On testing statistical hypothesis for mean and difference between mean based on revised signed distance under fuzzy environment

A. Hari Ganesh\(^1\)\(^*\) and N. Jaimaruthi\(^2\)

Abstract

In this paper, we deal with the problem of testing statistical hypothesis under fuzzy environment particularly for trapezoidal fuzzy data. In order to test the statistical hypothesis for trapezoidal fuzzy data, we extend the signed distance, with the suitable modification of average and width of triangular fuzzy data in the new signed distance introduced by Arefi, to trapezoidal fuzzy data.

The method introduced in this paper is also to be employed for testing hypothesis of various statistical parameters, that is, mean and variance. Finally, some numerical examples are given to demonstrate the feasibility of the revised signed distance for statistical hypothesis testing.

Keywords

Statistical hypothesis, fuzzy data, triangular & trapezoidal fuzzy numbers, testing hypothesis, signed distance.

AMS Subject Classification

03E72.

1. Introduction

In classical environments, testing statistical hypothesis depends on crisp data and crisp hypothesis. However, in realistic situations the data and the hypothesis may not be crisp. But, in it both may be fuzzy or hypothesis may be crisp and data may be fuzzy or hypothesis may be fuzzy and data may be crisp. In such situations to test the statistical hypotheses the classical environment are not give an exact result. So the problem of testing hypothesis has to be considered under fuzzy environment based on fuzzy set theory, introduced by Zadeh [14] in 1965.

Testing statistical hypotheses in fuzzy environments are done by many authors. Casals et al. [5] and Grzegorzewski [8] introduced methods for testing statistical hypothesis with fuzzy information in 1986 and 1997 respectively. In 2000, Grzegorzewski [9] introduced a new approach to test statistical hypothesis with vague data. In 2004, statistical hypotheses were tested with fuzzy data by Filzmoser and Viertl [7]. In 2011, Arefi and Taheri [3] introduced fuzzy based approach to testing statistical hypothesis. In 2012, Chachi et al. [6] used...
confidence interval for testing statistical hypothesis. Moreover, some other soft methods also were developed for testing statistical hypotheses in imprecise environment by Akbari and Rezaei [1], Buckley [4], Taheri and Hesamian [11, 12], and Yosef et al. [13].

Although, testing hypotheses based on fuzzy data are investigated by many authors. Arefi [2] introduced a new signed distance based method for testing statistical hypothesis under fuzzy data recently. In this article, we revisited the new signed distance introduced by Arefi with some changes in average and width of fuzzy numbers. Not only that, we also extend this method to trapezoidal fuzzy data with the suitable modification. As the triangular fuzzy number is a particular case of trapezoidal fuzzy number, we could test the statistical hypothesis for triangular fuzzy data based on the proposed approach.

The article is organized as follows: In Section 1, we have introduced some basic concepts of fuzzy set and numbers. In Section 2, we introduce the new revised sign distance for trapezoidal fuzzy number. In Section 3, the method is introduced to test for the single mean of a normal distribution (with known variance), in Section 4 the difference of means of two normal distributions (with known variances) with some numerical examples. Finally, in Section 5, a brief conclusion is provided based on the results obtained from the numerical example.

2. Revised Signed Distance between Fuzzy Numbers

**Definition 2.1.** The trapezoidal fuzzy number \( \tilde{A} = (x_0 - \delta, x_0, y_0 + \eta) \), with two defuzzifier \( x_0, y_0 \), and left fuzziness \( \delta > 0 \) and right fuzziness \( \eta > 0 \) is a fuzzy set where membership function is as

\[
\mu_{\tilde{A}}(x) = \begin{cases} 
\frac{1}{\delta}(x - x_0), & x_0 - \delta \leq x \leq x_0 \\
1, & x_0 \leq x \leq y_0 \\
\frac{1}{\eta}(y_0 - x + \eta), & y_0 \leq x \leq y_0 + \eta \\
0, & \text{otherwise}
\end{cases}
\]

and its parametric form is \( u(r) = x_0 - (1 - r)\delta \), \( \widetilde{u}(r) = y_0 + (1 - r)\eta \). It becomes a triangular fuzzy number if \( x_0 = y_0 \).

**Definition 2.2.** Let \( \tilde{A} = (x_0 - \delta, x_0, y_0 + \eta) \) be a trapezoidal fuzzy number with the r-cut \( \tilde{A}[r] = [u(r), \widetilde{u}(r)] \). The values of average and width of \( \tilde{A} \), respectively, are defined as follows:

\[
\begin{align*}
M(\tilde{A}) &= \frac{1}{2} \int_0^1 (u(r) + \widetilde{u}(r) + x_0 + y_0)f(r)dr \\
W(\tilde{A}) &= \frac{1}{2} \int_0^1 (\tilde{u}(r) - u(r) + y_0 - x_0)f(r)dr
\end{align*}
\]

where \( f(r) \) is nonnegative and increasing function on \([0, 1]\) with \( f(0) = 0, f(1) = 1 \) and \( \int_0^1 f(r)dr = \frac{1}{2} \). Here we consider \( f(r) = r \).

