, it is easily checked that the maximum and the minimum of these sets do coincide. (This is not true in general, as shown in [28].) On the other hand, the set of values $\mathscr{P}_{\tilde{T}_\alpha}(A)$ is not necessarily convex, because (as we checked in [29]) the set of probability measures $\mathscr{P}_{\tilde{T}_\alpha}$ is not so either. Generally speaking, when we replace this set by the set of probability measures associated to $(\mathrm{Bel}_\alpha, \mathrm{Pl}_\alpha)$ we may lose meaningful information as shown in [5]. Anyway, with respect to the information we want to provide with the current model, these differences are meaningless, because we are only interested in a pair of bounds (lower and upper) for the probability of every event, and our referential is finite, by hypothesis.

5.2. The average probability interval

In the literature, second order probability measures are used in general to represent the information given by an expert. They indicate subjective degrees of belief about the true probability measure that models a given phenomenon. Here we used a second-order possibility measure. Kyburg [27] argues that one should not combine regular (especially frequentist) probability values and higher order information. However a result in the form of a fuzzy subinterval in the unit interval is more difficult to interpret by a user interested in making a decision on the basis of the probability of an event of interest (like violating a safety threshold). A probability interval is a simpler concept.

There exists a strong relationship between the plausibility measure defined in Section 4 and the fuzzy set $\tilde{P}_{\tilde{T}}$ defined in [5]. In actual fact, we shall prove that, for every event A, the interval $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$ coincides with the mean value [13] (also called the average level [31]) of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$. This result implies, for instance, that, for all $x \in \mathbb{R}$, the interval $[F_*(x), F^*(x)] = [\mathrm{Bel}(-\infty, x], \mathrm{Pl}(-\infty, x]]$ (determined from our plausibility measure) coincides with the mean value and the average level of the fuzzy set determined by the nested family of intervals $\{[F_{*\alpha}(x), F^*_\alpha(x)]\}_{\alpha\in[0,1]}$ considered by Ferson and Ginzburg [18]. Before checking these results, we are going to recall the concepts of "mean value" and"average level" of a fuzzy set. In [13], Dubois and Prade define the "mean value" of a fuzzy number, π, as the interval:

$$M(\pi) = \{E(P) : P \leqslant \Pi\},$$

where $E(P)$ represents the expected value associated to the probability measure P, and Π is the possibility measure associated to the possibility distribution π. That interval represents the set of possible values for the expectation of the outcome of a certain random experiment. Let us recall that π represents a set of probability measures (the set of probability measures dominated by Π). So, $M(\pi)$ represents the set of all possible values for the expectation, when we only know that the probability measure that models the random experiment belongs to this set.

On the other hand, Ralescu defines in [31] the "average level" of a fuzzy number, $\pi : \mathbb{R} \to [0, 1]$, as the integral:

$$A(\pi) = \int_0^1 [\pi]_\alpha \, \mathrm{d}\alpha.$$

In this formula, it is considered the Kudo-Aumann integral of the multi-valued "level" mapping, $L_\pi : [0, 1] \to \mathbb{R}$, with respect to Lebesgue measure (the uniform distribution), over the unit interval. The "level" multi-valued mapping assigns, to each α in [0,1], the α-cut of π,

$$[\pi]_\alpha = \{x \in \mathbb{R} : \pi(x) \geqslant \alpha\}.$$

This last integral yields an interval of numbers. For example, when the fuzzy number is trapezoidal, its average level coincides with its 0.5-cut. In the general case, the author considers a uniform probability distribution over the class of α-cuts and calculates the "expected" α-cut (it is not a cut, in general).

Let us now prove that, for each event A, the interval $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$ considered in Section 4 coincides with the average level of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$ of the second order model.

Theorem 5.1. *Given an arbitrary event A, the interval $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$ coincides with the average level of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$.*

Proof. First of all, let us prove that the average level of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$ coincides with the average level of the fuzzy set associated to the nested family of intervals $\{[\mathrm{Bel}_\alpha(A), \mathrm{Pl}_\alpha(A)]\}_{\alpha\in[0,1]}$: on the one hand, as recalled before, the following equalities hold:

$$\max P_{\tilde{T}_\alpha}(A) = \mathrm{Pl}_\alpha(A) \quad \text{and} \quad \min P_{\tilde{T}_\alpha}(A) = \mathrm{Bel}_\alpha(A).$$

