International Journal of Applied Mathematics & Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 4, Issue 5, Aug - Sep 2015, 11-24
© IASET

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

TWO SAMPLE STATISTICAL HYPOTHESIS TEST FOR TRAPEZOIDAL FUZZY INTERVAL DATA

P. GAJIVARADHAN[1] & S. PARTHIBAN[2]

[1]Department of Mathematics, Pachaiyappa's College, Chennai, Tamil Nadu, India

[2]Research Scholar, Department of Mathematics, Pachaiyappa's College, Chennai, Tamil Nadu, India

ABSTRACT

Trapezoidal fuzzy numbers have numerous advantages over triangular fuzzy numbers as they have more generalized form. In this paper, two sample statistical test of hypothesis for means in normal population with interval data is given. The decision rules whether to accept or reject the null hypothesis or alternative hypothesis are given. Using numerical example, the test procedure is illustrated. The proposed test procedure has been extended to fuzzy valued statistical hypothesis testing for trapezoidal interval data.

KEYWORDS: Fuzzy Numbers, Trapezoidal Fuzzy Number (TFN), Trapezoidal Interval Data, Test of Hypothesis, Confidence Limits, Two Sample t - test

Mathematical Subject Classification: 62A86, 62F03, 97K80

1. INTRODUCTION

Hypothesis testing is a method for testing a claim or hypothesis about a parameter in a population by measuring the sample data. It is one of the most important areas of statistical analysis. In many situations, the statisticians are interested in testing hypothesis about the population parameter by using the available sample data. In classical testing procedure, the observations of sample are crisp and the corresponding statistical test leads to the binary decision like yes or no / positive or negative / accepted or rejected. But in practical life, we often come across the data in which most of them are vague or imprecise in nature. The statistical hypothesis testing procedure under such vague or fuzzy environments has been studied by many authors.

Arnold [4] discussed the fuzzy hypotheses testing with crisp data. Casals and Gil [8] and Son et al. analysed the Neyman-Pearson type of testing hypotheses [20]. Saade [18, 19] analysed the binary hypotheses testing and discussed the likelihood functions in the process of decision making.Akbari and Rezaei [2] analysed a notable method for inference about the variance based on fuzzy data. Grzegorzewski [12], Watanabe and Imaizumi [24] analysed the fuzzy tests for hypotheses testing with vague and ambiguous data. Wu [25] discussed and analysed the statistical hypotheses testing for fuzzy data by using the notion of degrees of optimism and pessimism. Viertl [22, 23] found some methods to construct confidence intervals and statistical test for fuzzy valued data. Wu [26] approached a new method to construct fuzzy confidence intervals for the unknown fuzzy parameter. Arefi and Taheri [3] found a new approach to test the fuzzy hypotheses upon fuzzy test statistic for imprecise and vague data. Chachi et al. [10] found a new method for the problem of testing statistical hypotheses for fuzzy data using the relationship between confidence intervals and hypotheses testing. Zadeh [27] analysed some notions and criterions about fuzzy probabilities. B. Asady [5] introduced a method to obtain the nearest trapezoidal approximation of fuzzy numbers. Abhinav Bansal [1] explored some arithmetic properties of arbitrary trapezoidal fuzzy numbers of the form (a, b, c, d).

In this paper, we perform a new statistical hypothesis testing procedure about the population means when the data of the given two samples are real intervals. And the decision rules to accept or reject the null hypothesis and alternative hypothesis are given. In this testing procedure, we split the given interval data into two different sets of crisp data namely, upper level data (X_U, Y_U) and lower level data (X_L, Y_L), then we find the test statistic values for these two sets of crisp data and then we obtain a decision about the population means in the light of decision rules. In this testing procedure, we do not use degrees of optimism and pessimism and h – level set. And one numerical example is given, further this test procedure has been extended to trapezoidal fuzzy interval data and we conclude the testing procedure with decision rules with an example.

2. PRELIMINARIES AND DEFINITIONS

Definition-2.1 Membership Function

A characteristic function μ_A of a crisp set $A \subseteq X$ assigns a value either 0 or 1 to each of the members in X. This function can be generalized to a function $\mu_{\tilde{A}}$ such that the value assigned to the element of the universal set X fall within the specified range. That is, $\mu_{\tilde{A}} : X \to [0, 1]$. The assigned value indicates the membership grade of the element in the set A. The function $\mu_{\tilde{A}}$ is called the 'membership function'.

Definition-2.2 Fuzzy Set

A fuzzy set $\tilde{A}$ of a universal X is defined by its membership function $\mu_{\tilde{A}} : X \to [0, 1]$ and we write $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in X\}$.

