A STATISTICAL ANALYSIS OF VARIANCE FOR ONE-WAY CLASSIFICATION USING TRAPEZOIDAL FUZZY TECHNIQUES AND RANKING METHODS

Vanitha. R[1], Pachamuthu. M[1*], Kirthik Vairamariappan. A[2] , Kokkilambigai. S[1] and Sona. S[1]

•

[1] Department of Statistics, Periyar University, India.
[2]Department of Statistics, Government Arts and Science College, India.

vanithapai50@gmail.com, pachamuthu@periyaruniversity.ac.in

Abstract

The Analysis of Variance (ANOVA) is a powerful tool for testing significance, with its primary purpose being to assess the homogeneity of several means. Variation is inherent in nature. The total variation in any set of numerical data arises from multiple causes, which are classified as assignable and chance causes. The variation due to assignable causes can be detected and measured, whereas variation due to chance is beyond human control. An ambiguous number is a normal number whose exact value is uncertain. Fuzziness occurs when the boundaries of information are unclear or imprecise. Sometimes, observational data may not be recorded, particularly during natural disasters. This paper presents a statistical analysis of variance for one-way classification using trapezoidal fuzzy techniques and the ranking method, demonstrated through a numerical example.

Keywords: *ANOVA*, Fuzzy Sets, Trapezoidal Fuzzy Numbers, Ranking Method, Decision Rule.

I. Introduction

The *ANOVA* is a powerful statistical tool for tests of significance. The term *ANOVA* was introduced by Prof. R. A. Fisher in 1920's to deal with problems in the analysis of agrochemical data. Variance is inherent in nature. The total variation in any set of numerical data is due to a number of causes which may be classified as assignable and chance causes. The variation due to assignable cause can be detected and measured whereas the variation due to assignable causes is beyond the control of human hand and cannot be traced separately. In general, *ANOVA* studies mainly the homogeneity of populations by separating the total variance into its various components. That is, this technique is to test the difference among the means of populations by studying the amount of variation within each of the samples relative to the amount of variance between the samples. Samples under employing in *ANOVA* model are assumed to be drawn from 'normal populations of equal variances. The variation of each value around its own grand mean should be independent for each value. The *ANOVA* is a powerful statistical tool for test of significance. The test of significance based on t - distribution is an adequate procedure only for testing the significance of the difference between two sample means. In a situation we have to study more than two samples to consider at a time an alternative producer is needed for testing the hypothesis that all the samples are drawn from same population. For example, six fertilizers are applied to five plots are given. We may be interested in finding out whether the effect of these fertilizers on the yield is significant different. In other words, test whether the samples have same population. The answer to this problem is provided by technique of *ANOVA*. In sometimes, the observations data is not recorded during a natural disaster. Therefore, fuzzy analysis is almost inevitable. The fuzzy logic was introduced in 1965 by Lofti A. Zadeh [15] in his research paper *Fuzzy Sets*. He is considered as the father of fuzzy logic. The theory of fuzzy logic Zadeh [15] gives an inference method under cognitive uncertainty, computational neural networks and fuzzy logic offer exciting benefits such as learning adaptation, fault tolerance, parallelism and generalization. Generally, our impression of the real world is pervaded by concepts that do not have precisely defined boundaries. For example; many, tall, much, larger than, smaller than, superior to, inferior to, younger than etc. And these adjectives are true only to some degree and they are false to some extent as well. These concepts can be called fuzzy or vague concepts; a human brain works with them, while computers may not do it because of the reason with a string of 0's (false) and 1's (true). Imprecision can be associated with quantitative data and qualitative data. In traditional hypothesis testing, the sample observations are crisp and a statistical test leads to a binary decision is either accept or reject the null hypothesis. However, in real life the data sometimes cannot be recorded or collected precisely in these situations fuzzy logic is most inevitable to the statistical analysis of hypothesis testing. In 1965s Zadeh [15] was introduced fuzzy sets as an extension of the set with classical notations. Fuzzy sets theory allows the estimated membership function in intervals belongs to [0, 1]. In the first case, due to technical problems, the response variable cannot be measured properly. So, in this case, the data cannot be clearly recorded with the exact numbers and the measurement errors are computed linguistically to justify the required tolerance. The second phenomenon is that the response variable is presented in terms of linguistic forms such as a special linguistic or variance reports. As for his products, they are not counted. In both of the above cases, the observed variable of the fuzziness often occurs the data can be represented by the concept of fuzzy sets for analyzing the test (Zadeh [16]. Liou and Wang has discussed the ranked fuzzy numbers with total integral value [2], [3] and [6]. Wang et.al. has presented the method for centroid formulae for a generalized fuzzy number [10]. Iuliana Carmen Barbacioru have discussed the statistical hypotheses testing using membership fuzzy numbers [5]. Salim Rezvani have analyzed the ranking function with trapezoidal fuzzy numbers [11],[12] and [13]. Wang have derived some different approach for ranking trapezoidal fuzzy numbers [7],[8] and [10]. Thorani et al. has discussed the ranking function of a trapezoidal fuzzy number with some modifications [9]. S. Abbas bandy and B. Asady has derived the nearest trapezoidal fuzzy number to a fuzzy quantity [1]. S. Parthiban and B. Gajiwardhan has proposed a one factor ANOVA model using trapezoidal fuzzy numbers by using α-cut interval method [4] and [14]. S. Chandrasekaran and A. Tamilmani has proposed the arithmetic operations of fuzzy numbers using alpha cut interval method [17]. H.C. Wu [18] has proposed the technique, ambiguous data as well, given the vague assumptions of the tests were imprecise data, along with two hypothesis tests replacing *ANOVA* models crisp data, models and results, getting after using the results obtained in terms of the provisions of the proposed decision of the population receive the decision. M. Mashinchi et.al. [19] has derived the *ANOVA* based on fuzzy observations and extended into a one-way *ANOVA*, which would be fuzzy numbers rather than real numbers Wang arrived some different approach for ranking TZFNs [10]. In this paper, propose a statistical analysis of variance for one-way classification with trapezoidal fuzzy techniques using ranking method.

II. Methods

I. Fuzzy Set

The fuzzy set theory was developed to handle the vagueness inherent in nature. Zadeh (1965) [15] extended the notion of binary membership to accommodate various degree of membership on the real continuous interval, where the endpoints of 0 and 1 refers to no membership and full membership respectively just as the indicator function does for crisp sets, but where the infinite number of values in between the endpoints 0 and 1 can represent various degrees of membership for an element in some set in the universe of discourse. The sets in the universe of discourse that can accommodate degrees of membership were termed by Zadeh (1965) [15] as fuzzy sets.

