CLASSICAL ANALYSIS OF RANDOMIZED BLOCK DESIGN USING AMBIGUOUS DATA

A. Mariappan*

*Research Scholar, Department of Statistics, Periyar University, Salem, Tamilnadu, India

M. Pachamuthu

Assistant Professor, Department of Statistics, Periyar University, Salem, Tamilnadu, India

Abstract

Experimental design plays an essential part in statistical analysis and data interpretation. The problem of a classical Randomized Block Design (*RBD*) test for Triangular Fuzzy Numbers (*TFN*) is discussed in this paper. In this proposed methodology, it is obvious that if the value of the observed fuzzy test statistics is similar to real numbers in the testing crisp hypotheses, then fuzzy *RBD* is very sensitive for making the determinations as to whether to accept or reject the fuzzy null hypotheses and also debates the application of the method for example.

Keywords: Classical *RBD*, Fuzzy *RBD*, *TFN*, Decision Rule.

Mathematics Subject Classification (2010): 62A86, 62F03, 97K80

1. INTRODUCTION

The Completely Randomized Design (*CRD*) was simple because the principle of local control was not used, and experimental material was assumed to be homogeneous, but it is noted that the experimental material is not absolutely homogeneous. A fertility gradient in one direction is often present in agricultural field experiments. The simple method of regulating the variability of the experimental material in such a situation consists of stratifying or grouping the entire experimental area into relatively homogeneous strata or subgroups (called blocks) perpendicular to the fertility gradient direction. These blocks are so designed that plots are homogeneous within a block, and heterogeneous between blocks. In other words, inside a block there might be less variation, and the main difference or variation between blocks. It should be held in mind that for an effective blocking of the content, familiarity with the design of experimental units is important. The method of dividing experimental material into a number of blocks gives rise to a design known as *RBD* that can be described as an arrangement of t treatments in r blocks such that each treatment takes place exactly once in each block. Fuzzy set theory [21] was extended to several areas that need to handle ambiguous and unclear data. These areas include estimated logic, decision-making, optimization, power, etc. The sample findings are crisp in conventional statistical research, and a statistical test leads to the binary decision. Many authors have studied the statistical theories that are evaluated in fuzzy environments using the fuzzy set theory principles introduced by Zadeh [22]. Chachi et al. [3] are proposing a new approach to the issue of evaluating statistical hypotheses. As a fuzzy subset of the real line, Dubois and Prade [5] identified some of the fuzzy numbers. Mikihiko Konishi et al. [14] suggested an Analysis of Variance (*ANOVA*) for the fuzzy interval data using the definition of the fuzzy set. Wu [19, 20] introduced hypothesis testing of a single factor *ANOVA* model for fuzzy data by solving optimization problems using the h-level and the notions of pessimistic degree and optimistic degree. The two-factor ANOVA test were analysed by Gajivaradhan and Parthiban [7] using an alpha cut interval method for trapezoidal fuzzy numbers. A bootstrap approach to the multi-sample test of means with imprecise data was suggested by Gil et al. [9]. When both the theories and the available data are fuzzy, Arefi and Taheri [1] formed the testing hypothesis. Filzmoser and Viertl [6], proposed to test hypotheses on the fuzzy *p* value with fuzzy data. Nakama et al. [15] Discuss derive statistical tests that are ideal for testing the null hypotheses, and develop a bootstrap scheme to estimate the *p* values of the test statistics observed. Mariappan and Pachamuthu [2020] suggested the statistical testing of hypotheses for fuzzy *CRD* using *TFN*. The classical *RBD* model for *TFN* is analyzed in this paper using a numerical example.

2. PRELIMINARIES

2.1 Triangular Fuzzy Number

A triangular fuzzy number $\tilde{A}$ is a fuzzy number fully specified by tribles (a, b, c) such that $a \le b \le c$ with membership function defined as

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & if\ x \le a \\ \dfrac{x-a}{b-a} & if\ a \le x \le b \\ \dfrac{x-c}{b-c} & if\ b \le x \le c \\ 0 & if\ x \ge c \end{cases}$$

where a is the indicates of lower point, b is the indicates of centre point and c is the indicates of upper point.