Definition-2.3 α - level Set of a Fuzzy Set $\tilde{A}$

The α - cut or α - level set of a fuzzy set $\tilde{A}$ is defined by $\tilde{A} = \{x : \mu_{\tilde{A}}(x) \geq \alpha\}$ where $x \in X$. And $\tilde{A}_0$ is the **closure** of the set $\{x : \mu_{\tilde{A}}(x) \neq 0\}$.

Definition-2.4 Normal Fuzzy Set

A fuzzy set $\tilde{A}$ is called normal fuzzy set if there exists an element (member) 'x' such that $\mu_{\tilde{A}}(x) = 1$.

Definition-2.5 Convex Fuzzy Set

A fuzzy set $\tilde{A}$ is called convex fuzzy set if $\mu_{\tilde{A}}(\alpha x_1 + (1 - \alpha) x_2) \geq \min\{\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)\}$ where $x_1, x_2 \in X$ and $\alpha \in [0, 1]$.

Definition-2.6 Fuzzy Number

A fuzzy set $\widetilde{A}$, defined on the universal set of real number R, is said to be 'fuzzy number' if its membership function has the following characteristics:

i. $\widetilde{A}$ is convex,

ii. $\widetilde{A}$ is normal,

iii. $\mu_{\widetilde{A}}$ is piecewise continuous.

Definition-2.7 Trapezoidal Fuzzy Number

A fuzzy number $\widetilde{A} = (a, b, c, d)$ is said to be a trapezoidal fuzzy number if its membership function is given by

$$\mu_{\widetilde{A}}(x) = \begin{cases} 0 & ; \ x < a \\ \dfrac{x - a}{b - a} & ; \ a < x \leq b \\ 1 & ; \ b < x < c \\ \dfrac{d - x}{d - c} & ; \ c \leq x < d \\ 0 & ; \ x > d \end{cases}$$

where $a \leq b \leq c \leq d$. A trapezoidal fuzzy number is a triangular fuzzy number if $b = c$.

Definition-2.8

Let $\widetilde{A}_1 = (a, b, c, d)$ and $\widetilde{A}_2 = (e, f, g, h)$ be two non - negative trapezoidal fuzzy numbers then,

i. $\widetilde{A}_1 \oplus \widetilde{A}_2 = (a, b, c, d) \oplus (e, f, g, h) = (a + e, b + f, c + g, d + h)$

ii. $\widetilde{A}_1 \Theta \widetilde{A}_2 = (a, b, c, d) \Theta (e, f, g, h) = (a - h, b - g, c - f, d - e)$

iii. $-\widetilde{A}_1 = -(a, b, c, d) = (-d, -c, -b, -a)$

iv. $\widetilde{A}_1 \otimes \widetilde{A}_2 = (a, b, c, d) \otimes (e, f, g, h) = (ae, bf, cg, dh)$

v. $k\widetilde{A} = [ka, kb]$ if 'k' is a positive real number.

vi. $k\widetilde{A} = [kb, ka]$ if 'k' is a negative real number.

vii. $A \otimes B = [p, q]$ where $p = \max \{ac, ad, bc, bd\}$ $q = \min \{ac, ad, bc, bd\}$

viii. $1/\widetilde{A} \cong \left(1/d,\ 1/c,\ 1/b,\ 1/a\right)$

Definition-2.9

Let $A = [a, b]$ and $B = [c, d] \in D$.

Then, (i) $A \le B$ if $a \le c$ and $b \le d$ (ii) $A \ge B$ if $a \ge c$ and $b \ge d$ (iii) $A=B$ if $a = c$ and $b = d$.

3. TWO – SAMPLE (STUDENT) T TEST

Let x_i and y_j, $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$ be two random samples from two different normal populations with sizes m and n such that $m + n \le 30$. And let $\overline{x}$ and $\overline{y}$ be the mean values and s_1 and s_2 be the sample standard deviations of the random variables x_i and y_j respectively and they are given by [13],

$$\overline{x} = \frac{1}{m}\left(\sum_{i=1}^{m} x_i\right) \text{ and } \overline{y} = \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)$$

$$s_1 = \sqrt{\left(\frac{1}{m-1}\right)\left(\sum_{i=1}^{m}\left(x_i - \overline{x}\right)^2\right)} \text{ and } s_2 = \sqrt{\left(\frac{1}{n-1}\right)\left(\sum_{j=1}^{n}\left(y_j - \overline{y}\right)^2\right)} \quad \rightarrow (1)$$

