Defuzzification of fuzzy p-values

Inés Couso[1] and Luciano Sánchez[2]

Abstract We provide a new description of the notion of fuzzy p-value, within the context of the theory of imprecise probabilities. The fuzzy p-value is viewed as a representation of a certain second-order possibility measure. According to Walley, any second-order possibility measure can be converted into a pair of lower and upper probabilities. Thus, we can convert the fuzzy p-value into an interval in the real line. We derive a construction of imprecise (but non fuzzy) tests, which are formally similar to recent tests used to manage with set-valued data.

Key words: Imprecise probabilities, hypothesis testing, fuzzy p-value, second-order possibility measure.

1 Introduction

Uncertainty about measurements arises naturally in a variety of circumstances (see [7] for a detailed description). This is the reason why the development of procedures for hypothesis testing with imprecise observations has recently gained increasing attention. When the data set contains intervals rather than points, we are not always able to take a clear decision about the null hypothesis. In the recent literature, imprecise tests are proposed to deal with such situations (see [7], for instance). According to this approach, an interval of upper and lower bounds of the critical value can be computed from the data set. When both bounds are on one side of the significance level, the decision (reject or accept) is clear. But when that interval and the significance threshold do overlap, we are not allowed to take a decision. In such situations, multi-valued test functions are defined. They can take the values $\{1\}$ (reject), $\{0\}$ (accept) and $\{0,1\}$ (undecided). This idea has been extended to the case of fuzzy-valued samples, under different approaches. Specifically, Filtzmoser & Viertl [8]

[1] Dept. of Statistics and O.R., University of Oviedo, Oviedo, Spain
[2] Dept. of Computer Sciences, University of Oviedo, Gijón, Spain

and Denœux et al. [6] independently introduce the concept of fuzzy p-value. The concept of fuzzy test is then derived in a natural way by Denœux et al. [6]. But what should we do when a crisp decision is needed? They propose a particular defuzzification of the test output, in order to take a decision. Here we will propose an alternative construction, based on an interval-valued assignation for the critical level. We will justify why such defuzzification of the fuzzy p-value makes sense. We will show that it is in accordance with the possibilistic interpretation of fuzzy random variables developed in [3].

2 Fuzzy p-values and fuzzy tests

2.1 Fuzzy p-value associated to a fuzzy random sample

Let $X^* : \Omega \to \mathbb{R}$ be a random variable with distribution function F^* and let $\mathbf{X}^* = (X_1^*, \ldots, X_n^*) : \Omega^n \to \mathbb{R}^n$ be a simple random sample of size n from F^* (a collection of n iid random variables with common distribution F^*. They represent n independent observations of X^*.) Let now the Borel-measurable mapping $\varphi : \mathbb{R}^n \to \{0,1\}$ represent a non-randomized test for

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1.$$

Both hypotheses refer to a certain parameter of the df F^*. We will denote by R the critical region of φ, i.e., $R = \{\mathbf{x} \in \mathbb{R}^n : \varphi(\mathbf{x}) = 1\}$. Let $\sup_{\theta\in\Theta_0} E_\theta(\varphi(\mathbf{X})) = \sup_{\theta\in\Theta_0} P_\theta(\text{Reject } H_0)$ denote the size of the test φ. Suppose that for every $\alpha \in (0,1)$ we have a size α test φ_α with rejection region R_α and let $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ a realization of the sample. The p-value of $\mathbf{x}^*$ is defined as the quantity $\mathrm{p_{val}}(\mathbf{x}^*) = \inf\{\alpha : \mathbf{x}^* \in R_\alpha\}$.

