Permutation k-sample Goodness-of-Fit Test for Fuzzy Data

Przemyslaw Grzegorzewski
Faculty of Mathematics and Information Science, Warsaw University of Technology,
Koszykowa 75, 00-662 Warsaw, Poland;
and *Systems Research Institute, Polish Academy of Sciences,*
Newelska 6, 01-447 Warsaw, Poland
pgrzeg@mini.pw.edu.pl

***Abstract*—The problem of testing goodness-of-fit for k distributions based on fuzzy data is considered. A new permutation test for fuzzy random variables is proposed. Besides the general constrution of the test an algorithm ready for the practical use is delivered. A case-study illustrating the applicability of the suggested testing procedure is also presented.**

***Index Terms*—fuzzy data, fuzzy number, fuzzy random variable, goodness-of-fit test, permutation test, random fuzzy number, trapezoidal fuzzy number**

I. Introduction

Most of statistical procedures are constructed with fairly specific assumptions regarding the underlying population distribution. In particular, one of most often used techniques, applied for comparing several treatments, i.e. analysis of variance (ANOVA), assumes not only independence of observations and that all populations are normally distributed but also the homogeneity of their variances. Obviously, so strong assumptions quite often are not satisfied. Unfortunately, ANOVA, like some other statistical tests is sensitive to violations of the fundamental model assumptions inherent in its derivation. In such case distribution-free methods, also called nonparametric, are very useful. In particular, the Kruskal-Wallis test could be used for comparing a few independent samples. This test, unlike ANOVA, requires neither normality nor homogeneity of variances and that is why it is sometimes called the nonparametric analogue of one-way analysis of variance.

Real-life data sets often consist of imprecise or vague observations. In particular, many human ratings based on opinions or associated with perceptions lead to data that cannot be expressed in a numerical scale. Such data consist of intrinsically imprecise or fuzzy elements. Thus, if they also appear as a realization of some random experiment, we are faced with random fuzzy variables that cannot be analyzed by classical statistical methods and require another adequate approach.

Several approaches have been developed in the literature for testing statistical hypotheses with fuzzy data. Depending on the context and whether data are perceived from the epistemic or the ontic view (see [4]), various test constructions appeared in the literature (for the overview we refer the reader e.g. to [8], [9], [12], [13]–[15], [17], [22], [26], [27], [28], [33], [32]). In particular, the problem of testing the equality of k samples against the so-called "simple-tree alternative" or "many-one problem" for fuzzy data based on the necessity index of strict dominance is considered in [17]. The bootstrap test for the equality of fuzzy means of k populations can be found in [9], while [33] contains the bootstrap procedure for testing the homoscedasticity of k populations.

One may wonder why the nonparametric Kruskal-Wallis test has not been generalized for fuzzy data. The reason is that the Kruskal-Wallis test is based on ranks which cannot be determined for fuzzy samples since fuzzy numbers are not linearly ordered. However, the Kruskal-Wallis test is actually the k-sample goodness-of-fit test, since its null hypothesis states that all k samples under study actually come from the same distribution (which is equivalent to the statement that all k populations are identically distributed). Therefore, one may consider another construction of the k-sample goodness of fit test, which does not need any ranks. Such construction based on permutations is suggested in this very contribution.

The paper is organized as follows: in Sec. II we introduce the notation and recall basic concepts related to fuzzy data modeling and operations on fuzzy numbers. Sec. III is devoted to fuzzy random variables. In Sec. IV we propose the general idea of the k-sample goodness-of-fit permutation test for fuzzy data. Besides the test construction we deliver testing algorithm ready for a practical use. Next, in Sec. V we adapt the general construction of the suggested test for the trapezoidal fuzzy numbers. Then we present some results of the simulation study (Sec. VI) and the case study (Sec. VII) with the proposed test. Finally, conclusions and some indications for the further research are given in Sec. VIII.

II. Fuzzy data

A **fuzzy number** is an imprecise value characterized by a mapping $A : \mathbb{R} \to [0, 1]$ (called a membership function) such that its α-cut defined by

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

is a nonempty compact interval for each $\alpha \in [0, 1]$. Operator cl in (1) stands for the closure. Thus every fuzzy number is completely characterized both by its memberschip function $A(x)$ and by a family of its α-cuts $\{A_\alpha\}_{\alpha\in[0,1]}$. Two α-cuts

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

are of special interest: $A_0 = \operatorname{supp}(A)$ called the **support** and $A_1 = \operatorname{core}(A)$ known as the **core** of fuzzy number A, respectively.

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

$$A(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & \text{if } a_1 \leqslant x < a_2, \\ 1 & \text{if } a_2 \leqslant x \leqslant a_3, \\ \frac{a_4-x}{a_4-a_3} & \text{if } a_3 < x \leqslant a_4, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where $a_1, a_2, a_3, a_4 \in \mathbb{R}$ such that $a_1 \leqslant a_2 \leqslant a_3 \leqslant a_4$. Such trapezoidal fuzzy number A is often denoted as $\operatorname{Tra}(a_1, a_2, a_3, a_4)$. Obviously, $a_1 = \inf \operatorname{supp}(A)$, $a_2 = \inf \operatorname{core}(A)$, $a_3 = \sup \operatorname{core}(A)$ and $a_4 = \sup \operatorname{supp}(A)$, which means that each trapezoidal fuzzy numbers is completely described by its support and core.

If $a_2 = a_3$ then A is said to be a **triangular fuzzy number**, while if $a_1 = a_2$ and $a_3 = a_4$ we have a so-called **interval** (or rectangular) fuzzy number. The families of all fuzzy numbers, trapezoidal fuzzy numbers, triangular fuzzy number and interval fuzzy numbers will be denoted by $\mathbb{F}(\mathbb{R})$, $\mathbb{F}^T(\mathbb{R})$, $\mathbb{F}^\Delta(\mathbb{R})$ and $\mathbb{F}^I(\mathbb{R})$, respectively. Obviously, $\mathbb{F}^I(\mathbb{R}) \subset \mathbb{F}^T(\mathbb{R})$, $\mathbb{F}^\Delta(\mathbb{R}) \subset \mathbb{F}^T(\mathbb{R})$ and $\mathbb{F}^T(\mathbb{R}) \subset \mathbb{F}(\mathbb{R})$.