Let μ_1 and μ_2 be the population means of X-sample and Y-sample respectively. In testing the null hypothesis $H_0 : \mu_1 = \mu_2$ assuming with **equal** population standard deviation, we generally use the test statistic:

$$t = \frac{\overline{x} - \overline{y}}{s\sqrt{\frac{1}{m} + \frac{1}{n}}} \text{ where } s = \sqrt{\frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}}$$

In testing the null hypothesis $H_0 : \mu_1 = \mu_2$ assuming with **unequal** population standard deviation, we commonly use the test statistic:

$$t = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

Now, the degrees of freedom used in this test is $\nu = n + m - 2$. Let α be the level of significance and let $t_{\alpha,\nu}$ be the tabulated value of ‘t’ for ν degrees of freedom at α level of significance. And the null hypothesis is given by $H_0 : \mu_1 = \mu_2$. And the rejection region for the alternative hypothesis H_A at α level is given below:

Alternative Hypothesis H_A	Rejection Region at α Level
$H_A : \mu_1 > \mu_2$	$t \geq t_{\alpha, m+n-2}$ (Upper tailed test)
$H_A : \mu_1 < \mu_2$	$t \leq -t_{\alpha, m+n-2}$ (Lower tailed test)
$H_A : \mu_1 \neq \mu_2$	$\lvert t \rvert \geq t_{\alpha/2,\ m+n-2}$ (Two tailed test)

If $\lvert t \rvert < t_{\alpha,\ m+n-2}$ (one tailed test), the difference between μ_1 and μ_2 is not significant at α level. Then the means of the populations are identical. That is, $\mu_1 = \mu_2$ at α level of significance. Therefore, the null hypothesis H_0 is accepted. Otherwise, the alternative hypothesis H_A is accepted.

If $\lvert t \rvert < t_{\alpha/2,\ m+n-2}$ (two tailed test), the difference between μ_1 and μ_2 is not significant at α level. Then the means of the populations are identical. That is, $\mu_1 = \mu_2$ at α level of significance. Therefore, the null hypothesis H_0 is accepted. Otherwise, the alternative hypothesis H_A is accepted.

Now, the $100(1-\alpha)\%$ confidence limits for the difference of population means μ_1 and μ_2 corresponding to the given samples are given by,

$$\left(\bar{x} - \bar{y}\right) - t_{\alpha/2,\ m+n-2}\left(s\sqrt{\frac{1}{m} + \frac{1}{n}}\right) < \left(\mu_1 - \mu_2\right)$$

$$< \left(\bar{x} + \bar{y}\right) + t_{\alpha/2,\ m+n-2}\left(s\sqrt{\frac{1}{m} + \frac{1}{n}}\right) \text{ (for equal standard deviations)}$$

Or

$$\left(\bar{x} - \bar{y}\right) - t_{\alpha/2,\ m+n-2}\left(\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}\right) < \left(\mu_1 - \mu_2\right)$$

$$< \left(\bar{x} + \bar{y}\right) + t_{\alpha/2,\ m+n-2}\left(\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}\right) \text{ (for unequal standard deviations)}$$

4. TEST OF HYPOTHESIS FOR INTERVAL DATA

Let $\left\{\left[a_i, b_i\right],\ i = 1, 2, ..., m\right\}$ be a random small sample (X-sample) with size m and $\left\{\left[c_j, d_j\right],\ j = 1, 2, ..., n\right\}$ be a random small sample (Y-sample) with size n. And let $\left[\eta_1, \mu_1\right]$ be mean of X from a normal population and let $\left[\eta_2, \mu_2\right]$ be mean of Y from another normal population.

Now, we test the null hypothesis H_0 such that the means of the population of the given samples are equal. That is $H_0 : [\eta_1, \mu_1] = [\eta_2, \mu_2] \Rightarrow \eta_1 = \eta_2$ and $\mu_1 = \mu_2$. And the alternative hypotheses are given by [11],

i. $H_A : [\eta_1, \mu_1] < [\eta_2, \mu_2] \Rightarrow \eta_1 < \eta_2$ and $\mu_1 < \mu_2$

ii. $H_A : [\eta_1, \mu_1] > [\eta_2, \mu_2] \Rightarrow \eta_1 > \eta_2$ and $\mu_1 > \mu_2$

iii. $H_A : [\eta_1, \mu_1] \neq [\eta_2, \mu_2] \Rightarrow \eta_1 \neq \eta_2$ or $\mu_1 \neq \mu_2$

Now the lower values and upper values for X-sample and Y-sample are given below:

X_L (Lower values of X-sample)	a_i ; $i = 1, 2, ..., m$
Y_L (Lower values of Y-sample)	c_j ; $j = 1, 2, ..., n$
X_U (Upper values of X-sample)	b_i ; $i = 1, 2, ..., m$
Y_U (Upper values of Y-sample)	d_j ; $j = 1, 2, ..., n$

Let $\bar{x}_L$ and $\bar{y}_L$ be the sample means, s_{x_L} and s_{y_L} be the sample standard deviation of X^L and Y^L respectively. Similarly let $\bar{x}_U$ and $\bar{y}_U$ be the sample means, s_{x_U} and s_{y_U} be the sample standard deviation of X^U and Y^U respectively.