II. α-Cut

A fuzzy set A in X and any actual number $\alpha \in [0,1]$, then the α-cut α-level or cut worthy set of $\tilde{A}$ denoted by A_α is the crisp set $A_\alpha = \{x \in X: \mu_A(x) \geq \alpha\}$. The strong α-cut $\tilde{A}_{\alpha+}$ is the crisp set $A_{\alpha_+} = \{x \in X: \mu_A(x) > \alpha\}$. For example, let A be a fuzzy set whose membership function is

$$\mu_A(x) = \begin{cases} \dfrac{x-a}{b-a}; & a \leq x \leq b \\ \dfrac{c-x}{c-b}; & b \leq x \leq c \end{cases}$$

To find the α-cut of A, first set $\alpha \in [0,1]$ to both left and right reference functions of A. That is $\alpha = \frac{x-a}{b-a}$ and $\alpha = \frac{c-x}{c-b}$. Expressing x in terms of α, have $x = a + (b-a)\alpha$ and $x = c - (c-b)\alpha$, which gives the α-cut of A is $A_\alpha = [a + (b-a)\alpha;\ c - (c-b)\alpha]$. Similarly, the *TZFN*s to both left and right reference functions of A.

III. Crisp Set

The support of a fuzzy set A defined on X is a crisp set defined as $\text{Supp}(A) = \{x \in X: \mu_A(x) > 0\}$

IV. h-level of a Fuzzy Set

The h-level of a fuzzy set $\tilde{A}$, denoted by $h(\tilde{A})$ is the supremum (maximum) of the membership grades obtained by any element in the set $h(\tilde{A}) = \sup_{\text{xix}} \mu_{\tilde{A}}(x)$.

V. Degrees of Membership in Fuzzy Logic

Let us consider the fuzzy set $A, A = \{(x, \mu_A(x)) \mid x \in X\}$ where $\mu_A(x)$ is called the membership function for the fuzzy set $A. X$ is referred to as the universe of discourse. The membership function associates each element $x \in X$ with a value in the interval [0,1] . In fuzzy sets, each element is mapped to [0,1] by membership function. That is, $\mu_A: X \in [0,1]$, where [0,1] means real numbers between 0 and 1 (including 0,1). Consequently, fuzzy set is with vague boundary set comparing with crisp set. The fuzzy set A can be alternatively denoted as follows: If X is discrete then $A = \sum \mu_A(x_i)/x_i$ and if X is continnous then $A = \int \mu_A(x)/x$. Here, $\mu_A(x)$ is the membership function. Value of this function is between 0 and 1. This value represents the degree of membership (membership value) of element x inset A. The members of a fuzzy set are members to some degree, known as a membership grade or degree of membership. The membership grade is the degree of belonging to the fuzzy set. The larger the number (in [0,1]) the more the degree of belonging. (This is not a probability). The translation from x to $\mu_A(x)$ is known as fuzzification. In the fuzzy theory, fuzzy set A of universe X is defined by a function $\mu_A(x)$ called the membership function of set A. An already discussed this point. $\mu_A(x): X \to [0,1]$, where $\mu_A(x) = 1$ it x is totally in $A; \mu_A(x) = 0$ it x is not in A; $0 < \mu_A(x) < 1$ it x is partly in A. This set allows a continutum of possible choices. For any element x of universe X, membership function $\mu_A(x)$ equals the degree to which x is an element of set A. This degree, a value between 0 and 1, represents the degree of membership, also called membership value, of element x inset A.

VI. Generalized Triangular Fuzzy Numbers (*GTFN*s)

A Generalized Triangular Fuzzy Numbers (*GTFN*s) $\tilde{A} = (a, b, c; w)$ is described as any fuzzy subset of the real line $\mathbb{R}$, whose membership function $\mu_A(x)$ satisfies the following conditions:

- $\mu_A(x)$ is a continuous mapping from $\mathbb{R}$ to the closed interval $[0, w]$; $0 \leq w \leq 1$
- $\mu_A(x) = 0$, for all $x \in (-\infty, a]$
- $\mu_L(x) = L_A(x)$ is strictly increasing on $[a, b]$
- $\mu_A(x) = w$, for all $[b], w$ is a constant and $0 \leq w \leq 1$
- $\mu_R(x) = R_{\dot{A}}(x)$ is strictly decreasing on $[b, c]$
- $\mu_A(x) = 0$, for all $x \in [c, \infty]$; a, b, c are real numbers such that $a \leq b \leq c$.

A *GTFN*s is denoted by $\tilde{A} = (a, b, c; w)$ and is defined with the membership function,

$$\mu_{\hat{A}}(x) = \begin{cases} 0 & ; \quad x \leq a \\ \dfrac{w(x-a)}{(b-a)} & ; \quad a \leq x \leq b \\ \dfrac{w(c-x)}{(c-b)} & ; \quad b \leq x \leq c \\ 0 & ; \quad c < x \end{cases}$$

VII. Triangular Fuzzy Numbers (*TrFZ*)

If $w = 1$ in *GTFN*s have Normalized Triangular Fuzzy Numbers (*NTFN*s) $\tilde{A} = (a, b, c, 1)$, simply denoted by $\tilde{A} = (a, b, c)$ and is defined with the membership function,

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

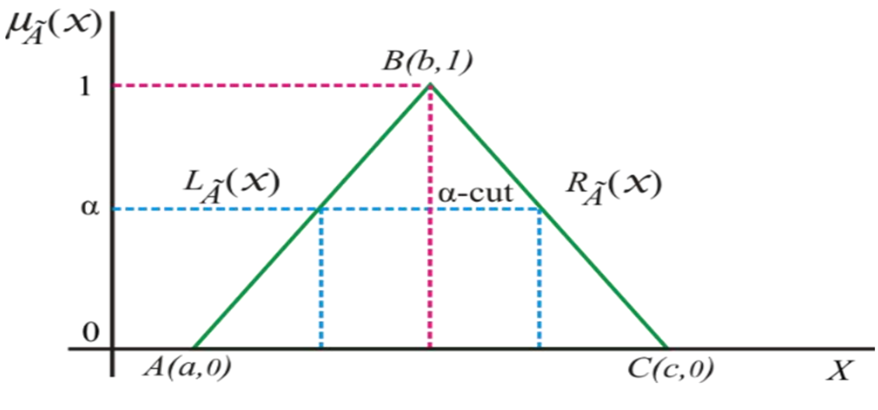

Figure 1: *Normalized Triangular Membership Function*

VIII. Normalized Trapezoidal Fuzzy Numbers (*NTZFN*s)

A Normalized Trapezoidal Fuzzy Numbers (*NTZFN*s) $\tilde{A} = (a, b, c, d; w)$ is described as any fuzzy subset of the real line $\mathbb{R}$, whose membership function $\mu_A(x)$ satisfies the following conditions:

