Higher order models for fuzzy random variables

Inés Couso[a,*], Luciano Sánchez[b]

[a]*Department of Statistics and O.R., University of Oviedo, Spain*
[b]*Department of Computer Sciences, University of Oviedo, Spain*

Abstract

A fuzzy random variable is viewed as the imprecise observation of the outcomes in a random experiment. Since randomness and vagueness coexist in the same framework, it seems reasonable to integrate fuzzy random variables into imprecise probabilities theory. Nevertheless, fuzzy random variables are commonly presented in the literature as classical measurable functions associated to a classical probability measure. We present here a higher order possibility model that represents the imprecise information provided by a fuzzy random variable. We compare it with previous classical models in the literature.

First, some aspects about the acceptability function associated to a fuzzy random variable are investigated. Secondly, we present three different higher order possibility models, all of them arising in a natural way. We investigate their similarities and differences, and observe that the first one (the fuzzy probability envelope) is the most informative. Finally we compare the fuzzy probability envelope with the (classical) probability measure induced by the fuzzy random variable. We conclude that the classical probability measure does not always contain all relevant information provided by a fuzzy random variable.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Imprecise probabilities; Second order possibility measure; Random set; Fuzzy random variable

1. Introduction

The concept of fuzzy random variable (frv) (also called "random fuzzy set") was introduced at the end of the 1970s (see [29]) in order to deal with situations where the outcomes of a random experiment are modelled by fuzzy sets. A frv is a mapping that associates a fuzzy set to each element of a universe (provided of probability space structure). In other words, it assigns a fuzzy subset of the final space to each possible result of a random experiment. This association expresses the available (imprecise) information about the relation between both universes. Thus, this concept generalizes the definition of random variable. But this generalization is not unique. Each definition in the literature differs from the others in the structure of the final space and the way the measurability condition is transferred to this context: Kwakernaak [29] and Puri and Ralescu [38], for instance, focus on the properties of the multi-valued mappings associated to the α-cuts. Kwakernaak assumes that the outcomes of the frv are fuzzy real subsets and the extreme points of their α-cuts are classical random variables. Puri and Ralescu require the α-cuts to be measurable (also different conditions for measurability of multi-valued mappings can be formulated). On the other hand, Klement et al. [24] and Diamond and Kloeden [17] define frv's, as classical measurable mappings. (They first define a particular metric

---

* Corresponding author.
E-mail addresses: couso@uniovi.es (I. Couso), luciano@uniovi.es (L. Sánchez).

on a class of fuzzy subsets, and then consider the induced Borel σ-algebra.) Krätschmer [26] revises all these previous works and examines some relationships between different measurability conditions. He investigates the properties of a special topology that can be defined on a class of fuzzy subsets proves that some of the previous definitions in the literature can be unified and viewed as particular cases of a new definition. In particular, some of the definitions based on the measurability properties of α-levels of the (frv for short) can be reformulated as classical measurability conditions, i.e., the frv can be viewed as an $\mathcal{A}$–$\mathcal{B}$ measurable mapping, where $\mathcal{B}$ is a particular σ-algebra (the Borel σ-algebra induced by certain topology) defined on a class of fuzzy subsets of the final space.

Within this framework, we can use the tools of general Probability Theory to define the probability distribution induced by a frv. We can also extend the concepts of expectation, variance, etc. by reproducing classical techniques. For instance, when the images of the frv are convex fuzzy subsets of $\mathbb{R}$, we can use fuzzy arithmetic to derive a method of construction of the expectation: a limit-based construction analogous to Lebesgue integral definition (using Zadeh's Extension Principle for the sums and products needed to define the expectation of "simple" frv) should lead to a definition of expectation which is consistent with Puri and Ralescu's one [38]. This expectation is a fuzzy subset of the final space and it plays the role of the "average value" of the frv. On the other hand, we can make a parallel construction of the variance: let us consider a particular metric defined over the class of fuzzy subsets of the final space. In this setting, we can define the variance of a frv as the mean (classical expectation of a random variable) of the squares of the distances from the images of the frv to the (fuzzy) expectation. The respective definitions of variance given by Feng [20], Körner [25] and Lubiano [31] fit this formulation. In this context the variance of a frv is a (precise) number that quantifies the degree of dispersion of the images of the frv.

Thus, from a purely formal point of view, this "classical" view of the probability model induced by a frv allows us to use the tools of Probability Theory to transfer different classical concepts and results from the Probability Theory to this new environment. Unfortunately, this kind of mathematical model is not useful in some problems where the images of the frv represent the imprecise observations of the outcomes of a random experiment [5,37,39–45]. It assigns a precise probability value to each (measurable) sub-class of fuzzy subsets. If, in particular, the frv has a finite number of images, probability values can be assigned to different fuzzy labels. For instance, the following model could be generated: the result is "approximately 8" with probability 0.5, "approximately 6" with probability 0.25 and "approximately 2" with probability 0.25, where "approximately 8", (resp. 6 and 2) are linguistic labels associated to particular fuzzy subsets of $\mathbb{R}$. This last model does not reflect the available (imprecise) information about the "true" probability distribution that governs the random experiment. Our imprecision should be reflected in our knowledge about the probability of each event. These events are in fact crisp subsets of the final space, but our information about their probability of occurrence is imprecise. Hence we should look for an imprecise model that assigns an imprecise quantity (a crisp or a fuzzy subset of the unit interval) to each particular event (measurable crisp set of the final space). Furthermore, the extensions of the concepts of expectation and variance should reflect, under this interpretation, our imprecise knowledge about the true (crisp) values of the expectation and variance. Both of them should be defined as imprecise quantities and not as precise numbers.

In this paper, we deeply investigate a model to represent the available imprecise information about the "true" probability distribution of a random experiment, when the imprecise observation of the outcomes is modelled by a frv. It is a higher order model integrated in the Theory of Imprecise Probabilities [46]. We have presented some initial ideas about it in [10]. It is based on the "possibilistic" interpretation of fuzzy sets (see [9,19] for a further explanation). This interpretation is useful when we deal with imprecise statistical data (see, for instance, [5,37,40,42,39]). In those papers, fuzzy sets are used to represent the available imprecise information about the true "values"[1] of the outcomes of certain experiment. Thus, the membership of x to the fuzzy set $\tilde{A}$ represents the degree of possibility that the true value of the outcome which is modelled by $\tilde{A}$ coincides with x. Each fuzzy set can be identified with a possibility measure, which is a particular type of "upper probability" in the Theory of Imprecise Probabilities. Furthermore, fuzzy sets arise as the natural representation of the imprecise information provided by an expert, when it is expressed by means of confidence levels. The family of strong α-cuts of the fuzzy set plays an important role in this context, as we show in [9].

Our model is related to Kruse and Meyer's interpretation [28] of frv's. In that context, a frv is associated to an "acceptability function" defined over the class of all (classical) random variables. In this setting, we can define, in a natural way, an acceptability function over the class of all possible probability distributions. It represents the available

---

[1] The word "value" is here used in a very wide sense. It can be a real number, vector of numbers or, more generally, an element of the universe of possible outcomes.

information about the true probability distribution of the random experiment under study. From a theoretical point of view, this model is equivalent to a second order possibility distribution [16]: the information about the true probability is represented by a possibility measure defined over the class of all probability measures. A possibility measure represents the same information as a set of probability measures (the set of probability measures it dominates). Therefore, a second order possibility measure is associated to a set of probability measures, each one of them is defined, in turn, over a set of probability measures. Thus, a second order possibility will allow us to state assertions like "the probability that the true probability of the value 7 is 0.5 is between 0.4 and 0.7." Along the paper, we will show the connections with other previous studies in the literature. We can see that it is quite related to nested families of p-boxes considered by Ferson et al. [21]. We have independently investigated this last concept in [8,13].

The paper is organized as follows. Section 2 provides the necessary technical background. In Section 3, we analyse some interesting aspects about the acceptability function of a frv. We will refer to them when studying different imprecise probability models associated to frv's. Section 4 presents three different higher order possibility measures associated to a frv, all of which arise in a natural way. We will investigate the relationships among them and show that the first one is the more suitable representation of the available (imprecise) information about the "true" probability measure. In Section 5, we observe that the (classical) probability measure induced by a frv does not determine, in general, this second order possibility measure. We derive some interesting conclusions about the imprecise information provided by a frv. We end the paper with some general concluding remarks and open problems.

2. Preliminary concepts and notation

In this section, some definitions needed in the rest of the paper are recalled.

Consider a probability space $(\Omega, \mathcal{A}, P)$, and an arbitrary measurable space $(\Omega', \mathcal{A}')$. Given an $\mathcal{A} - \mathcal{A}'$ measurable function $f : \Omega \to \Omega'$, we will denote by $P \circ f^{-1}$ the probability measure it induces on $\mathcal{A}'$, i.e.,

$$P \circ f^{-1}(A) := P(f^{-1}(A)), \quad \forall A \in \mathcal{A}',$$

where $f^{-1} : \wp(\Omega') \to \wp(\Omega)$ is the mapping defined as

$$f^{-1}(A) = \{\omega \in \Omega: f(\omega) \in A\}, \quad \forall A \subseteq \Omega'.$$

Let us now consider a measurable space, $(\Omega, \mathcal{A})$, the usual Borel σ-algebra on $\mathbb{R}^n$, $\beta_{\mathbb{R}^n}$, and a multi-valued mapping $\Gamma : \Omega \to \wp(\mathbb{R}^n)$ with non-empty images. Let $A \in \beta_{\mathbb{R}^n}$ an arbitrary Borel set. The *upper inverse* of A is the set

$$\Gamma^*(A) = \{\omega \in \Omega \: : \: \Gamma(\omega) \cap A \neq \emptyset\}.$$

We say that Γ is *strongly measurable* [36] when $\Gamma^*(A) \in \mathcal{A}$, $\forall A \in \beta_{\mathbb{R}^n}$.

Given a probability space $(\Omega, \mathcal{A}, P)$, we will say that the multi-valued mapping $\Gamma : \Omega \to \mathbb{R}^n$ is a *random set* when it is strongly measurable. We will denote by $S(\Gamma)$ the class of *measurable selections* of Γ:

$$S(\Gamma) = \{X : \Omega \to \mathbb{R}^n \: : \: X \text{ measurable and } X(\omega) \in \Gamma(\omega), \ \forall \omega \in \Omega\}.$$

Let us suppose that Γ represents the imprecise measurements of the outcomes of a (standard) random variable X_0. In other words, all we know about each outcome $X_0(\omega)$ is that it belongs to the set $\Gamma(\omega)$. Then, all we know about X_0 is that it is a measurable selection of Γ.

We will denote by $P_\Gamma^* : \beta_{\mathbb{R}^n} \to [0, 1]$ the Dempster upper probability associated to Γ,

$$P_\Gamma^*(A) = P(\Gamma^*(A)) = P(\{\omega \in \Omega \: : \: \Gamma(\omega) \cap A \neq \emptyset\}), \quad \forall A \in \beta_{\mathbb{R}^n}.$$

We can easily check that, for each measurable selection $X \in S(\Gamma)$, the following inclusion relations hold:

$$X^{-1}(A) \subseteq \Gamma^*(A), \quad \forall A \in \beta_{\mathbb{R}^n}.$$

Thus, the upper probability dominates every probability measure induced by a measurable selection of Γ, i.e.,

$$P_X(A) = P \circ X^{-1}(A) \leqslant P(\Gamma^*(A)) = P_\Gamma^*(A), \quad \forall A \in \beta_{\mathbb{R}^n}, \ X \in S(\Gamma).$$

On the other hand, we can observe that the upper probability is determined by the probability measure induced by Γ, $P \circ \Gamma^{-1}$ (considered as a classical measurable function). In fact, let us consider, for each $A \in \beta_{\mathbb{R}^n}$ the family of sets:

$$\wp_A = \{C \subseteq \mathbb{R}^n \ : \ C \cap A \neq \emptyset\}.$$

Let us now consider the σ-algebra, $\sigma_\wp$, generated by the class $\{\wp_A \ : \ A \in \beta_{\mathbb{R}^n}\}$ on the universe of "elements" $\wp(\mathbb{R}^n)$. We easily observe that Γ is strongly measurable if and only if it is $\mathcal{A} - \sigma_\wp$ measurable (regarded as a classical function and not as a multi-valued one). On the other hand, we can easily check that the probability measure induced by Γ on $\sigma_\wp$ (the probability measure $P \circ \Gamma^{-1}$) determines P_Γ^*. In fact, the following equalities hold:

$$P \circ \Gamma^{-1}(\wp_A) = P(\Gamma^{-1}(\wp_A)) = P(\Gamma^*(A)) = P_\Gamma^*(A), \quad \forall A \in \beta_{\mathbb{R}^n}.$$

And the converse is also true: the upper probability univocally determines $P \circ \Gamma^{-1}$, since the class $\{\wp_A^c \ : \ A \in \beta_{\mathbb{R}^n}\}$ is closed for intersections.