Case (i): If the population standard deviations are assumed to be **equal,** then under the null hypothesis $H_0 : [\eta_1, \mu_1] = [\eta_2, \mu_2]$, the test statistic is given by,

$$t_L = \frac{\bar{x}_L - \bar{y}_L}{s_L\sqrt{\frac{1}{m} + \frac{1}{n}}} \text{ and } t_U = \frac{\bar{x}_U - \bar{y}_U}{s_U\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

where $s_L = \sqrt{\frac{(m-1)s_{x_L}^2 + (n-1)s_{y_L}^2}{m+n-2}}$ and $s_U = \sqrt{\frac{(m-1)s_{x_U}^2 + (n-1)s_{y_U}^2}{m+n-2}}$

Case (ii): If the population standard deviations are assumed to be **unequal**, then under the null hypothesis $H_0 : [\eta_1, \mu_1] = [\eta_2, \mu_2]$, the test statistic is given by,

$$t_L = \frac{\bar{x}_L - \bar{y}_L}{\sqrt{\frac{s_{x_L}^2}{m} + \frac{s_{y_L}^2}{n}}} \text{ and } t_U = \frac{\bar{x}_U - \bar{y}_U}{\sqrt{\frac{s_{x_U}^2}{m} + \frac{s_{y_U}^2}{n}}}$$

where the standard deviations for upper and lower values of the samples of X and Y are given by the equation (1). And the rejection region of the alternative hypothesis H_A at α level of significance is given below:

Alternative Hypothesis H_A	**Rejection Region at** α **Level**
$H_A : [\eta_1, \mu_1] > [\eta_2, \mu_2]$	$t_L \geq t_{\alpha, m+n-2}$ and $t_U \geq t_{\alpha, m+n-2}$ (Upper tailed test)
$H_A : [\eta_1, \mu_1] < [\eta_2, \mu_2]$	$t_L \leq -t_{\alpha, m+n-2}$ and $t_U \leq -t_{\alpha, m+n-2}$ (Lower tailed test)
$H_A : [\eta_1, \mu_1] \neq [\eta_2, \mu_2]$	$\lvert t_L \rvert \geq t_{\alpha/2, m+n-2}$ or $\lvert t_U \rvert \geq t_{\alpha/2, m+n-2}$ (Two tailed test)

If $\lvert t_L \rvert < t_{\alpha, m+n-2}$ and $\lvert t_U \rvert < t_{\alpha, m+n-2}$ (one tailed test), then the difference between $[\eta_1, \mu_1]$ and $[\eta_2, \mu_2]$ is not significant at α level. Then the means of the populations are identical. That is, $[\eta_1, \mu_1] = [\eta_2, \mu_2]$ at α level of significance. Therefore, the null hypothesis H_0 is accepted. Otherwise, the alternative hypothesis H_A is accepted.

If $\lvert t_L \rvert < t_{\alpha/2, m+n-2}$ and $\lvert t_U \rvert < t_{\alpha/2, m+n-2}$ (two tailed test), the difference between $[\eta_1, \mu_1]$ and $[\eta_2, \mu_2]$ is not significant at α level. Then the means of the populations are identical. That is, $[\eta_1, \mu_1] = [\eta_2, \mu_2]$ at α level of significance. Therefore, the null hypothesis H_0 is accepted. Otherwise, the alternative hypothesis H_A is accepted.

And the $100(1-\alpha)\%$ confidence limits for the difference of lower limit and upper limit of the population means $[\eta_1, \mu_1]$ and $[\eta_2, \mu_2]$ corresponding to the given samples are given below:

$$\left(\bar{x}_L - \bar{y}_L\right) - t_{\alpha/2,\ m+n-2}\left(s_L\sqrt{\frac{1}{m}+\frac{1}{n}}\right) < \left(\eta_1 - \eta_2\right) < \left(\bar{x}_L - \bar{y}_L\right) + t_{\alpha/2,\ m+n-2}\left(s_L\sqrt{\frac{1}{m}+\frac{1}{n}}\right)$$

