Upper and lower probabilities induced by a fuzzy random variable

Inés Couso[a*], Luciano Sánchez[b]

University of Oviedo (Spain)
[a] Dep. of Statistics and O.R. couso@uniovi.es
[b] Dep. of Computer Sciences luciano@uniovi.es

Abstract

We review two existing interpretations of fuzzy random variables. In the first one, the fuzzy random variable is viewed as a linguistic random variable. In the second case, it represents some incomplete knowledge about an otherwise standard random variable. For each interpretation, the information provided by the frv is described by a specific model, namely a standard probability model and a second-order imprecise model, respectively. In this paper, we deal with an alternative interpretation. Guided by simple examples we will observe the usefulness of each interpretation when applied to particular situations. Then we will demonstrate that the new interpretation leads, in a natural way, to pair of order ∞ capacities. Furthermore, we show that they are formally related to the former models. These results can be applied in future works to make inferences from fuzzy sample data. The use of upper-lower models instead of second-order models will enable us in the future to reach crisp decisions in some specific statistical problems, without adding any arbitrary information, and taking into account the imprecision in data.

Keywords.- Imprecise probabilities, second-order possibility measure, fuzzy random variable, belief function, multi-valued mapping.

[*]Corresponding author

1 Introduction

A fuzzy random variable, $\tilde{X}$, is a mapping that assigns a fuzzy subset of the final space, $\tilde{X}(\omega) \in \mathcal{F}(\Omega_2)$, to any element of the initial space, $\omega \in \Omega_1$ (provided of probability space structure). This association expresses some imprecise information about the relation between the outcomes in both universes. Thus, the concept of fuzzy random variable extends the classical definition of random variable. In fact, a random variable assigns an element in Ω_2 to each outcome $\omega_1 \in \Omega_1$. So it expresses a deterministic relation between Ω_1 and Ω_2. (Once an element from Ω_1 has been selected, the element in Ω_2 is univocally determined.)

From a formal point of view, the concept of fuzzy random variable (frv for short) is not unique, and 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. On the other hand, fuzzy random variables have been given different interpretations in the literature (see [8] for a detailed discussion), independently from the specific formal definition. Let us now briefly recall two different interpretations existing in the literature.

In [40], Puri and Ralescu consider that the observations of some random experiments do not consist of numerical outputs, but are represented by vague linguistic terms. When, in particular, the fuzzy random variable has a finite number of images, probability values can be assigned to different "fuzzy labels". For example, the following model could be generated: the result is "high" with probability 0.5, "medium" with probability 0.25 and "low" with probability 0.25, where "high", "medium" and "low" are linguistic labels associated to fuzzy subsets of the final space.

Kruse and Meyer [24] offer an alternative interpretation of fuzzy random variables, as the representation if ill-known random variables. The fuzzy random variable induces an "acceptability function" defined over the class of all (classical) random variables. In this setting, an acceptability function over the class of all possible probability distributions is defined (see [8]) in a natural way. It represents the available information about the true probability distribution of the random experiment under study. From a theoretical point of view, this model is a second-order possibility distribution ([11, 13]): the information about the true probability is represented by a possibility measure defined over the class of all probability measures.

In this paper, we deal with an additional interpretation of fuzzy random variables, different from Puri-Ralescu and Kruse-Meyer approaches. The essential difference wrt Kruse-Meyer's approach focusses on the omission of the assumption about the existence of an underlying (classical) random variable. Under the new interpretation, the final outcome in Ω_2 is not assumed to be determined by the initial outcome $\omega_1 \in \Omega_1$, and so, the frv does not represent an acceptability function over the class of random variables anymore. Anyway, the new interpretation is in accordance with the possibilistic interpretation of fuzzy sets, as well as the Kruse and Meyer's approach. More specifically, suppose that we have partial information about the probability distribution that models a sequence of two random experiments on Ω_1 and Ω_2, respectively. Suppose, on the one hand, that the probability distribution that models the first one, P_1, is completely determined. On the other hand, the second experiment is only known via a family of conditional possibility measures. This family models our knowledge about the relationship between the outcome of the first sub-experiment and the possible outcomes of the second one. (If the result of the first experiment is ω, then the possibility degree of x occurring in the second one is $\tilde{X}(\omega)(x)$.) The combination of both sources of information leads us to describe, in a natural way, the available information about the probability distribution on Ω_2 (the probability distribution that rules the second sub-experiment) by means of an upper probability (a standard imprecise probability model, not an order-2 model, like the one described before.) This interpretation has been already introduced in [1, 3, 4, 5]. In [1], the specific situation where the frv is built from a crisp random variable and a fuzzy set in a finite environment have been studied. In [5], the notion of expectation is studied under this new perspective. But the differences among the three approaches becomes more clear for the concept of variance. In [3], the usefulness of each of the three definitions of variance when applied to particular situations is raised, under the light of simple examples. In the present paper, we study the formal relationships between the three models associated to each of the three interpretations (the probability measure associated to the frv, the second-order possibility distribution, and the upper-lower model). We also show that the upper and lower probabilities are, in fact, ∞-order capacities. Furthermore, we define a specific multi-valued mapping inducing such capacities, providing it with a meaningful interpretation. Finally we show that the upper and lower probabilities associated to this new approach admit an alternative interpretation under the Kruse-Meyer approach. It is based on a method developed by Walley [48], where an upper-lower probability model is derived from any second-order possibility measure, by using natural extension techniques.

The paper is organized as follows. In Section 2, we provide the necessary technical background. In Section 3, we recall Puri-Ralescu and Kruse-Meyer approaches. We show that these interpretations lead to a classical probability model and a second-order imprecise probability model, respectively. In Section 4, we introduce the third interpretation of fuzzy random variables as families of conditional possibility measures (the one described in the last paragraph). We describe in Section 5 the upper-lower model associated to it. We check in Section 6 that the upper and lower probabilities given in Section 5 are, in fact, $\infty-$order capacities, and we define a multi-valued mapping associated to them. We will also give an intuitive interpretation to this multi-valued mapping and show how it encompasses the same information as the initial fuzzy random variable. In Section 7, we will show the relationships between the upper and lower model given in Section 5, the classical and the second-order imprecise model recalled in Section 3. In particular we will show that the upper-lower model given in Section 5 coincides with the reduction (applying Walley's technique) of the second-order possibility model. We will derive some interesting consequences. We end the paper with some concluding remarks.

2 Preliminary concepts and notation

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

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 degree 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 degree 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) $\alpha-$cut of π:

$$\pi_\alpha = \{u \in U \,:\, \pi(u) \geq \alpha\}$$

For any $\alpha \in [0,1)$, $\pi_{\overline{\alpha}}$ will denote the strong $\alpha-$cut of π :

$$\pi_\alpha = \{u \in U \,:\, \pi(u) > \alpha\}$$

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 *graded set* [21] of U is a multi-valued mapping $\varphi : [0,1] \to \wp(U)$ satisfying

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

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

$$[\pi_\varphi]_{\overline{\alpha}} \subseteq \varphi(\alpha) \subseteq [\pi_\varphi]_\alpha, \forall\, \alpha \in (0,1).$$

Furthermore, it can be calculated from φ as:

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

Consider a probability space $(\Omega_1, \mathcal{A}_1, P_1)$ and an arbitrary measurable space $(\Lambda, \sigma_\Lambda)$. Given an $\mathcal{A}_1 - \sigma_\Lambda$ measurable mapping, $f : \Omega_1 \to \Lambda$, we will denote by $P_1 \circ f^{-1}$ the probability measure it induces on σ_Λ, i.e.:

$$P_1 \circ f^{-1}(B) := P_1(f^{-1}(B)),\ \forall\, B \in \sigma_\Lambda,$$

where $f^{-1} : \wp(\Lambda) \to \wp(\Omega_1)$ is the mapping defined as

$$f^{-1}(B) = \{\omega \in \Omega \,:\, f(\omega) \in B\},\ \forall\, B \subseteq \Lambda.$$

Let us now consider the measurable spaces, (Ω_1, σ_1) and (Ω_2, σ_2), and the multi-valued mapping $\Gamma : \Omega_1 \to \wp(\Omega_2)$ with non-empty images. Let $B \in \sigma_2$ be an arbitrary measurable set. The *upper inverse* of B is the set

$$\Gamma^*(B) = \{\omega \in \Omega_1 \,:\, \Gamma(\omega) \cap B \neq \emptyset\}.$$

The *lower inverse* of B is the set

$$\Gamma_*(B) = \{\omega \in \Omega_1 \,:\, \Gamma(\omega) \subseteq B\} = [\Gamma^*(B^c)]^c.$$

Γ is said to be *strongly measurable* [36] when $\Gamma^*(B) \in \sigma_1,\ \forall\, B \in \sigma_2$. We will say that the multi-valued mapping $\Gamma : \Omega_1 \to \wp(\Omega_2)$ is a *random set* when it is strongly measurable. We will denote by $S(\Gamma)$ the class of *measurable selections* of Γ:

$$S(\Gamma) = \{X : \Omega_1 \to \Omega_2,\ \sigma_1 - \sigma_2 \text{ measurable } \,:\, X(\omega) \in \Gamma(\omega),\ \forall\, \omega \in \Omega_1\}.$$

Given a probability measure P_1 on (Ω_1, σ_1), we will respectively denote by $P^*_\Gamma : \sigma_2 \to [0,1]$ and $P_{*\Gamma} : \sigma_2 \to [0,1]$ the *Dempster upper and lower probabilities* associated to Γ,

$$P^*_\Gamma(B) = P_1(\Gamma^*(B)) = P_1(\{\omega \in \Omega_1 \,:\, \Gamma(\omega) \cap B \neq \emptyset\}),\ \forall\, B \in \sigma_2,$$

$$P_{*\Gamma}(B) = P_1(\Gamma_*(B)) = P_1(\{\omega \in \Omega_1 \,:\, \Gamma(\omega) \subseteq B\}),\ \forall\, B \in \sigma_2.$$

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

$$\Gamma_*(B) \subseteq X^{-1}(B) \subseteq \Gamma^*(B), \ \forall B \in \sigma_2.$$

Thus, the upper probability dominates every probability measure induced by a measurable selection of Γ, i.e., for every $B \in \sigma_2$:

$$P_X(B) = P_1 \circ X^{-1}(B) = P_1(X^{-1}(B)) \leq P_1(\Gamma^*(B)) = P^*_\Gamma(B), \forall X \in S(\Gamma),$$

and the lower probability is dominated by it, i.e., for every $B \in \sigma_2$:

$$P_X(B) = P_1 \circ X^{-1}(B) = P_1(X^{-1}(B)) \geq P_1(\Gamma_*(B)) = P_{*\Gamma}(B), \forall X \in S(\Gamma).$$

Thus, the class of probability measures associated to the selections of Γ:

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

is included in the class of probability measures dominated by P^*_Γ,

$$\{P : \sigma_2 \to [0,1] \text{ prob. } \ : P(B) \leq P^*_\Gamma(B), \ \forall B \in \sigma_2\} =$$

$$\{P : \sigma_2 \to [0,1] \text{ prob. } \ : P_{*\Gamma}(B) \leq P(B) \leq P^*_\Gamma(B), \ \forall B \in \sigma_2\}$$

Further relationships between both classes are investigated in [26, 27, 30, 31, 32]. From now on, for every $B \in \sigma_2$, we will denote by $\mathcal{P}(\Gamma)(B)$ the subset of $[0,1]$ determined as follows:

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

On the other hand, the upper probability of Γ represents the same information as the probability measure induced by Γ (considered as a "classical" measurable function). In fact, consider, for each $B \in \sigma_2$, the family of sets:

$$\mathcal{C}_B = \{C \in \sigma_2 \ : \ C \cap B \neq \emptyset\}.$$

Also consider the $\sigma-$algebra, $\sigma_{\mathcal{C}}$, generated by the class $\mathcal{C} = \{\mathcal{C}_B \ : \ B \in \sigma_2\}$ on the universe of "elements" $\wp(\Omega_2)$. We can easily check that Γ is strongly measurable if and only if it is $\sigma_1 - \sigma_{\mathcal{C}}$ measurable (regarded as a classical function and not as a multi-valued one). Furthermore, the probability measure induced by Γ on $\sigma_{\mathcal{C}}$ (the probability measure $P_1 \circ \Gamma^{-1}$) determines P^*_Γ. In fact, the following equalities hold:

$$P_1 \circ \Gamma^{-1}(\mathcal{C}_B) = P_1(\Gamma^{-1}(\mathcal{C}_B)) = P_1(\Gamma^*(B)) = P^*_\Gamma(B), \ \forall B \in \sigma_2.$$