Let us now assume that we have got imprecise information about $\mathbf{x}^*$, and such imprecise information is given by means of a fuzzy subset of $\mathbb{R}^n$, $\tilde{\mathbf{x}} \in \mathscr{F}(\mathbb{R}^n)$. According to the possibilistic interpretation of fuzzy sets[1], $\tilde{\mathbf{x}}(\mathbf{x})$ represents the possibility grade that the "true" realization $\mathbf{x}^*$ coincides with the vector $\mathbf{x}$. Denœux et al. [6] and Filzmoser & Viertl [8] independently extend the concept of p-value, introducing the notion of fuzzy p-value. Each of those papers deals with a specific problem, but both definitions lead to the same general notion. We will call the *fuzzy p-value* of the fuzzy sample $\tilde{\mathbf{x}}$ to the fuzzy set $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})$ determined by the membership function:

$$\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})(p) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : \exists\, \mathbf{x} \in \mathbb{R}^n, \text{ with } \mathrm{p_{val}}(\mathbf{x}) = p\},\ \forall\, p \in [0,1]. \tag{1}$$

According to the possibilistic interpretation of fuzzy sets, the membership $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})(p)$ represents the possibility grade of the equality $\mathrm{p_{val}}(\mathbf{x}^*) = p$, accord-

[1] We show in [2, 3] some specific situations where such a membership function is derived from an imprecise perception of some $\mathbf{x}^*$.

ing to the imprecise information we have about $\mathbf{x}^*$ described by $\tilde{\mathbf{x}}$. The last fuzzy set is closely related to the nested family of sets $(\mathrm{p_{val}}(\tilde{\mathbf{x}}_\delta))_{\delta\in[0,1]}$ defined as follows:

$$\mathrm{p_{val}}(\tilde{\mathbf{x}}_\delta) = \{\mathrm{p_{val}}(\mathbf{x}) : \mathbf{x} \in \tilde{\mathbf{x}}_\delta\},\ \forall\, \delta \in [0,1].$$

For some particular situations studied in [6] and [8], it is the family of $\delta-$cuts of $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})$. In the general case, it is just a gradual representation of the fuzzy p-value. In other words, the membership function of $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})$ can be derived from such nested family as follows:

$$\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})(p) = \sup\{\delta : p \in \mathrm{p_{val}}(\tilde{\mathbf{x}}_\delta)\}.$$

But we should assume some continuity properties to assure that $(\mathrm{p_{val}}(\tilde{\mathbf{x}}_\delta))_{\delta\in[0,1]}$ is the family of $\delta-$cuts. In general, only the following relation holds:

$$[\mathrm{p_{val}}(\tilde{\mathbf{x}})]_{\overline{\delta}} \subseteq \mathrm{p_{val}}(\tilde{\mathbf{x}}_\delta) \subseteq [\mathrm{p_{val}}(\tilde{\mathbf{x}})]_\delta,\ \forall\, \delta,$$

where $[\mathrm{p_{val}}(\tilde{\mathbf{x}})]_{\overline{\delta}}$ and $[\mathrm{p_{val}}(\tilde{\mathbf{x}})]_\delta$ respectively denote the strong and the weak $\delta-$cut.

2.2 Fuzzy test associated to the fuzzy p-value

First of all, let us specify the meaning of the expression "fuzzy test" in our context: The null and the alternative hypotheses are referred to the distribution of the original random variable, F^*, so they are customary hypotheses in usual statistical problems. But the test is a fuzzy-valued function, i.e., it is a mapping that assigns, to each possible fuzzy sample $\tilde{\mathbf{x}} \in \mathscr{F}(\mathbb{R}^n)$, a fuzzy subset of $\{0,1\}$. That fuzzy subset reflects the possibility grades of rejection and acceptance of the null hypothesis, in accordance with the information provided by the fuzzy random sample. Some recent papers in the literature about statistics with imprecise data fit this formulation (see [6], for instance.) Let the reader notice that this approach is not related to other different works in the fuzzy statistics literature (see [9] for a detailed description), where the test functions are crisp, but they are referred to a certain parameter of the probability distribution induced by a fuzzy random variable on a certain σ-algebra of fuzzy events. This approach would not be useful in our context, where the frv represents the imprecise description of an otherwise standard random variable (see [1, 3, 4] for more detailed comments.)

