Inner and outer fuzzy confidence regions determined by low quality data

Inés Couso
University of Oviedo (Spain)
e-mail: couso@uniovi.es

Luciano Sánchez
University of Oviedo (Spain)
e-mail: luciano@uniovi.es

Abstract

We extend the notion of confidence region to fuzzy data, by defining a pair of fuzzy inner and outer confidence regions. We show the connection with previous proposals, as well as with recent studies on hypothesis testing with low quality data.

Keywords: Confidence region, Possibility measure, Hypothesis testing, Fuzzy rough set.

1 Introduction

Imprecise measurements arise very often in real world problems. Sometimes, this imprecision is small enough so that it can be safely ignored, and other times, it can be modeled by a probability distribution (e.g. additive random noise). But there is another kind of problems where the imprecision is significant, and a probability distribution is not a natural model (see [8], for a detailed discussion). In this paper, we assume that the data set is a collection of n inputs, each one of them is:

- A set of mutually exclusive values, one of which is the attribute value of the object under concern.
- A fuzzy subset of the real line, interpreted as a possibility distribution over the class of possible values for the attribute.

In the recent literature, several inferential procedures, have been performed to manage with this kind of imprecise data (see [5, 6, 9], for instance). In this paper, we generalize the concept of of confidence region. The most related precedent in the literature is the notion of "fuzzy confidence interval" introduced by Kruse and Meyer in [10]. There, a convex fuzzy subset of $\mathbb{R}$ is assigned to each fuzzy sample (each collection of n fuzzy numbers). Such fuzzy confidence interval will contain the "fuzzy perception" of the parameter θ, with a certain confidence $1-\alpha$. In our paper, we will proceed in a different way. We will start from a specific $1-\alpha$ confidence interval for the parameter θ and we will try to express the available information about such confidence interval. Such available information will be imprecise, due to the imprecision in the data set. Thus, when each one of the n inputs is a set of values, we will represent the information about the confidence intervals by means of a pair "outer" and "inner" intervals. When, in a more general setting, each input is represented by means of a fuzzy subset of possible values for the attribute, we will represent the imprecise information about the confidence interval by means of a pair of "inner" and "outer" fuzzy subsets of the real line.

Furthermore, we will check that the computation of the "outer" fuzzy region is also related to the fuzzy confidence interval defined by Kruse and Meyer, even when the interpretation is totally different. But we will show that the information provided by such fuzzy confidence interval is not enough in the decision stage, and that we also need the additional information provided by the "inner" fuzzy region. At the end of the paper, we will give some guidelines about the way to take decisions based on this kind of information.

2 Imprecise confidence regions associated to low quality data

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 set-valued mapping $\mathrm{Reg} : \mathbb{R}^n \to \wp(\mathbb{R})$ represent a $1-\alpha$-level confidence interval for a certain unknown parameter θ_{X^*} of the df F^*, i.e., let is satisfy the following restriction:

$$P_{\theta_{X^*}}(\{\mathbf{x} \in \mathbb{R}^n : \mathrm{Reg}(\mathbf{x}) \ni \theta_{X^*}\}) \geq 1 - \alpha.$$

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 $\mathbb{R}^n$, $\tilde{\mathbf{x}} \in \mathcal{F}(\mathbb{R}^n)$. 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 $\mathbf{x}$.

We will extend the mapping $\mathrm{Reg} : \mathbb{R}^n \to \wp(\mathbb{R})$ to the class of fuzzy subsets of $\mathbb{R}^n$. But we will not apply Zadeh's Extension Principle. If we would do so, we would define a mapping from $\mathcal{F}(\mathbb{R}^n)$ to $\mathcal{F}(\wp(\mathbb{R}))$. Instead, we will define two mappings $\widetilde{\mathrm{OutReg}}$ and $\widetilde{\mathrm{InnReg}}$ from $\mathcal{F}(\mathbb{R}^n)$ to $\mathcal{F}(\mathbb{R})$ as follows.

Definition 2.1 *We will call the* outer fuzzy confidence region associated to $\tilde{\mathbf{x}}$ *to the fuzzy set* $\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})$ *defined as follows:*

$$\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(y) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : \mathrm{Reg}(\mathbf{x}) \ni y\}, \ \forall y. \tag{1}$$

According to the possibilistic interpretation of the fuzzy sample $\tilde{\mathbf{x}}$, $\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(y)$ will represent the grade of possibility that $\mathbf{x}^*$ belongs to the family of samples:

$$\{\mathbf{x} \in \mathbb{R}^n : \mathrm{Reg}(\mathbf{x}) \ni y\}.$$

In other words, the membership value $\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(y)$ represents the possibility that the confidence region determined by $\mathbf{x}^*$ contains y, when our imprecise perception about $\mathbf{x}^*$ is represented by $\tilde{\mathbf{x}}$.

