

A Fuzzy Approach To The Test Of Hypothesis Using Pe Fuzzy Number

¹* K.S. Keerthika & ²S. Parthiban

^{1,2}Division of Mathematics, Department of Science and Humanities, Vignan's Foundation for Science, Technology and Research, Vadlamudi, Guntur, Andhra Pradesh, India.

E-mail: ¹keerthikasundaram.29@gmail.com, ²sparthiban2008@gmail.com

Abstract

Testing and deciding on hypotheses is a general and meaningful final part in all areas. Getting better results by testing statistical hypotheses is crucial in almost any sector such as industry, education, elections, transport etc. However, accurate sample observations are not always possible when the job is a real-time population. In practice, the observed models may not be correct. In the proposed work, imprecise rainfall data was collected in southern and northern India and the relevant samples were observed as fuzzy numbers. More generally compared to Pentagonal Fuzzy Numbers (PnFN). In addition, fuzzy numbers are eliminated by clear classification via the pivot points of PnFN. After the fuzzy process, when testing the hypotheses, a relevant statistical method was used to make a better decision.

Keywords: Test of Hypothesis, Fuzzy numbers, Pivotal Spot, Rank, Samples.

Introduction

The real decision problems are in most cases uncertain or vague. Fuzzy numbers are used in various fields i.e, fuzzy process modelling, control theory, decision making, reasoning by expert systems, etc., they are not intuitive. Since fuzzy numbers are represented by possibility distributions, they can overlap and therefore cannot be ordered. It is true that fuzzy numbers are generally sub-ranges and cannot be compared to real numbers, which can be ordered linearly. To classify the fuzzy quantities, each fuzzy quantity is converted to a real number, and a classification function is defined from the set of fuzzy numbers and compared to a set of real numbers corresponding to each fuzzy number in natural order.

Typically, a lot of information is lost if the entire scan is reduced to a single number. So we have to try to minimize this loss. Zadeh [20] first introduced fuzzy set theory, several classification methods were developed. The analysis of fuzzy numbers was first proposed by Jain [6] for decision making in fuzzy situations by representing an incorrectly defined quantity as an ambiguous set. Dubois and Prade [3] gave an ambiguous numerical mean. Lee and Li [9] presented a fuzzy number comparison based on the fuzzy event probability measure. Wu [17-19] suggested different possibilities to extend some fuzzy intervals for the fuzzy-fuzzy parameter. B. F. Arnold [1] examined the hypothesis test with accurate data. Y. M. Wang [16] examined the centres of gravity of fuzzy numbers. Apurba Panda et al., [12] carried out a study on the pentagonal fuzzy number and its corresponding matrices. J. Chachi et al., [2], performed a hypothesis test based on fuzzy confidence intervals. Emrah Akyar et al., [4], developed a new method for classifying triangular fuzzy numbers and the advantages of explaining them. S. Parthiban et al., did a lot of research on hypothesis testing with different fuzzy numbers [7, 10, 11]. Kok Chin Chai et al., proposed a new sign distance-based ranking method for fuzz number. Y.L.P. Thorani et al., [14] generalized trapezoidal fuzzy numbers. R. Helen [5] found a new operation and classification in pentagonal fuzzy numbers. G. Uthra et al., [15] defined a generalized intuitionistic pentagonal fuzzy number and proposed a new classification formula. P.Phani Bushan Rao et al., [13] used the distance method using the centre of the perimeter of the centroids and a modality index to find the classification of a fuzzy number. The excluded examples provide further measurable fuzzy theory proof for two-tailed t-tests with information about pentagonal fuzzy numbers. Another thought in this article is that we have an idea for research, and what will be the consequences when we reach the level of this relaxed

information that is essential? Therefore, we use positions in theoretical experiments derived from the suggested critical positions of pentagonal fuzzy numbers. Confidence levels and h-level phrases are not used in the selection guidelines for the proposed test procedure.

Preliminaries

Definition 2.1. A fuzzy set \tilde{A} of a universal X is characterized by its membership function $\mu_{\tilde{A}} : X \to [0, 1]$

and we write $\tilde{A} = \left\{ (x, \mu_{\tilde{A}}(x)) \colon x \in X \right\}$.

