Int. J. Appl. Math. Comput. Sci., 2020, Vol. 30, No. 2, 281–297
DOI: 10.34768/amcs-2020-0022

FLEXIBLE RESAMPLING FOR FUZZY DATA

PRZEMYSLAW GRZEGORZEWSKI [a,b,*], OLGIERD HRYNIEWICZ [a], MACIEJ ROMANIUK [a,c]

[a] Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
e-mail: pgrzeg@ibspan.waw.pl

[b] Faculty of Mathematics and Information Science
Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland

[c] Warsaw School of Information Technology
Newelska 6, 01-447 Warsaw, Poland

In this paper, a new methodology for simulating bootstrap samples of fuzzy numbers is proposed. Unlike the classical bootstrap, it allows enriching a resampling scheme with values from outside the initial sample. Although a secondary sample may contain results beyond members of the primary set, they are generated smartly so that the crucial characteristics of the original observations remain invariant. Two methods for generating bootstrap samples preserving the representation (i.e., the value and the ambiguity or the expected value and the width) of fuzzy numbers belonging to the primary sample are suggested and numerically examined with respect to other approaches and various statistical properties.

Keywords: bootstrap, fuzzy data, fuzzy numbers, fuzzy sample, imprecise data, resampling.

1. Introduction

Forty years ago, Bradley Efron published his seminal paper "Bootstrap methods: Another look at the jackknife" (Efron, 1979). The bootstrap is typically used to find standard errors of estimators, confidence intervals for unknown parameters or p-values for statistical tests. However, the ideas suggested by Efron turned out so important in modern statistics that George Casella on the silver anniversary of the bootstrap concluded: "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" (Casella, 2003).

The bootstrap usually works out in complicated models. This is also the case of imprecise data often modeled with fuzzy random variables. Since there are not yet suitable models for the distribution of fuzzy random variables, nor central limit theorems for fuzzy random variables that can be straightforwardly applied, the bootstrap turns out to be an invaluable help in statistical reasoning with fuzzy data. In particular, it was widely used in statistical tests with fuzzy data (Colubi *et al.*, 2002; Gil *et al.*, 2006; González-Rodríguez *et al.*, 2006; Ramos-Guajardo and Lubiano, 2012; Montenegro *et al.*, 2004), classification (Ramos-Guajardo and Grzegorzewski, 2016), fuzzy rating in questionnaires (Lubiano *et al.*, 2016; 2017), quality control in cheese manufacturing (Ramos-Guajardo *et al.*, 2019), fuzzy Shewhart control charts (Wang and Hryniewicz, 2015), etc.

The classical bootstrap involves drawing random samples with replacement from the initial sample of observations. Consequently, nearly every bootstrap sample contains repeated values. What is worse, if the original sample size is small, all bootstrap samples consist of only few distinct values, which gives a strongly unwanted effect especially if the unknown original distribution is continuous. To overcome this inconvenience, various improvements of the classical bootstrap were proposed, like the balanced bootstrap

*Corresponding author

(Davison *et al.*, 1986; Graham *et al.*, 1990) or the so-called smoothed bootstrap (Silverman and Young, 1987; Hall *et al.*, 1989; De Angelis and Young, 1992). In a fuzzy context various resampling methods based on incremental increases of α-cuts were given by Romaniuk and Hryniewicz (2019a; 2019b).

In this paper we propose another modification of the classical bootstrap to increase the diversity of simulated fuzzy outcomes. Its key idea is to generate fuzzy numbers which may differ from the original ones but preserve some critical characteristics (e.g., the value and ambiguity or the expected value and width) of fuzzy numbers forming the primary sample. This contribution contains a substantial extension of the method introduced by Grzegorzewski *et al.* (2019). We not only provide new bootstrap algorithms, but present also a broad study of statistical properties of the suggested procedures.

The paper is organized as follows. Basic definitions and general notation are provided in Section 2. The so-called flexible bootstrap algorithm is introduced in Section 3 and its relation to the classical bootstrap is considered. New resampling methods for triangular and for trapezoidal fuzzy numbers are thoroughly developed in Sections 4 and 5, respectively. Then, in Section 6, the proposed algorithms are numerically examined and compared with other existing approaches.

2. Fuzzy data

A **fuzzy number** A is a fuzzy set in $\mathbb{R}$ which is normal, fuzzy-convex, has upper semicontinuous membership function $A(x)$ and bounded support. An α-cut of a fuzzy number A, where $\alpha \in [0,1]$, is defined by

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

where cl stands for the closure operator. It is easily seen that the α-cut $A(\alpha)$ of a fuzzy number A is a closed interval $A(\alpha) = [A_L(\alpha), A_U(\alpha)]$.

The most often used fuzzy numbers are **trapezoidal fuzzy numbers** (sometimes called *fuzzy intervals*) with membership functions of the form

$$A(x) = \begin{cases} \dfrac{x - a_1}{a_2 - a_1} & \text{if } a_1 < x \leq a_2, \\ 1 & \text{if } a_2 \leq x \leq a_3, \\ \dfrac{a_4 - x}{a_4 - a_3} & \text{if } a_3 \leq x < a_4, \\ 0 & \text{otherwise}, \end{cases} \tag{1}$$

where $a_1, a_2, a_3, a_4 \in \mathbb{R}$ such that $a_1 \leq a_2 \leq a_3 \leq a_4$. A trapezoidal fuzzy number A will be further denoted as $[a_1, a_2, a_3, a_4]$. If $a_2 = a_3$, then A is said to be a **triangular fuzzy number** and we have $A = [a_1, a_2, a_4]$. The families of all fuzzy numbers, trapezoidal fuzzy numbers and triangular fuzzy number will be denoted by $\mathbb{F}(\mathbb{R})$, $\mathbb{F}^T(\mathbb{R})$ and $\mathbb{F}^\Delta(\mathbb{R})$, respectively. Obviously, $\mathbb{F}^\Delta(\mathbb{R}) \subset \mathbb{F}^T(\mathbb{R}) \subset \mathbb{F}(\mathbb{R})$.

Often, instead of declaring two points a_1 and a_4 describing the support of A and next two points a_2 and a_3 for its core, it is more convenient to use another parametrization through its location and the spread of its arm. Namely, let us define the following parameters:

$$c := \frac{a_2 + a_3}{2}, \qquad s := \frac{a_3 - a_2}{2},$$
$$l := a_2 - a_1, \qquad r := a_4 - a_3.$$

One can easily identify c and s as the center and the half of the core, respectively, while l and r stand for the spread of the left and the right arm of the membership function $A(x)$, respectively. Obviously, $c \in \mathbb{R}$, while $s, l, r \geq 0$. Using this notation, a trapezoidal fuzzy number A would be denoted as $A(c, s; l, r)$. Similarly, $A(c; l, r)$ stands for a triangular fuzzy number (since then $s = 0$).

To simplify the representation of fuzzy numbers, Delgado *et al.* (1998) suggested two parameters: value and ambiguity, which represent some basic features of fuzzy numbers and hence the called the *canonical representation of fuzzy numbers*.

A location of a fuzzy number A is characterized by its **value** defined as

$$\mathrm{Val}(A) = \int_0^1 \alpha(A_U(\alpha) + A_L(\alpha))\, \mathrm{d}\alpha, \tag{2}$$

whereas the **ambiguity** of A, given by

$$\mathrm{Amb}(A) = \int_0^1 \alpha(A_U(\alpha) - A_L(\alpha))\, \mathrm{d}\alpha, \tag{3}$$

is a measure of the global spread (or vagueness) of a fuzzy number A.

Since the value and ambiguity represent basic features of a fuzzy number, two fuzzy numbers with the same ambiguity and value might be considered similar (sometimes they are even treated as “almost equal”(see Delgado *et al.*, 1998)). One can easily find that the value and ambiguity of a trapezoidal fuzzy number $A(c, s; l, r)$ are given as follows:

$$\mathrm{Val}(A) = c + \frac{r - l}{6}, \tag{4}$$

$$\mathrm{Amb}(A) = s + \frac{r + l}{6}. \tag{5}$$

If $A(c; l, r)$ is a triangular fuzzy number, then its value is still given by (4), while its ambiguity reduces to

$$\mathrm{Amb}(A) = \frac{r + l}{6}. \tag{6}$$

Another important characteristic of a fuzzy number is its **expected interval** (Dubois and Prade, 1987; Heilpern, 1992), defined as

$$\mathrm{EI}(A) = \left[\int_0^1 A_L(\alpha)\,\mathrm{d}\alpha, \int_0^1 A_U(\alpha)\,\mathrm{d}\alpha \right]. \tag{7}$$

The expected interval of a fuzzy number has many interesting properties and is very useful in many situations, like defuzzification or approximation of fuzzy numbers (see, e.g., Ban *et al.*, 2015)).

The middle point of the expected interval is called the **expected value** of the fuzzy number and is defined by

$$\mathrm{EV}(A) = \frac{1}{2} \left[\int_0^1 A_L(\alpha)\,\mathrm{d}\alpha + \int_0^1 A_U(\alpha)\,\mathrm{d}\alpha \right]. \tag{8}$$

The expected value of a fuzzy number is a characteristic of its location, i.e., it shows a real value which is (in some sense) typical for the fuzzy notion modeled by a fuzzy number under discussion. Thus the expected value of a fuzzy number is a counterpart of the value (2). We have also a counterpart of the ambiguity, called the **width** of a fuzzy number (Chanas, 2001), defined by

$$\mathrm{w}(A) = \int_0^1 (A_U(\alpha) - A_L(\alpha))\,\mathrm{d}\alpha. \tag{9}$$

For the trapezoidal fuzzy number $A(c, s; l, r)$ we obtain

$$\mathrm{EV}(A) = c + \frac{r-l}{4}, \tag{10}$$

$$\mathrm{w}(A) = s + \frac{r+l}{4}. \tag{11}$$

If $A(c; l, r)$ is a triangular fuzzy number, then its expected value remains as in (10), while its width reduces to

$$\mathrm{w}(A) = \frac{r+l}{4}. \tag{12}$$

For more details on fuzzy numbers, their characteristics and approximations we refer the reader to the work of Ban *et al.* (2015), and for some examples of their applications to, e.g., those of Gao *et al.* (2013) or Grzegorzewski and Hryniewicz (2002).

3. Flexible resampling

The key idea of the classical **bootstrap** is to construct new samples drawing n times with replacement from the original dataset $x_1, \ldots, x_n \in \mathbb{R}$. This way, one can produce any number (say b) of bootstrap samples, as shown in Fig. 1, where $x^*_{ij} \in \{x_1, \ldots, x_n\}$ denotes the j-th element of the i-th sample.

The bootstrap has a serious disadvantage: it produces only values that belong to the input (primary) sample. Consequently, nearly every bootstrap sample contains repeated values. Furthermore, if the original sample is small, all bootstrap samples consist of only few distinct values, which might be strongly inadvisable, especially if the original distribution is continuous.

Actually, the heart of the problem is the oversimplified nature of the real-valued data. Any element x_i of the primary sample might be characterized only by its real value (this is why we call it one-dimensional). Consequently, any attempt to enrich resampling results beyond members of the primary sample is inextricably linked with changing its elements. Hence we have to accept that a secondary sample would consist of values x^*_i which do not necessarily appear in the original one. Of course, we should generate those new elements smartly to preserve some global properties of the whole sample.