And the converse is also true: the upper probability univocally determines $P \circ \Gamma^{-1}$, since the class $\{\mathcal{C}_B^c \,:\, B \in \sigma_2\}$ is a π-system (it is closed for intersections) so there cannot exist a different probability on $\sigma_{\mathcal{C}}$ agreeing with $P_1 \circ \Gamma^{-1}$ on $\mathcal{C}$, according to a well known result in Probability Theory (see [2], Theorem 3.3, for instance.) Let now $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be an arbitrary fuzzy–valued mapping. For each $\alpha \in [0,1]$, we will denote by $\tilde{X}_\alpha$ the α-cut of $\tilde{X}$, i.e., the multi-valued mapping $\tilde{X}_\alpha : \Omega_1 \to \wp(\Omega_2)$ that assigns, to each $\omega \in \Omega_1$, the (weak) $\alpha-$cut of $\tilde{X}(\omega)$. We will say that $\tilde{X}$ is a *fuzzy random variable* when every $\alpha-$cut, $\tilde{X}_\alpha$, is strongly measurable. This condition is equivalent to a "classical" measurability assumption. In fact, consider, for each $B \in \sigma_2$ and each $\alpha \in [0,1]$, the family of sets:

$$\mathcal{F}_B^\alpha = \{F \in \mathcal{F}(\Omega_2) \,:\, F_\alpha \cap B \neq \emptyset\}.$$

Now denote by $\sigma_{\mathcal{F}}$ the $\sigma-$algebra generated by the class:

$$\{\mathcal{F}_B^\alpha, \,:\, B \in \sigma_2, \alpha \in [0,1]\} \subseteq \wp(\mathcal{F}(\Omega_2)).$$

We can easily check that a fuzzy-valued mapping is a frv if and only if it is $\sigma_1 - \sigma_{\mathcal{F}}$ measurable.

Given an arbitrary non-negative mapping $f : \Omega_1 \to \mathbb{R}^+$, we will denote by $(C)\int f\,d\mu$ the Choquet integral of f with respect to the set-function $\mu : \sigma_1 \to [0,1]$:

$$(C)\int f\,d\mu = \int_0^\infty \mu(f > x)\,dx.$$

When μ is an alternating capacity of order 2, the following equality holds:

$$(C)\int f\,d\mu = \sup\left\{\int f\,dP \,:\, P \leq \mu\right\},$$

where $P \leq \mu$ means that P is dominated by μ i.e. $P(A) \leq \mu(A),\ \forall A \in \sigma_1$.

A *possibilistic probability* (or a "fuzzy probability") [11], $\tilde{P} : \sigma_2 \to \mathcal{F}([0,1])$, is a map taking each event[1] $B \in \sigma_2$ to a normal possibility distribution, $\tilde{P}(B)$ on $[0,1]$. Its value $\tilde{P}(B)(p)$ in a point $p \in [0,1]$ can be interpreted as the modeller's "upper betting rate" that the true probability of the event B is equal to p. (The supremum of the prices one is willing to pay for gaining 1 unit in probability currency, if the actual probability of B coincides with p).

[1] 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.

Let $\mathcal{P}_{\sigma_2}$ denote the class of all probability measures that can be defined on σ_2. A possibilistic probability $\tilde{P}$ is called *representable* [11] if there is a (second-order)[2] normal possibility distribution $\boldsymbol{\pi} : \mathcal{P}_{\sigma_2} \to [0,1]$ that represents $\tilde{P}$, i.e., such that for all $p \in [0,1]$ and $A \in \sigma_2$,

$$\tilde{P}(A)(p) = \sup\{\boldsymbol{\pi}(Q) \,:\, Q \in \mathcal{P}_{\sigma_2},\ Q(A) = p\}.$$

3 Different interpretations for fuzzy random variables in the literature

Fuzzy sets have been given different interpretations [16], therefore a fuzzy random variable admits various meanings. In the remaining part of this section, we are going to review two interpretations of fuzzy random variables outlined in the introduction of this paper. For each interpretation, we will describe the information provided by the fuzzy random variable by means of a different model. It will be shown that these interpretations lead to a classical probability model, and a second-order possibility model, respectively.

3.1 Linguistic random variables

As we pointed out in the introduction, Puri and Ralescu ([40]) claim that the observations of some random experiments do not consist of numerical outputs, but are represented by vague linguistic terms. According to this idea, the fuzzy random variable, $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ is a measurable function, in the classical sense, between certain σ-algebra of events in the original space, Ω_1, and a σ-algebra[3] defined over a class of fuzzy subsets in Ω_2, $\boldsymbol{\sigma}$. In this context, the probability distribution induced by the fuzzy random variable can be used to summarize the probabilistic information that the variable provides us. We will denote by $P_1 \circ \tilde{X}^{-1}$ the probability measure induced by $\tilde{X}$ on $\boldsymbol{\sigma}$, i.e.:

$$P_1 \circ \tilde{X}^{-1}(\mathcal{F}) := P_1(\{\omega \in \Omega_1 \,:\, \tilde{X}(\omega) \in \mathcal{F}\}),\ \forall \mathcal{F} \in \boldsymbol{\sigma}.$$

Within this framework, we can use the tools of general Probability Theory to extend classical concepts and results, by reproducing classical techniques. The following example illustrates this approach. It has been taken from [40].

[2]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).

[3]The $\sigma-$algebra $\sigma_{\mathcal{F}}$ considered in the last section represents a particular case of this situation.

Example 3.1. *Consider a person who is questioned about the weather in a particular city some winter day chosen at random. Some possible answers would be "cold", "more or less cold", "very cold", "extremely cold", and so on. A natural question arising with reference to this example are: What is the probability that the answer is "very / extremely cold" (In other words, which is the proportion of winter days where he would answer "very cold" or "extremely cold")? We can answer this question by means of the induced probability distribution.*

Unfortunately, this kind of mathematical model is not useful to represent the imprecise observations of the outcomes of a random experiment [22, 37, 38, 39, 41, 42, 43, 44, 45, 46, 47], since it provides numerical (crisp) probabilities for fuzzy events, that do not reflect ill (imprecise) knowledge about the probabilities of crisp events. In the next subsection, we will overview an alternative model proposed by Kruse and Meyer in [24].

3.2 Ill-known classical random variables

Kruse and Meyer [24] choose a possibilistic interpretation of fuzzy sets. Each fuzzy set is viewed as modeling incomplete knowledge about an otherwise precise value. These authors claim that the fuzzy random variable represents imprecise or vague knowledge about a *classical* random variable, $X_0 : \Omega_1 \to \Omega_2$, to which they refer to as the "original random variable." Therefore, the membership degree of a point x to the fuzzy set $\tilde{X}(\omega)$ will represent the possibility degree of the assertion

> "The true image of element ω, $X_0(\omega)$, coincides with x."

Furthermore, they define the "acceptability degree" of a random variable $X : \Omega_1 \to \Omega_2$ as the value:

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

The function "acc" can be regarded as the possibility distribution associated to a possibility measure, $\Pi_{\tilde{X}}$, defined over the set of all random variables. The acceptability degree $\mathrm{acc}(X)$ represents the possibility degree of X being the true random variable that models the studied experiment. If the fuzzy random variable were a multi-valued mapping (if its images were crisp subsets of Ω_2) the acceptability function would assign the value 1 to certain class of measurable mappings (the class of measurable selections of the random set), and the value 0 to the remaining ones. When, more in particular, the images of the frv are singletons, the acceptability function would express completely knowledge about the actual random variable that models the experiment.

Under this framework, we can build (see [7, 8]) a possibility measure over the set of all the probability distributions in Ω_2, $\mathbf{\Pi}_{\tilde{X}}$. The possibility distribution, $\boldsymbol{\pi}_{\tilde{X}}$, that characterizes such possibility measure is defined as follows:

$$\boldsymbol{\pi}_{\tilde{X}}(Q) = \sup\{\mathrm{acc}(X) \,:\, P_1 \circ X^{-1} = Q\} =$$
$$\Pi_{\tilde{X}}(\{X : \Omega_1 \to \Omega_2 \text{ measurable} \,:\, P_1 \circ X^{-1} = Q\}).$$

$\boldsymbol{\pi}_{\tilde{X}}(Q)$ represents the degree of possibility that the original random variable is one of those that induce the probability distribution Q on σ_2. The possibility measure $\mathbf{\Pi}_{\tilde{X}}$ is a "second-order possibility" formally equivalent to those considered in [13]. It is so called, because it is a possibility distribution defined over a set of probability measures. It is a *representation* ([13]) of the possibilistic probability $\tilde{P}_{\tilde{X}} : \sigma_2 \to \mathcal{F}([0,1])$, defined as follows:

$$\tilde{P}_{\tilde{X}}(A)(p) = \sup\{\boldsymbol{\pi}_{\tilde{X}}(Q) \,:\, Q \in \mathcal{P}_{\sigma_2},\ Q(A) = p\},\ \forall\, p \in [0,1],\ A \in \sigma_2.$$

The relationships between $\mathbf{\Pi}_X$ and $\tilde{P}_{\tilde{X}}$ are explained in detail in [8]. This possibilistic probability is called [8] the *fuzzy possibility assignation associated to $\tilde{X}$*. The value $\tilde{P}_{\tilde{X}}(A)(p)$ represents the degree of possibility that the true probability of the event A is p.

Example 3.2. *A person is asked about the weather in a particular city in a winter day chosen at random. He has a thermometer to measure the temperature, but it has some imprecision. So, in a particular day, he is able to make assertions like the following one:*

> *I am sure that the actual temperature is between 5 and 11 ^oC. Furthermore, with probability greater than or equal to 0.9, it is between 7 and 9 ^oC (I know that at least the 90% times my observations have a +/- 1 ^oC of precision.)*

According to [6], it can be represented by means of the fuzzy set

$$\pi(x) = \begin{cases} 1 & \text{if } x \in [7,9] \\ 0.1 & \text{if } x \in [5,7) \cup (9,11] \\ 0 & \text{otherwise.} \end{cases}$$

Assume he can give this kind of information every day in winter. A natural question is: What is the probability that the (true) temperature is higher *than 8 ^oC? (In other words, which is the proportion of winter days where the temperature is over 8 ^oC?) We can answer this question by means of a second-order possibility measure, which will represent our knowledge about this probability value by means of a fuzzy set. In [8] we give detailed explanations about how to describe this imprecise knowledge.*

4 An additional interpretation: conditional possibility measure

Also in accordance with the possibilistic interpretation of fuzzy sets, in this work we are going to proceed in a slightly different way, to describe the information provided by $\tilde{X}$. As we suggested in the introduction of this paper, suppose that we have partial information about the probability distribution that models a sequence of two random experiments whose sample spaces are Ω_1 and Ω_2, respectively. On the one hand, we assume that the probability distribution that models the first one, $P_1 : \sigma_1 \to [0,1]$, is completely determined (in the preceding expression, σ_1 denotes a σ-algebra of events over Ω_1.) On the other hand, the connection with the second sub-experiment is only known via a family of conditional possibility measures $\{\Pi(\cdot \mid \omega)\}_{\omega\in\Omega_1}$, each of them inducing the fuzzy set $\tilde{X}(\omega)$. More specifically:

- The marginal probability $P_1 : \sigma_1 \to [0,1]$ is completely known.
- There exists a transition probability that models the relationship between the outcomes of both sub-experiment, $P_1^2 : \sigma_2 \times \Omega_1 \to [0,1]$, i.e.:
  - $P_1^2(\cdot,\omega)$ is a probability measure, $\forall\, \omega \in \Omega_1$.
  - $P_1^2(B,\cdot)$ is $\sigma_1 - \beta_{[0,1]}$ measurable.
  - Our imprecise knowledge about P_1^2 is determined by the following inequalities:

$$P_1^2(B,\omega) \le \Pi(B|\omega) = \sup_{b\in B} \tilde{X}(\omega)(b),\ \forall\, \omega \in \Omega_1,\ \forall\, B \in \sigma_2.$$

The family of possibility measures $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega_1}$ models our knowledge about the relationship between the outcome of the first sub-experiment and the possible outcomes of the second one. (If the result of the first experiment is ω, then the possibility degree of x happening in the second one is $\tilde{X}(\omega)(x)$.) In other words, we know the probability measure that drives the primary random process but the measurement process of outcomes is tainted with uncertainty. The combination of both sources of information, will allow us to describe the available information about the probability distribution on σ_2 (the probability distribution that rules the second sub-experiment) by means of an upper probability (a standard imprecise probability model, not an order 2 model, like the one described in Section 3.2.) In fact, the probability measure that rules the joint experiment, is given by the formula:

$$P(C) = \int_{\Omega_1} P_1^2(C_\omega, \omega)\, dP_1(\omega),\ \forall\, C \in \sigma_1 \otimes \sigma_2,$$

$$\text{where } C_\omega = \{x \in \Omega_2 \,:\, (\omega, x) \in C\}.$$

Hence, the probability measure associated to the second sub-experiment is given by:

$$P_2(B) = P(\Omega_1 \times B) = \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega),\ \forall\, B \in \sigma_2.$$

Thus, according to our knowledge about P_1^2, all we know about the probability $P_2(B_0)$ is that it belongs to the class:

$$\left\{ \int_{\Omega_1} P_1^2(B_0, \omega)\, dP_1(\omega) \,:\, P_1^2(B, \omega) \leq \Pi(B|\omega),\ \forall\, \omega \in \Omega_1,\ \forall\, B \in \sigma_2 \right\} \quad (1)$$

In the following section, we will provide a general method to compute the upper and lower bounds of this set. We will also illustrate situations that match this interpretation of fuzzy random variables and we will calculate upper and lower probabilities in a particular example. This way, we will be able to state assertions like the following: "the probability of the outcome between 3 and 7 lies between 0.3 and 0.6."