When, in particular, $\tilde{\mathbf{x}}$ is a crisp set $A \subseteq \mathbb{R}^n$, then $\widetilde{\mathrm{OutReg}}(A)$ is the indicator function associated to the crisp set $\cup_{\mathbf{x}\in A}\mathrm{Reg}(\mathbf{x})$. In fact, if all we knew about $\mathbf{x}^*$ were that it belongs to the crisp set A, then $\widetilde{\mathrm{OutReg}}(A) = \cup_{\mathbf{x}\in A}\mathrm{Reg}(\mathbf{x})$ would be the most committed set that would contain $\mathrm{Reg}(\mathbf{x}^*)$ with certainty. So it is an *outer approximation of* $\mathrm{Reg}(\mathbf{x}^*)$.

Let us now consider inner approximations of $\mathrm{Reg}(\mathbf{x}^*)$ associated to imprecise perceptions of $\mathbf{x}^*$.

Definition 2.2 *We will call the inner fuzzy confidence region associated to* $\tilde{\mathbf{x}}$ *to the fuzzy set* $\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})$ *defined as follows:*

$$\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(y) = 1 - \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \notin \mathrm{Reg}(\mathbf{x})\}.$$

For an arbitrary $y \in \mathbb{R}$, the value $[\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})]^c(y) = 1 - \widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(y) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \notin \mathrm{Reg}(\mathbf{x})\}$ represents the grade of possibility that $\mathbf{x}^*$ belongs to the following family of samples:

$$\{\mathbf{x} \in \mathbb{R}^n : \mathrm{Reg}(\mathbf{x}) \not\ni y\}.$$

Thus, according to the duality between possibility and necessity measures, the membership $\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(y)$ represents the necessity of the complementary

$$\{\mathbf{x} \in \mathbb{R}^n : \mathrm{Reg}(\mathbf{x}) \ni y\}.$$

In other words, $\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(y)$ represents the degree of certainty (necessity) that $\mathrm{Reg}(\mathbf{x}^*)$ (the "true" confidence region) contains y.

Let us notice that if, in particular, $\tilde{\mathbf{x}}$ is a crisp set $A \subseteq \mathbb{R}^n$, $\widetilde{\mathrm{InnReg}}(A)$ is the indicator function of the crisp set $\cap_{\mathbf{x}\in A}\mathrm{Reg}(\mathbf{x})$. In other words, it is the largest set contained in $\mathrm{Reg}(\mathbf{x}^*)$ with certainty. So it is an *inner approximation of* $\mathrm{Reg}(\mathbf{x}^*)$.

[1] We show in [1, 3, 4] some specific situations where such a membership function is derived from an imprecise perception of some $\mathbf{x}^*$.

Example 2.1 *We have a container of apples and we are asked about their expected weight. We use a scale, but we do not fully trust the obtained measurement. Let us denote the displayed quantity for each apple ω by $d = D(\omega)$. We consider the scales are "under control" 90% of the time, and in such situation the measurements are within a 3g error margin. In the remaining 10% of the time, the scales are "out of control" and we can only guarantee an error lower than 15g. Apples are picked at random from the container, Ω. Let us denote by $x^* = X^*(\omega)$ the ill-known quantity describing the (true) weight of an arbitrary apple $\omega \in \Omega$. We suppose that the random variable X^* is normally distributed, with know variance $\sigma^2 = 100g^2$ and unknown expectation, $E(X^*) = \theta$. We want to provide some confidence-interval information about θ, on the basis of a sample of 100 displayed quantities, $(d_1, \ldots, d_{100})$. We have imprecise information about the true weights $\mathbf{x}^* = (x_1^*, \ldots, x_{25}^*)$ (based on the displayed quantities and our knowledge about the precision of the scale). According to [3], we can describe this information by means of the fuzzy set $\tilde{\mathbf{x}} \in \mathcal{F}(\mathbb{R}^n)$ whose membership is defined as follows for each $\mathbf{x} = (x_1, \ldots, x_{25})$:*

$$\tilde{\mathbf{x}}(\mathbf{x}) = \begin{cases} 1 & \textit{if } x_i \in [d_i - 3, d_i + 3],\ \forall i \\ 0.1 & \textit{if } x_i \in [d_i - 15, d_i + 15],\ \forall i, \\ & \textit{and } x_j \notin [d_j - 3, d_j + 3] \\ & \textit{for some } j \\ 0 & \textit{otherwise.} \end{cases}$$