- $\mu_A(x)$ is a continuous mapping from $\mathbb{R}$ to the closed interval [0,1]; $0 \leq w \leq 1$,
- $\mu_A(x) = 0$, for all $x \in (-\infty, a]$,
- $\mu_L(x) = L_{\hat{A}}(x)$ is strictly increasing on $[a, b]$,
- $\mu_A(x) = 1$, for all $b \leq x \leq c$,
- $\mu_R(x) = R_{\hat{A}}(x)$ is strictly decreasing on $[c, d]$,
- $\mu_{\hat{A}}(x) = 0$, for all $x \in [d, \infty]$; a, b, c, d are real numbers such that $a < b \leq c < d$.

It is known that for a *NTZFN*s. $\tilde{A} = (a, b, c, d; 1)$, there exist four numbers $a, b, c, d \in \mathbb{R}$ and two functions $L_{\hat{A}}(x), R_{\hat{A}}(x): \mathbb{R} \to [0,1]$, where $L_{\hat{A}}(x)$ and $R_{\hat{A}}(x)$ are non-decreasing and nonincreasing functions respectively. The functions $L_{\hat{A}}(x)$ and $R_{\hat{A}}(x)$ are also called the left and right side of the fuzzy number $\tilde{A}$ respectively. The membership function is defined as follows:

$$\mu_A(x) = \begin{cases} 0; & \text{otherwise} \\ \left(\frac{x-a}{b-a}\right); & a \leq x \leq b \\ 1; & b \leq x \leq c \\ \left(\frac{d-x}{d-c}\right); & c \leq x \leq d \\ 0; & \text{otherwise} \end{cases}$$

Similarly, it is known that for a *NTZFN*s. $\tilde{A} = (a, b, c, d; 1)$, there exist four numbers $a, b, c, d \in \mathbb{R}$ and two functions $L_{\hat{A}}(x), R_{\hat{A}}(x): \mathbb{R} \to [0,1]$, where $L_{\hat{A}}(x)$ and $R_{\hat{A}}(x)$ are nondecreasing and non-increasing functions respectively and its membership function is defined as follows:

$$\mu_{\tilde{A}}(x) = L_{\tilde{A}}(x) = (x-a)/(b-a) \text{ for } a \leq x \leq b\,;\ 1 \text{ for } b \leq x \leq c$$

$\mu_{\hat{A}}(x) = R_{\hat{A}}(x) = (x-c)/(d-c)$ for $c \leq x \leq d$ and 0 otherwise }. The functions $L_{\hat{A}}(x)$ and $R_{\tilde{A}}(x)$ are also called the left and right side of the fuzzy number $\tilde{A}$ respectively. In this proposed work, the assume that $\int_{-\infty}^{\infty} \tilde{A}(x)dx < +\infty$ and it is known that the α - cut of a fuzzy number is $\tilde{A}_\alpha = \{x \in \mathbb{R}/\mu_{\tilde{A}}(x) \geq \alpha\}$, for $\alpha \in (0,1]$ and $\tilde{A}_0 = cl\left(\cup_{\alpha\in(0,1]} \tilde{A}_\alpha\right)$, according to the definition of a fuzzy number, it is seen at once that every α - cut of a fuzzy number is a closed interval. Hence, for a fuzzy number $\tilde{A}$, we have $\tilde{A}(\alpha) = \left[\tilde{A}_L(\alpha), \tilde{A}_U(\alpha)\right]$ where $\tilde{A}_L(\alpha) = \inf\{x \in \mathbb{R}: \mu_{\tilde{A}}(x) \geq \alpha\}$ and $\tilde{A}_U(\alpha) = \sup\{x \in \mathbb{R}: \mu_{\tilde{A}}(x) \geq \alpha\}$. The left and right sides of the fuzzy number $\tilde{A}$ are strictly monotone, obviously, $\tilde{A}_L(\alpha)$ and $\tilde{A}_U(\alpha)$ are inverse functions of $L_{\tilde{A}}(x)$ and $R_{\tilde{A}}(x)$ respectively. The normal TZFNs $\tilde{A} = (a, b, c, d)$ and the corresponding α-cut is defined by

$$\tilde{A}_\alpha = [a + \alpha(b-a), d - \alpha(d-c)]; \alpha \in [0,1]$$

(i) If w = 1 in $GTZFN$ s, have a $NTZFN$ A$\tilde{A} = (a, b, c, d; 1)$ and can be simply denoted by $\tilde{A} = (a, b, c, d)$ and defined with the membership function,

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

where $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$.

$$\mu_A(x) = \begin{cases} 0 & ; x \leq a \\ L_A(x) = \dfrac{(x-a)}{(b-a)} & ; a \leq x \leq b \\ 1 & ; b \leq x \leq c \\ R_A(x) = \dfrac{(x-d)}{(d-c)} & ; c \leq x \leq d \\ 0 & ; d < x \end{cases}$$

where $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$.