Remark 4.1. *This interpretation of fuzzy random variables has something to do with the Kruse & Meyer approach, because both approaches are based on the possibilistic interpretation of fuzzy sets. But, let us emphasize the differences between them.*

According to the Kruse & Meyer approach (Section 3.2), the mapping $X_0 : \Omega_1 \to \Omega_2$ *represents a deterministic conditional probability:*

$$P(\{X_0(\omega)\})|\omega) = 1, \forall\, \omega \in \Omega_1.$$

Furthermore, our imprecise knowledge about such deterministic probability measure is described by means of a (second-order) possibility measure. (The *possibility degree that the deterministic probability $P(\cdot,\omega)$ is focussed on r is $\tilde{X}(\omega)(r)$.) In this second-order model, "belief degrees about the occurrence of events" are distinguished from "belief degrees about the values of the probability of events", and both types of uncertainty stay in two different levels.*

On the contrary, in the present model, the underlying conditional probability $P_1^2(\cdot,\omega)$ is not assumed to be deterministic and it is only known as restricted (dominated) by the possibility measure $\Pi(\cdot|\omega)$. (The same idea is suggested in [33] for the particular case of random sets.)

We refer the reader to [3], where a realistic example illustrating this new interpretation is given. It involves weight measurements with a noisy scale. The scale is assumed to be under control the 90% times. Those times, we can guarantee a precision of 10g. The remaining times, we can only assure a precision of 50g. Our knowledge about the actual weight of an object taken at random can be represented by means of a frv which is interpreted under Kruse & Meyer approach (ill-known random variable). The actual weight is a fixed (ill-know) for a fixed object. On the other side, if we pick the same object again, our (observed) measurement can change. We can describe our knowledge about possible future measurements by means of a frv interpreted as a conditional possibility measure.

5 Upper-lower probability model

In this section we will show that the last interpretation of a frv leads, in a natural way, to a first-order imprecise model. Thus, our knowledge about the probability of a crisp event will be given by a pair of upper and lower probabilities.

According to the model described in the last subsection, our information about the probability of occurrence of the event B_0 in the second subexperiment is described by the set of values of Equation 1. Thus, the most informative upper and lower bounds are, respectively:

$$\overline{P}(B_0) = \sup_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(B_0,\omega)\, dP_1(\omega) \text{ and } \underline{P}(B_0) = \inf_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(B_0,\omega)\, dP_1(\omega) \tag{2}$$

where

$$\mathcal{H} = \{P_1^2 \text{ transition probability } : P_1^2(B,\omega) \leq \Pi(B|\omega),\ \forall\, \omega \in \Omega_1,\ \forall\, B \in \sigma_2\}. \tag{3}$$

We shall call respectively $\overline{P}(B_0)$ and $\underline{P}(B_0)$ the *upper and lower probabilities of* B_0. As a particular case of Theorem 1 in [5], the following result holds:

Theorem 5.1. *Consider a probability space* $(\Omega_1, \sigma_1, P_1)$*, the Borel* σ*-field on* $\mathbb{R}^n$*,* $\beta_{\mathbb{R}^n}$*, and a fuzzy random variable* $\tilde{X} : \Omega_1 \to \mathcal{F}(\mathbb{R}^n)$*. Let* $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega_1}$ *denote the family of possibility measures associated to* $\tilde{X}$*, i.e.*

$$\Pi(B|\omega) = \sup_{b\in B} \tilde{X}(\omega)(b), \ \forall\, B \in \beta_{\mathbb{R}^n}, \ \omega \in \Omega_1.$$

Consider the class:

$$\mathcal{H} = \{P_1^2 \text{ transition probability } : \ P_1^2(B,\omega) \leq \Pi(B|\omega), \ \forall\, \omega \in \Omega_1, \ \forall\, B \in \beta_{\mathbb{R}^n}\}.$$

Then,

$$\overline{P}(B) = \sup_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(B,\omega)\, dP_1(\omega) = \int_{\Omega_1} \Pi(B|\omega)\, dP_1(\omega).$$

It means that $\int_\Omega \Pi(B,\omega)\, dP_1(\omega)$ is the smallest upper bound we can give to the probability of B, taking into account the information provided by P_1 and $\tilde{X}$. Although the theorem is focussed on the upper bound, we can establish a similar result with respect to the lower bound. That is, if we consider the family of conjugate necessity measures, $\{N(\cdot|\omega)\}_{\omega\in\Omega}$, we can easily check that

$$\underline{P}(B) = \inf_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(B,\omega)\, dP_1(\omega) = \int_{\Omega_1} N(B|\omega)\, dP_1(\omega).$$

Next we will give some remarks concerning the result given in the last theorem.

Remark 5.1. *When, in particular,* $\tilde{X}$ *takes a finite number of different fuzzy images,* $\pi_1, \ldots, \pi_r$*, with respective probabilities* $p_1, \ldots, p_r$*, the upper and lower probabilities of an event* B *are given by:*

$$\overline{P}(B) = \sum_{i=1}^{n} p_i\, \Pi_i(B), \ \text{and } \underline{P}(B) = \sum_{i=1}^{n} p_i\, N_i(B), \tag{4}$$

where Π_i *and* N_i *are the dual possibility and the necessity measures associated to the fuzzy set* $\pi_i, i = 1, \ldots, r$*. Hence, at least in this particular case,* $\overline{P}$ *and* $\underline{P}$ *can be written as functions of the (classical) probability distribution of* $\tilde{X}$ *on* $\mathcal{F}(\Omega_2)$*. If, furthermore,* Ω_2 *is finite, we can give an alternative expressions for* $\overline{P}(B)$ *and* $\underline{P}(B)$*. In fact, for each* $i \in \{1, \ldots, r\}$*, let*

$m_i : \wp(\Omega_2) \to [0,1]$ be the Möbius transform of Π_i (the basic mass assignment associated to it) and let $\mathcal{F}_i = \{A_{i1}, \ldots, A_{ik_i}\}$ be the family of focal sets. We will introduce the following notation:

$$m_{ij} := m_i(A_{ij}), \text{ and } \nu_{ij} := p_i\, m_{ij},\ \forall\, j = 1, \ldots, k_i,\ i = 1, \ldots, r.$$

Let us now define the basic mass assignment: $m : \wp(\Omega_2) \to [0,1]$ as follows:

$$m(A_{ij}) := \nu_{ij},\ j = 1, \ldots, k_i,\ i = 1, \ldots, r.$$

(It is well defined, since $\sum_{i=1}^{r} \sum_{j=1}^{k_i} \nu_{ij} = 1$.) We can easily check that $\overline{P}$ and $\underline{P}$ are, respectively the plausibility and belief measures associated to m.

In Sections 6 and 7, we will check that these results can be extended for infinite universes: On the one hand, we will observe that $\overline{P}$ and $\underline{P}$ can be written as functions of the (classical) probability measure induced by $\tilde{X}$. On the other hand, we will check that they coincide with the upper and lower probabilities associated to a multi-valued mapping. Furthermore, we will extend the result given in Theorem 5.1 to more general cases.

In the following example, we illustrate the ideas given in Sections 4 and 5. The situation described is in accordance with our new interpretation of fuzzy random variables. We will illustrate that the upper and lower probabilities given in Equation 4 represent the most accurate upper and lower bounds for the probabilities of occurrence of events.

Example 5.1. *A person tosses a dice, and then he writes a number in a piece of paper, which is related to the result of the dice, $i \in \{1, \ldots, 6\}$. You have imprecise information about the number he writes:*

- *You are completely sure that it is one of the numbers $i-1$, i or $i+1$.*
- *It coincides with i with probability higher or equal than 0.5.*

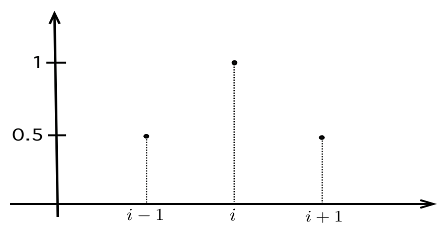
(Suppose for instance that he drops a coin after the dice is tossed. If the result is heads, he writes the same number of the dice, i. Otherwise, he chooses one of the numbers $\{i-1, i, i+1\}$ by a completely unknown procedure.) In other words, all you know about the conditional probability $P_i = P(\cdot|i)$ is that

$$P_i(\{i-1, i, i+1\}) = 1 \text{ and } P_i(\{i\}) \geq 0.5.$$

According to [6], the probability measures P_i satisfying the above restrictions are the probability measures dominated by the possibility measure $\Pi_i : \wp(\{0, \ldots, 7\}) \to [0,1]$ determined by $\pi_i(i-1) = \pi_i(i+1) = 0.5,\ \pi_i(i) = 1$.

($\pi_i(x)$ represent the possibility that he writes the number x if he observes the result i in the dice.)

Summarizing, our information (before the dice is tossed) about the number he is going to write in the paper is determined by the fuzzy random variable $\tilde{X} : \{1, \ldots, 6\} \to \wp(\{0, \ldots, 7\})$ given by: $\tilde{X}(i) = \pi_i, \ \forall\, i = 1, \ldots, 6$.



How can we describe our (imprecise) information about the probability of occurrence of an arbitrary event $A \subseteq \{0, \ldots, 7\}$? This probability is $P(A) = \sum_{i=1}^{6} \frac{1}{6} P_i(A)$, but we only have partial information about each P_i. All we know is that each one of them is dominated by the possibility measure Π_i. So, we know that

$$\sum_{i=1}^{6} \frac{1}{6} N_i(A) \leq P(A) \leq \sum_{i=1}^{6} \frac{1}{6} \Pi_i(A).$$

Furthermore, for each A we can find two families of probability measures, $\{P_i\}_{i=1}^{6}$ and $\{Q_i\}_{i=1}^{6}$ such that:

$$N_i(B) \leq P_i(B) \leq \Pi_i(B), \ \forall\, B \subseteq \{0, \ldots, 7\}, \ \forall\, i = 1, \ldots, 6,$$

$$N_i(B) \leq Q_i(B) \leq \Pi_i(B), \ \forall\, B \subseteq \{0, \ldots, 7\}, \ \forall\, i = 1, \ldots, 6,$$

$$P_i(A) = N_i(A), \text{ and } Q_i(A) = \Pi_i(A), \ \forall\, i = 1, \ldots, 6.$$

Thus, $\sum_{i=1}^{6} \frac{1}{6} N_i(A)$ and $\sum_{i=1}^{6} \frac{1}{6} \Pi_i(A)$ are the most accurate bounds for $P(A)$ according to our imprecise information.