In this paper, we will follow Denœux et al. [6] to construct a fuzzy test from a fuzzy p-value function. They specify the calculations for the Kendall and the Mann-Whitey-Wilcoxon tests. We will give here a more general description.

Let $(\varphi_\alpha)_{\alpha\in(0,1)}$ be a family of tests for H_0 against H_1, where $\varphi_\alpha : \mathbb{R}^n \to \{0,1\}$ is a test of size α, for each $\alpha \in (0,1)$. Let $\mathrm{p_{val}} : \mathbb{R}^n \to [0,1]$ and $\widetilde{\mathrm{ext}}(\mathrm{p_{val}}) : \mathscr{F}(\mathbb{R}^n) \to \mathscr{F}([0,1])$ respectively denote the crisp and the fuzzy p-value functions, in accordance with the formulae given in the last section. We can construct the fuzzy test $\varphi_{\widetilde{\mathrm{ext}}(\mathrm{p_{val}})}$ from $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})$ as follows:

$$\varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})(1) = \sup\{\widetilde{\text{ext}}(\text{p}_{\text{val}})(\tilde{\mathbf{x}})(p) : p \leq \alpha\}, \text{ and}$$
$$\varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})(0) = \sup\{\widetilde{\text{ext}}(\text{p}_{\text{val}})(\tilde{\mathbf{x}})(p) : p > \alpha\}.$$

According to the interpretation of $\widetilde{\text{ext}}(\text{p}_{\text{val}})(\tilde{\mathbf{x}})(p)$, the membership value $\varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})(1)$ represents the possibility grade that $\text{p}_{\text{val}}(\mathbf{x}^*)$ is less than or equal to α or, in other words, the possibility that $\mathbf{x}^*$ belongs to the rejection region. Similarly, $\varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})(0)$ represents the possibility of accepting (no rejecting) the null hypothesis. Thus, $\varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})$ represents a fuzzy decision. In the cases where a crisp decision is needed, this fuzzy subset may be defuzzified. Denœux et al. [6] suggest the following rule: rejecting the null hypothesis whenever $\varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})(1) > \varphi_{\widetilde{\text{ext}}(\text{p}_{\text{val}})}(\tilde{\mathbf{x}})(\tilde{x})(0)$ and accepting (no rejecting) it otherwise. In Section 3, we will propose a different rule based on the theory of imprecise probabilities. First, we need to give an alternative description of the fuzzy p-value.

2.3 An alternative approach to the concept of fuzzy p-value

Let us now give an alternative approach to the notion of fuzzy p-value. Let us first consider, for each particular realization $\mathbf{x} \in \mathbb{R}^n$, the Borel measurable mapping $D(\mathbf{x}) : \mathbb{R}^n \to \{0,1\}$ defined by:

$$D(\mathbf{x})(\mathbf{y}) = \begin{cases} 1 & \text{if } \text{p}_{\text{val}}(\mathbf{y}) < \text{p}_{\text{val}}(\mathbf{x}) \\ 0 & \text{otherwise.} \end{cases}$$

$D(\mathbf{x})(\mathbf{y})$ takes the value 1 when the sample $\mathbf{y}$ is "less compatible" with the null hypothesis than $\mathbf{x}$ is. Thus, for a fixed $\mathbf{x} \in \mathbb{R}^n$, we have:

$$\sup_{\theta \in \Theta_0} P_\theta(D(\mathbf{x}) = 1) = \sup_{\theta \in \Theta_0} P_\theta(\{\mathbf{y} \in \mathbb{R}^n : \text{p}_{\text{val}}(\mathbf{y}) < \text{p}_{\text{val}}(\mathbf{x})\}).$$

Let us now remind that φ_α is assumed to be a test of size α, i.e.,

$$\sup_{\theta \in \Theta_0} E_\theta(\varphi_\alpha(\mathbf{X})) = \sup_{\theta \in \Theta_0} P_\theta(R_\alpha) = \alpha.$$