**Definition 2.3.** Let \( \tilde{A} = (x_0 - \delta, x_0, x_0 + \eta) \) be a triangular fuzzy number with the r-cut \( \tilde{A}[r] = [u(r), \widetilde{u}(r)] \). The values of average and width of \( \tilde{A} \), respectively, are defined as follows:

\[
\begin{align*}
M(\tilde{A}) &= \frac{1}{2} \int_0^1 (u(r) + \widetilde{u}(r) + 2x_0)f(r)dr \\
W(\tilde{A}) &= \frac{1}{2} \int_0^1 (\widetilde{u}(r) - u(r))f(r)dr
\end{align*}
\]

where \( f(r) \) is nonnegative and increasing function on \([0, 1]\) with \( f(0) = 0, f(1) = 1 \) and \( \int_0^1 f(r)dr = \frac{1}{2} \). Here we consider \( f(r) = r \).

**Remark 2.4.** The signed distance between two trapezoidal fuzzy numbers is calculated based on the new signed distance introduced by Arefi in [2]. The triangular fuzzy number is as a special case of trapezoidal fuzzy number. So we can obtain the triangular fuzzy number by taking \( x_0 = y_0 \). The revised signed distance between such triangular fuzzy numbers is defined as follows:

**Definition 2.5.** Let \( \tilde{A} \) and \( \tilde{B} \) be two triangular fuzzy numbers. The new signed distance between \( \tilde{A} \) and \( \tilde{B} \) introduced in [2] is revised as follows:

\[
SD(\tilde{A}, \tilde{B}) = \begin{cases} 
M(\tilde{A}) - M(\tilde{B}) + 2|W(\tilde{A}) - W(\tilde{B})|, & M(\tilde{A}) \geq M(\tilde{B}) \\
M(\tilde{A}) - M(\tilde{B}) - 2|W(\tilde{A}) - W(\tilde{B})|, & M(\tilde{A}) < M(\tilde{B}) \\
|d(\tilde{A}, \tilde{B})|, & M(\tilde{A}) \geq M(\tilde{B}) \\
-d(\tilde{A}, \tilde{B}), & M(\tilde{A}) < M(\tilde{B}) 
\end{cases}
\]

where \( d(\tilde{A}, \tilde{B}) = |M(\tilde{A}) - M(\tilde{B})| + 2|W(\tilde{A}) - W(\tilde{B})| \) satisfies the properties of metric.

**Remark 2.6.**

1. The revised signed distance between fuzzy number and crisp number is reduced as follows:

\[
SD(\tilde{A}, b) = \begin{cases} 
M(\tilde{A}) + 2|W(\tilde{A})| - b, & M(\tilde{A}) \geq b \\
M(\tilde{A}) - 2|W(\tilde{A})| - b, & M(\tilde{A}) < b 
\end{cases}
\]

2. The revised signed distance between two crisp real numbers is reduced as follows:

\[
SD(a, b) = a - b
\]
Definition 2.7. Let $\bar{A}$ and $\bar{B}$ be two triangular or trapezoidal fuzzy numbers. Then, the ranking of $\bar{A}$ and $\bar{B}$ is expressed as
\[
\begin{align*}
D(\bar{A}, \bar{B}) > 0 & \iff \bar{A} > \bar{B} \\
D(\bar{A}, \bar{B}) < 0 & \iff \bar{A} < \bar{B} \\
D(\bar{A}, \bar{B}) = 0 & \iff \bar{A} \approx \bar{B}
\end{align*}
\]

Lemma 2.8. Let $\bar{A}$, $\bar{B}$, $\bar{C}$ be two triangular fuzzy numbers. Then, the signed distance introduced in Definition 2.2 satisfies the following properties:

(i) For $M(\bar{A}) \neq M(\bar{B})$, then $SD(\bar{A}, \bar{B}) = -SD(\bar{B}, \bar{A})$, and for $M(\bar{A}) = M(\bar{B})$, then $SD(\bar{A}, \bar{B}) = SD(\bar{B}, \bar{A})$.

(ii) We order $\bar{A}$, $\bar{B}$ and $\bar{C}$ such that $M(\bar{A}) \geq M(\bar{B}) \geq M(\bar{C})$, then $SD(\bar{A}, \bar{B}) + SD(\bar{B}, \bar{C}) \geq SD(\bar{A}, \bar{C})$.

(iii) $\bar{A} \approx \bar{B} \iff \bar{B} \approx \bar{A}$

Proof. (i) For $M(\bar{A}) \neq M(\bar{B})$, we have
\[
SD(\bar{A}, \bar{B}) = \begin{cases} 
M(\bar{A}) - M(\bar{B}) + 2 |W(\bar{A}) - W(\bar{B})|, & M(\bar{A}) > M(\bar{B}) \\
M(\bar{A}) - M(\bar{B}) - 2 |W(\bar{A}) - W(\bar{B})|, & M(\bar{A}) < M(\bar{B}) \\
- [M(\bar{B}) - M(\bar{A}) - 2 |W(\bar{B}) - W(\bar{A})|], & M(\bar{A}) < M(\bar{B}) \\
- [M(\bar{B}) - M(\bar{A}) + 2 |W(\bar{B}) - W(\bar{A})|], & M(\bar{A}) > M(\bar{B}) \\
\end{cases}
= -SD(\bar{B}, \bar{A})
\]

