

An analysis of the median of a fuzzy random variable based on Zadeh's extension principle

Sonia Pérez-Fernández and Beatriz Sinova

Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, 33071 Oviedo, Spain

Abstract

The median of a fuzzy random variable has been extended either by applying Zadeh's extension principle or by minimizing its mean distance w.r.t. a fuzzy number when a certain L^1 metric is considered. This paper aims to analyze connections between both approaches along with some properties of the first one, as well as some simulation studies to show that it also behaves better than the Aumann-type mean.

Keywords: 1-norm metric between fuzzy numbers, fuzzy random variable, median of a fuzzy number, random fuzzy number, Zadeh's extension principle

1. Introduction

Fuzzy random variables were introduced by Kwakernaak [14], and they were later re-formalized in an easier-to-follow and to-apply way by Kruse and Meyer [13], as a model for the fuzzy perception of real-valued random variables. In accordance with this approach, for each fuzzy random variable an underlying real-valued random variable (the so-called original) is assumed to actually exist, although it cannot be exactly observed and only a fuzzy perception from each observation/performance is available. In this way, and by following the ontic/epistemic classification of fuzzy sets by Dubois and Prade [9] (see also Couso and Dubois [4]), data from fuzzy random variables in Kwakernaak-Kruse and Meyer's sense correspond to the epistemic view.

In developing statistics with data coming from fuzzy random variables in Kwakernaak-Kruse and Meyer's sense, the most usual way to proceed has been to apply Zadeh's extension principle over the space of all the potential originals of the involved fuzzy random variable (see, for instance, Kruse and Meyer [13]).

Puri and Ralescu [15] have considered a different concept that has been originally coined as fuzzy random variable, although many authors and especially during the last years call it random fuzzy set (or *random fuzzy number* in the concerned one-dimensional case). This concept has been conceived as a model for the random mechanisms leading to fuzzy-valued data. More concretely, in accordance

with this approach there is no need for an underlying random variable to exist, but the random variable/element we observe is assumed to be intrinsically fuzzy-valued. By following the ontic/epistemic classification of fuzzy sets, data from random fuzzy numbers correspond to the ontic view.

In developing statistics with data coming from fuzzy random variables in Puri and Ralescu's sense, the most usual way to proceed is either to particularize some existing methods from Functional Data Analysis through the isometrical embedding given by the support function (see González-Rodríguez *et al.* [11]) or to establish *ad hoc* techniques (see Blanco-Fernández *et al.* [2, 3] for a recent review).

The median of a real-valued random variable was introduced as a summary measure of the location of the variable, showing a robust behaviour when outliers (atypical values) were present or some values are exchanged or modified. Thus, whereas some well-known location summary measures like the mean value are too sensitive to the presence of outliers or changes (which in some cases can be either erroneous or hardly representative of the population/sample), the median does not and this often makes it a more suitable description of the central tendency of the variable distribution.

Grzegorzewski [12] (see also Couso *et al.* [6, 5]) has stated a notion for the *median* of a fuzzy random variable in Kruse and Meyer's sense by using Zadeh's extension principle. The application of this principle has led to a unique fuzzy number which can be expressed and computed in terms of medians of certain real-valued random variables. Sinova *et al.* [17] has extended the notion of median to random fuzzy numbers as the fuzzy number(s) at which the mean distance from the random fuzzy number reaches the minimum, the considered distance being the 1-norm one by Diamond and Kloeden [7]. The solution for this minimization problem is not necessarily unique, and in case of non-uniqueness a convention has been suggested. This convention provides us with a fuzzy number which can be expressed and computed in terms of medians of certain real-valued random variables.

In this paper, we are first going to see that the measure by Grzegorzewski is a solution to the minimization problem by Sinova *et al.* In fact, the convention suggested by Sinova *et al.* would corre-

spond to that obtained by applying Zadeh's extension principle whenever the median of a real-valued random variable is constrained to consider such a usual convention for real-valued variables. The paper will also analyze some interesting properties for Grzegorzewski's median, and some comments about the practical coincidence of both measures will be given.

