

# Linear Rank Tests for Two-Sample Fuzzy Data: A P-Value Approach

G. Hesamian<sup>1,3,\*</sup>, S.M. Taheri<sup>2,3</sup>

<sup>1</sup>*Department of Statistics, Payame Noor University, 19395-3697, Tehran, Iran*

<sup>2</sup>*Department of Basic Engineering Science, College of Engineering, University of Tehran, Tehran, Iran*

<sup>3</sup>*Department of Mathematical Sciences, Isfahan University of Technology, Isfahan 84156-83111, Iran*

Received 15 December 2011; Revised 9 December 2012

## Abstract

The linear rank test for two-sample data, which is a popular non-parametric statistical test, is extended to the case when the available data are imprecise rather than crisp. To do this, using some elements of credibility theory, we suggest a ranking method among imprecise observations and apply it to extend the usual concept of p-value leading a degree to accept or reject the null hypothesis of interest. Some applied examples, in psychology and lifetime testing, are provided to clarify the proposed approach.

©2013 World Academic Press, UK. All rights reserved.

**Keywords:** linear rank test, fuzzy variable, credibility measure, interval p-value, psychology, lifetime

## 1 Introduction

Nonparametric statistical tests have advantages over their parametric counterparts because they have fewer underlying assumptions. Two-sample nonparametric tests are usual powerful tests in detecting population differences when underlying probability distributions are not completely assumed. In fact, two-sample tests are the most useful nonparametric methods for comparing two populations, as they are sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two populations. Such tests are commonly based on crisp (exact/nonfuzzy) observations. But, in the real world, there are many situations in which, due to some practical limitations and/or human judgments, the available data are fuzzy rather than crisp (precise). In such cases, we are often dealing with two types of uncertainty; randomness and fuzziness. Randomness involves only uncertainties in the outcomes of an experiment; fuzziness, on the other hand, involves uncertainties in the impreciseness/vagueness of the data. For instance, a virus may be active completely over a certain period but losing in effect for some time, and finally go dead completely at a certain time. In such a case, we may report the lifetimes as imprecise quantities such as: “approximately 40 (h)”, “approximately 55 (h)”, “essentially less than 30 (h)” and the like. As another example, the lifetime of a tire cannot be measured precisely. In this case, we can just report the lifetime of the tires by some imprecise quantities such as “about 32000 miles”, “essentially less than 41000 miles”, “approximately 33000 miles”, etc.

To deal with both types of uncertainties, it is necessary to incorporate uncertain concepts into statistical techniques. By inception of fuzzy set theory, a lot of researches were concerned about the generalization of statistical procedures to the fuzzy environments. However, as authors know, there have been few studies on the topic of non-parametric procedures for the fuzzy environment. Regarding the purposes of this article, we briefly review some of the literature on this topic. Grzegorzewski [5] introduced a method for inference about the median of a population based on fuzzy random variables. He demonstrated a straightforward generalization of some classical non-parametric tests for fuzzy random variables [6]. The last work relies on the quasi-ordering based on a metric in the space of fuzzy numbers. Also, he utilized the necessity-index of strict dominance, introduced by Zadeh [24], for ranking observations of vague data. Kahraman et al. [12] proposed some algorithms for non-parametric rank-sum tests based on fuzzy random variables. Grzegorzewski [7] proposed a two-sample fuzzy median test for fuzzy random variables. In this manner, he obtained a fuzzy test showing a degree of possibility and a degree of necessity for rejecting the underlying

\*Corresponding author. Email: g.hesamian@math.iut.ac.ir (G. Hesamian).

hypothesis. Denoeux et al. [2] introduced a method for some non-parametric rank-sum based tests (for large sample cases) that relies on the definition of a fuzzy partial ordering based on the necessity-index of strict dominance between fuzzy numbers. Applying Monte Carlo approximation methods, they extended the concept of p-value and proposed a degree of possibility and a degree of necessity to evaluate the null hypothesis of interest. Hryniewicz [10] investigated the fuzzy version of the Goodman-Kruskal  $\gamma$  statistic described by ordered categorical data. Hryniewicz and Szediw [11] investigated a heuristic algorithm for the calculation of fuzzy Kendall's  $\tau$  that makes the implementation of the proposed chart applicable in statistical quality control. Grzegorzewski [9] proposed a modification of the classical one-sided upper Sign test to cope with vague data modeled by intuitionists fuzzy set for testing crisp or imprecise hypotheses. He also studied the problem of testing the equality of k-samples against the so-called "simple-tree alternative" by generalizing the two-sample fuzzy median test [8]. Using a defuzzification method to compare the observations, and based on a crisp test statistic, Lin et al. [15] considered the problem of two-sample Kolmogorov-Smirnov test for certain fuzzy data. Arefi et al. [1] extended some methods (histogram, empirical c.d.f., and kernel methods) for estimating probability distribution functions (p.d.f.) when the available data are fuzzy. Taheri and Hesamian [21] introduced a version of the Goodman-Kruskal measure of association for a two-way contingency table when the observations are crisp but the categories are described by fuzzy sets, and developed a method for testing independence in such a two-way contingency table (see also [20]). They also proposed a generalized method of the Wilcoxon signed-rank test for imprecise observations, when the given significance level is a fuzzy set [22]. In their work, the usual concepts of the test statistic and critical value are extended to the concepts of the fuzzy test statistic and fuzzy critical value. For more on statistical methods for fuzzy observations, see for example [14], [18] and [23].