For $M(\bar{A}) = M(\bar{B})$, we have
\[
SD(\bar{A}, \bar{B}) = 2 |W(\bar{A}) - W(\bar{B})|
= 2 |W(\bar{B}) - W(\bar{A})|
= SD(\bar{B}, \bar{A})
\]

(ii) $M(\bar{A}) \geq M(\bar{B}) \geq M(\bar{C})$, we have
\[
SD(\bar{A}, \bar{C}) = M(\bar{A}) - M(\bar{C}) + 2 |W(\bar{A}) - W(\bar{C})|
= M(\bar{A}) - M(\bar{B}) + M(\bar{B}) - M(\bar{C})
+ 2 |W(\bar{A}) - W(\bar{B}) + W(\bar{B}) - W(\bar{C})|
\leq M(\bar{A}) - M(\bar{B}) + M(\bar{B}) - M(\bar{C})
+ 2 [W(\bar{A}) - W(\bar{B}) + W(\bar{B}) - W(\bar{C})]
= M(\bar{A}) - M(\bar{B}) + 2 |W(\bar{A}) - W(\bar{B})| + M(\bar{B}) - M(\bar{C})
\]

(iii) From item (i) and Definition 2.7, we have
\[
\bar{A} \approx \bar{B} \iff SD(\bar{A}, \bar{B}) = 0 = -SD(\bar{B}, \bar{A})
\]

Then $SD(\bar{B}, \bar{A}) = 0$, and $\bar{B} \approx \bar{A}$. $\Box$

Remark 2.9. The proof of Lemma 2.8 for trapezoidal fuzzy number is same as the proof of Lemma 2.6 in [2].

3. Testing Statistical Hypothesis Based on Fuzzy Data with One Parameter

3.1 Testing Simple Hypothesis Against Two-Sided Hypothesis

Suppose that we want to test
\[
\begin{cases}
H_0 : \bar{\xi} = \bar{\xi}_0 \\
H_1 : \bar{\xi} \neq \bar{\xi}_0
\end{cases}
\]

In the crisp case, the decision rule for testing a null hypothesis $H_0 : \bar{\xi} = \bar{\xi}_0$ against an alternative $H_1 : \bar{\xi} \neq \bar{\xi}_0$ at the significance level $\delta$ is as
\[
\begin{cases}
Z_0 \leq Z_{\delta/2} \text{ or } Z_0 \geq Z_{1-\delta/2}, & \Rightarrow \text{Reject } H_0 (RH_0) \\
Z_{\delta/2} < Z_0 \leq Z_{1-\delta/2}, & \Rightarrow \text{Accept } H_0 (AH_0)
\end{cases}
\]

where $Z_0 = Z_0(\bar{\xi}^*, \bar{\xi}_0)$ is the crisp test statistic value under null hypothesis, and $Z_0 = Z_0(\bar{\xi}^*, \bar{\xi}_0)$ is the $\delta$-quantile of the crisp test statistic. The above decision rule can be rewritten, if the signed distance $d(x, y) = x - y$ is available. Hence the above decision rule can be rewritten as follows:
\[
\begin{cases}
d(Z_0, Z_{\delta/2}) \leq 0 \text{ or } d(Z_0, Z_{1-\delta/2}) \geq 0, & \Rightarrow RH_0 \\
d(Z_0, Z_{\delta/2}) > 0 \text{ and } d(Z_0, Z_{1-\delta/2}) < 0, & \Rightarrow AH_0
\end{cases}
\]

Now, we extend the decision rule based on fuzzy data. First, the fuzzy test statistic under fuzzy data is obtained based on the method introduced by Arefi and Taheri [4] as follows:

(i) We obtain a fuzzy point estimation $\bar{\xi}^*$.

(ii) By substituting the $r$-cuts of the fuzzy point estimation $\bar{\xi}^*$ for the point estimation $\bar{\xi}_0$ in the crisp test statistic $(Z_0)$ and using the interval arithmetic, we obtain the $r$-cut of the so-called fuzzy statistic $\bar{Z}$ as follows:

\[
Z[r] = \left\{ Z_0(\bar{\xi}^*, \bar{\xi}_0) : \bar{\xi}^* \in \bar{\xi}_0[r] \right\}
\] (3.1)

Then the new decision rule for testing two-sided hypothesis based on the revised signed distance (Definition) is introduced as follows:
Then, we introduce the decision rule for testing null hypothesis which is follow as:

\[ Z_{\delta} = \frac{\bar{X} - \xi}{\sigma / \sqrt{n}} \]

\[ H_0 : \xi = \xi_0 \]

\[ H_1 : \xi > \xi_0 \]

In a precise environment, the decision rule for testing the above hypotheses, at \( \delta \) level of significance is of the form

\[ Z_0 \geq Z_{1-\delta} \Rightarrow RH_0 \]

\[ Z_0 < Z_{1-\delta} \Rightarrow AH_0 \]

where \( Z_0 \) is the value of the crisp test statistic under null hypothesis, and \( Z_{1-\delta} \) is the \((1 - \delta)\)-quantile of the distribution of the crisp test statistic. The above decision rule can be rewritten based on the signed distance \( d(x, y) = x - y \) which is presented as follows:

\[ d(Z_0, Z_{1-\delta}) \geq 0, \Rightarrow RH_0 \]

\[ d(Z_0, Z_{1-\delta}) < 0, \Rightarrow AH_0 \]