(for equal population standard deviations), and

$$\left(\bar{x}_U - \bar{y}_U\right) - t_{\alpha/2,\ m+n-2}\left(s_U\sqrt{\frac{1}{m}+\frac{1}{n}}\right) < \left(\eta_1 - \eta_2\right) < \left(\bar{x}_U - \bar{y}_U\right) + t_{\alpha/2,\ m+n-2}\left(s_U\sqrt{\frac{1}{m}+\frac{1}{n}}\right)$$

(for equal population standard deviations)

Or

$$\left(\bar{x}_L - \bar{y}_L\right) - t_{\alpha/2,\ m+n-2}\left(\sqrt{\frac{s_{x_L}^2}{m} + \frac{s_{y_L}^2}{n}}\right) < \left(\eta_1 - \eta_2\right)$$

$$< \left(\bar{x}_L - \bar{y}_L\right) + t_{\alpha/2,\ m+n-2}\left(\sqrt{\frac{s_{x_L}^2}{m} + \frac{s_{y_L}^2}{n}}\right)$$

(for unequal population standard deviations) and

$$\left(\bar{x}_U - \bar{y}_U\right) - t_{\alpha/2,\ m+n-2}\left(\sqrt{\frac{s_{x_U}^2}{m} + \frac{s_{y_U}^2}{n}}\right) < \left(\eta_1 - \eta_2\right)$$

$$< \left(\bar{x}_U - \bar{y}_U\right) + t_{\alpha/2,\ m+n-2}\left(\sqrt{\frac{s_{x_U}^2}{m} + \frac{s_{y_U}^2}{n}}\right)$$

(for unequal population standard deviations)

This test procedure has been illustrated using the following numerical examples.

Example-1

The following interval data are given the gain in weights (in lbs) of pet dogs fed on two kinds of diets A and B [13].

Diet-A	Diet-B	Diet-A	Diet-B
[18, 19]	[22, 26]	[19, 22]	[22, 28]
[16, 18]	[27, 31]	[20, 24]	[20, 24]
[30, 32]	[25, 28]	[27, 30]	[11, 15]
[28, 30]	[12, 16]	[18, 22]	[14, 17]
[22, 24]	[16, 20]	[21, 24]	[17, 21]
[14, 16]	[18, 22]	--	[25, 27]
[28, 32]	[26, 30]	--	[19, 22]
		--	[23, 25]

Now, we test if the two diets differ significantly on the basis of their nutrition effects on increase in the weight of the pet dogs.

Here the null hypothesis is, $H_0 : [\eta_1, \mu_1] = [\eta_2, \mu_2] \Rightarrow \eta_1 = \eta_2$ and $\mu_1 = \mu_2$.

$\Rightarrow$ There is no significant difference between the nutrition effects from diet A and diet B.

And the alternative hypothesis is $H_A : [\eta_1, \mu_1] \neq [\eta_2, \mu_2] \Rightarrow \eta_1 \neq \eta_2$ or $\mu_1 \neq \mu_2$ (Two tailed test).

$\Rightarrow$ The two kinds of the diets differ significantly on the basis of their nutrition effects.

We assume that the **standard deviations of the populations are not equal** and we use 5% level of significance. Here m=12 and n=15.

The tabulated value of 't' for m + n – 2 = 27 – 2 = 25 degrees of freedom at 5% level of significance is $T_\alpha = 2.06$.

Now, $\bar{x}_L = \frac{1}{m}\left(\sum_{i=1}^{m} x_{i_L}\right) \Rightarrow \bar{x}_L = 21.75$ and $\bar{y}_L = \frac{1}{n}\left(\sum_{i=1}^{n} y_{i_L}\right) \Rightarrow \bar{y}_L = 20.0667$ and

$$\bar{x}_U = \frac{1}{m}\left(\sum_{i=1}^{m} x_{i_U}\right) \Rightarrow \bar{x}_U = 24.4167 \text{ and } \bar{y}_U = \frac{1}{n}\left(\sum_{i=1}^{n} y_{i_U}\right) \Rightarrow \bar{y}_U = 23.7333$$

$$s_{x_L}^2 = \left(\frac{1}{m-1}\right)\sum_{i=1}^{m}\left(x_{i_L} - \bar{x}_L\right)^2 \Rightarrow s_{x_L}^2 = 27.8409 \text{ and } s_{y_L}^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\left(y_{i_L} - \bar{y}_L\right)^2 \Rightarrow s_{y_L}^2 = 27.0667$$

$$s_{x_U}^2 = \left(\frac{1}{m-1}\right)\sum_{i=1}^{m}\left(x_{i_U} - \bar{x}_U\right)^2 \Rightarrow s_{x_U}^2 = 30.0833 \text{ and } s_{y_U}^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\left(y_{i_U} - \bar{y}_U\right)^2 \Rightarrow s_{y_U}^2 = 26.6381$$

