Modelling And Initiating Knowledge Manegement Program Using Corralation Coeficient Between Fuzzy Numbers

Rahim SANEİFARD*

Department of Applied Mathematics, Urmia Branch Islamic Azad University, Oroumieh, IRAN

*Corresponding Author

e-mail: srsaneeifard@yahoo.com

Abstract

When we deal with crisp data, it is very common to find the correlation between variables. Here, we propose a method to calculate the correlation coefficient for fuzzy data, but rather than defining the correlation on the intuitionistic fuzzy sets like most of the previous works, we adopt the method from central interval. This interval can be used as a crisp set approximation with respect to a fuzzy quantity. This indices can be applied for comparison of fuzzy numbers namely fuzzy correlations in fuzzy environments and expert's systems. Finally some of their applications are mentioned.

Keywords : Fuzzy number, Correlation, Central interval, Modeling.

INTRODUCTION

Fuzzy numbers provide formalized tools for dealing with non-precise quantities possessing non-random imprecision or vagueness. An overview of important developments in the theory and applications of fuzzy numbers can be found in [1]. In order to summarize information contained in a fuzzy number, several characteristics of fuzzy numbers have been proposed. Some of them, e.g., the mean value, modal value or variance, have been motivated by descriptive statistics used in statistical analysis of crisp data. Dubois and Prade [2] defined the interval-valued mean of a fuzzy number under the random set interpretation of fuzzy quantities. Carlsson and Fuller [3] studied fuzzy numbers as possibilistic distributions. They introduced the notion of the possibilistic mean value and the variance of fuzzy numbers. Weighted possibilistic mean, variance, and covariance of fuzzy numbers can be found in the work of Fuller and Majlender [4]. Ralescu [5] defined the average level of a fuzzy set and applied this concept to fuzzy numbers. Any real number with the membership grade equal to 1 in a fuzzy number A is frequently called a modal value of A. Some other characterizations of fuzzy numbers can be found in [6]. In statistics, the center of distribution of a quantitative variable is characterized besides the mean and the modal value by the median also. The median value of a fuzzy number A was defined in [1] as the numerical value from the support of A that equally divides the area under the membership function of A. It is well known that the area under the membership function of A represents the cardinality of A. We will classify fuzzy numbers with respect to the distribution of their cardinality into four categories: fuzzy numbers with equally heavy tails, fuzzy numbers with light tails, fuzzy numbers with a heavy right tail and fuzzy numbers with a heavy left tail. We will use formulas for the location of the median value for trapezoidal fuzzy numbers and for some of their modifications. The interval-valued probabilistic mean and the interval-valued possibilistic mean are based on the lower and the upper means of a fuzzy number. We will define the interval-valued median (median interval) in a similar way. We will propose the lower and the upper median of a fuzzy number A with respect to the cardinality of the left and the right tails of A, respectively. We will investigate when the interval-valued median is a subset of the interval-valued possibilistic mean and when this relationship is reversed. Also, we propose a method to calculate the correlation coefficient for fuzzy data, but rather than defining the correlation on the intuitionistic fuzzy sets like most of the previous works, we adopt the method from central interval. This interval can be used as a crisp set approximation with respect to a fuzzy quantity. This indices can be applied for comparison of fuzzy numbers namely fuzzy correlations in fuzzy environments and expert's systems.

BASIC DEFINITIONS AND NOTATIONS

The basic definitions of a fuzzy number are given in [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] as follows.

Definition 2.1. A fuzzy number A is a mapping $\mu_A(x) : \Re \rightarrow [0,1]$ with the following properties:

1. μ_A is an upper semi-continuous function on $\Re$.

2. $\mu_A(x) = 0$ outside of some interval $[a_1, b_2] \subset \Re$.

3. There are real numbers a_1, a_2, b_1 and b_2 such that $a_1 \le a_2 \le b_1 \le b_2$ and

   3.1 $\mu_A(x)$ is a monotonic increasing function on $[a_1, a_2]$,

   3.2 $\mu_A(x)$ is a monotonic decreasing function on $[b_1, b_2]$,

   3.3 $\mu_A(x) = 1$ for all x x in $[a_2, b_1]$.