If the available data be as fuzzy (imprecise) the above decision rule extended. First, the fuzzy test statistic \( \bar{Z} \) is obtained based on relation (3.1). Then for testing one sided hypotheses we introduce new decision rule based on the signed distance, which is follow as:

**Definition 3.2 (New decision rule).** Let \( SD(\bar{Z}, Z_{\delta/2}) \) and \( SD(\bar{Z}, Z_{1-\delta/2}) \) be the signed distances between the fuzzy test statistic \( \bar{Z} \) and the quantiles \( Z_{\delta/2} \) and \( Z_{1-\delta/2} \) respectively. Then, we introduce the decision rule for testing null hypothesis \( H_0 : \xi = \xi_0 \) against one-sided hypothesis \( H_1 : \xi > \xi_0 \) is introduced as follows:

\[ SD(\bar{Z}, Z_{\delta/2}) \geq 0, \Rightarrow RH_0 \]

\[ SD(\bar{Z}, Z_{1-\delta/2}) < 0, \Rightarrow AH_0 \]

**3.3 Testing of hypotheses for the Mean**

Suppose that a random sample of size \( n \) is taken from \( N(\xi, \sigma^2) \) and we observe that the data in the form of fuzzy numbers \( \tilde{X}_1, \ldots, \tilde{X}_n \) instead of crisp data \( x_1, \ldots, x_n \). The usual point estimation for \( \xi \) is \( \bar{x} = \frac{\sum x_i}{n} \). By substituting the \( r \)-cuts of \( \tilde{X}_i, i = 1, \ldots, n \), \( (\tilde{X}_i^L, \tilde{X}_i^U) \) for \( x_i \) in the point estimation, the \( r \)-cuts of the fuzzy point estimation \( \tilde{X} \) is obtained as

\[ \tilde{X}_i[r] = \frac{1}{n} \sum_{i=1}^{n} [\tilde{X}_i^L, \tilde{X}_i^U] = \left[ \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_i^L, \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_i^U \right] = [\tilde{X}_i^L, \tilde{X}_i^U] \]

The crisp test statistic value under null hypotheses \( H_0 : \xi = \xi_0 \), the value of the crisp test statistic is \( z_0 = \frac{\bar{x} - \xi_0}{\sigma / \sqrt{n}} \). ByIn order to obtained the \( r \)-cuts of the fuzzy point estimation \( \tilde{X} \) for \( \xi \) and interval arithmetic is employed in \( z_0 \). The \( r \)-cuts of the fuzzy test statistic are presented as follows

\[ \tilde{Z}[r] = \frac{\tilde{X}_i[r] - \xi_0}{\sigma / \sqrt{n}} = \left[ \frac{\tilde{X}_i^L - \xi_0}{\sigma / \sqrt{n}}, \frac{\tilde{X}_i^U - \xi_0}{\sigma / \sqrt{n}} \right]. \]

Hence, the decision rule for testing the null hypothesis \( H_0 : \xi = \xi_0 \) against the hypothesis \( H_1 \) is as follows:

(i) for \( H_1 : \xi \neq \xi_0 \), we reject \( H_0 \), if \( SD(\bar{Z}, z_{\delta/2}) \leq 0 \) or \( SD(\bar{Z}, z_{1-\delta/2}) \geq 0 \)

(ii) for \( H_1 : \xi > \xi_0 \), we reject \( H_0 \), if \( SD(\bar{Z}, z_{1-\delta}) \geq 0 \)

(iii) for \( H_1 : \xi < \xi_0 \), we reject \( H_0 \), if \( SD(\bar{Z}, z_{\delta}) \leq 0 \)

where \( z_{\delta} \) is \( \delta \)-quantile of the standard normal distribution.

**3.4 Example**

Assume that, based on a random sample of size \( n = 25 \) from a population \( N(\xi, 9) \), the trapezoidal fuzzy data in Table 1. Now, the following hypotheses is to be tested at 5% level of significance.

\[ H_0 : \xi = 2 \]

\[ H_1 : \xi > 2 \]

Based on the centers of fuzzy data \( x_i \), the crisp test statistic is obtained as

\[ z_0 = \frac{\bar{x} - \xi_0}{\sigma / \sqrt{n}} = \frac{2.9868 - 2}{2/5} = 1.6447. \]

Since \( z_0 < z_{1-\delta} = 1.6449 \) and \( d(z_0, z_{1-\delta}) = -0.0002 < 0 \), we accept the hypothesis \( H_0 \). But in this case, the value of crisp statistic is close to the related quantile, and the decision to accept the null hypothesis \( H_0 \) is very sensitive.
If we use trapezoidal and triangular fuzzy data, then the \( r \)-cuts of fuzzy point estimation \( \tilde{X} \) are obtained as \( \tilde{X}_r = [X^L_r, X^U_r] \). For trapezoidal fuzzy data, we have

\[
\tilde{X}[r] = \frac{1}{n} \sum_{i=1}^{n} [X^L_i, X^U_i] = [1.6872 + 0.2996r, 4.4728 - 0.2860r]
\]