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 A and B is given by the Minkowski addition of corresponding α-cuts, i.e.

$$(A+B)_\alpha = \left[\inf A_\alpha + \inf B_\alpha, \sup A_\alpha + \sup B_\alpha\right],$$

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

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

It is worth noting that a sum of trapezoidal fuzzy numbers is also a trapezoidal fuzzy number, namely, if $A = \operatorname{Tra}(a_1, a_2, a_3, a_4)$ and $B = \operatorname{Tra}(b_1, b_2, b_3, b_4)$ then

$$A + B = \operatorname{Tra}(a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4). \tag{3}$$

Moreover, the product of a trapezoidal fuzzy number $A = \operatorname{Tra}(a_1, a_2, a_3, a_4)$ by a scalar θ is a trapezoidal fuzzy number

$$\theta \cdot A = \begin{cases} \operatorname{Tra}(\theta \cdot a_1, \theta \cdot a_2, \theta \cdot a_3, \theta \cdot a_4) & \text{if } \theta \geqslant 0, \\ \operatorname{Tra}(\theta \cdot a_4, \theta \cdot a_3, \theta \cdot a_2, \theta \cdot a_1) & \text{if } \theta < 0. \end{cases} \tag{4}$$

Unfortuntely, $\big(\mathbb{F}(\mathbb{R}), +, \cdot\big)$ has not linear but semilinear structure since in general $A + (-1 \cdot A) \neq \mathbb{1}_{\{0\}}$. Consequently, the Minkowski-based difference does not satisfy, in general, the addition/subtraction property that $(A + (-1 \cdot B)) + B = A$. To overcome this problem the so-called Hukuhara difference was proposed. It is defined as follows:

$$C := A -_H B \quad \text{if and only if} \quad B + C = A$$

Although now the desired properties $A -_H A = \mathbb{1}_{\{0\}}$ or $(A -_H B) + B = A$ are satisfied, the Hukuhara difference does not always exist. Therefore, one should be aware that there are critical problems with a subtraction in $\mathbb{F}(\mathbb{R})$.

To overcome some of the problems associated with the lack of a satisfying difference, especially in constructing tools for statistical reasoning based on fuzzy observations, an alternative approach utilizing metrics defined in $\mathbb{F}(\mathbb{R})$ has been developed [3].

One can define various metrics in $\mathbb{F}(\mathbb{R})$ but perhaps the most often used in statistical context is the one proposed by Gil et al. [6] and Trutschnig et al. [35]. Let λ denote a normalized measure associated with a continuous distribution with support in $[0,1]$ and let $\theta > 0$. Then for any $A, B \in \mathbb{F}(\mathbb{R})$ we define a metric D_θ^λ as follows

$$\begin{aligned} D_\theta^\lambda(A,B) = \Big(\int_0^1 &\big[(\operatorname{mid} A_\alpha - \operatorname{mid} B_\alpha)^2 \\ &+ \theta \cdot (\operatorname{spr} A_\alpha - \operatorname{spr} B_\alpha)^2\big] d\lambda(\alpha)\Big)^{1/2}, \end{aligned} \tag{5}$$

where $\operatorname{mid} A_\alpha = (\inf A_\alpha + \sup A_\alpha)/2$ and $\operatorname{spr} A_\alpha = (\sup A_\alpha - \inf A_\alpha)/2$ denote the mid-point and the radius of the α-cut A_α, respectively.

Both λ and θ correspond to some weighting. Namely, λ allows us to weight the influence of each α-cut. In practice, the most common choice of λ is the Lebesgue measure on $[0,1]$. On the other hand, each particular choice of θ allows us to weight the impact of the distance between the spreads (i.e. the deviation in vagueness) of the α-cuts in contrast to the distance between their mid-points (i.e. the deviation in location). Here, the most popular choice is $\theta = 1$ or $\theta = \frac{1}{3}$. Using $\theta = 1$ we obtain the following metric

$$\begin{aligned} D_1^\lambda(A,B) = \Big(\int_0^1 &\Big[\frac{1}{2}(\inf A_\alpha - \inf B_\alpha)^2 \\ &+ \frac{1}{2}(\sup A_\alpha - \sup B_\alpha)^2\Big] d\lambda(\alpha)\Big)^{1/2}, \end{aligned} \tag{6}$$

weighting uniformly the two squared Euclidean distances and equivalent to the distance considered e.g. in [5], [11]. On the other hand, for $\theta = \frac{1}{3}$ we obtain

$$D_{1/3}^\lambda(A,B) = \sqrt{\int_0^1 \Big(\int_0^1 \big[A_\alpha^{[t]} - B_\alpha^{[t]}\big]^2 dt\Big) d\lambda(\alpha)}, \tag{7}$$

where $A_\alpha^{[t]} = (1-t)\inf A_\alpha + t \sup A_\alpha$, so $D_{1/3}^\lambda(A,B)$ aggregates uniformly the squared Euclidean distances between the convex combination of points of the α-cuts representing A and B.

Whatever (λ, θ) is chosen, D_θ^λ has some important and useful properties: it is an L^2-type metric in $\mathbb{F}(\mathbb{R})$ and it is translational and rotational invariant, i.e., for all $A, B, C, \in \mathbb{F}(\mathbb{R})$

$$D_\theta^\lambda(A+C, B+C) = D_\theta^\lambda(A,B)$$
$$D_\theta^\lambda\big((-1)\cdot A, (-1)\cdot B\big) = D_\theta^\lambda(A,B)$$

Moreover, $(\mathbb{F}(\mathbb{R}), D_\theta^\lambda)$ is a separable metric space, and for each fixed λ all D_θ^λ are topologically equivalent.

For more details on fuzzy numbers, their types, characteristics, and approximations we refer the reader to [1].

III. Fuzzy random variables

When the output of an experiment consists of a random sample of imprecise data that could be satisfactorily described by fuzzy numbers we need a model which allow to grasp both aspects of uncertainty that appear in such data: randomness, associated with data generation mechanism and fuzziness, connected with data nature, i.e. their imprecision. To cope with this problem Puri and Ralescu [31] introduced the notion of a **fuzzy random variable** (random fuzzy number).

Definition 1. *Given a probability space $(\Omega, \mathcal{A}, P)$, a mapping $X : \Omega \to \mathbb{F}(\mathbb{R})$ is said to be a fuzzy random variable if for all $\alpha \in [0, 1]$ the α-level function is a compact random interval.*

In other words, X is a random fuzzy variable if and only if X is a Borel measurable function w.r.t. the Borel σ-field generated by the topology induced by D_θ^λ.

