Int. J. Appl. Math. Comput. Sci., 2022, Vol. 32, No. 2, 285–297
DOI: 10.34768/amcs-2022-0021

BOOTSTRAP METHODS FOR EPISTEMIC FUZZY DATA

PRZEMYSŁAW GRZEGORZEWSKI [a,b,*], MACIEJ ROMANIUK [b,c]

[a]Faculty of Mathematics and Information Science
Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland
e-mail: Przemyslaw.Grzegorzewski@pw.edu.pl

[b]Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland

[c]Warsaw School of Information Technology
Newelska 6, 01-447 Warsaw, Poland
e-mail: mroman@ibspan.waw.pl

Fuzzy numbers are often used for modeling imprecise perceptions of the real-valued observations. Such epistemic fuzzy data may cause problems in statistical reasoning and data analysis. We propose a universal nonparametric technique, called the epistemic bootstrap, which could be helpful when the existing methods do not work or do not give satisfactory results. Besides the simple epistemic bootstrap, we develop its several refinements that aim to reduce the variance in statistical inference. We also perform an extended simulation study to examine statistical properties of the approaches considered. The discussion of the results is supplemented by some hints for practical use.

Keywords: bootstrap, estimation, fuzzy data, fuzzy numbers, hypotheses testing, resampling.

1. Introduction

The bootstrap introduced by Efron (1979) turned out to be extremely valuable in countless applications. As noticed by Casella (2003), "the bootstrap has shown us how to use the power of the computer and iterated calculations to go where theoretical calculations cannot, which introduces a different way of thinking about all of statistics". It owes its great success and recognition to its low requirements (no assumptions are made about the distribution of the sample, like normality; samples do not have to be large, etc.), openness to different types of data, and general ease of use. Consequently, the bootstrap can be applied in many areas where the methods used so far do not work or do not give satisfactory results. In particular, the bootstrap turned out to be very useful in statistics with fuzzy data.

Fuzzy data have drawn increasing interest in recent years. They appear in various fields and applications. In particular, real-valued random variables are often imprecisely observed or they are so uncertain that the results are recorded as fuzzy numbers which model the precise outcomes of the experiment. There are also situations where the exact values of some variable are hidden deliberately because of the confidentiality reasons. In all such cases as mentioned above, fuzzy data represent the epistemic state of an agent so they are called *epistemic* (Couso and Dubois, 2014). On the other hand, there are situations when the experimental data appear as essentially fuzzy-valued, e.g., when we collect perceptions with no objective values behind or when we describe regions with intrinsically gradual boundaries, etc. Such data represent an objective entity and hence they are called *ontic* (Couso and Dubois, 2014).

So far, applications of the bootstrap in fuzzy data analysis have been limited to ontic fuzzy data. For instance, the bootstrap turned out to be very useful in hypotheses testing (Colubi *et al.*, 2002; Gil *et al.*, 2006; González-Rodríguez *et al.*, 2006; Montenegro *et al.*, 2004; Ramos-Guajardo and Lubiano, 2012), classification

*Corresponding author

(Ramos-Guajardo and Grzegorzewski, 2016), fuzzy rating in questionnaires (Lubiano *et al.*, 2016; 2017), quality control (Ramos-Guajardo *et al.*, 2019; Wang and Hryniewicz, 2015), and so on. Meanwhile, although the epistemic fuzzy data seem to be more natural in engineering and in other fields where we deal with imprecise measurement results, bootstrap methods have not been developed yet. The aim of this contribution is to fill this gap.

Some preliminary ideas on the possibility of the bootstrap application for epistemic fuzzy data were proposed by Grzegorzewski and Romaniuk (2021). In this paper, we develop and compare various methods (like the so-called antithetic approach, or the RSS-based resampling) to improve this simple epistemic bootstrap, especially intending to reduce the variance. Apart from broad numerical experiments related to various statistical fields (like the estimation of the standard error and the mean squared error, comparison of the power curves), some practical hints are also provided.

The paper is organized as follows. Basic notions on fuzzy data and the epistemic view on statistics with fuzzy data are recalled in Section 2. In Section 3 we propose how to perform the bootstrap in the framework of epistemic fuzzy data. Further on we present the results of the extended simulation study on various aspects of the epistemic bootstrap. We start from its statistical justification in Section 4. Next, in Sections 5 and 6, we consider the epistemic bootstrap in estimation and hypotheses testing with epistemic fuzzy data, respectively. We discuss different resampling methods and compare the results with other methods used so far. Theoretical considerations and analyses of the results are supplemented with guidance for practitioners.

2. Basic notions and notation

2.1. Fuzzy data. Traditionally, most of the experimental results are real-valued data and usually statistical methods refer to such data. However, real-life measurements are quite often imprecise. Moreover, in many situations where the outcomes are even real-valued, we are actually faced with their perceptions which are not necessarily precise but somehow vague. Fuzzy set theory delivers effective tools for modeling imprecision and its analysis. A natural counterpart of the real-valued outcomes $x_1, \ldots, x_n$, where each $x_i \in \mathbb{R}$ is just a real number, are fuzzy numbers $\widetilde{x}_1, \ldots, \widetilde{x}_n$.

More specifically, $\widetilde{x} : \mathbb{R} \to [0,1]$ is a *fuzzy number* if its α-cuts $(\widetilde{x})_\alpha$ are nonempty compact intervals for all $\alpha \in [0,1]$, where

$$(\widetilde{x})_\alpha = \begin{cases} \{x \in \mathbb{R} : \widetilde{x} \geq \alpha\} & \text{if } \alpha \in (0,1], \\ \mathrm{cl}\{x \in \mathbb{R} : \widetilde{x} > 0\} & \text{if } \alpha = 0, \end{cases}$$

and where ‘cl’ stands for the closure. A family of all fuzzy numbers will be denoted further on by $\mathbb{F}(\mathbb{R})$.

Two α-cuts of a fuzzy number $\widetilde{x}$ are of special interest: $(\widetilde{x})_{\alpha=0}$ called the *support* and $(\widetilde{x})_{\alpha=1}$ known as the *core*. The support contains all real values that are possible realizations (at least to some extent) of the object perceived as $\widetilde{x}$. On the other hand, the core contains all reals totally compatible with the notion described by $\widetilde{x}$.

A fuzzy number, as a function, may assume different shapes. However, the most common fuzzy numbers are the *trapezoidal fuzzy numbers* of the form

$$\widetilde{x}(x) = \begin{cases} \frac{x-a}{b-a} & \text{if } a < x \leq b, \\ 1 & \text{if } b \leq x \leq c, \\ \frac{d-x}{d-c} & \text{if } c \leq x < d, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $a, b, c, d \in \mathbb{R}$ such that $a \leq b \leq c \leq d$. Such a trapezoidal fuzzy number $\widetilde{x}$ is often denoted as $\mathrm{Tra}(a,b,c,d)$. If $b = c$ then $\widetilde{x}$ is said to be a *triangular fuzzy number*.

The reason for restricting attention to triangular or trapezoidal fuzzy numbers is their simplicity since they are easy to handle and have a natural interpretation. Moreover, even if the original data set consists of fuzzy numbers of other types, one may easily approximate them by such fuzzy numbers. The broad collection of approximation methods satisfying various requirements can be found in the work of Ban *et al.* (2015).