For the measurable space $(\mathbb{R}^n, \beta_{\mathbb{R}^n})$, we will denote as $\mathcal{P}_{\beta_{\mathbb{R}^n}}$ the class of all probability measures that can be defined on $\beta_{\mathbb{R}^n}$. We will call the *probability envelope* of Γ to the class of probability measures:

$$\mathcal{P}_1(\Gamma) = \{P_X \ : \ X \in S(\Gamma)\}.$$

According to the above interpretation of Γ as the imprecise perception of a random variable X_0, we can say that all we know about the probability measure $P \circ X_0^{-1}$ is that it belongs to the class $\mathcal{P}_1(\Gamma)$.

For an arbitrary event $A \in \beta_{\mathbb{R}^n}$, we will call the *set-valued probability assignation* of A to the set of values:

$$\mathcal{P}_1(\Gamma)(A) = \{Q(A) \ : \ Q \in \mathcal{P}_1(\Gamma)\}, \quad \forall A \in \beta_{\mathbb{R}^n}.$$

Based on this definition, we will denote by $\mathcal{P}_2(\Gamma)$ the class of probability measures:

$$\mathcal{P}_2(\Gamma) = \{Q \text{ probability} \ : \ Q(A) \in \mathcal{P}_1(\Gamma)(A), \forall A \in \beta_{\mathbb{R}^n}\},$$

On the other hand, we will denote by $\mathcal{P}_3(\Gamma)$ the class of probability measures dominated by the upper probability of Γ:

$$\mathcal{P}_3(\Gamma) = \{Q \text{ probability} \ : \ Q(A) \leqslant P_\Gamma^*(A), \ \forall A \in \beta_{\mathbb{R}^n}\}.$$

It is easily checked that

$$\mathcal{P}_1(\Gamma) \subseteq \mathcal{P}_2(\Gamma) \subseteq \mathcal{P}_3(\Gamma).$$

Let us furthermore notice that P_Γ^*, $\mathcal{P}_3(\Gamma)$ and the probability measure induced by Γ on $\sigma_\wp$ uniquely determine one another. On the other hand, $\mathcal{P}_1(\Gamma)$ is strictly included in $\mathcal{P}_3(\Gamma)$ in many situations. Furthermore, we can find two different random sets with the same probability distribution, but with different probability envelopes. In Section 3 we show some illustrative examples. Some additional relationships among these three classes of probability measures are studied in [33–35].

A *fuzzy set* π is identified with a membership function from a universe U to the unit interval. The value $\pi(u)$ is the membership grade of element u in the fuzzy set. In this paper, a fuzzy set is interpreted as a *possibility distribution*. Then $\pi(u)$ is interpreted as the possibility grade that u coincides with some imprecisely known element $u_0 \in U$. Throughout the paper, we will use the notation π to denote a possibility distribution and the membership function of its associated fuzzy set. For any $\alpha \in [0, 1]$ we will denote by π_α the (weak) α-cut of π, i.e., $\pi_\alpha = \{u : \pi(u) \geqslant \alpha\}$. We will use the notation Π to denote the *possibility measure* associated to π. ($\Pi(C) = \sup\{\pi(u) \ : \ u \in C\}, \forall C \subseteq U$.) We will say that Π is *normal* when $\Pi(U) = 1$. We will denote by $\mathcal{F}(U)$ the class of all fuzzy subsets of U.

A *possibilistic probability* (or a "fuzzy probability") [14], $\tilde{P} : \beta_{\mathbb{R}^n} \to \mathcal{F}([0, 1])$, is a map taking each event [2] $A \in \beta_{\mathbb{R}^n}$ to a normal possibility distribution, $\tilde{P}(A)$ on $[0, 1]$. Its value $\tilde{P}(A)(p)$ in a point $p \in [0, 1]$ can be interpreted as the modeller's "upper betting rate" that the true probability $\tilde{P}(A)$ for the event A is equal to p.

---

[2] The concept of possibilistic probability is a particularization of that of "possibilistic prevision". A possibilistic prevision is defined on a general set of "gambles" instead of a set of events.

A possibilistic probability $\tilde{P}$ is called *representable* [14] if there is a (second order) [3] normal possibility distribution $\pi : \mathcal{P}_{\beta_{\mathbb{R}^n}} \to [0, 1]$ that represents $\tilde{P}$, i.e., such that for all $p \in [0, 1]$ and $A \in \beta_{\mathbb{R}^n}$, $\tilde{P}(A)(p) = \sup\{\pi(Q) : Q(A) = p\}$. A *graded set* [23] of U is a multi-valued mapping $\varphi : [0, 1] \to \wp(U)$ satisfying

$$\forall \alpha, \beta \in [0, 1], \quad [\alpha \leqslant \beta \Rightarrow \varphi(\alpha) \supseteq \varphi(\beta)].$$

For an arbitrary graded set φ, there exists a unique fuzzy set, $\pi_\varphi : U \to [0, 1]$ that satisfies

$$\pi_\varphi(u) = \sup\{\alpha : u \in \varphi(\alpha)\}, \quad \forall u \in \mathcal{U}.$$

Consider a measurable space, $(\Omega, \mathcal{A})$ and a fuzzy-valued mapping $\tilde{X} : \Omega \to \mathcal{F}(\mathbb{R}^n)$. For each $\alpha \in [0, 1]$, we will denote by $\tilde{X}_\alpha$ to the α-cut of $\tilde{X}$, i.e., the multi-valued mapping $\tilde{X}_\alpha : \Omega \to \wp(\mathbb{R}^n)$ that assigns, to each ω the α-cut of $\tilde{X}(\omega)$. We will say that $\tilde{X}$ is a *frv* when every α-cut, $\tilde{X}_\alpha$ is strongly measurable. This condition is equivalent to a classical measurability assumption: let us first consider, for each $A \in \beta_{\mathbb{R}^n}$, the family of sets:

$$\mathcal{F}_A^\alpha = \{F \in \mathcal{F}(\mathbb{R}^n) : F_\alpha \cap A \neq \emptyset\}.$$

Let us now denote by $\sigma_\mathcal{F}$ the σ-algebra generated by the class:

$$\{\mathcal{F}_A^\alpha, : A \in \beta_{\mathbb{R}^n}, \alpha \in [0, 1]\} \subseteq \wp(\mathcal{F}(\mathbb{R}^n)).$$

We easily check that a fuzzy-valued mapping is a frv if and only if it is $\mathcal{A} - \sigma_\mathcal{F}$ measurable. According to the notation established at the beginning of this section, we will denote by $P \circ \tilde{X}^{-1}$ the probability measure induced by $\tilde{X}$ on $\sigma_\mathcal{F}$.

3. Acceptability function of a frv

As we pointed in the introduction of the paper, we will consider the possibilistic interpretation of fuzzy sets. Furthermore, in this model, we will follow Kruse and Meyer [28] interpretation of frv's. According to it, the frv represents the imprecise observation of the *original* random variable, $X_0 : \Omega \to \mathbb{R}^n$. Thus, for each $\omega \in \Omega$ and each $x \in \mathbb{R}^n$, the quantity $\tilde{X}(\omega)(x)$ represents the possibility degree for x to be the "true" image of ω. (In other words, the possibility that x coincides with $X_0(\omega)$.) According to this information, Kruse and Meyer define the degree of acceptability of an arbitrary measurable function, $X : \Omega \to \mathbb{R}^n$, as the quantity

$$\mathrm{acc}_{\tilde{X}}(X) = \inf_{\omega \in \Omega} \tilde{X}(\omega)(X(\omega)).$$

The function "$\mathrm{acc}_{\tilde{X}}$" takes values in the unit interval, so it can be regarded as the possibility distribution associated to a possibility measure, $\Pi_{\tilde{X}}$, defined over the set of all $\mathcal{A} - \beta_{\mathbb{R}^n}$ measurable functions from Ω to $\mathbb{R}^n$, $\mathcal{X}_{\mathcal{A} - \beta_{\mathbb{R}^n}}$. For each $\omega \in \Omega$ and an arbitrary measurable mapping X, Kruse and Meyer interpret the membership value $\tilde{X}(\omega)(X(\omega))$ as the acceptability of the proposition "$X(\omega)$ coincides with $X_0(\omega)$". Furthermore, the infimum $\inf\{\tilde{X}(\omega)(X(\omega)) : \omega \in \Omega\}$ is intended as the acceptability of "$X(\omega) = X_0(\omega), \forall \omega \in \Omega$" or, in other words "$X$ coincides X_0". Hence the quantity $\mathrm{acc}_{\tilde{X}}(X)$ is interpreted by the authors as the degree of acceptability of the proposition "X is the original variable". Next we give a further justification for the use of this infimum.

The following result is an immediate consequence of the "possibilistic extension theorem" proved in [4] and also mentioned in [15]. We will derive from it some interesting properties about the acceptability function.

Lemma 1. *Let U be a non-empty set and let $\mathcal{C}$ be a non-empty collection of subsets of U. Consider the set function $f : \mathcal{C} \to [0, 1]$. Then, the highest (least informative) possibility measure, $\Pi : \wp(U) \to [0, 1]$, which is dominated*

---

[3] The term "second order" reflects that this possibility distribution is defined over a set of probability measures. In the general setting, the initial space is the class of linear previsions (including σ-additive probability as particular cases).

by f is the one associated to the possibility distribution $\pi : U \to [0, 1]$ defined as

$$\pi(x) = \begin{cases} \inf_{C \in \mathcal{C}_x} f(C) & \text{if } \mathcal{C}_x \neq \emptyset, \\ 1 & \text{if } \mathcal{C}_x = \emptyset, \end{cases}$$

where $\mathcal{C}_x = \{C \in \mathcal{C} \, : \, x \in C\}$. If, furthermore, there exists a possibility measure, $\Pi' : \wp(U) \to [0, 1]$ such that $\Pi'(C) = f(C), \, \forall C \in \mathcal{C}$, then Π coincides with it on $\mathcal{C}$.

Let us now consider the class of all possibility measures on $\wp(\mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}})$ that satisfy the following inequalities:

$$\Pi(\{X \in \mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}} \, : \, X(\omega) = x\}) \leqslant \tilde{X}(\omega)(x), \quad \forall x \in \mathbb{R}^n, \ \omega \in \Omega. \tag{1}$$

According to last lemma, the acceptability function given by Kruse and Meyer is the possibility distribution that generates the highest possibility measure in this class. Thus, the acceptability function is associated to the highest (least informative) possibility measure, $\Pi_{\tilde{X}} : \wp(\mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}}) \to [0, 1]$, which is compatible with the available information about the images of X_0. The possibility distribution, $\mathrm{acc}_{\tilde{X}}$ is defined over the class of measurable functions from Ω to $\mathbb{R}^n$, $\mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}}$, and it represents the information the frv provides about the "original" random variable. For an arbitrary class of measurable mappings, $\mathcal{X} \subseteq \mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}}$, the quantity

$$\Pi_{\tilde{X}}(\mathcal{X}) = \sup_{X \in \mathcal{X}} \inf_{\omega \in \Omega} \tilde{X}(\omega)(X(\omega))$$

represents the degree of possibility that X_0 belongs to $\mathcal{X}$.

Remark 1. When the frv is, in particular, a random set, $\Gamma : \Omega \to \wp(\mathbb{R}^n)$, the acceptability function only takes the extreme values 0 and 1. In other words, it divides the class of measurable functions $\mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}}$ into two disjoint classes, $\mathcal{X}$ and $\mathcal{X}^c$, where $\mathcal{X}$ is, in fact, the class of measurable selections of Γ, i.e.,

$$\begin{aligned} \mathcal{X} &= \{X \in \mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}} \, : \, \mathrm{acc}_{\Gamma}(X) = 1\} \\ &= \{X \in \mathcal{X}_{\mathcal{A}-\beta_{\mathbb{R}^n}} \, : \, X(\omega) \in \Gamma(\omega), \quad \forall \omega \in \Omega\} = S(\Gamma). \end{aligned}$$

Hence, our information about the original random variable can be described as follows:
For each ω, all we know about $X_0(\omega)$ is that it belongs to the (crisp) set $\Gamma(\omega)$. In other words, all we know about the original random vector $X_0 : \Omega \to \mathbb{R}^n$ is that it belongs to the class $S(\Gamma)$.

Now we further explain the acceptability function in terms of confidence levels. First, we give two lemmas.