In the case of fuzzy data, the situation seems to be more conductive. Each fuzzy number $\widetilde{x}_i \in \mathbb{F}(\mathbb{R})$ has a much more complicated structure than a real one, unless it is a singletone. Therefore, it seems that we may enrich resampling by generating new values $\widetilde{x}^*_i$ which preserve some crucial properties of $\widetilde{x}_i$ but quite some other minor ones. This way, resampling may produce a new sample $(\widetilde{x}^*_1, \ldots, \widetilde{x}^*_n)$ of elements which may differ from the original one $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$, but which preserve both some local and global properties of the primary sample elements.

The distinction between more or less important characteristics of a fuzzy number is, of course, questionable. It might depend on the subjective preferences of the analyst or more objective reasons connected with a particular situation. But in a common feeling, parameters that characterize the location and the spread (vagueness) are considered the most important properties of fuzzy numbers, contrary to minor details in shape of their membership functions. Hence, such characteristics like the value and the ambiguity (or the expected value and the width), mentioned in Section 2, may be of interest. In this paper, we propose a modified bootstrap based on this idea.

To clarify the idea, let $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}_{in}(\mathbb{R}) \subseteq \mathbb{F}(\mathbb{R})$ denote the original fuzzy sample. We assume that observations are fuzzy numbers of some type, i.e., they belong to a given subfamily $\mathbb{F}_{in}(\mathbb{R})$ of all fuzzy numbers (or, possibly, they are arbitrary fuzzy numbers). In the case of the improved flexible bootstrap, its scheme looks like the classical one, shown in Fig. 1. However, now

$$\begin{array}{ccc}
 & \nearrow & (x^*_{11}, x^*_{12}, \ldots, x^*_{1n}) \\
 & \nearrow & (x^*_{21}, x^*_{22}, \ldots, x^*_{2n}) \\
(x_1, x_2, \ldots, x_n) & \longrightarrow & (x^*_{31}, x^*_{32}, \ldots, x^*_{3n}) \\
 & & \vdots \quad \vdots \quad \vdots \qquad \vdots \\
 & \searrow & (x^*_{b1}, x^*_{b2}, \ldots, x^*_{bn})
\end{array}$$

Fig. 1. Classical bootstrap scheme.

we substitute a very restrictive assumption that $\widetilde{x}^*_{ij} \in \{\widetilde{x}_1, \ldots, \widetilde{x}_n\}$ with the weaker one, which states that each $\widetilde{x}^*_{ij}$ is generated in such a way that it returns a fuzzy number having the identical location and spread as the one randomly selected from the original dataset $\{\widetilde{x}_1, \ldots, \widetilde{x}_n\}$.

To be more specific, suppose that one decides to preserve the value and the ambiguity. Moreover, let $\mathbb{F}_{out}(\mathbb{R}) \subset \mathbb{F}(\mathbb{R})$ denote a chosen subfamily of fuzzy numbers which is not necessarily equivalent to $\mathbb{F}_{in}(\mathbb{R})$. Now we select randomly an observation $\widetilde{x}_i$ from the original dataset $\{\widetilde{x}_1, \ldots, \widetilde{x}_n\}$ and compute its value and ambiguity. Then a new fuzzy number $\widetilde{x}^*_{ij} \in \mathbb{F}_{out}(\mathbb{R})$ is generated in such a way that its value and ambiguity remain the same as for $\widetilde{x}_i$, i.e., $\mathrm{Val}(\widetilde{x}^*_{ij}) = \mathrm{Val}(\widetilde{x}_i)$ and $\mathrm{Amb}(\widetilde{x}^*_{ij}) = \mathrm{Amb}(\widetilde{x}_i)$. Alternatively, one can generate a new fuzzy number $\widetilde{x}^*_{ij} \in \mathbb{F}_{out}(\mathbb{R})$ sothat it preserves the mean value and the width of the original observation $\widetilde{x}_i$, i.e., $\mathrm{EV}(\widetilde{x}^*_{ij}) = \mathrm{EV}(\widetilde{x}_i)$ and $\mathrm{w}(\widetilde{x}^*_{ij}) = \mathrm{w}(\widetilde{x}_i)$.

A key problem that arises here is how to choose a suitable subfamily of fuzzy numbers $\mathbb{F}_{out}(\mathbb{R})$. It seems reasonable to restrict our attention to triangular or trapezoidal fuzzy numbers only. Although one may ask why, the reason is straightforward. It has been noticed by many researchers that trapezoidal or triangular fuzzy numbers are most common in current applications mainly because they are both easy to handle and have a natural interpretation (Ban *et al.*, 2015; Pedrycz, 1994), since "the problems that arise with vague predicates are less concerned with precision and are more of a qualitative type; thus they are generally written as linearly as possible. Normally it is sufficient to use a trapezoidal representation, as it makes it possible to define them with no more than four parameters" (Jimenez and Rivas, 1998). Moreover, even if the original data set consists of fuzzy numbers which are neither triangular nor trapezoidal, one often approximates them by such fuzzy numbers before further processing. In particular, an approximation algorithm that preserves the value and the ambiguity of the original fuzzy number is given by Ban *et al.* (2011), while various algorithms for trapezoidal approximations of fuzzy numbers preserving the expected interval are also accessible (Grzegorzewski, 2008). A broad collection of approximation algorithms satisfying various requirements can be found in the work of Ban *et al.* (2015).

Further sections provide a detailed description of the suggested flexible bootstrap for creating secondary samples of triangular and trapezoidal fuzzy numbers.

4. Triangular fuzzy bootstrap

Let $\widetilde{X}_1, \ldots, \widetilde{X}_n$ denote a fuzzy random sample. Assume that each realization of this sample is given by fuzzy numbers $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}_{in}(\mathbb{R})$. Following Section 3, we will create bootstrap samples of triangular fuzzy numbers, i.e., $\mathbb{F}_{out}(\mathbb{R}) = \mathbb{F}^{\Delta}(\mathbb{R})$, which preserve some characteristics of the original observations.

Thus, given observation $\widetilde{x}_i$ we generate a new triangular fuzzy number $\widetilde{x}^*_{ij} = \widetilde{x}^*_{ij}(c^*_{ij}; l^*_{ij}, r^*_{ij})$ such that $\mathrm{Val}(\widetilde{x}^*_{ij}) = \mathrm{Val}(\widetilde{x}_i)$ and $\mathrm{Amb}(\widetilde{x}^*_{ij}) = \mathrm{Amb}(\widetilde{x}_i)$. Obviously, although the value and the ambiguity characterize nicely a fuzzy number, they do not identify it completely. By fixing the value and the ambiguity we impose some restrictions on a fuzzy number, but we have still some room for the choice of the particular membership function. Let us analyze how it works.

Given $(\mathrm{Val}(\widetilde{x}_i), \mathrm{Amb}(\widetilde{x}_i))$ and assuming that $(\mathrm{Val}(\widetilde{x}^*_{ij}), \mathrm{Amb}(\widetilde{x}^*_{ij})) = (\mathrm{Val}(\widetilde{x}_i), \mathrm{Amb}(\widetilde{x}_i))$ we will try to design formulae for parameters $c^*_{ij}; l^*_{ij}, r^*_{ij}$ describing univocally $\widetilde{x}^*_{ij}$. By (4) and (6), we obtain

$$\begin{cases} r^*_{ij} - l^*_{ij} = 6\mathrm{Val}(\widetilde{x}_i) - 6c^*_{ij}, \\ r^*_{ij} + l^*_{ij} = 6\mathrm{Amb}(\widetilde{x}_i); \end{cases}$$

moreover, by definition, $r^*_{ij}, l^*_{ij} \geq 0$. Some immediate transformations yield

$$\begin{cases} l^*_{ij} = 3(\mathrm{Amb}(\widetilde{x}_i) - \mathrm{Val}(\widetilde{x}_i) + c^*_{ij}), \\ r^*_{ij} = 3(\mathrm{Amb}(\widetilde{x}_i) + \mathrm{Val}(\widetilde{x}_i) - c^*_{ij}), \end{cases} \tag{13}$$

and hence, by $r^*_{ij}, l^*_{ij} \geq 0$, we obtain

$$\mathrm{Val}(\widetilde{x}_i) - \mathrm{Amb}(\widetilde{x}_i) \leq c^*_{ij} \leq \mathrm{Val}(\widetilde{x}_i) + \mathrm{Amb}(\widetilde{x}_i). \tag{14}$$

Now we are able to formulate the desired approach for generating b triangular bootstrap samples. Keeping in mind Eqns. (13) and (14), we obtain Algorithm 1.

As suggested in Section 3, one may prefer, for some reasons, basic characteristics of fuzzy number other then the value/ambiguity, like the expected value and the width. Then, given observation $\widetilde{x}_i$, we generate a new

Algorithm 1. VA method for triangular fuzzy numbers.

Require: Fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}_{in}(\mathbb{R})$
```
for i = 1 to n do
    Compute Val(x̃_i), Amb(x̃_i)
end for
for i = 1 to b do
    for j = 1 to n do
        Generate (with equal probabilities) a pair (Val*, Amb*) from
            {(Val(x̃_1), Amb(x̃_1)), ..., (Val(x̃_n), Amb(x̃_n))}
        Generate c*_ij from the uniform distribution on the
            interval [Val* − Amb*, Val* + Amb*]
        l*_ij ← 3[Amb* − Val* + c*_ij]
        r*_ij ← 3[Amb* + Val* − c*_ij]
        x̃*_ij ← x̃*_ij(c*_ij; l*_ij, r*_ij)
    end for
end for
```

triangular fuzzy number $\widetilde{x}^*_{ij} = \widetilde{x}^*_{ij}(c^*_{ij}; l^*_{ij}, r^*_{ij})$ such that $\mathrm{EV}(\widetilde{x}^*_{ij}) = \mathrm{EV}(\widetilde{x}_i)$ and $\mathrm{w}(\widetilde{x}^*_{ij}) = \mathrm{w}(\widetilde{x}_i)$. Similarly, as in the previous case, we will try to design formulae for parameters $c^*_{ij}; l^*_{ij}, r^*_{ij}$ describing univocally $\widetilde{x}^*_{ij}$. By (10) and (12), we obtain

$$\begin{cases} r^*_{ij} - l^*_{ij} = 4\mathrm{EV}(\widetilde{x}_i) - 4c^*_{ij}, \\ r^*_{ij} + l^*_{ij} = 4\mathrm{w}(\widetilde{x}_i), \end{cases}$$

where $r^*_{ij}, l^*_{ij} \geq 0$. Hence

$$\begin{cases} l^*_{ij} = 2(\mathrm{w}(\widetilde{x}_i) - \mathrm{EV}(\widetilde{x}_i) + c^*_{ij}), \\ r^*_{ij} = 2(\mathrm{w}(\widetilde{x}_i) + \mathrm{EV}(\widetilde{x}_i) - c^*_{ij}), \end{cases} \tag{15}$$

so by $r^*_{ij}, l^*_{ij} \geq 0$ we obtain

$$\mathrm{EV}(\widetilde{x}_i) - \mathrm{w}(\widetilde{x}_i) \leq c^*_{ij} \leq \mathrm{EV}(\widetilde{x}_i) + \mathrm{w}(\widetilde{x}_i). \tag{16}$$

Hence, by (15) and (16), a method for generating b triangular samples which preserve the expected value and the width of the primary sample is given in Algorithm 2.

Consider the following example illustrating the proposed algorithms.