The Test Statistics:

$$t_L = \frac{\bar{x}_L - \bar{y}_L}{\sqrt{\frac{s_{x_L}^2}{m} + \frac{s_{y_L}^2}{n}}} \Rightarrow t_L = 0.8288 \text{ and } t_U = \frac{\bar{x}_U - \bar{y}_U}{\sqrt{\frac{s_{x_U}^2}{m} + \frac{s_{y_U}^2}{n}}} \Rightarrow t_U = 0.3302$$

Since, $|t_L| < T_\alpha = 2.06$ and $|t_U| < T_\alpha = 2.06$, we accept the null hypothesis H_0.

$\Rightarrow$ There is no significant difference between the nutrition effects of the diets A and B at 5% level of significance.

5. TEST OF HYPOTHESIS FOR FUZZY DATA USING TRAPEZOIDAL FUZZY NUMBER (TFN)

Definition 5.1: Trapezoidal Fuzzy Number to Interval

Let a trapezoidal fuzzy number be defined as $\tilde{A} = (a, b, c, d)$, then the fuzzy interval (Superna Das and S. Chakraverty) in terms of α - cut interval is defined as follows [21]:

$$\tilde{A}_\alpha = \left[a + (b - a)\alpha, d - (d - c)\alpha\right];\ 0 \le \alpha \le 1 \quad \rightarrow(1)$$

Suppose that the given sample is a fuzzy data that are trapezoidal fuzzy numbers and we have to test the hypothesis about the population mean. Using the relation (1) and the proposed test procedure, we can test the hypothesis by transferring the fuzzy data into interval data.

Example-2

Two kinds of engine oils A and B for automobiles are under mileage test for some taxies, then we request the taxi drivers to record the consumption of fuel. Due to limited available source, the data are recorded as trapezoidal fuzzy numbers which are given in the following table. Suppose the random variables have normal distribution and their variances of both populations are **known and equal with one**. We now investigate the effects of the two kinds of engine oils on consumption of fuel at 5% level of significance [6, 14].

$\widetilde{A}$	$\widetilde{B}$
$(4, 4.5, 5, 6)$	$(5, 6.5, 7, 8)$
$(3.5, 4, 5, 6.5)$	$(4, 4.5, 5, 6)$
$(5, 5.5, 5.8, 6)$	$(5.5, 7, 8, 8.5)$
$(5.5, 5.8, 6, 6.5)$	$(5, 6, 6.5, 7)$
$(3, 3.5, 4, 5)$	$(6, 6.5, 7, 8)$
--	$(6, 7.5, 8.5, 9)$

Now the interval representation of the above trapezoidal data is given below:

$\left[\widetilde{A}_\alpha\right]$	$\left[\widetilde{B}_\alpha\right]$
$[4 + 0.5\alpha, 6 - \alpha]$	$[5 + 1.5\alpha, 8 - \alpha]$
$[3.5 + 0.5\alpha, 6.5 - 1.5\alpha]$	$[4 + 0.5\alpha, 6 - \alpha]$
$[5 + 0.5\alpha, 6 - 0.2\alpha]$	$[5.5 + 1.5\alpha, 8.5 - 0.5\alpha]$
$[5.5 + 0.3\alpha, 6.5 - 0.5\alpha]$	$[5 + \alpha, 7 - 0.5\alpha]$
$[3 + 0.5\alpha, 5 - \alpha]$	$[6 + 0.5\alpha, 8 - \alpha]$
--	$[6 + 1.5\alpha, 9 - 0.5\alpha]$

Lower Level Samples		**Upper Level Samples**	
x_L	y_L	x_U	y_U
$4 + 0.5\alpha$	$5 + 1.5\alpha$	$6 - \alpha$	$8 - \alpha$
$3.5 + 0.5\alpha$	$4 + 0.5\alpha$	$6.5 - 1.5\alpha$	$6 - \alpha$
$5 + 0.5\alpha$	$5.5 + 1.5\alpha$	$6 - 0.2\alpha$	$8.5 - 0.5\alpha$
$5.5 + 0.3\alpha$	$5 + \alpha$	$6.5 - 0.5\alpha$	$7 - 0.5\alpha$
$3 + 0.5\alpha$	$6 + 0.5\alpha$	$5 - \alpha$	$8 - \alpha$
--	$6 + 1.5\alpha$	--	$9 - 0.5\alpha$

Here, m = 5 and n = 6.