Let us notice that we consider the uniform probability distribution on [0,1] and it is non-atomic. Hence, although the sets in the family $\{P_{\tilde{T}_\alpha}(A)\}_{\alpha\in[0,1]}$ are not necessarily convex, the Aumann integral of the multi-valued mapping they determine is indeed an interval. Thus, the Aumann integral of the multi-valued mapping that assigns the set $P_{\tilde{T}_\alpha}(A)$ to each $\alpha \in [0,1]$, coincides with the Aumann integral of the multi-valued mapping that assigns, to each $\alpha \in [0,1]$, the set $[\mathrm{Bel}_\alpha(A), \mathrm{Pl}_\alpha(A)]$. On the other hand, we can observe that the family of nested sets $\{P_{\tilde{T}_\alpha}(A)\}_{\alpha\in[0,1]}$ determines the fuzzy set $\tilde{P}_{\tilde{T}}(A)$. In other words, the following equations are satisfied:

$$[\tilde{P}_{\tilde{T}}(A)]_{\overline{\alpha}} \subseteq P_{\tilde{T}_\alpha}(A) \subseteq [\tilde{P}_{\tilde{T}}(A)]_\alpha, \quad \forall \alpha \in [0,1),$$

where $[\pi]_{\overline{\alpha}} = \{x \in \mathbb{R} : \pi(x) > \alpha\}$; and then $\tilde{P}_{\tilde{T}}(A)(p) = \sup\{\alpha \in [0,1] : p \in P_{\tilde{T}_\alpha}(A)\}$. So, we conclude that the average level of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$ coincides with the Aumann integral:

$$\int_0^1 [\mathrm{Bel}_\alpha(A),\, \mathrm{Pl}_\alpha(A)]\,\mathrm{d}\alpha.$$

Finally, we easily check that this integral coincides with the closed interval:

$$\left[\int_0^1 \mathrm{Bel}_\alpha(A)\,\mathrm{d}\alpha, \int_0^1 \mathrm{Pl}_\alpha(A)\,\mathrm{d}\alpha\right].$$

(In this last equation, the Lebesgue integral is considered.) According to Eq. (13), the plausibility of an event A, is calculated as follows:

$$\mathrm{Pl}(A) = \sum_{i=1}^{m} \sum_{j:A\cap T_{ij}\neq\emptyset} p_i v_j. \tag{18}$$

By the commutative property of the sum, we can write it alternatively as

$$\mathrm{Pl}(A) = \sum_{j=0}^{q} v_j \left(\sum_{i:A\cap T_{ij}\neq\emptyset} p_i \right), \tag{19}$$

where $v_0 = \alpha_1$, $v_j = \alpha_{j+1} - \alpha_j$, $j = 1, \ldots, q-1$, $v_q = 1 - \alpha_q$ and $p_i = P(X = x_i)$, $i = 1, \ldots, m$. Now, for each j we can calculate the value $\mathrm{Pl}_{\alpha_j}(A)$ as follows:

$$\mathrm{Pl}_{\alpha_j}(A) = P(\{\omega \in \Omega : \tilde{T}_{\alpha_j}(\omega) \cap A \neq \emptyset\}) = \sum_{i:A \cap T_{ij} \neq \emptyset} p_i. \tag{20}$$

This way, we easily check that the quantity $\mathrm{Pl}(A)$ coincides with the Lebesgue integral $\int_0^1 \mathrm{Pl}_\alpha(A)\,\mathrm{d}\alpha$. We should check in an analogous way that $\mathrm{Bel}(A)$ coincides with the integral $\int_0^1 \mathrm{Bel}_\alpha(A)\,\mathrm{d}\alpha$. □

This last result improves one in [1], where we only proved that $\mathrm{Pl}(A)$ ranges between $\mathrm{Pl}_1(A)$ and $\mathrm{Pl}_0(A)$. Let us provide a direct proof that the interval $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$ also coincides with the "mean value" of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$.

Theorem 5.2. *Given an arbitrary event A, the interval $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$ coincides with the mean value of the fuzzy set $\tilde{P}_{\tilde{T}}(A)$.*

Proof. The mean value of that fuzzy set is given by the expression:

$$M(\tilde{P}_{\tilde{T}}(A)) = \left\{ \int \mathrm{id}\,\mathrm{d}P : P \leqslant \Pi_{\tilde{P}_{\tilde{T}}(A)} \right\},$$