Example 1. Suppose $\widetilde{x} = (6; 1, 2)$ is a randomly chosen triangular observation. Hence $\mathrm{Val}(\widetilde{x}) = 6\frac{1}{6}$ and $\mathrm{Amb}(\widetilde{x}) = \frac{1}{2}$. By Algorithm 1 the core c^* of the new fuzzy number $\widetilde{x}^*$ is randomly generated from the uniform distribution on the interval $\left(5\frac{2}{3}, 6\frac{2}{3}\right)$. Suppose, e.g., $c^* = 6\frac{1}{3}$ has been selected. Then by (13) we obtain $l^* = 2$ and $r^* = 1$, so the resulting fuzzy number is $\widetilde{x}^*_{VA} = \left(6\frac{1}{3}; 2, 1\right)$.

For the same initial observation we have $\mathrm{EV}(\widetilde{x}) = 6\frac{1}{4}$ and $\mathrm{w}(\widetilde{x}) = \frac{3}{4}$. Now, using Algorithm 2, the core c^* of $\widetilde{x}^*$ is randomly generated from $U\left(5\frac{1}{2}, 7\right)$. If, e.g., $c^* = 5\frac{3}{4}$ has been selected, then by (15) we obtain $l^* = \frac{1}{2}$ and $r^* = 2\frac{1}{2}$, so the EW-method produces $\widetilde{x}^*_{Ew} = \left(5\frac{3}{4}; \frac{1}{2}, 2\frac{1}{2}\right)$. All three fuzzy numbers: $\widetilde{x}$, $\widetilde{x}^*_{VA}$ and $\widetilde{x}^*_{EW}$, are shown in Fig. 2. ♦

Algorithm 2. EW method for triangular fuzzy numbers.

Require: Fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}_{in}(\mathbb{R})$
- **for** $i = 1$ to n **do**
  - Compute $\mathrm{EV}(\widetilde{x}_i), \mathrm{w}(\widetilde{x}_i)$
- **end for**
- **for** $i = 1$ to b **do**
  - **for** $j = 1$ to n **do**
    - Generate (with equal probabilities) a pair $(\mathrm{EV}^*, \mathrm{w}^*)$ from $\left\{\left(\mathrm{EV}(\widetilde{x}_1), \mathrm{w}(\widetilde{x}_1)\right), \ldots, \left(\mathrm{EV}(\widetilde{x}_n), \mathrm{w}(\widetilde{x}_n)\right)\right\}$
    - Generate c^*_{ij} from the uniform distribution on the interval $\left[\mathrm{EV}^* - \mathrm{w}^*, \mathrm{EV}^* + \mathrm{w}^*\right]$
    - $l^*_{ij} \leftarrow 2\left[\mathrm{w}^* - \mathrm{EV}^* + c^*_{ij}\right]$
    - $r^*_{ij} \leftarrow 2\left[\mathrm{w}^* + \mathrm{EV}^* - c^*_{ij}\right]$
    - $\widetilde{x}^*_{ij} \leftarrow \widetilde{x}^*_{ij}(c^*_{ij}; l^*_{ij}, r^*_{ij})$
  - **end for**
- **end for**

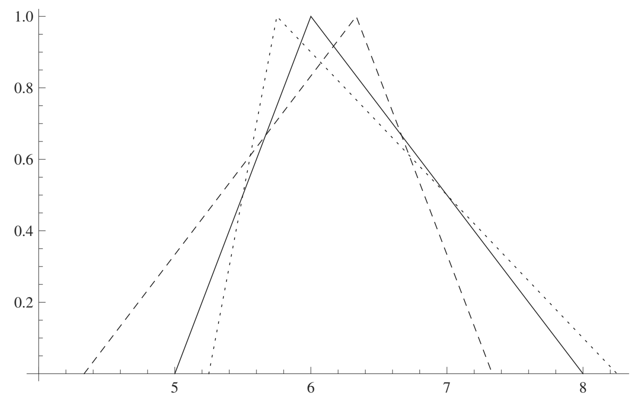

Fig. 2. Initial triangular observation $\widetilde{x}$ (solid line) and generated fuzzy numbers $\widetilde{x}^*_{VA}$ and $\widetilde{x}^*_{EW}$ for the VA-method (dashed line) and the EW-method (dotted line), respectively.

5. Trapezoidal fuzzy bootstrap

Similarly as in Section 4, we will create bootstrap samples of randomly generated trapezoidal fuzzy numbers $\widetilde{x}^*_{ij} = \widetilde{x}^*_{ij}(c^*_{ij}, s^*_{ij}; l^*_{ij}, r^*_{ij}) \in \mathbb{F}^T(\mathbb{R}) = \mathbb{F}_{out}(\mathbb{R})$ which preserve the value and the ambiguity of the original observation, i.e., $\mathrm{Val}(\widetilde{x}^*_{ij}) = \mathrm{Val}(\widetilde{x}_i)$ and $\mathrm{Amb}(\widetilde{x}^*_{ij}) = \mathrm{Amb}(\widetilde{x}_i)$. As in Section 4, starting from these fixed parameters, one has to determine $(c^*_{ij}, s^*_{ij}; l^*_{ij}, r^*_{ij})$.

Given $(\mathrm{Val}(\widetilde{x}_i), \mathrm{Amb}(\widetilde{x}_i))$, by (4) and (5) we have

$$\begin{cases} r^*_{ij} - l^*_{ij} = 6\mathrm{Val}(\widetilde{x}_i) - 6c^*_{ij}, \\ r^*_{ij} + l^*_{ij} = 6\mathrm{Amb}(\widetilde{x}_i) - 6s^*_{ij}, \end{cases}$$

where $s_{ij}, r_{ij}, l_{ij} \geq 0$, which is equivalent to

$$\begin{cases} l^*_{ij} = 3(\mathrm{Amb}(\widetilde{x}_i) - \mathrm{Val}(\widetilde{x}_i) + c^*_{ij} - s^*_{ij}), \\ r^*_{ij} = 3(\mathrm{Amb}(\widetilde{x}_i) + \mathrm{Val}(\widetilde{x}_i) - c^*_{ij} - s^*_{ij}). \end{cases} \tag{17}$$

Since $r_{ij}, l_{ij} \geq 0$, we get

$$\mathrm{Val}(\widetilde{x}_i) - \mathrm{Amb}(\widetilde{x}_i) + s^*_{ij} \leq c^*_{ij} \leq \mathrm{Val}(\widetilde{x}_i) + \mathrm{Amb}(\widetilde{x}_i) - s^*_{ij}, \tag{18}$$

where $s^*_{ij} \geq 0$. However, since the upper bound of (18) may not be smaller than its lower bound, we obtain additionally that

$$0 \leq s^*_{ij} \leq \mathrm{Amb}(\widetilde{x}_i). \tag{19}$$

Summing up the aforementioned considerations and Eqns. (17), (18) and (19), we obtain the bootstrap algorithm for trapezoidal fuzzy data (see Algorithm 3).

Algorithm 3. VA method for trapezoidal fuzzy numbers.

Require: Fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}_{in}(\mathbb{R})$
for $i = 1$ to n **do**
  Compute $\mathrm{Val}(\widetilde{x}_i), \mathrm{Amb}(\widetilde{x}_i)$
end for
for $i = 1$ to b **do**
  for $j = 1$ to n **do**
    Generate (with equal probabilities) a pair $(\mathrm{Val}^*, \mathrm{Amb}^*)$ from $\{(\mathrm{Val}(\widetilde{x}_1), \mathrm{Amb}(\widetilde{x}_1)), \ldots, (\mathrm{Val}(\widetilde{x}_n), \mathrm{Amb}(\widetilde{x}_n))\}$
    Generate s^*_{ij} from the uniform distribution on the interval $[0, \mathrm{Amb}^*]$
    Generate c^*_{ij} from the uniform distribution on the interval $[\mathrm{Val}^* - \mathrm{Amb}^* + s^*_{ij}, \mathrm{Val}^* + \mathrm{Amb}^* - s^*_{ij}]$
    $l^*_{ij} \leftarrow 3[\mathrm{Amb}^* - \mathrm{Val}^* + c^*_{ij} - s^*_{ij}]$
    $r^*_{ij} \leftarrow 3[\mathrm{Amb}^* + \mathrm{Val}^* - c^*_{ij} - s^*_{ij}]$
    $\widetilde{x}^*_{ij} \leftarrow \widetilde{x}^*_{ij}(c^*_i, s^*_i; l^*_i, r^*_i)$
  end for
end for

If one decides to generate bootstrap samples with the fixed expected value and the width, i.e., $\mathrm{EV}(\widetilde{x}^*_{ij}) = \mathrm{EV}(\widetilde{x}_i)$ and $\mathrm{w}(\widetilde{x}^*_{ij}) = \mathrm{w}(\widetilde{x}_i)$, then by (10) and (11) we have

$$\begin{cases} r^*_{ij} - l^*_{ij} = 4\mathrm{EV}(\widetilde{x}_i) - 4c^*_{ij}, \\ r^*_{ij} + l^*_{ij} = 4\mathrm{w}(\widetilde{x}_i) - 4s^*_{ij}, \end{cases}$$

where $s_{ij}, r_{ij}, l_{ij} \geq 0$, which is equivalent to

$$\begin{cases} l^*_{ij} = 2(\mathrm{w}(\widetilde{x}_i) - \mathrm{EV}(\widetilde{x}_i) + c^*_{ij} - s^*_{ij}), \\ r^*_{ij} = 2(\mathrm{w}(\widetilde{x}_i) + \mathrm{EV}(\widetilde{x}_i) - c^*_{ij} - s^*_{ij}). \end{cases} \tag{20}$$

Since $r_{ij}, l_{ij} \geq 0$, we get

$$\mathrm{EV}(\widetilde{x}_i) - \mathrm{w}(\widetilde{x}_i) + s^*_{ij} \leq c^*_{ij} \leq \mathrm{EV}(\widetilde{x}_i) + \mathrm{w}(\widetilde{x}_i) - s^*_{ij}, \tag{21}$$

where $s^*_{ij} \geq 0$. Because the upper bound of (21) may not be smaller than its lower bound, we additionally obtain

$$0 \leq s^*_{ij} \leq \mathrm{w}(\widetilde{x}_i), \tag{22}$$

which finally leads to Algorithm 4.

Let us illustrate the last two algorithms with the following example.

Example 2. Suppose that $\widetilde{x} = (6, \frac{1}{2}; 1, 2)$ is a randomly chosen trapezoidal observation from the initial sample; hence $\mathrm{Val}(\widetilde{x}) = 6\frac{1}{6}$ and $\mathrm{Amb}(\widetilde{x}) = 1$. By Algorithm 3, half of the core s^* of $\widetilde{x}^*$ is randomly generated from $U[0, 1]$. Suppose, e.g., that $s^* = \frac{1}{2}$ is selected, then by (21) c^* is generated from $U\left[5\frac{2}{3}, 6\frac{2}{3}\right]$. Let us assume that we obtain $c^* = 6\frac{1}{3}$. Then, by (17), we have $l^* = 2$ and $r^* = 1$, so the VA-method produces $\widetilde{x}^*_{VA} = \left(6\frac{1}{3}, \frac{1}{2}; 2, 1\right)$.

Algorithm 4. EW method for trapezoidal fuzzy numbers.