Definition 2.2. A fuzzy set \tilde{A} defined on the universal set of real number R, whose membership function μ is called the fuzzy number if it has the following characteristics.

- 1. \tilde{A} is convex.
- 2. \tilde{A} is normal.
- 3. \tilde{A} is piecewise continuous.

Definition 2.3. A Normalized Pentagonal Fuzzy Number (NPnFN) of a fuzzy set \tilde{A} as $\tilde{A}_p = p, q, r, s, t$ and its membership function is given by

$$\mu(x) = \begin{cases} 0 & \text{for } x t \end{cases}$$



Figure 1: Pentagonal Fuzzy Number

Hypothesis Test for Interval Data

$$H_0:[d_1, \mu_1] = [d_2, \mu_2] P n_1 = \mu_1 and n_2 = \mu_2$$

And the alternative hypothesis is given by

(i)
$$H_A: [d_1, \mu_1] = [d_2, \mu_2] \mathbb{P} \ d_1 < \mu_1 \text{ and } d_2 < \mu_2$$

(ii) $H_A: [d_1, \mu_1] = [d_2, \mu_2] \mathbb{P} \ d_1 > \mu_1 \text{ and } d_2 > \mu_2$
(iii) $H_A: [d_1, \mu_1] = [d_2, \mu_2] \mathbb{P} \ d_1^{-1} \ \mu_1 \text{ and } d_2^{-1} \ \mu_2$

Suppose that \tilde{x} and \tilde{y} are sample means, $S_{\tilde{x}}$ and $S_{\tilde{y}}$ are the sample standard deviations of $S_{\tilde{x}}$ and $S_{\tilde{y}}$ correspondingly.

(i) Let presumed the population standard deviation is to be same, then the null hypothesis

 $H_{\scriptscriptstyle 0}\!:\![\delta_{\scriptscriptstyle 1},\mu_{\scriptscriptstyle 1}]\!=\![\delta_{\scriptscriptstyle 2},\mu_{\scriptscriptstyle 2}]$ test will be

$$\tilde{t} = \frac{\tilde{x} - \tilde{y}}{s\sqrt{\frac{1}{m} + \frac{1}{n}}}$$
(3.1)

where

$$\tilde{s} = \sqrt{\frac{(m-1)s_{\tilde{x}}^2 + (n-1)s_{\tilde{y}}^2}{m+n-2}}$$
(3.2)

(ii) Let presumed the population standard deviation is to be in-equal, then the null hypothesis $H_0: [\delta_1, \mu_1] \neq [\delta_2, \mu_2]$ test will be

$$\tilde{t} = \frac{\tilde{x} - \tilde{y}}{\sqrt{\frac{s_{\tilde{x}}^2}{m} + \frac{s_{\tilde{y}}^2}{n}}}$$
(3.3)

where

$$s_{\tilde{x}}^{2} = \left(\frac{1}{m-1}\right) \left(\sum_{i=1}^{m} \left(\tilde{x}_{i} - \tilde{\overline{x}}\right)^{2}\right),$$

$$s_{\tilde{y}}^{2} = \left(\frac{1}{n-1}\right) \left(\sum_{j=1}^{n} \left(\tilde{y}_{j} - \tilde{\overline{y}}\right)^{2}\right)$$
(3.4)

Here, the parameters are denoted with tilde (\sim) symbols as they are defuzzified form connected with NPnFN.

My Proposed Work

Consider a Normalized pentagonal fuzzy number (NPnFN) $\tilde{p} = (g_1, g_2, g_3, g_4, g_5; \delta)$. The pentagon is divided as three plane figures, where PQS is a triangle, QSTU is a quadrilateral and QUR is a



Figure 2: Centroid of Pentagonal Fuzzy Number based Euler Line of Centroids

 D^{le} respectively. Let C_1, C_2 and C_3 will be the centroid of these three planes and this is taken as a point of reference to determine the classification of NPnFN.