To define basic arithmetic operations in $\mathbb{F}(\mathbb{R})$, we use natural α-cut-wise operations on intervals. In particular, the sum of two fuzzy numbers $\widetilde{x}$ and $\widetilde{y}$ is given by the Minkowski addition of the corresponding α-cuts, i.e.,

$$(\widetilde{x} + \widetilde{y})_\alpha = \big[\inf(\widetilde{x})_\alpha + \inf(\widetilde{y})_\alpha, \sup(\widetilde{x})_\alpha + \sup(\widetilde{y})_\alpha\big],$$

for all $\alpha \in [0,1]$. Similarly, the product of a fuzzy number $\widetilde{x}$ by a scalar $\theta \in \mathbb{R}$ is defined by the Minkowski scalar product for intervals, i.e., for all $\alpha \in [0,1]$

$$\begin{aligned}(\theta \cdot \widetilde{x})_\alpha = [&\min\{\theta \inf(\widetilde{x})_\alpha, \theta \sup(\widetilde{x})_\alpha\}, \\ &\max\{\theta \inf(\widetilde{x})_\alpha, \theta \sup(\widetilde{x})_\alpha\}] .\end{aligned}$$

It is worth noting that the sum of trapezoidal fuzzy numbers is also a trapezoidal fuzzy number. Indeed, if $\widetilde{x} = \mathrm{Tra}(a_1, b_1, c_1, d_1)$ and $\widetilde{y} = \mathrm{Tra}(a_2, b_2, c_2, d_2)$ then

$$\widetilde{x} + \widetilde{y} = \mathrm{Tra}(a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2).$$

Moreover, the product of a trapezoidal fuzzy number $\widetilde{x} = \mathrm{Tra}(a,b,c,d)$ by a scalar θ is a trapezoidal fuzzy number

$$\theta \cdot \widetilde{x} = \begin{cases} \mathrm{Tra}(\theta \cdot a, \theta \cdot b, \theta \cdot c, \theta \cdot d) & \text{if } \theta \geq 0, \\ \mathrm{Tra}(\theta \cdot d, \theta \cdot c, \theta \cdot b, \theta \cdot a) & \text{if } \theta < 0. \end{cases}$$

Unfortunately, $\big(\mathbb{F}(\mathbb{R}), +, \cdot\big)$ has only a semilinear structure since in general $\widetilde{x} + (-1 \cdot \widetilde{x}) \neq \mathbb{1}_{\{0\}}$.

Consequently, the Minkowski-based difference does not satisfy, in general, the addition/subtraction property that $(\widetilde{x} + (-1 \cdot \widetilde{y})) + \widetilde{y} = \widetilde{x}$. To overcome this problem, the so-called Hukuhara difference can be considered (Hukuhara, 1967). Although now the desired properties are satisfied, the Hukuhara difference does not always exist. Therefore, one should be aware that there are critical problems with a subtraction in $\mathbb{F}(\mathbb{R})$.

In general, due to the extension principle (Zadeh, 1973) any function $f : \mathbb{R}^n \to \mathbb{R}$ induces $\tilde{f} : \mathbb{F}(\mathbb{R})^n \to \mathbb{F}(\mathbb{R})$ defined by

$$\begin{aligned} \tilde{y}(y) &= [\tilde{f}(\tilde{x}_1, \ldots \tilde{x}_n)](y) \qquad (2) \\ &= \sup_{(x_1 \ldots, x_n) \in \mathbb{R}^n : y = f(x_1 \ldots, x_n)} \ \min_{i=1,\ldots,n} \tilde{x}(x), \end{aligned}$$

which allows us to extend any operation defined on real numbers to an operation on fuzzy numbers. Note that an attempt to define a different arithmetic on fuzzy sets designed to avoid some of its shortcomings was proposed by Piegat (2005).

2.2. Epistemic view on statistics with fuzzy data. When random experiments lead to imprecise data which can be properly described by fuzzy values, the mechanisms generating such elements can be treated as fuzzy-valued random variables (Kwakernaak, 1978; Kruse, 1982).

Definition 1. Given a probability space $(\Omega, \mathcal{F}, P)$, a mapping $\widetilde{X} : \Omega \to \mathbb{F}(\mathbb{R})$ is said to be a *fuzzy random variable* (f.r.v.) if for each $\alpha \in [0, 1]$ $(\inf \widetilde{X}_\alpha) : \Omega \to \mathbb{R}$ and $(\sup \widetilde{X}_\alpha) : \Omega \to \mathbb{R}$ are real-valued random variables on $(\Omega, \mathcal{F}, P)$.

Looking for an interpretation of a fuzzy random variable, we may consider $\widetilde{X}$ as a *fuzzy perception* of a usual random variable X, called the *original* of $\widetilde{X}$, which unfortunately remains unknown. Similarly, a *fuzzy random sample* $\widetilde{X}_1, \ldots, \widetilde{X}_n$ may be treated as a fuzzy perception of a random sample $X_1, \ldots, X_n$ of the usual real-valued random variables.

Following the extension principle (2) any statistic $T = T(X_1, \ldots, X_n)$ can be extended to a fuzzy statistic $\widetilde{T} = \widetilde{T}(\widetilde{X}_1, \ldots, \widetilde{X}_n)$. Given any realization $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of a fuzzy random sample $\widetilde{X}_1, \ldots, \widetilde{X}_n$, the corresponding value of the fuzzy statistic $\widetilde{T}$ is a fuzzy set $\widetilde{T}(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ described by the following family of its α-cuts

$$\begin{aligned} \widetilde{T}_\alpha &= \widetilde{T}(\widetilde{x}_1, \ldots, \widetilde{x}_n) \\ &= \big\{ T(x_1, \ldots, x_n) : x_1 \in (\widetilde{x}_1)_\alpha, \ldots, x_n \in (\widetilde{x}_n)_\alpha \big\}. \end{aligned}$$

If T is regular (monotone, continuous, etc.), then the α-cuts of $\widetilde{T}$ are intervals and so to obtain $\widetilde{T}_\alpha$ it is sufficient to find $\inf \widetilde{T}_\alpha$ and $\sup \widetilde{T}_\alpha$. Thus computing statistics under fuzzy uncertainty can be reduced to several problems of computing statistics under interval uncertainty. Sometimes it works smoothly. For instance, the average of a fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n$ is calculated immediately using the Minkowski sum and the scalar product; hence we obtain the following α-cuts:

$$\Big(\frac{1}{n} \sum_{i=1}^{n} \widetilde{x}_i \Big)_\alpha = \Big[\frac{1}{n} \sum_{i=1}^{n} \inf(\widetilde{x}_i)_\alpha, \frac{1}{n} \sum_{i=1}^{n} \sup(\widetilde{x}_i)_\alpha \Big],$$

where $(\widetilde{x}_i)_\alpha = [\inf(\widetilde{x}_i)_\alpha, \sup(\widetilde{x}_i)_\alpha]$, $i = 1, \ldots, n$.

Actual calculations are not always that straightforward even if strict mathematical formulas exist. For example, the sample variance of a fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n$ is defined by the extension principle as follows:

$$\widetilde{s}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \Big(\widetilde{x}_i - \Big(\frac{1}{n} \sum_{i=1}^{n} \widetilde{x}_i \Big) \Big)^2,$$

where

$$\left(\widetilde{s}^2\right)_\alpha = \left[\inf \left(\widetilde{s}^2\right)_\alpha, \sup \left(\widetilde{s}^2\right)_\alpha \right],$$

$\alpha \in [0, 1]$, denote its α-cuts. Unfortunately, computing $\sup \left(\widetilde{s}^2\right)_\alpha$ is NP-hard (Vavasis, 1991). Although there are algorithms that compute $\overline{S^2}$ in a more effective time (i.e., $O(n \log n)$ or even $O(n)$) but they impose some restrictions on the fuzzy sample (Nguyen *et al.*, 2012). A new approximate algorithm based on asymptotic reasoning for computing the upper bound of the sample variance which works in the $O(n)$ time was proposed recently (Kołacz and Grzegorzewski, 2019).

It is also worth emphasizing that even if obtaining $\widetilde{T}(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ involves no calculation problems, it may not be satisfying for practitioners, especially if the ranges of α-cuts of $\widetilde{T}$ are too large. Consequently, the solutions offered in such cases are too conservative and hence do not fully meet the expectations of potential users. Therefore, researchers are faced with the task of constructing statistical procedures which—despite imprecise input data—lead to “more precise” final decisions. In parametric models, i.e., when the population distribution is known up to a parameter value, several approaches to improve estimation based on imprecise data were discussed by Grzegorzewski and Goławska (2021). However, it appears that under imprecision nonparametric (distribution-free) methods would be much more desirable. Such a new promising nonparametric method is discussed in the next section.

3. Epistemic bootstrap

Suppose, our sample $X_1, \ldots, X_n$ consists of n independent and identically distributed random variables from the unknown distribution. However, instead of a real-valued realization $(x_1, \ldots, x_n)$ of this sample we observe only its imprecise perception modeled by a fuzzy sample $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$. We assume that for each $i = 1, \ldots, n$ the fuzzy set $\widetilde{x}_i$ contains an actual real-valued realization of the i-th observation but we do not know where it is precisely located. Instead, a membership function of $\widetilde{x}_i$ attributes to each point the possibility that this very point is the true realization x_i of the random variable X_i.