Require: Fuzzy sample $\widetilde{x}_1, \ldots, \widetilde{x}_n \in \mathbb{F}_{in}(\mathbb{R})$
for $i = 1$ to n **do**
  Compute $\mathrm{EV}(\widetilde{x}_i), \mathrm{w}(\widetilde{x}_i)$
end for
for $i = 1$ to b **do**
  for $j = 1$ to n **do**
    Generate (with equal probabilities) a pair $(\mathrm{EV}^*, \mathrm{w}^*)$ from $\{(\mathrm{EV}(\widetilde{x}_1), \mathrm{w}(\widetilde{x}_1)), \ldots, (\mathrm{EV}(\widetilde{x}_n), \mathrm{w}(\widetilde{x}_n))\}$
    Generate s^*_{ij} from the uniform distribution on the interval $[0, \mathrm{w}^*]$
    Generate c^*_{ij} from the uniform distribution on the interval $[\mathrm{EV}^* - \mathrm{w}^* + s^*_{ij}, \mathrm{EV}^* + \mathrm{w}^* - s^*_{ij}]$
    $l^*_{ij} \leftarrow 2[\mathrm{w}^* - \mathrm{EV}^* + c^*_{ij} - s^*_{ij}]$
    $r^*_{ij} \leftarrow 2[\mathrm{w}^* + \mathrm{EV}^* - c^*_{ij} - s^*_{ij}]$
    $\widetilde{x}^*_{ij} \leftarrow \widetilde{x}^*_{ij}(c^*_i, s^*_i; l^*_i, r^*_i)$
  end for
end for

For the same initial observation $\widetilde{x}$, we have $\mathrm{EV}(\widetilde{x}) = 6\frac{1}{4}$ and $\mathrm{w}(\widetilde{x}) = 1\frac{1}{4}$. Then, by Algorithm 4, s^* is generated from $U\left[0, 1\frac{1}{4}\right]$. If $s^* = \frac{1}{4}$ is randomly selected, then c^* is generated from $U\left[5\frac{1}{4}, 7\frac{1}{4}\right]$. Suppose that $c^* = 7$. Then, by (20), we get $l^* = 3\frac{1}{2}, r^* = \frac{1}{2}$, and the EW-method result is $\widetilde{x}^*_{EW} = \left(7, \frac{1}{4}; 3\frac{1}{2}, \frac{1}{2}\right)$. ♦

6. Simulation study

6.1. Employed models of fuzzy numbers. In our simulations we use initial samples which consist of various types of triangular fuzzy numbers, i.e., $\mathbb{F}_{in}(\mathbb{R}) = \mathbb{F}^{\Delta}(\mathbb{R})$, or trapezoidal fuzzy numbers, i.e., $\mathbb{F}_{in}(\mathbb{R}) = \mathbb{F}^{T}(\mathbb{R})$. Consequently, we adopt either $\mathbb{F}_{out}(\mathbb{R}) = \mathbb{F}^{\Delta}(\mathbb{R})$ or $\mathbb{F}_{out}(\mathbb{R}) = \mathbb{F}^{T}(\mathbb{R})$, respectively.

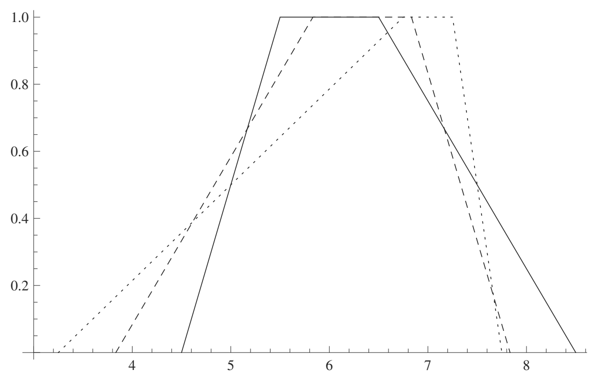

Fig. 3. Initial triapezoidal observation $\widetilde{x}$ (solid line) and generated fuzzy numbers $\widetilde{x}^*_{VA}$ and $\widetilde{x}^*_{EW}$ for the VA-method (dashed line) and the EW-method (dotted line), respectively.

To obtain random samples of triangular or trapezoidal fuzzy numbers, we simply have to generate independent tuples consisting of three or four reals corresponding to (c, l, r) or (c, s, l, r), respectively (Sinova *et al.*, 2012). Moreover, each element of a given tuple is an output of a random number generator for some specified distribution. We also assume that the elements of each tuple are generated independently. Therefore, choosing different distributions and their parameters, we may obtain easily various fuzzy numbers. The particular distributions applied in our study are summarized in Table 1. Most of them were previously used by Lubiano *et al.* (2017), Romaniuk and Hryniewicz (2019b), Romaniuk (2019) or Grzegorzewski *et al.* (2019) as models of the initial samples in numerical analyses of the bootstrapped versions of statistical tests.

The notation applied in Table 1 is self-explanatory, e.g., $\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$ indicates a triangular fuzzy number with the center generated from the standard normal distribution and the spread of the left and the right arms generated from the two i.i.d. chi-square distributions. On the other hand, $\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$ stands for a trapezoidal fuzzy number with its center generated from the normal distribution, half of its core generated from the exponential distribution and with spreads of the left and the right arms generated by two independent uniform distributions. The last type, $\mathbb{F}^{T}_{\mathrm{B,Ucon}}$, represents a more complex fuzzy number which involves the beta distribution describing the centers and a few different conditional uniform distributions for s, l, r, described in a more detailed way by Lubiano *et al.* (2017).

In the following, to limit the paper length, we present only results obtained for some selected types of fuzzy numbers (other results are available upon request). In all graphs, the results obtained with the VA-method, EW-method, d-method and classical bootstrap are marked by diamonds, triangles, squares, and circles, respectively.

6.2. Standard error estimation. One of the most widely considered applications of the bootstrap is the problem of the standard error estimation. Let $X_1, \ldots, X_n$ denote a random sample from the distribution p_θ, where $\theta \in \Theta$ is an unknown parameter. Moreover, let $\widehat{\theta}$ denote an estimator of θ and let $\mathrm{SE}(\theta) = \sqrt{\mathrm{Var}\widehat{\theta}}$ stand for its standard error. Since $\mathrm{SE}(\theta)$ can be calculated exactly only in some rare cases, it is usually estimated, and the bootstrap has appeared to be a very useful tool in this context.

Therefore, we included the problem of the standard error approximation for the mean estimation into our study. The algorithms introduced in Sections 4 and 5 are compared with the d-method proposed by Romaniuk and Hryniewicz (2019b) and the classical Efron bootstrap.

Fuzzy random samples of different sizes ($n = 5, 10, 30, 100$) and the types discussed in Section 6.1 were generated. Moreover, since different numbers of the bootstrap replications $b = 100, 200, 1000$ were applied, we could investigate a possible influence of n and b on the results. Each numerical experiment was iterated 100000 times to minimize the influence of randomness and strengthen our reasoning. To calculate the standard error, an estimator related to the Fréchet-type variance

$$\widehat{\mathrm{SE}}(\widehat{\theta}) = \sqrt{\frac{1}{b-1}\sum_{i=1}^{b} D_\theta^2\left(\widehat{\theta}^{*i}, \bar{\theta}^*\right)}, \tag{23}$$

where D_θ is the mid/spread distance with $\theta = 1$ (Casals *et al.*, 2013), $\widehat{\theta}^{*i}$ is an estimator of θ based on the i-th boostrap replication, and $\bar{\theta}^* = \sum_{i=1}^{b} \widehat{\theta}^{*i}$, was applied.

Some experimental results can be found in Tables 2–5 (others are available upon request). To facilitate the comparison of the results, the lowest values of the simulated standard errors are given in boldface.

Generally speaking, the results obtained for the different resampling algorithms do not differ substantially, especially if the sample size and the number of the bootstrap iterations are large enough. However, some conclusions are worth mentioning. Firstly, for $\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$ and $\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$, the VA-method is the only winner. For $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$, $\mathbb{F}^{\Delta}_{\mathrm{N,'U',U'}}$ and $\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$, there is no overall winner but the d-method is usually the best one. The situation seems to be more complex in the case of $\mathbb{F}^{T}_{\Gamma,\mathrm{U,E,E}}$, but here the VA-method usually leads to the lowest standard error.

To compare the resampling methods in a more synthetic way, a ranking table is also provided (see Table 6). The ranks are calculated according to a simple rule: the method giving the lowest standard error most often is the winner. Usually, either the VA-method or the d-method appears at the top. The EW-method seems to be worse, but it behaves in a very stable manner (it never drops below the third position, which happens both for the VA-method and the d-method). In our experiments, the classical bootstrap never wins. Moreover, when comparing the VA-method and the

Table 1. Description of types of simulated fuzzy numbers.

FN type	c	s	l	r
$\mathbb{F}^{\Delta}_{\mathrm{N,E,E}}$	N(0,1)	–	Exp(2)	Exp(4)
$\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$	N(0,1)	–	U(0,0.4)	U(0,0.4)
$\mathbb{F}^{\Delta}_{\mathrm{N',U',U'}}$	N(10,9)	–	U(0,1)	U(0,1)
$\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$	N(0,1)	–	$\chi^2(1)$	$\chi^2(1)$
$\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$	N(0,4)	Exp(4)	U(0,0.1)	U(0,0.2)
$\mathbb{F}^{T}_{\Gamma,\mathrm{U,E,E}}$	$\Gamma(2,2)$	U(0,0.2)	Exp(4)	Exp(2)
$\mathbb{F}^{T}_{\mathrm{B,Ucon}}$	$\beta(1,1)$	U (conditional)		
$\mathbb{F}^{T}_{\mathrm{B,Ucon'}}$	$\beta(5,5)$	U (conditional)		

Table 2. Empirical standard errors for $\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$.

n	5	10	30	100
b	100			
VA	**0.38545**	**0.29790**	**0.18145**	**0.10112**
EW	0.38637	0.29910	0.18224	0.10153
d-method	0.38687	0.29930	0.18226	0.10155
bootstrap	0.38645	0.29912	0.18221	0.10158
b	200			
VA	**0.38576**	**0.29871**	**0.18171**	**0.10124**
EW	0.38672	0.29975	0.18248	0.10167
d-method	0.38753	0.29961	0.18238	0.10175
bootstrap	0.38719	0.29928	0.18248	0.10167
b	1000			
VA	**0.38580**	**0.29874**	**0.18210**	**0.10134**
EW	0.38762	0.29965	0.18255	0.10182
d-method	0.38806	0.30049	0.18255	0.10179
bootstrap	0.38765	0.29993	0.18264	0.10181

Table 3. Empirical standard errors for $\mathbb{F}^{\triangle}_{\mathrm{N,U,U}}$.

n	5	10	30	100
b	100			
VA	0.37633	0.29181	0.17799	0.099224
EW	0.37613	0.29187	0.17806	0.099239
d-method	**0.37565**	**0.29166**	**0.17794**	**0.099190**
bootstrap	0.37623	0.29176	0.17794	0.099229
b	200			
VA	0.37673	0.29262	0.17826	0.09935
EW	0.37635	0.29256	0.17831	0.09938
d-method	**0.37635**	**0.29199**	**0.17806**	0.09939
bootstrap	0.37683	0.29200	0.17823	**0.09932**
b	1000			
VA	**0.37684**	0.29262	0.17857	0.09946
EW	0.37734	**0.29251**	0.17839	0.09953
d-method	0.37685	0.29290	**0.17822**	**0.09944**
bootstrap	0.37737	0.29260	0.17836	0.09946

Table 4. Empirical standard errors for $\mathbb{F}^{\triangle}_{\mathrm{N},\chi^2,\chi^2}$.

n	5	10	30	100
b	100			
VA	**0.47153**	**0.36174**	**0.21935**	**0.12224**
EW	0.48644	0.37417	0.22740	0.12680
d-method	0.49244	0.37866	0.22956	0.12798
Bootstrap	0.48825	0.37606	0.22932	0.12780
b	200			
VA	**0.47179**	**0.36189**	**0.21950**	**0.12229**
EW	0.48667	0.37466	0.22756	0.12698
d-method	0.49264	0.37831	0.22986	0.12803
Bootstrap	0.48848	0.37681	0.22947	0.12802
b	1000			
VA	**0.47263**	**0.36234**	**0.21976**	**0.12235**
EW	0.48737	0.37491	0.22783	0.12707
d-method	0.49276	0.37840	0.23008	0.12813
Bootstrap	0.48931	0.37738	0.22960	0.12805

Table 5. Empirical standard errors for $\mathbb{F}^{T}_{\Gamma,\mathrm{U,E,E}}$.

n	5	10	30	100
b	100			
VA	**1.02110**	0.80425	**0.49738**	**0.27978**
EW	1.02121	0.80491	0.49827	0.28001
d-method	1.02232	**0.80408**	0.49870	0.28005
bootstrap	1.02687	0.80569	0.49890	0.28002
b	200			
VA	1.02468	**0.80260**	0.49920	**0.28014**
EW	1.02258	0.80541	**0.49868**	0.28031
d-method	1.02262	0.80628	0.49951	0.28031
bootstrap	**1.02101**	0.80669	0.49940	0.28033
b	1000			
VA	1.02479	**0.80547**	0.49982	0.28052
EW	1.02491	0.80624	**0.49944**	**0.28050**
d-method	1.02599	0.80625	0.49955	0.28068
bootstrap	**1.02331**	0.80606	0.49962	0.28079

classical bootstrap, there are some significant differences between the estimated values of the standard errors, even about 3.5%.