Lemma 2. *The nested family of sets $\{S(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ is a gradual representation of $\mathrm{acc}_{\tilde{X}}(\tilde{X})$.*

Proof. Consider an arbitrary measurable mapping $X : \Omega \to \mathbb{R}^n$. Then

$$\begin{aligned} \mathrm{acc}_{\tilde{X}}(X) &= \inf_{\omega \in \Omega} \tilde{X}(\omega)(X(\omega)) = \inf_{\omega \in \Omega} \sup\{\alpha \in [0, 1] \, : \, X(\omega) \in \tilde{X}_\alpha(\omega)\} \\ &= \sup\{\alpha \in [0, 1] \, : \, X(\omega) \in \tilde{X}_\alpha(\omega), \ \forall \omega \in \Omega\} = \sup\{\alpha \in [0, 1] \, : \, X \in S(\tilde{X}_\alpha)\}. \quad \square \end{aligned}$$

This representation of $\mathrm{acc}_{\tilde{X}}$ as the fuzzy set associated to the nested family $\{S(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ is quite related to the interpretation of fuzzy sets as families of confidence sets, which is studied in detail in [9]. Let us briefly explain these ideas. First, we can easily check the following result:

Lemma 3. *Let us consider a nested family of subsets, $\{A_\alpha\}_{\alpha \in [0,1]}$, of a universe U. The class of probability measures that satisfy the inequalities*

$$P(A_\alpha) \geqslant 1 - \alpha, \quad \forall \alpha \in [0, 1]$$

coincides with the class of probability measures dominated by the possibility measure Π:

$$\Pi(C) = \sup\{\alpha \in [0, 1] : C \cap A_\alpha \neq \emptyset\}, \quad \forall C \subseteq U.$$

Based on this result, we can give $\text{acc}_{\tilde{X}}$ a "confidence degree" interpretation. Let us suppose that an expert provides the following information about the images of $X_0 : \Omega \to \mathbb{R}$:

For each $\alpha \in [0, 1]$, $X_0(\omega)$ belongs to $[\tilde{X}(\omega)]_\alpha$, $\forall \omega \in \Omega$, with probability greater than or equal to $1 - \alpha$.
In other words, the expert asserts that X_0 is a selection of $\tilde{X}_\alpha$ with probability greater than or equal to $1 - \alpha$.

According to Lemma 3, we observe that $\text{acc}_{\tilde{X}}$ represents this information. I.e., the class of probability measures (defined on a class of $\mathcal{A} - \beta_{\mathbb{R}^n}$ measurable functions [4]) that satisfy these inequalities coincides with the class of probability measures that are dominated by $\Pi_{\tilde{X}}$ (the possibility measure associated to $\text{acc}_{\tilde{X}}$).

4. Second order possibility measures associated to a frv

Based on this interpretation of a frv as an imprecise observation of the "original" variable, we will define three different possibility measures over the set of al probability measures on $\beta_{\mathbb{R}^n}$. We will denote them by $\Pi\Pi_{\tilde{X}}^1$, $\Pi\Pi_{\tilde{X}}^2$ and $\Pi\Pi_{\tilde{X}}^3$, respectively. Under the above interpretation of frv's, the first one of them will represent the available (imprecise) information about the *original* probability measure. [5] The other two possibility measures will also arise in a natural way, but they are less informative. In particular, $\Pi\Pi_{\tilde{X}}^2$ will be associated to a fuzzy-valued "probability" defined on $\beta_{\mathbb{R}^n}$. On the other hand, $\Pi\Pi_{\tilde{X}}^3$ will be related to the family of pairs of Dempster upper and lower probabilities associated to the α-cuts of $\tilde{X}$. The procedure described below is briefly outlined in [10].

Definition 1. Let us consider a probability space $(\Omega, \mathcal{A}, P)$, and a frv $\tilde{X} : \Omega \to \mathbb{R}^n$. Let us consider the class $\mathcal{P}_{\beta_{\mathbb{R}^n}}$ of all (σ-additive) probability measures over $\beta_{\mathbb{R}^n}$. The *fuzzy probability envelope associated to* $\tilde{X}$, $\pi_{\tilde{X}}^1 : \mathcal{P}_{\beta_{\mathbb{R}^n}} \to [0, 1]$, is defined as

$$\pi_{\tilde{X}}^1(Q) = \sup\{\text{acc}_{\tilde{X}}(X) : P_X = Q\}, \quad \forall Q \in \mathcal{P}_{\beta_{\mathbb{R}^n}}.$$

Under the last conditions, we will denote by $\Pi\Pi_{\tilde{X}}^1$ the possibility measure associated to the fuzzy probability envelope:

$$\Pi\Pi_{\tilde{X}}^1(\mathcal{P}) = \sup\{\text{acc}_{\tilde{X}}(X) : P_X \in \mathcal{P}\}, \quad \forall \mathcal{P} \subseteq \mathcal{P}_{\beta_{\mathbb{R}^n}}.$$

It represents the imprecise information that the frv provides about the original probability measure. For an arbitrary class of probability measures, $\Pi\Pi_{\tilde{X}}^1(\mathcal{P})$ represents the degree of possibility that the original probability belongs to the class $\mathcal{P}$, under the available information: let us observe that it can be written as $\Pi_{\tilde{X}}(\mathcal{X}(\mathcal{P}))$, where $\mathcal{X}(\mathcal{P})$ denotes the class of random vectors whose associated probability measure belongs to $\mathcal{P}$.

$\Pi\Pi_{\tilde{X}}^1$ is a second order possibility measure in the sense of [14,16], as it is a possibility measure defined on a class of linear previsions. In our particular model, only σ-additive probabilities have non-zero values of possibility.

Furthermore, the concept of fuzzy probability envelope is closely related to the definition of probability envelope recalled in Section 2. In fact, the (nested) family of sets of probability measures $\{\mathcal{P}_1(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ is a gradual representation of $\pi_{\tilde{X}}^1$ as we will prove next.

Proposition 1. $\{\mathcal{P}_1(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ *is a gradual representation of* $\pi_{\tilde{X}}^1$.

Proof. Let us take an arbitrary probability measure, Q. By Lemma 2, we know that $\pi_{\tilde{X}}^1(Q)$ coincides with the quantity:

$$\sup\{\alpha : \exists X \in S(\tilde{X}_\alpha) \text{ s.t. } Q = P_X\}.$$

---

[4] We can consider the σ-algebra generated by the nested family of classes $\{S(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$.

[5] From now on, we will call this way the probability measure associated to the original random variable, $P_{X_0} = P \circ X_0^{-1}$.

On the other hand, by definition of $\{\mathcal{P}_1(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ the class $\{\alpha \ : \ \exists \ X \in S(\tilde{X}_\alpha) \text{ s.t. } Q = P_X\}$ coincides with $\{\alpha \ : \ Q \in \mathcal{P}_1(\tilde{X}_\alpha)\}$. Hence, we deduce that $\pi_{\tilde{X}}^1(Q)$ coincides with $\sup\{\alpha \ : \ Q \in \mathcal{P}_1(\tilde{X}_\alpha)\}$. In words, $\{\mathcal{P}_1(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ is a gradual representation of $\pi_{\tilde{X}}^1$. $\square$

We deduce from this result that the concept of "fuzzy probability envelope" is closely related the concept of "expectation of a frv" introduced by Puri and Ralescu in [38]. In fact, the fuzzy expectation of a frv [38] is defined as the fuzzy set associated to the nested family $\{E(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$, where $E(\tilde{X}_\alpha) = \{\int X \, dP \ : \ X \in S(\tilde{X}_\alpha)\}$ represents the Aumann expectation of $\tilde{X}_\alpha$.

Furthermore, this representation of $\pi_{\tilde{X}}^1$ as the fuzzy set associated to the nested family $\{\mathcal{P}_1(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ is quite related to the interpretation of fuzzy sets as families of confidence sets [9] outlined in Section 3.

According to Lemma 3, we can give $\pi_{\tilde{X}}^1$ a "confidence degree" interpretation. Let us suppose that an expert provides the following information about the images of $X_0 : \Omega \to \mathbb{R}$:

For each $\alpha \in [0, 1]$, $X_0(\omega)$ belongs to $[\tilde{X}(\omega)]_\alpha$, $\forall \omega \in \Omega$, with probability greater than or equal to $1 - \alpha$.

In other words, the expert asserts that X_0 is a selection of $\tilde{X}_\alpha$ with probability greater than or equal to $1 - \alpha$. Under these conditions, we can represent the available (imprecise) information about the original probability measure as follows:

P_{X_0} belongs to $\mathcal{P}_1(\tilde{X}_\alpha)$ with probability greater than or equal to $1 - \alpha$, $\forall \alpha \in [0, 1]$. In other words, the available information about P_{X_0} is represented by the class of second order probability measures $\mathbb{P}$ such that:

$$\mathbb{P}(\mathcal{P}_1(\tilde{X}_\alpha)) \geqslant 1 - \alpha, \quad \forall \alpha \in [0, 1]. \tag{2}$$

According to Lemma 3, we observe that $\pi_{\tilde{X}}^1$ determines the same information as the set of probability measures satisfying Eq. (2). I.e., the class of second order probabilities that satisfy these inequalities coincides with the class of second order probabilities that are dominated by $\Pi_{\tilde{X}}^1$ (the possibility measure associated to $\pi_{\tilde{X}}^1$).

Based on the interpretation of $acc_{\tilde{X}}$ and $\Pi_{\tilde{X}}$, we can also define in a natural way the "fuzzy-valued probability assignment" of $\tilde{X}$ as a mapping, $\tilde{P}_{\tilde{X}}$, that assigns a fuzzy subset of the unit interval to each particular event $A \in \beta_{\mathbb{R}^n}$.

Definition 2. Let us consider a probability space $(\Omega, \mathcal{A}, P)$, and a frv $\tilde{X} : \Omega \to \mathbb{R}^n$. The *fuzzy probability assignment* [6] *associated to* $\tilde{X}$, $\tilde{P}_{\tilde{X}} : \beta_{\mathbb{R}^n} \to \mathcal{F}([0, 1])$, will be defined as

$$\tilde{P}_{\tilde{X}}(A)(p) = \sup\{acc_{\tilde{X}}(X) \ : \ P_X(A) = p\}, \quad \forall p \in [0, 1], \ \ \forall A \in \beta_{\mathbb{R}^n}.$$

For an arbitrary event $A \in \beta_{\mathbb{R}^n}$, the fuzzy value $\tilde{P}_{\tilde{X}}(A)$ represents, in this framework, the imprecise information the frv provides about the true probability of A, $P_{X_0}(A)$. Specifically, $\tilde{P}_{\tilde{X}}(A)(p)$ represents the degree of possibility that $P_{X_0}(A)$ coincides with p. From a formal point of view, $\tilde{P}_{\tilde{X}}$ is a possibilistic probability [14,16], so we can ask ourselves whether it is "representable" and which is its highest (least informative second order possibility measure) representation. We can easily check that $\tilde{P}_{\tilde{X}}(A)$ can be expressed as

$$\tilde{P}_{\tilde{X}}(A)(p) = \sup\{\pi_{\tilde{X}}^1(Q) \ : \ Q \text{ prob. with } Q(A) = p\}.$$

In other words, $\tilde{P}_{\tilde{X}}$ is, in fact, representable, and $\Pi_{\tilde{X}}^1$ is a representation of it. But it is not, in general, its highest representation: based on the formulae given in [14], we observe that the highest representation of $\tilde{P}_{\tilde{X}}$ is the second order possibility distribution $\pi_{\tilde{X}}^2$, defined as

$$\pi_{\tilde{X}}^2(Q) = \inf\{\tilde{P}_{\tilde{X}}(A)(Q(A)) \ : \ A \in \beta_{\mathbb{R}^n}\}, \quad \forall Q \in \mathcal{P}_{\beta_{\mathbb{R}^n}}.$$

Next we check that the (nested) family of classes of probability measures $\{\mathcal{P}_2(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ is a gradual set associated to the fuzzy set $\pi_{\tilde{X}}^2$.

---

[6] The term "fuzzy-valued probability" is used here in a very wide sense. In fact, $\tilde{P}_{\tilde{X}}$ does not satisfy in general the additivity property (the main property of probability measures).

Proposition 2. $\{\mathcal{P}_2(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$ *is a gradual representation of* $\pi^2_{\tilde{X}}$.