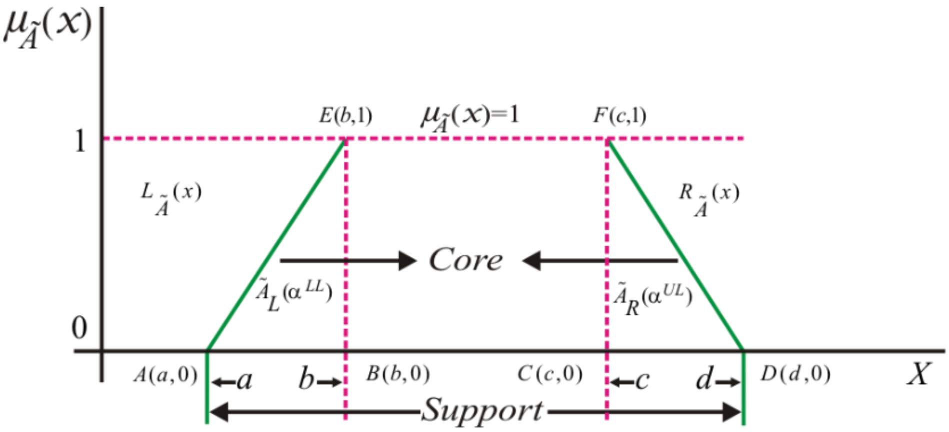

Figure 2: *Normalized Trapezoidal Membership Function*

(ii) If $w = 1$ in *GTZFNs*, have a *NTZFNs* $\tilde{A} = (a, b, c, d; 1)$ and can be simply denoted by $\tilde{A} = (a, b, c, d)$ dind uefined with the membership function is

$$\mu_A(x) = \begin{cases} 0 & ; x \leq a \\ L_A(x) = \dfrac{(x-b)}{(b-a)} & ; a \leq x \leq b \\ 1 & ; b \leq x \leq c \\ R_A(x) = \dfrac{(x-c)}{(d-c)} & ; c \leq x \leq d \\ 0 & ; d < x \end{cases}$$

where $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$

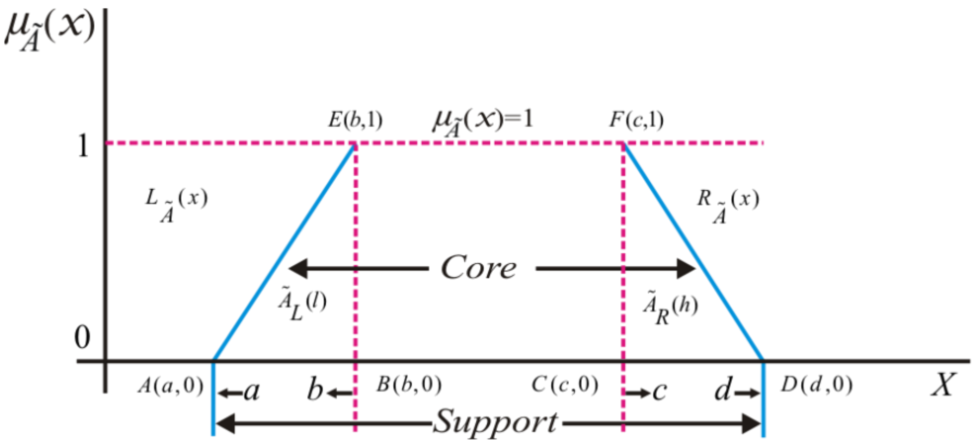

Figure 3: *Normalized Trapezoidal Membership Function*

The normal $TZFN$s $\tilde{A} = (a, b, c, d)$ and the corresponding without h level is defined by

$$\tilde{A} = [b - (b - a)l,\ c + (d - c)h];\ 0 \le h, l \le 1$$

IX. Trapezoidal Fuzzy Numbers (Tr$ZFNs$)

The membership function of a TZFNs $\widehat{A}$ is defined by

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{-(a-x)}{(b-a)}; a \le x \le b \\ 1; b \le x \le c \\ \frac{-(d-x)}{(c-d)}; c \le x \le d \end{cases}$$

Where, $a \le b \le c \le d$. A $TrFN$ becomes triangular fuzzy number if it satisfies b = c. In terms of $\alpha -$ cut interval, Tr FN is defined as follows:

$$\tilde{A} = [a + (b - a)\alpha, d - (d - c)\alpha]; 0 \le \alpha \le 1, \text{where, } a \le b \le c \le d$$

X. Wang's Centroid Point and Ranking Method

Wang et al. [14] has found that the centroid formulae proposed and have led to some misapplications such as by Liou, T. S [6]. They presented the correct method for centroid formulae for a generalized fuzzy number $\tilde{A} = (a, b, c, d; w)$ as;

$$(\bar{x}_0, \bar{y}_0) = \left[\left(\frac{1}{3}(a + b + c + d) - \left(\frac{dc-ab}{(d+c)-(a+b)}\right)\right), \left(\frac{w}{3}\right)\left(1 + \left(\frac{c-b}{(d+c)-(a+b)}\right)\right)\right] \quad (1)$$

and the ranking function associated with $\tilde{A}$ is

$$R(\tilde{A}) = \sqrt{\bar{x}_0^2 + \bar{y}_0^2}$$

For a normalized $TZFN$s, we put $w = 1$ in equations (1) so we have,

$$(\bar{x}_0, \bar{y}_0) = \left[\left(\frac{1}{3}(a + b + c + d) - \left(\frac{dc-ab}{(d+c)-(a+b)}\right)\right), \left(\frac{1}{3}\right)\left(1 + \left(\frac{c-b}{(d+c)-(a+b)}\right)\right)\right]$$

and the ranking function associated with $\tilde{A}$ is

$$R(\tilde{A}) = \sqrt{\bar{x}_0^2 + \bar{y}_0^2}$$

Let $\tilde{A}_i$and $\tilde{A}_j$be two fuzzy numbers, (i) $R(\tilde{A}_i) > R(\tilde{A}_j)$ then $\tilde{A}_i > \tilde{A}_j$ (ii) $R(\tilde{A}_i) < R(\tilde{A}_j)$ then $\tilde{A}_i < \tilde{A}_j$and $(iii) R(\tilde{A}_i) = R(\tilde{A}_j)$ then$\tilde{A}_i = \tilde{A}_j$.

XI. A Statistical Analysis of Variance for One-Way Classification with *TZFN*s using Ranking Method

The general linear model is

$$y_{ij} = \mu + \alpha_i + e_{ij}; \quad i = 1,2,\dots,k; \quad j = 1,2,\dots,n_i$$

The above general linear model of crisp *ANOVA* is classified into two interval fuzzy *ANOVA* models is Lower Limit Model (*LLM*) and Upper Limit Model (*ULM*). The *TZFN*s *LLM* is: $\left(\tilde{y}_{ij}\right)_l^{LL} = (\tilde{\mu})_l^{LL} + (\tilde{a}_i)_l^{LL} + \left(\tilde{e}_{ij}\right)_l^{LL}$, where $\left(\tilde{y}_{ij}\right)_l^{LL}$ is the j^{th} observations of the i^{th} treatment; $(\tilde{\mu})_l^{LL}$ is the general mean effect which is fixed; $(\tilde{a}_i)_l^{LL}$ is the fixed effect due to i^{th} treatment; and $\left(\tilde{e}_{ij}\right)_l^{LL}$ is the random error effect.