2. The approaches to the concepts of fuzzy random variables and some related medians

Let $\mathcal{F}_c(\mathbb{R})$ denote the space of (bounded) fuzzy numbers, which are the mappings $\tilde{U} : \mathbb{R} \rightarrow [0, 1]$ such that for each $\alpha \in [0, 1]$ the α -level set \tilde{U}_α is a nonempty compact interval, where

$$\tilde{U}_\alpha = \begin{cases} \{x \in \mathbb{R} : \tilde{U}(x) \geq \alpha\} & \text{if } \alpha > 0 \\ \text{closure}\{x \in \mathbb{R} : \tilde{U}(x) > 0\} & \text{otherwise} \end{cases}$$

Boundedness of the 0-level is removed in many studies, with hardly affecting most of the main ideas and results.

Consider a random experiment which is formalized within the probabilistic setting by means of a probability space (Ω, \mathcal{A}, P) , where Ω is the space of experimental outcomes, \mathcal{A} is the σ -field of the events associated with the experiment, and P is the associated probability measure. In accordance with Kruse and Meyer's approach

Definition 2.1 (Kruse and Meyer [13]) *Given the probability space (Ω, \mathcal{A}, P) , a mapping $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ is said to be a fuzzy random variable associated with this space if for all $\alpha \in [0, 1]$ both $\inf \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$ and $\sup \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$ are real-valued random variables, where for all $\omega \in \Omega$, $\{\mathcal{X}_\alpha(\omega) : \alpha \in [0, 1]\}$ denotes the level sets representation of $\mathcal{X}(\omega)$, i.e., $\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$ for each $\alpha \in [0, 1]$.*

As it has been already pointed out, fuzzy random variables in Kruse and Meyer's sense are viewed as perceptions of unknown real-valued random variables, and it is assumed we are not able to know which of the possible originals is the true original. Let $\mathfrak{D}_\mathcal{X}$ denote the set of all Borel measurable mappings $X : \Omega \rightarrow \mathbb{R}$ with respect to \mathcal{A} . By following Zadeh's extension principle [18], Kruse and Meyer [13] introduced the notion of fuzzy parameter of a fuzzy random variable \mathcal{X} as a fuzzy perception of a real-valued parameter of the original. Thus, if $\theta(X)$ denotes the parameter of the original X the fuzzy perception of θ for \mathcal{X} is given by the fuzzy number $\tilde{\vartheta}(\mathcal{X})$ such that for each $t \in \mathbb{R}$

$$\tilde{\vartheta}(\mathcal{X})(t) = \sup \left\{ \inf_{\omega \in \Omega} \mathcal{X}(\omega)(X(\omega)) : X \in \mathfrak{D}_\mathcal{X}, \theta(X) = t \right\}.$$

With this approach, the fuzzy perception of the median for \mathcal{X} corresponds to

Definition 2.2 (Grzegorzewski [12]) *If \mathcal{X} is a fuzzy random variable associated with the probability space (Ω, \mathcal{A}, P) , a fuzzy median of \mathcal{X} is the fuzzy number $\tilde{\Gamma}(\mathcal{X})$ such that for each $t \in \mathbb{R}$*

$$\tilde{\Gamma}(\mathcal{X})(t) = \sup \left\{ \inf_{\omega \in \Omega} \mathcal{X}(\omega)(X(\omega)) : X \in \mathfrak{D}_\mathcal{X}, t \in \text{Me}(X) \right\},$$

where $\text{Me}(X)$ denotes either the singleton or the interval of possible medians of the real-valued random variable X .

At this point, it should be recalled that $\text{Me}(X)$ is not necessarily uniquely defined. Actually, if F_X denotes the (cumulative) distribution function associated with X , then $\text{Me}(X)$ is a median of the distribution of X whenever $F_X(\text{Me}(X)-) \leq .5 \leq F_X(\text{Me}(X))$, so that if there is more than one value of X for which F_X takes on the value .5, then the median is not unique. A usual convention, in case there is a need for a choice, is to consider the midpoint of the interval of medians. But this means just a convention, not the definition of the median.