where $\Pi_{\tilde{P}_{\tilde{T}}(A)}$ represents the possibility measure determined by the fuzzy set $\tilde{P}_{\tilde{T}}(A)$. We can easily check that this mean value is a closed interval. Its minimum and maximum values respectively coincide with the Choquet integrals of the identity function with respect to $N_{\tilde{P}_{\tilde{T}}(A)}$ and $\Pi_{\tilde{P}_{\tilde{T}}(A)}$. This pair of dual necessity and possibility measures is associated to the focal sets $\mathscr{P}_{T_{\alpha_1}}(A), \ldots, \mathscr{P}_{T_{\alpha_q}}(A)$, with respective mass assignments $v_1, \ldots, v_q$. We deduce that the Choquet integral of the identity map with respect to $\Pi_{\tilde{P}_{\tilde{T}}(A)}$ coincides with the sum $\sum_{j=0}^{q} v_j \max \mathscr{P}_{T_{\alpha_j}}(A)$. On the other hand, as we pointed out above, each value $\max \mathscr{P}_{T_{\alpha_j}}(A)$ coincides with the plausibility $\mathrm{Pl}_{\alpha_j}(A)$. This way, we observe that the maximum of the set $M(\tilde{P}_{\tilde{T}}(A))$ coincides with the value $\sum_{j=0}^{q} v_j \mathrm{Pl}_{\alpha_j}(A)$. Now, according to Eqs. (18)–(20), we deduce that $\max M(\tilde{P}_{\tilde{T}}(A))$ coincides with $\mathrm{Pl}(A)$. We could check, in an analogous way, that the minimum of the interval $M(\tilde{P}_{\tilde{T}}(A))$ coincides with the value $\mathrm{Bel}(A)$. □

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. The necessity measure appears in a natural way from confidence levels. Hence, for each α, we assign the lower probability (degree of necessity) $1-\alpha$ to a set of probability measures. This order 2 necessity measure provides the same information as its dual possibility measure. Both of them are equivalent to a set of second order probability measures, as we recalled in Section 5.1. Let us consider a particular second order probability measure, $\mathbb{P}$, belonging to this set. Let us also fix an arbitrary event A. In this setting, we can define a random variable that takes each value $Q(A)$ with probability $\mathbb{P}(\{Q\})$.[1] If $\mathbb{P}$ was the "correct" second order probability 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

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

containing 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, Bel(A) and Pl(A) represent, in the average, the most precise bounds for the "true" probability of A, under the available information.

5.3. Relationship with the "classical" model of fuzzy random variables

In the last subsection we have compared our plausibility–belief model with the second order possibility distribution associated to the fuzzy random variable $\tilde{T}$: This second order model associates, to each event (crisp subset of the final space), a fuzzy set in the unit interval. However, another point of view in the literature leads to assigning a (crisp) value of probability to each possible (fuzzy) image of $\tilde{T}$, by considering a fuzzy random variable as a "classical" measurable mapping. Krätschmer [23] reviews all the previous definitions of fuzzy random variables in the literature, and offers a unified vision. He considers specific topologies defined on a certain class of fuzzy subsets of $\mathbb{R}$. A fuzzy random variable is then a measurable function. This "classical" vision of a fuzzy random variable viewing it as a measurable function agrees with the interpretation given by Puri and Ralescu [30]. 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 linguistic "values", its induced probability is determined by the mass function. Therefore, it will suffice to specify which are 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 (for example, we could generate a model of the following type: the probability that the result is "high" is 0.5, the probability of being "average" is 0.25 and the probability of being "low" is 0.25).

In our particular problem, the fuzzy random variable, $\tilde{T}$ has m different images (m different fuzzy sets), one per each value of X. As we assume at the beginning of Section 4, X takes a finite number of different values, $x_1, \ldots, x_m$ with respective probabilities $p_1, \ldots, p_m$. In other words, X is a "simple" mapping, $X = \sum_{i=1}^{m} x_i I_{C_i}$, where $\{C_1, \ldots, C_m\}$ is a partition of Ω and $P(C_i) = p_i$, $\forall i = 1, \ldots, m$. For each $i \in \{1, \ldots, m\}$, and each $\omega \in C_i$, the fuzzy set $\tilde{T}(\omega)$ coincides with the fuzzy set π^i, defined from f and π as indicated in Eq. (12). Thus, the fuzzy random variable $\tilde{T}$ takes its m "values" $\pi^1, \ldots, \pi^m$ with respective probabilities $p_1, \ldots, p_m$. These m fuzzy sets and their respective probabilities uniquely determine the probability distribution induced by $\tilde{T}$, considered as a classical measurable function.

If a question referring to the universe underlying the linguistic values of $\tilde{T}$ must be addressed, of the form "is $T \in A$?" , then if the knowledge about $\tilde{T}$ is π^i, one may provide the possibility and the necessity degrees of A according to π^i. In the fuzzy random setting here, the belief and plausibility functions defined in the previous section:

$$\mathrm{Pl}(A) = \sum_{i=1}^{m} p_i \sup_{t \in A} \pi^i(t), \quad \forall A,$$

$$\mathrm{Bel}(A) = \sum_{i=1}^{m} p_i \inf_{t \notin A} (1 - \pi^i(t)), \quad \forall A$$

represent the average possibility and necessity degrees for the event A on the universe underlying the random fuzzy set $\tilde{T}$.