The present paper aims to develop the non-parametric location and scale tests for imprecise observations based on extending the classical p-value approach. The proposed method has some novelty among them is that this method uses some concepts of credibility theory. Meanwhile, we develop the test function based on the concept of interval p-value.

This paper is organized as follows: In Section 2, we review some classical non-parametric location and scale tests. In Section 3, we recall some concepts of fuzzy variables and credibility measure. Then, using some ideas of credibility theory, we propose a method to rank imprecise observations. In Section 4, based on the proposed ranking method, we introduce a method to extend the linear rank statistics for imprecise observations. Finally, using an extension of the concept of p-value, we propose a procedure to test the hypothesis of interest based on a degree of acceptance and a degree of rejection. To explain the proposed method, some numerical examples are provided in Section 5. Section 6 concludes the paper.

## 2 Linear Rank Tests: A Brief Review

Suppose that two independent random samples  $X_1, \dots, X_n$  and  $Y_1, \dots, Y_m$  are drawn from populations with the continuous cumulative distribution functions  $F_X$  and  $F_Y$ , respectively. Many statistical tests applicable to the two-sample problem are based on the rank-sum statistics for the combined samples of size  $N = n + m$ . These tests can be classified as linear combinations of certain indicator variables for the combined ordered sample of observations  $W_1, W_2, \dots, W_N$ , where  $W_j = 1$  if the  $j$ th random variable in the combined ordered sample is an  $X$  and  $W_j = 0$  if it is a  $Y$ , for  $j = 1, 2, \dots, N$ . Such functions, often called linear rank statistics, are defined as follows

$$T_N = \sum_{j=1}^N a[j]W_j, \quad (2.1)$$

where the  $a[j]$  are given constants called weights or scores [3].

Under the null hypothesis  $H_0 : F_X = F_Y$ , we have for all  $j = 1, 2, \dots, N$ ,

$$E(T_N) = \frac{m}{N} \sum_{j=1}^N a[j],$$

$$Var(T_N) = \frac{mn}{N^2(N-1)} \left( N \sum_{j=1}^N a^2[j] - \left( \sum_{j=1}^N a[j] \right)^2 \right).$$

In this case, the distribution of  $T_N$  is symmetric about its mean if  $a[j] = a[N - j + 1]$ ,  $j = 1, 2, \dots, N$  (specially when  $m = n = N/2$ ). The distribution of the standardized linear rank statistic  $(T_N - E(T_N))/\sqrt{Var(T_N)}$  approaches to the standard normal distribution subject to certain regularity conditions [3, 4].

In testing the null hypothesis of identical distributions, there are generally two cases for the alternative hypothesis, location problem and scale problem, which are explained in below (for more details, see [3]).

### 2.1 Location Problem

In this case, the alternative hypothesis is that the populations are of the same form but with a different measure of central tendency. This can be expressed symbolically as follows,

$$\begin{cases} H_0 : F_Y(x) = F_X(x) \\ H_1 : F_Y(x) = F_X(x - \theta), \quad \theta \neq 0. \end{cases}$$

Note that, the random variable  $Y$  is stochastically larger than  $X$  when  $\theta > 0$ , and  $Y$  is stochastically smaller than  $X$  when  $\theta < 0$ . Thus, for example, when  $\theta < 0$ , the median of  $X$  ( $M_X$ ) is larger than the median of  $Y$  ( $M_Y$ ).

Some well-known linear rank statistics used in the two-sample location problem are given in Table 2.1. In the table,  $\Phi(x)$  and  $\zeta_j$  denote the cumulative standard normal distribution and the  $j$ th-order statistic from a standard normal population, respectively.

Table 2.1: Some well-known test statistics for two-sample location problem

Test Statistic	a[j]
Wilcoxon: $W_N$	$j$
Van-Der Vaerden: $V_N$	$\Phi^{-1}(\frac{j}{N+1})$
Terry-Hoeffing: $(TH)_N$	$E(\zeta_j)$

For example, the Wilcoxon’s large sample test statistic is given by

$$Z = \frac{W_N - \frac{mn}{2}}{\sqrt{\frac{mn(N+1)}{12}}},$$

whose distribution is approximately standard normal [3, 4]. At a given significance level  $\delta$ , we reject  $H_0$  if the observed  $p - value$  is less than  $\delta$ , and otherwise, we accept it.

