

Tests of Statistical Hypotheses with Respect to a Fuzzy Set

P. Pandian¹ & D. Kalpanapriya¹

¹ Department of Mathematics, School of Advanced Sciences, VIT University, Vellore-632 014, India

Correspondence: P. Pandian, Department of Mathematics, School of Advanced Sciences, VIT University, Vellore -632014, India. E-mail: pandian61@rediffmail.com

Received: October 31, 2013 Accepted: November 27, 2013 Online Published: December 17, 2013

doi:10.5539/mas.v8n1p35

URL: <http://dx.doi.org/10.5539/mas.v8n1p35>

Abstract

Tests of statistical hypotheses with crisp data using small samples are extended to with membership function of the fuzzy sets. The t-test statistic and the F-test statistic with respect to fuzzy sets are defined using the membership grades of the fuzzy sets. The rules for taking decision about the hypotheses are provided. In the proposed tests, the optimistic and pessimistic approach, h-level set, α -cut and fuzzy interval are not used. Numerical examples are provided for understanding the proposed testing procedures. The proposed tests of hypotheses may be useful to decision makers who are handling real life problems involving linguistic variables / fuzzy sets for taking suitable decisions in an acceptable manner.

Keywords: testing hypothesis, small sample, t-test, F-test, fuzzy set, membership function, confidence limits

1. Introduction

Statistical hypothesis testing is an applied statistical analysis in which inference of populations parameters are obtained using the numerical samples of the populations. The data analysts have interested to learn tests of statistical hypotheses for analyzing the population parameters. In conventional hypotheses testing (Devore (2008)), considering samples are crisp and the significance test leads to the binary decision. In real life situations, the sample data can not be recorded precisely always. So, imprecise data sample may be got for testing hypotheses. Many researchers (Arnold, 1998; Casals et al., 1986; Son et al., 1992; Saade & Schwarzlander, 1990; Saade, 1994; Casals & Gil, 1989, 1994) have proposed various tests of statistical hypotheses with imprecise samples. Using the fuzzy data sample, the fuzzy tests of hypotheses were discussed in Grzegorzewski (2000) and Watanabe and Imaizumi (1993). Niskanen (2001) studied statistical hypotheses in fields of human sciences. Wu (2005) developed hypotheses tests for fuzzy data using optimistic and pessimistic approach. Akbari and Rezaei (2009) proposed an approach to test the hypothesis about the variance using fuzzy data. For fuzzy data, hypotheses tests were studied by Viertl (2006, 2011) based on confidence intervals. Fuzzy confidence intervals for unknown fuzzy parameters were constructed by Wu (2009). For vague data, Arefi and Taheri (2011) tested the statistical fuzzy hypotheses. Based on confidence limits, the statistical hypotheses tests for fuzzy data is discussed by Chachi et al. (2012). The test of statistical hypothesis for comparing means with vague data was considered Baloui Jamkhaneh and Nadi Ghara (2010). Kalpanapriya and Pandian (2012, 2012) proposed tests of hypothesis for means of populations using imprecise samples.

In this paper, we propose four types of statistical hypothesis tests using small sample (or samples) based on the membership function (MF) of a fuzzy set (or fuzzy sets) namely, (i) testing of significance for difference of means of two populations with respect to a fuzzy set, (ii) testing of significance for difference of means of a population with respect to two fuzzy sets, (iii) to test the difference of variances of two populations with respect to a fuzzy set and (iv) to test the difference of variances of a population with respect to two fuzzy sets. The *t*-test statistic and *F*-test statistic are defined on the membership grades (MGs) of the fuzzy set over a random sample. The rules for taking decision about the hypotheses are provided. The optimistic and pessimistic approach, h-level set, α -cut and fuzzy interval are not used in the proposed tests. The procedures of the proposed tests of hypotheses are illustrated by means of numerical examples. The proposed tests of hypotheses can help decision makers in the linguistic hypotheses tests related issues of real life problems by aiding them in the decision making process and providing an appropriate decision rules in an acceptable manner.

2. Preliminaries

The following concepts related to fuzzy set and its MF are used which can be found in George J. Klir and Bo Yuan (2008) and Chiang and Lin (1999).

Let X and Y be two crisp set and let $\tilde{A}$ and $\tilde{B} \subseteq \mathcal{F}$ be fuzzy sets where $\mathcal{F}$ is a fuzzy space.

If a fuzzy set $\tilde{A}$ is defined on X with a MF $\mu_{\tilde{A}}(x)$, then $\tilde{A}$ can be represented as follows:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) / x \in X\}, \text{ where } \mu_{\tilde{A}} : X \rightarrow [0, 1].$$

If the two fuzzy sets $\tilde{A}$ and $\tilde{B}$ in $\mathcal{F}$ defined on X with MFs $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$, then, they are written as follows:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) / x \in X\} \text{ and } \tilde{B} = \{(x, \mu_{\tilde{B}}(x)) / x \in X\}, \text{ where } \mu_{\tilde{A}} \text{ and } \mu_{\tilde{B}} : X \rightarrow [0, 1].$$

If a fuzzy set $\tilde{A} \subseteq \mathcal{F}$ is defined on X and Y with MFs $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{A}}(y)$, then $\tilde{A}$ is written as follows:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) / x \in X\} \text{ and } \tilde{A} = \{(y, \mu_{\tilde{A}}(y)) / y \in Y\}, \text{ where } \mu_{\tilde{A}} : X \rightarrow [0, 1] \text{ and } \mu_{\tilde{A}} : Y \rightarrow [0, 1].$$

In Zadeh (1968), the probability of a fuzzy set $\tilde{A}$ defined on X with MF $\mu_{\tilde{A}}(x)$ is given by

$$P(\tilde{A}) = \int_X \mu_{\tilde{A}}(x) dP = E(\mu_{\tilde{A}}(x)) \quad (1)$$

where P is the probability measure over X . From (1), we can conclude that the probability of the happening of the fuzzy event $\tilde{A}$ is the expectation of $\mu_{\tilde{A}}(x)$ if the probability measure of X is known.