Puri and Ralescu [31] defined also the Aumann-type mean of a fuzzy random variable X as the fuzzy number $\mathcal{E}(X) \in \mathbb{F}(\mathbb{R})$ such that for each $\alpha \in [0, 1]$ the α-cut $\big(\mathcal{E}(X)\big)_\alpha$ is equal to the Aumann integral of X_α or, in other words,

$$\big(\mathcal{E}(X)\big)_\alpha = \big[\mathbb{E}(\mathrm{mid}\, X_\alpha) - \mathbb{E}(\mathrm{spr}\, X_\alpha), \mathbb{E}(\mathrm{mid}\, X_\alpha) + \mathbb{E}(\mathrm{spr}\, X_\alpha)\big]. \tag{8}$$

We can also define (see [25]) the D_θ^λ-Fréchet-type variance $\mathcal{V}(X)$ which is a nonnegative real number such that

$$\begin{aligned} \mathcal{V}(X) &= \mathbb{E}\Big(\big[D_\theta^\lambda(X, \mathcal{E}(X))\big]^2\Big) \\ &= \int_0^1 \mathrm{Var}(\mathrm{mid}\, X_\alpha) d\lambda(\alpha) + \theta \int_0^1 \mathrm{Var}(\mathrm{spr}\, X_\alpha) d\lambda(\alpha). \end{aligned} \tag{9}$$

Although (8) preserve the main properties known from the real-valued case one should be aware of the problems typical for statistical reasoning with fuzzy data. Firstly, as it was noted in Section II, there are problems with subtraction and division of fuzzy numbers, so it is advisable to avoid these operations whenever it is possible. Next source of possible problems is the lack of universally accepted total ranking between fuzzy numbers. Moreover, absence of suitable models for the distribution of fuzzy random variables makes the reasoning really hard. Finally, there are not yet Central Limit Theorems for fuzzy random variables that can be applied directly in decision making. All the above mentioned disadvantages impede the straightforward generalization of statistical methodology well established in the classical real-valued sample environment, including test construction, etc. In most cases of the inference based on fuzzy samples we are not able to find the null distribution of a test statistic and, consequently, to find the critical value or to compute the p-value required for making a decision. One possible way out is to use bootstrap [8], which has been successfully applied in many test constructions (like [9], [26], [27], [28], [33], [32]). However, in this paper we suggest another methodology based on permutations.

IV. K-sample test for fuzzy data – a general idea

Suppose, we observe independently k fuzzy random samples, where $k \geqslant 2$, drawn from populations with unknown distributions function $F_1, \ldots, F_k$, respectively, i.e.

$$\begin{array}{rcl} \mathbb{X}_1 = (X_{11}, \ldots, X_{1n_1}) & \sim & F_1 \\ \vdots & & \vdots \\ \mathbb{X}_k = (X_{k1}, \ldots, X_{1n_k}) & \sim & F_k. \end{array}$$

Let $N = n_1 + \ldots + n_k$ denote the total number of available. observations. We want to verify the null hypothesis that all k samples come from the same distribution, i.e.

$$H_0 : F_1(t) = \ldots = F_k(t) \quad \text{for all } t \in \mathbb{R}, \tag{10}$$

against the alternative hypothesis $H_1 : \neg H_0$ that at least two population distributions differ.

The crucial idea of the proposed test construction is that the null hypothesis implies total exchangeability of observed data with respect to groups. Indeed, if H_0 holds then all N available observations may be viewed as if they were randomly assigned to k groups but they come from the same population.

Let $\mathbb{V} = \mathbb{X}_1 \uplus \ldots \uplus \mathbb{X}_k$, where $\uplus$ stands for vector concatenation, so that the k samples are pooled into one, i.e. $V_i = X_{1i}$ if $1 \leqslant i \leqslant n_1$, $V_i = X_{2i}$ if $n_1 + 1 \leqslant i \leqslant n_1 + n_2$ and so on until $V_i = X_{k,i-(n_1+\ldots n_{k-1})}$ if $n_1 + \ldots n_{k-1} + 1 \leqslant i \leqslant N$.

Now, let $\mathbb{V}^*$ denote a permutation of the initial dataset $\mathbb{V}$. More formally, if $\nu = \{1, 2, \ldots, N\}$ and π_ν is a permutation of the integers ν, then $V_i^* = V_{\pi_\nu(i)}$ for $i = 1, \ldots, N$. Then the first n_1 elements of $\mathbb{V}^*$ is assigned to the first sample $\mathbb{X}_1^*$, next n_2 elements of $\mathbb{V}^*$ is assigned to the second sample $\mathbb{X}_2^*$, etc., and the remaining n_k elements to $\mathbb{X}_k^*$. In other words, it works like a random assignment of elements into k samples of the size $n_1, \ldots, n_k$, respectively. Each permutation corresponds to some relabeling of the combined dataset $\mathbb{V}$. Please, note that if H_0 holds then we are completely free to exchange the labels X_{ij} attributed to particular observations.

Since the total number of distinct rearrangements of N elements, where n_1 elements are labeled as $X_{1\cdot}$, n_2 elements are labeled as $X_{2\cdot}$, etc., and the remaining n_k as $X_{k\cdot}$ is multinomial, hence under H_0 the probability of each randomly selected $\mathbb{V}^*$ is

$$\mathbb{P}_{H_0}\big(\mathbb{V}^* = (v_1^*, \ldots, v_N^*)\big) = \frac{n_1! \ldots n_k!}{N!}. \tag{11}$$

The next step is to choose a test statistic which is supposed to discriminate between the null hypothesis and its alternative. Our test statistic would be based on the comparison of the sample means. Indeed, the arithmetic mean (average) is recognized as a suitable measure aggregation operator indicating the typical output of a sample. Obviously, in the case of fuzzy sample the mean is also a fuzzy number. If the null hypothesis holds we expect that due to exchangeability of observed data between groups all k sample means would not differ to much from the overall sample mean. On the other hand, a significant difference between sample means may indicate that the samples under study come from different distributions. To measure how much do the sample means differ we will use a suitable distance.

To be more specific, let $\overline{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$ denote the average of observations belonging to sample $\mathbb{X}_i$, where $i = 1,\ldots,k$, while $\overline{\overline{X}} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}$ stands for the overal mean of all N available observations.

Then, given $\mathbb{V} = \mathbb{X}_1 \uplus \ldots \uplus \mathbb{X}_k$, the test statistic is defined as

$$T(\mathbb{V}) = \sum_{i=1}^{k} n_i \cdot D_\theta^\lambda(\overline{X}_i, \overline{\overline{X}})^2, \tag{12}$$

where D_θ^λ is given by (6).

To decide whether the distance between the observed sample means is large enough to conclude as significant we need some knowledge on the null distribution of the test statistic. As a consequence of elements' exchangeability in $\mathbb{V}^*$ under H_0 we can estimate the distribution of T by considering all permutations of the initial dataset $\mathbb{V}$ and computing a value of the statistic $T(\mathbb{V}^*)$ corresponding to each permutation. Substituting given permutation $\mathbb{v}^*$ into (12) we obtain

$$T(\mathbb{v}^*) = \sum_{i=1}^{k} n_i \cdot D_\theta^\lambda(\overline{x}_i^*, \overline{\overline{x}})^2 \tag{13}$$

where

$$\overline{x}_i^* = \frac{1}{n_i} \sum_{j=n_1+\ldots+n_{i-1}+1}^{n_1+\ldots+n_i} v_j^*.$$

Obviously, $\overline{\overline{x}}^* = \frac{1}{k}\sum_{i=1}^{k}\overline{x}_i^* = \frac{1}{N}\sum_{i=1}^{N} v_i^* = \overline{\overline{x}}$ for any permutation $\mathbb{v}^*$.