$$\overline{x}_L = \frac{1}{m}\left(\sum_{i=1}^{m} x_{i_L}\right) \Rightarrow \overline{x}_L = 4.2 + 0.46\alpha \text{ and } \overline{x}_U = \frac{1}{m}\left(\sum_{i=1}^{m} x_{i_U}\right) \Rightarrow \overline{x}_U = 6 - 0.84\alpha$$

$$\overline{y}_L = \frac{1}{n}\left(\sum_{i=1}^{n} y_{i_L}\right) \Rightarrow \overline{y}_L = 5.25 + 1.083\alpha \text{ and } \overline{y}_U = \frac{1}{n}\left(\sum_{i=1}^{n} y_{i_U}\right) \Rightarrow \overline{y}_U = 7.75 - 0.75\alpha$$

$$s_{x_L}^2 = \left(\frac{1}{m-1}\right)\sum_{i=1}^{m}\left(x_{i_L} - \bar{x}_L\right)^2 = 0.008\alpha^2 - 0.13\alpha + 1.075 \text{ and}$$

$$s_{y_L}^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\left(y_{i_L} - \bar{y}_L\right)^2 = 0.2417\alpha^2 + 0.25\alpha + 0.5750$$

$$s_{x_U}^2 = \left(\frac{1}{m-1}\right)\sum_{i=1}^{m}\left(x_{i_U} - \bar{x}_U\right)^2 = 0.2402\alpha^2 + 0.375 \text{ and}$$

$$s_{y_U}^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\left(y_{i_U} - \bar{y}_U\right)^2 = 0.05\alpha^2 + 0.25\alpha - 0.05 \text{ and}$$

$$S_L^2 = \frac{(m-1)s_{x_L}^2 + (n-1)s_{y_L}^2}{(m+n-2)} = 0.1379\alpha^2 - 0.1967\alpha + 0.7972$$

$$S_U^2 = \frac{(m-1)s_{x_U}^2 + (n-1)s_{y_U}^2}{(m+n-2)} = 0.1346\alpha^2 + 0.1389\alpha + 0.1389$$

Now, the null hypothesis,

$\tilde{H}_0 : \tilde{\Lambda} \approx \tilde{\Omega} \Rightarrow$ The two kinds of engine oils on fuel consumption are same.

The alternative hypothesis,

$\tilde{H}_A : \tilde{\Lambda} \neq \tilde{\Omega} \Rightarrow$ The two kinds of engine oils on fuel consumption differ significantly.

Here, $\left[\tilde{\Lambda}\right] = \left[\eta_1, \mu_1\right]$ and $\left[\tilde{\Omega}\right] = \left[\eta_2, \mu_2\right]$.

And therefore, $\left[\tilde{H}_0\right] : \left[\tilde{\Lambda}\right] \approx \left[\tilde{\Omega}\right] \Rightarrow \tilde{H}_0 : \eta_1 = \eta_2$ and $\mu_1 = \mu_2$.

$\left[\tilde{H}_A\right] : \left[\tilde{\Lambda}\right] \neq \left[\tilde{\Omega}\right] \Rightarrow \tilde{H}_A : \eta_1 \neq \eta_2$ or $\mu_1 \neq \mu_2$ (Two tailed test).

Now, the tabulated value of 't' at 5% level of significance with 9 degrees of freedom is $T_\alpha = 2.262$.

Test statistics:

$$t_L = \frac{\bar{x}_L - \bar{y}_L}{\sqrt{\frac{s_{x_L}^2}{m} + \frac{s_{y_L}^2}{n}}} = \begin{cases} -1.9422 & \text{if } \alpha = 0 \\ -2.0814 & \text{if } \alpha = 0.1 \\ -2.2204 & \text{if } \alpha = 0.2 \\ -2.3578 & \text{if } \alpha = 0.3 \\ \cdots & \\ -3.2154 & \text{if } \alpha = 1 \end{cases} \qquad \Rightarrow \left|t_L\right| > T_\alpha \text{ for } 0.3 \le \alpha \le 1$$

$$t_U = \frac{\bar{x}_U - \bar{y}_U}{\sqrt{\frac{s^2_{x_U}}{m} + \frac{s^2_{y_U}}{n}}} = \begin{cases} -7.7537 & \text{if } \alpha = 0 \\ -4.7320 & \text{if } \alpha = 1 \end{cases} \qquad \Rightarrow |t_U| > T_\alpha \text{ for } 0 \le \alpha \le 1$$

CONCLUSIONS

Hence, $|t_L| > T_\alpha$ and $|t_U| > T_\alpha$ for $0.3 \le \alpha \le 1$ implies the null hypothesis $\widetilde{H}_0$ is rejected and we accept the alternative hypothesis $\widetilde{H}_A$. Therefore, the two kinds of engine oils for automobiles on consumption of fuel are not the same at 5% level of significance.