The set of all fuzzy numbers is denoted by F. Let $\Re$ be the set of all real numbers. We assume a fuzzy number A that can be expressed for all $x \in \Re$ in the form

$$A(x) = \begin{cases} g(x) & when \ x \in [a,b) \\ w & when \ x \in [b,c] \\ h(x) & when \ x \in (c,d] \\ 0 & otherwise. \end{cases} \quad (2.1)$$

where a, b, c and d are real numbers such that $a < b \le c < d$ and g and h are real valued functions such that g is increasing and right continuous and h is decreasing and left continuous. Based on the basic theories of fuzzy numbers, A is a normal fuzzy number if $w = 1$, whereas A is a non-normal fuzzy number if $0 < w \le 1$. Notice that (2.1) is an R fuzzy number. A normal fuzzy number A with shape function g and h defined by

$$g(x) = \left(\frac{x-a}{b-a}\right)^n. \quad (2.2)$$

And

$$h(x) = \left(\frac{d-x}{d-c}\right)^n, \quad (2.3)$$

respectively, where $n > 0$, will be denoted by $A = \langle a, b, c, d \rangle_n$. If A be non-normal fuzzy number, it will be denoted by $A = \langle a, b, c, d; w \rangle_n$. If $n = 1$, we simply write $A = \langle a, b, c, d \rangle$, which is known as a normal

trapezoidal fuzzy number and if $b = c$, is known as a normal triangular fuzzy number and represented by $A = \langle a, b, d \rangle$. If $n \ne 1$, a fuzzy number $A = \langle a, b, c, d \rangle_n$ is a concentration of A. If $0 < n < 1$, then A^* is a dilation of A. Concentration of A by $n = 2$ is often interpreted as the linguistic hedge "very". Dilation of A by $n = 0.5$ is often interpreted as the linguistic hedge "more or less". More about linguistic hedges can be found in [20].

Each fuzzy number A described by (2.1) has the following α-level sets (α-cuts) $A_\alpha = [a_\alpha, b_\alpha]$, $a_\alpha, b_\alpha \in \Re, \alpha \in [0,1]$.

1. $A_\alpha = [g^{-1}(\alpha) \ h^{-1}(\alpha)]$ for all $\alpha \in (0,1)$,

2. $A_1 = [b, c]$,

3. $A_0 = [a, d]$.

If $A = \langle a, b, c, d \rangle_n$ then for all $\alpha \in [0,1]$,

$$A_\alpha = \left[a + \alpha^{\frac{1}{n}}(b-a) \ d - \alpha^{\frac{1}{n}}(d-c) \right]. \quad (2.4)$$

Another important notion connected with fuzzy number A is a cardinality of a fuzzy number A. In this paper, the researchers will always refer to fuzzy number A described by (2.1).

MEDIAN INTERVAL AND CENTRAL VALUE

Various authors have studied the mean interval of a fuzzy number, also called the intervalvalued mean [2,3]. According to Dubios and Prade [2], the interval-valued probabilistic mean of a fuzzy number A with α-cuts $A_\alpha = [a_\alpha, b_\alpha]$ $\alpha \in [0,1]$ is the interval $E(A) = [E_*(A) \ E^*(A)]$, where

$$E_*(A) = \int_0^1 a_\alpha d\alpha \text{ and } E^*(A) = \int_0^1 b_\alpha d\alpha. \quad (3.5)$$

Carlsson [3] introduced the interval-valued possibilistic mean of a fuzzy number A as the interval $M(A) = [M_*(A) \ M^*(A)]$. The lower possibilistic mean $M_*(A)$, is the weighted average of the minima of the α-cuts of A. Similarly, the upper possibilistic mean $M^*(A)$ is the weighted average of the maxima of the α-cuts of A. If A is a fuzzy number characterized by (2.1), then

$$M_*(A) = 2\int_0^1 f(\alpha)a_\alpha d\alpha \qquad \text{and}$$

$$M^*(A) = 2\int_0^1 f(\alpha)b_\alpha d\alpha . \qquad (3.6)$$

and the function $f(\alpha)$ is nonnegative and increasing on $[0,1]$ with $f(0) = 0$ and $\int f(\alpha)d\alpha = 1$. The function $f(\alpha)$ is also called weighting function. They also proved that if A is a fuzzy number of LR type with strictly monotonous and continuous shape functions, then $M(A) \subset E(A)$. This reflects on the fact that real numbers with small membership grades in A are considered to be less important in the definition of lower and upper possibilistic mean value in the definition of probabilistic ones. We will introduce the median interval (interval-valued median) of A similarly to the definition of $E(A)$ and $M(A)$.