By (11), the null distribution od T is given by

$$\mathbb{P}_{H_0}(T = t) = \frac{\#\{\mathbb{v}^* : T(\mathbb{v}^*) = t\}}{\frac{N!}{n_1!\ldots n_k!}}, \tag{14}$$

where the numerator gives the number of permuations leading to the desired value of T.

Finally, the p-value of our test is defined as the proportion of cases when the test statistic values are greater or equal to the observed experimental value $t_0 = T(\mathbb{V})$, i.e.

$$\text{p-value} = \frac{n_1!\ldots n_k!}{N!} \cdot \sum_{\mathbb{v}^* \in \mathcal{P}(\mathbb{v})} \mathbb{1}(T(\mathbb{v}^*) \geqslant t_0), \tag{15}$$

where $\mathcal{P}(\mathbb{v})$ is a set of all distinct rearrangements of $\mathbb{v}$.

In practice, the exact distribution of T can be found for very small N only. Moreover, the numerator calculation for all cases would be also rather awkward. Therefore, instead of considering all possible permutations we consider an approximate distribution obtained by drawing randomly a large number of samples (permutations) with replacement.

Let M denote a fixed number of drawings. Then the approximate p-value of our test is given by

$$\text{p-value} \simeq \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}(T(\mathbb{v}_m^*) \geqslant t_0), \tag{16}$$

where M is usually not smaller than 1000.

To sum-up, the proposed permutation goodness-of-fit test for fuzzy random variables can be performed following Algorithm 1.

Algorithm 1

Require: Fuzzy samples $\mathbb{x}_1 = (x_{11},\ldots,x_{1n_1}),\ldots,\mathbb{x}_k = (x_{k1},\ldots,x_{kn_k})$
Ensure: Test p-value
$t_0 \longleftarrow \sum_{i=1}^{k} n_i \cdot D_\theta^\lambda(\overline{x}_i, \overline{\overline{x}})^2$
$s \longleftarrow 0$
Pool the data $\mathbb{v} = \mathbb{x}_1 \uplus \ldots \uplus \mathbb{x}_k$
for $m = 1$ to M **do**
  Take a permutation $\mathbb{v}^* = (v_1^*,\ldots,v_N^*)$ of the pooled data $\mathbb{v}$
  $l \longleftarrow 0$
  for $i = 1$ to k **do**
    $\overline{x}_i^* \longleftarrow \sum_{j=l+1}^{l+n_i} v_j^*/n_i$
    $l \longleftarrow l + n_i$
  end for
  $T \longleftarrow \sum_{i=1}^{k} n_i \cdot D_\theta^\lambda(\overline{x}_i^*, \overline{\overline{x}})^2$
  if $T \geqslant t_0$ **then**
    $s := s + 1$
  end if
end for
p-value $\longleftarrow s/M$

One can easily notice that the permutations may be divided into equivalence classes called *rearrangements*. Within each equivalence class the only differences between permutations lie in the arrangement of the observations. Hence, it seems that instead of *permutation test* it would be more adequate to term it the *rearrangement test* or *combinational test*, but the traditional name *permutation test* is already in a common usage.

Permutation tests, like the bootstrap, require extremly limited assumptions. Bootstrap tests usually rely on assumption that successive observations are independent. Permutation tests require only exchangeability (i.e., under the null hypothesis we can exchange the labels on the observations without affecting the results). Obviously, if the observations in a sample are independent and identically distributed then they are exchangeable.

There are two advantages of the permutation tests over the bootstrap tests. Firstly, permutation test are often more powerful than their bootstrap counterparts (see [10]). Secondly, permutation test are exact if all permutation are considered, while bootstrap tests are exact only for very large samples. Moreover, asymptotically permutation tests are usually as powerful as the most powerful parametric tests (see [2]).

However, one should remember that permutation tests are tools specialized for comparing distributions, whereas bootstrap tests are oriented on comparison between parameters. The bootstrap can also provide a reliable confidence intervals and standard errors estimators, beyond mere p-values delivered by permutation tests.

For more information on classical permutation tests we refer the reader to [10], [30],

V. GOODNESS-OF-FIT TEST FOR TRAPEZOIDAL DATA

In this section we adapt the general idea of the permutation goodness-of-fit test proposed in Section IV for trapezoidal data. Obviously, one may ask why do we restrict our attention to this subfamily of fuzzy numbers. Actually, it has been noticed by many researchers that trapezoidal (or triangular) fuzzy numbers are most common in various applications mainly because they are both easy to handle and have a natural interpretation (see [1], [29]). As noted in [24]: "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". Moreover, even if the original data set consists of fuzzy numbers which are not trapezoidal one often approximates them by such fuzzy numbers before further processing. One may ask why to consider a trapezoidal approximation and not simplify objects under study as much as possible, i.e. to defuzzify them. However, it is widely known that defuzzifying data too early we lose too much information and it is much better to process fuzzy information as long as possible. This is the case why we are looking for simplification to avoid difficulties in computation on the one hand and we do not want to simplify too much on the other hand. It seems that the trapezoidal approximation is a reasonable compromise between these two opposite tendencies (see, e.g., [16], [21], while for the broad collection of approximation algorithms satisfying various requirements we refer the reader to [1]). Finally, assuming trapezoidal fuzzy data we may express our permutation goodness-of-fit test in a closed analytic form.

Firstly, let us notice that the required formula (6) for the distance reduces meaningfully if we restrict our attention to trapezoidal fuzzy numbers. Indeed, for any $\alpha \in [0,1]$ the α-cut of $A = \mathrm{Tra}(a_1, a_2, a_3, a_4)$ is given by

$$A_\alpha = \big[a_1 + (a_2 - a_1)\alpha, a_4 - (a_4 - a_3)\alpha\big]. \tag{17}$$

Therefore, we easily obtain that

$$\begin{aligned} \mathrm{mid}\, A_\alpha &= \frac{1}{2}(\inf A_\alpha + \sup A_\alpha) \qquad (18)\\ &= \frac{a_1 + a_4}{2} + \Big(\frac{a_2 + a_3}{2} - \frac{a_1 + a_4}{2}\Big)\alpha \\ &= \mathrm{mid}\,\mathrm{supp}(A) + \big(\mathrm{mid}\,\mathrm{core}(A) - \mathrm{mid}\,\mathrm{supp}(A)\big)\alpha \\ &= \mathrm{mid}_0 A + \big(\mathrm{mid}_1 A - \mathrm{mid}_0 A\big)\alpha \end{aligned}$$

and

$$\begin{aligned} \mathrm{spr}\, A_\alpha &= \frac{1}{2}(\sup A_\alpha - \inf A_\alpha) \qquad (19)\\ &= \frac{a_4 - a_1}{2} + \Big(\frac{a_3 - a_2}{2} - \frac{a_4 - a_1}{2}\Big)\alpha \\ &= \mathrm{spr}\,\mathrm{supp}(A) + \big(\mathrm{spr}\,\mathrm{core}(A) - \mathrm{spr}\,\mathrm{supp}(A)\big)\alpha \\ &= \mathrm{spr}_0 A + \big(\mathrm{spr}_1 A - \mathrm{spr}_0 A\big)\alpha, \end{aligned}$$

where $\mathrm{mid}_0 A$ and $\mathrm{mid}_1 A$ denote the midpoint of the support and core of A, respectively, while $\mathrm{spr}_0 A$ and spr_1 stand for the spread of the support and core of A, respectively.