Therefore, a fuzzy perception $\widetilde{x}_i$ might be considered as a fuzzy neighbor of the desired but unknown real value x_i. Hence, by an appropriate selection of the element x_i^* from $\widetilde{x}_i$ which takes into account the degree of possibility that x_i^* is the true outcome of the experiment, we obtain an innovative technique to reconstruct a real-valued sample from the epistemic fuzzy sample. Its main idea is based on a random selection of elements from the input sample which resembles somehow the bootstrap. Since the suggested method can be adapted only to epistemic fuzzy samples (not ontic), it should come as no surprise that this approach has been called the *epistemic bootstrap* (Grzegorzewski and Romaniuk, 2021).

Suppose we have a fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}(\mathbb{R})$. To generate a single element x_i^* of a bootstrap sample, we need just one loop with two steps: firstly, we generate an α-cut, and secondly, we draw randomly a real value from this very α-cut. More specifically, we generate randomly

Step 1: a real number α_i from the uniform distribution on the unit interval $[0, 1]$, i.e., $\alpha_i \sim \mathrm{U}[0, 1]$,

Step 2: a real number x_i^* from the uniform distribution on the α-cut $(\widetilde{x}_i)_{\alpha_i}$, i.e., $x_i^* \sim \mathrm{U}\big[\inf(\widetilde{x}_i)_{\alpha_i}, \sup(\widetilde{x}_i)_{\alpha_i}\big]$.

Proceeding in this way n times, we complete the entire bootstrap sample $x_1^*, \ldots, x_n^*$. But to settle for a single bootstrap sample does not seem to be enough. Indeed, following the aforementioned two steps, we consider only one α-cut for each fuzzy observation. Here one may ask a natural question: Why should we limit ourselves to a single α-cut? Let us recall that the bootstrap world, as proposed by Efron, is very generous: we can generate as many bootstrap samples as we want or we have time for. Therefore, we may generate several (say, $B \geq 1$) α-cuts for each fuzzy set which provides a multiplicity of bootstrap samples, i.e., $(x_{1j}^*, \ldots, x_{nj}^*)$, where $j = 1, \ldots, B$, as it can be seen in Fig. 1. This general idea of drawing bootstrap samples from epistemic fuzzy data is also shown in Algorithm 1.

In solving practical problems we usually have only one random sample $(x_1, \ldots, x_n)$ of experimental outcomes. In turn, it gives a single value of the desired statistic $T = T(x_1, \ldots, x_n)$ necessary to achieve the desired goal, like an estimator, a test statistic, etc. However, in the bootstrap world each bootstrap sample provides another realization $T_j^*(x_{1j}^*, \ldots, x_{nj}^*)$ of the statistic of interest, called the bootstrap replication. Thus,

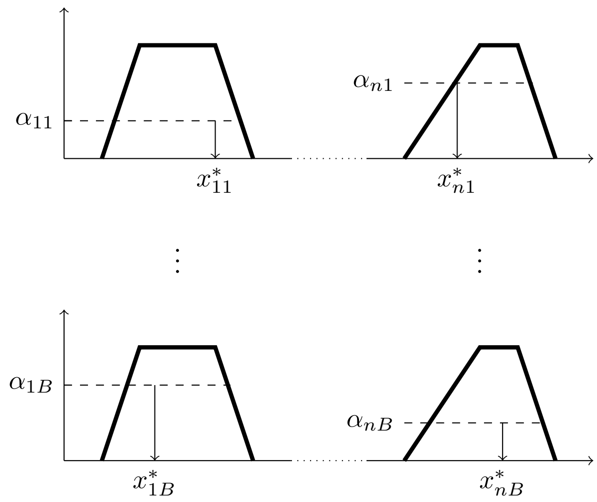

Fig. 1. General idea of the epistemic bootstrap, i.e., drawing bootstrap samples from epistemic fuzzy data.

Algorithm 1. Epistemic fuzzy bootstrap.

Require: Initial fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}(\mathbb{R})$.
Ensure: B bootstrap samples.
  for $j = 1$ to B **do**
    for $i = 1$ to n **do**
      Generate randomly a real number α_{ij} from the uniform distribution on the unit interval $[0, 1]$.
      Generate randomly a real number x_{ij}^* from the uniform distribution on the α-cut $(\widetilde{x}_i)_{\alpha_{ij}}$.
    end for
  end for
  Bootstrap samples $x_{1j}^*, \ldots, x_{nj}^*$, where $j = 1, \ldots, B$.

following the methodology introduced by Efron, the final bootstrap statistic T^* is obtained by some aggregation of the bootstrap replications $T_1^*, \ldots, T_B^*$. The most common solution is to aggregate the bootstrap replications by simple averaging, i.e., to consider $T^* = \frac{1}{B}\sum_{j=1}^{B} T_j^*$ in further inference, like to determine a bootstrap estimator or to evaluate the standard error, to design bootstrap confidence intervals or to verify hypotheses. Some applications of the proposed approach in reasoning with epistemic fuzzy data are discussed in Section 4.

Here a natural question arises on why the uniform distribution, not another, is applied in generating random values in both steps of the suggested method. It seems that several justifications can be made.

Firstly, in accordance with the principle of maximum entropy, if nothing is known about the distribution then the least informative one should be chosen as default. Thus information theory shows that the distribution with the greatest entropy is the desired one. Several motivations justify this claim. In particular, physical systems usually tend to a configuration that maximizes entropy over time. Moreover, by maximizing entropy we minimize the amount of prior information contained in the distribution. Now, keeping in mind that the bootstrap approach needs a distribution on a closed interval (both each α-cut $(\widetilde{x}_i)_{\alpha_i}$ is a closed interval and the range of admissible α-cuts is a closed interval $[0,1]$), we have to include the well-known result that the uniform distribution on a given interval is the maximum entropy distribution among all continuous distributions having support which coincide with this interval. This very property can be also related to the Laplace principle of indifference.

Going back for motivations for applying the uniform distribution one may also indicate a correspondence with Effron's classical bootstrap where the bootstrap samples are drawn randomly from the primary sample with equal probabilities, which means that they are generated from the uniform distribution, but on a finite set.

Finally, using the uniform distribution prevents repetitions from occurring which is the bane of the classical bootstrap where most of the so generated bootstrap samples contain repeated values. Even worse, if the sample size n is small, all resulting bootstrap samples usually consist of only a few distinct observations, which is a highly undesirable effect, especially if the unknown population distribution is continuous (so the probability of at least two identical observations equals zero). To overcome this problem one has to apply some modified versions of the bootstrap like the smoothed bootstrap for real-valued data (De Angelis and Young, 1992; Hall *et al.*, 1989; Silverman and Young, 1987) or a flexible resampling for ontic fuzzy data (Grzegorzewski *et al.*, 2019; 2020a; 2020b; Romaniuk, 2019; Romaniuk and Hryniewicz 2019; 2021), whereas the epistemic bootstrap we propose is free from such problems under any circumstances and regardless of the primary sample size.

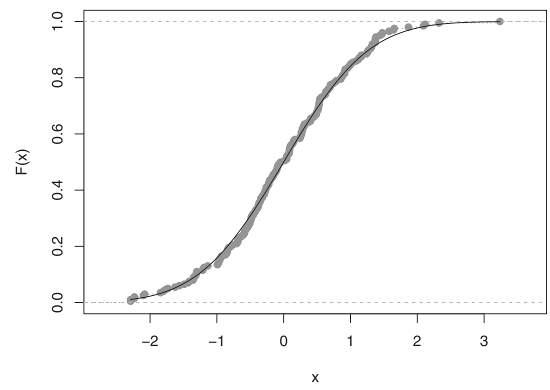

Fig. 2. Empirical vs. theoretical c.d.f. for the normal distribution and the epistemic bootstrap from $\mathbb{F}_{(\mathrm{N,U,U},1)}$ performed on a single α-cut, i.e., $B=1$.