Proof. First of all, let us consider, for each α and each $A \in \beta_{\mathbb{R}^n}$, the set of values $\mathcal{P}_1(\tilde{X}_\alpha)(A) := \{Q(A) \: : \: Q \in \mathcal{P}_1(\tilde{X}_\alpha)\}$. It is easily checked that the family $\{\mathcal{P}_1(\tilde{X}_\alpha)(A)\}_{\alpha \in [0,1]}$ is a gradual representation of $P_{\tilde{X}}(A)$, for each $A \in \beta_{\mathbb{R}^n}$. In other words, for any $A \in \beta_{\mathbb{R}^n}$ and Q, the following equality holds:

$$\tilde{P}_{\tilde{X}}(A)(Q(A)) = \sup\{\alpha \in [0,1] \: : \: Q(A) \in \mathcal{P}_1(\tilde{X}_\alpha)(A)\}.$$

On the other hand, we observe that the quantity

$$\inf\{\sup\{\alpha \in [0,1] \: : \: Q(A) \in \mathcal{P}_1(\tilde{X}_\alpha)(A)\} \: : \: A \in \beta_{\mathbb{R}^n}\}$$

can be written as

$$\sup\{\alpha \in [0,1] \: : \: Q(A) \in \mathcal{P}_1(\tilde{X}_\alpha)(A), \forall A \in \beta_{\mathbb{R}^n}\}.$$

We deduce that

$$\pi^2_{\tilde{X}}(Q) = \sup\{\alpha \in [0,1] \: : \: Q(A) \in \mathcal{P}_1(\tilde{X}_\alpha)(A), \quad \forall A \in \beta_{\mathbb{R}^n}\}$$
$$= \sup\{\alpha \in [0,1] \: : \: Q \in \mathcal{P}_2(\tilde{X}_\alpha)\}. \quad \square$$

Let us now compare the information provided by $\pi^1_{\tilde{X}}$ and the information provided by $\tilde{P}_{\tilde{X}}$. In other words, let us compare $\pi^1_{\tilde{X}}$ and $\pi^2_{\tilde{X}}$. On the one hand, as we checked above, the fuzzy probability assignation associated to $\tilde{X}$ is representable [14] by the (normal) possibility distribution $\pi^1_{\tilde{X}}$. On the other hand, its greatest representation is $\pi^2_{\tilde{X}}$. Then we derive that

$$\pi^1_{\tilde{X}}(Q) \leqslant \pi^2_{\tilde{X}}(Q), \quad \forall Q \in \mathcal{P}_{\mathbb{R}^n}.$$

In other words,

$$\Pi^1_{\tilde{X}}(\mathcal{P}) \leqslant \Pi^2_{\tilde{X}}(\mathcal{P}), \quad \forall \mathcal{P} \subseteq \mathcal{P}_{\mathbb{R}^n}. \tag{3}$$

Let us now remind that any possibility measure is equivalent to the class of probability measures it dominates. Hence, the information provided by $\Pi^1_{\tilde{X}}$ and $\Pi^2_{\tilde{X}}$ can be represented by individual classes of second order probability measures. Based on Eq. (3), the class of second order probability measures associated to $\Pi^1_{\tilde{X}}$ is included in the class (of second order probabilities) associated to $\Pi^2_{\tilde{X}}$. This means that the possibility distribution $\pi^1_{\tilde{X}}$ (or, equivalently, the possibility measure $\Pi^1_{\tilde{X}}$) provides at least as much information about the "original" probability measure as the fuzzy probability assignation does. The second order possibility distributions $\pi^1_{\tilde{X}}$ and $\pi^2_{\tilde{X}}$ do not coincide in general, as we will show in Example 2.

We will deal now with the third and last second order possibility distribution associated to $\tilde{X}$, $\pi^3_{\tilde{X}}$.

Definition 3. Let us consider a probability space $(\Omega, \mathcal{A}, P)$, and a frv $\tilde{X} : \Omega \to \mathbb{R}^n$. Let us consider the class $\mathcal{P}_{\beta_{\mathbb{R}^n}}$ of all (σ-additive) probability measures over $\beta_{\mathbb{R}^n}$. The *convex fuzzy probability envelope* is defined as the possibility distribution associated to the gradual set $\{\mathcal{P}_3(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$. In other words, $\pi^3_{\tilde{X}}$ is defined as

$$\pi^3_{\tilde{X}}(Q) = \sup\{\alpha \in [0,1] \: : \: Q(A) \leqslant P^*_{\tilde{X}_\alpha}(A), \forall A \in \beta_{\mathbb{R}^n}\}.$$

According to Propositions 1 and 2, the second order possibility measure $\pi^3_{\tilde{X}}$ dominates both $\pi^1_{\tilde{X}}$ and $\pi^2_{\tilde{X}}$, i.e.,

$$\pi^1_{\tilde{X}}(Q) \leqslant \pi^2_{\tilde{X}}(Q) \leqslant \pi^3_{\tilde{X}}(Q), \quad \forall Q \in \mathcal{P}(\beta_{\mathbb{R}^n}). \tag{4}$$

less information →

Fuzzy R.V.	$\pi^1_{\tilde{X}}$	$\pi^2_{\tilde{X}}$	$\pi^3_{\tilde{X}}$
Random Set	$\mathcal{P}_1(\Gamma)$	$\mathcal{P}_2(\Gamma)$	$\mathcal{P}_3(\Gamma)$

Fig. 1. Dominance relations between different second order possibility distributions.

In Fig. 1 we illustrate the dominance relations among these second order possibility distributions. We recall the notation we use for each one of them when $\tilde{X}$ is, in particular, a random set.

5. Relationships between the second order possibility measures and the probability measure induced by a frv

As we have stated at the beginning of this section, the frv represents here the imprecise observation of a (classical) random variable $X_0 : \Omega \to \mathbb{R}^n$. Each image $\tilde{X}(\omega)$ is a possibility distribution on $\mathbb{R}^n$ that represents the available information about the point $X_0(\omega)$. In this setting, the second order possibility measure $\pi^1_{\tilde{X}}$ arises in a natural way to represent the available information about P_{X_0}. Our main purpose in this section is to show that the "fuzzy probability envelope" associated to a frv is a good representation of this information, but the probability measure induced by $\tilde{X}$ on $\sigma_{\mathcal{F}}$ is not. In fact, we can find two different frv's inducing the same (classical) probability measure but with different "fuzzy probability envelopes".

In the following examples we will compare the information provided by $\pi^1_{\tilde{X}}$, $\pi^2_{\tilde{X}}$, $\pi^3_{\tilde{X}}$ and $P \circ \tilde{X}^{-1}$ about different aspects of P_{X_0}. In particular, we will deal with Shannon entropy (Examples 1 and 2), the variance (Examples 3 and 4) and the probability measure associated to a pair of independent identically distributed random variables (Examples 5 and 6). In each case, we will first deal with the particular case of random sets (Examples 1, 3 and 5), and then, we will derive some conclusions about the general case of frv's (Examples 2, 4 and 6).

So, let us now study the particular case of random sets (frv's whose images are crisp sets) and illustrate the differences among $\mathcal{P}_1(\Gamma)$, $\mathcal{P}_2(\Gamma)$ and $\mathcal{P}_3(\Gamma)$. We will then return to the general case (frv), and obtain some conclusions about the differences among $\pi^1_{\tilde{X}}$, $\pi^2_{\tilde{X}}$ and $\pi^3_{\tilde{X}}$.

Let us consider an arbitrary random set $\Gamma : \Omega \to \wp(\mathbb{R}^n)$. At first sight it seems that the differences among $\mathcal{P}_1(\Gamma)$, $\mathcal{P}_2(\Gamma)$ and $\mathcal{P}_3(\Gamma)$ are not relevant. In fact, as we pointed in Section 2, the following relations hold:

$$\mathcal{P}_1(\Gamma) \subseteq \mathcal{P}_2(\Gamma) \subseteq \mathcal{P}_3(\Gamma).$$

Furthermore, the information provided by $\mathcal{P}_3(\Gamma)$ is equivalent to the information provided by P^*_Γ. In most practical situations, the upper probability $P^*_\Gamma(A)$ coincides with the supremum:

$$\sup\{P_X(A) : X \in S(\Gamma)\} = \sup\{Q(A) : Q \in \mathcal{P}_1(\Gamma)\} = \sup\mathcal{P}_1(\Gamma)(A).$$

(See [33–35] for a detailed study.) Nevertheless, the classes of probability measures $\mathcal{P}_1(\Gamma)$, $\mathcal{P}_2(\Gamma)$ and $\mathcal{P}_3(\Gamma)$ may be quite different, even in the situations where both quantities $P^*_\Gamma(A)$ and $\sup\mathcal{P}_1(\Gamma)(A)$ do coincide. These differences are reflected on the information about parameters as the entropy or the variance, as we will, respectively, show in Examples 1 and 3.

Example 1. Let Ω be a universe with only two elements, $\Omega = \{\omega_1, \omega_2\}$. Let us consider the probability measure $P : \wp(\Omega) \to [0, 1]$ such that $P(\{\omega_1\}) = 2/3$ (and $P(\{\omega_2\}) = 1/3$). Let us define the random set $\Gamma : \Omega \to \wp(\mathbb{R})$ as follows:

$$\Gamma(\omega_1) = \{0, k\} \quad \text{and} \quad \Gamma(\omega_2) = \{-k, 0, k\}.$$
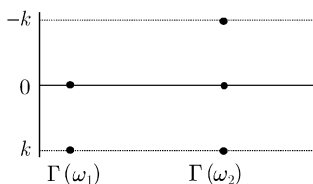
Let us identify each probability measure, Q, on the final space with the triplet $(Q(\{-k\}), Q(\{0\}), Q(\{k\}))$. Let us first calculate the expressions for $\mathcal{P}_1(\Gamma)$, $\mathcal{P}_2(\Gamma)$ and $\mathcal{P}_3(\Gamma)$.

- Calculation of $\mathcal{P}_1(\Gamma)$:

$$\mathcal{P}_1(\Gamma) = \{(1/3, 2/3, 0), (0, 1, 0), (0, 2/3, 1/3), (1/3, 0, 2/3),$$
$$(0, 1/3, 2/3), (0, 0, 1)\}.$$

- Calculation of $\mathcal{P}_2(\Gamma)$:
  For the calculation of $\mathcal{P}_2(\Gamma)$, we previously need to calculate $\mathcal{P}_1(\Gamma)(A)$, for each $A \subseteq \{-k, 0, k\}$.

$$\mathcal{P}_1(\Gamma)(\{0\}) = \mathcal{P}_1(\Gamma)(\{k\}) = \{0, 1/3, 2/3, 1\},$$
$$\mathcal{P}_1(\Gamma)(\{-k\}) = \{0, 1/3\},$$
$$\mathcal{P}_1(\Gamma)(\{-k, 0\}) = \{0, 1/3, 2/3, 1\},$$
$$\mathcal{P}_1(\Gamma)(\{-k, k\}) = \{0, 1/3, 2/3, 1\},$$
$$\mathcal{P}_1(\Gamma)(\{0, k\}) = \{2/3, 1\} \quad \text{and}$$
$$\mathcal{P}_1(\Gamma)(\{-k, 0, k\}) = \{1\}.$$

Hence, we deduce that

$$\mathcal{P}_2(\Gamma) = \mathcal{P}_1(\Gamma) \cup \{(1/3, 1/3, 1/3)\}.$$

- Calculation of $\mathcal{P}_3(\Gamma)$:
  For the calculation of $\mathcal{P}_3(\Gamma)$, we previously need to calculate the expression of P_Γ^*. We have

$$P_\Gamma^*(A) = P(\{\omega \in \Omega : \Gamma(\omega) \cap A \neq \emptyset\})$$
$$= \begin{cases} 1 & \text{if } A \cap \{0, k\} \neq \emptyset, \\ 1/3 & \text{if } A \cap \{0, k\} = \emptyset \quad \text{and } -k \in A, \\ 0 & \text{if } A = \emptyset. \end{cases}$$

Hence, we observe that $\mathcal{P}_3(\Gamma) = \{(p_1, p_2, p_3) : p_1 \leqslant 1/3\}$. In fact, $\mathcal{P}_3(\Gamma)$ is the convex hull of $\mathcal{P}_1(\Gamma)$. I.e.,

$$\mathcal{P}_1(\Gamma) \subsetneq \mathcal{P}_2(\Gamma) \subsetneq \mathcal{P}_3(\Gamma) = Conv(\mathcal{P}_1(\Gamma)).$$

Thus, if we need to calculate the closest upper and lower bounds for the true probability of an event, the differences among $\mathcal{P}_1(\Gamma)$, $\mathcal{P}_2(\Gamma)$ and $\mathcal{P}_3(\Gamma)$ are not relevant. Nevertheless, when we replace $\mathcal{P}_1(\Gamma)$ by $\mathcal{P}_2(\Gamma)$ or $\mathcal{P}_3(\Gamma)$ we loose important information about other properties of the original probability, P_{X_0}. For instance, the differences between $\mathcal{P}_1(\Gamma)$ and $\mathcal{P}_2(\Gamma)$ have effects, in the calculus of Shannon's entropy. $\mathcal{P}_1(\Gamma)$ provides much more information than $\mathcal{P}_2(\Gamma)$ about the entropy of the original probability measure, as we show below.