Remark

The obtained result from the above test procedure in Example-2 differs by the lower value of α by 0.3 when compared with the result in Baloui Jamkhaneh and Nadi Gara [6] and Kalpanapriya et al. [14] which is $0 \le \alpha \le 1$ when performing this test procedure using trapezoidal fuzzy interval data.

REFERENCES

1. Abhinav Bansal, Trapezoidal Fuzzy Numbers (a, b, c, d): Arithmetic Behavior, International Journal of Physical and Mathematical Sciences (2011), 39-44.
2. M. G. Akbari and A. Resaei, Bootstrap statistical inference for the variance based on fuzzy data. Austrian Journal of Statistics, 38 (2009), 121-130.
3. M. Arefi and S. M. Taheri, Testing fuzzy hypotheses using fuzzy data based on fuzzy test statistic. Journal of Uncertain Systems, 5 (2011), 45-61.
4. B. F. Arnold, Testing fuzzy hypotheses with crisp data, Fuzzy Sets and Systems 94 (1998), 323-333.
5. Asady, B. Trapezoidal Approximation of a Fuzzy Number Preserving the Expected Interval and Including the Core, American Journal of Operations Research, 3 (2013), 299-306.
6. E. Baloui Jamkhaneh and A. Nadi Ghara, Testing statistical hypotheses for compare means with vague data, International Mathematical Forum, 5 (2010) 615-620.
7. James J. Buckley, Fuzzy Probability and Statistics, Springer-Verlag Berlin Heidelberg 2006.
8. M. R. Casals, M. A. Gil, A note on the operativeness of Neyman-Pearson tests with fuzzy information, Fuzzy Sets and Systems 30 (1989) 215-220.
9. 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 Operational Research 27 (1986), 371-382.
10. J. Chachi, S. M. Taheri and R. Viertl, Testing statistical hypotheses based on fuzzy confidence intervals, Forschungsbericht SM-2012-2, Technische Universitat Wien, Austria, 2012.
11. J. L. Devore, Probability and Statistics for Engineers, Cengage, 2008.

12. P. Grzegorzewski, Testing statistical hypotheses with vague data, Fuzzy sets and Systems, 112 (2000), 501-510.
13. S. C. Gupta and V. K. Kapoor, Fundamentals of Mathematical Statistics (A Modern Approach), Sultan Chand & Sons, New Delhi.
14. D. Kalpanapriya and P. Pandian, Two sample statistical hypothesis test for means with imprecise data, International Journal of Engineering Research and Applications, Vol.2 Issue 3 (May-June 2012), 3210-3217.
15. George J. Klir and Bo Yuan, Fuzzy sets and fuzzy logic, Theory and Applications, Prentice-Hall, New Jersey, 2008.
16. R. E. Moore, Method and applications of interval analysis, SLAM, Philadelphia, PA, 1979.
17. V. A. Niskanen, Prospects for soft statistical computing: describing data and inferrings from data with words in the human sciences, Information Sciences, 132 (2001), 83-131.
18. J. J. Saade and H. Schwarzlander, Fuzzy hypothesis testing with hybrid data, Fuzzy Sets and Systems, 35 (1990), 197-212.
19. J. J. Saade, Extension of fuzzy hypothesis testing with hybrid data, Fuzzy Sets and Systems, 63 (1994), 57-71.
20. J. Ch. Son, I. Song and H. Y. Kim, A fuzzy decision problem based on the generalized Neymen-Pearson criterion, Fuzzy Sets and Systems, 47 (1992), 65-75.
21. Superna Das and S. Chakraverty, Numerical Solution of Interval and Fuzzy Systems of Linear Equations, Applications and Applied Mathematics: An International Journal (AAM): Vol. 7, Issue 1 (June 2012), pp. 334-356.
22. R. Viertl, Univariate statistical analysis with fuzzy data, Computational Statistics and Data Analysis, 51 (2006), 33-147.
23. R. Viertl, Statistical methods for fuzzy data, John Wiley and Sons, Chichester, 2011.
24. N. Watanabe and T. Imaizumi, A fuzzy statistical test of fuzzy hypotheses, Fuzzy Sets and Systems, 53 (1993) 167-178.
25. H. C. Wu, Statistical hypotheses testing for fuzzy data, Information Sciences, 175 (2005), 30-56.
26. H. C. Wu, Statistical confidence intervals for fuzzy data, Expert Systems with Applications, 36 (2009), 2670-2676.
27. Zadeh, L. Fuzzy Probabilities, Information Processing and Management 20 3 (1984), 363-372.