Definition 3.1. [20]. Let A be a fuzzy number characterized by (2.1). Let $m_L \in (a,b)$ and $m_R \in (c,d)$ be such that

$$\int_a^{m_L} A(x)dx = \int_{m_L}^b A(x)dx ,$$

and

$$\int_c^{m_R} A(x)dx = \int_{m_R}^d A(x)dx$$

respectively. Then $M_e(A) = [m_L, m_R]$ is called the median interval (interval-valued median) of A. For a trapezoidal fuzzy number A and for its modifications by selected linguistic hedges we will provide formulas for the location of the median interval.

Proposition 3.1. [20]. Let $A = \langle a,b,c,d \rangle_n$. Then $M_e(A) = [m_L, m_R]$, where

$$m_L = a + \frac{b-a}{\sqrt[n+1]{2}} ,$$

and

$$m_R = d - \frac{d-c}{\sqrt[n+1]{2}} .$$

Corollary 3.1. [20]. Let $A = \langle a,b,c,d \rangle_n$. Then $M_e(A) = A_\alpha$, where $\alpha = 2^{\frac{-n}{n+1}}$ and $\alpha \in (0.5,1)$. The median value m_A is always in the median interval $M_e(A)$.

Corollary 3.2. [3]. Let $A = \langle a,b,c,d \rangle$ be trapezoidal fuzzy number. Then $M_e(A) = A_\alpha$, where $\alpha = 0.707$.

In general, if A can not be expressed in the form $A = \langle a,b,c,d \rangle_n$, then $M_e(A)$ is not an α-cut of A [20].

Proposition 3.2. [20]. Let $A = \langle a,b,c,d \rangle_n$. Then

$$M_e(A) \subset M(A) \subset E(A) \quad \text{if}, \quad 0 < n \le 1.72 \quad \text{and}$$
$$M(A) \subset M_e(A) \subset E(A) \quad \text{if}. \quad n \ge 1.73$$

Corollary 3.3. [20]. Let A be a trapezoidal fuzzy number. Then $M_e(A) \subset M(A) \subset E(A)$.

Example 3.1. Let $A = \langle 0,10,11,12 \rangle$ be a trapezoidal fuzzy number. Then, $E(A) = [5,11.5] = A_\alpha$ where $\alpha = 0.5$. Similarly $M(A) = [6.66,11.33] = A_\alpha$ where $\alpha = 0.6$ and, $M_e(A) = [7.07,11.29] = A_\alpha$ where $\alpha = 0.707$. Therefore $M(A) \subset M_e(A) \subset E(A)$.

Definition 3.2. [20]. Let be a fuzzy number characterized by 2.1, and $E(A)$, $M(A)$ and $M_e(A)$ be the interval-valued probabilistic mean, interval-valued possibilistic mean and the interval-valued median of A, respectively. Then the interval $C(A) = E(A) \cap M(A) \cap M_e(A)$ is called the central interval of A. The central interval $C(A) = [C_*(A) \ C^*(A)]$ has the lower bound
$$C^*(A) = \max\{E^*(A) \ M^*(A) \ m_R(A)\},$$
and the upper bound
$$C^*(A) = \max\{E^*(A) \ M^*(A) \ m_R(A)\} .$$
Also, $core(A) \subset C(A) \subset Supp(A)$. If $E(A)$, $M(A)$ and $M_e(A)$ are associated with α-cuts of A A, then $C(A)$ is equal to one of them. If A is a fuzzy number of LR type with strictly monotonous and continuous shape functions, then $C(A) = M_e(A) \cap M(A)$. From Proposition 3.2, it follows that if $A = \langle a,b,c,d \rangle_n$ then for $0 < n \le 1.72$, $C(A) = M_e(A)$ and for $n \ge 1.73$, there is $C(A) = M(A)$.

APPLICATIONS

In this Section, the researchers introduce some of applications of the weighted interval approximation of fuzzy numbers. This indices can be applied for comparison of fuzzy numbers namely fuzzy correlations in fuzzy environments and expert's systems.