6. Joint propagation of probabilistic and ill-known deterministic information

In this section, more information available information about T is available than in Section 4: now the value of Y is a constant y_0, which is only partially known, but of course does not depend on X. According to case (b) of the motivating example, the information given by the fuzzy random variable $\tilde{T}$ previously defined from Π and X is compatible with our vague knowledge about T. Nevertheless, the new assumption leads to more precise results.

The random variable $X : \Omega \to \mathbb{R}$, again takes known values $x_1, \ldots, x_m$ with respective known probabilities $p_1, \ldots, p_m$. Let us consider, on the other hand an imprecisely known fixed number $y_0 \in \mathbb{R}$. Let us suppose that our information about y_0 is given by means of subjective "confidence levels". Thus, we will assume that there is a family of nested sets, $A_1 \supseteq \cdots \supseteq A_q$, with respective confidence degrees $1 - \alpha_1 \geqslant \cdots \geqslant 1 - \alpha_q$. As in the previous section, our information about y_0 is given by the possibility measure Π in Eq. (11). Our information about Y is now more precise than in Section 4: Not only is the *information about* $Y(\omega)$ constant, independent of ω but, furthermore we know that the real value of Y is constant (even if unknown) : $Y(\omega) = y_0$, $\forall \omega \in \Omega$. Let us now consider the random variable $T = f(X, y_0)$, where $f : \mathbb{R}^2 \to \mathbb{R}$ is a known mapping. The objective probability measure P_T attached to $T : \Omega \to \mathbb{R}$ only depends on X. The random variable T satisfies again all the properties of T in Section 4. Thus, we know that its probability distribution is dominated by the plausibility measure there defined. But now we know in addition that Y is constant. So, we will look for a more precise upper probability. In other words, we will find a set function P^* that satisfies the inequalities:

$$P_T(A) \leqslant P^*(A) \leqslant \mathrm{Pl}(A), \quad \forall A.$$

6.1. Main result

Let us denote by Γ the "most precise set *we know that* contains y_0". It is a *random* set that takes the "values" $A_0 = \mathbb{R} \supseteq A_1 \supseteq \cdots \supseteq A_q$ with respective probabilities

$$m(\mathbb{R}) = v_0 = \alpha_1, \quad m(A_j) = v_j = \alpha_{j+1} - \alpha_j, \;\; j = 1, \ldots, q-1, \quad m(A_q) = v_q = 1 - \alpha_q.$$

m is the basic mass assignment associated to Π. Following [32], each quantity $m(C)$ is understood to be the fraction of belief that is committed exactly to C, and to nothing smaller. Note that the conditional probability $P_T(A|A_i)$ is the probability that $T \in A$, given that all we know about y_0 is that $y_0 \in A_i$. So we can write it $P(T \in A|\Gamma = A_j)$. Moreover, $P_T(A|y) = P(f(X, Y) \in A|y) = P(\{\omega \in \Omega : f(X(\omega), y) \in A\})$. Thus, for each event A, we have

$$\begin{aligned} P_T(A) = P(T \in A) &= \sum_{j=0}^{q} P(T \in A|\Gamma = A_j) \cdot P(\Gamma = A_j) \\ &\leqslant \sum_{j=0}^{q} \sup\{P(\{\omega \in \Omega : f(X(\omega), y) \in A\}) : y \in A_j\} \cdot v_j = P^*(A), \end{aligned} \tag{21}$$

since, in general $P_T(A|B) \leqslant \sup_{y\in B} P_T(A|y)$. The upper probability $P^*(A)$ coincides with the Choquet integral of $g_A : \mathbb{R} \to [0,1]$ with respect to Π, where g_A is given by $g_A(y) = P(\{\omega \in \Omega : f(X(\omega), y) \in A\})$, as suggested in the motivating example case (b).

We easily observe that P^* is dominated by the plausibility function, Pl defined in Section 4: As we check in Eq. (19), Pl can be written as follows:

$$\mathrm{Pl}(A) = \sum_{j=0}^{q} v_j \left(\sum_{i:A\cap T_{ij}\neq\emptyset} p_i \right).$$