For an arbitrary class of probability measures, let us denote as $H(\mathcal{P})$ to the set of values of entropy:

$$H(\mathcal{P}) = \{H(Q) : Q \in \mathcal{P}\}.$$

We easily check that

$$H(\mathcal{P}_1(\Gamma)) = \{0, \log_2 3 - 2/3\},$$

$$H(\mathcal{P}_2(\Gamma)) = \{0, \log_2 3 - 2/3, \log_2 3\} \quad \text{and}$$

$$H(\mathcal{P}_3(\Gamma)) = [0, \log_2 3].$$

Let us summarize the conclusions we derive from this example. We have considered a random experiment whose possible results are ω_1 and ω_2, with respective probabilities $2/3$ and $1/3$. The random variable $X_0 : \Omega \to \mathbb{R}$ represents certain characteristic of these two elements. But it is imprecisely observed. In fact, all we know about $X_0(\omega_i)$ is that it belongs to the set of values $\Gamma(\omega_i)$, $i = 1, 2$. Hence, we know that P_{X_0} is dominated by P_Γ^*, i.e., $P_{X_0} \in \mathcal{P}_3(\Gamma)$. We derive that the uncertainty about the image of X_0 (the entropy of X_0) is less than $\log_2 3$ bit. But let us go further on.

We do not only know that P_{X_0} is dominated by P_Γ^*, but we also know that it belongs to $\mathcal{P}_1(\Gamma)$. Based on this (more precise) information, we derive that the entropy of X_0 is not greater than $(\log_2 3 - 2/3)$ bit.

We conclude that the information provided by $\mathcal{P}_1(\Gamma)$ about P_{X_0} is, in general, more precise than the information determined by P_Γ^*. We have shown here that these differences are reflected in the information about the entropy. Notice that P_Γ^* determines the probability distribution of Γ (regarded as a $\wp(\Omega) - \wp(\wp(\{-k, 0, k\}))$ measurable function). Hence we derive that the probability distribution of Γ, $P \circ \Gamma^{-1}$, does not provide all the relevant information when Γ represents the imprecise observation of a random variable.

Remark 2. We can slightly modify the conditions in last example and define another random set with the same probability distribution as Γ, but different probability envelope. Let us consider the initial space $([0, 1], \beta_{[0,1]}, \lambda_{[0,1]})$ (the unit interval, with the usual Borel σ-algebra and the uniform probability distribution). Let us now consider the random set $\Gamma' : [0, 1] \to \wp(\mathbb{R})$, defined as

$$\Gamma'(\omega) = \begin{cases} \Gamma(\omega_1) = \{0, k\} & \text{if } \omega \leqslant 2/3, \\ \Gamma(\omega_2) = \{-k, 0, k\} & \text{if } \omega > 2/3. \end{cases}$$

We easily check that $\mathcal{P}_1(\Gamma') = \mathcal{P}_3(\Gamma') = \mathcal{P}_3(\Gamma) \neq \mathcal{P}_1(\Gamma)$. We derive that the upper probability distributions of both random sets coincide ($P_{\Gamma'}^* = P_\Gamma^*$). Hence, they induce the same probability distribution on $\sigma_\wp$, i.e., $P \circ \Gamma^{-1} = P \circ \Gamma'^{-1}$. We conclude that, in general, the probability distribution induced by a random set does not determine its probability envelope.

Remark 3. The formulation of $H(\mathcal{P}_1(\Gamma))$ as the class of entropy values associated to the measurable selections of Γ can be extended to other types of parameters (variance, expectation, etc.) Aumann expectation [1], Kruse variance [27], and the imprecise distribution function studied in [13] fit this formulation. This procedure can be generalized to frv's. For a particular parameter $\theta(P_{X_0})$, the fuzzy set $\theta(\pi_{\tilde{X}}^1)$ defined as

$$\theta(\pi_{\tilde{X}}^1)(x) = \sup\{\text{acc}_{\tilde{X}}(X) : \theta(P_X) = x\} = \sup\{\pi_{\tilde{X}}^1(Q) : \theta(Q) = x\}$$

represents the available imprecise information about $\theta(P_{X_0})$. Puri and Ralescu expectation [38], second order variance [6,11] and the fuzzy probability assignment are compatible with this definition. Under the above interpretation of frv's, $\theta(P_X)(x)$ represents the possibility grade that x coincides with $\theta(P_{X_0})$. Next we will slightly modify the last example and consider a frv instead of a random set. We will study the differences among $\pi_{\tilde{X}}^1$, $\pi_{\tilde{X}}^2$ and $\pi_{\tilde{X}}^3$ about the entropy of the original probability $H(P_{X_0})$. We will compare the fuzzy sets $H(\pi_{\tilde{X}}^1)$, $H(\pi_{\tilde{X}}^2)$ and $H(\pi_{\tilde{X}}^3)$, each one defined as

$$H(\pi_{\tilde{X}}^i)(x) = \sup\{\pi_{\tilde{X}}^i(Q) : H(Q) = x\}, \quad \forall x \in \mathbb{R}^+, \quad i = 1, 2, 3.$$

Example 2. Let Ω be the same universe as in Example 1. Let us define the frv $\tilde{X} : \Omega \to \mathcal{F}(\mathbb{R})$ as follows:

$$\tilde{X}(\omega_1)(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0.5 & \text{if } x = k, \\ 0 & \text{otherwise,} \end{cases} \qquad \tilde{X}(\omega_2)(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0.5 & \text{if } x \in \{-k, k\}, \\ 0 & \text{otherwise.} \end{cases}$$
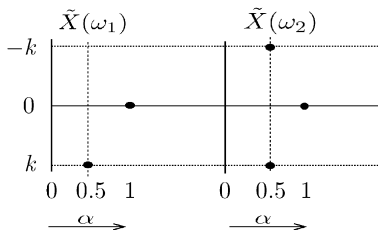
We observe that the α-cuts of $\tilde{X}$, $\tilde{X}_\alpha : \Omega \to \wp(\mathbb{R})$ are: $\tilde{X}_\alpha = \Gamma$, $\forall \alpha \in (0, 0.5]$ and $\tilde{X}_\alpha(\omega_i) = \{0\}$, $i = 1, 2$, $\forall \alpha \in (0.5, 1]$.

Hence, the classes of probability measures $\mathcal{P}_1(\tilde{X}_\alpha)$, $\mathcal{P}_2(\tilde{X}_\alpha)$ and $\mathcal{P}_3(\tilde{X}_\alpha)$ are

$$\mathcal{P}_1(\tilde{X}_\alpha) = \mathcal{P}_i(\Gamma), \quad i = 1, 2, 3, \ \alpha \in (0, 0.5],$$

$$\mathcal{P}_1(\tilde{X}_\alpha) = \{(0, 1, 0)\}, \quad i = 1, 2, 3, \ \alpha \in (0.5, 1].$$

So,

$$\mathcal{P}_1(\tilde{X}_\alpha) \subsetneq \mathcal{P}_2(\tilde{X}_\alpha) \subsetneq \mathcal{P}_3(\tilde{X}_\alpha) = Conv(\mathcal{P}_3(\tilde{X}_\alpha)), \quad \forall \alpha \in (0, 0.5] \quad \text{and}$$
$$\mathcal{P}_1(\tilde{X}_\alpha) = \mathcal{P}_2(\tilde{X}_\alpha) = \mathcal{P}_3(\tilde{X}_\alpha), \quad \forall \alpha \in (0.5, 1].$$

Let us now compare of $\pi_{\tilde{X}}^1$, $\pi_{\tilde{X}}^2$ and $\pi_{\tilde{X}}^3$. From the calculations of $\mathcal{P}_1(\tilde{X}_\alpha)$ and $\mathcal{P}_2(\tilde{X}_\alpha)$ for each α, we observe that $\pi_{\tilde{X}}^2$ strictly dominates $\pi_{\tilde{X}}^1$. In fact, $\pi_{\tilde{X}}^2((1/3, 1/3, 1/3)) = 0.5$ and $\pi_{\tilde{X}}^1((1/3, 1/3, 1/3)) = 0$. Furthermore, $\pi_{\tilde{X}}^3$ strictly dominates $\pi_{\tilde{X}}^2$. These differences are reflected in the calculus of the entropy. In fact, we easily check that $H(\pi_{\tilde{X}}^1)(\log_2 3) = 0$. But let us now consider the fuzzy set $H(\pi_{\tilde{X}}^2)$ defined as

$$H(\pi_{\tilde{X}}^2)(x) = \sup\{\pi_{\tilde{X}}^2(Q) \, : \, H(Q) = x\}, \quad \forall x \in \mathbb{R}^+.$$

Hence we observe that $H(\pi_{\tilde{X}}^2)(\log_2 3) = 0.5$. We deduce that the second order possibility measure $\pi_{\tilde{X}}^2$ provides less precise information about the original probability P_{X_0} than $\pi_{\tilde{X}}^1$ does. Hence, the fuzzy set $H(\pi_{\tilde{X}}^3)$ defined as

$$H(\pi_{\tilde{X}}^3)(x) = \sup\{\pi_{\tilde{X}}^3(Q) \, : \, H(Q) = x\}, \quad \forall x \in \mathbb{R}^+$$

is also more imprecise than $H(\pi_{\tilde{X}}^1)$.)

Remark 4. We can slightly modify the conditions in last example and define another frv with the same probability distribution as $\tilde{X}$, but with a different fuzzy probability envelope, as we did in Remark 2 for the particular case of random sets. Let us consider the initial space $([0, 1], \beta_{[0,1]}, \lambda_{[0,1]})$ (the unit interval, with the usual Borel σ-algebra and the uniform probability distribution). Let us now consider the frv $\tilde{X}' : [0, 1] \to \mathcal{F}(\mathbb{R})$, defined as:

$$\tilde{X}'(\omega) = \begin{cases} \tilde{X}(\omega_1) & \text{if } \omega \leqslant 2/3, \\ \tilde{X}(\omega_2) & \text{if } \omega > 2/3. \end{cases}$$

We easily check that $\pi_{\tilde{X}'}^1 = \pi_{\tilde{X}'}^3 = \pi_{\tilde{X}}^3 \neq \pi_{\tilde{X}}^1$. Furthermore, both frv's induce the same probability distribution on $\sigma_{\mathcal{F}}$, i.e., $P \circ \tilde{X}^{-1} = P \circ \tilde{X}'^{-1}$. We conclude that, in general, the probability distribution induced by a frv does not determine its fuzzy probability envelope.

We have here investigated the information provided about the entropy. Next we will make a parallel study about the variance. In Example 3, we will deal with the particular case of random sets. Then (Example 4) we will go back to the general case of frv's.

Example 3. In this example, we will consider two random sets with the same probability distribution (when regarded as classical measurable mappings). Nevertheless, their probability envelopes are totally different. When both random sets represent individual random variables, the information provided by them is completely different.

Let us consider two random sets, each one defined on a different space:

(a) Let the first universe $\Omega_1 = \{a\}$ be a singleton. It represents a deterministic experiment. Let $\Gamma_1 : \Omega_1 \to \wp(\mathbb{R})$ be the multi-valued mapping defined as $\Gamma_1(a) = (-\infty, k]$, where k is an arbitrary positive number. Γ_1 represents the imprecise information about certain constant $X_1(a) = x_1$.

(b) Let Ω_2 be the unit interval. Let us consider the usual Borel σ-algebra and the uniform distribution defined on it. Let $\Gamma_2 : \Omega_2 \to \wp(\mathbb{R})$ be the constant multi-valued mapping defined as $\Gamma_2(\omega) = (-\infty, k]$, $\forall \omega \in [0, 1]$. Γ_2 represents the imprecise information about a random variable, $X_2 : [0, 1] \to \mathbb{R}$. All we know about it is an upper bound (k) for its images ($X_2(\omega) \leqslant k$, $\forall \omega \in [0, 1]$).

Both multi-valued mappings are strongly measurable with respect to the respective initial σ-algebras and $\beta_{\mathbb{R}}$. Hence, it can be checked that they are measurable mappings for the σ-algebra generated by the class $\{\wp_A : A \in \beta_{\mathbb{R}}\}$, where $\wp_A = \{C \in \mathbb{R}^n : C \cap A \neq \emptyset\}$. They induce the same probability measure on this σ-algebra. In fact, both of them take the "value" $(-\infty, k]$ with probability 1. Hence, they upper probability measures, $P^*_{\Gamma_1}$ and $P^*_{\Gamma_2}$, coincide, and hence also $\mathcal{P}_3(\Gamma_1) = \mathcal{P}_3(\Gamma_2)$ Nevertheless the classes $\mathcal{P}_1(\Gamma_1)$ and $\mathcal{P}_1(\Gamma_2)$ are not the same. We easily check that

$$\mathcal{P}_1(\Gamma_1) = \{\delta_c : c \leqslant k\}.$$

(For each $c \in \mathbb{R}$, δ_c represents the probability distribution degenerated on c.)