We will calculate, for instance, the bounds for the probability of the event $A = \{1, 2\}$. We first compute the quantities $\Pi_i(A)$ and $N_i(A) = 1 - \Pi_i(\overline{A}), \forall\, i = 1, \ldots, 6$. We can easily check that:

$$\Pi_i(A) = \begin{cases} 1 & \text{if } \ i\text{=}1,2 \\ 0.5 & \text{if } \ i\text{=}3 \\ 0 & \text{if } \ i\text{=}4,5,6 \end{cases} \qquad \Pi_i(\overline{A}) = \begin{cases} 0.5 & \text{if } \ i\text{=}1,2 \\ 1 & \text{if } \ i\text{=}3,4,5,6 \end{cases}$$

$$\text{Hence, } \underline{P}(A) = \sum_{i=1}^{6} \frac{1}{6} N_i(A) = \sum_{i=1}^{6} \frac{1}{6}[1 - \Pi_i(\overline{A})] = 1/6$$

$$\text{and } \overline{P}(A) = \sum_{i=1}^{6} \frac{1}{6} \Pi_i(A) = 5/12.$$

Thus, the probability that the number he is going to write belongs to $A = \{1, 2\}$ *lies between 1/6 and 5/12.*

Remark 5.2. *According to the "classical" probability model, we can say that the frv* $\tilde{X}$ *takes each fuzzy value* $\pi_1, \ldots, \pi_6$ *with probability 1/6. Thus* $\underline{P}$ *and* $\overline{P}$ *can be written as functions of the probability distribution induced by* $\tilde{X}$ *on* $\mathcal{F}(\Omega_2)$*, since*

$$\overline{P}(A) = \sum_{i=1}^{6} \frac{1}{6} \sup_{a \in A} \pi_i(a) \text{ and}$$

$$\underline{P}(A) = 1 - \overline{P}(A^c), \ \forall\, A \subseteq \{0, \ldots, 7\}.$$

Nevertheless, this "classical" probability distribution does not explicitly represent our (imprecise) information about the probability of each (crisp) event.

Remark 5.3. *Assume that, when the result of the coin is tails, then three possible numbers* $i-1$*,* i *and* $i+1$ *are equiprobable. In that case,* P_i *should satisfy the equalities* $P_i(\{i-1\}) = P_i(\{i-1\}) = 1/6$ *and* $P_i(\{i\}) = 2/3$*, and hence the probability* P *should be known with total precision. It should be given by:*

$$\begin{array}{llll} P(\{0\}) = 1/36 & P(\{1\}) = 5/36 & P(\{2\}) = 1/6 & P(\{3\}) = 1/6 \\ P(\{4\}) = 1/6 & P(\{5\}) = 1/6 & P(\{6\}) = 5/36 & P(\{7\}) = 1/36. \end{array}$$

But this assumption is unsupported by the actual information, and it should reflect additional artificial information. The procedure followed to choose one of the numbers $i-1$*,* i *or* $i+1$ *is completely unknown. The person could decide to write, for instance, the number* $i-1$ *every time. In that case, the respective probabilities of the possible outcomes would be different from above:*

$$\begin{array}{llll} P(\{0\}) = 1/12 & P(\{1\}) = 1/6 & P(\{2\}) = 1/6 & P(\{3\}) = 1/6 \\ P(\{4\}) = 1/6 & P(\{5\}) = 1/6 & P(\{6\}) = 1/12 & P(\{7\}) = 0. \end{array}$$

6 Multi-valued mapping associated to a frv

As we pointed out in Remark 5.1, when the universe of the second sub-experiment is finite, the upper and lower probabilities considered in Equation 2 (beginning of Section 5) are, respectively a plausibility and a belief measure. Furthermore, we recall in Theorem 5.1 that they can be calculated as the average (with respect to the probability measure P_1 defined on Ω_1) of the family of possibility measures $\{\Pi(\cdot|\omega)\}_{\omega\in\Omega_1}$, when the final space is $\mathbb{R}^n$. Our concerns in this section are threefold:

1. We will extend Theorem 5.1 and check that the equalities

$$\sup_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(B,\omega)\, dP_1(\omega) = \int_{\Omega_1} \Pi(B|\omega)\, dP_1(\omega),\ \forall\, B \in \sigma_2$$

and

$$\inf_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(B,\omega)\, dP_1(\omega) = \int_{\Omega_1} N(B|\omega)\, dP_1(\omega),\ \forall\, B \in \sigma_2.$$

also hold for more general topological structures, which will be listed in Corollary 6.4. Thus, the bounds

$$\overline{P}(\cdot) = \sup_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(\cdot,\omega)\, dP_1(\omega) \text{ and } \underline{P}(\cdot) = \inf_{P_1^2\in\mathcal{H}} \int_{\Omega_1} P_1^2(\cdot,\omega)\, dP_1(\omega)$$

will be proved to be ∞-order capacities under fairly general conditions.

2. We will also check the equalities

$$\int_{\Omega_1} \Pi(B|\omega)\, dP_1(\omega) = \int_0^1 P^*_{\tilde{X}_\alpha}(B)\, d\alpha \text{ and}$$

$$\int_{\Omega_1} N(B|\omega)\, dP_1(\omega) = \int_0^1 P_{*\,\tilde{X}_\alpha}(B)\, d\alpha.$$

(The right hand side expressions will be used in Section 7).

3. In order to extend Theorem 5.1, we will explicitly define a multi-valued mapping, Γ, whose (Dempster-)upper and lower probabilities do coincide with the above integrals, and we will make use of some recent results about the equality between the supremum of the set of probabilities associated to the measurable selections of a random set and its upper probability. Another concern in the section is providing Γ with a meaningful interpretation within this new interpretation of fuzzy random variables.

Let the probability space $(\Omega_1, \sigma_1, P_1)$ represent an arbitrary random experiment. Consider a measurable space (Ω_2, σ_2) and a fuzzy random variable $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ that represents our available information about the connection between the outcome in Ω_1 and the class of possible outcomes in Ω_2. According to the interpretation of fuzzy random variables described in Section 4, for a fixed $\omega_0 \in \Omega$, $\tilde{X}(\omega_0)$ represents a possibility measure, $\Pi(\cdot|\omega_0) : \sigma_2 \to [0,1]$. Hence, for each $\omega \in \Omega_1$, we assume that there exists a transition probability $P_1^2 : \sigma_2 \times \Omega_1 \to [0,1]$ such that

$$N(B|\omega) \leq P_1^2(B, \omega) \leq \Pi(B|\omega),\ \forall\, B \in \sigma_2,\ \omega \in \Omega_1,\ \text{where}$$

$$\Pi(B|\omega) = \sup_{b \in B} \tilde{X}(\omega)(b),\ \text{and}\ N(B|\omega) := 1 - \Pi(B^c|\omega),\ \forall\, B \in \sigma_2, \omega \in \Omega_1.$$

We can alternatively express the above information by means of a multi-valued mapping Γ_{ω_0}. Consider the usual Borel $\sigma-$algebra $\beta_{[0,1]}$ on the unit interval $[0,1]$, and the uniform distribution defined on it, $\lambda : \beta_{[0,1]} \to [0,1]$ and let $\Gamma_{\omega_0} : [0,1] \to \wp(\Omega_2)$ be the multi-valued mapping defined as follows:

$$\Gamma_{\omega_0}(\alpha) = \tilde{X}_\alpha(\omega_0),\ \forall\, \alpha \in [0,1].$$

According to [12], Γ_{ω_0} is $\beta_{[0,1]} - \wp(\Omega_2)$ strongly measurable, and its upper probability coincides with $\Pi(B|\omega_0)$, i.e.,

$$\lambda(\{\alpha \in [0,1] \,:\, \Gamma_{\omega_0}(\alpha) \cap B \neq \emptyset\}) = \Pi(B|\omega_0).$$

We must recall that, for a fixed $\omega_0 \in \Omega$, the information provided by $\Pi(\cdot|\omega_0)$ can be expressed by means of a nested family of confidence intervals. (We do not refer to the usual notion of (random) confidence interval in statistics. A $1 - \alpha$ confidence interval is defined here as a (fixed) interval whose probability is known to be lower bounded by $1-\alpha$.) In fact, according to [6], the family of probability measures $P_1^2(\cdot, \omega_0) : \sigma_2 \to [0,1]$ satisfying the restriction $P_1^2(B, \omega_0) \leq \Pi(B|\omega_0),\ \forall\, B \in \sigma_2$ coincides with the set of probability measures[4]:

$$P_1^2(\tilde{X}_\alpha(\omega_0)|\omega_0) \geq 1 - \alpha,\ \forall\, \alpha \in (0,1). \tag{5}$$

In our context, Equation 5 is interpreted as follows:

> The probability that the outcome of the second experiment belongs to $\tilde{X}_\alpha(\omega_0)$, knowing that the outcome of the first one has been ω_0, is greater than or equal to $1 - \alpha$, for each α.

[4] Provided that $\tilde{X}_\alpha(\omega_0) \in \sigma_2,\ \forall\, \alpha, \omega_0$.

For any fixed α^*, the set of α's such that $\Gamma_{\omega_0}(\alpha)$ is included in $X_{\alpha^*}(\omega_0)$ contains (possibly strictly) the interval $[\alpha, 1]$. So, the images of Γ_{ω_0} are contained in $\tilde{X}_{\alpha^*}(\omega_0)$ with probability greater than or equal to $1-\alpha^*$. Thus, the above information about the outcome of the second experiment can be expressed by means of an additional experiment ruled on the unit interval: we take at random an $\alpha \in [0, 1]$ (according to the uniform distribution) and we observe an object belonging to the set $\Gamma_{\omega_0}(\alpha) = \tilde{X}_\alpha(\omega_0)$.

We have expressed above our information about the outcome of the second experiment, when the outcome of the first one is ω_0. But such outcome is indeed taken at random, according to the probability measure P_1. So it seems natural to express our information by means of the multi-valued mapping $\Gamma : \Omega_1 \times [0, 1] \to \wp(\Omega_2)$ defined as

$$\Gamma(\omega, \alpha) := \tilde{X}_\alpha(\omega) = \{x \in \Omega_2 \,:\, \tilde{X}(\omega)(x) \geq \alpha\},\ \forall (\omega, \alpha) \in \Omega_1 \times [0, 1].$$

This multi-valued mapping univocally determines $\tilde{X}$, since the family of $\alpha-$cuts of a fuzzy sets univocally determines it. It is an alternative way to express the same information. Let now $P_1 \otimes \lambda$ denote the product probability, i.e., the unique probability measure on $\sigma_1 \otimes \beta_{[0,1]}$ satisfying the equalities:

$$(P_1 \otimes \lambda)(A \times B) = P_1(A) \cdot \lambda(B),\ \forall\, A \in \sigma_1,\ B \in \beta_{[0,1]}.$$

In the rest of this section, we will show that the upper probability induced by Γ coincides with the upper probability $\overline{P}$ defined in Equation 2, under fairly general conditions. This means that the following equality holds for any $B \in \sigma_2$:

$$P_1 \otimes \lambda(\{(\omega, \alpha) \in \Omega_1 \times [0, 1] \,:\, \Gamma(\omega, \alpha) \cap B\}) = \sup_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega).$$

So, our information about the probability that rules the second experiment can be expressed alternatively by means of Γ : we take at random an element $\omega \in \Omega_1$ (according to the probability measure P_1) and, independently, some $\alpha \in [0, 1]$ (according to the uniform distribution). The subset $\Gamma(\omega, \alpha) \subseteq \Omega_2$ will represent our information about the outcome of the second experiment.

We point out in Lemma 6.1 that Γ is strongly measurable, so it induces an upper probability on σ_2.

Lemma 6.1. *Let (Ω_1, σ_1) and (Ω_2, σ_2) be two measurable spaces. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ a fuzzy random variable. Let $([0, 1], \beta_{[0,1]})$ represent the unit interval with the usual Borel $\sigma-$algebra. Consider the product $\sigma-$algebra*

$\sigma_1 \otimes \beta_{[0,1]}$ and the multi-valued mapping $\Gamma : \Omega_1 \times [0,1] \to \wp(\Omega_2)$ defined as follows:

$$\Gamma(\omega, \alpha) := \tilde{X}_\alpha(\omega),\ \forall\, (\omega, \alpha) \in \Omega \times [0,1].$$

Then Γ is $\sigma_1 \otimes \beta_{[0,1]} - \sigma_2$ strongly measurable.