Now we will provide the imprecise information we have about the 0.95-confidence interval about $\theta_{X^*} = E(X^*)$, $\mathrm{Reg}\ (\mathbf{x}^*) = (\overline{x} - 3.92, \overline{x} + 1.96)$, based on our imprecise information about the realization $\mathbf{x}^*$. According to Equation 1, we will define the fuzzy set $\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})$ as follows:

$\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(y) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \in \mathrm{Reg}(\mathbf{x})\} = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \in (\overline{x} - 3.96, \overline{x} + 3.92)\} =$

$$\begin{cases} 1 & \textit{if } y \in (\overline{d} - 6.92, \overline{d} + 6.92), \\ 0.1 & \textit{if } y \in (\overline{d} - 18.92, \overline{d} + 18.92), \\ & \textit{but } y \notin (\overline{d} - 6.92, \overline{d} + 6.92) \\ 0 & \textit{otherwise.} \end{cases}$$

As we pointed out at the beginning of this section, $\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(y)$ represents the degree of possibility that y belongs to $\mathrm{Reg}(\mathbf{x}^*)$.

Similarly, we can define the fuzzy set $\widetilde{\mathrm{InnReg}}(\tilde{x})$ as follows:

$\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(y) = 1 - \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \notin \mathrm{Reg}(\mathbf{x})\} = 1 - \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \in (-\infty, \overline{x} - 3.92) \cup (\overline{x} + 3.92, \infty)\} =$

$$\begin{cases} 0 & \textit{if } y \in (-\infty, \overline{d} - 0.92) \cup (\overline{d} + 0.92, \infty) \\ 0.0 & \textit{if } y \in [\overline{d} - 0.92, \overline{d} + 0.92], \end{cases}$$

$\widetilde{\mathrm{InnReg}}(\tilde{x})(y)$ represents the degree of certainty that y belongs to $\mathrm{Reg}(\mathbf{x}^*)$.

2.1 Relationship with Kruse and Meyer approach

First of all, we need to recall the notion of "fuzzy perception" of a parameter in Kruse and Meyer's context. According to the last section, let us denote θ_{X^*} a certain parameter associated to the df F^* of the random variable $X^* : \Omega \to \mathbb{R}$. Let $\tilde{X}$ denote the fuzzy perception of X^* in the sense that, for any random variable X, $\tilde{X}(X)$ denotes the grade of possibility that X coincides with the "true random variable", X^*. The *fuzzy perception of* θ_{X^*} is the fuzzy set $\tilde{\theta}_{\tilde{X}}$ defined as follows:

$$\tilde{\theta}_{\tilde{X}}(y) = \sup\{\tilde{X}(X) : \theta_X = y\}.$$

The membership $\tilde{\theta}_{\tilde{X}}(y)$ represents the grade of possibility that the true value of the parameter θ_{X^*} coincides with y. This possibility degree is calculated on the basis of the vague perception of X^*, but let us recall that such uncertainty is not related at all to the idea of random sample. Let us now recall the definition of a *fuzzy confidence interval* introduced in [10]. The authors say that the convex fuzzy subset $\Pi \in \mathcal{F}(\mathbb{R})$ is a "fuzzy confidence interval" when it satisfies the restrictions

$$P(\{\omega \in \Omega : (\tilde{\theta}_{\tilde{X}})_\delta \subseteq \Pi_\delta\}) \geq 1 - \alpha,\ \forall\, \delta \in (0, 1).$$

Let now $\mathrm{Reg} = (T_1, T_2) : \mathbb{R}^n \to \wp(\mathbb{R})$ denote a $1 - \alpha$-confidence interval for θ_{X^*} based on a specific realization $\mathbf{x}^*$ and let $\tilde{\mathbf{x}}$ denote the fuzzy perception of $\mathbf{x}^*$. The authors check that the convex fuzzy set $\Pi(\tilde{\mathbf{x}})$ defined in Equation 2 satisfies the above definition.