Now, we need the following definitions of the sample mean and the sample variance of the MGs of a fuzzy set which can be found in Chiang and Lin (1999).

Definition 2.1: Let $\{x_1, x_2, \dots, x_n\}$ be a random sample of size n from a crisp set X with the MGs of a fuzzy set $\tilde{A}$ where $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) / x \in X\}$. Then, the average MGs of fuzzy set $\tilde{A}$ over the random sample or the sample mean of the MF of the fuzzy $\tilde{A}$ denoted by $\bar{\mu}_{\tilde{A}}(x)$ is defined as follows:

$$\bar{\mu}_{\tilde{A}}(x) = \frac{1}{n} \left(\sum_{i=1}^n \mu_{\tilde{A}}(x_i) \right)$$

Definition 2.2: Let $\{x_1, x_2, \dots, x_n\}$ be a random sample of size n from a crisp set X with the MGs of a fuzzy set $\tilde{A}$ where $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) / x \in X\}$. Then, the variance of the MGs of fuzzy set $\tilde{A}$ over the random sample or the sample variance of the MF of the fuzzy $\tilde{A}$ denoted by $S_{\tilde{A}}^2(x)$ is defined as follows:

$$S_{\tilde{A}}^2(x) = \frac{1}{n-1} \left(\sum_{i=1}^n (\mu_{\tilde{A}}(x_i) - \bar{\mu}_{\tilde{A}}(x))^2 \right)$$

and the standard deviation of the MGs of fuzzy set $\tilde{A}$ over the random sample or the sample standard deviation of the MF of the fuzzy $\tilde{A}$, $S_{\tilde{A}}(x) = \sqrt{S_{\tilde{A}}^2(x)}$.

3. Testing of Significance for Difference of two Population Means with Respect to Fuzzy Sets

In this section, we propose the following two types of tests of statistical hypotheses:

- (i) Test for the difference of means of two populations using their small samples with respect to a fuzzy set.
- (ii) Test for the difference of means of a population using its small sample with respect to two fuzzy sets.

3.1 Testing of Significance for the Difference of two Population Means with Respect to a Fuzzy Set

Let X and Y be two crisp population and $\tilde{A}$ be a fuzzy set defined on X and Y . Let $\{x_1, x_2, \dots, x_m\}$ be a linguistic random sample of X with MGs $\mu_{\tilde{A}}(x_i)$, $i=1, 2, \dots, m$ and $\{y_1, y_2, \dots, y_n\}$ be another linguistic random sample of Y with MGs $\mu_{\tilde{A}}(y_j)$, $j=1, 2, \dots, n$ such that $m+n < 30$. Based on the samples, we test that the mean of the population X with respect to $\tilde{A}$, $\bar{\mu}(\tilde{A}, X)$ and the mean of the population Y with respect to $\tilde{A}$, $\bar{\mu}(\tilde{A}, Y)$ are the same.

Let the sample mean of the MF of $\tilde{A}$ defined on X be $\bar{\mu}_{\tilde{A}}(x)$, the sample mean of the MF of $\tilde{A}$ defined on Y be $\bar{\mu}_{\tilde{A}}(y)$, the sample variance of the MF of $\tilde{A}$ defined on X be $S_{\tilde{A}}^2(x)$ and the sample variance of the MF of $\tilde{A}$ defined on Y be $S_{\tilde{A}}^2(y)$.

Now, we have the null hypothesis (NH), $H_0 : \bar{\mu}(\tilde{A}, X) = \bar{\mu}(\tilde{A}, Y)$.

If both populations have same standard deviation with respect to $\tilde{A}$, we use the test statistic for testing the NH,

$$t = \frac{\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)}{S_{\tilde{A}} \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

$$\text{where } S_{\tilde{A}} = \sqrt{\frac{(m-1)S_{\tilde{A}}^2(x) + (n-1)S_{\tilde{A}}^2(y)}{m+n-2}}.$$

If both populations standard deviations with respect to $\tilde{A}$ are not the same, we use the test statistic for testing the H_0 ,

$$t = \frac{\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)}{\sqrt{\frac{S_{\tilde{A}}^2(x)}{m} + \frac{S_{\tilde{A}}^2(y)}{n}}}.$$

Now, the degrees of freedom (df) used in this test is $\nu = n + m - 2$. Let the level of significance (LOS) be α and let $t_{\alpha, \nu}$ denote the table value of t for ν df at α level.

Now, for α LOS, the critical region of the alternative hypothesis (AH), H_A is given below:

Alternative Hypothesis	Critical Region
$\bar{\mu}(\tilde{A}, X) > \bar{\mu}(\tilde{A}, Y)$ (upper tailed test)	$t \geq t_{\alpha, \nu}$
$\bar{\mu}(\tilde{A}, X) < \bar{\mu}(\tilde{A}, Y)$ (lower tailed test)	$t \leq -t_{\alpha, \nu}$
$\bar{\mu}(\tilde{A}, X) \neq \bar{\mu}(\tilde{A}, Y)$ (two tailed test)	$ t \geq t_{\alpha/2, \nu}$

If $|t| < t_{\alpha, \nu}$ (one tailed test), the difference between $\bar{\mu}(\tilde{A}, X)$ and $\bar{\mu}(\tilde{A}, Y)$ at α level is not significant. That is, $\bar{\mu}(\tilde{A}, X) = \bar{\mu}(\tilde{A}, Y)$ (the means of populations with respect to $\tilde{A}$ are identical) at α level. Therefore, the H_0 is accepted. Otherwise, the AH ($\bar{\mu}(\tilde{A}, X) > \bar{\mu}(\tilde{A}, Y)$ (for upper tailed test) or $\bar{\mu}(\tilde{A}, X) < \bar{\mu}(\tilde{A}, Y)$ (for lower tailed test)) is accepted.