6.3. Bootstrap in hypothesis testing. In contrast to real data usually there are not suitable models for the distribution of fuzzy random variables. Moreover, central limit theorems for fuzzy random variables often cannot be directly applied in statistical inference. Fortunately, the bootstrap appears here as a powerful tool to redeem the situation. In particular, it is often used in hypothesis testing with fuzzy data to support the decision on rejection or acceptance of the hypothesis under study. This was the reason to examine the suggested resampling methods also in this field.

In this section, we present experimental results concerning the test for the mean equipped with different resampling algorithms. As the respective statistical test, we used the procedure developed by Colubi (2009) and then algorithmically summarized by Lubiano *et al.* (2016). From now on it will be called the C-test. For other examples of statistical tests concerning fuzzy data, see, e.g., the works of Gil *et al.* (2006), González-Rodríguez *et al.* (2006), Ramos-Guajardo and Lubiano (2012) or Montenegro *et al.* (2004).

Consider a fuzzy random sample $(\widetilde{X}_1, \ldots, \widetilde{X}_n)$ and

Table 6. Ranking table of resampling methods for empirical standard errors.

FN type	VA	EW	d-method	bootstrap
$\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$	1	2	3	4
$\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$	2	2	1	4
$\mathbb{F}^{\Delta}_{\mathrm{N,'U','U'}}$	2	3	1	4
$\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$	1	2	4	3
$\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$	4	3	1	2
$\mathbb{F}^{T}_{\Gamma,\mathrm{U,E,E}}$	1	2	4	3

the following hypothesis testing the mean problem:

$$H_0 : \mathbb{E}\,\widetilde{X} = \widetilde{v} \quad \text{vs.} \quad H_1 : \mathbb{E}\,\widetilde{X} \neq \widetilde{v}, \tag{24}$$

where $\mathbb{E}\,\widetilde{X}$ is the Aumman-type mean (see Puri and Ralescu, 1986) and $\widetilde{v} \in \mathbb{F}(\mathbb{R})$ is a fixed fuzzy number corresponding to the true population mean.

6.3.1. Empirical size of a test. As the essential benchmark we use the empirical size of the test $\hat{\alpha}$ (i.e., the percentage of null hypothesis rejections when it is true) and its relation to the nominal significance level α (we assume the standard value of 0.05).

Both small and medium initial sample sizes and different numbers of bootstrap replications ($n = 5, 10, 30, 100$ and $b = 100, 200, 1000$) were used. In each experiment the whole resampling procedure was iterated 10^5 times. A similar approach was considered by Gil *et al.* (2006), González-Rodríguez *et al.* (2006), Ramos-Guajardo and Lubiano (2012), Montenegro *et al.* (2004), Romaniuk and Hryniewicz (2019b), Romaniuk (2019) or Grzegorzewski *et al.* (2019).

Selected results of our simulations can be found in Tables 7–11 (other results are available upon request). To emphasize some significant differences, the empirical size $\hat{\alpha}$ closest to the true value $\alpha = 0.05$ is printed in boldface. It is easily seen that the resampling methods considered do not differ vastly and no method is the overall winner. However, one may also conclude that the classical bootstrap is usually the worst, especially for the lower values of n and b.

A kind of ranking of the methods considered (similar to the one shown in Section 6.2) is given in Table 12. Here, a method giving empirical sizes closest to the true α most often is considered the winner. In Table 12, the classical bootstrap never occupies the first position and is the second one only in the single case. The other resampling methods are far better, especially the d-method and the VA-method. The EW-method again behaves in a relatively stable manner and never drops below the third position. It should be pointed out that the relative differences for $\hat{\alpha}$ are quite significant, even about 0.006 (more than 10% of the nominal significance level α), when the classical bootstrap approach is compared with other resampling algorithms.

Table 7. Empirical C-test size $\hat{\alpha}$ for $\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$.

n	5	10	30	100
b	100			
VA	0.03223	0.04789	**0.05381**	**0.04939**
EW	0.03049	0.04738	0.05641	0.05436
d-method	**0.03499**	**0.05050**	0.05716	0.06052
bootstrap	0.02877	0.04821	0.05702	0.06018
b	200			
VA	0.02613	0.04299	**0.05008**	0.04542
EW	0.02620	0.04184	0.05123	**0.05089**
d-method	**0.03029**	**0.04410**	0.05249	0.05531
bootstrap	0.02454	0.04398	0.05330	0.05514
b	1000			
VA	0.02365	0.03812	0.04562	0.04173
EW	0.02379	0.03933	0.04717	0.04598
d-method	**0.02761**	**0.04187**	0.05124	**0.05076**
bootstrap	0.02229	0.03967	**0.04962**	0.05111

Table 8. Empirical C-test size $\hat{\alpha}$ for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$.

n	5	10	30	100
b	100			
VA	0.03279	**0.04978**	0.05803	0.05945
EW	0.03154	0.04848	0.05816	**0.05932**
d-method	**0.05066**	0.05050	0.05705	0.06011
bootstrap	0.04920	0.04821	**0.05688**	0.06080
b	200			
VA	0.02651	**0.04511**	0.05391	**0.05404**
EW	**0.02686**	0.04273	0.05407	0.05492
d-method	0.02613	0.04343	**0.05262**	0.05553
bootstrap	0.02451	0.04436	0.05337	0.05521
b	1000			
VA	0.02431	0.04014	0.04930	0.05040
EW	**0.02443**	0.04075	0.04852	**0.05002**
d-method	0.02411	**0.04123**	0.05124	0.05094
bootstrap	0.02261	0.04021	**0.04957**	0.05121

6.3.2. Power analysis. The next step of our investigation is a power study of the C-test. To examine the power of this test, we estimate the number of rejections under increasing shift $\epsilon \in \mathbb{R}$ of realizations of the initial fuzzy sample, namely, $\epsilon = 0.1, 0.2, 0.3, 0.4, 0.5$.

To shorten the paper we provide detailed results only for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$ (see Tables 13–15 and Figs. 4–5). Generally, the power of the C-test equipped with different resampling techniques is rather similar, especially for bigger values of n and b (like $n = 100$ and $b = 1000$; see Fig. 5).

Table 9. Empirical C-test size $\hat{\alpha}$ for $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

n	5	10	30	100
b	100			
VA	0.02709	0.04533	**0.05433**	**0.04850**
EW	**0.02896**	0.04770	0.05623	0.05549
d-method	0.02745	0.04775	0.05735	0.05916
bootstrap	0.02714	**0.04902**	0.05804	0.06018
b	200			
VA	0.02293	0.04164	0.04818	0.04455
EW	**0.02336**	0.04362	**0.05094**	**0.05082**
d-method	0.02295	0.04261	0.05271	0.05539
bootstrap	0.02321	**0.04380**	0.05322	0.05421
b	1000			
VA	0.02059	0.03719	0.04581	0.03973
EW	**0.02184**	0.03840	0.04732	0.04796
d-method	0.02014	**0.0388**	0.05047	**0.05048**
bootstrap	0.02036	0.03878	**0.04959**	0.05115

Table 10. Empirical C-test size $\hat{\alpha}$ for $\mathbb{F}^T_{\Gamma,\mathrm{U,E,E}}$.

n	5	10	30	100
b	100			
VA	0.05154	0.06741	**0.05869**	**0.05569**
EW	**0.05062**	0.06725	0.06086	0.05724
d-method	0.05408	0.06748	0.06310	0.06156
bootstrap	0.05104	**0.06634**	0.0609	0.05997
b	200			
VA	0.04724	0.06181	**0.05526**	**0.05242**
EW	0.04520	0.06164	0.05801	0.05348
d-method	**0.04941**	**0.06135**	0.05819	0.05408
bootstrap	0.04658	0.06223	0.05833	0.05446
b	1000			
VA	0.04371	0.05765	**0.05061**	0.04747
EW	0.04140	0.05819	0.05425	0.04982
d-method	**0.04489**	**0.05733**	0.05428	**0.04999**
bootstrap	0.04224	0.05864	0.05528	0.05065

Some significant differences appear for smaller n and b. To emphasize them, the highest power in each experiment is given in boldface. The results are then summarized in the form of the ranking list in Table 15. One can notice that the VA-method and the EW-method usually take the first or second positions. The power curves for $n = 5$, $b = 100$ and $n = 100$, $b = 1000$ (i.e., for the smallest and the largest values of n and b, respectively) are given in Figs. 4–5.

Additionally, the power curves for the $\mathbb{F}^T_{\mathrm{N,E,U,U}}$ and $\mathbb{F}^T_{\mathrm{B,Ucon}}$ models are given in Figs. 6 and 7, respectively. It seems that the classical bootstrap leads to lower power especially for smaller n and b, which are quite common in real-life applications.

Table 11. Empirical C-test size $\hat{\alpha}$ for $\mathbb{F}^T_{\mathrm{B,Ucon'}}$.

n	5	10	30	100
b	100			
VA	**0.03774**	0.04790	**0.05315**	**0.04746**
EW	0.03601	**0.04995**	0.05860	0.05883
d-method	0.03496	0.05058	0.05707	0.05974
bootstrap	0.02797	0.04780	0.05742	0.05852
b	200			
VA	**0.03308**	**0.04552**	**0.04969**	0.04292
EW	0.03141	0.04532	0.05371	0.05453
d-method	0.02940	0.04484	0.05433	0.05446
bootstrap	0.02435	0.04443	0.05349	**0.05431**
b	1000			
VA	**0.03065**	0.04068	0.04548	0.04095
EW	0.02923	**0.04246**	**0.05004**	0.05042
d-method	0.02622	0.04216	0.05006	0.05057
bootstrap	0.02248	0.03896	0.04967	**0.05016**

Table 12. Ranking table of resampling methods for empirical C-test size $\hat{\alpha}$.