Furthermore, for each $j \in \{0,\dots,q\}$, the quantity $\sum_{i:A\cap T_{ij}\neq\emptyset} p_i$ coincides with the probability

$$P(\{\omega \in \Omega : f(X(\omega), A_j) \cap A \neq \emptyset\}).$$

And we easily check, for each $j = 0,\dots,q$, that

$$P(\{\omega \in \Omega : f(X(\omega), A_j) \cap A \neq \emptyset\}) \geqslant \sup_{y\in A_j} P(\{\omega \in \Omega : f(X(\omega), y) \in A\}),$$

since $\{\omega \in \Omega : f(X(\omega), A_j) \cap A \neq \emptyset\} = \cup_{y\in A_i}\{\omega \in \Omega : f(X(\omega), y) \in A\}$. Hence, we have found an upper probability measure, P^*, that dominates P_T and is more precise than Pl. In other words,

$$P_T(A) \leqslant P^*(A) \leqslant \mathrm{Pl}(A), \quad \forall A.$$

Its dual lower probability measure, P_* is given by the formula

$$P_*(A) = \sum_{j=0}^{q} \inf\{P(\{\omega \in \Omega : f(X(\omega), y) \in A\}) : y \in A_j\} \cdot v_j, \tag{22}$$

and it satisfies the inequalities

$$\mathrm{Bel}(A) \leqslant P_*(A) \leqslant P_T(A), \quad \forall A.$$

6.2. Relation with second order models

Let us consider, for each $y \in \mathbb{R}$, the random variable $T_y : \Omega \to \mathbb{R}$, given by $T_y(\omega) = f(X(\omega), y)$, $\forall \omega \in \Omega$. We can say that $T \in \mathscr{S}_j = \{T_y : y \in A_j\}$ with confidence $1 - \alpha_j$, $\forall j = 1,\dots,q$. Thus, with confidence $1 - \alpha_j$ we know that P_T belongs to the set

$$\mathscr{P}_j = \{P_{T_y} : y \in A_j\}.$$

So, we can describe our information about P_T by a family of nested sets of probability measures. The class of (meta-) probabilities, $\mathbb{P}$, satisfying

$$\mathbb{P}(\mathscr{P}_j) \geqslant 1 - \alpha_j, \quad \forall j = 1,\dots,q$$

coincides with the set of (meta-) probabilities dominated by the (second order) possibility distribution, $\tilde{P}$, given by

$$\tilde{P}(Q) = \begin{cases} \alpha_1 & \text{if } Q \notin \mathscr{P}_1, \\ \alpha_{j+1} & \text{if } Q \in \mathscr{P}_j \setminus \mathscr{P}_{j+1}, \quad j = 1,\dots,q-1. \\ 1 - \alpha_q & \text{if } Q \in \mathscr{P}_q. \end{cases}$$

This second order possibility measure is dominated by the one considered in Section 5.1. (This is easily checked, as each set $\mathscr{P}_j$ is included in $\mathscr{P}_{T_{\alpha_j}}$.)

Furthermore, for a particular event A, the probability value $P_T(A)$ can be described by the fuzzy subset of $[0,1]$, $\tilde{P}(A)$:

$$\tilde{P}(A)(p) = \begin{cases} \alpha_1 & \text{if } p \notin B_1, \\ \alpha_{j+1} & \text{if } p \in B_j \setminus B_{j+1}, \quad j = 1,\ldots,q-1, \\ 1-\alpha_q & \text{if } p \in B_q, \end{cases}$$

where $\{B_1,\ldots,B_q\}$ is the family of nested sets defined as

$$B_j = \{P_{T_y}(A) | y \in A_j\} = \{g_A(y) | y \in A_j\}, \quad \forall j = 1,\ldots,q$$

in the sense that $\tilde{P}(A)(p)$ represents the possibility that $P_T(A)$ coincides with p, $\forall p \in [0,1]$. We easily observe that this last fuzzy set is included in $P_{\tilde{T}}(A)$. Moreover, the (first order) model considered in Section 6 coincides with the mean value of this fuzzy set. In other words, $[P_*(A), P^*(A)] = M(\tilde{P}(A)), \forall A$, as we prove in the following theorem.

Theorem 6.1. *Given an arbitrary event A, the interval $[P_*(A), P^*(A)]$ coincides with the mean value of the fuzzy set $\tilde{P}(A)$.*

Proof. Let us consider an arbitrary event, A. First of all, let us notice that the possibility distribution $\tilde{P}(A)$ is associated to the mass assignment $m_{\tilde{P}(A)}$:

$$m_{\tilde{P}(A)}([0,1]) = \nu_0 = \alpha_1,$$
$$m_{\tilde{P}(A)}(B_j) = \nu_j = \alpha_{j+1} - \alpha_j, \quad j = 1,\ldots,q-1,$$
$$m_{\tilde{P}(A)}(B_q) = \nu_q = 1 - \alpha_q.$$

Furthermore, let us observe that all the focal subsets have maximum and minimum value: $B_0 = [0,1]$ is a closed interval and B_j is finite, for all $j \in \{1,\ldots,q\}$ (each B_j can be written as a set of elements of the form $\sum_{i\in I} p_i$, with $I \subseteq \{1,\ldots,m\}$).