And, on the other hand,

$$\mathcal{P}_1(\Gamma_2) = \{Q \text{ probability} : Q((k, \infty)) = 0\}.$$

This reflects, for instance, that we know much more about the dispersion of X_1 than about the dispersion of X_2. In fact, we know that X_1 is a constant (its variance is null). However, we have no information about the variance of X_2 (under the available information, it can be any non-negative number).

We conclude that the probability distribution of a random set does not keep all the relevant information when it is regarded as the imprecise observation of a random variable.

Example 4. We can modify last example and consider two frv's, instead of random sets. Let each one of them be defined on a different space:

(a) Let the first universe $\Omega_1 = \{a\}$ be a singleton. Let $\tilde{X}_1 : \Omega_1 \to \mathcal{F}(\mathbb{R})$ be the multi-valued mapping defined as

$$\tilde{X}_1(a)(x) = \begin{cases} \dfrac{x}{k} & \text{if } x \in [0, k), \\ \dfrac{2k - x}{k} & \text{if } x \in [k, 2k], \end{cases}$$

where k is an arbitrary positive number. This frv represents the available imprecise information about certain constant $X_1(a) = x_1$.

(b) Let Ω_2 be the unit interval. Let us consider the usual Borel σ-algebra and the uniform distribution defined on it. Let $\tilde{X}_2 : \Omega_2 \to \wp(\mathbb{R})$ be the constant fuzzy-valued mapping defined as $\tilde{X}_2(\omega) = \tilde{X}_1(a)$, $\forall \omega \in [0, 1]$. $\tilde{X}_2$ represents the imprecise information about a random variable, $X_2 : [0, 1] \to \mathbb{R}$.

Both of them take the "value" $\tilde{X}_1(a)$ with probability 1. (They induce the same probability measure on a σ-algebra defined on $\mathcal{F}(\mathbb{R})$.) Nevertheless the second order possibility measures $\pi^1_{\tilde{X}_1}$ and $\pi^1_{\tilde{X}_2}$ are not the same. The information provided by $\pi^1_{\tilde{X}_1}$ is much more restrictive. In fact, it assigns null values to each non-degenerate probability.

Now we will not compare the information provided about certain parameter of the original distribution, but we will study the information about the joint probability measure associated to a pair of independent identically distributed random variables. We will describe the information about each one of them by a random set (Example 5) or a frv (Example 6). We will observe that the information provided by $\mathcal{P}_3(\Gamma)$ (resp., $\pi^3_{\tilde{X}}$) is much more imprecise than the information contained in $\mathcal{P}_1(\Gamma)$ (resp. $\pi^1_{\tilde{X}}$.)

Example 5. Suppose that we have an urn with three numbered balls. We know that all of them are coloured either red or white. Furthermore, we know that ball number 1 is red and ball number 3 is white. We do not know the colour of the second ball. We can represent our information about the colour of each ball by the multi-valued mapping $\Gamma : \{1, 2, 3\} \to \wp(\{r, w\})$ defined as $\Gamma(\{1\}) = \{r\}$, $\Gamma(\{2\}) = \{r, w\}$ and $\Gamma(\{3\}) = \{w\}$. Let us identify each probability Q on $P(\{r, w\})$

by the pair $(Q(\{r\}), Q(\{w\}))$. The probability envelope associated to Γ is $\mathcal{P}_1(\Gamma) = \{(1/3, 2/3), (2/3, 1/3)\}$. This set of probability measures represents the available information about the experiment consisting in drawing a ball at random and observing its colour.

Let us now suppose that a ball is drawn at random from the urn and replaced, and then a second ball is drawn at random (and the two drawings are independent). Let us identify now each probability measure Q on the product space by the tuple $(Q(\{rr\}), Q(\{rw\}), Q(\{wr\}), Q(\{ww\}))$. All we know about the true joint probability is that it belongs to the set

$$\{(p^2, p(1-p), (1-p)p, (1-p)^2) : (p, 1-p) \in \mathcal{P}_1(\Gamma)\}$$
$$\{(1/9, 2/9, 2/9, 4/9), (4/9, 2/9, 2/9, 1/9)\}.$$

On the other hand, the set of probability measures on $\wp(\{r, w\})$ that are dominated by P_Γ^* is the convex hull of $\mathcal{P}_1(\Gamma)$,

$$\mathcal{P}_3(\Gamma) = \{(p, 1-p) : p \in [1/3, 2/3]\}.$$

Nevertheless, the class

$$\{(p^2, p(1-p), (1-p)p, (1-p)^2) : (p, 1-p) \in \mathcal{P}_3(\Gamma)\}$$

is not included in the convex hull of

$$\{(p^2, p(1-p), (1-p)p, (1-p)^2) : p \in \mathcal{P}_1(\Gamma)\}.$$

In fact, $(0.25, 025, 0.25, 0.25)$ belongs to the first one, but not to the convex hull of the latter. In this situation, we are sure that the probability of $\{rw\}$ is 2/9. Nevertheless, when we consider the set $\mathcal{P}_3(\Gamma)$ as the class of probability measures compatible with the experiment we lose some information about this quantity: in that case, we only know that it is between 2/9 and 0.25. Hence, the information provided by the probability distribution induced by Γ (regarded as a $\wp(\{1, 2, 3\}) - \wp(\wp(\{r, w\}))$ measurable mapping) is not enough in this example.

Example 6. Let us now add some information to last example. Assume, that, in addition to the information given in last example, we now, with confidence 0.7 that the second ball is red. According to [9], the information about the experiment of drawing a ball at random can be now modelled by the frv: $\tilde{X} : \{1, 2, 3\} \to \Gamma(\{r, w\})$:

$$\tilde{X}(\{1\})(x) = \begin{cases} 1 & \text{if } x = r, \\ 0 & \text{if } x = w, \end{cases}$$

$$\tilde{X}(\{2\})(x) = \begin{cases} 1 & \text{if } x = r, \\ 0.3 & \text{if } x = w, \end{cases}$$

$$\tilde{X}(\{3\})(x) = \begin{cases} 1 & \text{if } x = w, \\ 0 & \text{if } x = r. \end{cases}$$

Let us first compare the second order possibility distributions $\pi_{\tilde{X}}^1$ and $\pi_{\tilde{X}}^3$. Based on the calculations in last example, they are defined as follows:

$$\pi_{\tilde{X}}^1((2/3, 1/3)) = 1, \pi_{\tilde{X}}^1((1/3, 2/3)) = 0 \quad \text{and} \quad \pi_{\tilde{X}}^1((p, 1-p)) = 0, \ \forall p \notin \{1/3, 2/3\}.$$

On the other hand,

$$\pi_{\tilde{X}}^3((p, 1-p)) = 0.3, \quad \forall p \in [1/3, 2/3] \quad \text{and} \quad \pi_{\tilde{X}}^3((2/3, 1/3)) = 1.$$

Let us now assume, as in Example 5, that a ball is drawn at random from the urn and replaced, and then a second ball is drawn again at random. The information about the true joint probability distribution is determined by the second order possibility measure prod $\pi_{\tilde{X}}^1$ defined as

$$\text{prod } \pi_{\tilde{X}}^1((p^2, p(1-p), (1-p)p, (1-p)^2)) = \pi_{\tilde{X}}^1((p, 1-p)), \quad \forall p \in [0, 1].$$

This information is quite different from the information described by

$$\text{prod } \pi_{\tilde{X}}^3((p^2, p(1-p), (1-p)p, (1-p)^2)) = \pi_{\tilde{X}}^3((p, 1-p)), \quad \forall p \in [0, 1].$$

For instance,

$$\text{prod } \pi_{\tilde{X}}^3(0.25, 0.25, 0.25, 0.25) = 0.3 \quad \text{nevertheless}$$

$$\text{prod } \pi_{\tilde{X}}^1(0.25, 0.25, 0.25, 0.25) = 0.$$

In last examples, we have compared the information provided by $\pi_{\tilde{X}}^1$, $\pi_{\tilde{X}}^2$ and $\pi_{\tilde{X}}^3$ about P_{X_0}. We have observed that, under the Kruse and Meyer interpretation, $\pi_{\tilde{X}}^1$ represents all the available information about this probability distribution. We have also noticed that, when we replace $\pi_{\tilde{X}}^1$ by $\pi_{\tilde{X}}^2$ or $\pi_{\tilde{X}}^3$, we can lose some relevant information.

For the particular case of random sets, the second order possibility measures $\pi_{\tilde{X}}^1$, $\pi_{\tilde{X}}^2$ and $\pi_{\tilde{X}}^3$ only take the extreme values 0 and 1. We identify them with the (crisp) sets of probability measures $\mathcal{P}_1(\Gamma)$, $\mathcal{P}_2(\Gamma)$ and $\mathcal{P}_3(\Gamma)$, respectively. We have also shown that $\mathcal{P}_3(\Gamma)$ represents the same information as the probability measure that the random set induces on $\sigma_\wp$. Hence, the probability measure induced by the random set (considered as a measurable mapping) does not contain all the relevant information about the original probability, P_{X_0}.

Returning to the general case of frv, we can also try to find the relationships between $\pi_{\tilde{X}}^3$ and the probability measure induced by $\tilde{X}$ on $\sigma_{\mathcal{F}}$. We prove below that the probability distribution $P \circ \tilde{X}^{-1}$ determines $\pi_{\tilde{X}}^3$. But as we shall demonstrate, the converse is not true: we will show an example of two frv's with the same convex fuzzy probability envelope, but inducing different probability measures.

Proposition 3. *The second order possibility distribution $\pi_{\tilde{X}}^3$ is determined by the induced probability distribution,* $P \circ \tilde{X}^{-1}$.

Proof. Let us first remind that $\pi_{\tilde{X}}^3$ is given by the nested family of sets $\{\mathcal{P}_3(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$. In other words, it is determined by the family of upper probabilities $\{P_{\tilde{X}_\alpha}^*\}_{\alpha \in [0,1]}$. Furthermore, each one of these upper probabilities is, as well, determined by $P \circ \tilde{X}^{-1}$. In fact:

$$P_{\tilde{X}_\alpha}^*(A) = P \circ \tilde{X}^{-1}(\mathcal{C}_\alpha^A). \quad \square$$

As we have pointed out, $\pi_{\tilde{X}}^3$ does not determine in general the induced probability distribution, $P \circ \tilde{X}^{-1}$. (It does when, in particular, $\tilde{X}$ is a random set, but not in general.) In the following example, we will define two frv's, $\tilde{X}$ and $\tilde{Y}$ associated to the same family of upper probabilities, $\mathcal{P}_3(\tilde{X}_\alpha) = \mathcal{P}_3(\tilde{Y}_\alpha)$, $\forall \alpha$. Nevertheless, their induced probability distributions do not coincide.

Example 7. Let us consider a probability space, $(\Omega, \mathcal{A}, P)$ where $\Omega = \{\omega_1, \omega_2\}$, $\mathcal{A} = \wp(\Omega)$ and $P(\{\omega_1\}) = 0.5 = P(\{\omega_2\})$. Let us consider the frv's $\tilde{X} : \Omega \to \mathcal{F}(\mathbb{R})$ and $\tilde{Y} : \Omega \to \mathcal{F}(\mathbb{R})$, whose α-cuts are defined as

$$\tilde{X}_\alpha(\omega_1) = [0, 3], \quad \alpha \leqslant 0.5, \quad \tilde{X}_\alpha(\omega_1) = [1, 2], \quad \alpha > 0.5,$$

$$\tilde{X}_\alpha(\omega_2) = [1, 3], \quad \alpha \leqslant 0.5, \quad \tilde{X}_\alpha(\omega_2) = [2, 3], \quad \alpha > 0.5,$$

$$\tilde{Y}_\alpha(\omega_1) = [0, 3], \quad \alpha \leqslant 0.5, \quad \tilde{Y}_\alpha(\omega_1) = [2, 3], \quad \alpha > 0.5,$$

$$\tilde{Y}_\alpha(\omega_2) = [1, 3], \quad \alpha \leqslant 0.5, \quad \tilde{Y}_\alpha(\omega_2) = [1, 2], \quad \alpha > 0.5.$$

We observe (Fig. 2) that they induce different probability distributions on $(\mathcal{F}(\mathbb{R}), \sigma_{\mathcal{F}})$. On the other hand their associated second order possibility distributions coincide, i.e.,

$$\pi_{\tilde{X}}^i = \pi_{\tilde{Y}}^i, \quad \forall i = 1, 2, 3.$$
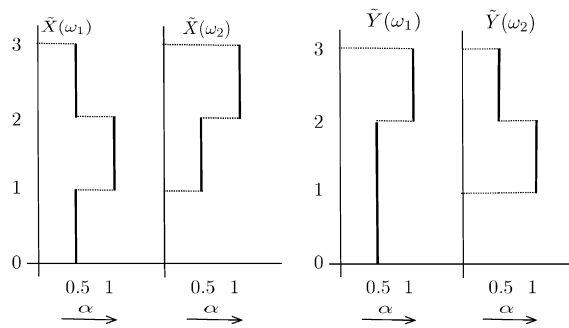
Fig. 2. Two fuzzy random variables with different probability distributions associated to the same nested family of envelopes.