Hence, we can prove that $D(\mathbf{x})$ satisfies the equality:

$$\sup_{\theta \in \Theta_0} P_\theta(D(\mathbf{x}) = 1) = \text{p}_{\text{val}}(\mathbf{x}).$$

For the sake of simplicity, let us assume that the sizes of the $\alpha-$tests are associated to a certain value of the parameter $\theta_0 \in \Theta_0$, i.e., let us assume that:

$$\sup_{\theta \in \Theta_0} P_\theta(R_\alpha) = P_{\theta_0}(R_\alpha) = \alpha, \; \forall \alpha \in (0,1).$$

(The above condition holds, for instance, when the null hypothesis is simple and also for the most common unilateral and bilateral tests.) In that case, $D(\mathbf{x})$ is a Bernoulli random variable with parameter $\mathrm{p_{val}}(\mathbf{x})$, under the distribution F_{θ_0}. In other words, $\mathrm{p_{val}}(\mathbf{x}) = P_{\theta_0}(\{D(\mathbf{x}) = 1\}), \; \forall \mathbf{x} \in \mathbb{R}^n$. (The p-value of $\mathbf{x}$ represents the probability, under the null hypothesis, of getting a sample which is "less compatible" with H_0 than $\mathbf{x}$ is.) Let $\mathscr{X}$ represent the class of binary random variables that can be defined on $\mathbb{R}^n$ and let us now use the extension principle to extend $D : \mathbb{R}^n \to \mathscr{X}$ to $\mathscr{F}(\mathbb{R}^n)$. I.e., let us define the mapping $\widetilde{\mathrm{ext}}(D) : \mathscr{F}(\mathbb{R}^n) \to \mathscr{F}(\mathscr{X})$ as follows:

$$\widetilde{\mathrm{ext}}(D)(\tilde{\mathbf{x}})(Z) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : D(\mathbf{x}) = Z\}, \; \forall Z \in \mathscr{X}.$$

Let us note that $\widetilde{\mathrm{ext}}(D)(\tilde{\mathbf{x}})$ is a possibility distribution over $\mathscr{X}$ and represents our imprecise information about $D(\mathbf{x}^*)$, according to our imprecise perception of the realization $\mathbf{x}^*$, represented by $\tilde{\mathbf{x}}$. More specifically, for each binary random variable $Z \in \mathscr{X}$, $\widetilde{\mathrm{ext}}(D)(\tilde{\mathbf{x}})(Z)$ represents the possibility grade that $D(\mathbf{x}^*)$ coincides with Z. Each binary random variable induces a Bernoulli distribution, $B(p)$. Thus, according to [3], we can derive a possibility distribution on the class of the Bernoulli measures. From now on, we will denote the class of all Bernoulli distributions by $\mathscr{P}_{\wp(\{0,1\})}$, since it is the class of probability measures that can be defined over $\wp(\{0,1\})$. This possibility measure, $\boldsymbol{\Pi}_{\tilde{\mathbf{x}}}$, is determined by the possibility distribution $\boldsymbol{\pi}_{\tilde{\mathbf{x}}} : \mathscr{P}_{\wp(\{0,1\})} \to [0,1]$:

$$\boldsymbol{\pi}_{\tilde{\mathbf{x}}}(B(p)) = \sup\{D(\tilde{\mathbf{x}})(Z) : P_Z \equiv B(p)\}, \; \forall p \in [0,1].$$