Thus, $M(\tilde{P}(A))$ has a maximum and a minimum value (the maximum value is the expected value of the probability measure that associates the mass ν_j to $\max B_j$, $\forall j = 0,\ldots,q$. The minimum is the expectation of the probability measure that associates the mass ν_j to $\min B_j$, $\forall j = 0,\ldots,q$. In other words,

$$\max M(\tilde{P}(A)) = \sum_{j=0}^{q} \nu_j \max B_j, \quad \min M(\tilde{P}(A)) = \sum_{j=0}^{q} \nu_j \min B_j.$$

These two values respectively coincide with $P^*(A)$ and $P_*(A)$, as we observe in Eqs. (21) and (22). On the other hand, $M(\tilde{P}(A))$ is convex (the mean value of any fuzzy number is convex.) Hence, it coincides with the closed interval $[P_*(A), P^*(A)]$. □

The intuitive meaning of this last result is similar to the interpretation of the result given in Theorem 5.2.

7. A general setting for joint possibility–probability propagation

In this section, a more general dependence setting is assumed that encompasses the third scenario of the motivating example as a particular case. Let us consider again a random variable $X : \Omega \to \mathbb{R}$, that takes the values $x_1, \ldots, x_m$ with respective probabilities $p_1, \ldots, p_m$. In other words, X is a "simple" mapping, $X = \sum_{i=1}^{m} x_i I_{C_i}$, where $\{C_1, \ldots, C_m\}$ is a partition of Ω and $P(C_i) = p_i$, $\forall i = 1, \ldots, m$. Assume that we have imprecise information about a random variable $Y : \Omega \to \mathbb{R}$ and it is also represented by a possibility distribution π. But now, this possibility distribution is obtained as follows: let us consider another partition of Ω, $\{D_0, \ldots, D_q\}$. For each $\omega \in D_j$, we know that $Y(\omega)$ belongs to the set A_j. Then, the information about Y is determined by the random set $\Gamma_Y : \Omega \to \mathbb{R}$ defined as $\Gamma(\omega) = A_j$, $\forall \omega \in D_j$, $j = 0, \ldots, q$. (All we know about Y is that it is a measurable selection of Γ_Y). The upper probability of Γ_Y is associated to the mass assignment m_{Γ_Y}:

$$m_{\Gamma_Y}(A_j) = P(D_j), \quad j = 0, \ldots, q.$$

If, in addition, the images of Γ_Y are nested sets, $A_0 = \mathbb{R} \supseteq A_1 \supseteq \cdots \supseteq A_q$, then this upper probability is a possibility measure. In this case, all we know about P_Y is that it is dominated by this possibility measure.

The main assumption in this model is that the pair $(X(\omega), Y(\omega))$ is generated by a single occurrence ω, and the information about it is in the form of a random set $\Gamma_{X,Y}$ such that if $\omega \in C_i \cap D_j$ then $\Gamma_{X,Y}(\omega) = \{x_i\} \times A_j$. Let T be the random variable given by $T(\omega) = f(X(\omega), Y(\omega))$, $\forall \omega \in \Omega$.

Under these conditions, all we know about the random variable $T = f(X, Y) : \Omega \to \mathbb{R}$ is that it is a selection of the random set $\Gamma_T : \Omega \to \mathscr{P}(\mathbb{R})$ defined as

$$\Gamma_T(\omega) = f(X(\omega), \Gamma(\omega)) = \{f(X(\omega), Y(\omega)) : Y(\omega) \in \Gamma(\omega), \quad \forall \omega \in \Omega\}.$$

The basic mass assignment of Γ_T assigns, to each focal $T_{ij} = f(x_i, D_j)$, the mass

$$m_{\Gamma_T}(T_{ij}) = P(C_i \cap D_j), \quad \forall i, j.$$

This mass assignment is associated to the plausibility function:

$$\mathrm{Pl}_{\Gamma_T}(A) = \sum_{i=1}^{m} \sum_{j : f(x_i, A_j) \cap A \neq \emptyset} p_{ij}, \tag{23}$$

where $p_{ij} = P(C_i \cap D_j)$, $\forall i$, j. Under the available information, Pl_{Γ_T} is the most precise upper probability that dominates P_T.