The triangular fuzzy number is represented diagramatically as

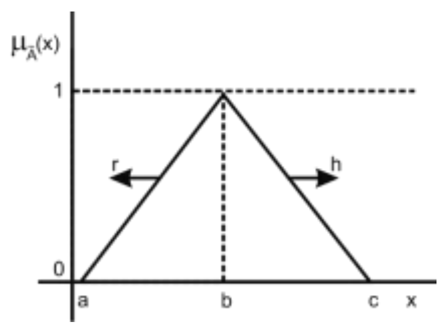

Fig. 1. Triangular Fuzzy Numbers

The form of a fuzzy interval number can be expressed as a triangular fuzzy number follows: $\tilde{A} = \left[\left\{(b-a)r + a\right\}^{L}; \left\{(b-c)h + c\right\}^{U} \right]; 0 \le h, r \le 1$

where r is the level of pessimistic and h is the level of optimistic of the fuzzy numbers $\tilde{A} = (a, b, c)$.

3. CLASSICAL *RBD*

Through proving the local control (blocking) measure in the design, an increase in *CRD* can be obtained. One such design is entirely *RBD*. *ANOVA* technique for two-way data classification is applicable to the *RBD* layout experiment. The data obtained from the experiment is graded by two factors namely treatments and blocks according to different levels. For *RBD*, the linear model is described by

$$y_{ij} = \mu + a_i + b_j + e_{ij}; i = 1, 2, ..., t;\ j = 1, 2, ..., b$$

where, y_{ij} is the observations corresponding to the i^{th} treatment and j^{th} block, μ is the general mean effect which is fixed, a_i is the fixed effect due to the i^{th} treatment, b_j is the fixed effect due to the j^{th} block and e_{ij} is the random error effect. To determine whether the factor level means μ_i equal or not. The following testing hypotheses are known as

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \text{ against } H_1 : \text{not all } \mu_i \text{ are equal}$$

Let $\sum_{ij}^{tb} y_{ij} = y_{..} = G$ be the grand total of tb observations, $\sum_{ij}^{tb} y_{ij} = y_{i.} = T_i$ be the i^{th} treatment total, $\sum_{ij}^{tb} y_{ij} = y_{.j} = B_j$ is the j^{th} block total and also $cf = \frac{G^2}{kr}$. Then, the various sum of squares, mean sum of squares and F Ratio listed given below:

$SST = Q_{SST} = \sum_{ij}^{tb} y_{ij}^2 - \frac{G^2}{kr}$ which has $(tb-1)\,df$, $SSBTR = Q_{SSBTR} = \sum_{ij}^{tb} \frac{T_i^2}{r} - \frac{G^2}{kr}$ which has $(t-1)$ df, $SSB = Q_{SSB} = \sum_{ij}^{tb} \frac{B_j^2}{k} - \frac{G^2}{kr}$ which has $(b-1)$ df, $SSE = Q_{SSE} = Q_{SST} - Q_{SSBTR} - Q_{SSBB}$ which has $(t-1)(b-1)$ df , $MSBTR = \frac{SSBTR}{(t-1)}, MSBB = \frac{SSBB}{(b-1)}, MSE = \frac{SSE}{(t-1)(b-1)}$, $F_{BTR} = \frac{MSBTR}{MSE}$ and $F_{BB} = \frac{MSBB}{MSE}$.

In the ANOVA table, all these values are referred to and inferences are drawn.

Table – 1. *ANOVA* Table for Crisp *RBD*

SV	*df*	*SS*	*MSS*	*F Ratio*
Between Treatments	$(t-1)$	Q_{SSBTR}	$MSBTR$	F_{BTR}
Between Blocks	$(b-1)$	Q_{SSBB}	$MSBB$	F_{BB}
Experimental Error	$(t-1)(b-1)$	Q_{SSE}	MSE	-
Total	$(tb-1)$	Q_{SST}	-	-

3.1 Decision Rule of Between Treatments and Between Blocks

The decision rules of F test to accept or reject between treatments and between blocks at $\alpha\%$ significance level the null hypothesis and alternative hypothesis. Suppose that if $F_T < F_C$, [where F_T is the tabulated value for $(t-1),(t-1)(b-1)$ and $(b-1),(t-1)(b-1)$ degrees of freedom, and F_C is the calculated value], then the null hypothesis H_0 is rejected. Otherwise, alternative hypothesis H_0 is rejected.