$$\Pi(\tilde{\mathbf{x}})(y) = \sup\{\delta \in [0,1] : y \in (\Pi^1_\delta(\tilde{\mathbf{x}}), \Pi^2_\delta(\tilde{\mathbf{x}}))\}, \quad (2)$$

where $\Pi^1_\delta(\tilde{\mathbf{x}})$ and $\Pi^2_\delta(\tilde{\mathbf{x}})$ are defined as follows:

$$\Pi^1_\delta(\tilde{\mathbf{x}}) = \inf\{T_1(\mathbf{x}) : \mathbf{x} \in \tilde{\mathbf{x}}_\delta\} \text{ and}$$
$$\Pi^2_\delta(\tilde{\mathbf{x}}) = \sup\{T_2(\mathbf{x}) : \mathbf{x} \in \tilde{\mathbf{x}}_\delta\}.$$

When the extremes of the confidence interval, T_1 and T_2, are continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$, and the δ-cuts $\tilde{\mathbf{x}}_\delta$ are closed, our outer fuzzy region $\widetilde{\text{OutReg}}(\tilde{\mathbf{x}})$ is a convex fuzzy set and it coincides with Kruse & Meyer's fuzzy interval $\Pi(\tilde{\mathbf{x}})$. But our interpretation is not the same: while they try to cover the imprecise (fuzzy) perception, $\tilde{\theta}_{\tilde{X}}$ of the parameter θ_{X^*}, we aim to describe the available imprecise information about the (crisp) confidence region $\text{Reg}(\mathbf{x}^*)$. Furthermore, Kruse and Meyer definition does not consider the inner fuzzy region $\widetilde{\text{InnReg}}(\tilde{\mathbf{x}})$. In Section 2.2, we will show that we need to take it into account, when we aim to construct "fuzzy tests" (see [5, 6, 8]) from fuzzy regions.

2.2 Fuzzy tests induced by fuzzy confidence regions

In classical statistics, we find a strong connection between the construction of parametric tests and confidence regions. In this subsection, we will show how this connection can be extended to the case of fuzzy imprecise perceptions of the sample. More specifically, we will show the relation between the inner and outer fuzzy regions considered in the last subsection and the fuzzy tests considered in some recent papers as [5, 6, 8], for instance.

Let $X^* : \Omega \to \mathbb{R}$ be a random variable and let us state the hypothesis

$$H_0 : \theta_{X^*} = \theta_0 \text{ against } H_1 : \theta_{X^*} \in \Theta_1, \quad (3)$$

where θ_{X^*} is a parameter that depends on the probability distribution F^* induced by X^*. (The last equation contains, as particular cases, all one-sided and two-sided parametrical tests). Let $\varphi : \mathbb{R}^n \to \{0,1\}$ be a non-randomized test that represents the decision rule that will lead to a decision to accept or reject the null hypothesis. The critical region associated to φ is:

$$C = \{\vec{x} \in \mathbb{R}^n : \varphi(\vec{x}) = 1\}.$$

The mapping φ is said to be a test with level of significance α, $0 \le \alpha \le 1$, when

$$E_{\theta_0}(\varphi) = P_{\theta_0}(C) \le \alpha$$

Let now $\text{Reg} : \mathbb{R}^n \to \wp(\mathbb{R})$ denote a $1-\alpha$-confidence region, in the sense that

$$P_{\theta_{X^*}}\left(\{\mathbf{x} \in \mathbb{R}^n : \text{Reg}(\mathbf{x}) \ni \theta_{X^*}\}\right) \ge 1 - \alpha.$$

It is well known that the test φ_{Reg} defined as follows:

$$\varphi_{\text{Reg}}(\mathbf{x}) = \begin{cases} 1 & \text{if } \text{Reg}(\mathbf{x}) \not\ni \theta_0 \\ 0 & \text{if } \text{Reg}(\mathbf{x}) \ni \theta_0 \end{cases} \quad (4)$$

is a test of size α for testing the hypotheses given in Equation 3.

On the other hand, in some recent papers ([5, 6, 8]) the notion of α-test has been extended to the case where the sample realization $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ is perceived with imprecision. Let us recall the construction proposed in those papers.