FN type	VA	EW	d-method	bootstrap
$\mathbb{F}^T_{\mathrm{N,E,U,U}}$	2	3	1	3
$\mathbb{F}^\Delta_{\mathrm{N,U,U}}$	2	1	3	4
$\mathbb{F}^\Delta_{\mathrm{N},\chi^2,\chi^2}$	4	2	1	3
$\mathbb{F}^T_{\mathrm{N,E,U,U}}$	4	1	3	2
$\mathbb{F}^T_{\Gamma,\mathrm{U,E,E}}$	1	3	1	4
$\mathbb{F}^T_{\mathrm{B,Ucon}}$	2	3	1	4
$\mathbb{F}^T_{\mathrm{B,Ucon'}}$	1	2	4	3

6.4. Real-life example. Lubiano *et al.* (2016) considered the data from the TIMSS-PIRLS 2011 questionnaire on reading, mathematics and science. Statistical tests are conducted for different hypotheses, e.g., to show potential differences between the Likert and were fuzzy rating scales.

In the following we use the same data (i.e., the responses to the item M.2: *My math teacher is easy to understand*) and the C-test to compare the four bootstrap algorithms. We verify the hypothesis (24) for a few different values of $\widetilde{v} \in \mathbb{F}^T(\mathbb{R})$ or $\widetilde{v} \in \mathbb{F}^\Delta(\mathbb{R})$. The number of the bootstrap replications b is set to 100, 200, and 1000 as in the previous experiments. For the estimated p-values we refer the reader to Table 16.

If we choose $\widetilde{v}$ close to the sample mean, like $\widetilde{v} = [6, 7, 8, 9]$ or $\widetilde{v} = [7, 8, 9]$, the p-value corresponding to each bootstrap method is large enough to indicate no reason for rejecting the null hypothesis. It can be also noticed that the estimated p-values for the EW-method, the d-method and the classical bootstrap are quite similar and smaller than the results obtained for the VA-method.

Table 13. C-test power analysis for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$.

n	5	10	30	100
ϵ	0.1			
b	100			
VA	**0.03500**	**0.05887**	0.09293	0.18191
EW	0.03352	0.05664	**0.09482**	0.18201
d-method	0.03330	0.05861	0.09193	0.18292
bootstrap	0.03125	0.05702	0.09177	**0.18431**
b	200			
VA	**0.02858**	**0.05359**	**0.08812**	**0.17566**
EW	0.02890	0.05154	0.08706	0.17523
d-method	0.02853	0.05118	0.08646	0.17558
bootstrap	0.02577	0.05197	0.08576	0.17556
b	1000			
VA	**0.02606**	0.04844	0.08035	0.16967
EW	0.02644	**0.04864**	**0.08226**	0.16732
d-method	0.02538	0.04805	0.08214	0.16771
bootstrap	0.02458	0.04844	0.08112	**0.17015**
ϵ	0.2			
b	100			
VA	**0.04151**	**0.08519**	0.19961	0.52537
EW	0.04011	0.0832	**0.20072**	**0.52775**
d-method	0.03991	0.08482	0.19891	0.52207
bootstrap	0.03811	0.08311	0.19921	0.52352
b	200			
VA	0.03482	**0.07752**	0.19022	0.5168
EW	**0.03499**	0.07639	**0.19158**	**0.51866**
d-method	0.03377	0.07649	0.19045	0.51572
bootstrap	0.03217	0.07689	0.18731	0.51777
b	1000			
VA	**0.03110**	0.07258	0.18230	0.50922
EW	0.03165	**0.07376**	0.18321	0.50867
d-method	0.03062	0.07234	**0.18358**	0.50806
bootstrap	0.02972	0.07336	0.18139	**0.51107**

On the other hand, the VA-method produces p-values that are very stable with respect to the number of the bootstrap replications b. If $\widetilde{v}$ is not too close to the sample mean, like $\widetilde{v} = [3\frac{1}{3}, 6\frac{2}{3}, 6\frac{2}{3}, 10]$, then all methods suggest definitely the rejection of the null hypothesis.

6.5. Statistical comparison of samples. There are examples of goodness-of-fit statistical tests that can be used to compare fuzzy samples in a non-parametric way in the literature (e.g., Denoeux *et al.*, 2005). Unfortunately, because of the complexity, their practical usefulness is questionable. However, since we restrict our attention to trapezoidal (triangular) fuzzy numbers which are completely defined through their supports and cores, in this paper we apply the special version of the Kolmogorov–Smirnov two-sample test for interval-valued

Table 14. C-test power analysis for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$ (continuation).

ϵ	0.3			
b	100			
VA	**0.05246**	**0.13078**	0.37106	0.84990
EW	0.05171	0.12954	0.37092	**0.85083**
d-method	0.05037	0.13050	0.36915	0.84885
bootstrap	0.04923	0.12823	**0.37298**	0.84888
b	200			
VA	0.04417	0.12067	0.35726	0.84675
EW	**0.04437**	0.12004	**0.36242**	0.84709
d-method	0.04365	0.12016	0.36062	**0.84745**
bootstrap	0.04174	**0.12114**	0.35801	0.84660
b	1000			
VA	0.04004	0.11457	0.35105	0.84271
EW	**0.04055**	**0.11519**	**0.35334**	**0.84440**
d-method	0.03953	0.11296	0.35249	0.84266
bootstrap	0.03783	0.11454	0.35300	0.84279
ϵ	0.4			
b	100			
VA	0.06693	**0.19473**	**0.57805**	0.97781
EW	**0.06749**	0.19259	0.57495	**0.97782**
d-method	0.06569	0.19438	0.57175	0.97735
bootstrap	0.06471	0.19064	0.57627	0.97709
b	200			
VA	0.05831	0.18223	0.56558	0.97749
EW	**0.05843**	0.18206	0.56591	0.97697
d-method	0.05730	0.18250	0.56568	0.97720
bootstrap	0.05531	**0.18330**	**0.56672**	**0.97760**
b	1000			
VA	0.05237	0.17436	0.55527	0.97705
EW	**0.05362**	**0.17481**	**0.56094**	**0.97779**
d-method	0.05208	0.17243	0.5583	0.97691
bootstrap	0.05042	0.17401	0.5593	0.97663
ϵ	0.5			
b	100			
VA	0.08682	**0.27607**	**0.76183**	0.99857
EW	**0.08712**	0.27458	0.75829	**0.99860**
d-method	0.08618	0.27571	0.75891	0.99853
bootstrap	0.08504	0.27360	0.76030	0.99846
b	200			
VA	0.07619	0.26176	0.75311	0.99866
EW	**0.07641**	0.26190	**0.75442**	**0.99877**
d-method	0.07511	**0.26334**	0.75222	0.99851
bootstrap	0.07230	0.26191	0.75333	0.99872
b	1000			
VA	0.06859	0.25156	0.74835	0.99849
EW	**0.07068**	**0.25188**	**0.75235**	0.99862
d-method	0.06901	0.24994	0.74815	**0.99875**
bootstrap	0.06700	0.25185	0.74892	0.99861

data, proposed by Grzegorzewski (2018). From now on, it will be denoted as the K–S–G test.

Consider two fuzzy random samples: $\widetilde{X}_1, \dots, \widetilde{X}_1$ from the initial distribution $F^{(in)}$ and the bootstrap sample $\widetilde{X}_1^*, \dots, \widetilde{X}_1^*$ from the initial distribution $F^{(out)}$. We are interested in verifying the hull hypothesis $H_0 : F^{(in)} = F^{(out)}$ of no difference between those two distributions against the alternative hypothesis $H_1 : F^{(in)} \neq F^{(out)}$ that the distributions differ significantly. Using the K–S–G test we will actually consider a slightly more specific alternative, namely, that at least one of the following equalities does not hold:

$$F_{\mathrm{mid}_0}^{(in)} = F_{\mathrm{mid}_0}^{(out)}, \qquad F_{\mathrm{spr}_0}^{(in)} = F_{\mathrm{spr}_0}^{(out)},$$
$$F_{\mathrm{mid}_1}^{(in)} = F_{\mathrm{mid}_1}^{(out)}, \qquad F_{\mathrm{spr}_1}^{(in)} = F_{\mathrm{spr}_1}^{(out)},$$

where $F_{\mathrm{mid}_0}^{(\cdot)}$ and $F_{\mathrm{spr}_0}^{(\cdot)}$ denote the distributions of the midpoint and spread of the support of $\widetilde{X}$ and $\widetilde{X}^*$, respectively, while $F_{\mathrm{mid}_1}^{(\cdot)}$ and $F_{\mathrm{spr}_1}^{(\cdot)}$ denote the distributions of the midpoint and spread of the core of $\widetilde{X}$ and $\widetilde{X}^*$, respectively.

This way, our test is a composition of four one-dimensional goodness-of-fit tests which produce four p-values: p_{mid_0}, p_{spr_0}, p_{mid_1}, p_{spr_1}. Following Grzegorzewski (2018), to make a final decision we have to aggregate these p-values. Obviously, one may apply various aggregation operators to calculate the overal p-value. The most restrictive one is the minimum, i.e.,

$$p = \min\{p_{\mathrm{mid}_0}, p_{\mathrm{spr}_0}, p_{\mathrm{mid}_1}, p_{\mathrm{spr}_1}\}. \tag{25}$$

In Table 17 we show some examples of p-values obtained for different simulated initial samples and different bootstrap methods. In our study the initial samples consist of 50 elements and the secondary ones have 100 elements.

It is worth noting that the condition (25) is very restrictive. For instance, in the case of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$ and for the EW method we have $p_{\mathrm{mid}_0} = 0.99975$, $p_{\mathrm{spr}_0} = 0.13892$, $p_{\mathrm{mid}_1} = 0.99999$ and $p_{\mathrm{spr}_1} = 0.05880$, which finally gives $p = 0.05880$. However, if we choose some less restrictive aggregation operator, like the mean, we will obtain $p = 0.54937$. Anyway, as can be seen, assuming the significance level 0.05 and using even the most restrictive criteria (25), there are no reasons to reject

Table 15. Ranking table of resampling methods for the C-test power size for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$.

(n, b)	VA	EW	d-method	bootstrap
$(5, 100)$	1	2	3	4
$(10, 100)$	1	3	2	4
$(30, 100)$	1	2	4	3
$(100, 100)$	3	1	4	2
$(10, 200)$	2	1	3	4
$(30, 200)$	2	4	3	1
$(100, 200)$	3	1	4	2
$(200, 200)$	3	1	4	2
$(10, 1000)$	2	1	3	4
$(30, 1000)$	3	1	4	2
$(100, 1000)$	4	1	2	3
$(200, 1000)$	4	1	3	2
overall	2	1	4	3

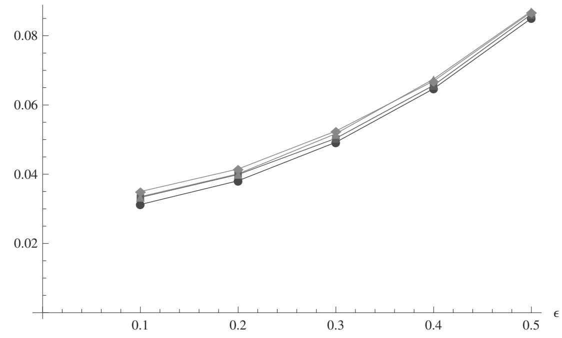

Fig. 4. Power curves of the C-test for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$ fuzzy numbers for $n = 5, b = 100$.