If $|t| < t_{\alpha/2, \nu}$ (two tailed test), the difference between $\bar{\mu}(\tilde{A}, X)$ and $\bar{\mu}(\tilde{A}, Y)$ at α level is not significant. That is, $\bar{\mu}(\tilde{A}, X) = \bar{\mu}(\tilde{A}, Y)$ (the population means of the MF of $\tilde{A}$ are identical) at α level. Therefore, the H_0 is accepted. Otherwise, the AH, that is, $\bar{\mu}(\tilde{A}, X) \neq \bar{\mu}(\tilde{A}, Y)$ is accepted.

Now, the $100(1-\alpha)\%$ confidence limits for the difference of the population means of the MF of $\tilde{A}$, $\bar{\mu}(\tilde{A}, X)$ and $\bar{\mu}(\tilde{A}, Y)$ corresponding to the given samples are given below:

$$(\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)) - t_T \left(S_{\tilde{A}} \sqrt{\frac{1}{m} + \frac{1}{n}} \right) < \bar{\mu}(\tilde{A}, X) - \bar{\mu}(\tilde{A}, Y) < (\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)) + t_T \left(S_{\tilde{A}} \sqrt{\frac{1}{m} + \frac{1}{n}} \right)$$

(or)

$$(\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)) - t_T \left(\sqrt{\frac{S_{\tilde{A}}^2(x)}{m} + \frac{S_{\tilde{A}}^2(y)}{n}} \right) < \bar{\mu}(\tilde{A}, X) - \bar{\mu}(\tilde{A}, Y) < (\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)) + t_T \left(\sqrt{\frac{S_{\tilde{A}}^2(x)}{m} + \frac{S_{\tilde{A}}^2(y)}{n}} \right)$$

where $t_T = t_{\alpha/2, \nu}$.

Now, with the help of the numerical example given below, the procedure of the above said testing of hypothesis is explained.

Example 3.1: Let $X = \{\text{Students in a Government University}\}$ and $Y = \{\text{Students in a Private University}\}$ be the two populations. Let the fuzzy set, $\tilde{A} = \{\text{comfort ability}\}$ be defined on X and Y . Now, we are going to test that the $\tilde{A}$ in X is better than the $\tilde{A}$ in Y , that is, $\bar{\mu}(\tilde{A}, X) > \bar{\mu}(\tilde{A}, Y)$.

Let $S_1 = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ be the sample of size eight taken from the population X and $S_2 = \{y_1, y_2, y_3, y_4, y_5, y_6, y_7\}$ be the sample of size seven taken from the population Y .

Then, the membership grades of the given two samples based on their information concerning the fuzzy set $\tilde{A}$ are given below.

x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$\mu_{\tilde{A}}(x_i)$	0.91	0.85	0.82	0.79	0.89	0.76	0.81	0.78

and

y_j	y_1	y_2	y_3	y_4	y_5	y_6	y_7
$\mu_{\tilde{A}}(y_j)$	0.74	0.76	0.62	0.79	0.86	0.65	0.58

Now, the sample average membership grades of $\tilde{A}$ and the sample variance of $\tilde{A}$ over these two samples are $\bar{\mu}_{\tilde{A}}(x) = 0.826$; $\bar{\mu}_{\tilde{A}}(y) = 0.714$; $S_{\tilde{A}}^2(x) = 0.002827$ and $S_{\tilde{A}}^2(y) = 0.010129$.

$$\text{Now, } S_{\tilde{A}} = \sqrt{S_{\tilde{A}}^2} = \sqrt{\frac{n_1 S_{\tilde{A}}^2(x) + n_2 S_{\tilde{A}}^2(y)}{8 + 7 - 2}} = 0.085.$$

Now the NH, $H_o: \bar{\mu}_{\tilde{A}}(X) = \bar{\mu}_{\tilde{A}}(Y)$ against the AH, $H_A: \bar{\mu}_{\tilde{A}}(X) > \bar{\mu}_{\tilde{A}}(Y)$.

We take LOS, $\alpha = 5\%$ and the table value of t for 13 df at 5% LOS (one tailed test) is 2.16.

$$\text{Now, the test statistic, } t = \frac{\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{A}}(y)}{S_{\tilde{A}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{0.826 - 0.714}{0.085 \sqrt{\frac{1}{8} + \frac{1}{7}}} = 2.545 > 2.16.$$

Therefore, the NH is rejected and the AH is accepted. Thus, on the basis of the sample, at 5% LOS, the government university students are more comfortable than private university students.

3.2 Testing of Significance for the Difference of Means of a Population with Respect to two Fuzzy Sets

Let X be a crisp population and $\tilde{A}$ and $\tilde{B}$ be two fuzzy sets defined on X . Let $\{x_1, x_2, \dots, x_m\}$ be a linguistic random sample of size m from a normal population with MGs $\mu_{\tilde{A}}(x_i)$ and $\mu_{\tilde{B}}(x_i)$, $i=1, 2, \dots, m$ over $\tilde{A}$ and $\tilde{B}$ respectively. Based on the sample, we test that the mean of the population X related to MGs of $\tilde{A}$, $\bar{\mu}(\tilde{A}, X)$ and the mean of the population X related to MGs of $\tilde{B}$, $\bar{\mu}(\tilde{B}, X)$ are the same.

Let the sample means of the MF of $\tilde{A}$ and $\tilde{B}$ defined on X be $\bar{\mu}_{\tilde{A}}(x)$ and $\bar{\mu}_{\tilde{B}}(x)$ respectively and the sample variances of the MF of $\tilde{A}$ and $\tilde{B}$ defined on X be $S_{\tilde{A}}^2(x)$ and $S_{\tilde{B}}^2(x)$.

Now, we test the NH, $H_o: \bar{\mu}(\tilde{A}, X) = \bar{\mu}(\tilde{B}, X)$.