In words, $\boldsymbol{\pi}_{\tilde{\mathbf{x}}}(B(p))$ represents the degree of possibility that the probability measure $B(\mathrm{p_{val}}(\mathbf{x}^*))$ induced by $D(\mathbf{x}^*)$ coincides with $B(p)$. In other words, $\boldsymbol{\pi}_{\tilde{\mathbf{x}}}(B(p))$ represents the degree of possibility of the equality $\mathrm{p_{val}}(\mathbf{x}^*) = p$. Mathematically,

$$\begin{aligned} \boldsymbol{\pi}_{\tilde{\mathbf{x}}}(B(p)) &= \sup\{D(\tilde{\mathbf{x}})(Z) : P_Z \equiv B(p)\} = \sup\{D(\tilde{\mathbf{x}})(Z) : P(Z=1) = p\} \\ &= \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : P(D(\mathbf{x}) = 1) = p\} = \widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})(p), \; \forall p \in [0,1]. \end{aligned}$$

Summarizing, the fuzzy p-value is closely related to a certain second-order possibility measure [5]. Section 3 will be based on this alternative description of the fuzzy p-value.

3 Defuzzification of the fuzzy p-value

In Section 2.1 we have shown how the fuzzy p-value can be interpreted in terms of a second order possibility measure. In fact, $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})$ represents a possibility distribution over the class of possible values of the parameter of a Bernoulli random variable, and we have identified it with a second-order possibility measure $\boldsymbol{\Pi}_{\tilde{\mathbf{x}}}$ defined over the class of all Bernoulli distributions. According to Section 2.1, $\boldsymbol{\Pi}_{\tilde{\mathbf{x}}}$ and $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})$ are connected by the formula:

$$\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})(p) = \boldsymbol{\pi}_{\tilde{\mathbf{x}}}(B(p)) = \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}(\{B(p)\}) \tag{2}$$

According to Walley [10], any second-order possibility measure (which is an upper probability over the class of standard probabilities) can be reduced into a pair of upper and lower probabilities. Let us briefly describe Walley's procedure in our particular situation. We will consider the product space $\mathscr{P}_{\wp(\{0,1\})} \times \wp(\{0,1\})$ and:

- the possibility measure $\boldsymbol{\Pi}_{\tilde{\mathbf{x}}}$ on $\mathscr{P}_{\wp(\{0,1\})}$. (In our particular problem, it represents our imprecise knowledge about the probability distribution of the random variable $D(\mathbf{x}^*)$.)
- the "transition probability" $\mathbb{P}_2^1 : \mathscr{P}_{\wp(\{0,1\})} \times \wp(\{0,1\}) \to [0,1]$ given by the formula:
$$\mathbb{P}_2^1(A,P) := P(A),\ \forall A \in \wp(\{0,1\}), P \in \mathscr{P}_{\wp(\{0,1\})}.$$
(It represents the following conditional probability information: if P were the true Bernoulli distribution associated to $D(\mathbf{x}^*)$, then the probability of occurrence of the event $D(\mathbf{x}^*) \in A$ should be $P(A)$. In particular, for $A = \{1\}$, and $P = B(p)$, the quantity $\mathbb{P}_2^1(\{1\}, B(p)) = p$ represents the probability of occurrence of the event $D(\mathbf{x}^*) = 1$ according to the conditional information "$D(\mathbf{x}^*)$ induces the probability measure $B(p)$".)

In this setting, Walley constructs, by means of natural extension techniques, an upper-lower joint model. Thus, the available information about the marginal distribution on the second space $\wp(\{0,1\})$ is described, in a natural way, by a pair of lower and upper probabilities, $\underline{P}_W$ and $\overline{P}_W$. In particular, $\underline{P}_W(\{1\})$ and $\overline{P}_W(\{1\})$ will represent the tightest bounds for the probability of the event $D(\mathbf{x}^*) = 1$ or, in other words, the tightest bounds for the p-value, $\mathrm{p_{val}}(\mathbf{x}^*)$. To specify how this reduction is made, let us first recall that the second-order possibility measure $\boldsymbol{\Pi}_{\tilde{\mathbf{x}}}$ can be identified with the class of second-order probability measures $\{\mathbb{P} : \mathbb{P} \le \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}\}$. If $\mathbb{P}$ were the "true" second-order probability that governs the "random"[2] experiment associated to the choice of the "true" Bernoulli distribution, then the probability of occurrence of the event $\{1\}$ (i.e., the "true" p-value) should be computed as follows (if we combine degrees of belief about events and about probabilities of events into the same model):