Grzegorzewski [12] has proved that

Proposition 2.1 (Grzegorzewski [12]) *The fuzzy median $\tilde{\Gamma}(\mathcal{X})$ of the fuzzy random variable \mathcal{X} is uniquely given by the fuzzy number such that*

$$(\tilde{\Gamma}(\mathcal{X}))_\alpha = [\underline{\text{Me}}(\inf \mathcal{X}_\alpha), \overline{\text{Me}}(\sup \mathcal{X}_\alpha)],$$

where $\underline{\text{Me}}$ denotes the median of the corresponding real-valued random variable with the convention (if the median is not unique) of taking the infimum of the interval of medians, and $\overline{\text{Me}}$ denotes the median of the corresponding real-valued random variable with the convention (if the median is not unique) of taking the supremum of the interval of medians.

On the other hand, in accordance with Puri and Ralescu's approach

Definition 2.3 (Puri and Ralescu [15]) *Given the probability space (Ω, \mathcal{A}, P) , a mapping $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ is said to be a fuzzy random variable (from now on a random fuzzy number) associated with this space if for all $\alpha \in [0, 1]$ the α -level mapping $\mathcal{X}_\alpha : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ (with $\mathcal{K}_c(\mathbb{R})$ denoting the class of nonempty compact intervals) is a random compact interval, i.e., a measurable function with respect to the Borel σ -algebra generated by the Hausdorff topology induced on $\mathcal{K}_c(\mathbb{R})$.*

As it has been already pointed out, random fuzzy numbers are viewed as fuzzy-valued random elements, and they model mechanisms generating at random fuzzy numbers. In fact, they can be equivalently defined as Borel-measurable mappings w.r.t. certain metrics in such a way that one can properly refer to the induced distribution of a random fuzzy number or the stochastic independence of random

fuzzy numbers. Furthermore, it should be highlighted that although motivations and situations to be modelled for the two approaches differ, and random fuzzy sets have been originally formalized in a much more general way, in case of dealing with fuzzy numbers both definitions have been proved to be mathematically equivalent.

This approach follows quite closely the ideas for random elements stated by Fréchet [10], where random elements are viewed as Borel-measurable mappings taking on values of a very wide nature. Actually, it seems that Fréchet foresaw in some senses the notion of random fuzzy set. In this respect, it seems rational to extend parameters by following Fréchet's ideas, what for the median can be done by extending the idea of looking for the fuzzy number(s) minimizing the mean distance between it(them) and the random fuzzy number. At this point, the extension should depend on the choice for the distance. A rather natural choice is to consider an L^1 metric between fuzzy numbers as the 1-norm one defined by Diamond and Kloeden [7] as

$$\rho_1(\tilde{U}, \tilde{V}) = \frac{1}{2} \int_{[0,1]} \left[|\inf \tilde{U}_\alpha - \inf \tilde{V}_\alpha| + |\sup \tilde{U}_\alpha - \sup \tilde{V}_\alpha| \right] d\alpha.$$

With this metric and approach, the 1-norm median for \mathcal{X} corresponds to

Definition 2.4 (Sinova *et al.* [17]) *If \mathcal{X} is a random fuzzy number associated with the probability space (Ω, \mathcal{A}, P) , a fuzzy median of \mathcal{X} is any fuzzy number $\tilde{\text{Me}}(\mathcal{X})$ such that*

$$\tilde{\text{Me}}(\mathcal{X}) = \arg \min_{\tilde{U} \in \mathcal{F}_c(\mathbb{R})} E \left(\rho_1(\mathcal{X}, \tilde{U}) \right),$$

where E denotes the mean value of the real-valued random variable it operates on.

The preceding minimization problem has been shown to have at least one operational solution. Thus, Sinova *et al.* [17] have proved that

Proposition 2.2 (Sinova *et al.* [17]) *A fuzzy median $\tilde{\text{Me}}(\mathcal{X})$ of the random fuzzy number \mathcal{X} is given by the fuzzy number such that*

$$(\tilde{\text{Me}}(\mathcal{X}))_\alpha = [\text{Me}^m(\inf \mathcal{X}_\alpha), \text{Me}^m(\sup \mathcal{X}_\alpha)],$$

where Me^m denotes the median of the corresponding real-valued random variable with the convention (if the median is not unique) of taking the mid-point of the interval of medians.