If population standard deviations with respect to two fuzzy sets are the same, we use the test statistic for testing the NH,

$$t = \frac{\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)}{S(x) \sqrt{\frac{2}{m}}}$$

$$\text{where } S(x) = \sqrt{\frac{(m-1)(S_{\tilde{A}}^2(x) + S_{\tilde{B}}^2(x))}{2m-2}}.$$

If population standard deviations with respect to two fuzzy sets are not the same, we use the test statistic for testing the NH,

$$t = \frac{\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)}{\sqrt{\frac{S_{\tilde{A}}^2(x) + S_{\tilde{B}}^2(x)}{m}}}.$$

Now, the df used in this test is $\nu = 2m - 2$.

Let the LOS be α and let $t_{\alpha, \nu}$ be the table value of t for ν df at α level.

Now, for α LOS, the critical regions of the AH, H_A for different types of tests are given below:

Alternative Hypothesis	Critical Region
$\bar{\mu}(\tilde{A}, X) > \bar{\mu}(\tilde{B}, X)$ (upper tailed test)	$t \geq t_{\alpha, v}$
$\bar{\mu}(\tilde{A}, X) < \bar{\mu}(\tilde{B}, X)$ (lower tailed test)	$t \leq -t_{\alpha, v}$
$\bar{\mu}(\tilde{A}, X) \neq \bar{\mu}(\tilde{B}, X)$ (two tailed test)	$ t \geq t_{\alpha/2, v}$

If $|t| < t_{\alpha, v}$ (one tailed test), the difference between $\bar{\mu}(\tilde{A}, X)$ and $\bar{\mu}(\tilde{B}, X)$ at α level is not significant. That is, the mean of the population X with respect to $\tilde{A}$ and the mean of the population X with respect to $\tilde{B}$ are the same ($\bar{\mu}(\tilde{A}, X) = \bar{\mu}(\tilde{B}, X)$) at α level. Therefore, the NH is accepted. Otherwise, the AH ($\bar{\mu}(\tilde{A}, X) > \bar{\mu}(\tilde{B}, X)$ (for upper tailed test) or $\bar{\mu}(\tilde{A}, X) < \bar{\mu}(\tilde{B}, X)$ (for lower tailed test)) is accepted.

If $|t| < t_{\alpha/2, v}$ (two tailed test), the difference between $\bar{\mu}(\tilde{A}, X)$ and $\bar{\mu}(\tilde{B}, X)$ at α level is not significant. That is, the means of the population X with respect to $\tilde{A}$ and $\tilde{B}$ are identical ($\bar{\mu}(\tilde{A}, X) = \bar{\mu}(\tilde{B}, X)$) at α level. Therefore, the NH is accepted. Otherwise, the AH, that is, $\bar{\mu}(\tilde{A}, X) \neq \bar{\mu}(\tilde{B}, X)$ is accepted.

Now, the $100(1 - \alpha)\%$ confidence limits for the difference of the population means $\bar{\mu}(\tilde{A}, X)$ and $\bar{\mu}(\tilde{B}, X)$ corresponding to the given samples are given below:

$$(\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)) - t_D \left(S \sqrt{\frac{2}{m}} \right) < \bar{\mu}(\tilde{A}, X) - \bar{\mu}(\tilde{B}, X) < (\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)) + t_D \left(S \sqrt{\frac{2}{m}} \right)$$

(or)

$$(\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)) - t_D \left(\sqrt{\frac{S_{\tilde{A}}^2(x) + S_{\tilde{B}}^2(x)}{m}} \right) < \bar{\mu}(\tilde{A}, X) - \bar{\mu}(\tilde{B}, X) < (\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)) + t_D \left(\sqrt{\frac{S_{\tilde{A}}^2(x) + S_{\tilde{B}}^2(x)}{m}} \right)$$

where $t_D = t_{\alpha/2, 2m-2}$.

Now, the numerical example given below is used to illustrate the above said hypothesis test procedure..

Example 3.2: Let $X = \{\text{Doctors in a city}\}$ be the population. Let the fuzzy sets, $\tilde{A} = \{\text{Compassionate}\}$ and $\tilde{B} = \{\text{Contentment}\}$ be defined on X. We are going to test that in the city, compassionate doctors and contentment doctors are the same, that is, $\bar{\mu}_{\tilde{A}}(X) = \bar{\mu}_{\tilde{B}}(X)$.

Let $S = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ be the sample of size ten taken randomly. The MGs of the given sample concerning fuzzy sets $\tilde{A}$ and $\tilde{B}$ based on the sample information are obtained as.

Doctor	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
$\mu_{\tilde{A}}(x_i)$	0.91	0.85	0.82	0.79	0.89	0.76	0.81	0.78	0.86	1.00
$\mu_{\tilde{B}}(x_i)$	0.74	0.76	0.62	0.79	0.72	0.82	0.58	0.72	0.67	0.45

Now, the sample averages MGs of $\tilde{A}$ and $\tilde{B}$ and the sample variances of the fuzzy set $\tilde{A}$ and $\tilde{B}$ over the sample are $\bar{\mu}_{\tilde{A}} = 0.847$; $\bar{\mu}_{\tilde{B}} = 0.687$; $S_{\tilde{A}}^2 = 0.005201$ and $S_{\tilde{B}}^2 = 0.012334$.

$$\text{Now, } S = \sqrt{S^2} = \sqrt{\frac{n_1 S_{\tilde{A}}^2 + n_2 S_{\tilde{B}}^2}{10 + 10 - 2}} = 0.0987.$$

Now, the NH, $H_0 : \bar{\mu}_{\tilde{A}}(X) = \bar{\mu}_{\tilde{B}}(X)$ and the AH, $H_A : \bar{\mu}_{\tilde{A}}(X) \neq \bar{\mu}_{\tilde{B}}(X)$.