Null hypothesis: $\tilde{H}_0^{LL}: \tilde{\mu}_1^{LL} = \tilde{\mu}_2^{LL} = \cdots = \tilde{\mu}_r^{LL}$ against $\tilde{H}_1^{LL}: \tilde{\mu}_1^{LL} \neq \tilde{\mu}_2^{LL} \neq \cdots \neq \tilde{\mu}_r^{LL}$.

The *LLM* and *ULM* from the null hypothesis of acceptance levels of the direction in the pessimistic and optimistic of *TZFN*s. Using trapezoidal fuzzy *LLM* formula Is $b_{ij} - (b_{ij} - a_{ij})l$, where $0 \le i \le k, 0 \le j \le n_i$ and *ULM* formula is $c_{ij} + \left(d_{ij} - c_{ij}\right)h$, where $0 \le i \le k, 0 \le j \le n_i$.

The layout of *LLM* using *TZFN*s without -h level concept can be assigned as follows:

$$\begin{pmatrix} b_{11} - (b_{11} - a_{11})l, b_{12} - (b_{12} - a_{12})l, \dots, b_{1n_i} - \left(b_{1n_i} - a_{1n_i}\right)l \\ b_{21} - (b_{21} - a_{21})l, b_{22} - (b_{22} - a_{22})l, \dots, b_{2n_i} - \left(b_{2n_i} - a_{2n_i}\right)l \\ \vdots \qquad \vdots \qquad \dots \qquad \vdots \\ b_{1i} - (b_{1i} - a_{1i})l, b_{2i} - (b_{2i} - a_{2i})l, \dots, b_{in_i} - \left(b_{in_i} - a_{in_i}\right)l \\ \vdots \qquad \vdots \qquad \dots \qquad \vdots \\ b_{k1} - (b_{k1} - a_{k1})l, b_{k2} - (b_{t2} - a_{k2})l, \dots, b_{kn_i} - \left(b_{tn_i} - a_{kn_i}\right)l \end{pmatrix}$$

Then the required sum of squares for *LLM* is given below:

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left[\left(\tilde{y}_{ij}^2\right)_l^{LL}\right] = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left[b_{ij} - \left(b_{ij} - a_{ij}\right)l\right]^2 \text{ where } \left[\left(\tilde{y}_{ij}^2\right)_l^{IL}\right] = \left[b_{ij} - \left(b_{ij} - a_{ij}\right)l\right]^2$$

$$(cf)_l^{LL} = \frac{\left[(\tilde{y}^2)_l^{LL}\right]}{N}$$

$$(SS_T)_l^{LL} = \sum_{i=1}^{k} \sum_{j=1}^{n} \left[\left(\tilde{y}_{ij}^2\right)_l^{LL}\right] - \frac{\left[(\tilde{y}^2)_l^{LL}\right]}{N} \sim (N-1)\ df$$

$$(SS_{\tilde{T}})_l^{LL} = \sum_{i=1}^{k} \frac{\left[(\tilde{y}_i^2)_l^{LL}\right]}{n_i} - \frac{\left[(\tilde{y}^2)_l^{LL}\right]}{N} \sim (k-1)\ df$$

$$(SS_E)_l^{IL} = (SS_T)_l^{IL} - (SS_{T\tau})_l^{IL} \sim (N-k)df$$

$$\left(MS_{T_r}\right)_l^{IL} = \frac{(SS_{T\tau})_l^{LL}}{(k-1)}$$

$$(MS_E)_l^{IL} = \frac{(SS_E)_l^{IL}}{(N-k)}$$

$$\left(\tilde{F}_{TT}\right)_l^{IL} = \frac{(MS_{T\tau})_l^{IL}}{(MS_E)_l^{IL}}$$

The fuzzy *ANOVA* one-way table for *LLM* shown below;

Table 1: *ANOVA Table for ULM of Trapezoidal Fuzzy ANOVA One-Way*

SV	df	SS	MSS	$\tilde{F}^{LL}$ − Ratio
Treatments	$(k-1)$	$(SS_{T\tau})_l^{IL}$	$(MS_{T\tau})_l^{IL}$	$\frac{(MS_{T\tau})_l^{IL}}{(MS_E)_l^{LL}}$
Error	$(N-k)$	$(SS_E)_l^{LL}$	$(MS_E)_l^{IL}$	-
Total	$(N-1)$	$(SS_T)_l^{LL}$	-	-

Similarly, the trapezoidal fuzzy ULM is linear model is: $\left(\tilde{y}_{ij}\right)_h^{UI} = (\tilde{\mu})_h^{UL} + (\tilde{a}_i)_h^{UL} + \left(\tilde{e}_{ij}\right)_h^{UL}$ in which $\left(\tilde{y}_{ij}\right)_h^{UL}$ is the j^{th} observations of the i^{th} treatment; $(\tilde{\mu})_h^{UL}$ is the general mean effect which is fixed; $(\tilde{a}_i)_l^{LL}$ and $(\tilde{a}_i)_h^{UL}$ is the fixed effect due to i^{th} treatment; and $\left(\tilde{e}_{ij}\right)_h^{UL}$ is the random error effect; $i = 1,2,\dots,k$ and $j = 1,2,\dots,n_i$.

Null hypothesis: $\tilde{H}_0^{UI}: \tilde{\mu}_1^{UL} = \tilde{\mu}_2^{UL} = \cdots = \tilde{\mu}_r^{UI}$ against $\tilde{H}_1^{UL}: \tilde{\mu}_1^{UI} \neq \tilde{\mu}_2^{UI} \neq \cdots \neq \tilde{\mu}_r^{UL}$.