Note that $\sum_{i=1}^{m} p_{ij} = m_{\Gamma_Y}(A_j)$, $\forall j$ and $\sum_{j=0}^{q} p_{ij} = p_i, \forall i$. In practice, only the marginals of the joint mass assignment are known [12], because no assumption is made about the relationship between the observation processes. If, in particular, observations C_i and D_j are independent (hence the sets always intersect), $\forall i$, j then, this model reduces to the one given in Section 4. It corresponds to the case where both pieces of information about X and Y are independent of one another. In any other case, the plausibility Pl_{Γ_T} does not dominate, neither dominates the plausibility measure considered in Section 4. Conversely, if $C_i = D_i$, $\forall i$, and $p_i = m_{\Gamma_Y}(A_i), \forall i$, then we recover the scenario of total dependence of sources of information in case (c) of the motivating example.

Actually, we used neither the fact that Pl_{Γ_Y} is a possibility measure, nor that X is a pure random variable, to build this last propagation model. Thus, when both pieces of information Γ_X and Γ_Y about X and Y are arbitrary random sets (whose respective images are not necessarily nested) the plausibility measure Pl_{Γ_T} is still given by formula (23), where the argument of the summation is replaced by $f(X_i, A_j) \cap A \neq \emptyset$, and X_i is an imprecisely perceived realization of X. This is the framework already proposed by Dubois and Prade [14] for encompassing both the calculus of random variables (when X and Y are independent random variables) and the extension principle for fuzzy intervals (when X and Y are ill-known quantities described by possibility distributions, whose dependence is not known, but that are informed by fully dependent sources). This is also the framework more recently adopted by Krieger and Held [25] in the study of climate prediction, using the terminology of belief functions for the propagation of imprecise cumulative distribution functions.

8. Conclusion and open problems

This paper tries to provide some interpretative and formal foundations to some techniques of uncertainty propagation that were used in risk assessment by various authors (Guyonnet et al [21], Ferson and Ginzburg [19], Baudrit [1], Fetz and Oberguggenberger [17], Krieger and Held [25], and others) for combining incomplete and random pieces of information. Our approach encompasses several dependence assumptions between ill-known quantities. Assuming independence between sources of information yields a result consisting of a belief function (or a random set) which averages the fuzzy random variable that is produced from a function having as arguments random variables and fuzzy intervals. The higher order model described in Section 5.1 is more faithful to the actual information and does not keep imprecision and randomness separate, since it gives a fuzzy-interval-valued probability. The aggregated model in Section 4 does not destroy all higher order information in the end since an interval-valued probability still displays information in terms of frequency (in the case of objective variability) and amount of subjective ignorance.

We pointed out that using a possibility measure to describe our knowledge about an ill-known quantity presupposes nothing about the nature of this quantity: it can be an ill-known random variable as described by an expert or observed from an imperfect statistical experiment, or it can refer to a fixed ill-known quantity. We studied the joint propagation of probability and possibility under this multiple-facetted view. If Y is an ill-known random variable, then the knowledge about Y, when provided by a human expert, generally does not depend on the occurrences of other random variables. However, knowing that Y is a fixed quantity has useful impact on the propagation process: If X is a known random variable and Y a fixed quantity, then we know that Y does not depend on X (as in scenario (b) of the motivating example), a piece of information that further reduces the uncertainty of the result.

Future works should try to further formalize notions of independence in the presence of variability and imprecision. It is clear that in a pure probabilistic approach it is not very convenient to formally distinguish between independent variables and independent observations of these variables (unless resorting to very complex higher order probability approaches). The framework of random sets in the sense of Dempster [9] (in terms of probability bounds) and the notion of uncertain evidence by Shafer [32] make it possible to lay bare this important distinction. Clarifying these issues, following the works of Couso et al. [7], Ben Yaghlane et al. [37], and Cozman [8], is an important line of further research.

Acknowledgements

The first and third authors acknowledge partial support by French risk analysis institutes BRGM, IRSN, INERIS. The first author also benefited from a travel grant from Toulouse III University.

References

[1] C. Baudrit, D. Dubois, D. Guyonnet, H. Fargier, Joint treatment of imprecision and randomness in uncertainty propagation, in: B. Bouchon-Meunier, G. Coletti, R.R. Yager (Eds.), Modern Information Processing: From Theory to Applications, Elsevier, 2006, pp. 37–47.

[2] G. De Cooman, A behavioural model for vague probability assessments, Fuzzy Sets and Systems 154 (3) (2005) 305–358.

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

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

[5] I. Couso, S. Montes, P. Gil, Second order possibility measure induced by a fuzzy random variable, in: C. Bertoluzza, M.A. Gil, D.A. Ralescu (Eds.), Statistical Modeling, Analysis and Management of Fuzzy Data, Springer, Heidelberg, 2002, pp. 127–144.