### 2.2 Scale Problem

Assume that, we are interested in detecting differences in variability between two populations. This can be expressed symbolically as follows,

$$\begin{cases} H_0 : F_{Y-M}(x) = F_{X-M}(x) \\ H_1 : F_{Y-M}(x) = F_{X-M}(\theta x), \quad \theta > 0, \theta \neq 1, \end{cases}$$

where  $M$  is interpreted to be the median and  $\theta = \sigma_X/\sigma_Y$ . The alternative hypothesis  $H_1$  appropriately called the scale alternative because the cumulative distribution function of the  $Y$  is the same as that of the  $X$  but with a compressed or enlarged scale according to  $\theta > 1$  or  $\theta < 1$ , respectively. Some well-known linear rank statistics used in the two-sample scale problem are given in Table 2.2 [3, 4].

Similar to the location problem, at a given significance level  $\delta$ , we reject  $H_0$  if  $p - value < \delta$ , and otherwise, we accept it.

In Section 4, using some ideas from credibility theory, we will extend the linear rank tests for the location and scale problems to the case when the available data are provided as fuzzy observations.

Table 2.2: Some well-known test statistics for two-sample scale problem

Test Statistic	a[j]
Mood: $M_N$	$(j - (N + 1)/2)^2$
Ansari-Bradly: $A_N$	$ j - (N + 1)/2 $
Klotz Normal-Scores: $K_N$	$(\Phi^{-1}(j/(N + 1)))^2$
Siegel-Tukey: $S_N$	$\begin{cases} 2j & j \text{ even, } 1 < j \leq N/2, \\ 2j - 1 & j \text{ odd, } 1 \leq j \leq N/2, \\ 2(N - j) + 2 & j \text{ even, } N/2 < j \leq N, \\ 2(N - j) + 1 & j \text{ odd, } N/2 < j \leq N. \end{cases}$

### 3 Credibility Measure: Some Elementary Definitions and Results

Credibility theory, founded by Liu [17], is a branch of mathematics for studying the behavior of uncertain phenomena. Since we will use some elements of credibility theory in our proposed procedure, in this section we present some main concepts and results of this topic.

**Definition 1** [17] Let  $\Omega$  be a nonempty set, and  $\mathcal{A}$  the power set of  $\Omega$ . Generally, each element in  $\mathcal{A}$  is called an event. A set function  $Cr : \mathcal{A} \rightarrow [0, 1]$  is called a credibility measure on  $(\Omega, \mathcal{A})$  if it satisfies the following four axioms:

1. Axiom 1. (Normality)  $Cr\{\Omega\} = 1$ .
2. Axiom 2. (Monotonicity)  $Cr\{A\} \leq Cr\{B\}$  whenever  $A \subseteq B$ .
3. Axiom 3. (Self-Duality)  $Cr\{A\} + Cr\{A^c\} = 1$  for any event  $A$ .
4. Axiom 4. (Maximality)  $Cr\{\cup_i A_i\} = \sup_i Cr\{A_i\}$  for any events  $\{A_i\}$  with  $\sup_i Cr\{A_i\} < 0.5$ .

**Definition 2** [17] A fuzzy variable is a measurable function from a credibility space  $(\Omega, \mathcal{A}, Cr)$  to the set of real numbers. For a fuzzy variable  $\tilde{A}$ , its membership function is defined by

$$\mu_{\tilde{A}}(x) = (2Cr\{\tilde{A} \in \{x\}\}) \wedge 1, x \in \mathcal{R},$$

where the symbol  $\wedge$  stands for the minimum.

In the following, by  $\{\tilde{A} \in C\}$  we mean the set  $\{w \in \Omega : \tilde{A}(w) \in C\}$ .

A fuzzy variable  $\tilde{A}$  with  $\sup \mu_{\tilde{A}}(x) = 1$  is called a normal fuzzy variable. By a triangular fuzzy variable we mean the fuzzy variable fully determined by the triplet  $(a_L, a, a_U)$  (briefly,  $\tilde{A} = (a_L, a, a_U)_T$ ) of crisp numbers with  $a_L < a < a_U$ , whose membership function is given by

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-a_L}{a-a_L} & \text{if } a_L \leq x < a \\ \frac{a_U-x}{a_U-a} & \text{if } a \leq x \leq a_U \\ 0 & \text{if } x < a_L \text{ or } x > a_U. \end{cases}$$

**Example 1:** ([17], p.180) Let  $\tilde{A}$  be a normal fuzzy variable and  $r$  be a real number. Then, the credibility of  $\{\tilde{A} \in (-\infty, r]\}$  is

$$Cr\{\tilde{A} \in (-\infty, r]\} = \frac{1}{2}(\sup_{x \leq r} \mu_{\tilde{A}}(x) + 1 - \sup_{x > r} \mu_{\tilde{A}}(x)). \quad (3.1)$$