3.2 Fuzzy Technique of RBD

The triangular fuzzy approach to the fuzzy statistical analysis of *RBD*. Throughout this case, the data recorded as well as the observations are regarded as *TFN*. Below is the mathematical general linear model:

$$\tilde{y}_{ij} = \tilde{\mu} + \tilde{a}_i + \tilde{b}_j + e_{ij};\; i = 1,2,...,t;\; j = 1,2,...,b$$

In the fuzzy interval *RBD* models, the general linear model of classical *RBD* is classified; the fuzzy lower and upper level models are regarded as: $\tilde{y}_{ij}^{L} = (\tilde{\mu})_r^L + (\tilde{a}_i)_r^L + (\tilde{b}_j)_r^L + (\varepsilon_{ij})_r^L$ and $\tilde{y}_{ij}^{U} = (\tilde{\mu})_h^U + (\tilde{a}_i)_h^U + (\tilde{b}_j)_h^U + (\varepsilon_{ij})_h^U$ in which $(\tilde{y}_{ij})_r^L$ and $(\tilde{y}_{ij})_h^U$ is the observation corresponding to the i^{th} level of factor A and j^{th} level of factor B. $(\tilde{\mu})_r^L$ and $(\tilde{\mu})_h^U$ is the general mean effect which is fixed. $(\tilde{a}_i)_r^L$ and $(\tilde{a}_i)_h^U$ is the fixed effect due to the i^{th} level of factor A. $(\tilde{b}_j)_r^L$ and $(\tilde{b}_j)_h^U$ is the fixed effect due to the j^{th} level of factor B. $(\tilde{\varepsilon}_{ij})_r^L$ and $(\tilde{\varepsilon}_{ij})_h^U$ is the random error effect which is independent identically distributed (iid) with mean is 0 and constant variance is $\sigma^2; i = 1, 2, ..., t$ and $j = 1, 2, ..., b$. After this, to test the lower and upper level model and the fuzzy null hypotheses and fuzzy alternative hypotheses respectively, utilizing classical RBD methods. The simplest the fuzzy null hypotheses $\tilde{H}_0 : \tilde{\mu}_1 = \tilde{\mu}_2 = ... = \tilde{\mu}_r$ against the fuzzy alternative hypotheses $\tilde{H}_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2 \neq ... \neq \tilde{\mu}_r$. This implies the following two sets (Lower and Upper levels) of hypotheses are given below

3.3 Fuzzy Hypotheses of Lower and Upper Level Models

The fuzzy null hypotheses of between treatment and between block is $\tilde{H}_0^L : \tilde{\mu}_1^L = \tilde{\mu}_2^L = ... = \tilde{\mu}_r^L$ against the fuzzy alternative hypotheses between treatment and between block is $\tilde{H}_1^L : \tilde{\mu}_1^L \neq \tilde{\mu}_2^L \neq ... \neq \tilde{\mu}_r^L$.

The fuzzy null hypotheses of between treatment and between block is $\tilde{H}_0^U : \tilde{\mu}_1^U = \tilde{\mu}_2^U = ... = \tilde{\mu}_r^U$ against the fuzzy alternative hypotheses of between treatment and between block is $\tilde{H}_1^U : \tilde{\mu}_1^U \neq \tilde{\mu}_2^U \neq ... \neq \tilde{\mu}_r^U$.

In *TFN* pessimistic and optimistic for the fuzzy lower and upper level models from the null hypothesis of acceptance or rejection direction levels. Through the use of triangular fuzzy lower and upper levels formulas are $(b_{ij} - a_{ij})r + a_{ij}$ where $0 \le i \le t; 0 \le j \le b$ and $(b_{ij} - c_{ij})h + c_{ij}$ where $0 \le i \le t; 0 \le j \le b$. (Note that $r^L = 1$ and $h^U = 1$, centre level). Then the required formula for lower level of fuzzy *RBD* is given below:

$$SST_r^L = \sum_{i=1}^{t}\sum_{j=1}^{b}[(\tilde{y}_{ij})_r^L] - \frac{[(\tilde{y}_{..}^2)_r^L]}{tb}, \quad SSBTR_r^L = \sum_{i=1}^{t}\frac{[(\tilde{y}_{i.}^2)_r^L]}{b} - \frac{[(\tilde{y}_{..}^2)_r^L]}{tb}$$