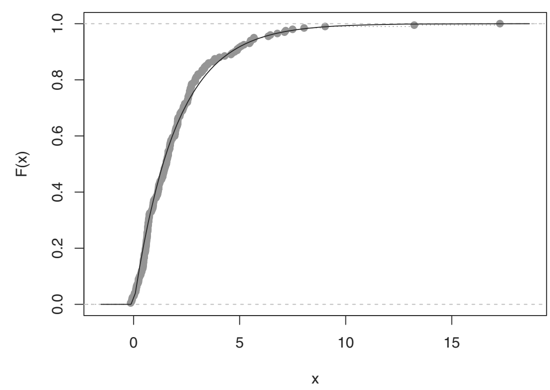

Fig. 3. Empirical vs. theoretical c.d.f. for the exponential distribution and the epistemic bootstrap from $\mathbb{F}_{(\mathrm{E,U,U},1)}$ performed on $B=10$ α-cuts.

4. Empirical bootstrap and its statistical justification

Before applying our new bootstrap method in statistical reasoning let us check whether it can approximate the actual population distribution.

Let $X_1,\dots,X_n$ denote a sample of independent and identically distributed (i.i.d.) random variables with the cumulative distribution function (c.d.f.) F. The Glivenko–Cantelli lemma proves that the empirical distribution function (e.d.f.) $\widehat{F}$ based on $X_1,\dots,X_n$ converges with probability 1 to F as n tends to infinity. This means that the empirical distribution based solely on the empirical results approaches the actual c.d.f. if the sample size is large enough.

Let $(\widetilde{x}_1,\dots,\widetilde{x}_n)$ denote a fuzzy perception of the random sample and let $(x^*_{1j},\dots,x^*_{nj})$, where $j=1,\dots,B$, be a collection of the bootstrap samples obtained from the initial fuzzy sample $(\widetilde{x}_1,\dots,\widetilde{x}_n)$ according to Algorithm 1.

Set $\widehat{F}^*_j$ as the bootstrap counterpart of the e.d.f. $\widehat{F}$ based on the j-th bootstrap sample $(x^*_{1j},\dots,x^*_{nj})$, where

$$\widehat{F}^*_j(t)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{(-\infty,x^*_{ij})}(t).$$

Many simulation experiments were performed to check if the e.d.f. $\widehat{F}^*$ estimates F well. Some exemplary results are given in Figs. 2 and 3 showing both the theoretical c.d.f. and the e.d.f. obtained from fuzzy samples with the epistemic bootstrap. A comparison pictured in Fig. 2 corresponds to data from the normal distribution, while in Fig. 3 we illustrate results obtained for the exponential distribution. In both cases, the sample size $n=200$ and all simulations were performed using the R package.

Fuzzy samples $\widetilde{x}_1,\dots,\widetilde{x}_n$ used in simulations consist of trapezoidal fuzzy numbers of the form (1).

Obviously, the primary reason for using trapezoidal fuzzy numbers is numerical simplicity since to generate $\widetilde{x}_i \in \mathbb{F}(\mathbb{R})$ one needs only four real numbers which characterize the endpoints of its support $[a_i, d_i]$ and core $[b_i, c_i]$, respectively. However, this is not the only reason why we use such fuzzy numbers. Trapezoidal fuzzy numbers are not only easy to handle but they have a natural interpretation (Ban *et al.*, 2015) which usually suffices in fuzzy modeling of uncertainty (Pedrycz, 1994). Moreover, if the original fuzzy sample consists of arbitrary fuzzy numbers, one often approximates them by trapezoidal fuzzy numbers before further processing (for a review of approximation algorithms satisfying various requirements, we refer the reader to Ban *et al.* (2015)). All this makes trapezoidal fuzzy numbers standard tools in statistical simulation and many other numerical computations.

Thus, to obtain a fuzzy number $\widetilde{x}_i$, we generated $a_i \le b_i \le c_i \le d_i$, for $i = 1, \dots, n$, according to the following formulas:

$$\begin{aligned} a_i &= X - S^l - C^l, & b_i &= X - C^l, \\ c_i &= X + C^r, & d_i &= X + C^r + S^r, \end{aligned}$$

where X is a random variable corresponding to the "true" population distribution, while C^l, C^r, S^l and S^r are some random variables applied for modeling fuzzy perception of the crisp input. These five random variables were generated independently from various distributions described in Table 1. For example, $\widetilde{x}_i \in \mathbb{F}_{(\mathrm{N,U,U,1})}$ means that X is simulated from the standard normal distribution, while both random variables C^l and C^r are generated from the uniform distribution on the interval $(0, 0.6)$ and both random variables S^l and S^r are simulated from the uniform distribution on the interval $(0, 0.8)$. The last two types included in Table 1, i.e., $\mathbb{F}_{(\beta,\mathrm{Ucond},1)}$ and $\mathbb{F}_{(\beta,\mathrm{Ucond},2)}$ are simulated in a more complex way described in detail by Lubiano *et al.* (2017).

The same distributions for C^l, C^r (and S^l, S^r, respectively) are used because in the non-parametric setting we cannot assume additional knowledge about some "asymmetric" tendency for X. If these distributions are not the same, it leads to a permanent left- or right-hand shift of X compared with its simulated fuzzy counterpart $\widetilde{x}$, and to a possible bias of the obtained estimator. In such a case, we need additional assumptions about the value of this shift to remove this bias, e.g., by adding some weights in the averaging procedure. For a similar approach concerning simulations of fuzzy variables, see, e.g., the work of Lubiano *et al.* (2017), Romaniuk and Hryniewicz (2019) or Grzegorzewski *et al.* (2020b).

Here we actually touch on the problem of a possible bias in data. Generally, it is usually difficult to tell if the available data are biased, especially if the data are human-generated. Moreover, bias in data is a complex phenomenon and there are various sources and types of it including the bias called historical, the representation bias, measurement bias, and so on (Suresh and Guttag, 2021). To prevent bias, one may try to make sure that samples are representative and not convenient ones (when in doubt, to use additional randomization), to use diverse data sources if possible, etc. Anyway, identifying bias in data is often difficult or sometimes even not tractable.

Table 1. Simulation scenarios for fuzzy samples.

Type	X	C^l, C^r	S^l, S^r
$\mathbb{F}_{(\mathrm{N,U,U,1})}$	N(0,1)	U(0,0.6)	U(0,0.8)
$\mathbb{F}_{(\mathrm{N,U,U,2})}$	N(0,2)	U(0,1)	U(0,2)
$\mathbb{F}_{(\mathrm{E,U,U,1})}$	Exp(0.5)	U(0,0.4)	U(0,0.6)
$\mathbb{F}_{(\mathrm{E,U,U,2})}$	Exp(1)	U(0,0.6)	U(0,1.2)
$\mathbb{F}_{(\mathrm{N,E,U})}$	N(0,1)	Exp(1)	U(0,0.8)
$\mathbb{F}_{(\Gamma,\mathrm{U,U,1})}$	$\Gamma(1,1)$	U(0,0.5)	U(0,0.8)
$\mathbb{F}_{(\Gamma,\mathrm{U,U,2})}$	$\Gamma(2,2)$	U(0,0.5)	U(0,0.8)
$\mathbb{F}_{(\Gamma,\mathrm{E,E})}$	$\Gamma(2,2)$	Exp(0.5)	Exp(1)
$\mathbb{F}_{(\mathrm{W,E,E})}$	Weibull (1, 1.5)	Exp(2)	Exp(2)
$\mathbb{F}_{(\mathrm{W,U,U})}$	Weibull (1, 1.5)	U(0,0.6)	U(0,0.8)
$\mathbb{F}_{(\beta,\mathrm{Ucond},1)}$	$\beta(2,2)$	U (conditional)	
$\mathbb{F}_{(\beta,\mathrm{Ucond},2)}$	$\beta(4,2)$	U (conditional)	

Figures 2 and 3 show that if the sample size is large enough then, even if the bootstrap samples are drawn using a single α-cut, our method provides quite a satisfying approximation of the population distribution.

Although the figures shown in this section illustrate a very limited situation (two distributions only), the other experiments also confirmed the desired behavior of the e.d.f. based on the epistemic bootstrap that if the epistemic fuzzy sample size is large enough then $\widehat{F}^*$ tends to the population c.d.f. F.