Assume that $\ell = \lambda$ is the Lebesgue measure on $[0,1]$ and $A, B \in \mathbb{F}^T(\mathbb{R})$. Then, substituting (18) and (19) into (6) we obtain the following formula for the distance D_θ^λ between trapezoidal fuzzy numbers

$$\begin{aligned} D_\theta^\ell(A,B) = \Bigg[&\frac{\big(\mathrm{mid}_0 A - \mathrm{mid}_0 B\big)^2 + \big(\mathrm{mid}_1 A - \mathrm{mid}_1 B\big)^2}{3} \\ &+ \frac{\big(\mathrm{mid}_0 A - \mathrm{mid}_0 B\big)\big(\mathrm{mid}_1 A - \mathrm{mid}_1 B\big)}{3} \\ &+ \theta\,\frac{\big(\mathrm{spr}_0 A - \mathrm{spr}_0 B\big)^2 + \big(\mathrm{spr}_1 A - \mathrm{spr}_1 B\big)^2}{3} \\ &+ \theta\,\frac{\big(\mathrm{spr}_0 A - \mathrm{spr}_0 B\big)\big(\mathrm{spr}_1 A - \mathrm{spr}_1 B\big)}{3} \Bigg]^{1/2}. \end{aligned} \tag{20}$$

Now, we are able to perform our data analysis. Let $\mathbb{x} = (x_1, \ldots, x_n)$ and $\mathbb{y} = (y_1, \ldots, y_m)$ denote realizations (i.e. actual observations) of the fuzzy random samples $\mathbb{X} = (X_1, \ldots, X_n)$ and $\mathbb{Y} = (Y_1, \ldots, Y_m)$, respectively. Further on we assume that all our imprecise observations are modeled by trapezoidal fuzzy numbers, i.e. $x_i \in \mathbb{F}^T(\mathbb{R})$ for each $i = 1, \ldots, n$ and $y_j \in \mathbb{F}^T(\mathbb{R})$ for each $j = 1, \ldots, m$.

By (3) and (4) it is obvious that the arithmetic mean of trapezoidal fuzzy numbers is also a trapezoidal fuzzy number. Indeed, the sample mean of trapezoidal observations is completely described by averages of the lower and upper bounds of the supports and cores of observations. Thus, consequently, the mid-point of the support of the sample mean is equal to the average of the mid-points of supports of observations, i.e. $\mathrm{mid}_0 \overline{x} = \overline{\mathrm{mid}_0 x}$ and similar results hold for the core $\mathrm{mid}_1 \overline{x} = \overline{\mathrm{mid}_1 x}$ and both with respect to the radii of the support and core, i.e. $\mathrm{spr}_0 \overline{x} = \overline{\mathrm{spr}_0 x}$ and $\mathrm{spr}_1 \overline{x} = \overline{\mathrm{spr}_1 x}$.

Therefore, given aforementioned observations $\mathbb{x} = (x_1, \ldots, x_n)$ and $\mathbb{y} = (y_1, \ldots, y_m)$ of the two fuzzy random samples and assuming that all observations are trapezoidal fuzzy numbers, i.e. $x_i, y_j \in \mathbb{F}^T(\mathbb{R})$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, the value of our test statistic (12) can be calculated using the following formula

$$t_0 = T(\mathbb{v}) = \sum_{i=1}^{k} n_i \cdot D_\theta^\ell(\overline{x}_i, \overline{\overline{x}})^2, \tag{21}$$

where, by (20)

$$\begin{aligned} D_\theta^\ell(\overline{x}_i, \overline{\overline{x}}) = \Bigg[&\frac{\big(\mathrm{mid}_0 \overline{x}_i - \mathrm{mid}_0 \overline{\overline{x}}\big)^2 + \big(\mathrm{mid}_1 \overline{x}_i - \mathrm{mid}_1 \overline{\overline{x}}\big)^2}{3} \\ &+ \frac{\big(\mathrm{mid}_0 \overline{x}_i - \mathrm{mid}_0 \overline{\overline{x}}\big)\big(\mathrm{mid}_1 \overline{x}_i - \mathrm{mid}_1 \overline{\overline{x}}\big)}{3} \\ &+ \theta\,\frac{\big(\mathrm{spr}_0 \overline{x}_i - \mathrm{spr}_0 \overline{\overline{x}}\big)^2 + \big(\mathrm{spr}_1 \overline{x}_i - \mathrm{spr}_1 \overline{\overline{x}}\big)^2}{3} \\ &+ \theta\,\frac{\big(\mathrm{spr}_0 \overline{x}_i - \mathrm{spr}_0 \overline{\overline{x}}\big)\big(\mathrm{spr}_1 \overline{x}_i - \mathrm{spr}_1 \overline{\overline{x}}\big)}{3} \Bigg]^{1/2}. \end{aligned} \tag{22}$$

Next, we take M permutations of $\mathbb{v} = \mathbb{x}_1 \uplus \ldots \uplus \mathbb{x}_k$ and split each permutation $\mathbb{v}_m^*$ into k parts of size $n_1, \ldots, n_k$, respectively, such that $\mathbb{v}_m^* = \mathbb{x}_1^* \uplus \ldots \uplus \mathbb{x}_k^*$, where $\mathbb{x}_i^* = (x_{i1}^*, \ldots, x_{in_i}^*)$, $i = 1, \ldots, k$, we compute the desired averages $\overline{x}_i^*$ and substitute them into (22) instead of $\overline{x}_i$ to obtain $D_\theta^\ell(\overline{x}_i^*, \overline{\overline{x}})$. Finally, following (16), we will be able to compute the p-value of our test, i.e.

$$\text{p-value} = \frac{1}{M} \sum_{m=1}^{M} \mathbb{1}\Big(\sum_{i=1}^{k} n_i \cdot D_\theta^\ell(\overline{x}_i^*, \overline{\overline{x}})^2 \geqslant t_0 \Big). \qquad (23)$$