Next we give two alternative ways for the calculation of the upper probability induced by Γ.

Lemma 6.2. *Let (Ω_1, σ_1) and (Ω_2, σ_2) be two measurable spaces. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable. Let $([0,1], \beta_{[0,1]})$ represent the unit interval with the usual Borel $\sigma-$ algebra. Let us now consider the product $\sigma-$algebra $\sigma_1 \otimes \beta_{[0,1]}$ and the multi-valued mapping $\Gamma : \Omega_1 \times [0,1] \to \wp(\Omega_2)$ considered in Lemma 6.1. Then*

$$P^*_\Gamma(B) = \int_{\Omega_1} \Pi(B|\omega)\, dP_1(\omega) = \int_0^1 P^*_{\tilde{X}_\alpha}(B)\, d\alpha.$$

According to the last lemma, the Dempster upper probability $P^*_\Gamma(B)$ can be alternatively calculated as the expectations of the random variables $\omega \hookrightarrow \Pi(B|\omega)$ or $\alpha \hookrightarrow P^*_{\tilde{X}_\alpha}(B)$, with respect to the probability measures P_1 and λ, respectively. So, we average the upper probabilities of the random sets Γ_ω with respect to P_1 (first case) and the α-level-wise upper probabilities of the random sets $\tilde{X}_\alpha$ (second case). We interpret such calculations as follows: in the first case, we average the possibility of occurrence B, conditional to each possible outcome of the first sub-experiment. In the second case, we average the probability of not discarding B, conditional to the appearance of the (randomly selected) point $\alpha \in [0,1]$.

In Theorem 6.3, we will point out that the class of probability measures associated to the measurable selections of Γ is included in the family of integrals $\left\{\int_{\Omega_1} P_1^2(\cdot, \omega)\, dP_1(\omega) \,:\, P_1^2 \in \mathcal{H}\right\}$, which is in fact the family of probabilities that represents our knowledge about the outcomes of the random experiment on Ω_2.

Theorem 6.3. *Let $(\Omega_1, \sigma_1, P_1)$ be a probability space and let (Ω_2, σ_2) be a measurable space. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable. Let $([0,1], \beta_{[0,1]})$ represent the unit interval with the usual Borel $\sigma-$ algebra. Consider now the product $\sigma-$algebra $\sigma_1 \otimes \beta_{[0,1]}$, and the product probability $P_1 \otimes \lambda$. Let $\Gamma : \Omega_1 \times [0,1] \to \wp(\Omega_2)$ be the multi-valued mapping considered in Lemma 6.1. Then*

$$\mathcal{P}(\Gamma)(B) \subseteq \left\{ \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega) \,:\, P_1^2 \in \mathcal{H} \right\},$$

where $\mathcal{H}$ is the class of transition probabilities defined in Equation 3.

From the above result, we derive that the upper and lower bounds

$$\overline{P}(B) = \sup\left\{\int_{\Omega_1} P_1^2(B,\omega)\,dP_1(\omega) \;:\; P_1^2 \in \mathcal{H}\right\}$$

and

$$\underline{P}(B) = \inf\left\{\int_{\Omega_1} P_1^2(B,\omega)\,dP_1(\omega) \;:\; P_1^2 \in \mathcal{H}\right\}$$

satisfy the following inequalities:

$$\sup \mathcal{P}(\Gamma)(B) \leq \sup\left\{\int_{\Omega_1} P_1^2(B,\omega)\,dP_1(\omega) \;:\; P_1^2 \in \mathcal{H}\right\} \leq \int_{\Omega_1} \Pi(B|\omega)\,dP_1(\omega)$$

and

$$\int_{\Omega_1} N(B|\omega)\,dP_1(\omega) \leq \inf\left\{\int_{\Omega_1} P_1^2(B,\omega)\,dP_1(\omega) \;:\; P_1^2 \in \mathcal{H}\right\} \leq \inf \mathcal{P}(\Gamma)(B).$$

Thus, in those cases where $\sup \mathcal{P}(\Gamma)$ coincides with P_Γ^*, those inequalities become into equalities. It has been pointed out in [34] that such supremum is, in general, a maximum, but it does necessarily coincide with P_Γ^*. Sufficient conditions for the equality between P_Γ^* and $\max \mathcal{P}(\Gamma)$ are given in [30] and [34] and [35]. Based on that, we can derive the following corollary from Lemma 6.2, and Theorems 6.3:

Corollary 6.4. *Let (Ω_1, σ_1) a measurable space. Let (Ω_2, τ_2) be a topological space and let σ_2 be the Borel $\sigma-$ algebra associated to it. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable. If any of the following conditions hold:*

- *τ_2 is induced by a metric separable space and $\tilde{X}_\alpha(\omega)$ is open for all $\omega \in \Omega_1$ and all $\alpha \in [0,1]$.*
- *τ_2 is induced by a $\sigma-$compact metric space and $\tilde{X}_\alpha(\omega)$ is closed for all $\omega \in \Omega_1$ and all $\alpha \in [0,1]$.*
- *(Ω_2, τ_2) is a Polish space and $\tilde{X}_\alpha(\omega)$ is closed for all $\omega \in \Omega_1$ and all $\alpha \in [0,1]$.*
- *(Ω_2, τ_2) is induced by a metric separable space and $\tilde{X}_\alpha(\omega)$ is compact for all $\omega \in \Omega_1$ and all $\alpha \in [0,1]$.*

- *$[x] \in \sigma_2,\ \forall\, x \in \Omega_2$ and $\{A^* : A \in \sigma_2\}$ is numerable, where A^* denotes the set:*

$$A^* = \{(\omega, \alpha) \in \Omega_1 \times [0,1] \,:\, \tilde{X}_\alpha(\omega) \cap A \neq \emptyset\},\ \forall\, A \in \sigma_2$$

and $[x]$ denotes the intersection $[x] = \cap\{A \in \sigma_2 \mid x \in A\}$.

- *$[x] \in \sigma_2,\ \forall\, x \in \Omega_2$ and the class $\{\tilde{X}_\alpha(\omega) \,:\, (\omega, \alpha) \in \Omega_1 \times [0,1]\}$ is countable.*

Then, for an arbitrary event $B \in \sigma_2$, the upper probability $\overline{P}(B) = \sup_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega)$ is, in fact, a maximum, and the lower probability $\underline{P}(B) = \inf_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega)$ is a minimum. Furthermore, the following equalities hold:

$$\max_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega) = \int_{\Omega_1} \Pi(B|\omega)\, dP_1(\omega) = \int_0^1 P^*_{\tilde{X}_\alpha}(B)\, d\alpha$$

$$\min_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega) = \int_{\Omega_1} [1 - \Pi(B^c|\omega)]\, dP_1(\omega) = \int_0^1 P_{*\,\tilde{X}_\alpha}(B)\, d\alpha.$$

Remark 6.1. *Corollary 6.4 generalizes the result recalled in Theorem 5.1 to cases where the final space is not necessarily $\mathbb{R}^n$.*

Remark 6.2. *Assume that σ_1 and τ_2 fulfill some of the above hypotheses, and consider the particular case where $\tilde{X}$ represents a multi-valued mapping. In other words, assume that, for each $\omega \in \Omega_1$, the fuzzy set $\tilde{X}(\omega)$, considered as a membership function, only takes the values 0 and 1, so it can be identified with a crisp subset of Ω_2, $\tilde{X}(\omega) \subseteq \Omega_2$. In this case, the upper and lower probabilities defined in Equation 2 coincide with Dempster's upper and lower probabilities associated to $\tilde{X}$, when we consider it as a multi-valued mapping.*

7 Relationships with other models in the literature

In this section we will show the connections between the upper-lower model described above (Sections 5 and 6) and the other precedent models for fuzzy random variables from the literature. In particular we will observe that our upper-lower model can be written as a function of the classical model (Section 3.1) and also of the second-order imprecise model (Section 3.2).

7.1 Relationships between the upper-lower probability model and the classical model

As we checked in Section 2, a fuzzy random variable can be viewed as a standard measurable function. Let us first consider, for each $B \in \sigma_2$ and each $\alpha \in [0,1]$, the family of sets:

$$\mathcal{F}_B^\alpha = \{F \in \mathcal{F}(\Omega_2) \,:\, F_\alpha \cap B \neq \emptyset\}.$$

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

$$\{\mathcal{F}_B^\alpha, \,:\, B \in \sigma_2, \alpha \in [0,1]\} \subseteq \wp(\mathcal{F}(\Omega_2)).$$

As we pointed out in Section 2, we can easily check that a fuzzy-valued mapping is a frv if and only if it is $\sigma_1 - \sigma_{\mathcal{F}}$ measurable. So, we can consider the probability measure it induces over $\sigma_{\mathcal{F}}$, $P_1 \circ \tilde{X}^{-1}$. Now, we will check that our upper and lower probabilities (Equation 2) can be written as functions of $P_1 \circ \tilde{X}^{-1}$ under fairly general conditions.

Theorem 7.1. *Let (Ω_1, σ_1) a measurable space. Let (Ω_2, τ_2) be a topological space and let σ_2 be the Borel $\sigma-$algebra associated to it. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable with closed α-cuts. Let $P_1 \circ \tilde{X}^{-1}$ represent the probability measure it induces on $\sigma_{\mathcal{F}}$. Let $\underline{P}$ and $\overline{P}$ the upper and lower probabilities given by:*

$$\underline{P}(B) = \inf_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega), \quad \overline{P}(B) = \sup_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega)$$

Under any of the conditions considered in Corollary 6.4, $\underline{P}$ and $\overline{P}$ can be written as functions of $P_1 \circ \tilde{X}^{-1}$. Specifically,

$$\overline{P}(B) = \int_0^1 P_1 \circ \tilde{X}^{-1}(\{\Gamma \,:\, \Gamma_\alpha \cap B \neq \emptyset\}\, d\alpha)$$

and

$$\underline{P}(B) = \int_0^1 P_1 \circ \tilde{X}^{-1}(\{\Gamma \,:\, \Gamma_\alpha \subseteq B\}\, d\alpha).$$

Remark 7.1. *Consider the particular case where $\tilde{X}$ is a random set and let $B \in \sigma_2$ be an arbitrary event. Under fairly general conditions, the upper probability defined in Equation 2, $\overline{P}(B) = \sup_{P_1^2 \in \mathcal{H}} \int_{\Omega_1} P_1^2(B, \omega)\, dP_1(\omega)$ coincides with Dempster's upper probability, $P^*_{\tilde{X}}(B)$, according to Remark 6.2.*

As we have recalled in the preliminaries section, Dempster's upper probability univocally determines the probability induced by the random set. Thus, we conclude that the upper probability $\overline{P}$ determines the probability measure induced by $\tilde{X}$ on $\sigma_{\mathcal{F}}$, $P_1 \circ \tilde{X}^{-1}$. Summarizing, when the images of the fuzzy random variables are, in particular, crisp sets, the upper probability, $\overline{P}$, and the induced probability measure, $P_1 \circ \tilde{X}^{-1}$ univocally determine each other. Furthermore, when the second space, Ω_2 is finite, $P \circ \tilde{X}^{-1}$ coincides with the basic probability assignment, while $\overline{P}$ coincides with the plausibility measure.

7.2 Relationships between the upper-lower probability model and the second-order possibility model

According to Walley [48], any second order possibility model can be reduced into a first-order imprecise model, by means of natural extension techniques. To adapt these ideas to our particular situation, let us replace linear previsions by probability measures, and the sets of bounded real functions by σ–algebras of subsets of the referential. Walley considers the product space $\mathcal{P}_{\sigma_2} \times \Omega_2$ and he assumes that the following items are available:

- A (second-order) possibility measure, $\Pi\!\Pi$, on the first space $\mathcal{P}_{\sigma_2}$. (It indicates degrees of possibility over probabilities.)
- The transition probability $I\!P_2^1 : \sigma_2 \times \mathcal{P}_{\sigma_2} \to [0,1]$ given by the formula: $I\!P_2^1(B,P) := P(B),\ \forall B \in \sigma_2, P \in \mathcal{P}_{\sigma_2}$. (It indicates the following information: if the probability P rules the experiment on Ω_2, then the probability of the occurrence of B is $P(B)$.