5. Point estimation based on the epistemic bootstrap

Some simulation results of point estimation with the epistemic bootstrap were published by Grzegorzewski and Romaniuk (2021). Although different simulation scenarios (like those given in Table 1), sample sizes n, and the numbers B of the α-cuts considered were discussed, the results confirmed that estimators were consistent. An interesting and somewhat surprising result is that increasing B does not substantially improve the estimation quality. Unfortunately, it turned out that the standard error of the estimators considered was relatively large. This can be explained by the overlap of the two sources of variability: the first one inherently connected with randomness, and the second, caused by imprecision.

Hence the resulting standard error of an estimator actually measures the overall effect instead of focusing on that related to randomness. Thus, the immediate conclusion was to enrich estimation with a variance reduction method or to learn how to exclude the diminished effects of data imprecision.

Following this idea, we performed an extensive numerical experiment to examine several improvements of our standard epistemic bootstrap (further on called simple bootstrap) with various refinements including the so-called antithetic sampling, the ranked set sampling (RSS) (Wolfe, 2004), usual jackknife and the jackknife combined with the antithetic approach. In the above-mentioned antithetic method, two random numbers are generated based on the "the same" draw α_{ij} for $\widetilde{x}_i$, the first one from the α_{ij}-cut and the second one from its complement, i.e., the $1-\alpha_{ij}$-cut, respectively (see Algorithm 2). This idea resembles in some way the antithetic approach which is aimed at reducing the variance of the MC methods (Kroese *et al.*, 2011).

We considered different point estimators and various population distributions as described in Table 1. Some simulation results can be found in Tables 2–5. To make the comparison easier, the best estimates in each category are given in boldface.

We were mostly interested in estimating the population mean and variance. We estimated the mean with the average $\overline{X}$, while to estimate the variance we used both the well-known sample variance S^2 as well as its corrected version W of the form

$$\mathrm{varW} = \frac{1}{B}\sum_{b=1}^{B}\frac{1}{n-1}\sum_{i=1}^{n}\left(x^*_{ib}-\frac{1}{n}\sum_{k=1}^{n}x^*_{kb}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{B-1}\sum_{b=1}^{B}\left(x^*_{ib}-\frac{1}{B}\sum_{j=1}^{B}x^*_{ij}\right)^2, \quad (3)$$

which, in some way, separates the within- and between-group variations.

Each experiment in the simulation study was repeated $m = 1000$ times. The estimated standard error (SE) and the mean squared error (MSE) of each estimator are calculated from the formulas

$$\mathrm{SE}(\hat{\theta}) = \frac{1}{m}\sum_{i=1}^{m}\sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\theta}^*_{ib}-\frac{1}{B}\sum_{j=1}^{B}\hat{\theta}^*_{ij}\right)^2}, \quad (4)$$

$$\mathrm{MSE}(\hat{\theta}) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\theta}^*_{ib}-\theta\right)^2, \quad (5)$$

where the first subscript in $\hat{\theta}^*_{ib}$ stands for the i-th the repetition of the experiment while the second subscript indicates the b-th bootstrap sample.

Algorithm 2. Antithethic epistemic fuzzy bootstrap.

Require: Initial fuzzy sample $\widetilde{x}_1,\dots,\widetilde{x}_n \in \mathbb{F}(\mathbb{R})$.
Ensure: B bootstrap samples.
  for $j = 1$ to B **do**
    for $i = 1$ to n **do**
      Generate randomly a real number α_{ij} from the uniform distribution on the unit interval $[0,1]$.
      Generate randomly two real numbers: x'_{ij} from the uniform distribution on the α-cut $(\widetilde{x}_i)_{\alpha_{ij}}$, and x''_{ij} from the uniform distribution on the $1-\alpha$-cut $(\widetilde{x}_i)_{1-\alpha_{ij}}$.
      Let $x^*_{ij} = \frac{1}{2}(x'_{ij}+x''_{ij})$.
    end for
  end for
  Bootstrap samples $x^*_{1j},\dots,x^*_{nj}$, where $j = 1,\dots,B$.

The variance estimator for the RSS was obtained using function `varRSS` (with `Montip`) from the `RSSampling` library (Sevinc *et al.*, 2019). Therefore, to compare properly results obtained for the RSS with other methods instead of (4), the standard error was estimated from the formula

$$\mathrm{AcSE}(\hat{\theta}) = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{\theta}^*_i-\frac{1}{m}\sum_{j=1}^{m}\hat{\theta}^*_j\right)^2}, \quad (6)$$

where $\hat{\theta}^*_i$ denotes the bootstrap estimator obtained in the i-th experiment. Formula (6) was also applied for the jackknife method and the estimator W including a correction for the variance.

Let us discuss the results of the simulation study. We shall start from the comparison of point estimators based on the bootstrap methods, i.e., the simple epistemic bootstrap, the antithetic approach, and the RSS-driven method.

Although the estimates obtained with these three methods usually do not differ significantly, some differences can be noted for the dispersion estimators, especially for the RSS-driven approach. Here one should keep in mind that this method is more computationally demanding. Thus, it appears that the antithetic approach might be recommended especially if we have computational constraints. Moreover, W generally leads to lower values than S^2 when estimating the variance. Actually, it is not surprising since the observed dispersion stems both from randomness related to the population distribution and the imprecise perception. Hence this lowering effect of W looks promising.

More visible differences between estimators reveal their standard errors. Assuming SE as the quality criterion, the antithetic approach seems to be favored. Indeed, if we compare the simple bootstrap with its

Table 2. Comparison of the estimators for $\mathbb{F}_{(\mathrm{N},\mathrm{U},\mathrm{U},1)}$.

Type	simple	antithetic	RSS	jack+std	jack+anti
$\overline{X}$	0.0022	0.0019	0.0018	0.0014	**0.0013**
S	1.0307	**1.0069**	0.9871	1.0620	1.0381
S^2	1.1230	1.0719	**1.0196**	1.1221	1.0719
W	1.0246	**1.0240**	NA	NA	NA
$\mathrm{SE}(\overline{X})$	0.0959	**0.0669**	0.0686	0.0965	0.0670
$\mathrm{SE}(S)$	0.0984	**0.0698**	NA	0.1017	0.0719
$\mathrm{SE}(S^2)$	0.2035	0.1407	NA	0.2030	**0.1398**
$\mathrm{AcSE}(S^2)$	0.4888	0.4867	**0.4392**	0.4898	0.4869
$\mathrm{AcSE}(W)$	0.4886	0.4866	NA	NA	NA
$\mathrm{MSE}(\overline{X})$	0.0949	0.0950	**0.0948**	0.0967	0.0950
$\mathrm{MSE}(S)$	0.0505	0.0517	**0.0447**	0.0568	0.0573
$\mathrm{MSE}(S^2)$	0.2470	0.2355	**0.1911**	0.2435	0.2346
$\mathrm{MSE}(W)$	0.2324	**0.2303**	NA	NA	NA

Table 3. Comparison of the estimators for $\mathbb{F}_{(\mathrm{N},\mathrm{U},\mathrm{U},2)}$.

Type	simple	antithetic	RSS	jack+std	jack+anti
$\overline{X}$	-0.0017	-0.0032	-0.0020	-0.0017	-0.0032
S	2.0757	2.0301	**1.9879**	2.1415	2.0913
S^2	4.5395	4.3401	**4.1182**	4.5509	4.3334
W	**4.1346**	4.1432	NA	NA	NA
$\mathrm{SE}(\overline{X})$	0.1944	**0.1375**	0.1399	0.1956	0.1362
$\mathrm{SE}(S)$	0.1993	**0.1407**	NA	0.2102	0.1480
$\mathrm{SE}(S^2)$	0.8314	**0.5708**	NA	0.8458	0.5813
$\mathrm{AcSE}(S^2)$	1.9376	1.9294	**1.7239**	1.9390	1.9200
$\mathrm{AcSE}(W)$	1.9364	**1.9301**	NA	NA	NA
$\mathrm{MSE}(\overline{X})$	0.4087	0.4020	0.4028	0.4061	**0.4002**
$\mathrm{MSE}(S)$	0.2067	0.2067	**0.1763**	0.2358	0.2319
$\mathrm{MSE}(S^2)$	4.0499	3.7611	**3.0257**	4.0142	3.7923
$\mathrm{MSE}(W)$	3.7596	**3.6616**	NA	NA	NA

antithetic counterpart, then SE can be reduced even by about 20–30%. Lower values of $\mathrm{AcSE}(S^2)$ were obtained for the RSS approach. The differences obtained for $\mathrm{AcSE}(W)$ are not significant but the antithetic approach is still favored.