The last choice corresponds also to what has been suggested by Grzegorzewski to compute the sample fuzzy median, but it is not always the same as the one obtained on the basis of Zadeh's extension principle (Proposition 2.1). As a simple counterexample illustrating the last assertion one can consider a fuzzy random variable \mathcal{X} on a population in which it

takes on 4 different triangular fuzzy values, namely, $\text{Tri}(1, 1, 2)$, $\text{Tri}(1, 2, 3)$, $\text{Tri}(2, 3, 4)$ and $\text{Tri}(3, 4, 4)$, with respective probabilities .2, .3, .3 and .2, that is, with distribution given in Table 1.

values				
probabilities	.2	.3	.3	.2

Table 1: Induced distribution and values of a fuzzy random variable

Then, one can easily see that Zadeh's extension principle-based median is given by the trapezoidal fuzzy number $\tilde{\Gamma}(\mathcal{X}) = \text{Tra}(1, 2, 3, 4)$ (see Figure 1), whereas the 1-norm median (with the considered convention) is given by the triangular fuzzy number $\tilde{\text{Me}}(\mathcal{X}) = \text{Tri}(1.5, 2.5, 3.5)$ (see Figure 2).

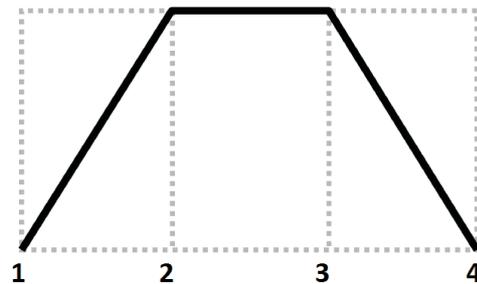


Figure 1: Zadeh's extension principle-based median of \mathcal{X}

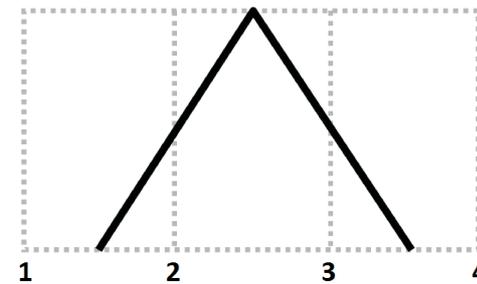


Figure 2: 1-norm median of \mathcal{X}

Remark 2.1 It should be remarked that in case the medians of the infimum or the supremum of any of the α -level mappings are not unique, there exist a continuum of valid fuzzy medians according to Definition 2.4. More concretely, at least the choice of any convex linear combination of the extremes of the interval of possible medians, $\lambda \text{inf Me}(X) + (1 - \lambda) \text{sup Me}(X)$, for an arbitrary $0 \leq \lambda \leq 1$, leads to a well-defined fuzzy median. One of these valid fuzzy medians corresponds to the one proposed by Grzegorzewski and agreeing with Zadeh's extension principle (Proposition 2.1).

On the other hand, by arguing like in Grzegorzewski [12], if we try to apply Zadeh's extension principle and, instead of employing the gen-

eral definition for the median of a real-valued random variable, we simply extend the more restrictive one (involving the usual convention of choosing the mid-point of the interval of medians in case of non-uniqueness), then we would obtain the median in Proposition 2.2.

3. Some basic properties of Zadeh's extension principle-based median of a fuzzy random variable

As it has been seen for the 1-norm median in Sinova *et al.* [17], Grzegorzewski's population fuzzy median also inherits several properties from the median in the real-valued case in connection with the arithmetic of fuzzy numbers based on Zadeh's extension principle. These properties can be summarized in the following

Proposition 3.1 $\tilde{\Gamma}$ is equivariant under 'linear' transformations, that is, if $\lambda \in \mathbb{R}$, $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$ and \mathcal{X} is a fuzzy random variable, then

$$\tilde{\Gamma}(\lambda \cdot \mathcal{X} + \tilde{U}) = \lambda \cdot \tilde{\Gamma}(\mathcal{X}) + \tilde{U}.$$

Consequently, if \mathcal{X} is a fuzzy random variable associated with the probability space (Ω, \mathcal{A}, P) and the distribution of \mathcal{X} is degenerate at a fuzzy number $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$ (i.e., $\mathcal{X} = \tilde{U}$ a.s. $[P]$), then $\tilde{\Gamma}(\mathcal{X}) = \tilde{U}$.