In this proposed work, the calculation of the centroid of pentagon and the Euler line that passes through the planes C_1, C_2 and C_3 is considered to find the best balancing point for NPnFN. From the figure: 2. The centroids of these three plane figures are

$$C_1 = \left(\frac{g_1 + g_2 + g_3}{3}, \frac{\delta}{3}\right), C_2 = \left(\frac{g_2 + g_3 + g_4}{3}, \frac{\delta + 1}{3}\right), C_3 = \left(\frac{g_3 + g_4 + g_5}{3}, \frac{\delta}{3}\right)$$

The line connecting C_1C_3 is $\delta/3$ and C_2 does not lie on C_1C_3 . This is because C_1, C_2 and C_3 are not on the same line. Equation C_1, C_2 is $\delta/3$ and C_2 didn't lie on the line C_1C_2 . So C_1C_2 and C_3 are not collinear and form a triangle. We derive the Euler line from the triangle with the vertices C_1, C_2 and C_3 from NPnFN $\tilde{A}=(g_1, g_2, g_3, g_4, g_5; \delta)$ as

$$y - \frac{(12\delta + (3g_1 - 3g_4)g_5 + 3g_2g_4 - 3g_1g_2 + 5)}{(5g_5 + 7g_4 + 12g_3 + 7g_2 + 5g_1)} x = \frac{+g_1g_2 - 1)g_4 + (3g_1g_2 - g_1^2)g_5 - 2g_2g_4^2 + (-3g_2g_3 - 2g_2^2)g_5 - 2g_2g_4 - 2g_2g_4$$

and the equation of the line joining ${\it T}$ and ${\it G}\,$ is

$$y + \frac{(2x)}{(g_1 + g_5 - 2g_3)} = \frac{(g_1 + g_5 + g_3)}{(g_1 + g_5 - 2g_3)}$$
(4.2)

The point of intersection of (4.1) and (4.2) is evaluated as

$$S_{\tilde{A}}(\bar{x}_0, \bar{y}_0) = [A_1, B_1]$$
(4.3)

where

$$\begin{aligned} &((g_5^2 + (-g_4 - 2g_3 - g_2 + 2g_1)g_5 + (2g_3 - g_1)g_4 + (2g_2 - 2g_1)g_3 - g_1g_2 + g_1^2)\delta \\ &+ (g_4 - g_1)g_5^3 + (2g_4^2 + (g_3 + g_2)g_4 - g_1g_3 - g_1g_2 - 2g_1^2 - 15)g_5^2 + ((-4g_3 - 2g_2)g_1 + 2g_1)g_4^2 + (-6g_3^2 + (5g_1 - 5g_2)g_3 - 2g_2^2 + 2g_1g_2 - g_1^2 - 22)g_4 + 6g_1g_3^2 + (5g_1g_2)g_4^2 + (2g_2g_3^2 + 2g_1g_2 - g_1^2 - 22)g_4 + 2g_1g_2^2 + 2g_1g_2 - 2g_1g_2 + 2g_1g_2 - 2g_1g_2 + 2g_1g_2 - 2g_1g_2 + 2g_1g_2 - 2g_1g_2 + 2g_1g_2 - 2g_1g_2)g_4^2 + (6g_2g_3^2)g_4 + (4g_2^2 - 5g_1g_2 + 23)g_3 - 2g_1g_2^2 + g_1^2g_2 - 22g_1)g_4 + (38 - 6g_1g_2)g_3^2 + (-4g_1g_2^2)g_4 + (4g_2^2 - 5g_1g_2 - 22g_1)g_3 + 2g_1^2g_2^2 + (g_1^3 - 22g_1)g_2 - 15g_1^2) \\ &((36g_5 - 72g_3 + 36g_1)\delta + (9g_1 - 9g_4)g_5^2 + ((18g_3 + 9g_2 - 9g_1)g_4 + (18g_1g_2 + 42)g_3 + (42 - 9g_1^2)g_2 + 45g_1) \end{aligned}$$