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

Since all we know about $\mathbb{P}$ is that it is dominated by the possibility measure $\boldsymbol{\Pi}_{\tilde{\mathbf{x}}}$, the lowest upper bound for the probability of occurrence of the event $D(\mathbf{x}^*) = \{1\}$ is determined by

[2] Note that we are here interpreting the uncertainty associated to the perception of $\mathbf{x}^*$ as "randomness", since this imprecise perception is described by a possibility measure, which is, in turn, an upper probability.

$$\overline{P}_W(\{1\}) = \sup_{\mathbb{P}\leq \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}} \int \mathbb{P}_2^1(\{1\},P)\, d\mathbb{P}(P) = \sup_{\mathbb{P}\leq \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}} \int P(\{1\})\, d\mathbb{P}(P).$$

Similar arguments lead us to represent the highest lower bound of the probability by:

$$\underline{P}_W(\{1\}) = \inf_{\mathbb{P}\leq \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}} \int \mathbb{P}_2^1(\{1\},P)\, d\mathbb{P}(P) = \inf_{\mathbb{P}\leq \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}} \int P(\{1\})\, d\mathbb{P}(P).$$

Thus, the Walley reduction allows us to convert the fuzzy p-value into the crisp interval $[\underline{\mathrm{p_{val}}}(\tilde{\mathbf{x}}), \overline{\mathrm{p_{val}}}(\tilde{\mathbf{x}})] = [\underline{P}_W(\{1\}), \overline{P}_W(\{1\})]$. Furthermore, according to Walley [10], these upper and lower bounds can be alternatively computed as follows:

$$\overline{P}_W(\{1\}) = \int_0^1 \overline{P}_\delta(\{1\})\, d\delta, \quad \underline{P}_W(\{1\}) = \int_0^1 \underline{P}_\delta(\{1\})\, d\delta,$$

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

$$\overline{P}_\delta(\{1\}) = \sup\{Q(\{1\}) : Q \in \mathscr{P}_{\wp(\{0,1\})}, \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}(\{Q\}) \geq \delta\} \text{ and}$$
$$\underline{P}_\delta(\{1\}) = \inf\{Q(\{1\}) : Q \in \mathscr{P}_{\wp(\{0,1\})}, \boldsymbol{\Pi}_{\tilde{\mathbf{x}}}(\{Q\}) \geq \delta\}.$$

Theorem 1.

$$\overline{P}_\delta(\{1\}) = \sup[\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})]_\delta \ \ \textit{and} \ \underline{P}_\delta(\{1\}) = \inf[\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})]_\delta,\ \forall\, \delta \in [0,1].$$

According to the last theorem, the combination of first and second-order probabilities into the same model converts the fuzzy p-value, $\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})$ into the interval:

$$\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}} = [\underline{\mathrm{p_{val}}}(\tilde{\mathbf{x}}), \overline{\mathrm{p_{val}}}(\tilde{\mathbf{x}})] = \left[\int_0^1 \inf[\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})]_\delta\, d\delta, \int_0^1 \sup[\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})]_\delta\, d\delta\right]. \quad (3)$$

The extreme points of such interval represent the most accurate bounds for the true p-value, $\mathrm{p_{val}}(\mathbf{x}^*)$, based on our imprecise knowledge of $\mathbf{x}^*$. Let us denote by $\varphi_{\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}}}$ the multi-valued $\alpha-$test associated to such interval

$$\varphi_{\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}}}(\tilde{\mathbf{x}}) = \begin{cases} \{0\} & \text{if } \underline{\mathrm{p_{val}}}(\tilde{\mathbf{x}}) = \int_0^1 \inf[\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})]_\delta\ d\delta > \alpha \\ \{1\} & \text{if } \overline{\mathrm{p_{val}}}(\tilde{\mathbf{x}}) = \int_0^1 \sup[\widetilde{\mathrm{ext}}(\mathrm{p_{val}})(\tilde{\mathbf{x}})]_\delta\ d\delta \leq \alpha \\ \{0,1\} & \text{otherwise.} \end{cases}$$