Quite similar conclusions result from the comparison of the mean squared errors. Differences in MSE between estimators of the mean are not significant, unlike the case of measures of dispersion, where the MSE can be reduced even by about 20–30%. Considering MSE as a quality criterion, the RSS-driven approach seems to be the best, but if we are interested in smaller computational costs, rather the antithetic approach is recommended.

If, instead of the bootstrap, the resampling methods related to the jackknife are used, the general conclusions are nearly unchanged. The jackknife combined with the simple bootstrap in some cases can slightly lower SE and MSE or produce estimates closer to original distribution parameters. However, this does not happen in general. Better results are achieved if the jackknife is used together with the antithetic approach, similar to the case of the above-mentioned bootstrap methods.

To sum up, taking into account various statistical properties as well as the computational costs, the antithetic approach (combined with the bootstrap or jackknife method) should be recommended.

We also numerically investigated the similar examples with bigger initial samples (i.e., $n = 100$). In general, the obtained differences between estimators were lower but the conclusions remained unchanged.

In our numerical study we also compared the quality of the estimators considered with the results provided by the fuzzy version of the classical EM algorithm, known as FEM (Denœux, 2011)) and available in `EM.Fuzzy` package (Parchami, 2018) implemented in R. In this study, only small ($n = 10$) samples were considered because

Table 4. Comparison of the estimators for $\mathbb{F}_{(\mathrm{N,E,U})}$.

Type	simple	antithetic	RSS	jack+std	jack+anti
$\overline{X}$	-0.0074	-0.0120	-0.0114	-0.0133	-0.0095
S	1.4164	**1.3087**	1.3889	1.4694	1.3493
S^2	2.1746	**1.8514**	2.0144	2.1697	1.8253
W	1.5125	**1.5220**	NA	NA	NA
$\mathrm{SE}(\overline{X})$	0.2467	0.1730	**0.1502**	0.2441	0.1737
$\mathrm{SE}(S)$	0.2639	**0.1982**	NA	0.2857	0.2118
$\mathrm{SE}(S^2)$	0.7897	0.5514	NA	0.7948	**0.5423**
$\mathrm{AcSE}(S^2)$	0.9868	0.9148	0.8762	0.9770	**0.8744**
$\mathrm{AcSE}(W)$	0.8583	**0.8516**	NA	NA	NA
$\mathrm{MSE}(\overline{X})$	0.1514	0.1494	**0.1442**	0.1497	0.1446
$\mathrm{MSE}(S)$	0.2582	**0.1827**	0.2389	0.3222	0.2201
$\mathrm{MSE}(S^2)$	2.1899	1.4426	1.8220	2.2758	**1.4007**
$\mathrm{MSE}(W)$	**0.8831**	0.9071	NA	NA	NA

Table 5. Comparison of the estimators for $\mathbb{F}_{(\beta,\mathrm{Ucond},1)}$.

Type	simple	antithetic	RSS	jack+std	jack+anti
$\overline{X}$	0.4976	**0.4977**	**0.4977**	**0.4977**	0.4974
S	0.2227	0.2185	0.2133	0.2282	**0.2231**
S^2	0.0515	**0.0495**	0.0469	0.0517	0.0494
W	0.0475	**0.0476**	NA	NA	NA
$\mathrm{SE}(\overline{X})$	0.0193	0.0133	0.0138	0.0191	**0.0131**
$\mathrm{SE}(S)$	0.0183	**0.0130**	NA	0.0187	0.0131
$\mathrm{SE}(S^2)$	0.0081	**0.0056**	NA	0.0081	**0.0056**
$\mathrm{AcSE}(S^2)$	0.0179	0.0178	**0.0160**	0.0178	0.0179
$\mathrm{AcSE}(W)$	0.0180	**0.0179**	NA	NA	NA
$\mathrm{MSE}(\overline{X})$	0.0048	0.0048	0.0048	0.0048	0.0048
$\mathrm{MSE}(S)$	0.0015	0.0016	**0.0014**	0.0015	0.0015
$\mathrm{MSE}(S^2)$	0.0003	0.0003	**0.0002**	0.0003	0.0003
$\mathrm{MSE}(W)$	0.0003	0.0003	NA	NA	NA

`EM.Fuzzy` package is intended for such samples. As in previous experiments we set $B = 10$ and $m = 1000$. The default accuracy $\varepsilon = 0.001$ of the FEM algorithm was applied (Parchami, 2018). In general, our resampling methods provided better results than the FEM. Sometimes the FEM produced strange outputs, e.g., the MSE of the variance estimator in $\mathbb{F}_{(\mathrm{N,U,U,2})}$ was more than 10 times greater than the MSE of estimators based on our resampling methods. Moreover, since the FEM algorithm is a parametric method, it requires knowledge of the population distribution which is not that common in the case of fuzzy data. In this context methods suggested in this paper have an important advantage of their distribution-free nature.

6. Epistemic bootstrap in hypothesis testing

The bootstrap methods considered in this contribution were also applied in hypotheses testing based on epistemic fuzzy data.

Suppose $X_1, \ldots, X_n$ are i.i.d. random variables from the normal distribution $\mathrm{N}(\mu, \sigma)$ with the unknown mean μ and known standard deviation σ. Let us consider the null hypothesis $H_0 : \mu \leq \mu_0$ against the alternative $H_1 : \mu > \mu_0$ on the significance level 0.05. To solve the given problem, we used various resampling methods (including the simple epistemic bootstrap and its modification, i.e., the antithetic approach, as well as the usual jackknife and jackknife combined with the antithetic approach) followed by the classical one-sided Z-test for the mean. This method was compared with two well-known approaches: the fuzzy test based on fuzzy confidence intervals (Grzegorzewski, 2000) and the test utilizing the inner and outer approximations of the confidence interval (Couso and Sánchez, 2011). Finally, the power analysis of all resampling methods considered is given.

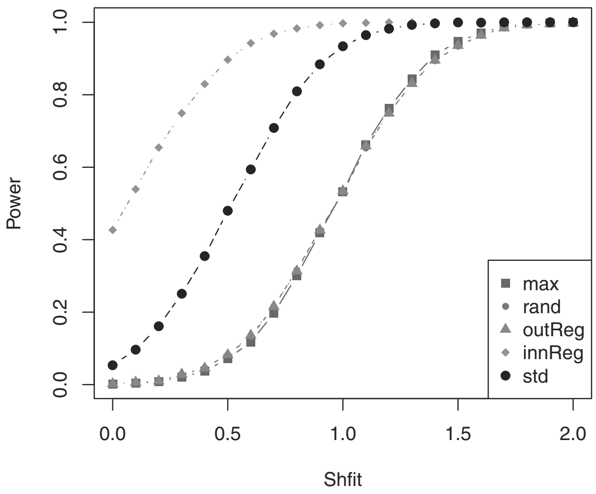

Fig. 4. Simulated power curves for the small sample size $n = 10$ and fuzzy data from $\mathbb{F}_{(\mathrm{N,U,U,1})}$.

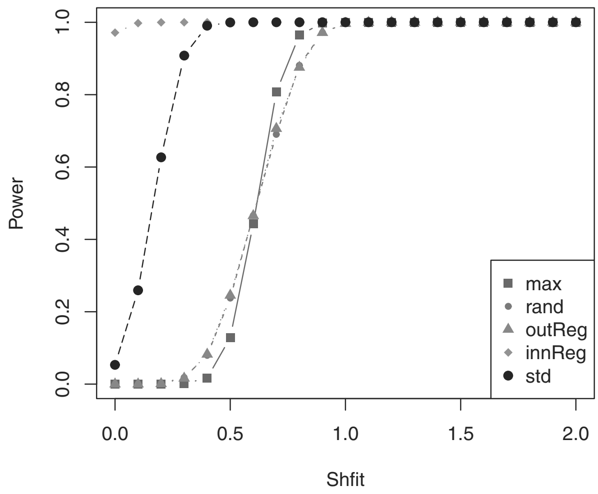

Fig. 5. Power curves for the moderate sample size $n = 100$ and fuzzy data from $\mathbb{F}_{(\mathrm{N,U,U,1})}$.