The median of a real-valued random variable is often defined as a 'middle position' value with respect to a specified ranking. Although fuzzy numbers cannot be ranked through a universally acceptable total ordering, it can be verified that Grzegorzewski's population fuzzy median of a fuzzy random variable can be also formalized as a certain 'middle position' value with respect to the *fuzzy max partial order*, whenever this order applies. The fuzzy max order on $\mathcal{F}_c(\mathbb{R})$ was introduced by Dubois and Prade [8]. It is the natural level-wise extension through the inf/sup representation of the product order on \mathbb{R}^2 , so that $\tilde{U} \preceq \tilde{V}$ if and only if for all $\alpha, \lambda \in [0, 1]$ one has that

$$\lambda \sup \tilde{U}_\alpha + (1 - \lambda) \inf \tilde{U}_\alpha \leq \lambda \sup \tilde{V}_\alpha + (1 - \lambda) \inf \tilde{V}_\alpha.$$

The main practical drawback for this ranking lies in the fact that it only leads to a partial ordering and many fuzzy numbers cannot be compared with it. However, it is often viewed as a widely accepted ranking criterion and as a pattern which should be preserved for any other suggested partial or total ranking. On the basis of this ranking, one can state that

Proposition 3.2 For any finite population $(\omega_1, \dots, \omega_n)$ for which the values of a fuzzy random variable \mathcal{X} satisfy that

$$\mathcal{X}(\omega_1) \preceq \dots \preceq \mathcal{X}(\omega_n)$$

we have that

- if n is odd, then

$$\tilde{\Gamma}(\mathcal{X}) = \mathcal{X}(\omega_{(n+1)/2}),$$

- if n is even, then

$$\tilde{\Gamma}(\mathcal{X}) = \text{any fuzzy number within the interval between } \mathcal{X}(\omega_{n/2}) \text{ and } \mathcal{X}(\omega_{(n/2)+1}),$$

where by the interval between fuzzy numbers \tilde{U} and \tilde{V} with $\tilde{U} \preceq \tilde{V}$ we mean the fuzzy number $[\tilde{U}, \tilde{V}]$ such that for each $\alpha \in [0, 1]$ we have that $[\tilde{U}, \tilde{V}]_\alpha = [\inf \tilde{U}_\alpha, \sup \tilde{V}_\alpha]$.

Remark 3.1 Although Proposition 3.2 links Grzegorzewski's approach with a well-known partial order on $\mathcal{F}_c(\mathbb{R})$, it should be recalled that the aim of this paper is not to propose medians that are coherent with the different ways of ordering fuzzy numbers. As it has been already said, there is not a universally acceptable total ordering on this space. Therefore, such kind of approaches may not be always applicable (if based on a partial order) or useful (due to the lack of agreement on total orderings), in contrast to the two fuzzy medians presented in this paper. Also notice that a lot of probabilistic concepts induced from the use of distances in the definition of the median should be explicitly introduced when using orderings and also the proof of the statistical and inferential properties of the corresponding median would be much more difficult to obtain.

Another interesting property in examining the adequacy of Grzegorzewski's population fuzzy median for fuzzy random variables as a location measure is now discussed by examining its behaviour in case of symmetrically distributed fuzzy random variables. Thus, in the real-valued case a well-known result is that the median of a symmetric random variable coincides with the point the variable is symmetric about whenever it is unique. Sinova *et al.* [16] have proved that the 1-norm median shows a suitable central tendency behaviour since it leads to a fuzzy number which is symmetric about the symmetry point. This property is also shared by Grzegorzewski's population fuzzy median.

Proposition 3.3 Let (Ω, \mathcal{A}, P) be a probability space, and let \mathcal{X} be a symmetric fuzzy random variable about $c \in \mathbb{R}$ (that is, $\mathcal{X} - c$ and $c - \mathcal{X}$ are identically distributed). Then, Grzegorzewski's population fuzzy median of \mathcal{X} is a symmetric fuzzy number about c .