The following relation between $\varphi_{\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}}}$ and the Denœux et al. [6] defuzzification of $\varphi_{\widetilde{\mathrm{ext}}(\mathrm{p_{val}})}$ holds:

Theorem 2. $\mathrm{defuz_{DMH}}(\varphi_{\widetilde{\mathrm{ext}}(\mathrm{p_{val}})}) \subseteq \varphi_{\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}}}$.

According to this result, the multi-valued test proposed in this paper is more times inconclusive than the Denœux et al. defuzzification is. I.e., whenever $\varphi_{\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}}}$ leads us to a clear decision (reject or accept the null hypothesis), $\mathrm{defuz}(\varphi_{\widetilde{\mathrm{ext}}(\mathrm{p_{val}})})$ also leads to the same decision. But, for some fuzzy samples $\mathrm{defuz_{DMH}}(\varphi_{\widetilde{\mathrm{ext}}(\mathrm{p_{val}})})$ is conclusive and $\varphi_{\underline{\overline{\mathrm{p_{val}}(\tilde{\mathbf{x}})}}}$ is not. This could be viewed as an argument against the use of $\varphi_{\underline{\overline{p_{val}(\tilde{\mathbf{x}})}}}$. Nevertheless, it is not clear whether a higher number of inconclusive tests is a disadvantage or an improvement. The dependence between the degree of imprecision of the data-set and how many times a given test is inconclusive is not clear, and should be further studied in future works.

4 Concluding remarks

We have proposed a new construction of crisp tests from fuzzy data, based on the theory of imprecise probabilities. The new tests are obtained as functions of the fuzzy p-values associated to the fuzzy samples, but they cannot be obtained as direct defuzzifications of the initial fuzzy tests.

Acknowledgements This work has been supported by the FEDER-MEC Grants TIN2007-67418-C03-03, TIN2005-08386-C05-05 and MTM2007-61193.

References

1. Couso I, Dubois D, Montes S, Sánchez L (2007) On various definitions of the variance of a fuzzy random variable. In: De Cooman G, Vejnarova J, Zaffalon M (eds) Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'07, Prague, Czech Republic). Action M Agency, SIPTA, Prague, pp 135–144
2. Couso I, Montes S, Gil P (2001) The necessity of the strong $\alpha-$cuts of a fuzzy set. Int J Uncertain Fuzziness Knowl-Based Syst 9:249–262
3. Couso I, Sánchez L (2008) Higher order models for fuzzy random variables. Fuzzy Sets Syst 159:237–258
4. Couso I, Sánchez L (2008) Mark-recapture techniques in statistical tests for imprecise data. Submitted for publication
5. De Cooman G (2005) A behavioural model for vague probability assessments. Fuzzy Sets Syst 154:305–358
6. Denœux T, Masson MH, Hébert PA (2005) Nonparametric rank-based statistics and significance test for fuzzy data. Fuzzy Sets Syst 153:1–28
7. Ferson S, Kreinovich V, Hajagos J, Oberkampf W, Ginzburg L (2007) Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty. Technical Report SAND2007-0939. Sandia National Laboratories, Albuquerque, New Mexico
8. Filzmoser P, Viertl R (2004) Testing hypotheses with fuzzy data: The fuzzy p-value. Metrika 59:21–29
9. Körner R (1997) Linear Models with Random Fuzzy Variables. PhD Dissertation, Technische Universität Bergakademie Freiberg
10. Walley P (1997) Statistical inferences based on a second-order possibility distribution. Int J Gen Syst 26:337–384