Now, we take LOS, $\alpha = 5\%$. and the table value of t for 18 df at 5% LOS (two tailed test) is 2.101.

$$\text{Now, the test statistic, } t = \frac{\bar{\mu}_{\tilde{A}}(x) - \bar{\mu}_{\tilde{B}}(x)}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{0.847 - 0.687}{0.0987 \sqrt{\frac{1}{10} + \frac{1}{10}}} = 3.636 > 2.101.$$

Therefore, NH is rejected and the AH is accepted, that is, the compassionate doctors need not be contentment in the city at 5% LOS.

4. Testing of Significance for Difference of two Variances with Respect to Fuzzy Sets

The following two types of tests of hypotheses are discussed in this section:

- (i) To test the difference of variances of two populations using their small samples with respect to a fuzzy set.
- (ii) To test the difference of variances of one population using its small samples with respect to two fuzzy sets.

4.1 To Test the Difference of Variances of two Populations with Respect to one Fuzzy Set

Let X and Y be two populations with variances σ_1^2 and σ_2^2 respectively. Let $\tilde{A}$ be a fuzzy set defined on X and Y . Let $S_1 = \{x_1, x_2, \dots, x_m\}$ be a random sample of size m from the population X with MF $\mu_{\tilde{A}}(x)$ and $S_2 = \{y_1, y_2, \dots, y_n\}$ be a random sample of size n from the population Y with MF $\mu_{\tilde{A}}(y)$ such that $m+n < 30$. Let $s_{\tilde{A}}^2(x)$ and $s_{\tilde{A}}^2(y)$ be sample variances of X and Y with respect to $\tilde{A}$.

Now, the test statistic

$$F = \frac{\left(\frac{s_{\tilde{A}}^2(x)}{s_{\tilde{A}}^2(y)} \right)}{\left(\frac{\sigma_{\tilde{A}}^2(X)}{\sigma_{\tilde{A}}^2(Y)} \right)} \quad (\text{or}) \quad F = \frac{\left(\frac{s_{\tilde{A}}^2(y)}{s_{\tilde{A}}^2(x)} \right)}{\left(\frac{\sigma_{\tilde{A}}^2(Y)}{\sigma_{\tilde{A}}^2(X)} \right)}.$$

Now, we have to test the hypothesis that the variance of X and the variance of Y are the same with respect to $\tilde{A}$, that is, the $H_0: \sigma_{\tilde{A}}^2(X) = \sigma_{\tilde{A}}^2(Y)$.

Now, the test statistic for H_0 ,

$$F = \begin{cases} \frac{s_{\tilde{A}}^2(x)}{s_{\tilde{A}}^2(y)} \text{ with } (v_1, v_2) \text{ df if } s_{\tilde{A}}^2(x) \geq s_{\tilde{A}}^2(y). \\ \frac{s_{\tilde{A}}^2(y)}{s_{\tilde{A}}^2(x)} \text{ with } (v_2, v_1) \text{ df if } s_{\tilde{A}}^2(x) < s_{\tilde{A}}^2(y). \end{cases}$$

where $v_1 = m-1$ and $v_2 = n-1$.

Let the LOS be α .

Now, the critical region of the AH, H_A for α LOS is given below:

Alternative hypothesis	Rejection region
$\sigma_{\tilde{A}}^2(X) > \sigma_{\tilde{A}}^2(Y)$ (one tailed test)	$F = \frac{s_{\tilde{A}}^2(x)}{s_{\tilde{A}}^2(y)} \geq F_{\alpha, (m-1, n-1)}$
$\sigma_{\tilde{A}}^2(X) < \sigma_{\tilde{A}}^2(Y)$ (one tailed test)	$F = \frac{s_{\tilde{A}}^2(y)}{s_{\tilde{A}}^2(x)} \geq F_{\alpha, (n-1, m-1)}$
$\sigma_{\tilde{A}}^2(x) \neq \sigma_{\tilde{A}}^2(y)$ (two tailed test)	$F = \frac{s_{\tilde{A}}^2(x)}{s_{\tilde{A}}^2(y)} \geq F_{\alpha/2, (m-1, n-1)}$ or $F = \frac{s_{\tilde{A}}^2(y)}{s_{\tilde{A}}^2(x)} \geq F_{\alpha/2, (n-1, m-1)}$

If $F = \frac{s_{\tilde{A}}^2(x)}{s_{\tilde{A}}^2(y)} < F_{\alpha, (m-1, n-1)}$ (one tailed test), the difference between the variances of X and Y with respect to $\tilde{A}$ at

α level is not significant. That is, the population variances with respect to $\tilde{A}$ are identical ($\sigma_{\tilde{A}}^2(X) = \sigma_{\tilde{A}}^2(Y)$) at α level. Therefore, the H_0 is accepted. Otherwise, the H_A , that is, $\sigma_{\tilde{A}}^2(X) > \sigma_{\tilde{A}}^2(Y)$, is accepted.

If $F = \frac{s_{\tilde{A}}^2(y)}{s_{\tilde{A}}^2(x)} < F_{\alpha/(n-1, m-1)}$ (one tailed test), the difference between the variances of X and Y with respect to $\tilde{A}$

at α level is not significant. That is, the population variances with respect to $\tilde{A}$ are identical ($\sigma_{\tilde{A}}^2(X) = \sigma_{\tilde{A}}^2(Y)$) at α level. Therefore, the NH is accepted. Otherwise, the AH, that is, $\sigma_{\tilde{A}}^2(X) < \sigma_{\tilde{A}}^2(Y)$, is accepted.