In this setting, he constructs, by means of natural extension techniques, an upper-lower joint model. Thus, the available information about the marginal distribution on the second space is described, in a natural way, by a pair of lower and upper probabilities:

$$\overline{P}_W(B) = \sup_{I\!P \leq \Pi\!\Pi} \int I\!P_2^1(B,P)\ dI\!P(P) = \sup_{I\!P \leq \Pi\!\Pi} \int P(B)\, dI\!P(P) \quad \text{and}$$

$$\underline{P}_W(B) = \inf_{I\!P \leq \Pi\!\Pi} \int I\!P_2^1(B,P)\ dI\!P(P) = \inf_{I\!P \leq \Pi\!\Pi} \int P(B)\, dI\!P(P)$$

Furthermore, Walley proves that these upper and lower probabilities can be alternatively calculated as follows:

$$\overline{P}_W(B) = \int_0^1 \overline{P}_\alpha(B)\, d\alpha, \quad \underline{P}_W(B) = \int_0^1 \underline{P}_\alpha(B)\, d\alpha, \tag{6}$$

where, for each index, $\alpha \in [0,1]$, $\overline{P}_\alpha$ and $\underline{P}_\alpha$ are defined as follows:

$$\overline{P}_\alpha(B) = \sup\{Q(B) \,:\, Q \in \mathcal{P}, \amalg(\{Q\}) \geq \alpha\} \text{ and}$$
$$\underline{P}_\alpha(B) = \inf\{Q(B) \,:\, Q \in \mathcal{P}, \amalg(\{Q\}) \geq \alpha\}.$$

Next, we will consider the second-order possibility measure derived from Kruse & Meyer approach ("ill-known classical random variable"), $\amalg_{\tilde{X}}$, . We will prove that, if we apply Walley's reduction to this specific second-order possibility (making use of the transition probability $I\!P_2^1$ mentioned at the beginning of this section), the resulting upper-lower pair coincides with the upper-lower model associated to the third interpretation of a frv, i.e., the model described in Sections 5 and 6.

Theorem 7.2. *Let (Ω_1, σ_1) be a measurable space. Let (Ω_2, τ_2) be a topological space and let σ_2 be the Borel $\sigma-$algebra associated to it. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable. Let $\tilde{P}_{\tilde{X}}(B_0)$ represent the fuzzy possibility assignation of the event B_0. Let $\overline{P}(B_0)$ and $\underline{P}(B_0)$ be the upper and lower probabilities defined in Equation 2. Under the hypotheses in Corollary 6.4, the following equalities hold:*

$$\overline{P}(B_0) = \overline{P}_W(B_0) \text{ and } \underline{P}(B_0) = \underline{P}_W(B_0) \tag{7}$$

Remark 7.2. *According to the last theorem, we can follow two alternative ways to build the same upper-lower model.*

1. *In the first case (upper-lower model suggested in Equation 2) we consider:*
   - *A probability measure defined over the first space Ω_1.*
   - *A conditional possibility measure, $\Pi(\cdot|\cdot)$ determined by the frv. It represents the imprecise perception of a transition probability P_1^2. This transition probability models the random relationship between the outcome of the first experiment, $\omega \in \Omega_1$ and the outcome in the second space, Ω_2. For each $\omega \in \Omega$, our imprecise knowledge about $P_1^2(\cdot, \omega)$ is determined by the possibility measure $\Pi(\cdot|\omega)$ associated to the fuzzy set $\tilde{X}(\omega)$.*
2. *In the second case (Walley's reduction of the second-order model) we consider the product space $\mathcal{P}_{\sigma_2} \times \Omega_2$ and:*
   - *A (second-order) possibility measure on $\mathcal{P}_{\sigma_2}$, that characterizes the knowledge about the probability measure induced by the ill-known random variable.*

- *A transition probability measure $P_2^1 : \sigma_2 \times \mathcal{P}_{\sigma_2} \to [0,1]$. For each pair $(B, \mathbb{P}) \in \sigma_2 \times \mathcal{P}_{\sigma_2}$ $\mathbb{P}(B,P)$ expresses the following information: if P were the true probability measure induced by the ill-known random variable, the probability of occurrence of B would be $P(B)$.*

Since any possibility measure can be represented by a class of probability measures, we can represent both situations by means of upper-lower models. Accoding to the last theorem, both procedures lead to the same pair of upper and lower probabilities.

The upper and lower probabilities given in Equation 2 are also related to the concept of mean value of a fuzzy set. Let us recall this concept. Dubois and Prade [15] define the "mean value" of a fuzzy set, π, as the interval:

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

where $E(P)$ represents the expected value associated to the probability measure P, and Π is the possibility measure associated to the possibility distribution π. That interval represents the set of possible values for the expectation of the outcome of a certain random experiment, when we only know that the probability measure that models it is dominated by Π. Next we will prove that the upper probability defined in Equation 2 coincides with the upper bound of the mean value of the fuzzy probability envelope of B_0. Based on this lemma and Corollary 6.4, we can prove the following result.

Theorem 7.3. *Let (Ω_1, σ_1) a measurable space. Let (Ω_2, τ_2) be a topological space and let σ_2 be the Borel $\sigma-$algebra associated to it. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable with closed α-cuts. Let $\tilde{P}_{\tilde{X}}(B_0)$ represent the fuzzy possibility assignation of the event B_0. Let $\overline{P}(B_0)$ and $\underline{P}(B_0)$ be the upper and lower probabilities defined in Equation 2. Under the hypotheses in Corollary 6.4, the following equalities hold:*

$$\overline{P}(B_0) = \sup M(\tilde{P}_{\tilde{X}}(B_0)) \text{ and } \underline{P}(B_0) = \inf M(\tilde{P}_{\tilde{X}}(B_0)) \tag{8}$$

Remark 7.3. *The idea behind Walley's reduction is closely related to the concept of mean value of a fuzzy set, when this fuzzy set represents the fuzzy probability determined by a second-order possibility measure. On the one hand, the upper and lower probabilities associated to Walley's reduction are determined as follows:*

$$\overline{P}_W(B) = \sup_{\mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}} \int \mathbb{P}_2^1(B, P)\ d\mathbb{P}(P) = \sup_{\mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}} \int P(B)\, d\mathbb{P}(P) \quad \text{and}$$

$$\underline{P}_W(B) = \inf_{\mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}} \int \mathbb{P}_2^1(B, P)\ d\mathbb{P}(P) = \inf_{\mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}} \int P(B)\, d\mathbb{P}(P)$$

Let us now consider, for an arbitrary $B \in \sigma_2$*, the random variable* $X_B : \mathcal{P}_{\sigma_2} \to [0,1]$ *defined as follows:*

$$X_B(P) = \mathbb{P}_2^1(B, P) = P(B),\ \forall\, P \in \mathcal{P}_{\sigma_2}.$$

This random variable represents the "probability of occurrence of B*". Under the second-order model, such probability value is viewed as a random quantity. Walley's reduction combines both kind's of probabilities into the same model. Thus, if the "true" meta-probability is* $\mathbb{P}$*, then the probability of occurrence of* B *is computed as*

$$\int \mathbb{P}_2^1(B, P)\ d\mathbb{P}(P) = \int X_B(P)\, d\mathbb{P}(P) = E_{\mathbb{P}}(X_B).$$

Thus, Walley's upper and lower probabilities can be written as follows:

$$\overline{P}_W(B) = \sup_{\mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}} \int X_B(P)\, d\mathbb{P}(P) \ \text{and}$$

$$\underline{P}_W(B) = \inf_{\mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}} \int X_B(P)\, d\mathbb{P}(P).$$

Thus, Walley upper and lower probabilities represent the tightest upper and lower bounds for the probability of occurrence of B *under the Kruse and Meyer interpretation, when both kinds of probabilities (second-order and standard probabilities) are combined into the same model.*

The set-function $\mathbb{\Pi}$ *is, in particular, an order 2 alternating capacity, and hence, the above suprema can be written as Choquet integrals:*

$$\overline{P}_W(B) = (C) \int X_B(P)\, d\mathbb{\Pi}(P) \ \text{and}$$

$$\underline{P}_W(B) = (C) \int X_B(P)\, d\mathrm{Nec}_{\tilde{X}}(P),$$

where Nec is the dual of the second-order possibility $\mathbb{\Pi}$*.*

The above formulae do coincide with the supremum and the infimum of the mean value of the fuzzy set $\tilde{P}_{\tilde{X}}(B)$. To clarify this point, let us consider the possibility measure, $\Pi'_B = \mathbb{\Pi}_{\tilde{X}} \circ X_B^{-1}$, defined as follows:

$$\Pi'_B(C) = \mathbb{\Pi}_{\tilde{X}}(X_B^{-1}(C)) = \mathbb{\Pi}_{\tilde{X}}(\{P \in \mathcal{P}_{\sigma_2} \,:\, X_B(P) \in C\}) =$$

$$\mathbb{\Pi}_{\tilde{X}}(\{P \in \mathcal{P}_{\sigma_2} \,:\, P(B) \in C\}).$$

We can easily check that it coincides with the possibility associated to the fuzzy set $\tilde{P}_{\tilde{X}}(B)$. In fact,

$$\mathbb{\Pi}_{\tilde{X}}(\{P \in \mathcal{P}_{\sigma_2} \,:\, P(B) \in C\}) = \sup_{\{P \in \mathcal{P}_{\sigma_2} \,:\, P(B) \in C\}} \boldsymbol{\pi}_{\tilde{X}}(P) =$$

$$\sup_{x \in C} \sup_{\{P \in \mathcal{P}_{\sigma_2} \,:\, P(B) = x\}} \boldsymbol{\pi}_{\tilde{X}}(P) = \sup_{x \in C} \tilde{P}_{\tilde{X}}(B)(x) = \Pi_B(C).$$

Furthermore, the above Choquet integrals can be written as follows:

$$\overline{P}_W(B) = (C) \int X_B(P)\, d\mathbb{\Pi}(P) = (C) \int x\, d\Pi_B(x) \text{ and}$$

$$\underline{P}_W(B) = (C) \int X_B(P)\, d\mathrm{Nec}_{\tilde{X}}(P) = (C) \int x\, d\mathrm{Nec}_B(x).$$

By definition, these integrals coincide with the supremum and the infimum of the mean value of the fuzzy set $\tilde{P}_X(B)$.

Thus, we follow parallel procedures in both approaches. On the one hand, according to Walley reduction, each second-order probability, $\mathbb{P}$ is identified with the standard probability $P_{\mathbb{P}}$, defined as

$$P_{\mathbb{P}}(B) = \int X_B \, d\mathbb{P}, \ \forall\, B \in \sigma_2.$$

Then we consider the pair of upper and lower probabilities on σ_2 associated to the class

$$\{P_{\mathbb{P}} \,:\, \mathbb{P} \leq \mathbb{\Pi}_{\tilde{X}}\}.$$

On the other hand, to take the mean value of a fuzzy probability, we first fix an event $B \in \sigma_2$, and consider the fuzzy probability assignation $\tilde{P}_{\tilde{X}}(B)$. It is a fuzzy set, so it is associated to a possibility measure on the unit interval, Π_B. We identify each probability measure $P_B \leq \Pi_B$ with the quantity $E(P_B)$ Then we consider the upper and lower bounds of the class of such expectations,

$$\{E(P_B) \,:\, P_B \leq \Pi_B\}.$$

In both cases, we convert second-order probabilities into standard probability measures.

8 Concluding remarks

Most studies in the literature about fuzzy random variables have a formal foundation in (classical) Probability Theory. 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 $\mathcal{B}$. Nevertheless, since randomness and vagueness coexist in the same framework, it seems reasonable to integrate fuzzy random variables into imprecise probabilities theory. In [8] we present a second-order possibility model that represents the imprecise information provided by a fuzzy random variable about the underlying probability distribution. We show there that the classical probability measure induced by the frv on $\mathcal{B}$ does not univocally determine this second-order possibility.