[6] I. Couso, E. Miranda, G. de Cooman, A possibilistic interpretation of the expectation of a fuzzy random variable, in: M. López-Díaz, M.A. Gil, P. Grzegorzewski, O. Hryniewicz, J. Lawry (Eds.), Soft Methodology and Random Information Systems, Springer, Heidelberg, 2004, pp. 133–140.

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

[8] F.G. Cozman, Graphical models for imprecise probabilities, International Journal of Approximate Reasoning 39 (2–3) (2005) 167–184.

[9] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics 38 (1967) 325–339.

[10] P. Diamond, P. Kloeden, Metric Spaces of Fuzzy Sets, World Scientific, Singapore, 1994.

[11] D. Dubois, H. Prade, Evidence measures based on fuzzy information, Automatica 21 (5) (1985) 547–562.

[12] D. Dubois, H. Prade, On the unicity of Dempster rule of combination, International Journal of Intelligent Systems 1 (2) (1986) 133–142.

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

[14] D. Dubois, H. Prade, Random sets and fuzzy interval analysis, Fuzzy Sets and Systems 42 (1991) 87–101.

[15] D. Dubois, H. Prade, When upper probabilities are possibility measures, Fuzzy Sets and Systems 49 (1992) 65–74.

[16] D. Dubois, H. Prade, The three semantics of fuzzy sets, Fuzzy Sets and Systems 90 (1997) 141–150.

[17] T. Fetz, M. Oberguggenberger, Propagation of uncertainty through multivariate functions in the framework of sets of probability measures, Reliability Engineering and System Safety 85 (2005) 73–87.

[18] S. Ferson, L.R. Ginzburg, Hybrid Arithmetic, in: Proceedings of ISUMA-NAFIPS 95, IEEE Computer Society Press, Los Alamitos, CA, 1995, pp. 619–623.

[19] S. Ferson, L.R. Ginzburg, Different methods are needed to propagate ignorance and variability, Reliability Engineering and System Safety 54 (1996) 133–144.

[20] M.A. Gil (Ed.), Fuzzy random variables, Information Sciences 133 (1–2) (2001) (special issue).

[21] D. Guyonnet, D. Dubois, B. Bourgine, H. Fargier, B. Côme, J.-P. Chilès, Hybrid method for addressing uncertainty in risk assessments, Journal of Environmental Engineering 129 (2003) 68–78.

[22] E.P. Klement, M.L. Puri, D.A. Ralescu, Limit theorems for fuzzy random variables, Proceedings of the Royal Society of London A 407 (1986) 171–182.

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

[24] V. Krätschmer, When fuzzy measures are upper envelopes of probability measures, Fuzzy Sets and Systems 138 (2003) 455–468.

[25] E. Kriegler, H. Held, Utilizing belief functions for the estimation of future climate change, International Journal of Approximate Reasoning 39 (2–3) (2005) 185–209.

[26] H. Kwakernaak, Fuzzy random variables, definition and theorems, Information Sciences 15 (1989) 1–29.

[27] H.E. Kyburg, Higher order probabilities and intervals, International Journal of Approximate Reasoning 2 (1988) 195–209.

[28] E. Miranda, I. Couso, P. Gil, Upper probabilities attainable by distributions of measurable selections, Technical Report TR04/05 on statistics and decision sciences, University Rey Juan Carlos, 2004.

[29] E. Miranda, I. Couso, P. Gil, Random sets as imprecise random variables, Journal of Mathematical Analysis and Applications 307 (1) (2005) 32–47.

[30] M.L. Puri, D. Ralescu, Fuzzy Random Variables, Journal of Mathematical Analysis and Applications 114 (1986) 409–422.
[31] D. Ralescu, Average level of a fuzzy set, in: C. Bertoluzza, M.A. Gil, D.A. Ralescu (Eds.), Statistical Modeling, Analysis and Management of Fuzzy Data, Springer, Heidelberg, 2002, pp. 119–126.
[32] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[33] P. Smets, The normative representation of quantified beliefs by belief functions, Artificial Intelligence 92 (1997) 229–242.
[34] J. Yen, Generalizing the Dempster–Shafer theory to fuzzy sets, IEEE Transactions on Systems, Man and Cybernetics 20 (1990) 559–570.
[35] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall (1991).
[36] P. Walley, Statistical inference based on a second-order possibility distribution, International Journal of General Systems 26 (1997) 337–383.
[37] B. Ben Yaghlane, P. Smets, K. Mellouli, Belief function independence: I. The marginal case; II. The conditional case, International Journal of Approximate Reasoning 31 (2002) 31–75.
[38] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Information Sciences (1975), Part I: 8, 199–249; Part II: 8, 301–357; Part III: 9, 43–80.
[39] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.