Consequently, one can assert that for symmetric random fuzzy numbers about c , the Aumann-type mean, the 1-norm median and Grzegorzewski's population fuzzy median are symmetric fuzzy numbers about c . However, they do not necessarily coincide. This can be checked considering again as counterexample the fuzzy random variable in Table 1, which is symmetric with respect to 2.5. Both

medians have been displayed in Figures 1 and 2 and, although they are symmetric fuzzy numbers with respect to 2.5, they do not coincide. Recall that the well-known Aumann-type mean (see Aumann [1]) can be expressed, whenever the random fuzzy number \mathcal{X} associated with the probability space (Ω, \mathcal{A}, P) is integrably bounded (that is to say, $\max\{|\inf \mathcal{X}_\alpha|, |\sup \mathcal{X}_\alpha|\} \in L^1(\Omega, \mathcal{A}, P)$), as the fuzzy value $\tilde{E}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ such that for all $\alpha \in [0, 1]$ $(\tilde{E}(\mathcal{X}))_\alpha = [E(\inf \mathcal{X}_\alpha), E(\sup \mathcal{X}_\alpha)]$. By a straightforward computation, the Aumann-type mean is given by the triangular fuzzy number $\text{Tri}(1.7, 2.5, 3.3)$, which is also symmetric with respect to 2.5 but differs from both fuzzy medians.

4. Empirical analysis of the behaviour of Grzegorzewski's population fuzzy median

In this section, an empirical comparison of the behaviour of the Aumann-type mean and Grzegorzewski's fuzzy median will be presented. The aim is to show the reaction of these measures to the appearance of outliers or data changes. For some additional comments, the 1-norm median has also been considered in these simulations.

To ease the computations, only trapezoidal fuzzy numbers will be used, following the scheme detailed in [17] and randomly generated by means of the real-valued random variables $X_1 = (\inf \mathcal{X}_1 + \sup \mathcal{X}_1)/2$, $X_2 = (\sup \mathcal{X}_1 - \inf \mathcal{X}_1)/2$, $X_3 = \inf \mathcal{X}_1 - \inf \mathcal{X}_0$ and $X_4 = \sup \mathcal{X}_0 - \sup \mathcal{X}_1$. Although two different sample sizes ($n = 100$ and $n = 10000$) have been considered, only the results obtained with $n = 10000$ will be shown here, since the conclusions in the other situation are analogous. Notice that even sample sizes have been considered because the 1-norm median and Grzegorzewski's population fuzzy median coincide for odd sample sizes, and, therefore, this study would coincide with the one in [17].

Two cases have been examined when generating the fuzzy numbers. In Case 1, X_i are the following independent random variables: $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim \chi_1^2$, while in Case 2, the chosen distributions are dependent: $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim 1/(X_1^2 + 1)^2 + 0.1 \cdot \chi_1^2$. First, the population values of the three central tendency measures (Aumann-type mean, 1-norm median and Grzegorzewski's median) have been approximated using the Monte Carlo method with 1000 iterations.

Secondly, the procedure of introducing contamination in the study will be explained. The parameters c_p and C_D will denote the proportion and the size of the contamination, respectively. To analyze the effect of the perturbations on each measure, the (1-norm) distance between the corresponding population and sample values will be approximated by Monte Carlo with 1000 iterations. In other words, once fixed the parameters within their range ($c_p \in \{0, 0.1, 0.2, 0.4\}$ and $C_D \in \{0, 1, 5, 10, 100\}$), 1000 samples have been randomly generated accord-

ing to either Case 1 or Case 2. Afterwards, $[n \cdot c_p]$ observations in each sample, where $[\cdot]$ denotes the floor function, have been perturbed by using the distributions $X_1 \sim \mathcal{N}(0, 3) + C_D$ and/or

- $X_2, X_3, X_4 \sim \chi_4^2 + C_D$ in Case 1, or
- $X_2, X_3, X_4 \sim 1/(X_1^2 + 1)^2 + 0.1 \cdot \chi_1^2 + C_D$ in Case 2.