$$SSBB_r^L = \sum_{j=1}^{b}\frac{[(\tilde{y}_{i.}^2)_h^U]}{t} - \frac{[(\tilde{y}_{..}^2)_h^U]}{tb}, \quad SSE_r^L = SST_r^L - SSBTR_r^L - SSBB_r^L$$

$$MSBTR_r^L = \frac{SSBTR_r^L}{(t-1)}, \quad MSBB_r^L = \frac{SSBB_r^L}{(b-1)} \text{ and } MSE_r^L = \frac{SSE_r^L}{(t-1)(b-1)}$$

$$(\tilde{F}_{BTR})_r^L = \frac{MSBTR}{MSE} \text{ and } (\tilde{F}_{BB})_r^L = \frac{MSBB}{MSE}$$

In the lower level of *ANOVA* table for fuzzy *RBD* table, all these values are represented and fuzzy decision rule is drawn.

Table – 2. *ANOVA* Table for Lower Level of Fuzzy *RBD*

SV	*df*	*SS*	*MSS*	$\tilde{F}-Ratio$
Between Treatments	$(t-1)$	$SSBTR_r^L$	$MSBTR_r^L$	$(\tilde{F}_{BTR})_r^L$
Between Blocks	$(b-1)$	$SSBB_r^L$	$MSBB_r^L$	$(\tilde{F}_{BB})_r^L$
Experimental Error	$(t-1)(p-1)$	SSE_r^L	MSE_r^L	-
Total	$(tb-1)$	SST_r^L	-	-

Similarly, upper level models of ambiguous *RBD* can write formulas and tables so formulas and tables could be avoided.

3.4 Fuzzy Decision Rule of (Lower and Upper levels) Between Treatments and Between Blocks

Suppose that if $F_T < F_C$, [where F_T is the tabulated value for $(t-1),(t-1)(b-1)$ and $(b-1),(t-1)(b-1)$ degrees of freedom, and F_C is the calculated value (using 3.1)], then the fuzzy null hypotheses of lower level for $\tilde{H}_0^L$ and fuzzy null hypotheses of upper level for $\tilde{H}_0^U$ is rejected for $0 \le r^L \le r_T$ where $0 \le r_T \le 1$ and $0 \le h^U \le h_T$ where $0 \le h_T \le 1$. Otherwise, fuzzy alternative hypotheses of lower level for $\tilde{H}_0^L$ and fuzzy alternative hypotheses of upper level for $\tilde{H}_0^U$ is rejected for $0 \le h^U \le h_T$ where $0 \le h_T \le 1$.

The proposed classical technique for evaluating *RBD* model fuzzy hypotheses with fuzzy data is illustrated with an example below.

4. APPLICATIONS

In our study, to collect the yields of primary data groundnut varieties at Omalur, Salem District of Tamilnadu. Three replicates of various groundnut varieties (TMV 2, TMV 7, VRI 2) in kilograms and four yields of (Y1, Y2, Y3, Y4). Via an *RBD*, with four replications of groundnut in kilograms for yields per plot, three varieties of crops are tested, the layout being *TFN* due to certain work friction is given as below.

Table – 3. Table for Classical *RBD* using *TFN*

Varieties of Groundnut	Yields in kilograms			
	Y1	Y2	Y3	Y4
TMV 2	56,58,60	54,58,62	53,56,59	54,58,62
TMV 7	58,60,62	53,58,63	56,59,62	57,60,63
VRI 2	59,62,65	58,60,62	57,60,63	57,59,61

To test if there is any substantial difference between the production of the groundnut varieties in the yields in kilograms per plot. Let $\tilde{\mu}_i$ be the mean number of yields in kilograms per plots for the i^{th} varieties of groundnut. Now, the null hypothesis, $\tilde{H}_0 : \tilde{\mu}_1 = \tilde{\mu}_2 = \tilde{\mu}_3 = \tilde{\mu}_4$ and the alternative hypothesis, $\tilde{H}_1 :$ not all $\tilde{\mu}_i$'s are equal.

$\tilde{H}_0 :$ To test whether groundnut varieties do not vary significantly with respect to yields.

$\tilde{H}_1$: To test if groundnut varieties vary significantly with respect to yields.