For triangular fuzzy data, we have

\[
\tilde{X}[r] = \frac{1}{n} \sum_{i=1}^{n} [X^L_i, X^U_i] = [2.6872 + 0.2996r, 3.2864 - 0.2996r]
\]

Also based on Subsection 3.2, the \( r \)-cuts of fuzzy test statistic for trapezoidal and triangular fuzzy data are obtained as

For trapezoidal fuzzy data, we have

\[
Z[r] = \frac{\tilde{X} - \xi_0}{\frac{\sigma}{\sqrt{n}}} = [-0.5213 + 0.4993r, 4.1213 - 0.4767r]
\]

For triangular fuzzy data, we have

\[
\tilde{Z}[r] = \frac{\tilde{X} - \xi_0}{\frac{\sigma}{\sqrt{n}}} = [1.1454 + 0.4993r, 2.1440 - 0.4993r]
\]

Hence, based on the revised signed distance, we have

For trapezoidal fuzzy data, we have

\[
SD(Z,z_{1-\delta}) = M(Z) - |W(Z)| - z_{0.95}
= 2.4472 - 1.5480 - 1.6449
= -0.7457 < 0
\]

For triangular fuzzy data, we have

\[
SD(Z,z_{1-\delta}) = M(Z) - 2|W(Z)| - z_{0.95}
= 1.6447 - 2(0.24965) - 1.6449
= -0.4995 < 0
\]

Since for both the fuzzy data \( SD(Z,z_{1-0.05}) < 0 \), we surely accept the hypothesis \( H_0 \).

### 3.5 Example

Assume that, based on a random sample of size \( n = 50 \) from \( N(\xi, \sigma^2 = 9) \). The trapezoidal fuzzy data is observed in Table 2. Now, suppose that the following hypothesis is wanted to be tested at 5% level of significance.

\[
\begin{align*}
H_0 : \bar{x} & = 2 \\
H_1 : \bar{x} & > 2
\end{align*}
\]

Based on the trapezoidal fuzzy data \( \tilde{X}, i = 1, \ldots, n \), the \( r \)-cuts of the fuzzy point estimation \( \tilde{X} \) is obtained as

\[
\tilde{X}[r] = \left[ \tilde{X}^L, \tilde{X}^U \right]
\]

\[
\tilde{X}[r] = \frac{1}{n} \sum_{i=1}^{n} [X^L_i, X^U_i] = \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_r = \frac{1}{n} \sum_{i=1}^{n} \left[ X^L_i + (1-r)X^U_i \right]
\]

\[
[1.77, 0.39, 0.28, 0.32] \quad [1.94, 0.39, 0.32, 0.27] \quad [1.68, 0.88, 0.27, 0.33] \quad [1.87, 0.47, 0.29, 0.31]
\]

\[
[1.61, 0.81, 0.26, 0.22] \quad [2.08, 0.82, 0.31, 0.28] \quad [2.05, 0.25, 0.30, 0.27] \quad [2.22, 0.40, 0.32, 0.28]
\]

\[
[1.24, 0.44, 0.30, 0.28] \quad [1.78, 0.98, 0.28, 0.3] \quad [2.14, 3.01, 0.31, 0.34]
\]

Also based on Subsection 3.2, the \( r \)-cuts of fuzzy test statistic for trapezoidal and triangular fuzzy data are obtained as

For trapezoidal fuzzy data, we have

\[
Z[r] = \frac{\tilde{X} - \xi_0}{\frac{\sigma}{\sqrt{n}}} = [-0.5213 + 0.4993r, 4.1213 - 0.4767r]
\]

For triangular fuzzy data, we have

\[
\tilde{Z}[r] = \frac{\tilde{X} - \xi_0}{\frac{\sigma}{\sqrt{n}}} = [1.1454 + 0.4993r, 2.1440 - 0.4993r]
\]

Hence, based on the revised signed distance, we have

For trapezoidal fuzzy data, we have

\[
SD(Z,z_{1-\delta}) = M(Z) - |W(Z)| - z_{0.95}
= 2.4472 - 1.5480 - 1.6449
= -0.7457 < 0
\]

For triangular fuzzy data, we have

\[
SD(Z,z_{1-\delta}) = M(Z) - 2|W(Z)| - z_{0.95}
= 1.6447 - 2(0.24965) - 1.6449
= -0.4995 < 0
\]

Since for both the fuzzy data \( SD(Z,z_{1-0.05}) < 0 \), we surely accept the hypothesis \( H_0 \).

In this example, suppose that we want to test the following hypothesis, at the significance level \( \delta = 0.05 \)

\[
\begin{align*}
H_0 : \bar{x} & = 2 \\
H_1 : \bar{x} & \neq 2
\end{align*}
\]
On testing statistical hypothesis for mean and difference between mean based on revised signed distance under fuzzy environment — 407/409

Based on the above fuzzy test statistic, we have

\[ SD(\hat{Z}, \tilde{z}_{\delta/2}) = SD\left(\hat{Z}, z_{0.025}\right) \]
\[ = M(\hat{Z}) + |W(\hat{Z})| - z_{0.025} \]
\[ = 0.8235 + 0.9797 + 1.9600 \]
\[ = 3.7632 > 0 \]

\[ SD(\hat{Z}, z_{1-\delta/2}) = SD\left(\hat{Z}, z_{0.975}\right) \]
\[ = M(\hat{Z}) - |W(\hat{Z})| - z_{0.025} \]
\[ = 0.8235 - 0.9797 - 1.9600 \]
\[ = -2.1162 < 0 \]

Since, \( SD(\hat{Z}, \tilde{z}_{\delta/2}) > 0 \) and \( SD(\hat{Z}, z_{1-\delta/2}) < 0 \). Hence based on the decision rule the hypothesis \( H_0 \) is accepted, at the significance level \( \delta = 0.05 \).