**Example 2:** Suppose that  $\tilde{A} = (a_L, a, a_U)_T$  is a triangular fuzzy variable. For any real number  $r$ , we can calculate  $Cr\{\tilde{A} \in (-\infty, r]\}$  as follows

$$Cr\{\tilde{A} \in (-\infty, r]\} = \begin{cases} 0 & \text{if } r < a_L \\ \frac{r-a_L}{2(a-a_L)} & \text{if } a_L \leq r \leq a \\ \frac{a_U-2a+r}{2(a_U-a)} & \text{if } a < r \leq a_U \\ 1 & \text{if } r > a_U. \end{cases}$$

For example, let  $\tilde{A} = (-2, 0, 1)_T$ , then

$$Cr\{\tilde{A} \in (-\infty, r]\} = \begin{cases} \frac{r+2}{4} & \text{if } -2 \leq r \leq 0 \\ \frac{1+r}{2} & \text{if } 0 < r \leq 1 \\ 0 & \text{elsewhere.} \end{cases}$$

To perform the non-parametric two-sample tests for location and scale problems based on imprecise observations, we need a suitable method of ranking fuzzy variables. Here, we introduce a method for ranking fuzzy variables, which will be used in this article.

**Definition 3** Let  $\tilde{A}$  and  $\tilde{B}$  be two normal fuzzy variables, and  $\alpha \in (0, 1]$  be a prescribed real number. Then, at the level of  $\alpha$ ,  $\tilde{B}$  is said to be smaller than  $\tilde{A}$ , denoted by  $\tilde{B} <_\alpha \tilde{A}$ , if  $\tilde{B}_\alpha < \tilde{A}_\alpha$ , where,  $\tilde{B}_\alpha = \inf\{r : Cr\{\tilde{B} \in (-\infty, r] \geq \alpha\}$  and  $\tilde{A}_\alpha = \inf\{r : Cr\{\tilde{A} \in (-\infty, r] \geq \alpha\}$ . Also, at the level of  $\alpha$ ,  $\tilde{B}$  is said to be equal to  $\tilde{A}$ , denoted by  $\tilde{B} =_\alpha \tilde{A}$ , if  $\tilde{B}_\alpha = \tilde{A}_\alpha$ . Finally, at the level of  $\alpha$ ,  $\tilde{B}$  is said to be smaller than or equal to  $\tilde{A}$ , denoted by  $\tilde{B} \leq_\alpha \tilde{A}$ , if and only if  $\tilde{B} <_\alpha \tilde{A}$  or  $\tilde{B} =_\alpha \tilde{A}$ .

Note that  $\tilde{A}_\alpha$  is called  $\alpha$ -optimistic value of  $\tilde{A}$  and  $\tilde{A}^\alpha = \sup\{r : Cr\{\tilde{B} \in [r, \infty) \geq \alpha\}$  is called  $\alpha$ -pessimistic value of  $\tilde{B}$  [19].

**Example 3:** Suppose that  $\tilde{A} = (a_L, a, a_U)_T$  is a triangular fuzzy variable. Then, from Example 3 we have

$$\tilde{A}_\alpha = \begin{cases} a_L + 2\alpha(a - a_L) & \text{for } 0 \leq \alpha \leq 0.5 \\ 2\alpha(a_U - a) - (a_U - 2a) & \text{for } 0.5 < \alpha \leq 1. \end{cases}$$

For instance, if  $\tilde{B} = (0, 2, 6)_T$  and  $\tilde{A} = (1, 4, 5)_T$ , then  $\tilde{B} \leq_\alpha \tilde{A}$  for any  $\alpha \in [0, 0.833]$ .

The proposed ranking method has some properties, which are given in the following theorem.

**Theorem 1** Let  $\tilde{A}$ ,  $\tilde{B}$  and  $\tilde{C}$  be some normal fuzzy variables. At any level of  $\alpha \in (0, 1]$ ,

1. Reflexivity:  $\tilde{A} \leq_\alpha \tilde{A}$ .
2. Symmetry:  $\tilde{A} \leq_\alpha \tilde{B}$  and  $\tilde{B} \leq_\alpha \tilde{A}$ , implies  $\tilde{B} =_\alpha \tilde{A}$ .
3. Transitivity:  $\tilde{A} \leq_\alpha \tilde{B}$  and  $\tilde{B} \leq_\alpha \tilde{C}$ , imply  $\tilde{A} \leq_\alpha \tilde{C}$ .
4. For two arbitrary fuzzy variables  $\tilde{A}$  and  $\tilde{B}$ ,  $\tilde{A} \leq_\alpha \tilde{B}$  or  $\tilde{A} >_\alpha \tilde{B}$ .