According to Kyburg [25] and Walley [48] any second-order probability can be re-interpreted as a first-order one. Those authors claim that second-order and standard probabilities can be combined into a joint model, despite they express different kinds of beliefs. (Let us recall that first-order models represent degrees of belief about occurrence of events, while second-order ones concern statements of first-order probabilities.) Thus, the second-order possibility model introduced in [8] can be converted into a first-order model. According to [48] this model can be then used as prior information for decision making and statistical reasoning. First-order models produce intervals, instead of fuzzy sets, so they are easier to manage from a computational point of view. Furthermore, in this paper we have proven that the upper and lower probabilities obtained from second-order possibility model, when Walley's reduction is applied are in fact ∞-order capacities. Moreover, they are the natural representation of an additional interpretation of fuzzy random variables.

But, as a counterpart, this computational simplification has a price. In fact, our upper and lower probabilities can be written as functions of the probability measure induced by the frv, when it is regarded as a classical measurable mapping. This means that they do not univocally determine the second-order possibility measure associated to the frv. So, we can find two different fuzzy random variables inducing different second possibilities, but with the same upper-lower model. As we show in [8], these differences on their second-order possibility measures are relevant when we want to represent the information that the frv provides about certain parameters such as the variance or the entropy of the underlying random variable (the original random variable under the Kruse and Meyer's interpretation). So, when should we use the lower-upper model and when should we work with the second-order possibility measure? The upper-lower model can be used under two different kinds of situations:

(a) On the one hand, it is the natural representation when the frv is interpreted as a family of conditional probability measures. This interpretation is based on the possibilistic interpretation of fuzzy sets but it differs from the Kruse and Meyer approach: under their approach, the outcome of the second sub-experiment is univocally determined by the result of the first sub-experiment. Thus, $\tilde{X}(\omega)(x)$ represents the degree of possibility that the outcome associated to ω coincides with x. On the contrary, under the new interpretation, the result of the second sub-experiment is not determined by ω. The quantity $\tilde{X}(\omega)(x)$ represents the degree of possibility of x occurring if the result of the first experiment is ω.

(b) Even when the Kruse & Meyer approach is considered, the upper and lower model is useful. Assume that we can combine both kinds of beliefs (about events and about probabilities of events) and, furthermore, we want to give a range of values for a certain parameter that can be written as a linear function of the probability measure (e.g. the probability of an event or the expectation). Then we should use the first-order model. On the contrary, when we want to give the range of values for non-linear parameters as the variance or the entropy, the first-order model misses relevant information as well as the induced probability measure does.

Future works would apply these ideas to making decisions in the presence of vague data. Some recent papers in the literature are devoted to testing hypothesis on the basis of fuzzy samples (see, for instance, [14, 18, 20]). Those papers follow the Kruse and Meyer approach on fuzzy random variables, and use Zadeh's extension principle to define "fuzzy decisions". In particular, Filzmoser and Viertl [18] and Denoeux et al. [14] independently generalize the concept of set-valued critical level to the case of fuzzy data and they define the fuzzy critical level. But they claim that a defuzzification is needed to take a crisp decision. According to our first-order model, we can apply natural extension techniques to get an interval-valued generalization of the concept of p-value, in the presence of fuzzy sample data. Thus, we will be able to follow the natural approach described in [17], to reach a crisp decision: when the upper and lower bounds of the p-value are on one side of the $\alpha-$ significance level, the decision of the hypothesis test is clear. But when the bounds straddle the threshold, the test is inconclusive, since the imprecision in the data prevents us to make a clear determination. We plan to compare this approach with the defuzzifications of fuzy tests proposed in those papers.

Acknowledgements

This work has been supported by MEC-FEDER grants TIN2008-06681-C06-04 and TIN2010-20900-C04-01. This financial support is gratefully acknowledged. We thank two anonymous the referees for their detailed comments and corrections that helped us to significantly improve the presentation and content.

References

[1] C. Baudrit, I. Couso and D. Dubois, Joint propagation of probability and possibility in risk analysis: Towards a formal framework, Int. Journal of Approximate Reasoning 45 (2007) 82-105.

[2] P. Billingsley, Probability and Measure, Wiley, New York, 1979.

[3] I. Couso and D. Dubois, On the variability of the concept of variance for fuzzy random variables, IEEE Transactions on Fuzzy Systems 17 (2009) 1070-1080.

[4] I. Couso, D. Dubois, S. Montes, L. Sánchez, On various definitions of the variance of a fuzzy random variable, Proccedings of ISIPTA07, Prague, Czech Republic, 2007, 135–144.

[5] I. Couso, E. Miranda and G. de Cooman, A possibilistic interpretation of the expectation of a fuzzy random variable, in: M. López-Díaz, M. A. Gil, P. Grzegorzewski, O. Hyrniewicz, and J. Lawry (Eds.), Soft methodology and random information systems, Springer, Heidelberg, 2004, 133–140.

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

[7] I. Couso, S. Montes and P. Gil, Second-order possibility measure induced by a fuzzy random variable, in: C. Bertoluzza, M. A. Gil and

D. A. Ralescu (Eds.), Statistical modeling, analysis and management of fuzzy data, Physica-Verlag, Heidelberg, 2002, 127-144.

[8] I. Couso and L. Sánchez, Higher order models for fuzzy random variables, Fuzzy Sets and Systems 159 (2008) 237–258.

[9] I. Couso and L. Sánchez, Mark-recapture techniques in statistical tests for imprecise data, International Journal In Approximate Reasoning, d.o.i.: 10.1016/j.ijar.2010.07.009.

[10] I. Couso, L. Sánchez and P. Gil, Imprecise distribution function associated to a random set, Information Sciences, 159 (2004) 109–123.

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

[12] G. de Cooman and D. Aeyels, Supremum preserving upper probabilities, Information Sciences, 118 (1999) 173–212.

[13] G. de Cooman and P. Walley, An imprecise hierarchical model for behaviour under uncertainty, Theory and Decision, 52 (2002) 327–374.

[14] T. Denoeux, M. H. Masson and P.A. Hebert, Nonparametric Rank-Based Statistics and Significance Test for Fuzzy Data, Fuzzy Sets and Systems, 153 (2005) 1–28.

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

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

[17] S. Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf and L. Ginzburg, Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty, Technical Report, SAND2007-0939, 2007.

[18] P. Filzmoser and R. Viertl, Testing hypotheses with fuzzy data: The fuzzy p-value, Metrika 59 (2004) 21–29.

[19] M. Grabisch, H. T. Nguyen and E. A. Walker. Fundamentals of uncertainty calculi with applications to fuzzy inference, Kluwer Academic Publishers, Dordrecht, 1995.

[20] P. Grzegorzewski, Testing statistical hypotheses with vague data, Fuzzy Sets and Systems 112 (2000) 501–510.

[21] J.A. Herencia, Graded sets and points: A stratified approach to fuzzy sets and points, Fuzzy Sets and Systems 77 (1996) 191–202.

[22] M. J. del Jesus, F. Hoffmann, L. Junco and L. Sánchez, Induction of fuzzy-rule-based classifiers with evolutionary boosting algorithms, IEEE Transactions on Fuzzy Systems 12 (2004) 296–308.

[23] R. Körner, Linear models with random fuzzy variables, Thesis, Technische Universität Bergakademie Freiberg, 1997.

[24] R. Kruse, K.D. Meyer, Statistics with vague data, D. Reidel Publishing Company, 1987.

[25] H.E. Kyburg, Higher order probabilities and intervals, Int. J. of Approximate Reasoning 2 (1988) 195–209.

[26] E. Miranda, I. Couso, P. Gil, Upper probabilities and selectors of random sets. in: P. Gzegorzewski, O. Hryniewicz, M.A. Gil (Eds.), Soft methods in probability, statistics and data analysis, Physica-Verlag, Heidelberg, Germany, 2002.

[27] E. Miranda, I. Couso, P. Gil, Study of the probabilistic information of a random set, Proceedings of the 3rd ISIPTA Conference, Lugano, Switzerland, 2003.

[28] E. Miranda, I. Couso and P. Gil, Relationships between possibility measures and nested random sets, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10 (2002) 1–15.

[29] E. Miranda, I. Couso and P. Gil, A random set characterisation of possibility measures, Information Sciences, 168 (2004) 51–75.

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

[31] E. Miranda, I. Couso and P. Gil, Random intervals as a model for imprecise information, Fuzzy Sets and Systems 154 (2005) 386–412.

[32] E. Miranda, I. Couso and P. Gil, Approximation of upper and lower probabilities by measurable selections. Information Sciences 180 (2010), 1407-1417.

[33] E. Miranda, G. de Cooman and I. Couso, Imprecise probabilities induced by multi-valued mappings, J. Stat. Plann. Inference 133 (2005) 173–197.

[34] E. Miranda, I. Couso, P. Gil, Upper Probabilities Attainable by Distributions of Measurable Selections, in: C. Sossai, G. Chemello (Eds.) Symbolic and Quantitative Approaches to Reasoning with Uncertainty (LNAI 5590), Springer, 2009, 335-346.

[35] E. Miranda, I. Couso, P. Gil, Approximation of upper and lower probabilities by measurable selections, Information Sciences 180 (2010) 1407-1417.

[36] H.T. Nguyen, On random sets and belief functions, J. Math. Analysis and Applications 63 (1978) 531–542.

[37] A. Palacios, L. Sánchez, I. Couso, Extending a simple Genetic Cooperative-Competitive learning fuzzy classifier to low quality datasets, Evolutionary Intelligence 2 (2009) 73-84.

[38] A.M. Palacios, L. Sánchez and I. Couso, Diagnosis of dyslexia with low quality data with genetic fuzzy systems, International Journal of Approximate Reasoning 51(2010) 993-1009.

[39] A. Palacios, L. Sánchez, I. Couso, Future performance modelling in athletism with low quality data-based GFSs, Journal of Multivalued Logic and Soft Computing (accepted).

[40] M.L. Puri and D. Ralescu, Fuzzy Random Variables, J. Math. Anal. Appl. 114 (1986) 409–422.

[41] L. Sánchez and I. Couso, Advocating the use of Imprecisely Observed Data in Genetic Fuzzy Systems, IEEE Trans. Fuzzy Systems 15 (2007) 551–562.

[42] L. Sánchez, I. Couso Obtaining Fuzzy Rules from Interval Censored Data with Genetic Algorithms and a Random Sets-based Semantic of the Linguistic Labels. Soft Computing, d.o.i.: 10.1007/s00500-010-0627-6.

[43] L. Sánchez, I. Couso and J. Casillas, Genetic learning of fuzzy rules based on low quality data, Fuzzy Sets and Systems 160 (2009) 2524–2552.

[44] L. Sánchez, M.R. Suárez, I. Couso. A fuzzy definition of Mutual Information with application to the design of Genetic Fuzzy Classifiers. International Conference on Machine Intelligence (ACIDCA-ICMI05). Tozeur (Tunisia, 2005) 602–609.

[45] L. Sánchez, M.R. Suárez, J.R. Villar and I. Couso, Some Results about Mutual Information-based Feature Selection and Fuzzy Discretization of Vague Data, Proceedings of the 16th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE07), London, United Kingdom, 2007, 1958–1963.

[46] L. Sánchez, M.R. Suárez, J. R. Villar and I. Couso, Mutual information-based feature selection and partition design in fuzzy rule-based classifiers from vague data, International Journal in Approximate Reasoning 49 (2008) 607–622.

[47] Villar, J. R., Otero, A., Otero, J., Sánchez, L. Taximeter verification with GPS and soft computing techniques. Soft Computing. 14(4) 405-418. (2009)

[48] P. Walley, Statistical inferences based on a second-order possibility distribution, International Journal of General Systems 26 (1997) 337–384.