### 4. Testing Hypotheses for Difference between Means of two Normal Populations with Known Variances

Suppose that the two independent random samples of sizes \( n_1 \) and \( n_2 \) have been taken from \( N(\mu_1, \sigma_1^2) \) and \( N(\mu_2, \sigma_2^2) \) respectively (with \( \sigma_1^2 \) and \( \sigma_2^2 \) known) and observed the fuzzy numbers have been observed \( \tilde{X}_1, \ldots, \tilde{X}_{n_1} \) and \( \tilde{Y}_1, \ldots, \tilde{Y}_{n_1} \) instead of crisp numbers \( x_1, \ldots, x_{n_1} \) and \( y_1, \ldots, y_{n_1} \). The usual point estimation for \( \tilde{\xi} = \mu_1 - \mu_2 \) is \( \tilde{\xi}^* = \bar{X} - \bar{Y} \). Under the fuzzy data, the \( \rho \)-cuts of the point estimations of \( \tilde{\xi}^* \) is given by

\[ \tilde{\xi}^*[\rho] = [\tilde{\xi}^{U\rho}, \tilde{\xi}^{L\rho}] \]
\[ = \bar{X}^\rho - \bar{Y}^\rho \]
\[ = \left[ \frac{1}{n_1} \sum_{i=1}^{n_1} X_i^\rho - \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i^\rho \right] \]

Under the null hypothesis \( H_0 : \mu_1 - \mu_2 = \tilde{\xi}_0 \), the value of the crisp test statistic,

\[ Z_0 = \frac{(\bar{X} - \bar{Y}) - \tilde{\xi}_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \]

(distributed according to standard normal \( N(0, 1) \).

By substituting the \( \rho \)-cuts of the fuzzy point estimation \( \tilde{\xi}^* \) for \( \tilde{\xi}^* \), and using the interval arithmetic, the \( \rho \)-cuts of the fuzzy test statistics are obtained

\[ \bar{X}^\rho - \bar{Y}^\rho \]

Hence, we introduce the decision rule based on the new signed distance for testing the null hypothesis \( H_0 : \mu_1 - \mu_2 = \tilde{\xi}_0 \) against the hypothesis \( H_1 \) as follows:

(i) For \( H_1 : \mu_1 - \mu_2 \neq \tilde{\xi}_0 \) we reject \( H_0 \)
\[ \text{if } SD(\bar{X}, \tilde{z}_{\delta/2}) \leq 0 \text{ or } SD(\bar{X}, x_{1-\delta/2}) \geq 0, \]

(ii) For \( H_1 : \mu_1 - \mu_2 > \tilde{\xi}_0 \) we reject \( H_0 \)
\[ \text{if } SD(\bar{X}, \tilde{z}_{1-\delta}) \geq 0 \]

(iii) For \( H_1 : \mu_1 - \mu_2 < \tilde{\xi}_0 \) we reject \( H_0 \)
\[ \text{if } SD(\bar{X}, \tilde{z}_{\delta}) \leq 0 \]