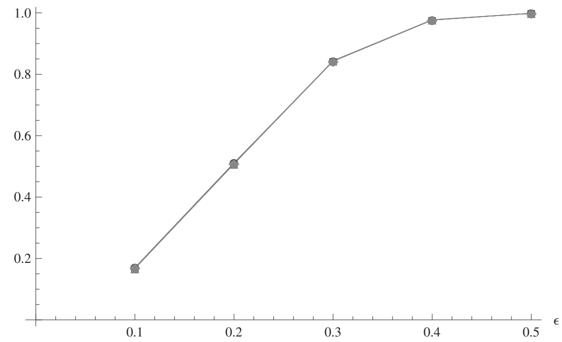

Fig. 5. Power curves of the C-test for $\mathbb{F}^{\Delta}_{\mathrm{N,U,U}}$ fuzzy numbers for $n = 100, b = 1000$.

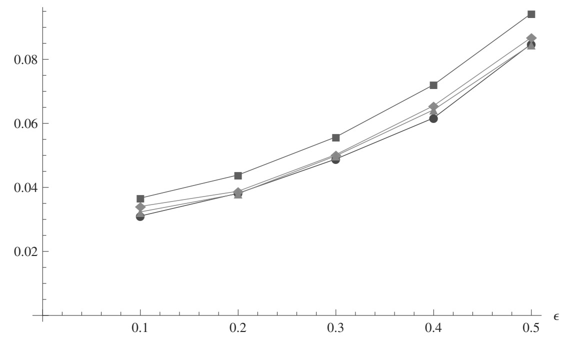

Fig. 6. Power curves of the C-test for $\mathbb{F}^{T}_{\mathrm{N,E,U,U}}$ fuzzy numbers for $n = 5, b = 100$.

the null hypothesis that there is no significant difference between the initial and the output distribution for all the types of fuzzy numbers considered in Table 17.

6.6. Graphs of means and variances. We also present sample means and variances obtained for various bootstrap methods as functions of the sample size n of $\mathbb{F}_{out}(\mathbb{R})$ both for small (5 elements) and moderate (50 elements) initial samples. To shorten the paper, we provide the results only for $\mathbb{F}^T_{\mathrm{N,E,U,U}}$ (see Figs. 8–17).

Since the Aumman-type means of the simulated distributions are trapezoidal fuzzy numbers, we provide separate plots for the lower and upper bounds of the supports, and the lower and upper bounds of the cores. In each of these plots the horizontal thick line corresponds to the respective means for the bounds of the initial sample and the x-axes are located at the y-axis exactly at the value of the mean of the given model of the fuzzy number. On the other hand, the Fréchet-type variances based on the mid/spread distance D_θ with $\theta = 1$ (Casals

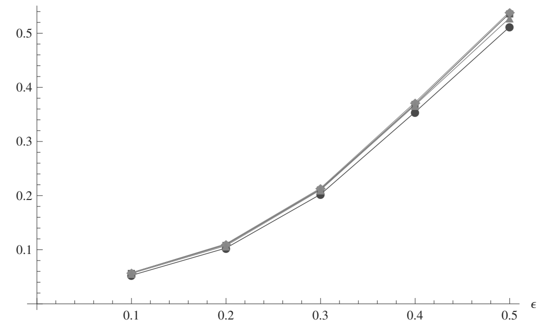

Fig. 7. Power curves of the C-test for $\mathbb{F}^T_{\mathrm{B,Ucon}}$ fuzzy numbers for $n = 5, b = 100$.

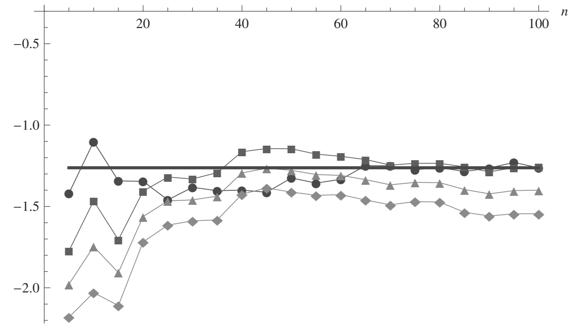

Fig. 8. Means for the left endpoint of the support for the small sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

Table 16. Empirical p-values for the C-test of the item M.2 and different null hypotheses (***$p < 0.001$).

b	100	200	1000
$\widetilde{v}$	$[6, 7, 8, 9]$		
VA	0.300	0.356	0.321
EW	0.240	0.185	0.186
d-method	0.240	0.175	0.164
bootstrap	0.160	0.185	0.166
$\widetilde{v}$	$[7, 8, 9]$		
VA	0.310	0.355	0.332
EW	0.240	0.190	0.192
d-method	0.250	0.175	0.174
bootstrap	0.160	0.200	0.169
$\widetilde{v}$	$[3\frac{1}{3}, 6\frac{2}{3}, 6\frac{2}{3}, 10]$		
VA	0.000***	0.000***	0.000***
EW	0.000***	0.000***	0.000***
d-method	0.000***	0.000***	0.000***
bootstrap	0.000***	0.000***	0.000***

Table 17. Empirical p-values for testing the difference between the initial and the output distribution.

FN type	VA	EW	d-method	bootstrap
$\mathbb{F}^T_{\mathrm{N,E,U,U}}$	0.89278	0.72315	0.72315	0.95000
$\mathbb{F}^\Delta_{\mathrm{N,U,U}}$	0.95000	0.95000	0.95000	0.95000
$\mathbb{F}^\Delta_{\mathrm{N},\chi^2,\chi^2}$	0.23030	0.44131	0.53072	0.89278
$\mathbb{F}^T_{\mathrm{N,E,U,U}}$	0.05880	0.05880	0.89278	0.99675
$\mathbb{F}^T_{\Gamma,\mathrm{U,E,E}}$	0.62623	0.07937	0.23030	0.89278
$\mathbb{F}^T_{\mathrm{B,Ucon}'}$	0.05880	0.29037	0.36077	0.89278

et al., 2013) are illustrated on single graphs since they are real numbers.

Generally speaking, the sample means calculated for all bootstrap algorithms considered tend to their population means especially if the initial samples are small. The sample means generated by the classical bootstrap are very close to the means of $\mathbb{F}_{in}(\mathbb{R})$, while those generated by other methods are more diversified. This conclusion is confirmed also by the graphs of the variances, where the results obtained for methods other than the classical bootstrap seem to be higher in the case of small $\mathbb{F}_{in}(\mathbb{R})$ (but not necessarily in the case of moderate sample sizes). Interestingly, the means for moderate $\mathbb{F}_{in}(\mathbb{R})$ seem to be closer to the population mean more often if the d-method, the VA-method or the EW-method are compared with the classical bootstrap.

6.7. Graphs of variability of the estimator. To succeed with our study, we analyse variabilities of the estimator (i.e., the average) based on various bootstrap methods as functions of the sample size n of $\mathbb{F}_{out}(\mathbb{R})$ both for the small and moderate initial samples. To shorten the paper, we provide the results only for $\mathbb{F}^\Delta_{\mathrm{N},\chi^2,\chi^2}$ (see Figs. 18–21).

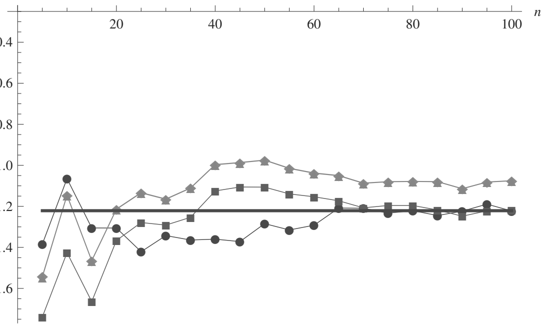

Fig. 9. Means for left endpoint of the core for the small sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

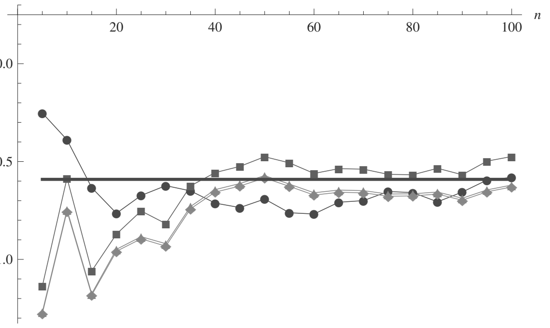

Fig. 10. Means for the right endpoint of the core for the small sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

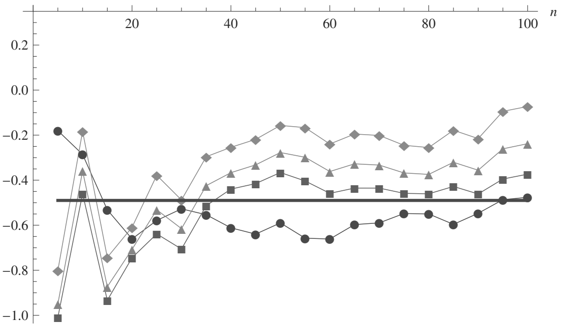

Fig. 11. Means for the right endpoint of the support for the small sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

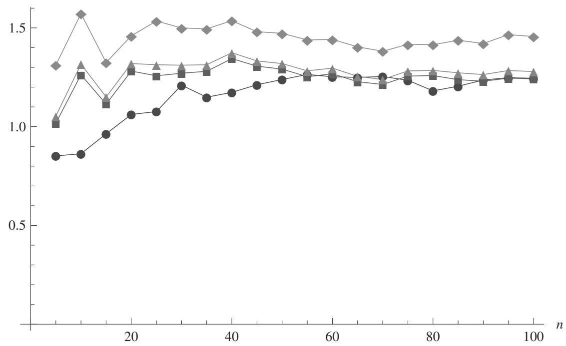

Fig. 12. Variances for the small sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

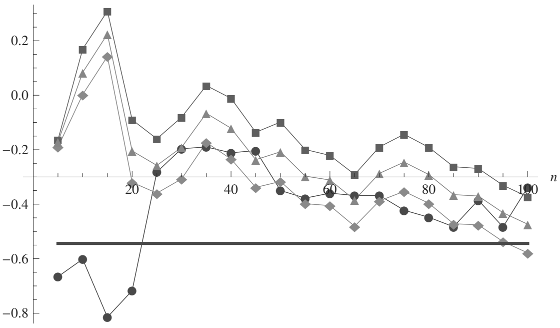

Fig. 13. Means for the left endpoint of the support for the moderate sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

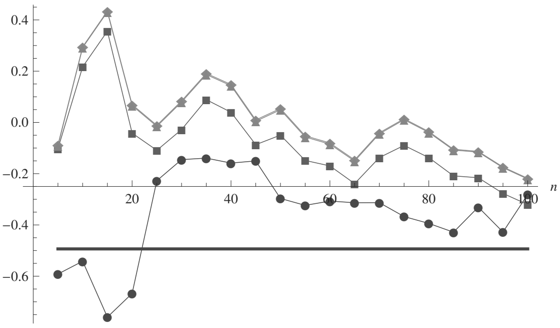

Fig. 14. Means for left endpoint of the core for the moderate sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

We calculate the variabilities related to the Fréchet-type variance (i.e., based on D_θ with $\theta = 1$; see also (23)) and the Aumman-type mean $\mathbb{E}\,\widetilde{X}$ (see also (24)), whose true value (i.e., estimated value) is known for the model considered, using

$$d^{(1)}_{\mathrm{var}}(n) = \frac{1}{n-1}\sum_{i=1}^{n} D^2_\theta\left(\tilde{X}^*_i, \mathbb{E}\,\widetilde{X}\right), \tag{26}$$

$$d^{(2)}_{\mathrm{var}}(n) = D^2_\theta\Big(\frac{1}{n}\sum_{i=1}^{n}\tilde{X}^*_i, \mathbb{E}\,\widetilde{X}\Big). \tag{27}$$

Both (26) and (27) tend to be similar for all sampling methods, but in some cases they are significantly smaller for both the VA-method and the EW-method than for the classical boostrap even for lower values of n.