If $F = \frac{s_{\tilde{A}}^2(y)}{s_{\tilde{A}}^2(x)} < F_{\alpha/2, (n-1, m-1)}$ or $F = \frac{s_{\tilde{A}}^2(x)}{s_{\tilde{A}}^2(y)} < F_{\alpha/2, (m-1, n-1)}$ (two tailed test), the difference between the variances

of the population X and Y with respect to $\tilde{A}$ at α level is not significant. That is, the variance of the population X with respect to $\tilde{A}$ and the variance of the population Y with respect to $\tilde{A}$ are the same ($\sigma_{\tilde{A}}^2(X) = \sigma_{\tilde{A}}^2(Y)$) at α level. Therefore, the NH is accepted. Otherwise, the AH, that is, $\sigma_{\tilde{A}}^2(X) \neq \sigma_{\tilde{A}}^2(Y)$, is accepted.

Now, the $100(1 - \alpha)\%$ confidence limits for the quotient of variances $\sigma_{\tilde{A}}^2(X)$ and $\sigma_{\tilde{A}}^2(Y)$ corresponding to the given samples are given below:

$$\frac{\left(\frac{S_{\tilde{A}}^2(x)}{S_{\tilde{A}}^2(y)} \right)}{F_{\alpha/2, (m-1, n-1)}} < \frac{\sigma_{\tilde{A}}^2(X)}{\sigma_{\tilde{A}}^2(Y)} < F_{\alpha/2, (n-1, m-1)} \left(\frac{S_{\tilde{A}}^2(x)}{S_{\tilde{A}}^2(y)} \right)$$

(or)

$$\frac{\left(\frac{S_{\tilde{A}}^2(y)}{S_{\tilde{A}}^2(x)} \right)}{F_{\alpha/2, (n-1, m-1)}} < \frac{\sigma_{\tilde{A}}^2(Y)}{\sigma_{\tilde{A}}^2(X)} < F_{\alpha/2, (m-1, n-1)} \left(\frac{S_{\tilde{A}}^2(y)}{S_{\tilde{A}}^2(x)} \right)$$

Now, we explain the procedure of the above said hypothesis test with a numerical example.

Example 4.1: Let $X = \{\text{All girls in a city}\}$ and $Y = \{\text{All girls in a town}\}$. Now, we test that the variability of the prettiness among girls in both places are the same.

Now, a sample of six girls was taken at random from the city $(x_1, x_2, x_3, x_4, x_5, x_6) = (\text{Mary, Judy, Linda, Susan, Betty, Julia})$ and a sample of six girls was taken at random from the town $(y_1, y_2, y_3, y_4, y_5, y_6) = (\text{Maya, Jasmine, Latha, Shela, Bindu, Jaya})$.

Now, let us define a fuzzy set over the crisp set X and the crisp set Y , $\tilde{A} = \{\text{Pretty girl}\}$.

Now, the membership grades of these two sets of six girls concerning the fuzzy set $\tilde{A}$ based on the known information are obtained as

x	Mary	Judy	Linda	Susan	Betty	Julia
$\mu_{\tilde{A}}(x)$	0.79	0.96	0.65	0.84	1.00	0.88

and

y	Maya	Jasmine	Latha	Shela	Bindu	Jaya
$\mu_{\tilde{A}}(y)$	0.70	0.98	0.60	0.90	0.83	0.86

Now, the sample averages of $\tilde{A}$ are $\bar{\mu}_{\tilde{A}}(x) = 0.85$ and $\bar{\mu}_{\tilde{A}}(y) = 0.81$.

Now, the sample variances of $\tilde{A}$ are given $S_{\tilde{A}}^2(x) = 0.0158$ and $S_{\tilde{A}}^2(y) = 0.0192$.

Now, the NH, $\sigma_{\tilde{A}}^2(X) = \sigma_{\tilde{A}}^2(Y)$ against the AH, $\sigma_{\tilde{A}}^2(X) \neq \sigma_{\tilde{A}}^2(Y)$.

Now, the test statistic: $F = \frac{S_{\tilde{A}}^2(y)}{S_{\tilde{A}}^2(x)}$.

We take $\alpha = 5\%$ and the table value of F for $v=(5,5)$ at 5% level (two tailed test) = 5.05.

$$\text{Now, } F = \frac{S_A^2(y)}{S_A^2(x)} = \frac{0.0192}{0.0158} = 1.22 < 5.05.$$

Therefore, the NH is accepted at 5% level.

Thus, city girls and town girls have the same degrees of variations of the prettiness, according to the random samples of girls.

4.2 To Test Variances of one Population with Respect to two Fuzzy Sets

Let X be a population with variance σ^2 . Let two fuzzy sets $\tilde{A}$ and $\tilde{B}$ be defined on X . Let $S = \{x_1, x_2, \dots, x_m\}$ be a random sample of the population X of the size $m < 30$ with MFs $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$. Let $s_A^2(x)$ and $s_B^2(x)$ be sample variances of X with respect to $\tilde{A}$ and $\tilde{B}$.

Now, the test statistic

$$F = \frac{\left(\frac{s_A^2(x)}{s_B^2(x)} \right)}{\left(\frac{\sigma_A^2(X)}{\sigma_B^2(X)} \right)} \quad (\text{or}) \quad F = \frac{\left(\frac{s_B^2(x)}{s_A^2(x)} \right)}{\left(\frac{\sigma_B^2(X)}{\sigma_A^2(X)} \right)}.$$

Now, we have to test the hypothesis that the variances of the population X with respect to $\tilde{A}$ and $\tilde{B}$ are the same, that is, the NH, $H_o: \sigma_A^2(X) = \sigma_B^2(X)$. Then, the test statistic for H_o ,

$$F = \begin{cases} \frac{s_A^2(x)}{s_B^2(x)} \text{ with } (v_1, v_2) \text{ df if } s_A^2(x) \geq s_B^2(x). \\ \frac{s_B^2(x)}{s_A^2(x)} \text{ with } (v_2, v_1) \text{ df if } s_A^2(x) < s_B^2(x). \end{cases}$$

where $v_1 = m-1$ and $v_2 = m-1$.

Let the LOS be α .