The test proposed by Grzegorzewski (2000) does not produce a binary answer (i.e., to reject or accept H_0, as we are used to) but it provides a degree of rejection and acceptance of the null hypothesis under study. A method to communicate and interpret results of a fuzzy test in a user-friendly manner is proposed by Grzegorzewski and Hryniewicz (2002). If the crisp decision is required, one has to combine this fuzzy test with an appropriate defuzzification method, like those indicated by Grzegorzewski (2001). In our simulation study two defuzzification operators were considered: the maximum value, denoted further on as "max", and the randomized operator, denoted as "rand."

The testing procedure based on the double intervals (Couso and Sánchez, 2011) gives the possibility of acceptance of the null hypothesis (which is related to the outer region) and the possibility of its rejection (based on the inner region), which results in two power curves (denoted further as "outReg" and "innReg," respectively). It is worth noting that the applied defuzzification rule mimics the "rand" operator (one can imagine this as an asymmetric coin toss to receive a decision when the probability of rejecting H_0 is equal to its possibility).

Some results of the extended simulation study are given in Table 6 and Fig. 4 (for the small $n = 10$ sample size) and in Fig. 5 (for the moderate $n = 100$ sample size). Each numerical experiment was repeated 10000 times.

Firstly, we noticed that the power curves obtained for the applied resampling methods did not differ significantly. One can see in Table 6 that the estimated sizes of the corresponding four tests are close to the set significance level (i.e., 0.05), as well as the power for some values of the shift in mean are very similar. Therefore, to simplify the plots, only the results for the simple bootstrap approach are presented in Figs. 4 and 5 (depicted in black line and circles).

Secondly, it can be seen immediately that the power of the test based on the simple epistemic bootstrap dominates significantly over power curves corresponding to the Grzegorzewski test (depicted by rectangles for the "max" operator and by bullets for the "rand" rule) and the test related to the outer approximation of the confidence interval (depicted by triangles point-up). One might think that the test based on the inner approximation of the confidence interval (depicted by diamonds) is the winner due to the fact that its graph is located well above the others. However, both Table 6 and Figs. 4–5 show that this effect is ostensible since that this test does not hold the assumed significance level (it is bigger than 0.4 for $n = 10$ and 0.97 for $n = 100$ whereas the assumed significance level is equal to 0.05). Therefore, the test based on the inner approximation of the confidence interval is simply so conservative that it makes it practically useless.

To sum up, numerical experiments show that the proposed bootstrap for epistemic fuzzy data followed by some classical tests provides better results than fuzzy tests.

7. Final remarks, conclusions and further research

Fuzzy data appear in many real-life situations to model imprecision, vagueness, and some other data weakness which makes standard modeling difficult or even impossible. When fuzzy numbers are used for modeling imprecise perceptions of the real-valued

Table 6. Comparison of power values for $n = 10$ and fuzzy data from $\mathbb{F}_{(\mathrm{N,U,U,1})}$.

Shift	0	0.1	0.2	0.3	0.4	0.5
simple	0.0537	0.0962	0.1610	0.2507	0.3549	0.4800
antithetic	0.0555	0.0978	0.1612	0.2522	0.3585	0.4810
jackknife+std	0.0546	0.0965	0.1592	0.2490	0.3576	0.4801
jackknife+anti	0.0542	0.0954	0.1592	0.2503	0.3568	0.4785
max	0.0014	0.0041	0.0082	0.0200	0.0377	0.0711
rand	0.0023	0.0050	0.0109	0.0264	0.0466	0.0828
outReg	0.0022	0.0065	0.0101	0.0290	0.0454	0.0825
innReg	0.4268	0.5393	0.6543	0.7491	0.8295	0.8967

observations, we deal with the so-called epistemic fuzzy data. Since such data often cause problems in statistical reasoning, a new bootstrap technique was suggested.

Simulation results show that the proposed bootstrap method may be useful in different fields of statistical inference. Moreover, besides the simple epistemic bootstrap, we have developed several of its refinements aimed at improving certain properties, like the variance reduction. In addition to theoretical investigations and discussion of the results, comments and tips are provided that should prove useful to practitioners.

The main goal of this contribution is to suggests a general methodology of the epistemic bootstrap and to indicate some of its applications in statistical reasoning with fuzzy data. Obviously, this contribution does not end discussion on the epistemic bootstrap. In particular, it would be interesting to find new application areas for the introduced methodology and to deal with the problems that may arise there.

In further considerations, one may ask for theoretical justifications of the proposed approach, like conditions to be met to ensure the consistency of the particular bootstrap procedures. When considering statistical consistency in the fuzzy environment, one should realize that this problem does not only concern the proposed methodology but is primarily dependent on the data structure. Here we mean both the shape of membership functions describing fuzzy data and their relative location. To make it easier to imagine, assume that our fuzzy sample consists of rectangular fuzzy numbers that are isomorphic with interval data. Then both computability and statistical properties of the procedures considered depend on whether available intervals are wide (“puffy”) or narrow (“skinny”) as well as if these intervals are disjoint or they have a few intersections only or have arbitrarily many intersections. It is also important if they are of the same precision or not, if they are nesting or not, etc. For a list of possible cases related to imprecise data and their taxonomy, see the works of Ferson *et al*. (2007) or Nguyen *et al*. (2012).

Returning to consistency, consider, e.g., a sample from the beta distribution and suppose that all imprecise observations are modeled as identical unit intervals. Then the empirical distribution obtained with the epistemic bootstrap, obviously, cannot be a consistent estimator of the actual distribution function. On the other hand, if the sample consists of intervals having no intersections, then the epistemic bootstrap provides the consistent estimator. Naturally, both situations are in some sense extreme and hence the conclusions are trivial, whereas in practical situations a more subtle analysis is necessary.

When moving from interval data to non-rectangular fuzzy numbers, we may distinguish much more possible situations, e.g., for some α-cuts we may have no intersections at all while they may appear for other α-cuts. Therefore, it seems that a comprehensive analysis of all possible fuzzy data structures in the context of resolving the issue would require a separate article, if such a resolution was possible at all. Even if we restrict our considerations to main data structures, this would require a separate discussion for each type of inference. Here, some general techniques for proving the consistency of the bootstrap given by Shao and Tu (1995) might be helpful, as well as the results developed for generalized random elements by Giné and Zinn (1990) or Gil *et al*. (2006).

Anyway, wherever we use bootstrap, we should always remember that this technique is not a recipe for all possible problems. Indeed, “bootstrap methods are intended to help avoid tedious calculations based on questionable assumptions, and this they do. But they cannot replace clear critical thought about the problem, appropriate design of the investigation and data analysis, and incisive presentation of conclusions” (Davison and Hinkley, 1997).

References

Ban, A., Coroianu, L. and Grzegorzewski, P. (2015). *Fuzzy Numbers: Approximations, Ranking and Applications*, Polish Academy of Sciences, Warsaw.

Casella, G. (2003). Introduction to the silver anniversary of the bootstrap, *Statistical Science* **18**(2): 133–134.

Colubi, A., Fernández-García, C. and Gil, M. (2002). Simulation of random fuzzy variables: An empirical approach to statistical/probabilistic studies with fuzzy experimental data, *IEEE Transactions on Fuzzy Systems* **10**(3): 384–390.

Couso, I. and Dubois, D. (2014). Statistical reasoning with set-valued information: Ontic vs. epistemic views, *International Journal of Approximate Reasoning* **55**(7): 1502–1518.

Couso, I. and Sánchez, L. (2011). Inner and outer fuzzy approximations of confidence intervals, *Fuzzy Sets and Systems* **184**(1): 68–83.

Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*, Cambridge University Press, Cambridge.

De Angelis, D. and Young, G.A. (1992). Smoothing the bootstrap, *International Statistical Review* **60**(1): 45–56.

Denœux, T. (2011). Maximum likelihood estimation from fuzzy data using the EM algorithm, *Fuzzy Sets and Systems* **183**(1): 72–91.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *Annals of Statistics* **7**(1): 1–26.