| \((x_1, y_1, \mu, \eta)_{\tilde{T}_P}\) | \((x_1, y_1, \mu, \eta)_{\tilde{L}_P}\) | \((x_1, y_1, \mu, \eta)_{\tilde{U}_P}\) | \((x_1, y_1, \mu, \eta)_{\tilde{L}_P}\) |
|------------------|------------------|------------------|------------------|
| (1.4, 2.0, 2.0, 6) | (6.9, 7.5, 1.5, 1.9) | (3.4, 4.0, 0.4, 0.8) | (-0.8, -0.2, 2.0, 6) |
| (2.5, 3.1, 0.4, 0.8) | (4.7, 5.3, 1.1, 4) | (4.4, 6.0, 0.4, 0.8) | (0.5, 1.0, 2.0, 6) |
| (0.5, 1.0, 0.2, 0.6) | (0.9, 1.5, 0.2, 0.6) | (5.6, 6.2, 1.2, 1.6) | (-1.6, -1.0, 2.0, 6) |
| (3.3, 3.9, 0.7, 1.1) | (1.7, 2.3, 0.2, 0.6) | (0.9, 1.5, 0.3, 0.7) | (-3.2, -2.6, 0.4, 0.8) |
| (0.8, 1.4, 0.2, 0.6) | (1.1, 6.0, 1.0, 5) | (3.3, 6.0, 4.0, 8) | (1.2, 1.8, 0.3, 0.7) |
| (1.3, 1.9, 0.3, 0.7) | (2.6, 3.2, 0.6, 1) | (5.4, 6.1, 1.1, 4) | (0.5, 1.1, 0.2, 0.6) |
| (0.1, 0.7, 0.1, 0.5) | (0.1, 0.5, 0.1, 0.5) | (-2.5, -1.9, 0.2, 0.6) | (5.3, 5.9, 1.1, 4) |
| (3.3, 6.0, 0.5, 0.9) | (1.2, 1.8, 0.2, 0.6) | (0.6, 1.2, 0.2, 0.6) | (-1.1, -0.5, 0.1, 0.5) |
| (-1.7, -1.1, 0.2, 0.6) | (6.5, 7.1, 1.1, 4) | (0.4, 1.0, 1.0, 5) | |
| (-3.8, -3.2, 0.4, 0.8) | (1.4, 2.0, 2.0, 6) | (4.6, 5.2, 1.1, 4) | |
| (2.2, 6.0, 1.0, 5) | (4.5, 5.1, 1.1, 4) | (2.6, 3.2, 0.6, 1) | |
| (1.5, 2.1, 0.3, 0.7) | (2.4, 3.0, 3.0, 7) | (6.5, 7.1, 1.2, 1.6) | |
| (5.6, 6.2, 0.8, 1.2) | (-5.4, 4.0, 3.0, 7) | (5.8, 6.4, 1.1, 5) | |
| (2.7, 3.3, 0.4, 0.8) | (6.4, 7.1, 4.1, 8) | (5.4, 6.1, 1.1, 5) | |
4.1 Example
Assume that based on two independent random samples of sizes $n_1 = 25$ and $n_2 = 16$ from $N(\mu_1, 9)$ and $N(\mu_2, 4)$, the non-symmetric normal fuzzy numbers are obtained under uncertain environment in Table 3. Now, we assume that the following hypotheses are wanted to test 5% level of significance.

$$
\begin{align*}
H_0: \mu_1 - \mu_2 &= 2, \\
H_1: \mu_1 - \mu_2 &> 2
\end{align*}
$$

Based on fuzzy data given in Table 3, we calculate $\tilde{X}[r]$ and $\tilde{Y}[r]$ as follows:

$$
\begin{align*}
\tilde{X}[r] &= \left[ \frac{1}{25} \sum_{i=1}^{25} (x_i - \delta_i) \sqrt{-\ln(r)}, \frac{1}{25} \sum_{i=1}^{25} (x_i - \eta_i) \sqrt{-\ln(r)} \right] \\
&= [6.222 - 0.6344 \sqrt{-\ln(r)}, 7.422 + 0.7438 \sqrt{-\ln(r)}] \\
\tilde{Y}[r] &= \left[ \frac{1}{16} \sum_{i=1}^{16} (y_i - \delta_i) \sqrt{-\ln(r)}, \frac{1}{16} \sum_{i=1}^{16} (y_i - \eta_i) \sqrt{-\ln(r)} \right] \\
&= [4.0206 - 0.4021 \sqrt{-\ln(r)}, 5.2206 + 0.5221 \sqrt{-\ln(r)}]
\end{align*}
$$

These the $r$-cuts of fuzzy point estimations are obtained as

$$
\xi^+[r] = [\tilde{\xi}^+L, \tilde{\xi}^+U] = \left[ 1.0014 - 1.1565 \sqrt{-\ln(r)}, \right]
$$

hence, the $r$-cuts of the fuzzy test statistic are obtained as

$$
\tilde{Z}[r] = \left[ \tilde{Z}^L, \tilde{Z}^U \right] = \left[ \frac{\tilde{\xi}^+L - \eta_0}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - \frac{\tilde{\xi}^+U - \eta_0}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \right] \\
= \left[ -0.9986 - 1.1565 \sqrt{-\ln(r)} - 2, \right] \\
\left\lfloor \sqrt{\frac{9}{25}} + \frac{4}{16} \right\rfloor \\
1.4014 + 1.1459 \sqrt{-\ln(r)} - 2 \\
1.7944 + 1.4672 \sqrt{-\ln(r)}
$$

Using Maple software $I(\tilde{Z})$ and $D(\tilde{Z})$ are calculated as

$$
\begin{align*}
I(\tilde{Z}) &= \frac{1}{2} \int_{0}^{1} (Z^U_{\alpha} + Z^L_{\alpha}) d\alpha = 0.2519 \\
D(\tilde{Z}) &= \int_{0}^{1} (Z^U_{\alpha} - Z^L_{\alpha}) d\alpha d\alpha = 2.4601
\end{align*}
$$

Since $I(\tilde{Z}) = 0.2519 < z_{0.95} = 1.6449$, we obtain

$$
T(\tilde{Z}, z_{1-\delta}) = I(\tilde{Z}) - |D(\tilde{Z})| = 0.2519 - 2.4601 - 1.6449 \\
= -3.8532
$$

Hence $T(\tilde{Z}, z_{1-\delta}) < 0$ and we accept the null hypotheses, at the significance level $\delta = 0.05$. 