Correlation Between Fuzzy Numbers

In many applications the correlation between fuzzy numbers is of interest. Several authors have proposed different measures

of correlation between membership functions, intuitionistic fuzzy sets and correlation [4,21]. Hung and Wu [22] defined a correlation by means of expected interval. They defined the correlation coefficient between fuzzy numbers A and B as follows:

$$\rho(A,B) = \frac{E_*(A)E_*(B) + E^*(A)E^*(B)}{\sqrt{E_*^2(A) + E^{*2}(A)}\sqrt{E_*^2(B) + E^{*2}(B)}} \quad (4.7)$$

where

$$\left[E_*(A)\ E^*(B)\right] = \left[\int_0^1 a_\alpha d\alpha, \int_0^1 b_\alpha d\alpha\right] \quad (4.8)$$

This correlation coefficient shows not only the degree of relationship between the fuzzy numbers but also whether these fuzzy numbers are positively or negatively related. The researchers extend (4.7) by interval $C(A) = \left[C_*(A)\ C^*(A)\right]$ instead of (4.8), then,

$$\rho_C(A,B) = \frac{C_*(A)C_*(B) + C^*(A)C^*(B)}{\sqrt{C_*^2(A) + C^{*2}(A)}\sqrt{C_*^2(B) + C^{*2}(B)}} \quad (4.9)$$

$\rho_C(A,B)$ is called the weighted correlation coefficient between two fuzzy numbers A and B. This correlation coefficient lies in $\left[-1,1\right]$ and gives us more information compared to correlation coefficient in [21,23] and some others, which lie within $\left[0,1\right]$. It has all mentioned properties from correlation coefficient that introduced in [23], the researchers review properties $\rho_C(A,B)$ as follows:

Proposition 4.1. For any fuzzy numbers A and B we have:

1. $\rho_C(A,B) = \rho_C(B.A)$,

2. If $A = B$ then $\rho_C(A,B) = 1$,

3. If $\qquad A = B$ ` $\qquad$ for $\qquad$ some $c > 0 \Rightarrow \rho(A,B) = 1$,

4. $\left|\rho_C(A,B)\right| \leq 1$.

5.

Proof: The proof is obvious.

Example 4.1. Let $A_\alpha = \left[a_1 + (a_2 - a_1)\alpha, a_4 - (a_4 - a_3)\alpha\right]$ and $B_\alpha = \left[b_1 + (b_2 - b_1)\alpha, b_4 - (b_4 - b_3)\alpha\right]$ are two trapezoidal fuzzy numbers and $f(\alpha) = \frac{(k+1)}{2}\alpha^k, k = 1,2,\cdots$, is the weighted function. Then

$$\rho_C(A,B) = \frac{((k+1)a_2 + a_1)((k+1)b_2 + b_1) + ((k+1)a_3 + a_4)((k+1)b_3 + b_4)}{\sqrt{((k+1)a_2 + a_1)^2 + ((k+1)a_3 + a_4)^2}\sqrt{((k+1)b_2 + b_1)^2 + ((k+1)b_3 + b_4)^2}}$$

The above example shows that the weighted function interacts on the correlation coefficient between two fuzzy numbers such that for large values of k this article has:

$$\lim_{k \to \infty} \rho_C(A,B) = \frac{a_2 b_2 + a_3 b_3}{\sqrt{a_2^2 + a_3^2}\sqrt{b_2^2 + b_3^2}}.$$

Moreover if A and B be two triangular fuzzy numbers, whenever $k \to \infty$, we have

$$\rho_C(A,B) = \frac{a_2 b_2}{|a_2 b_2|},\ a_2, b_2 \neq 0.$$

Example 4.2. Yen [24] used to six grade levels (6th-grade to 11th-grade) as student's mathematical learning progress. He assigned the linguistic values $G_6, G_7, G_8, G_9, G_{10}$ and G_{11} to these grade levels, respectively, and transferred these linguistic values to corresponding reasonable normal fuzzy numbers $G_6, G_7, G_8, G_9, G_{10}$ and G_{11} with triangular membership functions as follows:(see Fig. 1.)

$$\mu_{G_6} = \begin{cases} 1 - 5x, & 0 \leq x \leq 0.2, \\ 0, & 0.2 \leq x \leq 1. \end{cases}$$

$$\mu_{G_7} = \begin{cases} 5x, & 0 \leq x \leq 0.2, \\ 2 - 5x, & 0.2 \leq x \leq 0.4, \\ 0, & 0.4 \leq x \leq 1. \end{cases}$$

$$\mu_{G_8} = \begin{cases} 0, & 0 \leq x \leq 0.2, \\ 5x - 1, & 0.2 \leq x \leq 0.4, \\ 3 - 5x, & 0.4 \leq x \leq 0.6, \\ 0, & 0.6 \leq x \leq 1. \end{cases}$$

$$\mu_{G_9} = \begin{cases} 0, & 0 \leq x \leq 0.4, \\ 5x - 2, & 0.4 \leq x \leq 0.6, \\ 4 - 5x, & 0.6 \leq x \leq 0.8, \\ 0, & 0.8 \leq x \leq 1. \end{cases}$$

$$\mu_{G_{10}} = \begin{cases} 0, & 0 \leq x \leq 0.6, \\ 5x - 3, & 0.6 \leq x \leq 0.8, \\ 5 - 5x, & 0.8 \leq x \leq 1. \end{cases}$$

$$\mu_{G_{11}} = \begin{cases} 0, & 0 \leq x \leq 0.8, \\ 5x - 4, & 0.8 \leq x \leq 1. \end{cases}$$