Let us first suppose that such imprecise perceptions are given by sets mutually exclusive points in $\mathbb{R}^n$. Then, given an α-test $\varphi : \mathbb{R}^n \to \{0,1\}$, it seems natural to extend it to $\wp(\mathbb{R})$ as follows:

$$\varphi(A) = \{\varphi(\mathbf{x}) : \mathbf{x} \in A\} =$$
$$\begin{cases} \{1\} & \text{if } \varphi(\mathbf{x}) \in C, \ \forall \mathbf{x} \in A, \\ \{0\} & \text{if } \varphi(\mathbf{x}) \in C^c, \ \forall \mathbf{x} \in A, \\ \{0,1\} & \text{otherwise.} \end{cases} \quad (5)$$

The assignation $\varphi(A) = \{0,1\}$ means that our perception of $\mathbf{x}^*$ is too imprecise and prevents us to take a clear decision (rejecting (1) or no rejecting (0)). So we would need further information to be able to take a decision.

Let us now suppose the more general case were the imprecise perception of $\mathbf{x}^*$ is represented by a fuzzy set $\tilde{\mathbf{x}}$. Then, according to Zadeh's Extension Principle (that extends the construction proposed in Equation 5) φ can be extended from $\mathbb{R}^n$ to $\mathcal{F}(\mathbb{R}^n)$ (see [5, 6]) as follows:

$$\tilde{\varphi}(\tilde{\mathbf{x}})(1) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : \varphi(\mathbf{x}) = 1\} \text{ and}$$

$$\tilde{\varphi}(\tilde{\mathbf{x}})(0) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : \varphi(\mathbf{x}) = 0\}. \tag{6}$$

Having into account the possibilistic interpretation of the fuzzy set $\tilde{\mathbf{x}}$, $\tilde{\varphi}(\tilde{\mathbf{x}})(1)$ represents the possibility that the null hypothesis would be rejected, had the sample realization $\mathbf{x}^*$ been precisely observed. Similarly, $\tilde{\varphi}(\tilde{\mathbf{x}})(0)$ represents the possibility that H_0 would not be rejected, had the data been precisely observed. Equivalently, it is equal to one minus the necessity of rejection of the null hypothesis.

When, in particular, the crisp test, φ_{Reg} is derived (according to Equation (4)) from a confidence region Reg, the associated fuzzy test, $\tilde{\varphi}_{\mathrm{Reg}}$, can be expressed in terms of the outer and inner fuzzy regions considered in the last subsection. In fact, we can easily check that the fuzzy test $\tilde{\varphi}_{\mathrm{Reg}}$ obtained by applying Equation (6) to the classical φ_{Reg} can be alternatively expressed as follows:

$$\tilde{\varphi}_{\mathrm{Reg}}(\tilde{\mathbf{x}})(0) = \widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(\theta_0)$$

$$\tilde{\varphi}_{\mathrm{Reg}}(\tilde{\mathbf{x}})(1) = 1 - \widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(\theta_0)$$

Note that we need to take into account not only the outer, but also the inner fuzzy region. Thus, the information provided by Kruse and Meyer fuzzy region would not be enough to take such kind of fuzzy decisions.

Example 2.2 *Let us consider again the situation described in Example 2.1 and let us suppose that the mean of the displayed quantities has been* $\overline{d} = 95g$*, so the outer and the inner fuzzy regions are:*

$$\widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(y) = \sup\{\tilde{\mathbf{x}}(\mathbf{x}) : y \in \mathrm{Reg}(\mathbf{x})\} =$$

$$\begin{cases} 1 & \text{if } y \in (88.04, 101.96), \\ 0.1 & \text{if } y \in (78.04, 88.04) \cup (101.96, 111.96), \\ 0 & \text{if } y < 78.04 \text{ or } y > 111.96. \end{cases}$$

and

$$\widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(y) =$$

$$\begin{cases} 0.9 & \text{if } y \in (94.08, 95.92), \\ 0 & \text{otherwise.} \end{cases}$$

If, for instance, we want to test the null hypothesis $H_0 : E(X^*) = 95.5$ *against* $H_1 : E(X^*) \neq 95.5$*, the fuzzy test associated to the above fuzzy confidence regions is the following fuzzy subset of* $\{0, 1\}$*:*

$$\tilde{\varphi}_{\mathrm{Reg}}(\tilde{\mathbf{x}})(0) = \widetilde{\mathrm{OutReg}}(\tilde{\mathbf{x}})(95.5) = 1$$

$$\tilde{\varphi}_{\mathrm{Reg}}(\tilde{\mathbf{x}})(1) = 1 - \widetilde{\mathrm{InnReg}}(\tilde{\mathbf{x}})(95.5) = 0.1.$$