$$\begin{split} &(-(-38g_5-2g_4-36g_3-2g_2+38g_1)\delta-(7g_1-7g_4)g_5^2\\ &-(4g_4^2+(15g_3+11g_2-11g_1)g_4-15g_1g_3-11g_1g_2+7g_1^2\\ &+15)g_5+4g_2g_4^2-(-15g_2g_3-4g_2^2+11g_1g_2-2)g_4\\ B_1 = \frac{-(15g_1g_2-17)g_3-4g_1g_2^2-(-7g_1^2-2g_2)-15g_1)}{((36g_5-72g_3+36g_1)\delta+(9g_1-9g_4)g_5^2+((18g_3+9g_2)-9g_1)g_4-18g_1g_3-9g_1g_2+9g_1^2+45)g_5+(-18g_2g_3)\\ &+9g_1g_2+42)g_4+(18g_1g_2+42)g_3+(42-9g_1^2)g_2+45g_1) \end{split}$$

The ranking function of the NPnFN $\tilde{A}=(g_1,g_2,g_3,g_4,g_5;\delta)$ which maps the set of all fuzzy numbers to a set of all real numbers is defined as,

$$R(\tilde{\mathbf{A}}) = r_x^{\tilde{\mathbf{A}}} \ge r_y^{\tilde{\mathbf{A}}}$$
$$= \{\mathbf{A}_1 \ge \mathbf{B}_1\}$$
(4.4)

This is the area between the original point and the point of intersection of the Euler line and the vertical line joining the midpoints of the horizontal lines of the NPnFN.

Numerical Example

Here the data considered is the total rainfall happened in the North and South India. The study on intensity of the rainfall based on five stages is recorded as normalized pentagonal fuzzy numbers that has been given below. For this problem, the test of hypothesis has been conducted to test the homogeneity of the severity of rainfall in to regions. Here given data are in the form of NPnFN in Table: 1. These NPnFN are transformed to interval data by using pivotal spot defuzzifying formula (4.3)

(4	•	3)	•	

South India	South India Rainfall	North India	North India Rainfall
States		States	
Andhra Pradesh	$\tilde{F}_1 = (0.2, 0.3, 0.5, 0.7, 0.8)$	Chattisgarh	$\tilde{F}_8 = (0.1, 0.2, 0.4, 0.6, 0.7)$
Andaman & Nicobar	$\tilde{F}_2 = (0.2, 0.4, 0.6, 0.9, 0.95)$	Delhi	$\tilde{F}_9 = (0.1, 0.2, 0.4, 0.5, 0.6)$
Karnataka	$\tilde{F}_3 = (0.3, 0.5, 0.7, 0.75, 0.8)$	Haryana	$\tilde{F}_{10} = (0.1, 0.2, 0.3, 0.5, 0.8)$
Kerala	$\tilde{F}_4 = (0.5, 0.6, 0.7, 0.8, 0.9)$	Himachal Pradesh	$\tilde{F}_{11} = (0.3, 0.4, 0.5, 0.7, 0.9)$
Lakshadweep	$\tilde{F}_5 = (0.3, 0.4, 0.5, 0.6, 0.7)$	Jammu & Kashmir	$\tilde{F}_{12} = (0.15, 0.2, 0.4, 0.5, 0.6)$
Puducherry	$\tilde{F}_6 = (0.2, 0.3, 0.4, 0.7, 0.8)$	Punjab	$\tilde{F}_{13} = (0.05, 0.125, 0.3, 0.35, 0.4)$
Tamil Nadu	$\tilde{F}_7 = (0.1, 0.2, 0.3, 0.4, 0.5)$	Rajasthan	$\tilde{F}_{14} = (0.2, 0.3, 0.4, 0.6, 0.8)$
		Uttarkhand	$\tilde{F}_{15} = (0.1, 0.3, 0.4, 0.7, 0.9)$
		Uttar Pradesh	$\tilde{F}_{16} = (0.3, 0.4, 0.5, 0.8, 0.85)$



Hypothesis Test:

The Null Hypothesis:

 \tilde{H}_0 : The intensity of the rainfall between two sides.

The Alternative Hypothesis:

 \hat{H}_A : There is a significant difference between the rainfalls. The test statistics are given in the Table: 2. Based on the interval data calculated above, the test of hypothesis will be evaluated as per the test statics given in section: 5.2.