Similarly, the layout of *ULM* using *TZFN*s without *h*- level concept can be assigned as follows:

$$\begin{pmatrix} c_{11} + (d_{11} - c_{11})h, c_{12} + (d_{12} - c_{12})h, \dots, c_{1n_i} + \left(d_{1n_i} - c_{1n_i}\right)h \\ c_{21} + (d_{21} - c_{21})h, c_{22} + (d_{22} - c_{22})h, \dots, c_{2n_i} + \left(d_{2n_i} - c_{2n_i}\right)h \\ \vdots \qquad \vdots \qquad \cdots \qquad \vdots \\ c_{i1} + (d_{i1} - c_{i1})h, c_{i2} + (d_{i2} - c_{i2})h, \dots, c_{in_i} + \left(d_{in_i} - c_{in_i}\right)h \\ \vdots \qquad \vdots \qquad \cdots \qquad \vdots \\ c_{k1} + (d_{k1} - c_{k1})h, c_{k2} + (d_{k2} - c_{k2})h, \dots, c_{kn_i} + \left(d_{tn_i} - c_{kn_i}\right)h \end{pmatrix}$$

Then the required sum of squares for *ULM* is given bel

$$\sum_{i=1}^k \sum_{j=1}^n \left[\left(\tilde{y}_{ij}^2\right)_h^{UL}\right] = \sum_{i=1}^k \sum_{j=1}^n \left[c_{ij} + \left(d_{ij} - c_{ij}\right)h\right]^2 \; \left[\left(\tilde{y}_{ij}^2\right)_h^{UL}\right] = \left[c_{ij} + \left(d_{ij} - c_{ij}\right)h\right]^2$$

$$(cf)_h^{UL} = \frac{\left[\left(\tilde{y}^2\right)_h^{UL}\right]}{N} = \frac{\left[\left(G^2\right)_h^{UL}\right]}{N}$$

$$(SS_T)_h^{UL} = \sum_{i=1}^k \sum_{j=1}^n \left[\left(\tilde{y}_{ij}^2\right)_h^{UL}\right] - \frac{\left[\left(\tilde{y}^2\right)_h^{UL}\right]}{N} \sim (N-1)df$$

$$(SS_{T\tau})_h^{UL} = \sum_{i=1}^k \frac{\left[\left(\tilde{y}_i^2\right)_h^{UL}\right]}{n_i} - \frac{\left[\left(\tilde{y}^2\right)_h^{UL}\right]}{N} \sim (k-1)df$$

$$(SS_E)_h^{UL} = (SS_T)_h^{UL} - (SS_{T\tau})_h^{UL} \sim (N-k)df$$

$$(MS_{T\tau})_h^{UL} = \frac{(SS_{T\tau})_h^{UL}}{(k\text{-}1)}$$

$$\left(\tilde{F}_{TF}\right)_h^{UL} = \frac{(MS_{T\tau})_h^{UI}}{(MS_E)_h^{UI}}$$

The fuzzy *ANOVA* one-way table for *ULM* shown below;

Table 2: *ANOVA Table for ULM of Trapezoidal Fuzzy ANOVA One-Way*

SV	df	SS	MSS	$\tilde{F}^{UL}$ − Ratio
Treatments	$(k-1)$	$(SS_{TT})_h^{UL}$	$(MS_{T\tau})_h^{UL}$	$\frac{(MS_{T\tau})_h^{UL}}{(MS_E)_h^{UL}}$
Error	$(N-k)$	$(SS_E)_h^{UI}$	$(MS_E)_h^{UL}$	-
Total	$(N-1)$	$(SS_T)_h^{UL}$	-	-

Decision Rules for *LLM* and *ULM*

- If $(MS_{Tr})_l^{LL} > (MS_E)_l^{LL}$ or $\left(\tilde{F}_{T_r}\right)_l^{LL} = \frac{(MS_{T_r})_l^{LL}}{(MS_k)_l^{LL}} < F_T$, where $\left(\tilde{F}_{Tr}\right)_l^{IL}$ is the calculated value and F_T is the tabulated value of F for $(k-1), (N-k)$ df at α% loss, then the null hypothesis $\tilde{H}_0^{LL}$ is accepted, otherwise the alternative hypothesis $\tilde{H}_1^{LL}$ is accepted for all $0 \leq l^{II} \leq 1$.
- If $(MS_{TV})_h^{UL} > (MS_E)_h^{UL}$ or $\left(\tilde{F}_{T_r}\right)_h^{tL} = \frac{(MS_T)_k^{UL}}{(MS_E)_h^{ULL}} < F_T$, where $\left(\tilde{F}_{TT}\right)_h^{UL}$ is the calculated value and F_T is the tabulated value of F for $(k-1), (N-k)$ dif at α% lo.s, then the null hypothesis $\tilde{H}_0^{UI}$ is accepted, otherwise the alternative hypothesis $\tilde{H}_1^{UI}$ is accepted for all $0 \leq h^{ul} \leq 1$.

Note: If when l=0 or h=0, that is, centre level, the null hypothesis of the *ANOVA* model is accepted at 5% l.o.s. This study is about the acceptance levels of null hypothesis and its directions of pessimistic and optimistic. In this method, the notions of pessimistic degree and optimistic degree are not used in without h-level concept and the decision obtained.

III. Results

I. Application

The data was collected for yields of paddy on Salem District of Tamil Nadu in India. In the yields of four different types (Y_1, Y_2, Y_3, Y_4) in kilograms with three replications of paddy (*ADT, IR, ASD*) each by using *ANOVA* method. Stipulation we cannot record the precise number of yields in kilograms in a sample due to approximately an unexpected situation happen, there the proposed fuzzy data for number yields in kilograms. The trapezoidal fuzzy data is performed below;

Table 3: *Observations*

Yields in kilograms (*i*)	Varieties of paddy (*j*)		
	ADT	IR	ASD
Y_1	8, 10, 11, 14	11, 13, 15, 18	-
Y_2	7, 9, 12, 15	8, 9, 11, 14	8, 10, 12, 14
Y_3	12, 14, 16, 18	12, 13, 15, 19	14, 16, 19, 23
Y_4	12, 13, 16, 19	19, 20, 22, 24	-

To test whether there is any significant difference in the varieties of paddy performance of the yields in kilograms.