On the one hand, the second order possibility distributions associated to a frv, $\tilde{X}$, do not determine its induced probability measure. On the other hand, this last one ($P \circ \tilde{X}^{-1}$) determines the second order possibility distribution $\pi_{\tilde{X}}^3$. But, as we derive from Remark 4 and Example 4, $P \circ \tilde{X}^{-1}$ does not determine the second order possibilities $\pi_{\tilde{X}}^1$ and $\pi_{\tilde{X}}^2$.

In last examples we have shown the differences between the information provided by the fuzzy probability envelope and the classical probability distribution induced by a frv. Next, we will illustrate the main ideas in the paper with an additional example. We will show the usefulness of the fuzzy probability envelope as the representation of the available imprecise information about the true probability distribution in a random experiment. We will represent the imprecise information about the weight of certain objects chosen at random by means of a frv. First we will show how the frv is derived from the imprecise data. We will then calculate the associated second order possibilities, as well as the induced probability measure. We will compare the information that these different models provide about the true probability distribution associated to the random experiment.

Example 8. Let us consider a population of objects, Ω, that we weigh with a scale. This scale is not completely precise. Let us suppose that 90% times, measurements are in a 5% error margin. But in general, we guarantee an error lower than 10%. Let us denote by $X_0(\omega)$ the (true) weight of an arbitrary object $\omega \in \Omega$ and let $D(\omega)$ represent the number displayed by the scale. Under the information above, we know that

$$[0.95\,X_0(\omega), 1.05\,X_0(\omega)] \ni D(\omega) \quad \text{with 90\% probability, and}$$
$$[0.9\,X_0(\omega), 1.1\,X_0(\omega)] \ni D(\omega) \quad \text{with certainty.}$$

Thus, we are sure that $X_0(\omega)$ is within $0.91D(\omega)$ and $1.11D(\omega)$. Furthermore, with probability 0.9, the true weight $X_0(\omega)$ is included in the interval from $0.95D(\omega)$ to $1.05D(\omega)$.

The information about the true weight $X_0(\omega)$ can be described by the random set $\Gamma_\omega : [0, 1] \to \mathbb{R}$ defined as

$$\Gamma_\omega(\alpha) = \begin{cases} [0.95D(\omega), 1.05D(\omega)] & \text{if } \alpha \leqslant 0.9, \\ [0.91D(\omega), 1.11D(\omega)] & \text{if } \alpha > 0.9, \end{cases}$$

where the uniform distribution is considered over $[0, 1]$. This random set reflects the following information: 90% times, the true weight is within $0.95D(\omega)$ and $1.05D(\omega)$. The remaining 10% times we can only assure that it is within $0.91D(\omega)$ and $1.11D(\omega)$. This information about $X_0(\omega)$ can be also described by the set of probabilities:

$$\{Q \,:\, Q([0,95D(\omega), 1.05D(\omega)]) \geqslant 0.95 \text{ and } Q([0.91D(\omega), 1.11\,D(\omega)]) = 1\}.$$

This class is the set of probability measures associated to the selections of Γ_ω, i.e., it coincides with

$$\mathcal{P}_1(\Gamma_\omega) = \{Q \text{ probability } :\, Q \in S(\Gamma_\omega)\}.$$

According to the results in [9], this class is the family of probability measures dominated by Π_ω:

$$\Pi_\omega(B) = \begin{cases} 0 & \text{if } B \cap [0.91D(\omega), 1.11D(\omega)] = \emptyset, \\ 0.1 & \text{if } B \cap [0.91D(\omega), 1.11D(\omega)] \neq \emptyset, \quad B \cap [0.95D(\omega), 1.05D(\omega)] = \emptyset, \\ 1 & \text{if } B \cap [0.95D(\omega), 1.05D(\omega)] \neq \emptyset. \end{cases}$$

It is the possibility measure associated to the possibility distribution (or fuzzy set) $\pi_\omega = \tilde{X}(\omega) : \mathbb{R} \to [0, 1]$ defined as

$$\tilde{X}(\omega)(x) = \begin{cases} 0 & \text{if } x \notin [0.91D(\omega), 1.11D(\omega)] = \emptyset, \\ 0.1 & \text{if } x \in [0.91D(\omega), 1.11D(\omega)] \setminus [0.95D(\omega), 1.05D(\omega)], \\ 1 & \text{if } x \in [0.95D(\omega), 1.05D(\omega)]. \end{cases}$$

Summarizing, the information about random variable $X_0 : \Omega \to \mathbb{R}$ is described by the frv $\tilde{X} : \Omega \to \mathcal{F}(\mathbb{R})$, where $\tilde{X}(\omega)$ is defined above for each $\omega \in \Omega$. For each pair (ω, x), $\tilde{X}(\omega)(x)$ represents the grade of possibility that the true weight of the object ω, $X_0(\omega)$, coincides with x. In other words, for a confidence level $1 - \alpha$ all we now about X_0 is that it belongs to the set

$$S(\tilde{X}_\alpha) = \{X : \Omega \to \mathbb{R} \text{ random variable } : X(\omega) \in [\tilde{X}(\omega)]_\alpha, \; \forall \omega \in \Omega\}.$$

Once we have represented the available information about X_0 by means of a frv, let us describe the (imprecise) available information about its induced probability measure, P_{X_0}. Let us suppose that Ω contains only 10 objects, $\omega_1, \ldots, \omega_{10}$ and the respective values displayed by the scale have been

500, 300, 400, 500, 500, 400, 400, 300, 300, 400.

Hence, the random variable D takes the values $d_1 = 300$, $d_2 = 400$, and $d_3 = 500$ with respective probabilities 0.3, 0.4, and 0.3. Thus, the frv $\tilde{X}$ takes also three different "values" in $\mathcal{F}(\mathbb{R})$, π_1, π_2 and π_3. Their membership functions are, respectively,

$$\pi_1(x) = \begin{cases} 0 & \text{if } x \notin [0.91\,d_1, 1.11\,d_1] = \emptyset, \\ 0.1 & \text{if } x \in [0.91\,d_1, 1.11\,d_1] \setminus [0.95\,d_1, 1.05\,d_1], \\ 1 & \text{if } x \in [0.95\,d_1, 1.05\,d_1], \end{cases}$$

$$\pi_2(x) = \begin{cases} 0 & \text{if } x \notin [0.91\,d_2, 1.11\,d_2] = \emptyset, \\ 0.1 & \text{if } x \in [0.91\,d_2, 1.11\,d_2] \setminus [0.95\,d_2, 1.05\,d_2], \\ 1 & \text{if } x \in [0.95\,d_2, 1.05\,d_2], \end{cases}$$

$$\pi_3(x) = \begin{cases} 0 & \text{if } x \notin [0.91\,d_3, 1.11\,d_3] = \emptyset, \\ 0.1 & \text{if } x \in [0.91\,d_3, 1.11\,d_3] \setminus [0.95\,d_3, 1.05\,d_3], \\ 1 & \text{if } x \in [0.95\,d_3, 1.05\,d_3]. \end{cases}$$

$\tilde{X}$ takes these "values" with respective probabilities $p_1 = 0.3$, $p_2 = 0.4$, and $p_3 = 0.3$. The mass function $p(\pi_i) = p_i$, $i = 1, 2, 3$ determines the probability distribution associated to $\tilde{X}$, $P \circ \tilde{X}^{-1}$. But this probability measure does not describe the information that $\tilde{X}$ captures about P_{X_0}. In fact, and according to the information given above, for each confidence level $1 - \alpha$, all we now about P_{X_0} is that it belongs to

$$\mathcal{P}_1(\tilde{X}_\alpha) = \{P_X \text{ probability } : X \in S(\tilde{X}_\alpha)\}.$$

Hence, the available information about P_{X_0} should be described by means of the second order possibility distribution $\pi^1_{\tilde{X}}$ (the fuzzy set associated to the graded representation $\{\mathcal{P}_1(\tilde{X}_\alpha)\}_{\alpha \in [0,1]}$. Let us describe $\mathcal{P}_1(\tilde{X}_\alpha)$ for each $\alpha \in [0, 1]$.

- For $\alpha \leqslant 0.1$, $Q \in \mathcal{P}_1(\tilde{X}_\alpha)$ if and only if:

$$\exists \, x_1, x_2, x_3 \in [0.91 \, d_1, 1.11 \, d_1] : \sum_{i=1}^{3} Q(\{x_i\}) = 0.3,$$

$$\exists \, x_4, x_5, x_6, x_7 \in [0.91 \, d_2, 1.11 \, d_2] : \sum_{i=4}^{7} Q(\{x_i\}) = 0.4,$$

$$\exists \, x_8, x_9, x_{10} \in [0.91 \, d_3, 1.11 \, d_3] : \sum_{i=8}^{10} Q(\{x_i\}) = 0.3.$$

- For $\alpha > 0.1$, $Q \in \mathcal{P}_1(\tilde{X}_\alpha)$ if and only if:

$$\exists \, x_1, x_2, x_3 \in [0.95 \, d_1, 1.05 \, d_1] : \sum_{i=1}^{3} Q(\{x_i\}) = 0.3,$$

$$\exists \, x_4, x_5, x_6, x_7 \in [0.95 \, d_2, 1.05 \, d_2] : \sum_{i=4}^{7} Q(\{x_i\}) = 0.4,$$

$$\exists \, x_8, x_9, x_{10} \in [0.95 \, d_3, 1.05 \, d_3] : \sum_{i=8}^{10} Q(\{x_i\}) = 0.3.$$

In other words, $\pi_{\tilde{X}}^1$ takes the values 0, 0.1 and 1. These values are assigned as follows:

$$\pi_{\tilde{X}}^1(Q) = \begin{cases} 0 & \text{if } Q \notin \mathcal{P}_1(\tilde{X}_{0.1}), \\ 0.1 & \text{if } Q \in \mathcal{P}_1(\tilde{X}_{0.1}) \setminus \mathcal{P}_1(\tilde{X}_1), \\ 1 & \text{if } Q \in \mathcal{P}_1(\tilde{X}_1). \end{cases}$$

It is easy to check that $\pi_{\tilde{X}}^2$ coincides with $\pi_{\tilde{X}}^1$ in this example. On the other hand $\pi_{\tilde{X}}^3$ is strictly greater than them. We can observe that

$$\pi_{\tilde{X}}^3(Q) = \begin{cases} 1 & \text{if } Q(A_{11}) = 0.3, \, Q(A_{21}) = 0.4 \text{ and } Q(A_{31}) = 0.3, \\ 0.1 & \text{if } Q(A_{12}) = 0.3, \, Q(A_{22}) = 0.4 \text{ and } Q(A_{32}) = 0.3 \text{ but } \pi_{\tilde{X}}^3(Q) \neq 1, \\ 0 & \text{otherwise}, \end{cases}$$

where

$$A_{11} = [0.95d_1, 1.05d_1]), \quad A_{21} = [0.95d_2, 1.05d_2], \quad A_{31} = [0.95d_3, 1.05d_3],$$
$$A_{12} = [0.91d_1, 1.11d_1]), \quad A_{22} = [0.91d_2, 1.11d_2], \quad A_{32} = [0.91d_3, 1.11d_3].$$

Notice that the information provided by $\pi_{\tilde{X}}^1$ is more precise than the information provided by $\pi_{\tilde{X}}^3$. In fact, $\pi_{\tilde{X}}^3$ does not distinguish whether the original probability distribution is discrete or continuous. The same happens with the probability distribution induced by $\tilde{X}$, $P \circ \tilde{X}^{-1}$. It does not have into account the nature of the initial space. The same probability distribution could be generated for a non-atomic initial space. In that case, the original probability measure could be continuous, and $\pi_{\tilde{X}}^1$ should be convex, i.e.,

$$\pi_{\tilde{X}}^1(\lambda \, Q_1 + (1 - \lambda) \, Q_2) \geqslant \lambda \pi_{\tilde{X}}^1(Q_1) + (1 - \lambda) \, Q_2, \quad \forall \lambda \in [0, 1].$$

However, in our example, $\pi_{\tilde{X}}^1$ is not convex. It reflects that the original probability is focused on no more than 10 points.