Ferson, S., Kreinovich, V., Hajagos, J., Oberkampf, W. and Ginzburg, L. (2007). Experimental uncertainty estimation and statistics for data having interval uncertainty, *Technical Report SAND2007-0939*, Applied Biomathematics, New York.

Gil, M., Montenegro, M., González-Rodríguez, G., Colubi, A. and Casals, M. (2006). Bootstrap approach to the multi-sample test of means with imprecise data, *Computational Statistics and Data Analysis* **51**(1): 148–162.

Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures, *Annals of Probability* **18**(2): 851–869.

González-Rodríguez, G., Montenegro, M., Colubi, A. and Gil, M. (2006). Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data, *Fuzzy Sets and Systems* **157**(19): 2608–2613.

Grzegorzewski, P. (2000). Testing statistical hypotheses with vague data, *Fuzzy Sets and Systems* **112**(3): 501–510.

Grzegorzewski, P. (2001). Fuzzy tests—defuzzification and randomization, *Fuzzy Sets and Systems* **118**(3): 437–446.

Grzegorzewski, P. and Goławska, J. (2021). In search of a precise estimator based on imprecise data, *Joint Proceedings of the IFSA-EUSFLAT-AGOP 2021 Conferences, Bratislava, Slovakia*, pp. 530–537.

Grzegorzewski, P. and Hryniewicz, O. (2002). Computing with words and life data, *International Journal of Applied Mathematics and Computer Science* **12**(3): 337–345.

Grzegorzewski, P., Hryniewicz, O. and Romaniuk, M. (2019). Flexible bootstrap based on the canonical representation of fuzzy numbers, *Proceedings of EUSFLAT 2019, Prague, Czech Republic*, pp. 490–497.

Grzegorzewski, P., Hryniewicz, O. and Romaniuk, M. (2020a). Flexible bootstrap for fuzzy data based on the canonical representation, *International Journal of Computational Intelligence Systems* **13**(1): 1650–1662.

Grzegorzewski, P., Hryniewicz, O. and Romaniuk, M. (2020b). Flexible resampling for fuzzy data, *International Journal of Applied Mathematics and Computer Science* **30**(2): 281–297, DOI: 10.34768/amcs-2020-0022.

Grzegorzewski, P. and Romaniuk, M. (2021). Epistemic bootstrap for fuzzy data, *Joint Proceedings of the IFSA-EUSFLAT-AGOP 2021 Conferences, Bratislava, Slovakia*, pp. 538–545.

Hall, P., DiCiccio, T. and Romano, J. (1989). On smoothing and the bootstrap, *Annals of Statistics* **17**(2): 692–704.

Hukuhara, M. (1967). Integration des applications measurables dont la valeur est un compact convexe, *Funkcialaj Ekvacioj* **10**: 205–223.

Kołacz, A. and Grzegorzewski, P. (2019). Asymptotic algorithm for computing the sample variance of interval data, *Iranian Journal of Fuzzy Systems* **16**(4): 83–96.

Kroese, D.P., Taimre, T. and Botev, Z.I. (2011). *Handbook of Monte Carlo Methods*, Wiley, Hoboken.

Kruse, R. (1982). The strong law of large numbers for fuzzy random variables, *Information Sciences* **28**(3): 233–241.

Kwakernaak, H. (1978). Fuzzy random variables, Part I: Definitions and theorems, *Information Sciences* **15**(1): 1–15.

Lubiano, M.A., Montenegro, M., Sinova, B., de la Rosa de Sáa, S. and Gil, M.A. (2016). Hypothesis testing for means in connection with fuzzy rating scale-based data: Algorithms and applications, *European Journal of Operational Research* **251**(3): 918–929.

Lubiano, M.A., Salas, A., Carleos, C., de la Rosa de Sáa, S. and Gil, M.A. (2017). Hypothesis testing-based comparative analysis between rating scales for intrinsically imprecise data, *International Journal of Approximate Reasoning* **88**: 128–147.

Montenegro, M., Colubi, A., Casals, M. and Gil, M. (2004). Asymptotic and bootstrap techniques for testing the expected value of a fuzzy random variable, *Metrika* **59**: 31–49.

Nguyen, H., Kreinovich, V., Wu, B. and Xiang, G. (2012). *Computing Statistics under Interval and Fuzzy Uncertainty*, Springer, Berlin/Heidelberg.

Parchami, A. (2018). EM algorithm for maximum likelihood estimation by non-precise information, `https://cran.r-project.org/package=EM.Fuzzy`.

Pedrycz, W. (1994). Why triangular membership functions?, *Fuzzy Sets and Systems* **64**(1): 21–30.

Piegat, A. (2005). A new definition of the fuzzy set, *International Journal of Applied Mathematics and Computer Science* **15**(1): 125–140.

Ramos-Guajardo, A., Blanco-Fernández, A. and González-Rodríguez, G. (2019). Applying statistical methods with imprecise data to quality control in cheese manufacturing, *in* P. Grzegorzewski *et al.* (Eds), *Soft Modeling in Industrial Manufacturing*, Springer, Berlin/Heidelberg, pp. 127–147.

Ramos-Guajardo, A. and Grzegorzewski, P. (2016). Distance-based linear discriminant analysis for interval-valued data, *Information Sciences* **372**: 591–607.

Ramos-Guajardo, A. and Lubiano, M. (2012). k-Sample tests for equality of variances of random fuzzy sets, *Computational Statistics and Data Analysis* **56**(4): 956–966.

Romaniuk, M. (2019). On some applications of simulations in estimation of maintenance costs and in statistical tests for fuzzy settings, *in* A. Steland *et al.* (Eds), *Stochastic Models, Statistics and Their Applications*, Springer, Cham, pp. 437–448.

Romaniuk, M. and Hryniewicz, O. (2019). Interval-based, nonparametric approach for resampling of fuzzy numbers, *Soft Computing* **23**: 5883–5903.

Romaniuk, M. and Hryniewicz, O. (2021). Discrete and smoothed resampling methods for interval-valued fuzzy numbers, *IEEE Transactions on Fuzzy Systems* **29**(3): 599–611.

Sevinc, B., Cetintav, B., Esemen, M. and Gurler, S. (2019). RSSampling: A pioneering package for ranked set sampling, *The R Journal* **11**(1): 401–415.

Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*, Springer, New York.

Silverman, B.W. and Young, G.A. (1987). The bootstrap: To smooth or not to smooth?, *Biometrika* **74**(3): 469–479.

Suresh, H. and Guttag, J.V. (2021). A framework for understanding sources of harm throughout the machine learning life cycle, *Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21), New York, USA*.

Vavasis, S.A. (1991). *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York.

Wang, D. and Hryniewicz, O. (2015). A fuzzy nonparametric Shewhart chart based on the bootstrap approach, *International Journal of Applied Mathematics and Computer Science* **25**(2): 389–401, DOI: 10.1515/amcs-2015-0030.

Wolfe, D.A. (2004). Ranked set sampling: An approach to more efficient data collection, *Statistical Science* **19**(4): 636–643.

Zadeh, L.A. (1973). Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Transactions on Systems, Man and Cybernetics* **SMC-3**(1): 28–44.

Przemysław Grzegorzewski received his MSc in mathematics from the University of Warsaw. Then he obtained his PhD with distinction and DSc (habilitation) in computer science from the Systems Research Institute of the Polish Academy of Sciences. In 2018 he received the state professorial title. He is currently a full professor at the Faculty of Mathematics and Information Science of the Warsaw University of Technology as well as the Systems Research Institute of the Polish Academy of Sciences. His areas of expertise include mathematical statistics, statistical decisions with imprecise data, data mining, fuzzy sets, fuzzy logic, soft computing, etc.

Maciej Romaniuk received his MSc degree in mathematics from the Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw, Poland in 2001. He earned his PhD and DSc (habilitation) degrees in computer science from the Systems Research Institute, Polish Academy of Sciences, in 2007 and 2018, respectively. The main topics of his current research are Monte Carlo simulations, simulations of events under circumstances of uncertain and imprecise information, financial mathematics, actuarial mathematics, statistics, and fuzzy numbers.

Received: 18 October 2021
Revised: 31 January 2022
Re-revised: 21 March 2022
Accepted: 20 April 2022