Let us consider the lower level model is given below

4.1 Lower Level Model

Table – 4. Table for Upper Level Model

Varieties of Groundnut	Yields in kilograms			
	Y1	Y2	Y3	Y4
TMV 2	$2r+56$	$4r+54$	$3r+53$	$4r+54$
TMV 7	$2r+58$	$5r+53$	$3r+56$	$3r+57$
VRI 2	$3r+59$	$2r+58$	$3r+57$	$2r+57$

$$SST_r^L = 10r^2 - 30r + 46,\ SSBTR_r^L = 1.5r^2 - 10.5r + 24.5$$

$$SSBB_r^L = 2.67r^2 - 10.67r + 12.67,$$

$$SSE_r^L = 5.83r^2 - 8.83r + 8.83\ MSBTR_r^L = 0.75r^2 - 5.25r + 12.25,$$

$$MSBB_r^L = 0.89r^2 - 3.56r + 4.22, MSE_r^L = 0.97r^2 - 1.47r + 1.47$$

$$(\tilde{F}_{BTR})_r^L = \frac{0.75r^2 - 5.25r + 12.25}{0.97r^2 - 1.47r + 1.47},\ (\tilde{F}_{BB})_r^L = \frac{0.89r^2 - 3.56r + 4.22}{0.97r^2 - 1.47r + 1.47}$$

4.2 Fuzzy Decision Rule of Between Treatments

If $\tilde{F}_r^L > F_T$, for all $r; 0 \le r \le 1$ where $F_T = 5.14$ is the F table value of α at 5% level of significance with $(2,6)\,df$ then, the fuzzy null hypotheses $\tilde{H}_0^L$ is rejected for the $r; 0 \le r \le 1$. Thus, the disparity between the treatments is substantial. Therefore, groundnut varieties vary greatly in yields.

4.3 Fuzzy Decision Rule of Between Blocks

If $\tilde{F}_r^L < F_T$, for all $r; 0 \le r \le 1$ where $F_T = 4.76$ is the F table value of α at 5% level of significance with $(3,6)\,df$ then, the fuzzy null hypotheses $\tilde{H}_0^L$ is accepted for the $r; 0 \le r \le 1$. Furthermore, the difference between treatments is not significant. Therefore, the groundnut varieties are not substantially different in terms of yields.

Let us consider the upper level model is given below

4.4 Upper Level Model

Table – 5. Table for Upper Level Model

Varieties of Groundnut	Yields in kilograms			
	Y1	Y2	Y3	Y4
TMV 2	$-2h+60$	$-4h+62$	$-3h+59$	$-4h+62$
TMV 7	$-2h+62$	$-5h+63$	$-3h+62$	$-3h+63$
VRI 2	$-3h+65$	$-2h+62$	$-3h+63$	$-2h+61$

Likewise, upper level models of ambiguous *RBD* use formula and table, thus avoiding calculation.

4.5 Fuzzy Decision Rule of Between Treatments

If $F_T > \tilde{F}_h^U$, for all $h; 0 \le h \le 1$ where $F_T = 5.14$ is the F table value of α at 5% level of significance with $(2,6)\,df$ then, the fuzzy null hypotheses $\tilde{H}_0^U$ is accepted for the $h; 0 \le h \le 1$.

Consequently, the disparity between treatments is not significant. Therefore, with regard to yields in kilograms, groundnut varieties do not vary significantly.

4.6 Fuzzy Decision Rule of Between Blocks

If $F_T > \tilde{F}_h^U$, for all $h; 0 \le h \le 1$ where $F_T = 8.14$ is the F table value of α at 5% level of significance with $(6,3)\,df$ then, the null hypothesis of the $\tilde{H}_0^U$ is accepted for the $h; 0 \le h \le 1$. Therefore, the discrepancy between treatments is not significant. Consequently, with regard to yields in kilograms, groundnut varieties do not vary significantly.