**Proof.** (1) and (2) follow immediately from Definition 3.

(3) Since  $\tilde{A} \leq_\alpha \tilde{B}$  and  $\tilde{B} \leq_\alpha \tilde{C}$ , we have four cases as 1)  $\tilde{A}_\alpha < \tilde{B}_\alpha$  and  $\tilde{B}_\alpha < \tilde{C}_\alpha$ , 2)  $\tilde{A}_\alpha < \tilde{B}_\alpha$  and  $\tilde{B}_\alpha = \tilde{C}_\alpha$ , 3)  $\tilde{A}_\alpha = \tilde{B}_\alpha$  and  $\tilde{B}_\alpha < \tilde{C}_\alpha$ , 4)  $\tilde{A}_\alpha = \tilde{B}_\alpha$  and  $\tilde{B}_\alpha = \tilde{C}_\alpha$ . For each case, we have  $\tilde{A}_\alpha < \tilde{C}_\alpha$  or  $\tilde{A}_\alpha = \tilde{C}_\alpha$ . Thus, the transitivity property is verified.

(4) Since  $\tilde{A}_\alpha \leq \tilde{B}_\alpha$  or  $\tilde{A}_\alpha > \tilde{B}_\alpha$ , the result follows immediately. □

## 4 Linear Rank Tests with Imprecise Observations

In this section, we extend the linear rank tests to examine the hypotheses about the differences in location or variability between two populations to the case when the observations are imprecise rather than crisp.

As we mentioned in Section 2, in the classical approach to test the null hypothesis  $H_0 : F_Y = F_X$ , the observed linear rank statistic is given by

$$T_N = \sum_{j=1}^N a[j]w_j. \tag{4.1}$$

Now, consider the problem of non-parametric location and scale tests based on the imprecise observations  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$  and  $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m$ . In this situation, to verify the null hypothesis  $H_0$ , we have to combine the imprecise observations as  $\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_N$ . Since the expression “less than or equal to” is a vague concept in fuzzy environment, so we apply the ranking method proposed in Section 3 for realizing whether  $\tilde{w}_j$  could be regarded as “less than or equal to”  $\tilde{w}_i$ . For a given  $\alpha \in (0, 1]$ , assume  $\tilde{w}_{\alpha 1}, \tilde{w}_{\alpha 2}, \dots, \tilde{w}_{\alpha N}$  denotes the ordered combined samples  $\tilde{x}_{\alpha 1}, \tilde{x}_{\alpha 2}, \dots, \tilde{x}_{\alpha n}$  and  $\tilde{y}_{\alpha 1}, \tilde{y}_{\alpha 2}, \dots, \tilde{y}_{\alpha m}$ , where  $\tilde{x}_{\alpha i}$  and  $\tilde{y}_{\alpha j}$  are introduced in Definition 3 and  $\tilde{w}_{\alpha j} = 1$  if the  $j$ th element in the combined ordered sample is an  $x$  and  $\tilde{w}_{\alpha j} = 0$  if it is a  $y$ , for  $j = 1, 2, \dots, N$ . However, since for every  $\alpha \in (0, 1]$ , the value

$$T_{N\alpha} = \sum_{j=1}^N a[j] \tilde{w}_{\alpha j}$$

is a candidate to be a test statistic, therefore, we have a set of linear rank test statistics as

$$\tilde{T}_N = \{\tilde{T}_N^L, \tilde{T}_N^L + 1, \dots, \tilde{T}_N^U\}, \quad (4.2)$$

in which

$$\tilde{T}_N^L = \inf_{\alpha \in (0,1]} \sum_{j=1}^N a[j] \tilde{w}_{\alpha j}, \quad \tilde{T}_N^U = \sup_{\alpha \in (0,1]} \sum_{j=1}^N a[j] \tilde{w}_{\alpha j}. \quad (4.3)$$

Since the extended test statistic is a set, hence to construct a procedure for testing the interested hypothesis, we define an extended p-value as an interval in the following way.

**Definition 4** In testing the null hypothesis  $H_0 : F_Y = F_X$  based on imprecise observations, the p-value is defined by an interval  $\tilde{p}$  – value =  $[\tilde{p}^L, \tilde{p}^U]$  as follows

- for the case of alternative hypothesis  $H_1 : \theta < 0$ ,

$$\tilde{p}^L = \inf_{w \in \tilde{T}_N} \mathbf{P}_{H_0}(T_N \geq w) = \mathbf{P}_{H_0}(T_N \geq \tilde{T}_N^U), \quad (4.4)$$

$$\tilde{p}^U = \sup_{w \in \tilde{T}_N} \mathbf{P}_{H_0}(T_N \geq w) = \mathbf{P}_{H_0}(T_N \geq \tilde{T}_N^L), \quad (4.5)$$

- for the case of alternative hypothesis  $H_1 : \theta > 0$ ,

$$\tilde{p}^L = \inf_{w \in \tilde{T}_N} \mathbf{P}_{H_0}(T_N \leq w) = \mathbf{P}_{H_0}(T_N \leq \tilde{T}_N^L), \quad (4.6)$$

$$\tilde{p}^U = \sup_{w \in \tilde{T}_N} \mathbf{P}_{H_0}(T_N \leq w) = \mathbf{P}_{H_0}(T_N \leq \tilde{T}_N^U), \quad (4.7)$$