In the case of triangular fuzzy numbers, i.e. $x_{i,j} \in \mathbb{F}^\Delta(\mathbb{R})$ for all $j = 1, \ldots, n_i$ and $j = 1, \ldots, k$, the formula for distance (22) simplifies into

$$D_\theta^\ell(\overline{x}_i, \overline{\overline{x}}) = \Bigg[\frac{\big(\text{mid}_0 \overline{x}_i - \text{mid}_0 \overline{\overline{x}}\big)^2 + \big(\text{mid}_1 \overline{x}_i - \text{mid}_1 \overline{\overline{x}}\big)^2}{3} + \frac{\big(\text{mid}_0 \overline{x}_i - \text{mid}_0 \overline{\overline{x}}\big)\big(\text{mid}_1 \overline{x}_i - \text{mid}_1 \overline{\overline{x}}\big)}{3} + \theta \frac{\big(\text{spr}_0 \overline{x}_i - \text{spr}_0 \overline{\overline{x}}\big)^2}{3} \Bigg]^{1/2}.$$

Obviously, we have even more simplification for intervals, i.e. for $x_{i,j} \in \mathbb{F}^I(\mathbb{R})$ for all $j = 1, \ldots, n_i$ and $j = 1, \ldots, k$, we obtain

$$D_\theta^\ell(\overline{x}_i, \overline{\overline{x}}) = \Big[\big(\text{mid}_0 \overline{x}_i - \text{mid}_0 \overline{\overline{x}}\big)^2 + \frac{\theta}{3} \big(\text{spr}_0 \overline{x}_i - \text{spr}_0 \overline{\overline{x}}\big)^2 \Big]^{1/2}.$$

VI. Simulations

Some simulations were conducted to illustrate the behavior of the proposed test.

To generate fuzzy samples from a trapezoidal-valued fuzzy random variable $X = \text{Tra}(\xi_1, \xi_2, \xi_3, \xi_4)$, where $\xi_1, \xi_2, \xi_3, \xi_4$ are real-valued random variables such that $\xi_1 \leqslant \xi_2 \leqslant \xi_3 \leqslant \xi_4$, the following characterization appears to be useful (see [27], [34]): $c = \frac{1}{2}(\xi_3 + \xi_2) = \text{mid}_1 X$, $s = \frac{1}{2}(\xi_3 - \xi_2) = \text{spr}_1 X$, $l = \xi_2 - \xi_1$ and $r = \xi_4 - \xi_3$. Hence, conversely, we have $\text{Tra}\langle c, s, l, r\rangle = \text{Tra}(c - s - l, c - s, c + s, c + s + r)$.

Thus, in our study we generated fuzzy observations $\mathbb{x}_{ij}$ by simulating the four real-valued random variables $\langle c_{ij}, s_{ij}, l_{ij}, r_{ij}\rangle$ with the last three ones random variables in each quartet being nonnegative. In particular, we generated trapezoidal-valued fuzzy random variables using the following real-valued random variables: c_{ij} from the normal distribution and s_{ij}, l_{ij}, r_{ij} from the uniform or chi-square distribution.

We examined the proposed permutation test with respect to its type 1 error. Therefore, 10 000 simulations of the test performed on three independent fuzzy samples comming from the same distribution were generated at the significance level 0.05. In each test $M = 1000$ permutations were drawn. Then empirical percentages of rejections under H_0 were determined. The results both for equal and nonequal sample sizes are gathered in Tables I and II, respectively.

It is seen that the empirical type 1 error of the proposed test seems to be stable and is close to the preassumed level. It becomes a little more liberal for large sample sizes.

TABLE I: Empirical percentages of rejections under H_0 for equal sample sizes $n_1 = n_2 = n_3 = n$.

n	percentage of rejections
10	4.87
15	4.78
20	4.88
50	5.20
100	5.31

TABLE II: Empirical percentages of rejections under H_0 for unequal sample sizes

(n_1, n_2, n_3)	percentage of rejections
$(10, 15, 20)$	5.06
$(10, 20, 10)$	5.01
$(20, 50, 100)$	4.96
$(50, 50, 100)$	5.11

VII. Illustrative example

The Gamonedo cheese is a kind of a blue cheese produced Asturias, Spain (see, e.g., [7]). It experiences a smoked process and later on is let settle in natural caves or a dry place. Keeping the quality of a cheese, due to its sensitivity and complexity, requires a solid tasting system. Usually, the experts (or tasters) express their subjective perceptions about different characteristics of the cheese, like visual parameters (shape, rind and appearance), texture parameters (hardness and crumbliness), olfactory-gustatory parameters (smell intensity, smell quality, flavour intensity, flavour quality and aftertaste) and an overall impression of the cheese.

So far, the experts (tasters) provide an ordinal number ranging from 1 to 7 to describe their perceptions about different cheese characteristics. However, recently some of the tasters were proposed to express their subjective perceptions about the quality of the Gamonedo cheese by using trapezoidal fuzzy numbers. These fuzzy sets were determined in the following way: the set of values considered by the expert to be fully compatible with his/her opinion led to $\alpha = 1$-cut, while the set of values that he/she considered to be compatible with his/her opinion at some extent (i.e., the taster thought that it was not possible that the quality was out of this set) led to $\alpha = 0$-cut of a fuzzy number. Then these two α-cuts were linearly interpolated to get the trapezoidal fuzzy set representing exppert's personal valuation.

Some statistical analyses of fuzzy data obtained in this way for the Gamonedo cheese quality inspection was performed by Ramos-Guajardo and Lubiano [33] and Ramos-Guajardo et al. [32]. In particular, some considerations based on the bootstrap generalization of the Levene test for k variances are given in [33]. On the other hand, a construction of robust summary measures for the fuzzy opinions of tasters and their applicability to the sensory evaluation of the Gamonedo cheese can be found in [32].

Here we utilize some data given in [32] to compare the opinions of the three experts about the overall impression of the Gamonedo cheese (the trapezoidal fuzzy sets corresponding to their opinions are gathered in Table III).

Thus we have three independent fuzzy samples of sizes $n_1 = 40$, $n_2 = 38$ and $n_3 = 42$, comming from the unknown distributions F_1, F_2 and F_3, respectively. Our problem is to check whether there is a general agreement between these experts. To reach the goal we verify the following null hypothesis

$$H_0 : F_1 = F_2 = F_3,$$

stating there is no significant difference between experts' opinions, against $H_1 : \neg H_0$ that their opinions on the cheese quality differ.