Therefore, so because fuzzy null hypotheses between treatments $\tilde{H}_0^L$ and $\tilde{H}_0^U$ of the lower level data is rejected and upper level data is accepted for all $r; 0 \le r \le 1$ and $h; 0 \le h \le 1$ (note that null hypotheses of accepted or rejected at $r = 1$ and $h = 1$, that is the centre level), the between blocks of fuzzy null hypothesis $\tilde{H}_0^L$ and $\tilde{H}_0^U$ of the lower and upper level data is accepted for all $r; 0 \le r \le 1$ and $h; 0 \le h \le 1$ the fuzzy null hypotheses $\tilde{H}_0$ of the fuzzy *RBD* model is accepted and rejected for all $r; 0 \le r \le 1$ and $h; 0 \le h \le 1$. Thus, we conclude that four yields of groundnut kilograms are equal if $r; 0 \le r \le 1$ and $h; 0 \le h \le 1$.

V. CONCLUSION

A statistical test of the hypothesis for *RBD* model using *TFN* for fuzzy data is suggested in this study. Can make a decision on the fuzzy *RBD* model hypothesis based on the hypothesis determinations of two crisp *RBD* models. Since our fuzzy test is rather than standard significance tests, it appears to be a useful method in circumstances with imprecise data and also extend the crisp *RBD* to *LSD*, *BIBD*, *PBIBD* for fuzzy environments.

REFERENCES

[1] M. Arefi and S.M. Taheri, Testing Fuzzy Hypotheses Using Fuzzy Data Based on Fuzzy Test Statistic, Journal of Uncertain Systems Vol.5, No.1, pp.45-61, (2011).

[2] S. Chanas, On the interval approximation of a fuzzy numbers, Fuzzy sets and Systems, 102, (1999), 221-226.

[3] 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).

[4] D. Dubois and H. Prade, Fuzzy Sets and Systems: Theory and Application, Academic Press, New York, (1980).

[5] D. Dubois and H. Prade, Operations on fuzzy numbers, International Journal of System Science, 9, (1978), 613-626.

[6] P. Filzmoser and R. Viertl, Testing Hypotheses with Fuzzy Data: The Fuzzy P value, Metrika, 59, (2009), 21-29.

[7] P. Gajivaradhan and S. Parthiban, A Comparative Study of Two Factor ANOVA Model Under Fuzzy Environments using Trapezoidal Fuzzy Numbers, International Journal of Fuzzy Mathematical Archive, Vol. 10, No. 1, (2016), 1-25.

[8] George.J. Klir and Bo Yuan, Fuzzy sets and fuzzy logic, Theory and Applications, Prentice-Hall, New Jersey, (2008).

[9] Gil et al. Bootstrap Approach to the Multi-sample Test of Means with Imprecise Data, Computer Statistics and Data Analysis 51, (2006), 148-162.

[10] S.C. Gupta and V.K. Kapoor, Fundamentals of applied statistics, Sultan Chand & Sons, New Delhi, India, (2007).

[11] R.R. Hocking, Methods and applications of linear models: regression and the analysis of variance, New York: John Wiley & Sons, (1996).

[12] K.G. Manton et al., Statistical applications using fuzzy sets, New York: John Wiley & Sons, (1994).

[13] A. Mariappan and M. Pachamuthu, Statistical Testing of Hypotheses for Fuzzy Completely Randomized Design, Journal of Xidian University, Vol. 14, Issue 9, (2020).

[14] Mikihiko Konishi, Tetsuji Okuda and Kiyoji Asai, Analysis of Variance based on Fuzzy Interval Data using Moment Correction Method, International Journal of Innovative Computing, Information and Control, 2, (2006), 83-99.

[15] T. Nakama, A. Colubi, and M.A. Lubiano, Two-Way Analysis of Variance for Interval-Valued Data, Combining Soft Computing & Stats. Methods, AISC 77, pp. 475–482, (2010).

[16] H.T. Nguyen and E.A. Walker, A First Course in Fuzzy Logic (3rd ed.), Paris: Chapman Hall/CRC, (2005).

[17] R. Viertl, “Univariate statistical analysis with fuzzy data”, Computational Statistics and Data Analysis, Vol. 51, 2006, 33-147.

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

[19] H.C. Wu, Statistical hypotheses testing for fuzzy data, Information Sciences, 175, (2005), 30-56.

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

[21] L.A. Zadeh, Fuzzy sets, Information and Control, 8, (1965), 338-353.

[22] L.A. Zadeh, The Concept of a Linguistic Variable and its application to approximate Reasoning, Information Sciences, 8, (1975), 199-249.