- for the case of alternative hypothesis  $H_1 : \theta \neq 0$ ,

$$\tilde{p}^L = \min_{t \in \tilde{T}_N} \left\{ \min \{2 \min \{\mathbf{P}_{H_0}(T_N \leq t), \mathbf{P}_{H_0}(T_N \geq t)\}, 1\} \right\}, \quad (4.8)$$

$$\tilde{p}^U = \max_{t \in \tilde{T}_N} \left\{ \min \{2 \min \{\mathbf{P}_{H_0}(T_N \leq t), \mathbf{P}_{H_0}(T_N \geq t)\}, 1\} \right\}. \quad (4.9)$$

Now, by inception the idea given by Grzegorzewski [9], decision making to accept or reject the null hypothesis is provided based on the following test function.

**Definition 5** Consider the problem of linear rank test with imprecise observations at a given significance level  $\delta \in (0, 1]$ . Then, the test function is defined by

$$\phi_\delta(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n, \tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m) = \begin{cases} 0 & \text{if } \delta \leq \tilde{p}^L \\ \frac{\delta - \tilde{p}^L}{\tilde{p}^U - \tilde{p}^L} & \text{if } \tilde{p}^L < \delta \leq \tilde{p}^U \\ 1 & \text{if } \delta > \tilde{p}^U. \end{cases} \quad (4.10)$$

In which, test function,  $\phi_\delta(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n, \tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m)$  is interpreted as “degree of rejection” of  $H_0$  and  $1 - \phi_\delta(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n, \tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m)$  is interpreted as “degree of acceptance” of  $H_0$ .

Table 5.1: Data set in Example 5

Supplier A	Supplier B
(42 , 46 , 53 ) <sub>T</sub>	(30 , 35 , 42 ) <sub>T</sub>
(53 , 56 , 62 ) <sub>T</sub>	(56 , 66 , 73 ) <sub>T</sub>
(57 , 60 , 68 ) <sub>T</sub>	(48 , 58 , 63 ) <sub>T</sub>
(42 , 49 , 57 ) <sub>T</sub>	(77 , 83 , 88 ) <sub>T</sub>
-	(65 , 71 , 76 ) <sub>T</sub>

Based on the proposed method in this paper, therefore, the steps of linear rank test for imprecise observations can be summarized as follows:

1. Determine the fuzzy test statistic, using Equation (4.2).
2. Determine the interval p-value  $\tilde{p} - value = [\tilde{p}^L, \tilde{p}^U]$  for the hypotheses of interest, using Definition 4.
3. At the given significance level of  $\delta$ :
  - if  $\delta \leq \tilde{p}^L$ , then accept  $H_0$  with degree of one;
  - if  $\tilde{p}^U < \delta$ , then reject  $H_0$  with degree of one;
  - if  $\tilde{p}^L < \delta \leq \tilde{p}^U$ , then reject  $H_0$  with degree of  $\phi_\delta = \frac{\delta - \tilde{p}^L}{\tilde{p}^U - \tilde{p}^L}$  and accept it with degree of  $1 - \phi_\delta$ .

**Remark 1:** If the imprecise observations reduce to the crisp values, then  $Cr\{\tilde{w}_j \in (-\infty, r]\} = \frac{1}{2}(I(w_j \leq r) + 1 - I(w_j > r))$ , and so for any  $\alpha \in (0, 1]$  we have  $\tilde{w}_{\alpha j} = \inf\{r : Cr\{\tilde{w}_{\alpha j} \in (-\infty, r]\} \geq \alpha\} = w_j$ . Therefore,  $\tilde{T}_N^L = \tilde{T}_N^U = \sum_{j=1}^N a[j]w_j = T_N$ , and then  $\tilde{p} - value = p - value$ . So, the proposed test is reduced to the classical non-parametric location and scale tests (Section 2).

**Remark 2:** It should be mentioned that, using a concept of fuzzy test statistic, Grzegorzewski [7, 8] investigated median tests for vague data. He utilized the necessity-index introduced by Zadeh [24] for ranking fuzzy data. At a crisp significance level, he constructed a fuzzy test based on the classical critical region, in which the result of the test is presented by two possibility-necessity-based indices. In addition, Kahraman et al. [12] considered the problem of some non-parametric tests for fuzzy observations. Based on some fuzzy ranking methods and some representative values, they obtained a crisp test statistic to evaluate the hypotheses of interest. In addition, they constructed a fuzzy non-parametric test for which the result of the test is given by a 0-1 decision. In contrast, our method lead to a degree between zero and one, to evaluate the null hypothesis of interest.