South	n India Rain Fall (x)	North	India Rain Fall (y)
$ ilde{F}_1$	[-0.2500, -0.1354]	$ ilde{F}_8$	[-0.2000, -0.1354]
$ ilde{F}_2$	[-0.2782, -0.1266]	$ ilde{F}_9$	[-0.1554, -0.1094]
$ ilde{F}_3$	[-0.2121, -0.0807]	$ ilde{F}_{10}$	[-0.2686, -0.2090]
$ ilde{F}_4$	[-0.3500, -0.1288]	$ ilde{F}_{11}$	[-0.3334, -0.1663]
$ ilde{F}_5$	[-0.2500, -0.1288]	$ ilde{F}_{12}$	[-0.1781, -0.1251]
$ ilde{F}_6$	[-0.2834, -0.1659]	$ ilde{F}_{13}$	[-0.0804, -0.0725]
$ ilde{F}_7$	[-0.1500, -0.1288]	$ ilde{F}_{14}$	[-0.2827, -0.1732]
		$ ilde{F}_{15}$	[-0.2830, -0.1693]
		$ ilde{F}_{16}$	[-0.3137, -0.1509]

Table 2: Interval Data of Fuzzy Samples

Rank of Normalized Pentagonal Fuzzy Numbers

The rank of the NPnFN is defined by its geometric mean of the *pivotal points*: $\tilde{\overline{x}}_N$ and $\tilde{\overline{y}}_S$,

 $R(\tilde{F}) = \sqrt{\tilde{x}^2 + \tilde{y}^2}$. Now, the ranks of NPnFN will be evaluated for the above numerical example by using the following formula:

$$R(\tilde{F}) = \sqrt{\tilde{x}^2 + \tilde{y}^2}$$

The calculated rank values are tabulated below,

South India Rain Fall (x)	North India Rain Fall (y)
0.284312	0.241523
0.305651	0.190046
0.226934	0.340334
0.372947	0.372574
0.281228	0.217646
0.328388	0.108261
0.197710	0.331538
	0.329775
	0.348107

Table 3: Interval Data of Fuzzy Samples

$$\tilde{\bar{x}}_{S} = \left(\frac{1}{m}\right) \left(\sum_{i=1}^{m} \left(\tilde{x}_{i_{S}}\right)\right) = 0.285310; \quad \tilde{\bar{y}}_{N} = \left(\frac{1}{n}\right) \left(\sum_{j=1}^{n} \left(\tilde{y}_{j_{N}}\right)\right) = 0.275534$$

where $\tilde{\overline{x}}_{S}$ and $\tilde{\overline{y}}_{N}$ are the extremes of the given interval data.

$$s_{\tilde{x}_{S}}^{2} = \left(\frac{1}{m-1}\right) \left(\sum_{i=1}^{m} \left(\tilde{x}_{i_{S}} - \tilde{x}_{S}\right)^{2}\right) = 0.000501$$
$$s_{\tilde{y}_{N}}^{2} = \left(\frac{1}{n-1}\right) \left(\sum_{j=1}^{n} \left(\tilde{y}_{j_{N}} - \tilde{y}_{N}\right)^{2}\right) = 0.000899$$

$$\tilde{t} = \frac{\frac{\bar{x}_{S} - \bar{y}_{N}}{\sqrt{s_{\tilde{x}_{S}}^{2} + \frac{s_{\tilde{y}_{N}}^{2}}{m}}} = 56.99574$$

The tabulated value of 't' at 5 % significance level with 14 degree of freedom is $t_{\alpha} = 1.76$

Conclusion

The "calculated value" of 't' is |t| = 56.99574 and the "tabulated value" of 't' is $t_{\alpha} = 1.76$. Here,

 $|t| > t_{\alpha} \Rightarrow$ The null hypothesis H_0 ' is rejected and it can be concluded that there is significant difference between the rainfalls. While the best solution can be obtained using a classification method derived from the fuzzy key point, this is not the final solution for all problems. Furthermore, it opened the door to test the hypothesis involving PnFN with their defuzzified forms as range, etc. And, of course, it requires further refinement and research to get a more accurate result.