7. Conclusions

A new methodology for flexible generation of the bootstrap fuzzy samples was proposed. Contrary to the classical bootstrap, our new algorithms generate samples that do not necessarily consist of observations forming the primary sample only, but they are somehow more diversified. The key idea of the suggested algorithms is to generate fuzzy numbers that preserve some crucial characteristics of the original observations (i.e., the value

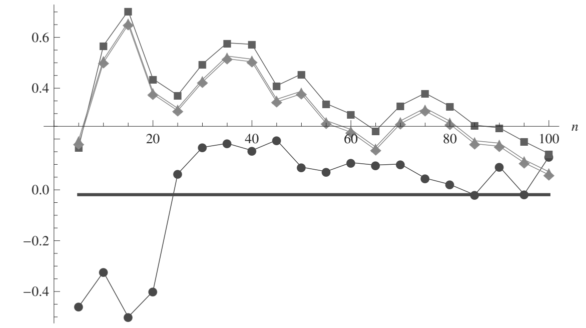

Fig. 15. Means for the right endpoint of the core for the moderate sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

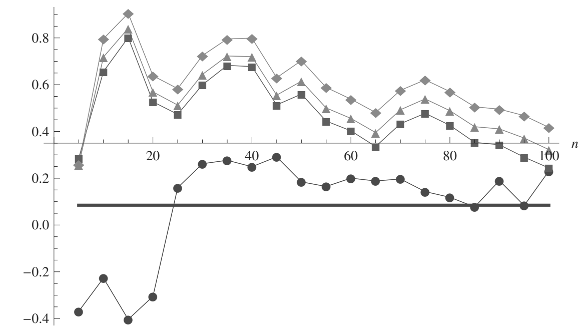

Fig. 16. Means for the right endpoint of the support for the moderate sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

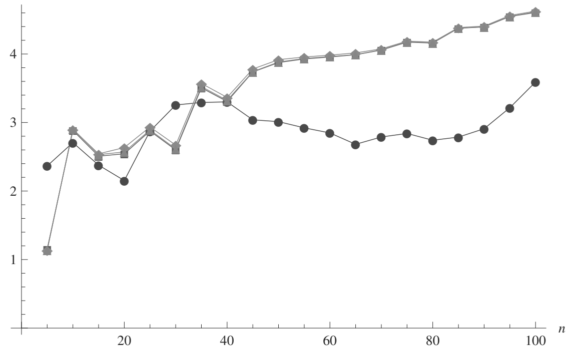

Fig. 17. Variances for the moderate sample of $\mathbb{F}^T_{\mathrm{N,E,U,U}}$.

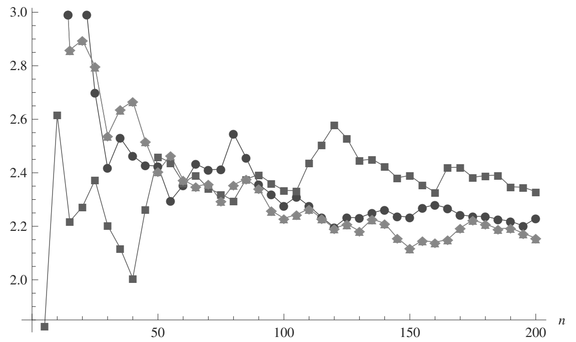

Fig. 18. Distance $d^{(1)}_{\mathrm{var}}(n)$ for the small sample of $\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$

and the ambiguity or the expected value and the width), but ignore other minor ones.

The paper delivers four bootstrap algorithms ready for direct use by practitioners. However, it is worth noting that the suggested methodology can be also applied in respectively modified jackknife algorithms.

An extended simulation study to examine various statistical properties and approaches (like the standard error estimation, benchmarking based on the empirical size of a statistical test, a power analysis, a goodness-of-fit statistical test between the initial and secondary samples, and graphs of means and variances) of the proposed bootstrap algorithms was performed, also in the case of real-life data. The results of this study, as well as the simplicity of new algorithms, indicate that the suggested approaches turn out to be a remarkable and powerful tool for making an inference and supporting decisions with fuzzy data.

References

Ban, A., Brândaş, A., Coroianu, L., Negruţiu, C. and Nica, O. (2011). Approximations of fuzzy numbers by trapezoidal fuzzy numbers preserving the ambiguity and value, *Computers and Mathematics with Applications* **61**(5): 1379–1401.

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

Casals, M.R., Corral, N., Gil, M. Á., López, M.T., Lubiano, M.A., Montenegro, M., Naval, G. and Salas, A. (2013). Bertoluzza *et al.'s* metric as a basis for analyzing fuzzy data, *METRON* **71**(3): 307–322.

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

Chanas, S. (2001). On the interval approximation of a fuzzy number, *Fuzzy Sets and Systems* **122**(2): 353–356.

Colubi, A. (2009). Statistical inference about the means of fuzzy random variables: Applications to the analysis of fuzzy- and real-valued data, *Fuzzy Sets and Systems* **160**(3): 344–356.

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.

Davison, A.C., Hinkley, D.V. and Schechtman, E. (1986). Efficient bootstrap simulation, *Biometrika* **73**(3): 555–566.

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

Delgado, M., Vila, M. and Voxman, W. (1998). On a canonical representation of a fuzzy number, *Fuzzy Sets and Systems* **93**(1): 125–135.

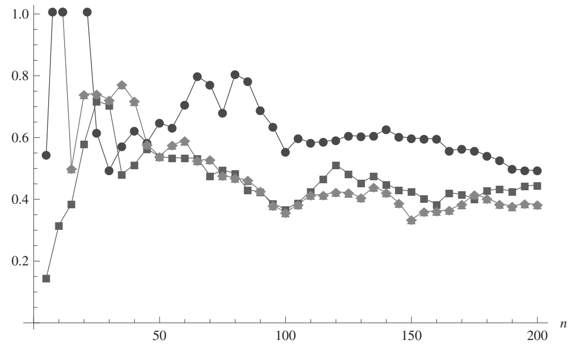

Fig. 19. Distance $d^{(2)}_{\mathrm{var}}(n)$ for the small sample of $\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$

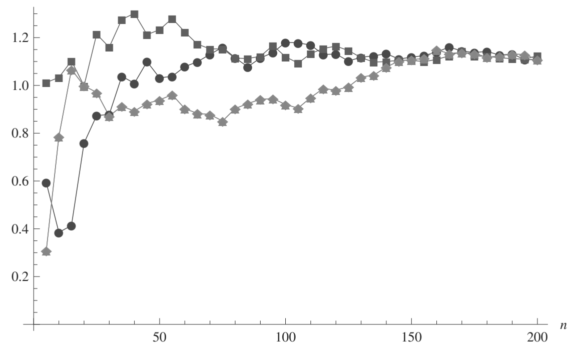

Fig. 20. Distance $d^{(1)}_{\mathrm{var}}(n)$ for the moderate sample of $\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$

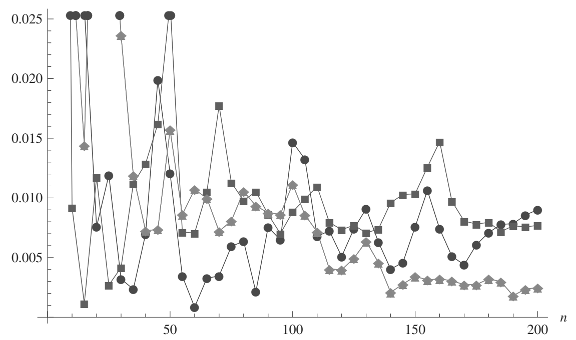

Fig. 21. Distance $d^{(2)}_{\mathrm{var}}(n)$ for the moderate sample of $\mathbb{F}^{\Delta}_{\mathrm{N},\chi^2,\chi^2}$

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

Dubois, D. and Prade, H. (1987). The mean value of a fuzzy number, *Fuzzy Sets and Systems* **24**(3): 279–300.

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

Gao, J.-Q., Fan, L.-Y., Li, L. and Xu, L.-Z. (2013). A practical application of kernel-based fuzzy discriminant analysis, *International Journal of Applied Mathematics and Computer Science* **23**(4): 887–903, DOI: 10.2478/amcs-2013-0066.

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.

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.

Graham, R., Hinkley, D.V., John, P.W.M. and Shi, S. (1990). Balanced design of bootstrap simulations, *Journal of the Royal Statistical Society B* **52**(1): 185–202.

Grzegorzewski, P. (2008). Trapezoidal approximations of fuzzy numbers preserving the expected interval—Algorithms and properties, *Fuzzy Sets and Systems* **159**(11): 1354–1364.

Grzegorzewski, P. (2018). The Kolmogorov–Smirnov goodness-of-fit test for interval-valued data, *in* E. Gil *et al.* (Eds), *The Mathematics of the Uncertain: A Tribute to Pedro Gil*, Springer International Publishing, Cham, pp. 615–627.

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.

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

Heilpern, S. (1992). The expected value of a fuzzy number, *Fuzzy Sets and Systems* **47**(1): 81–86.

Jimenez, M. and Rivas, J.A. (1998). Fuzzy number approximation, *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems* **6**(1): 68–78.

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. and 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**(1): 31–49.

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

Puri, M. and Ralescu, D.A. (1986). Fuzzy random variables, *Journal of the Mathematical Analysis and Applications* **114**(2): 409–422.

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, Cham, 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 International Publishing, Cham, pp. 437–448.

Romaniuk, M. and Hryniewicz, O. (2019a). Discrete and smoothed resampling methods for interval-valued fuzzy numbers, *IEEE Transactions on Fuzzy Systems*, DOI: 10.1109/TFUZZ.2019.2957253, (in press).

Romaniuk, M. and Hryniewicz, O. (2019b). Interval-based, nonparametric approach for resampling of fuzzy numbers, *Soft Computing* **23**(14): 5883–5903.

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

Sinova, B., Gil, M.A., Colubi, A. and Aelst, S.V. (2012). The median of a random fuzzy number. The 1-norm distance approach, *Fuzzy Sets and Systems* **200**: 99–115.

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.

Przemyslaw Grzegorzewski holds an MSc in mathematics from the University of Warsaw, a PhD with distinction and a DSc (habilitation) in computer science from the Systems Research Institute of the Polish Academy of Sciences. In 2018 he received a full professorial title. He is currently a full professor at the Faculty of Mathematics and Information Science of the Warsaw University of Technology and the Systems Research Institute of the Polish Academy of Sciences. His research interests include mathematical statistics, statistical decisions with imprecise data, data mining, fuzzy sets, fuzzy logic, soft computing, statistical quality control, etc.

Olgierd Hryniewicz holds an MSc degree in electronics from the Warsaw University of Technology, a PhD degree in cybernetics from the Institute of Management and Automatic Control, Polish Academy of Sciences, and a DSc (habilitation) in statistics from the Cracow University of Economics. Since 1996 he has held a full professorial title. He is the head of the Department of Stochastic Methods of the Systems Research Institute, Polish Academy of Sciences, and an elected member of the Committee on Informatics of the Polish Academy of Sciences. The main topics of his research are fuzzy modeling of random events and applications of fuzzy statistics in reliability and statistical quality control.

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: 16 November 2019
Revised: 20 March 2020
Accepted: 29 April 2020