Grzegorzewski also proposed a modification of the classical one-sided upper Sign test to cope with vague data modeled by intuitionists fuzzy set for testing crisp or imprecise hypotheses [9]. Based on his approach, the output of the Sign test statistic is an interval  $\tilde{T}_N = [T^L, T^U]$ . As a counterpart of the traditional p-value, he considered an interval p-value as  $\tilde{p} = [\tilde{p}^L, \tilde{p}^U]$ . At a significance level  $\delta$ , he made the following decision rules:

- if  $\tilde{p}^U \leq \delta$ , then we reject  $H_0$ ;
- if  $\tilde{p}^L > \delta$ , then we accept  $H_0$ ;
- if  $\tilde{p}^L \leq \delta < \tilde{p}^U$ , then the test is not decisive (see [9]).

In contrast, in our proposed test, using a new ranking method and defining a test statistic  $\tilde{T}_N = \{\tilde{T}_N^L, \dots, \tilde{T}_N^U\}$ , we suggested a different method to that of Grzegorzewski's method. Moreover, for making decision to reject or accept a given hypothesis, we also proposed an extended version of Grzegorzewski's decision rule by introducing the concepts of degree of acceptance and degree of rejection.

## 5 Numerical Examples

To clarify our proposed method, some numerical examples are provided in this section.

**Example 4:** ([3], p.348) Two potential suppliers of street lighting equipment, *A* and *B*, want to present their bids to a city manager. Two independent random samples of sizes 5 and 4 street lighting equipments

Table 5.2: Data set in Example 5

Married	Single
(51, 57, 60) <sub>T</sub>	(55, 58, 58) <sub>T</sub>
(38, 44, 47) <sub>T</sub>	(45, 45, 45) <sub>T</sub>
(59, 60, 60) <sub>T</sub>	(45, 50, 55) <sub>T</sub>
(40, 44, 49) <sub>T</sub>	(47, 54, 57) <sub>T</sub>
(34, 39, 43) <sub>T</sub>	(42, 49, 56) <sub>T</sub>
(60, 60, 60) <sub>T</sub>	(32, 39, 45) <sub>T</sub>
(35, 44, 50) <sub>T</sub>	(50, 50, 50) <sub>T</sub>
-	(42, 51, 58) <sub>T</sub>

were tested from the suppliers. Since, under some unexpected situations, we cannot measure the lifetimes, precisely, we can just obtain them around a number. The lifetimes are reported to be triangular fuzzy variables as shown in Table 5.1. At the significance level  $\delta = 0.05$  we wish to test  $H_0 : M_A = M_B$  v.s.  $H_1 : M_A \neq M_B$ .

One of the most commonly used tests for the location problem is the Wilcoxon test. From Equation (4.3), calculations show that

$$\widetilde{W}_N^L = \inf_{\alpha \in (0,1]} \sum_{j=1}^9 j \widetilde{w}_{\alpha j} = 15, \quad \widetilde{W}_N^U = \sup_{\alpha \in (0,1]} \sum_{j=1}^9 j \widetilde{w}_{\alpha j} = 18,$$

and so  $\widetilde{W}_N = \{15, 16, 17, 18\}$ . Based on [3], Table J, pp. 585, and using Equations (4.8) and (4.9), we obtain,  $\widetilde{p}\text{-value} = [0.286, 0.730]$ . From Equation (4.10), since  $\widetilde{p}^L \geq 0.05$ , we conclude that, with degree of acceptance equal to one, there is no difference in the locations of the populations A and B.

**Example 5:** Consider the previous example. Assume that, we wish to test the null hypothesis whether the lifetime of suppliers A and B have equal variability or not (i.e.  $H_0 : \sigma_A = \sigma_B$  v.s.  $H_1 : \sigma_A \neq \sigma_B$ ), at significance level 0.05. For such a case, the most frequently used test statistic is Siegel-Tukey statistic (Table 2.2). Using Equation (4.3), we obtained  $\widetilde{S}_N = \{23, 24, 25, 26\}$ . We also obtained  $\widetilde{p}\text{-value} = [0.180, 0.256]$ . Now, since  $\widetilde{p}^L \geq 0.05$ , therefore, at the significance level  $\delta = 0.05$ , we conclude that, with degree of one, there is no basis for difference between variability suppliers A and B.

**Example 6:** ([13], p.221) A psychologist studies the degree of happiness of people at various stages in life. His measure of general happiness varies from 0 to 60. In a certain study he compared the happiness of married and single men aged 25. Each response is reported by a triangular fuzzy number as shown in Table 5.2. We wish to test if singles are happier than marries (i.e.,  $H_0 : M_S = M_M$  v.s.  $H_1 : M_S > M_M$ ).

From Equation (4.2), we obtained  $\widetilde{W}_N = \{55, 56, 57, 58, 59, 60, 61, 62\}$  and therefore, using Equations (4.4-4.5), we obtained  $\widetilde{p}\text{-value} = [0.268, 0.567]$ . At the significance level of  $\delta = 0.10$ , since  $\widetilde{p}^L \geq 0.10$ , we conclude that with degree of one there is no difference between the two groups.