Notice that this way of modifying the observations covers all the possible situations: a perturbation in either 'location' or 'imprecision' (or both). For each of the 1000 polluted samples, the Aumann-type mean, and the 1-norm and the Grzegorzewski's medians have been computed, as well as the distance between these values and the population ones already approximated by Monte Carlo. Finally, the mean of these distances over the 1000 samples has been written in Tables 2 and 3.

c_p	C_D	Mean	1-norm	Grzegorzewski
0.0	0	0.0147198	0.0149612	0.0173286
0.0	1	0.0154872	0.0153962	0.0177638
0.0	5	0.0151089	0.0150653	0.0174278
0.0	10	0.0154439	0.0160397	0.0185496
0.0	100	0.0148953	0.0151692	0.0176065
0.1	0	0.3375262	0.1400632	0.1426220
0.1	1	0.4407420	0.1527727	0.1569648
0.1	5	0.9247706	0.1722154	0.1825132
0.1	10	1.4942703	0.1737683	0.1860467
0.1	100	12.117506	0.1761080	0.1874358
0.2	0	0.6775427	0.3130868	0.3179072
0.2	1	0.9083152	0.3589331	0.3671523
0.2	5	1.8160240	0.3957859	0.4218930
0.2	10	3.0367900	0.4114539	0.4389965
0.2	100	24.200746	0.4091132	0.4369104
0.4	0	1.3346467	0.7964349	0.8115211
0.4	1	1.8226241	1.0182947	1.0462244
0.4	5	3.6717042	1.2770609	1.3736092
0.4	10	5.9722961	1.3759261	1.5030230
0.4	100	48.700371	1.4018313	1.5197380

Table 2: Mean distances, approximated by Monte Carlo with 1000 iterations, between the sample and the population (also approximated by Monte Carlo) Aumann-type mean, 1-norm median and Grzegorzewski's median in Case 1

As it has been already commented, the main goal of this study was to show that, whereas the Aumann-type mean is highly influenced by outliers, Grzegorzewski's median is not: when increasing the contamination by means of the parameters c_p and C_D , the mean distance between the sample and the population Aumann-type mean gets very large, but the sample and population Grzegorzewski's medians remain much closer. Therefore, the behaviour of both medians is much more robust towards contamination than the Aumann-type mean's one. A remark is that if we also take into account the obtained outputs for the other fuzzy median, we see that the 1-norm median behaves a little better than Grzegorzewski's proposal in the sense that the sample estimate is closer (in ρ_1 -sense) to the corresponding population value. Notice that the use of the 1-norm metric is not especially favourable to

c_p	C_D	Mean	1-norm	Grzegorzewski
0.0	0	0.00910095	0.00469351	0.00558056
0.0	1	0.00908627	0.00455644	0.00543028
0.0	5	0.00899310	0.00456553	0.00544852
0.0	10	0.00910947	0.00462771	0.00552922
0.0	100	0.00924808	0.00460354	0.00548662
0.1	0	0.01542908	0.00514968	0.00625408
0.1	1	0.10764232	0.01940603	0.02179313
0.1	5	0.57892527	0.04007421	0.04428870
0.1	10	1.16919254	0.04238534	0.04646708
0.1	100	11.8587389	0.04320223	0.04683409
0.2	0	0.02534476	0.00697175	0.00837692
0.2	1	0.21594809	0.04087282	0.04585473
0.2	5	1.14670266	0.08619019	0.09660657
0.2	10	2.33501080	0.09155654	0.10191756
0.2	100	23.4408095	0.09191485	0.10143283
0.4	0	0.04638929	0.01315370	0.01510015
0.4	1	0.43555509	0.10004493	0.11397409
0.4	5	2.30052148	0.24407904	0.27057380
0.4	10	4.65427067	0.26474110	0.29121839
0.4	100	47.0408072	0.26254160	0.28475122

Table 3: Mean distances, approximated by Monte Carlo with 1000 iterations, between the sample and the population (also approximated by Monte Carlo) Aumann-type mean, 1-norm median and Grzegorzewski’s median in Case 2

the 1-norm median due to its definition, since also Grzegorzewski’s median minimizes the mean ρ_1 distance to the values the random fuzzy number takes, as already outlined in Remark 2.1.