To clarify our approach a bit more, let us modify a bit this example. Let us now suppose that the scales are always under control, so we know with certainty that the sample of the 25 true weights $\mathbf{x}^*$ *falls into the rectangle* $A = \Pi_{i=1}^{25}[d_i - 3, d_i + 3]$*. In such a case, the outer and the inner confidence regions are respectively the crisp sets:*

$$\mathrm{OutReg}(A) = (88.04, 101.96)$$

and

$$\mathrm{InnReg}(A) = (94.08, 95.98).$$

In that case, if we wanted to to test again the null hypothesis $H_0 : E(X^*) = 95.5$ *against* $H_1 : E(X^*) \neq 95.5$*, the imprecise test* $\varphi_{\mathrm{Reg}}(A)$ *would return the crisp set* $\{0, 1\}$*. So, we would not be able to take any decision (reject/no reject), due the scales imprecision. But if, for instance, we test the null hypothesis* $H_0 : E(X^*) = 96$ *against* $H_1 : E(X^*) \neq 96$*, then our imprecise test would be informative enough to reject the null hypothesis. (The imprecise test would return the set* $\{1\}$*). Notice that we have made use of our knowledge about the inner confidence interval* $\mathrm{InnReg}(A) = (94.08, 95.98)$ *to make such distinction between both situations.*

3 Concluding remarks and future work

We have extended the notion of confidence region of a parameter to the case where the sample is a collection of n crisp or fuzzy subsets of the real line. We suppose that each (fuzzy) subset represents some imprecisely observed quantity, interpreted as a possibility distribution (when the subset is crisp, such possibility distribution is 0-1 valued). We have defined a pair of inner and outer (fuzzy) confidence regions that represent our knowledge about the confidence region determined by the “true” sample, and they remind us to the concept of (fuzzy) upper and lower approximations in rough set theory. In a future, we will try to check whether these outer and inner regions can be expressed as upper and lower fuzzy-rough-set approximations associated to some similarity relation [7] or if, at least, they can be integrated into a broader context described in [2], where the similarity relation is non-transitive.

The fuzzy tests considered in Section 2.2 define “fuzzy decisions”. In the case where a crisp decision is absolutely needed, these fuzzy subsets may be defuzzified. In [6] the following defuzzification is proposed: the null hypothesis is rejected whenever the possibility of rejection is greater than the possibility of nonrejection ($\tilde{\phi}(\tilde{\mathbf{x}})(1) > (\tilde{\phi}(\tilde{\mathbf{x}})(0))$. In [5], we propose an alternative defuzzification. It is based on an interval-valued assignation for the critical level. Such defuzzification of the fuzzy p-value makes sense within the theory of Imprecise Probabilities [11], in accordance with the possibilistic interpretation of fuzzy random variables developed in [4]. In a expanded version of the paper, we plan to formalize the way that such defuzzifications transform our pair of fuzzy inner and outer regions into a pair of crisp regions, and how the new crisp regions can be interpreted in both cases.

Acknowledgements

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

References

[1] I. Couso, D. Dubois, On the variability of the concept of variance for fuzzy random variables, IEEE Transactions on Fuzzy Systems, accepted for publication.

[2] I. Couso, D. Dubois, Rough Sets, Coverings and Incomplete Information, First International Workshop on Rough Sets theory (RST09), accepted for presentation.

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

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

[5] I. Couso, L. Sánchez, Defuzzification of fuzzy p-values, In: Soft Methods for Handling Variability and Imprecision, D. Dubois et al. (Eds.) Springer, 2008.

[6] T. Denoeux, M.H. Masson, P. A. Hébert, Nonparametric rank-based statistics and significance tests for fuzzy data, Fuzzy Sets and Systems 153 (2005) 1-28.

[7] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems, 17, 191-209 (1990).

[8] S. Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf and L. Ginzburg, Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty, SAND2007-0939, Sandia National Laboratories, Albuquerque, New Mexico. http://www.ramas.com/intstats.pdf.

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

[10] R. Kruse, K.D. Meyer, Confidence intervals for the parameters of a linguistic random variable, in: J. Kacprzyk, M. Fedrizzi, (Eds.), Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making, Springer, Berlin, 1988, 113-123.

[11] P. Walley, Statistical reasoning with imprecise probabilities, Chapman and Hall, London, 1991.