## 6 Conclusion

In the present paper, we proposed a generalization of the non-parametric two-sample tests for imprecise observations. We considered mainly location and scale tests that utilized a proposed ranking method based on some preliminary concepts of the credibility measure. To do this, the usual concept of p-value is extended to an interval p-value. Finally, we introduce the concepts of “degree of acceptance” and “degree of rejection”, to evaluate the hypothesis of interest.

The study of the power of the proposed test and developing the linear rank test for fuzzy hypotheses are some potential subjects for further research.

## Acknowledgements

The authors would like to thank referees for their useful comments and suggestions. Part of this work were completed during the first author’s stay in the Department of Mathematical Sciences at Tsinghua University, China. He is grateful for the hospitality of this department.



## References

- [1] Arefi, M., R. Viertl, and S.M. Taheri, Fuzzy density estimation, *Metrika*, vol.75, no.1, pp.5–22, 2012.
- [2] Denoeux, T., M.H. Masson, and P.H. Herbert, Non-parametric rank-based statistics and significance tests for fuzzy data, *Fuzzy Sets and Systems*, vol.153, no.1, pp.1–28, 2005.
- [3] Gibbons, J.D., and S. Chakraborti, *Non-parametric Statistical Inference*, 4th Edition, Marcel Dekker, New York, 2003.
- [4] Govindarajulu, Z., *Non-parametric Inference*, 2nd Edition, World Scientific, Hackensack, 2007.
- [5] Grzegorzewski, P., Statistical inference about the median from vague data, *Control and Cybernetics*, vol.27, no.3, pp.447–464, 1998.
- [6] Grzegorzewski, P., Distribution-free tests for vague data, *Soft Methodology and Random Information Systems*, Springer, Heidelberg, pp.495–502, 2004.
- [7] Grzegorzewski, P., Two-sample median test for vague data, *Proceedings of the 4th Conference of the European Society for Fuzzy Logic and Technology-Eusflat*, Barcelona, pp.621–626, 2005.
- [8] Grzegorzewski, P., K-sample median test for vague data, *International Journal of Intelligent Systems*, vol.24, no.5, pp.529–539, 2009.
- [9] Grzegorzewski, P., A Bi-Robust test for vague data, *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Torremolinos (Malaga), Spain, pp.138–144, 2008.
- [10] Hryniewicz, O., Goodman-Kruskal  $\gamma$  measure of dependence for fuzzy ordered categorical data, *Computational Statistics and Data Analysis*, vol.51, no.1, pp.323–334, 2006.
- [11] Hryniewicz, O., and A. Szediw, Fuzzy Kendall  $\tau$  statistic for autocorrelated data, *Soft Methods for Handling Variability and Imprecision*, vol.48, pp.155–162, 2008.
- [12] Kahraman, C., C.F. Bozdog, and D. Ruan, Fuzzy sets approaches to statistical parametric and non-parametric tests, *International Journal of Intelligent Systems*, vol.19, no.11, pp.1069–1078, 2004.
- [13] Kenny, D.A., *Statistics for the Social and Behavioral Sciences*, Little Brown, Boston, 1987.
- [14] Kruse, R., and K.D. Meyer, *Statistics with Vague Data*, Reidel Publishing Company, Dordrecht, 1987.
- [15] Lin, P., B. Wu, and J. Watada, Kolmogorov-Smirnov two sample test with continuous fuzzy data, *Integrated Uncertainty Management and Applications*, vol.68, pp.175–186, 2010.
- [16] Liu, B., and Y. Liu, Expected value of fuzzy variable and fuzzy expected value models, *IEEE Transactions on Fuzzy Systems*, vol.10, no.4, pp.445–450, 2002.
- [17] Liu, B., *Uncertainty Theory*, Springer-Verlag, Berlin, 2004.
- [18] Nguyen, H., and B. Wu, *Fundamentals of Statistics with Fuzzy Data*, Springer, Netherlands, 2006.
- [19] Peng, J., and B. Liu, Some properties of optimistic and pessimistic values of fuzzy variables, *Proceedings of the 13th IEEE International Conference on Fuzzy Systems*, Budapest, Hungary, pp.745–750, 2004.
- [20] Taheri, S.M., G. Hesamian, and R. Viertl, Contingency table with fuzzy categories, *Proceedings of the 11th International Conference on Intelligent Thechnology*, Thailand, pp.100–104, 2010.
- [21] Taheri, S.M., and G. Hesamian, Goodman-Kruskal measure of association for fuzzy-categorized variables, *Kybernetika*, vol.47, no.1, pp.110–122, 2011.
- [22] Taheri, S.M., and G. Hesamian, A generalization of the Wilcoxon signed-rank test and its applications, *Statistical Papers*, vol.54, no.2, pp.457–470, 2013.
- [23] Viertl, R., *Statistical Methods for Fuzzy Data*, Wiley, Chichester, 2011.
- [24] Zadeh, L.A., A theory of approximate reasoning, *Mathematical Frontiers of the Social and Policy Sciences*, Westview Press, Boulder, pp.69–129, 1979.