5. Concluding remarks

The discussion in this paper motivates an immediate open direction, namely:

- the search of all possible solutions of the minimization problem in Definition 2.4,
- the study of the basic properties of these solutions,
- and the formal analysis of the statistical properties (mainly robustness and consistency) of the sample versions of different solutions.

Acknowledgements

The research in this paper has been partially supported by/benefited from the Spanish Ministry of Economy and Competitiveness Grant MTM2013-44212-P and the Principality of Asturias Grant GRUPIN14-101. Their financial support is gratefully acknowledged.

References

[1] R.J. Aumann, Integrals of set-valued functions, *Journal of Mathematical Analysis and Applications*, 12: 1–12, Elsevier, 1965.
[2] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárcana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo,

S. de la Rosa de Súa and B. Sinova, A distance-based statistical analysis of fuzzy number-valued data, *International Journal of Approximate Reasoning*, 55: 1487–1501, Elsevier, 2014.
[3] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárcana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo, S. De la Rosa de Súa and B. Sinova, Rejoinder on “A distance-based statistical analysis of fuzzy number-valued data”, *International Journal of Approximate Reasoning*, 55: 1601–1605, Elsevier, 2014.
[4] I. Couso and D. Dubois, Statistical reasoning with set-valued information: ontic vs. epistemic views, *International Journal of Approximate Reasoning*, 55: 1502–1518, Elsevier, 2014.
[5] I. Couso, S. Montes and P. Gil, Central tendency measures for random sets, *Proceedings of EUROFUSE-SIC’99*, 3–8, 1999.
[6] I. Couso, S. Montes and P. Gil, Función de distribución y mediana de variables aleatorias difusas, *Proceedings of Estylf’98*, 279–284, 1998.
[7] P. Diamond and P. Kloeden, Metric spaces of fuzzy sets, *Fuzzy Sets and Systems*, 35: 241–249, Elsevier, 1990.
[8] D. Dubois and H. Prade, Systems of linear fuzzy constraints, *Fuzzy Sets and Systems*, 3: 37–48, Elsevier, 1980.
[9] D. Dubois and H. Prade, Gradualness, uncertainty and bipolarity: making sense of fuzzy sets, *Fuzzy Sets and Systems*, 192: 3–24, Elsevier, 2012.
[10] M. Fréchet, Les éléments aléatoires de nature quelconque dans un espace distancié, *Annales de l’Institut Henri Poincaré*, 10: 215–310, Gauthier-Villars, 1948.
[11] G. González-Rodríguez, A. Colubi, M.A. Gil, Fuzzy data treated as functional data: A one-way ANOVA test approach, *Computational Statistics & Data Analysis*, 56: 943–955, Elsevier, 2012.
[12] P. Grzegorzewski, Statistical inference about the median from vague data, *Control and Cybernetics*, 27: 447–464, Polish Academy of Sciences, 1998.
[13] R. Kruse and K.D. Meyer, *Statistics with Vague Data*, D. Riedel Publishing Company, 1987.
[14] H. Kwakernaak, Fuzzy random variables. Part I: definitions and theorems, *Information Sciences*, 15: 1–29, Elsevier, 1978; Part II: algorithms and examples for the discrete case, *Information Sciences*, 17: 253–278, Elsevier, 1979.
[15] M.L. Puri and D.A. Ralescu, Fuzzy random variables, *Journal of Mathematical Analysis and Applications*, 114:409–422, Elsevier, 1986.
[16] B. Sinova, M.R. Casals and M.A. Gil, Central tendency for symmetric random fuzzy num-

- bers, *Information Sciences*, 278: 599–613, Elsevier, 2014.
- [17] B. Sinova, M.A. Gil, A. Colubi and S. Van Aelst, The median of a random fuzzy number. The 1-norm distance approach, *Fuzzy Sets and Systems*, 200: 99–115, Elsevier, 2012.
- [18] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning. Part 1, *Information Sciences*, 8: 199–249, Elsevier, 1975; Part 2, *Information Sciences*, 8: 301–353, Elsevier, 1975; Part 3, *Information Sciences*, 9: 43–80, Elsevier, 1975.