Let's take the initial set value (8,10,11,14). using LLM formula is $b-(b-a)la = 8, b = 10$, $10-(10-8)l, (10-2l)$. Similarly, using the same procedure get the remaining values as given below:

Table 4: *LLM using TZFNs without h Level*

Yields in kilograms (i)	Varieties of paddy (j)			
	ADT	IR	ASD	
Υ_1	$10-2l$	$13-2l$	-	$23-4l$
Υ_2	$9-2l$	$9-l$	$10-2l$	$28-5l$
Υ_3	$14-2l$	$13-l$	$16-2l$	$43-5l$
Υ_4	$13-l$	$20-l$	-	$33-2l$
				$127-16l$

$\tilde{H}_0^{LL}$: The varieties of paddy do not differ significantly with respect to the yields in kilograms The sum of squares is given below;

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\tilde{y}_{ij}^2\right)_l^{LL} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[b_{ij}-\left(b_{ij}-a_{ij}\right)l\right]^2 = (10-2l)^2+\cdots+(20-l)^2$$

$$(cf)_k^{UL} = \frac{[(\tilde{y}_{-}^2)_h^{LL}]}{N} = \frac{(127-16l)^2}{10} = 25.6l^2-406.4l+1612.9$$

$$(SS_r)_j^{LL} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[\left(\tilde{y}_{ij}^2\right)_l^{LL}\right]-\frac{[(\tilde{y}_{-}^2)_l^{LL}]}{N} = (28l^2-398l+1721)-(25.6l^2-406.4l+1612.9)$$

$$(SS_T)_l^{LL} = 2.4l^2+8.4l+108.1$$

$$(SS_{TT})_l^{IL} = \sum_{i=1}^{k}\frac{[(\tilde{y}_i^2)_l^{LU}]}{n_i}-\frac{[(\tilde{y}^2)_l^{LL}]}{N} = \left\{\begin{matrix}\left(\frac{(23-4l)^2}{2}+\frac{(28-5l)^2}{3}+\frac{(43-5l)^2}{3}+\frac{(33-2l)^2}{2}\right)\\ -(25.6l^2-406.4l+1612.9)\end{matrix}\right\}$$

$$(SS_E)_l^{LL} = 1.33l^2-3.33l+34.33$$

$$(MS_{TT})_l^{IL} = \frac{\left(SS_{T_f}\right)_l^{IL}}{(4-1)} = \frac{1.37l^2+11.73l+73.77}{3} = 0.36l^2+3.91l+24.59$$

$$(MS_E)_l^{IL} = \frac{(SS_E)_l^{IL}}{(10-4)} = \frac{1.33l^2-3.33l+34.33}{6} = 0.22l^2-0.56l+5.72$$

$$\left(\tilde{F}_{Tr}\right)_l^{LL} = \frac{(MS_{Tr})_l^{IL}}{(MS_E)_l^{LL}} = \frac{0.36l^2+3.91l+24.59}{0.22l^2-0.56l+5.72}$$

$$\text{Put } l = 0, \left(\tilde{F}_{TT}\right)_l^{LL} = \frac{0.36l^2 + 3.91l + 24.59}{0.22l^2 - 0.56l + 5.72} = 4.3, \dots, l = 1, \left(\tilde{F}_R\right)_l^{LL} = 5.36.$$

The sum of squares is summarized given below;

Table 5: *ANOVA Table for LLM of Fuzzy ANOVA One Way*

SV	df	SS	MSS	$\tilde{F}^{LL}$ − Ratio
Treatments	3	$1.07l^2 + 11.73l + 73.77$	$0.36l^2 + 3.91l + 24.59$	$\frac{0.36l^2 + 3.91l + 24.59}{0.22l^2 - 0.56l + 5.72}$
Error	6	$1.33l^2 - 3.33l + 34.33$	$0.22l^2 - 0.56l + 5.72$	-
Total	9	$2.4l^2 + 8.4l + 108.1$	-	-

Decision Rules for LLM : Since, the $(\tilde{F})_l^{LL} > F_T$, where $F_T = 4.76$ is the table value at 5% at 0.5 with (3,6) dif Then, the null hypothesis $\tilde{H}_0^{LL}$ of the LLM is accepted for all $0 \leq l^{LL} \leq 0.42$; otherwise, the null hypothesis $\tilde{H}_0^{UL}$ is rejected for all $043 \leq l^{IL} \leq 1$. Therefore, the varieties of paddy do not differ significantly with respect to the yields in kilograms. Let's take the initial set value (8,10,11, 14). using ULM formula is $c + (d - c)h$ $c = 11, d = 14, 11 + (14 - 11)h, 11 + 3h$. Similarly, using the same procedure get the remaining values as given below

Table 6: *ULM using TZFNs without h Level*

Yields in kilograms (i)	Varieties of paddy (j)			
	ADT	IR	Total	
Y1	$11 + 3h$	$15 + 3h$	-	$26 + 6h$
Y2	$12 + 3h$	$11 + 3h$	$12 + 2h$	$35 + 8h$
Y3	$16 + 2h$	$15 + 4h$	$19 + 4h$	$50 + 10h$
Y4	$16 + 3h$	$22 + 2h$	-	$38 + 5h$
				$149 + 29h$

$\tilde{H}_0^{UL}$: The varieties of paddy do not differ significantly with respect to the yields in kilograms
The sum of squares is given below;

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\tilde{y}_{ij}^2\right)_h^{uL} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[c_{ij} + \left(d_{jy} - c_{ij}\right)h\right]^2 = (11 + 3h)^2 + \cdots + (22 + 2h)^2$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\tilde{y}_{ij}^2\right)_h^{iL} = 89h^2 + 862h + 2337$$

$$(S_T)_h^{UL} = \sum_{l=1}^{k}\sum_{j=1}^{n_i}\left[\left(\tilde{y}_{ij}^2\right)_h^{eL}\right] - \frac{[(\tilde{y}^2)_h^{ULL}]}{N}$$

$$(SS_T)_h^{UL} = 4.9h^2 - 2.2h + 116.9$$

$$(SS_{Tr})_h^{ULL} = 1.07h^2 + 1.8h + 81.57$$

$$(SS_E)_h^{UL} = (4.9h^2 - 2.2h + 116.9) - (1.07h^2 + 1.8h + 81.57)$$

$$(SS_E)_h^{UL} = 3.83h^2 - 4h + 35.33$$

$$(MS_{TT})_h^{UL} = \frac{(SS_{T\tau})_h^{UL}}{(4-1)} = \frac{1.07h^2 + 1.8h + 81.57}{3} = 0.36h^2 + 0.6h + 27.19$$

$$(MS_E)_h^{UL} = \frac{(SS_E)_h^{UI}}{(10-4)} = \frac{3.83h^2 - 4h + 35.33}{6} = 0.64h^2 - 0.67h + 5.89$$

$$\left(\tilde{F}_{V_r}\right)_h^{UL} = \frac{\left(MS_{r_r}\right.}{(MS_t)_h^{UL}} = \frac{0.36h^2 + 0.6h + 27.19}{0.64h^2 - 0.67h + 5.89}$$