Now, for α LOS, the critical regions of the AH, H_A of various types of tests are given below:

Alternative hypothesis	Critical region
$\sigma_A^2(X) > \sigma_B^2(X)$ (one tailed test)	$F = \frac{s_A^2(x)}{s_B^2(x)} \geq F_{\alpha, (m-1, m-1)}$
$\sigma_A^2(X) < \sigma_B^2(X)$ (one tailed test)	$F = \frac{s_B^2(x)}{s_A^2(x)} \geq F_{\alpha, (m-1, m-1)}$
$\sigma_A^2(X) \neq \sigma_B^2(X)$ (two tailed test)	$F = \frac{s_A^2(x)}{s_B^2(x)} \geq F_{\alpha/2, (m-1, m-1)}$ or $F = \frac{s_B^2(x)}{s_A^2(x)} \geq F_{\alpha/2, (m-1, m-1)}$

If $F = \frac{s_A^2(x)}{s_B^2(x)} < F_{\alpha, (m-1, m-1)}$ (one tailed test), the difference between the variances of X with respect to $\tilde{A}$ and $\tilde{B}$ at α level is not significant. That, the population variances with respect to $\tilde{A}$ and $\tilde{B}$ are identical

$(\sigma_A^2(X) = \sigma_B^2(X))$ at α level. Therefore, the NH is accepted. Otherwise, the AH, that is, $\sigma_A^2(X) > \sigma_B^2(X)$, is accepted.

If $F = \frac{s_B^2(x)}{s_A^2(x)} < F_{\alpha/(m-1, m-1)}$ (one tailed test), the difference between the variances X with respect to $\tilde{A}$ and $\tilde{B}$

at α level is not significant. That is, the population variances with respect to $\tilde{A}$ and $\tilde{B}$ are identical $(\sigma_A^2(X) = \sigma_B^2(X))$ at α level. Therefore, the NH is accepted. Otherwise, the AH, that is, $\sigma_A^2(X) < \sigma_B^2(X)$, is accepted.

If $F = \frac{s_B^2(x)}{s_A^2(x)} < F_{\alpha/2, (m-1, m-1)}$ or $F = \frac{s_A^2(x)}{s_B^2(x)} < F_{\alpha/2, (m-1, m-1)}$ (two tailed test), the difference between the

variances of the population X with respect to $\tilde{A}$ and $\tilde{B}$ at α level is not significant. That is, the variance of the population X with respect to $\tilde{A}$ and the variance of the population X with respect to $\tilde{B}$ are the same $(\sigma_A^2(X) = \sigma_B^2(X))$ at α level. Therefore, the NH is accepted. Otherwise, the AH, that is, $\sigma_A^2(X) \neq \sigma_B^2(X)$, is accepted.

Now, the $100(1-\alpha)\%$ confidence limits for the quotient of variances $\sigma_A^2(X)$ and $\sigma_B^2(X)$ corresponding to the given samples are given below:

$$\frac{\left(\frac{S_A^2(x)}{S_B^2(x)} \right)}{F_{\alpha/2, (m-1, m-1)}} < \frac{\sigma_A^2(X)}{\sigma_B^2(X)} < F_{\alpha/2, (m-1, m-1)} \left(\frac{S_A^2(x)}{S_B^2(x)} \right)$$

(or)

$$\frac{\left(\frac{S_B^2(x)}{S_A^2(x)} \right)}{F_{\alpha/2, (m-1, m-1)}} < \frac{\sigma_B^2(X)}{\sigma_A^2(X)} < F_{\alpha/2, (m-1, m-1)} \left(\frac{S_B^2(x)}{S_A^2(x)} \right)$$

Now, we illustrate the procedure of the above said hypothesis test with help of a numerical example.

Example 4.2: Let $X = \{\text{Faculties in a University}\}$ be the population. Let $S = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ be the random sample of X . Let the fuzzy sets $\tilde{A} = \{\text{Punctuality}\}$ and $\tilde{B} = \{\text{Availability}\}$ be defined on X . Now, the MGs of the sample concerning fuzzy sets $\tilde{A}$ and $\tilde{B}$ based on the collected information are given below.

Faculty	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$\mu_{\tilde{A}}(x_i)$	0.96	0.85	0.82	0.79	0.89	0.76	0.81	0.88
$\mu_{\tilde{B}}(x_i)$	0.74	0.76	0.62	0.79	0.64	0.82	0.58	0.72

We test the hypothesis that faculties in the University have no variability related to the behaviour of punctuality and availability.

Now, the NH, $H_0 : \sigma_A^2(X) = \sigma_B^2(X)$ against the AH, $H_A : \sigma_A^2(X) \neq \sigma_B^2(X)$

Now, we take $\alpha = 5\%$ and for $(7,7)$ df at 5% level of significance, the table value of F is 3.78.

Now, the sample average MGs of $\tilde{A}$ and $\tilde{B}$ and the sample variance of $\tilde{A}$ and $\tilde{B}$ over the sample are $\bar{\mu}_{\tilde{A}} = 0.845$; $\bar{\mu}_{\tilde{B}} = 0.709$, $S_{\tilde{A}}^2 = 0.004086$ and $S_{\tilde{B}}^2 = 0.007413$.

Now, the test statistic, $F = \frac{S_{\tilde{B}}^2}{S_{\tilde{A}}^2} = \frac{0.007413}{0.004086} = 1.814 < 3.78$.

Therefore, we accept the NH, that is, faculties in the University have no variability related to the behaviour of punctuality and availability.

5. Conclusion

Four types of tests of statistical hypotheses based on the MF of fuzzy sets which are totally different from conventional statistical hypothesis testing are proposed in this article. In the proposed tests of hypotheses, the differences of means and variances of the populations are studied with the help of fuzzy sets and small samples of the populations. The rules for decision taken about the hypotheses are provided. We can easily observe that the each proposed test of statistical hypothesis is a characteristic or attribute based test on the population. The optimistic and pessimistic approach, h-level set, α -cut and fuzzy interval are not used in the proposed hypotheses tests. The proposed statistical hypotheses tests can help decision makers in tests of hypotheses related issues of real life problems for choosing an appropriate decision with satisfaction.