TABLE III: Sample of the opinions of Expert 1, 2 and 3 concerning the overall impression of the Gamonedo cheese (see [32])

Opinion	Expert 1	Expert 2	Expert 3
1	$(65, 75, 85, 85)$	$(50, 50, 63, 75)$	$(60, 63, 67, 72)$
2	$(35, 37, 44, 50)$	$(39, 47, 52, 60)$	$(53, 58, 63, 68)$
3	$(66, 70, 75, 80)$	$(60, 70, 85, 90)$	$(43, 47, 54, 58)$
4	$(70, 74, 80, 84)$	$(50, 56, 64, 74)$	$(70, 76, 83, 86)$
5	$(65, 70, 75, 80)$	$(39, 45, 53, 57)$	$(54, 60, 65, 70)$
6	$(45, 50, 57, 65)$	$(55, 60, 70, 76)$	$(76, 80, 83, 86)$
7	$(60, 66, 70, 75)$	$(50, 50, 57, 67)$	$(65, 68, 73, 80)$
8	$(65, 65, 70, 76)$	$(65, 67, 80, 87)$	$(77, 80, 86, 90)$
9	$(60, 65, 75, 80)$	$(50, 50, 65, 75)$	$(76, 80, 85, 90)$
10	$(55, 60, 66, 70)$	$(50, 55, 64, 70)$	$(70, 76, 80, 85)$
11	$(60, 65, 70, 74)$	$(39, 46, 53, 56)$	$(50, 51, 55, 64)$
12	$(30, 46, 44, 54)$	$(19, 29, 41, 50)$	$(43, 47, 51, 58)$
13	$(60, 65, 75, 75)$	$(40, 47, 52, 56)$	$(50, 55, 60, 64)$
14	$(70, 75, 85, 85)$	$(54, 55, 65, 76)$	$(65, 67, 73, 80)$
15	$(44, 45, 50, 56)$	$(59, 65, 75, 85)$	$(65, 70, 75, 80)$
16	$(51, 56, 64, 70)$	$(50, 52, 57, 60)$	$(50, 55, 60, 65)$
17	$(40, 46, 54, 60)$	$(60, 60, 70, 80)$	$(65, 70, 75, 80)$
18	$(55, 60, 65, 70)$	$(50, 54, 61, 67)$	$(74, 80, 85, 90)$
19	$(80, 85, 90, 94)$	$(40, 46, 50, 50)$	$(46, 50, 55, 60)$
20	$(80, 84, 90, 90)$	$(44, 50, 56, 66)$	$(50, 57, 64, 70)$
21	$(65, 70, 76, 80)$	$(60, 64, 75, 85)$	$(65, 74, 80, 84)$
22	$(75, 80, 86, 90)$	$(54, 56, 64, 75)$	$(55, 58, 64, 70)$
23	$(65, 70, 73, 80)$	$(50, 50, 60, 66)$	$(65, 73, 80, 85)$
24	$(70, 80, 84, 84)$	$(44, 46, 55, 57)$	$(54, 57, 62, 70)$
25	$(55, 64, 70, 70)$	$(59, 63, 74, 80)$	$(73, 80, 85, 90)$
26	$(64, 73, 80, 84)$	$(49, 50, 54, 58)$	$(54, 60, 65, 70)$
27	$(50, 56, 64, 70)$	$(55, 60, 70, 75)$	$(50, 55, 60, 64)$
28	$(55, 55, 60, 70)$	$(44, 47, 53, 60)$	$(65, 74, 80, 84)$
29	$(60, 70, 75, 80)$	$(19, 20, 30, 41)$	$(40, 47, 53, 60)$
30	$(64, 71, 80, 80)$	$(40, 44, 50, 60)$	$(46, 50, 57, 64)$
31	$(50, 50, 55, 65)$	$(50, 50, 59, 66)$	$(55, 60, 65, 74)$
32	$(50, 54, 60, 65)$	$(50, 53, 60, 66)$	$(50, 57, 63, 70)$
33	$(65, 75, 80, 86)$	$(50, 52, 58, 61)$	$(40, 47, 53, 60)$
34	$(50, 55, 60, 66)$	$(60, 65, 72, 80)$	$(65, 70, 76, 80)$
35	$(40, 44, 50, 50)$	$(50, 50, 55, 60)$	$(55, 60, 65, 70)$
36	$(70, 76, 85, 85)$	$(30, 34, 43, 47)$	$(70, 74, 83, 90)$
37	$(44, 50, 53, 60)$	$(19, 25, 36, 46)$	$(60, 66, 74, 81)$
38	$(34, 40, 46, 46)$	$(53, 63, 74, 80)$	$(64, 70, 75, 80)$
39	$(40, 45, 51, 60)$		$(40, 44, 51, 56)$
40	$(84, 90, 95, 95)$		$(35, 40, 46, 50)$
41			$(35, 44, 50, 55)$
42			$(66, 70, 75, 85)$

Substituting data from Table III into formula (21) we obtain $t_0 = 2259.436$. Then, after combining samples and generating $M = 10\,000$ random permutations we have obtained the p-value of 0.0011. Hence, assuming the typical 5-percent significant level we may conclude that there is no general aggreement between experts' opinion on the overal impression of the Gamonedo cheese. In Figure 1 one can find the empirical null distribution of the permutation goodness-of-fit test with red vertical line indicating the value t_0 of the test statistic.

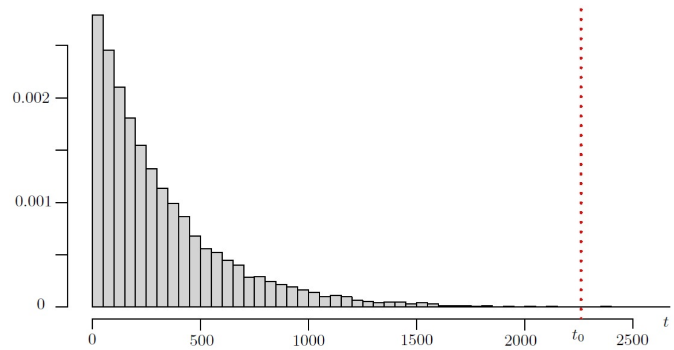

Fig. 1: Empirical null distribution of the permutation goodness-of-fit test with red vertical line indicating the value of the test statistic.

VIII. Conclusions

Hypothesis testing with samples which consist of fuzzy numbers generated by fuzzy random variables is neither easy nor straightforward. Till now tests based on the bootstrap were mainly developed and used in this area. Another approach for constracting statistical tests for the aforementioned data was proposed in this paper. Namely, the k-sample permutation goodness-of-fit for fuzzy data was developed. Moreover, some simulation studies and the case study dedicated to fuzzy rating problem were performed. Although the results obtained seem to be promissing, further research including power studies and a comparison with other tests are intended in the nearest future.

References

[1] A. I. Ban, L. Coroianu and P. Grzegorzewski, "Fuzzy Numbers: Approximations, Ranking and Applications," Polish Academy of Sciences, Warsaw, 2015.