---

Table 3. Trapezoidal Fuzzy Data from a Normal Population in Example of Section 4.1

| $(x_i, y_i, \delta_i, \eta_i)_T$ | $(x_i, y_i, \delta_i, \eta_i)_T$ | $(x_i, y_i, \delta_i, \eta_i)_T$ | $(x_i, y_i, \delta_i, \eta_i)_T$ |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| (0.75, 7.5, 8.7, 0.87) | (0.76, 7.62, 8.82, 0.88) | (0.40, 3.99, 5.19, 0.52) | (0.38, 3.81, 5.01, 0.50) |
| (0.16, 1.57, 2.77, 0.28) | (0.43, 4.26, 5.46, 0.55) | (0.75, 7.5, 8.7, 0.87) | (0.37, 3.66, 4.86, 0.49) |
| (0.99, 9.95, 11.15, 1.12) | (1.2, 11.7, 12.9, 1.3) | (0.51, 13.6, 33.0, 0.63) | (0.14, 13.9, 25.9, 0.26) |
| (0.14, -1.4, -0.2, 0.02) | (0.25, 2.45, 3.65, 0.37) | (0.20, 2.03, 3.23, 0.32) |  |
| (0.59, 5.89, 7.09, 0.71) | (0.98, 9.76, 10.96, 1.1) | (0.41, 4.07, 5.27, 0.53) |  |
| (0.49, 4.93, 6.13, 0.61) | (0.63, 6.29, 7.49, 0.75) | (0.49, 4.86, 6.06, 0.61) |  |
| (0.85, 8.59, 7.97, 0.97) | (0.83, 8.31, 9.51, 0.95) | (0.45, 4.45, 5.65, 0.57) |  |
| (1.12, 12.01, 13.21, 1.3) | (0.88, 8.79, 9.98, 1.0) | (0.33, 3.25, 4.45, 0.45) |  |
| (0.74, 7.39, 8.59, 0.86) | (0.82, 8.23, 9.43, 0.94) | (0.16, 1.57, 2.77, 0.28) |  |
| (0.95, 9.53, 10.73, 1.1) | (0.47, 4.68, 5.88, 0.59) | (0.66, 6.64, 7.84, 0.78) |  |
| (0.48, 4.83, 6.03, 0.60) | (0.21, 2.12, 3.32, 0.33) | (0.18, 1.84, 3.04, 0.30) |  |
| (0.26, 2.63, 3.83, 0.38) |  | (0.44, 4.39, 5.59, 0.56) |  |
| (0.62, 2.42, 4.44, 0.74) |  | (0.58, 5.75, 6.95, 0.70) |  |
| (0.18, 1.78, 2.98, 0.30) |  |  |  |
5. Conclusion

In this work, we have approached the problem of testing statistical hypothesis under fuzzy environment. In order to test the statistical hypothesis which is crisp in nature, a decision rule is provided based on new signed distance of trapezoidal fuzzy number in the proposed approach. The main advantage of this approach is that the statistical hypothesis can be tested when the available data are of both triangular and trapezoidal fuzzy numbers. Finally, the numerical examples are given in order to demonstrate the feasibility of the proposed approach.

References

[1] M. Gh. Akbari and A. Rezaei, Bootstrap testing fuzzy hypotheses and observations on fuzzy statistic, Expert Systems with Applications, 37(2010), 5782–5787.
[2] M. Arefi, Testing Statistical Hypotheses under Fuzzy Data and based on a New Signed Distance, Iranian Journal of Fuzzy Systems, 15(3)(2018), 153–176.
[3] M. Are and S. M. Taheri, A fuzzy-based approach to testing statistical hypotheses, International Journal of Intelligent Technologies and Applied Statistics, 4(1)(2011), 109–132.
[4] J. J. Buckley, Fuzzy statistics: hypothesis testing, Soft Computing, 9(2004), 512–518.
[5] M. R. Casals, M. A. Gil, and P. Gil, The fuzzy decision problem: an approach to the problem of testing statistical hypotheses with fuzzy information, European Journal of Operation Research, 27(1986), 371–382.
[6] J. Chachi, S. Taheri, and R. Viertl, Testing statistical hypotheses based on fuzzy confidence intervals, Austrian Journal of Statistics, 41(4)(2012), 267–286.
[7] P. Filzmoser and R. Viertl, Testing hypotheses with fuzzy data: the fuzzy P-value, Metrika, 59(2004), 21–29.
[8] P. Grzegorzewski, and O. Hryniewicz, Testing statistical hypotheses in fuzzy environment, Mathware & Soft Computing, 4(1997), 203–217.
[9] P. Grzegorzewski, Testing statistical hypotheses with vague data, Fuzzy Sets and Systems, 112(2000), 501–510.
[10] P. Grzegorzewski, Testing fuzzy hypotheses with vague data, In: Statistical Modeling, Analysis and Management of Fuzzy Data, C. Bertoluzza et al. (Eds.), Springer, Heidelberg, (2002), 213–225.
[11] S. M. Taheri and G. Hesamian A generalization of the Wilcoxon signed-rank test and its applications, Statistical Papers, 54(2)(2013), 457–470.
[12] S. M. Taheri and G. Hesamian, Goodman-Kruskal measure of association for fuzzy categorized variables, Kybernetika, 47(1)(2011), 110–122.
[13] Sh. Yosefi, M. Are, and M. Gh. Akbari, A new approach for testing fuzzy hypotheses based on likelihood ratio statistic, Statistical Papers, 57(2016), 665–688.
[14] L. A. Zadeh, Fuzzy sets, Information and Control, 8(1965), 338–353.