Acknowledgements

The authors thank to Dr. M. A. M. Sahul Hameed, Professor in English, VIT University, Vellore, India for checking the language and the presentation in this article.

References

- Akbari, M. G., & Rezaei, A. (2009). Bootstrap statistical inference for the variance based on fuzzy data. *Austrian Journal of Statistics*, 38, 121-130.
- Arefi, M., & Taheri, S. M. (2011). Testing fuzzy hypotheses using fuzzy data based on fuzzy test statistic. *Journal of Uncertain Systems*, 5, 45-61. <http://dx.doi.org/10.1007/S00500-008-0339-3>.
- Arnold, B. F. (1998). Testing fuzzy hypotheses with crisp data. *Fuzzy Sets and Systems*, 94, 323-333. [http://dx.doi.org/PII S0165-01 14\(96\)00258-8](http://dx.doi.org/PII S0165-01 14(96)00258-8).
- Baloui, J. E., & Nadi, G. A. (2010). Testing statistical hypotheses for compare means with vague data. *International Mathematical Forum*, 5, 615-620.
- Casals, M. R., Gil, M. A., & Gil, P. (1986). The Fuzzy decision problem: an approach to the problem of testing statistical hypotheses with fuzzy information. *European Journal of Operational Research*, 27, 371-382. [http://dx.doi.org/10.1016/0377-2217\(86\)90333-4](http://dx.doi.org/10.1016/0377-2217(86)90333-4)
- Casals, M. R., & Gil, M. A. (1989). A note on the operativeness of Neyman-Pearson tests with fuzzy information. *Fuzzy Sets and Systems*, 30, 215-220. [http://dx.doi.org/10.1016/0165-0114\(89\)90082-1](http://dx.doi.org/10.1016/0165-0114(89)90082-1)
- Casals, M. R., & Gil, P. (1994). Bayesian sequential test for fuzzy parametric hypotheses from fuzzy information. *Information Sciences*, 80, 283-298. [http://dx.doi.org/10.1016/0020-0255\(94\)90080-9](http://dx.doi.org/10.1016/0020-0255(94)90080-9)
- Chachi, J., Taheri, S. M., & Viertl, R. (2012). *Testing statistical hypotheses based on fuzzy confidence intervals*. Forschungsbericht SM-2012-2, Technische Universitat Wien, Austrisa.
- Chiang, D. A., & Lin, N. P. (1999). Correlation of fuzzy sets. *Fuzzy Sets and Systems*, 2, 221-226.
- Devore, J. L. (2008). *Probability and Statistics for Engineers*. Cengage.
- George, J. K., & Bo Yuan. (2008). *Fuzzy Sets and Fuzzy Logic, Theory and Applications*. Prentice-Hall, New Jersey.
- Grzegorzewski, P. (2000). Testing statistical hypotheses with vague data. *Fuzzy Sets and Systems*, 112, 501-510. [http://dx.doi.org/10.1016/S0165-0114\(98\)00061-X](http://dx.doi.org/10.1016/S0165-0114(98)00061-X)
- Kalpanapriya, D., & Pandian, P. (2012). Two-sample statistical hypothesis test for means with imprecise data. *International Journal of Engineering Research and Applications*, 2, 3210-3217.
- Kalpanapriya, D., & Pandian, P. (2012). Statistical hypotheses testing with imprecise data. *Applied Mathematical Sciences*, 6, 5285-5292.
- Niskanen, V. A. (2001). Prospects for soft statistical computing: describing data and inferences from data with words in the human sciences. *Information Sciences*, 132, 83-131. [http://dx.doi.org/10.1016/S0020-0255\(01\)00060-3](http://dx.doi.org/10.1016/S0020-0255(01)00060-3)
- Saade, J. J., & Schwarzlander, H. (1990). Fuzzy hypothesis testing with hybrid data. *Fuzzy Sets and Systems*, 35, 197-212. [http://dx.doi.org/10.1016/0165-0114\(90\)90193-A](http://dx.doi.org/10.1016/0165-0114(90)90193-A)
- Saade, J. J. (1994). Extension of fuzzy hypothesis testing with hybrid data. *Fuzzy Sets and Systems*, 63, 57-71. [http://dx.doi.org/10.1016/0165-0114\(94\)90145-7](http://dx.doi.org/10.1016/0165-0114(94)90145-7)
- Son, J. Ch., Song, I., & Kim, H.Y. (1992). A fuzzy decision problem based on the generalized Neymen-Pearson criterion. *Fuzzy Sets and Systems*, 47, 65-75.

- Viertl, R. (2006). Univariate statistical analysis with fuzzy data. *Computational Statistics and Data Analysis*, 51, 33-147. <http://dx.doi.org/10.1016/j.csda.2006.04.002>
- Viertl, R. (2011). *Statistical methods for fuzzy data*. John Wiley and Sons, Chichester. <http://dx.doi.org/10.1002/9780470974414.ch9>
- Watanabe, N., & Imaizumi, T. (1993). A fuzzy statistical test of fuzzy hypotheses. *Fuzzy Sets and Systems*, 53, 167-178.
- Wu, H. C. (2005). Statistical hypotheses testing for fuzzy data. *Information Sciences*, 175, 30-56. <http://dx.doi.org/10.1016/j.ins.2003.12.009>.
- Wu, H. C. (2009). Statistical confidence intervals for fuzzy data. *Expert Systems with Applications*, 36, 2670-267. <http://dx.doi.org/10.1016/j.eswa.2008.01.022>
- Zadeh, L. A. (1968). Probability measures of fuzzy events. *J. Math. Anal. Appl.*, 23, 421-427.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).