[2] P. M. Bickel and W. R. Van Zwet, "Asymptotic expansion for the power of distribution-free tests in the two-sample problem," Ann. Stat., vol. 6, pp. 987–1004, 1170–1171, 1987

[3] A. Blanco-Fernández, M. R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M. A. Gil, G. González-Rodríguez, M. T. López, M. A. Lubiano, M. Montenegro, A. B. Ramos-Guajardo, S. de la Rosa de Sáa and B. Sinova, "A distance-based statistic analysis of fuzzy number-valued data" (with Rejoinder), Int. J. Approx. Reason., vol. 55, pp. 1487-1501, 1601-1605, 2014.

[4] I. Couso and D. Dubois, "Statistical reasoning with set-valued information: Ontic vs. epistemic views," Int. J. Approx. Reason., vol. 55, pp. 1502-1518, 2014.

[5] P. Diamond and P. Kloeden, "Metric spaces of fuzzy sets," Fuzzy Sets Syst., vol. 100, pp. 63-71, 1999.

[6] M. A. Gil, M. A. Lubiano, M. Montenegro and M. T. López, "Least squares fitting of an affine function and strength of association for interval-valued data," Metrika, vol. 56, pp. 97-111, 2002.

[7] D. Gonzalez de Llano, M. Ramos, A. Rodríguez, A. Montill and M. Juárez, "Microbiological and physicochemical characteristics of gamonedo blue cheese during ripening," Int. Dairy J., vol. 2, pp. 121-135, 1992.

[8] G. González-Rodríguez, M. Montenegro, A. Colubi and M. A. Gil, "Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data," Fuzzy Sets Syst., vol. 157, pp. 2608-2613, 2006.

[9] G. González-Rodríguez, A. Colubi and M. A. Gil, "Fuzzy data treated as functional data. A one-way ANOVA test approach," Comput. Stat. Data An., vol. 56, pp. 943-955, 2012.

[10] P. Good, "Permutation, Parametric and Bootstrap Tests of Hypotheses," Springer, 2005.

[11] P. Grzegorzewski, "Metrics and orders in space of fuzzy numbers," Fuzzy Sets Syst., vol. 97, pp. 83-94, 1998.

[12] P. Grzegorzewski, "Statical inference about the median from vague data," Control Cybern., vol. 27, pp. 447-464, 1998.

[13] P. Grzegorzewski, "Testing statistical hypotheses with vague data," Fuzzy Sets Syst., vol. 112, pp. 501-510, 2000.

[14] P. Grzegorzewski, "Fuzzy tests - defuzzification and randomization," Fuzzy Sets Syst., vol. 118, pp. 437-446, 2001.

[15] P. Grzegorzewski, "Two-sample median test for vague data," in: Proc. 4th Conference of the European Society for Fuzzy Logic and Technology - Eusflat 2005, Barcelona, 2005, pp. 621-626.

[16] P. Grzegorzewski, "Trapezoidal approximations of fuzzy numbers preserving the expected interval - algorithms and properties," Fuzzy Sets Syst., vol. 159, pp. 1354-1364, 2008.

[17] P. Grzegorzewski, "k-sample median test for vague data," Int. J. Intell. Syst., vol. 24, pp. 529-539, 2009.

[18] P. Grzegorzewski, "The Kolmogorov goodness-of-fit test for interval-valued data," 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1-6, DOI: 10.1109/FUZZ-IEEE.2017.8015557.

[19] P. Grzegorzewski, "The Kolmogorov–Smirnov goodness-of-fit test for interval-valued data," in: Gil E. et al. (Eds.), The Mathematics of the Uncertain, Springer 2018, pp. 615-627.

[20] P. Grzegorzewski and A. Jędrej, "Chi-square goodness-of-fit test for vague data," in: Proc. 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems - IPMU 2006, Paris, 2006, pp. 952-956.

[21] P. Grzegorzewski and E. Mrówka, "Trapezoidal approximations of fuzzy numbers," Fuzzy Sets Syst., vol. 153, pp. 115-135, 2005.

[22] P. Grzegorzewski and H. Szymanowski, "Goodness-of-fit tests for fuzzy data," Inf. Sci., vol. 288, pp. 374-386, 2014.

[23] G. Hesamian and S. M. Taheri, "Fuzzy empirical distribution function: properties and application," Kybernetika, vol. 49, pp. 962-982, 2013.

[24] M. Jimenez and J. A. Rivas, "Fuzzy number approximation," Int. J. Uncertain. Fuzz., vol. 6, pp. 68-78, 1998.

[25] M. A. Lubiano, M. A. Gil, M. López-Díaz and M. T. López-García, "λ-mean squared dispersion associated with a fuzzy random variable," Fuzzy Sets Syst., vol. 111, pp. 307-317, 2000.

[26] M. A. Lubiano, M. Montenegro, B. Sinova, S. de la Rosa de Sáa and M. A. Gil, "Hypothesis testing for means in connection with fuzzy rating scale-based data: algorithms and applications," Europ. J. Oper. Res., vol. 251, pp. 918-929, 2016.

[27] M. A. Lubiano, A. Salas, C. Carleos, S. de la Rosa de Sáa and M. A. Gil, "Hypothesis testing-based comparative analysis between rating scales for intrinsically imprecise data," Int. J. Approx. Reason., vol. 88, pp. 128-147, 2017.

[28] M. Montenegro, A. Colubi, M. R. Casals and M. A. Gil, "Asymptotic and Bootstrap techniques for testing the expected value of a fuzzy random variable," Metrika, vol. 59, pp. 31-49, 2004.

[29] W. Pedrycz, "Why triangular membership functions?," Fuzzy Sets Syst., vol. 64, pp. 21-30, 1994.

[30] F. Pesarin, "Multivariate Permutation Tests," Wiley, 2001.

[31] M. L. Puri and D. A. Ralescu, "Fuzzy random variables," J. Math. Anal. Appl., vol. 114, pp. 409-422, 1986.

[32] A. B. Ramos-Guajardo, A. Blanco-Fernández and G. González-Rodríguez, "Applying statistical methods with imprecise data to quality control in cheese manufacturing," in: P. Grzegorzewski, A. Kochanski and J. Kacprzyk J. (Eds.), Soft Modeling in Industrial Manufacturing, Springer 2019, pp. 127-147.

[33] A. B. Ramos-Guajardo and M. A. Lubiano M.A., "K-sample tests for equality of variances of random fuzzy sets," Comput. Stat. Data An., vol. 56, pp. 956-966, 2012.

[34] B. Sinova, M. A. Gil, A. Colubi and S. Van Aelst, "The median of a random fuzzy number. The 1-norm distance approach," Fuzzy Sets Syst., vol. 200, pp. 99-115, 2012.

[35] W. Trutschnig, G. González-Rodríguez, A. Colubi and M. A. Gil, "A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread," Inf. Sci., vol. 179, pp. 3964-3972, 2009.