6. Concluding remarks

Most studies in the literature about fuzzy random variables have a formal foundation in (classical) Probability Theory. There, the fuzzy random variable can be viewed as a classical $\mathcal{A}$–$\mathcal{B}$ measurable mapping, where $\mathcal{B}$ is a particular

σ-algebra defined over a class of fuzzy subsets of the final space. Within this framework, we can consider the probability distribution induced by the fuzzy random variable on $\mathcal{B}$. Classical techniques in Probability Theory can be used to extend some common definitions (expectation, variance, correlation coefficient, etc.) and results to this new environment. The present paper connects previous studies about fuzzy random variables with the Theory of Imprecise Probabilities. The second order model here described represents the information about the probability distribution induced by an imprecisely observed random variable. Along the paper, we compare, in several situations, the information provided by this second order model and the (classical) probability measure induced by the frv. We show that, under a possibilistic interpretation of fuzzy sets, the probability measure induced by the frv is not the most suitable way to describe this type of imprecise information.

The subject of the paper is closely related to some discussions about convexity in the context of Imprecise Probabilities (see [30] for instance). We have argued that sets of probabilities should not be convex in general, when modelling frv's and random sets in particular. The convex-hull of the set of probabilities associated to a random set loses some relevant information about the possible values of the entropy or the variance, as we have illustrated in the text. When the available information about the likelihood function associated to a simple random sample is being modelled, significant information is also lost. This last comment is based on the ideas suggested in Examples 5 and 6 and it is somehow related to the concept to "repetition independency" proposed in [12]. Nevertheless the bounds of the set $\mathcal{P}_1(\tilde{X}_\alpha)(A)$ and $\mathcal{P}_3(\tilde{X}_\alpha)(A)$ are coincident under somewhat general conditions, provided that we are not devising a model for a product space. This is not true in the general case, though, as we have demonstrated in [35].

Sigma-additivity is used, because we intend to show the shortcomings of the probability measure induced by a frv in the original context of all the works in this area. Nevertheless, we think that it should be interesting to make a similar comparison when we drop σ-additivity in the future.

We should notice that the model here presented is quite related to the concept of nested family of probability boxes [21]. A probability box (or just, p-box) [22] represents a pair of upper and lower cumulative distribution functions. Thus, a nested family of p-boxes (each one of them associated to a different confidence degree $1-\alpha$) represents a second order model: let us consider, for an arbitrary α-cut, the upper and lower cumulative distribution functions associated to $\tilde{X}_\alpha$, i.e.,

$$F_{\tilde{X}_\alpha}^*(x) := P_{\tilde{X}_\alpha}^*((-\infty, x]) \quad \text{and} \quad F_{*\tilde{X}_\alpha}(x) := P_{*\tilde{X}_\alpha}((-\infty, x]), \quad \forall x \in \mathbb{R}.$$

The nested family of p-boxes $[F_{*\tilde{X}_\alpha}, F_{\tilde{X}_\alpha}^*]_{\alpha \in [0,1]}$ is the gradual representation of a second order possibility distribution, $\pi_{\tilde{X}}^F$. We easily observe that it dominates $\pi_{\tilde{X}}^3$, and so

$$\pi_{\tilde{X}}^F(Q) \geqslant \pi_{\tilde{X}}^3(Q) \geqslant \pi_{\tilde{X}}^1(Q), \quad \forall Q.$$

For the particular case of random sets, this means that the class of probability measures associated to the p-box $[F_{*\Gamma}, F_\Gamma^*]$, contains the class $\mathcal{P}_3(\Gamma)$, and hence it also contains $\mathcal{P}_1(\Gamma)$. In other words, it provides less precise information than the probability envelope of Γ. More detailed studies about these relationships can be found in [13]. Nevertheless, p-boxes and nested families of p-boxes are useful representations in other different situations, when we have imprecise information about the "true" probability associated to a random experiment, but not about the images of a random variable. In some recent studies, the information given through a p-box is combined with different kinds of information about the "shape" of the true density function, the mean, the mode, etc.

The second order possibility model here described does not aggregate imprecision and randomness. But, a different interpretation of fuzzy random variables aggregating both sources of uncertainty (randomness and imprecision) leads to a pair of upper and lower probabilities (a first order model), as we show in [2,3,6,7,11,32]. Future works should try to compare the second order model here studied with this kind of upper–lower probability model. We could verify whether the intervals of upper–lower probabilities coincide with the mean value [18] and/or with the Walley's reduction [47] of the fuzzy values of probability determined by the second order model. Furthermore, the upper and lower probabilities considered in [2,3] are, in fact, a pair of belief and plausibility functions. We should verify if, for infinite universes, they are also ∞-monotone/resp. alternating Choquet capacities. In that case, we should try to find some random set associated to these capacities, and find the relationships between it and the initial fuzzy random variable.

Acknowledgements

References

 [1] J. Aumann, Integral of set valued functions, J. Math. Anal. Appl. 12 (1965) 1–12.
 [2] C. Baudrit, I. Couso, D. Dubois, Probabilities of events induced by fuzzy random variables, in: Proc. Internat. Conf. in Fuzzy Logic and Technology (EUSFLAT 05), Barcelona, Spain, 2005.
 [3] C. Baudrit, I. Couso, D. Dubois, Joint propagation of probability and possibility in risk analysis: towards a formal framework, Internat. J. Approx. Reasoning 45 (2007) 82–105.
 [4] L. Boyen, G. de Cooman, E.E. Kerre, On the extension of P-consistent mappings, in: G. de Cooman, D. Ruan, E.E. Kerre (Eds.), Foundations and Applications of Possibility Theory—Proc. of FAPT'95, World Scientific, Singapore, 1995, pp. 88–98.
 [5] J. Casillas, L. Sánchez, Knowledge extraction from fuzzy data for estimating consumer behavior models, in: Proc. 2006 IEEE Internat. Conf. on Fuzzy Systems (FUZZ-IEEE'06), Vancouver, Canada, 2006, pp. 572–578.
 [6] I. Couso, D. Dubois, S. Montes, L. Sánchez, On various definitions of the variance of a fuzzy random variable, in: Proc. Fifth Internat. Symp. on Imprecise Probabilities: Theory and Applications (ISIPTA 07), Prague, Czech Republic, 2007, pp. 135–144.
 [7] 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, Berlin, 2004, pp. 133–140.
 [8] I. Couso, S. Montes, P. Gil, Función de distribución y mediana de variables aleatorias difusas, in: Proc. VIII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF'98), Pamplona, Spain, 1998, pp. 279–284 (in Spanish).
 [9] I. Couso, S. Montes, P. Gil, The necessity of the strong alpha-cuts of a fuzzy set, Internat. J. Uncertainly Fuzziness and Knowledge-Based Systems 9 (2001) 249–262.
[10] 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.
[11] I. Couso, S. Montes, L. Sánchez, Varianza de una variable aleatoria difusa. Estudio de distintas definiciones, in: Proc. of XII Congreso Español de Tecnologías y Lógica Difusa (ESTYLF'04), Jaén, Spain, 2004. (in Spanish).
[12] I. Couso, S. Moral, P. Walley, A survey of concepts of independence for imprecise probabilities, Risk, Decision and Policy 5 (2000) 165–181.
[13] I. Couso, L. Sánchez, P. Gil, Imprecise distribution function associated to a random set, Inform. Sci. 159 (2004) 109–123.
[14] G. de Cooman, A behavioural model for vague probability assessments, Fuzzy Sets and Systems 154 (2005) 305–358.
[15] G. de Cooman, D. Aeyels, Supremum preserving upper probabilities, Inform. Sci. 118 (1999) 173–212.
[16] G. de Cooman, P. Walley, An imprecise hierarchical model for behaviour under uncertainty, Theory and Decision 52 (2002) 327–374.
[17] P. Diamond, P. Kloeden, Metric Spaces of Fuzzy Sets, World Scientific, Singapore, 1994.
[18] D. Dubois, H. Prade, The mean value of a fuzzy number, Fuzzy Sets and Systems 24 (1987) 279–300.
[19] D. Dubois, H. Prade, The three semantics of fuzzy sets, Fuzzy Sets and Systems 90 (1997) 141–150.
[20] Y. Feng, L. Hu, H. Shu, The variance and covariance of fuzzy random variables and their applications, Fuzzy Sets and Systems 120 (2001) 487–497.
[21] S. Ferson, L. Ginzburg, V. Kreinovich, H.T. Nguyen, S.A. Starks, Uncertainty in risk analysis: towards a general second-order approach combining interval, probabilistic, and fuzzy techniques, in: Proc. 2002 IEEE Internat. Conf. on Fuzzy Systems (FUZZ-IEEE'02), Honolulu, Hawaii, USA, 2002, pp. 1342–1347.
[22] S. Ferson, V. Kreinovich, L. Ginzburg, K. Sentz, D.S. Myers, Constructing probability boxes and Dempster–Shafer structures, Sandia National Laboratories, SAND2002-4015, Albuquerque, NM, USA, 2003.
[23] J.A. Herencia, Graded sets and points: a stratified approach to fuzzy sets and points, Fuzzy Sets and Systems 77 (1996) 191–202.
[24] E.P. Klement, M.L. Puri, D.A. Ralescu, Limit theorems for fuzzy random variables, Proc. Roy. Soc. London A 407 (1986) 171–182.
[25] R. Körner, On the variance of fuzzy random variables, Fuzzy Sets and Systems 92 (1997) 83–93.
[26] V. Krätschmer, A unified approach to fuzzy random variables, Fuzzy Sets and Systems 123 (2001) 1–9.
[27] R. Kruse, On the variance of random sets, J. Math. Anal. Appl. 122 (1987) 469–473.
[28] R. Kruse, K.D. Meyer, Statistics with Vague Data, D. Reidel Publishing Company, Dordrecht, 1987.
[29] H. Kwakernaak, Fuzzy random variables: definition and theorems, Inform. Sci. 15 (1978) 1–29.
[30] H.E. Kyburg Jr., M. Pittarelli, Set-based Bayesianism, IEEE Trans. Systems Man Cybernet. 26 (1996) 324–339.
[31] M.A. Lubiano, Variation measures for imprecise random elements, Ph.D. Thesis, Universidad de Oviedo, Spain, 1999 (in Spanish).
[32] E. Miranda, G. de Cooman, I. Couso, Lower previsions induced by multi-valued mappings, J. Stat. Plann. Inference 133 (2005) 173–197.
[33] E. Miranda, I. Couso, P. Gil, Study of the probabilistic information of a random set, in: Proc. Third Internat. Symp. on Imprecise Probabilities and Their Applications (ISIPTA'03), Lugano, Switzerland, 2003.
[34] E. Miranda, I. Couso, P. Gil, Random intervals as a model for imprecise information, Fuzzy Sets and Systems 154 (2005) 386–412.
[35] E. Miranda, I. Couso, P. Gil, Random sets as imprecise random variables, J. Math. Anal. Appl. 307 (2005) 32–47.
[36] H.T. Nguyen, On random sets and belief functions, J. Math. Anal. Appl. 65 (1978) 531–542.
[37] A. Otero, J. Otero, L. Sánchez, J.R. Villar, Longest path estimation from inherently fuzzy data acquired with GPS using genetic algorithms, in: Second Internat. Symp. on Evolving Fuzzy Systems 2006 (EFS06), Ambleside Lake District, UK, 2006, pp. 300–306.
[38] M. Puri, D. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. 114 (1986) 409–422.

[39] L. Sánchez, I. Couso, Advocating the use of imprecisely observed data in genetic fuzzy systems, IEEE Trans. Fuzzy Systems 15 (2007) 551–562.

[40] L. Sánchez, I. Couso, J. Casillas, A multiobjective genetic fuzzy system with imprecise probability fitness for vague data, in: Second Internat. Symp. on Evolving Fuzzy Systems 2006 (EFS06), Ambleside, Lake District, UK, 2006, pp. 131–136.

[41] L. Sanchez, I. Couso, J. Casillas. Knowledge extraction from vague data with genetic fuzzy systems under a combination of crisp and imprecise criteria, IEEE Trans. on Fuzzy Systems 2007, submitted for publication.

[42] L. Sánchez, J. Otero, J.R. Villar, Boosting of fuzzy models for high-dimensional imprecise datasets, in: Proc. 11th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference (IPMU), Paris, France, 2006.

[43] L. Sánchez, M.R. Suárez, I. Couso, A fuzzy definition of mutual information with application to the design of genetic fuzzy classifiers, in: Proc. Internat. Conf. on Machine Intelligence (ACIDCA-ICMI05), Tozeur, Tunisia, 2005, pp. 602–609.

[44] L. Sánchez, M.R. Suárez, J.R. Villar, I. Couso, Some results about mutual information-based feature selection and fuzzy discretization of vague data, in: Proc. 16th IEEE Internat. Conf. on Fuzzy Systems (FUZZ-IEEE07), London, UK, 2007.

[45] L. Sánchez, M.R. Suárez, J.R. Villar, I. Couso, Mutual information-based feature selection and fuzzy discretization of vague data, Soft Comput. 2007, submitted for publication.

[46] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman & Hall, London, 1991.

[47] P. Walley, Statistical inferences based on a second-order possibility distribution, Internat. J. General Systems 26 (1997) 337–384.