9 Appendix

Proof of Lemma 6.1. Let $B \in \sigma_2$ an arbitrary measurable subset of Ω_2. We must check that $\Gamma^*(B) \in \sigma_1 \otimes \beta_{[0,1]}$. We will see that $\Gamma^*(B)$ coincides with a numerable intersection $\cap_{n\in I\!N}\Gamma^*_n(B)$, where each $\Gamma^*_n(B)$ is a measurable set. For each $n \in I\!N$, consider the partition of the unit interval $\{I_n^1, \ldots, I_n^{2^n}\}$, where $I_n^i = [\frac{i-1}{2^n}, \frac{i}{2^n})$, $i = 1, \ldots, 2^n - 1$, $I_n^{2^n} = [\frac{2^n-1}{2^n}, 1]$. Let us now define the multi-valued mapping $\Gamma_n : \Omega_1 \times [0,1] \to \wp(\Omega_2)$ as follows:

$$\Gamma_n(\omega, \alpha) = \tilde{X}_{\frac{i-1}{2^n}}(\omega), \text{ if } (\omega, \alpha) \in \Omega \times I_n^i, \; i = 1, \ldots, 2^n.$$

Let us check that $\Gamma^*_n(B)$ belongs to $\sigma_1 \otimes \beta_{[0,1]}$. In fact, $\Gamma^*_n(B) = \{(\omega, \alpha) \in \Omega_1 \times [0,1] \,:\, \Gamma_n(\omega, \alpha) \cap B \neq \emptyset\} = \cup_{i=1}^{2^n} \tilde{X}^*_{\frac{i-1}{2^n}}(B) \times I_n^i$. Since $\tilde{X}^*_{\frac{i}{2^n}}(B) \in \sigma_1$ and $I_n^i \in \beta_{[0,1]}, \forall\, i = 1, \ldots, 2^n$, we derive that $\Gamma^*_n(B)$ belongs to the product $\sigma-$algebra $\sigma_1 \otimes \beta_{[0,1]}$. Now, it only remains to check the equality $\Gamma^*(B) = \cap_{n\in I\!N}\Gamma^*_n(B)$.

($\subseteq$) Let (ω, α) be an arbitrary element in $\Gamma^*(B)$. For each $n \in I\!N$, there exists an index i_n such that $\alpha \in I_n^{i(n)}$. And we observe that $\omega \in \tilde{X}^*_\alpha(B) \subseteq \tilde{X}^*_{\frac{i(n)-1}{2^n}}(B)$. We conclude that $\Gamma^*(B) \subseteq \cap_{n\in I\!N}\Gamma^*_n(B)$.

($\supseteq$) For an arbitrary $(\omega, \alpha) \in \cap_{n\in I\!N}\Gamma^*_n(B)$, there exists a sequence of indices $(i(n))_{n\in I\!N}$ such that $\omega \in \tilde{X}^*(B)_{\frac{i(n)-1}{2^n}}$ and $(\frac{i(n)-1}{2^n})_{n\in I\!N}$ upward converging to α : Thus, according to the properties of $\alpha-$cuts, $X^*(B)_\alpha = \cap_{n\in I\!N}X^*(B)_{\frac{i(n)-1}{2^n}}$. Hence, $\omega \in \tilde{X}^*_\alpha(B)$, or equivalently, $(\omega, \alpha) \in \Gamma^*(B)$. □

To check the result given in Lemma 6.2, the following lemma will be used. The proof is immediate.

Lemma 9.1. *Let (Ω_1, σ_1) and (Ω_2, σ_2) be two measurable spaces. Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ a fuzzy random variable. For each $\omega \in \Omega_1$, let $\Pi(\cdot|\omega_1) : \sigma_2 \to [0,1]$ denote the possibility measure associated to the fuzzy set $\tilde{X}(\omega)$, i.e., $\Pi(B|\omega) = \sup_{b\in\ B} \tilde{X}(\omega)(b),\ \forall\, B \in \sigma_2$. Then, the following equalities hold:*

$$\Pi(B|\omega) = \sup\{\alpha \in [0,1]\ :\ \tilde{X}_\alpha(\omega) \cap B \neq \emptyset\} = \lambda(\{\alpha \in [0,1]\ :\ \omega \in \tilde{X}^*_\alpha(B)\}).$$

Proof of Lemma 6.2. It is an immediate consequence of Fubini's theorem and Lemma 9.1.

Proof of Theorem 6.3. Consider an arbitrary measurable selection of Γ, $g \in S(\Gamma)$. Let us define the mapping $P_1^2 : \sigma_2 \times \Omega_1 \to [0,1]$ as

$$P_1^2(C, \omega) := \lambda(g_\omega^{-1}(C)) = \lambda([g^{-1}(C)]^2_\omega), \forall\, C \in \sigma_2,\ \forall\, \omega \in \Omega_1.$$

We will next prove that it belongs to the class $\mathcal{H}$.

- We observe that $P_1^2(\cdot, \omega)$ is a probability measure: In fact, it is the probability measure induced by the measurable function g_ω, $\forall\, \omega \in \Omega_1$.
- We can easily check $P_1^2(C, \cdot)$ is $\sigma_1 - \beta_{[0,1]}$ measurable, since it can be written as $P_1^2(C, \omega) = \lambda([g^{-1}(C)]^2_\omega)$, $\forall\, \omega \in \Omega$.
- Let us now check that $P_1^2(C, \omega) \leq \Pi(C|\omega),\ \forall\, \omega \in \Omega_1,\ \forall\, C \in \sigma_2$.

  In fact, $P_1^2(C, \omega) = \lambda(\{\alpha\ :\ g(\omega, \alpha) \in C\}) \leq \lambda(\{\alpha\ :\ \Gamma(\omega, \alpha) \cap C \neq \emptyset\})) = \lambda([\Gamma^*(C)]^1_\omega)$. Furthermore, $\lambda([\Gamma^*(C)]^1_\omega)$ coincides with $\Pi(C|\omega)$.

Finally, the equality $\int_\Omega P_1^2(B, \omega)\, dP_1(\omega) = (P_1 \otimes \lambda)(g^{-1}(B))$ is straightforward. □

Proof of Theorem 7.1. Under the conditions given in Corollary 6.4, the following equality holds:

$$\overline{P}(B) = \int_0^1 P^*_{\tilde{X}_\alpha}(B)\, d\alpha,\ \forall\, B \in \sigma_2.$$

Furthermore, for each $\alpha \in [0,1]$, the upper probability $P^*_{\tilde{X}_\alpha}$ can be written as a function of $P_1 \circ \tilde{X}^{-1}$. In fact:

$$P^*_{\tilde{X}_\alpha}(B) = P_1 \circ \tilde{X}^{-1}(\mathcal{F}^\alpha_B),\ \forall\, B \in \sigma_2,$$

$$\text{where } \mathcal{F}^\alpha_B = \{F \in \mathcal{F}(\Omega_2)\,:\, F_\alpha \cap B \neq \emptyset\}.$$

Hence, $\overline{P}$ can be written as a function of $P_1 \circ \tilde{X}^{-1}$. Furthermore, we can easily check that $\underline{P}$ satisfies the equalities $\underline{P}(B) = 1 - \overline{P}(B^c),\ \forall\, B \in \sigma_2$. Hence, it can be also written as a function of $P_1 \circ \tilde{X}^{-1}$. $\square$

The proof of Theorem 7.2 will be supported on these two lemmas. The proof of the first one is straightforward. The second one is given in [8].

Lemma 9.2. *Let $\pi : [0,1] \to [0,1]$ be an arbitrary fuzzy set on the unit interval. Let $\varphi : [0,1] \to \wp([0,1])$ be a graded set associated to it, i.e., satisfying*

$$\pi_{\overline{\alpha}} \subseteq \varphi(\alpha) \subseteq \pi_\alpha,\ \forall\, \alpha \in [0,1],$$

Then

$$\int_0^1 \sup \pi_\alpha d\alpha = \int_0^1 \sup \varphi(\alpha) d\alpha \text{ and}$$

$$\int_0^1 \inf \pi_\alpha d\alpha = \int_0^1 \inf \varphi(\alpha) d\alpha.$$

Lemma 9.3. *Let us consider a probability space $(\Omega_1, \sigma_1, P_1)$ and a measurable space (Ω_2, σ_2). Let $\tilde{X} : \Omega_1 \to \mathcal{F}(\Omega_2)$ be a fuzzy random variable. Let $\tilde{P}_{\tilde{X}}(B_0)$ represent the fuzzy possibility assignation of the event B_0. Then:*

$$[\tilde{P}_{\tilde{X}}(B)]_{\overline{\alpha}} \subseteq \mathcal{P}(\tilde{X}_\alpha)(B) \subseteq [\tilde{P}_{\tilde{X}}(B)]_\alpha,\ \forall\, B \in \sigma_2,\ \alpha \in [0,1].$$

Proof of Theorem 7.2. By Lemmas 9.2 and 9.3, the following equalities hold:

$$\int_0^1 \sup [\tilde{P}_{\tilde{X}}(B)]_\alpha\, d\alpha = \int_0^1 \sup \mathcal{P}(\tilde{X}_\alpha)(B)\, d\alpha$$

and

$$\int_0^1 \inf [\tilde{P}_{\tilde{X}}(B)]_\alpha\, d\alpha = \int_0^1 \inf \mathcal{P}(\tilde{X}_\alpha)(B)\, d\alpha.$$

Furthermore, according to Theorem [35]:

$$\sup \mathcal{P}(\tilde{X}_\alpha)(B) = P^*_{\tilde{X}_\alpha}(B) \text{ and } \inf \mathcal{P}(\tilde{X}_\alpha)(B) = P_{*\,\tilde{X}_\alpha}(B),\ \forall\, B \in \sigma_2.$$

By Corollary 6.4, the upper and lower probabilities of B can be calculated as follows:

$$\overline{P}(B) = \int_0^1 P^*_{\tilde{X}_\alpha}(B)\, d\alpha \text{ and } \underline{P}(B) = \int_0^1 P_{*\,\tilde{X}_\alpha}(B)\, d\alpha,\ \forall\, B \in \sigma_2.$$

Now, according to Equation 6, the upper and lower probabilities associated to Walley's reduction can be expressed as follows:

$$\overline{P}_W(B) = \int_0^1 \sup[\tilde{P}_{\tilde{X}}(B)]_\alpha\, d\alpha \text{ and}$$

$$\underline{P}_W(B) = \int_0^1 \inf[\tilde{P}_{\tilde{X}}(B)]_\alpha\, d\alpha,\ \forall\, B \in \sigma_2.$$

Thus, we can easily derive the result given in Equation 7. □

The following lemma will be used to prove Theorem 7.3.

Lemma 9.4. *[19] Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(\Omega', \mathcal{A}')$ be a measurable space and let $\Gamma : \Omega \to \mathcal{P}(\Omega')$ be a random set. For any bounded random variable $f : \Omega' \to \mathbb{R}$, let $f \circ \Gamma : \Omega \to \wp(\mathbb{R})$ denote the random set whose image is $f(\Gamma(\omega)) := \{f(\omega') : \omega' \in \Omega'\}$, for each $\omega \in \Omega$. Then, the following equality holds*

$$(C) \int_{\Omega'} f\, dP^* = \int_\Omega \sup(f \circ \Gamma)\, dP.$$

Proof of Theorem 7.3. Let us prove the first equality. We must notice that the supremum of $M(\tilde{P}_{\tilde{X}}(B_0))$ coincides with the Choquet integral of the identity function with respect to the possibility measure associated to the fuzzy set $\tilde{P}_{\tilde{X}}(B_0)$, i.e.

$$\sup M(\tilde{P}_{\tilde{X}}(B_0)) = (C) \int \mathrm{id}\, d\Pi_{B_0},$$

where Π_{B_0} denotes the possibility measure associated to the fuzzy set $\tilde{P}_{\tilde{X}}(B_0)$. This is so because any possibility measure is an order 2 alternating capacity. Furthermore, consider the nested random set $\Gamma_{B_0} : [0,1] \to \wp([0,1])$ defined as follows:

$$\Gamma_{B_0}(\alpha) = \{p \in [0,1] \,:\, \tilde{P}_{\tilde{X}}(B_0)(p) \geq \alpha\},\ \forall\, \alpha.$$

According to [12], its Dempster's upper probability coincides with the possibility measure associated to $\tilde{P}_{\tilde{X}}(B_0)$. Thus, according to Lemma 9.4, the following equalities hold:

$$\sup M(\tilde{P}_{\tilde{X}}(B_0)) = \int_0^1 \sup(\mathrm{id} \circ \Gamma_{B_0})\, d\alpha = \int_0^1 \sup \mathcal{P}(\tilde{X}_\alpha)(B_0)\, d\alpha.$$

According to Corollary 6.4 and Lemma 9.2, we get the following equality:

$$\int_0^1 \sup \mathcal{P}(\tilde{X}_\alpha)(B_0)\, d\alpha = \overline{P}(B_0),$$

and so the equality $\sup M(\tilde{P}_{\tilde{X}}(B_0)) = \overline{P}(B_0)$ is fulfilled. By duality, we can check the equality $\inf M(\tilde{P}_{\tilde{X}}(B_0)) = \underline{P}(B_0)$. $\square$