Reference

- 1. B. F. Arnold, "Testing fuzzy hypotheses with crisp data", Fuzzy Sets and Systems, Vol. 94, pp. 323-333, 1998.
- 2. 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.
- 3. D. Dubois and H. Prade, "The mean value of a fuzzy number", Fuzzy Sets and Systems, Vol. 24, pp. 279-300, 1987.
- 4. Emrah Akyar, Handan Akyar and Serkan Ali Duzce, "A new method for ranking triangular fuzzy numbers", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 20, pp. 729-740, 2012.
- 5. R. Helen and G. Uma, "A new operation and ranking on Pentagonal Fuzzy Numbers, International Journal of Mathematical Sciences and Applications", Vol. 5, no. 2, pp. 341-346, 2015.
- R. Jain, "Decision making in the presence of fuzzy variable", IEEE Transactions on Systems Man, and Cybrenctics, Vol. 6, pp. 698 – 703, 1976.
- 7. K.S. Keerthika and S. Parthiban, "Statistical hypothesis test to the fuzzy samples from biomedical observations by pivotal spot of trapezoidal fuzzy numbers", Journal of Green Engineering, Vol. 10, pp. 1218-1231, 2020.
- 8. Kok Chin Cha, Kai Meng Tay and Chee Peng Lim, "A new sign distance-based ranking method for fuzzy numbers", Proceedings of the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems, Vol.2, pp.55-64, 2015.

- 9. E.S.Lee and R-J.Li, "Comparison of fuzzy numbers based on the probability measure of fuzzy events", Computers & Mathematics with Applications, Vol. 15, pp. 887-896, 1998.
- S. Parthiban and P. Gajivaradhan, "Statistical Hypothesis on Industrial Applications through Ranks from COG of TrFNs", International Journal of Recent Technology and Engineering (IJRTE), Vol. 6, pp. 1116-1118, 2019.
- 11. S. Parthiban and P. Gajivaradhan, "Statistical hypothesis test in three factor ANOVA model under fuzzy environments using trapezoidal fuzzy numbers", Bulletin of Mathematical Sciences and Applications, Vol.14, pp. 23-42, 2016.
- 12. A. Panda and MadhumangalPal, "A study on pentagonal fuzzy number and its corresponding matrices", Pacific Science Review B: Humanities and Social Sciences, Vol. 1, pp. 131-139, 2015.
- 13. Phani Bushan Rao and N. Ravi Shankar, "Ranking Fuzzy Numbers with a Distance Method using Circumcenter of centroids and an Index Modality", Advances in Fuzzy Systems, pp. 1-7, 2011.
- 14. Y. L. P. Thorani, P. Phani Bushan Rao and N. Ravi Shankar, "Ordering Generalized Trapezoidal Fuzzy Numbers", International Journal Contemporary Mathematical Sciences, Vol. 7, no. 12, pp. 555-573, 2012.
- 15. G. Uthra, K. Thangavelu, S. Shunmugapriya, "Ranking Generalized Intuitionistic Pentagonal Fuzzy Number by Centroid Approach", Vol. 5, no. 4, pp. 589-593, 2017.
- 16. Y. M. Wang, J.B. Yang, D. L. Xu and K. S. Chin, "On the centroids of fuzzy numbers", Fuzzy Sets and Systems, Vol. 157, pp. 919-926, 2006.
- 17. H. C. Wu, "Analysis of variance for fuzzy data", International Journal of Systems Science, Vol. 38, pp. 235-246, 2007.
- 18. H. C. Wu, "Statistical confidence intervals for fuzzy data", Expert Systems with Applications, Vol. 36, pp. 2670-2676, 2009.
- 19. H. C. Wu, "Statistical hypotheses testing for fuzzy data", Information Sciences, Vol. 175, pp. 30-56, 2005.
- 20. L. A. Zadeh, "Fuzzy sets", Information and Control, Vol. 8, pp. 338-353, 1965.