Put $h = 0, \left(\tilde{F}_{T\tau}\right)_h^{UL} = \frac{0.36(0)^2+0.6(0)+27.19}{0.64(0)^2-0.67(0)+5.89} = 4.62 \ldots h = 1\left(\tilde{F}_R\right)_l^{LL} = 4.8.$

The sum of squares is summarized given below;

Table 7: *ANOVA Table for ULM of Fuzzy ANOVA One Way*

SV	df	SS	MSS	$\tilde{F}^{UL}$ – Ratio
Treatments	3	$1.07h^2 + 1.8h + 81.57$	$0.36h^2 + 0.6h + 27.19$	$\frac{0.36h^2 + 0.6h + 27.19}{0.64h^2 - 0.67h + 5.89}$
Error	6	$3.83h^2 - 4h + 35.33$	$0.64h^2 - 0.67h + 5.89$	-
Total	9	$4.9h^2 - 2.2h + 116.9$	-	-

Decision Rules for *ULM*: Since, the $(\tilde{F})_k^{UL} > F_T$, where $F_T = 4.76$ is the table value at 5% $\mathrm{l.o.s}$ with (3,6) *df*. Therefore, the null hypothesis $\tilde{H}_0^{UL}$ of the *ULM* is accepted for all $0 \leq h'^{IN} \leq 0.30$, otherwise the null hypothesis $\hat{H}_0^{UL}$ is rejected for all $0.31 \leq h^{JN} \leq 1$. Hence, the varieties of paddy do not differ significantly with respect to the yields in kilograms.

IV. Conclusion

This study applied fuzzy *ANOVA* techniques using both the *LLM* and *ULM* to evaluate the performance of different paddy varieties in terms of yield. The findings indicate that the null hypothesis is accepted across all levels of significance (α %) for both models, suggesting that the yields of the paddy varieties do not differ significantly. The null hypothesis $\tilde{H}_0^{LL}$ is accepted for all $0 \leq l^{lL} \leq 0.42$ and $\tilde{H}_0^{\mathrm{ULL}}$ is accepted for all $0 \leq h^{\mathrm{LL}} \leq 0.30$. Therefore, the fuzzy null hypothesis $\tilde{H}_0$ of the fuzzy *ANOVA* one-way model is accepted. It says that the varieties of paddy performance of the yields in kilograms are equal only if $0 \leq l^{tL} \leq 0.42$ and $0 \leq h^{iL} \leq 0.30$. But some specific levels of the maximum level of pessimistic value is 0.42 and the maximum level of optimistic value is 0.30. Finally, it is inference that there is no relation between the different yields in kilograms and paddy varieties as decision rules of the conclusions reached by both *LLM* and *ULM*. Furthermore, the analysis highlights specific fuzzy levels, with a maximum pessimistic value of 0.42 and a maximum optimistic value of 0.30, indicating no substantial variation in yield across the varieties. The consistency of the conclusions drawn from both LLM and ULM strengthens the inference that paddy varieties yield similar performance in kilograms. These results underscore the utility of fuzzy *ANOVA* in agricultural studies and provide a robust framework for analysing performance variability under uncertainty. Future research may explore the application of these fuzzy models to other crops and environmental conditions to validate their broader applicability. For future works, one can use the approach of this paper to extend other experimental designs such as two-way ANOVA, $LSD, BIBD, PBIBD$, etc.

References

[1] Abbasbandy, S., and Asady, B. (2004). The nearest trapezoidal fuzzy number to a fuzzy quantity. *Applied mathematics and computation*, 156(2): 381-386.

[2] Abbasbandy, S., and Amirfakhrian, M. (2006). The nearest approximation of a fuzzy quantity in parametric form. *Applied Mathematics and Computation*, 172(1): 624-632.

[3] Abbasbandy, S., and Amirfakhrian, M. (2006). The nearest trapezoidal form of a generalized left right fuzzy number. *International Journal of Approximate Reasoning*, 43(2): 166-178.

[4] Buckley, J. J., and Jowers, L. J. (2008). Monte Carlo methods in fuzzy optimization. *Berlin: Springer.*

[5] Barbacioru, I. C. (2012). Statistical hypothesis testing using fuzzy linguistic variables. *Fiability & Durability/Fiabilitate si Durabilitate*. (1): 336-342.

[6] Liou, T. S., and Wang, M. J. J. (1992). Ranking fuzzy numbers with integral value. *Fuzzy sets and systems*, 50(3): 247-255.

[7] Salahshour, S., et al. (2011). Ranking fuzzy numbers using fuzzy maximizing-minimizing points. In *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology* (pp. 763-769). Atlantis Press.

[8] Rezvani, S., and Molani, M. (2014). Representation of trapezoidal fuzzy numbers with shape function. *Ann. Fuzzy Math. Inform*, 8(1): 89-112.

[9] Thorani, Y. L. P. et al. R. (2012). Ordering generalized trapezoidal fuzzy numbers. *Int. J. Contemp. Math. Sciences*, 7(12): 555-573.

[10] Wang, Y. M., et al. (2006). On the centroids of fuzzy numbers. Fuzzy sets and systems, 157(7): 919-926.

[11] Viertl, R. (2011). Statistical methods for fuzzy data. *John Wiley and Sons*.

[12] Viertl, R. (2006). Univariate statistical analysis with fuzzy data. *Computational Statistics & Data Analysis*, *51*(1): 133-147.

[13] Rezvani, S. (2013). Ranking generalized trapezoidal fuzzy numbers with Euclidean distance by the incentre of centroids. *Mathematica Aeterna*, 3(2): 103-114.

[14] Parthiban, S., and Gajivaradhan, P. (2016). One-factor ANOVA model using trapezoidal fuzzy numbers through alpha cut interval method. *Annals of Pure and Applied Mathematics*, 11(1): 45-61.

[15] Zadeh, L. A. (1996). On fuzzy algorithms. *In fuzzy sets, fuzzy logic, and fuzzy systems: selected papers By Lotfi A Zadeh*.

[16] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information sciences*, 8(3): 199-249.

[17] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information sciences*, 8(3): 199-249.

[18] Wu, H. C. (2007). Analysis of variance for fuzzy data. *International Journal of Systems Science*, 38(3): 235-246.

[19] Nourbakhsh, M., et al. (2013). Analysis of variance based on fuzzy observations. *International Journal of Systems Science*, 44(4): 714-726.