The researchers use the weighted correlation coefficient to compute $\rho_p(G_6, G_i)$ $i = 7,8,9,10,11$. The result of comparison is summarized in Table 1. Throughout this section the researchers assumed that $f(\alpha) = \frac{(k+1)}{2}\alpha^k$. The above example shows that $\rho_C(G_6, G_7)$ is decreasing in

Table 1. Comparative results of Example 4.2

Method	$\rho_C(G_6,G_7)$	$\rho_C(G_6,G_8)$	$\rho_C(G_6,G_9)$	$\rho_C(G_6,G_0)$	$\rho_C(G_6,G_1)$
Hong and Wang	0.852	0.778	0.778	0.778	0.857
Murty et. al	0.667	0.500	0.500	0.500	0.500
Sahnoun et. al	0.224	0.000	0.000	0.000	0.000
Proposed method with k=1	0.949	0.857	0.814	0.789	0.743
Proposed method with k=2	0.894	0.814	0.781	0.763	0.731
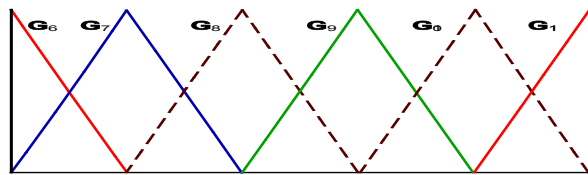
Fig.1.

i intuitively. But the methods of [25,26,27] have not decreasing behavior. For our method weighted function f interact on correlation value between two fuzzy numbers. Hence, there is reasonable advantage in using the proposed formula (4.9).

CONCLUSION

In this study, the researchers introduce the nearest weighted interval approximation of a fuzzy number. A new type of correlation coefficient between fuzzy numbers is proposed and it is compared with some other methods. These indices can be used in many times series and regression models in a fuzzy environment.

REFERENCES

[1] Dubois D, Kerre E, Mesiar R. 2000. Fuzzy interval analysis, in: Dubois D., Prade H. (Eds.). Fundamentals of Fuzzy Sets, Handbook Series of Fuzzy Sets. 483-582.

[2] Dubois D, Prade H. 1987. The mean value of a fuzzy number. Fuzzy Sets and Systems. 24:279-300.

[3] Carlsson C, Fuller R. 2002. Fuzzy reasoning in decision making and optimization. physica Verlag, heidelberg.

[4] Carlsson C, Full´er F, Majlender P. 2005. On possibilistic correlation. Fuzzy Sets and Systems. 155:425-445.

[5] Ralescu DA. 2002. Average level of a fuzzy set, in: C. Bertoluzza, M.A. Gil, D.A. Ralescu (Eds.). Statistical Modeling, Analysis and Management of Fuzzy Data, Physica-Verlag, Heidelberg. 119-126.

[6] Delgado M, Vila MA, Voxman W. 1998. On canonical representation of fuzzy numbers. Fuzzy Sets and Systems. 93:125-135.

[7] Heilpern S. 1992. The expected value of a fuzzy number. Fuzzy Sets and Systems. 47: 81-86.

[8] Kauffman A, Gupta M. M. 1991. Introduction to Fuzzy Arithmetic: Theory and Application. Van Nostrand Reinhold, New York.

[9] Zimmermann H. 1991. Fuzzy sets theory and its applications. Kluwer Academic Press, Dordrecht.

[10] Zadeh LA. 1972. A fuzzy set theoretic interpretation of linguistic hedgs. Journal of Cybernetics. 2:4-34.

[11] Zeng W, Li H. 2007. Weighted triangular approximation of fuzzy numbers. International Journal of Approximate Reasoning. 46:137-150.

[12] Saneifard R, Allahviranloo T, Hosseinzadeh F, Mikaeilvand N. 2007. Euclidean ranking DMUs with fuzzy data in DEA. Applied Mathematical Sciences. 60:2989-2998.

[13] Saneifard R, Ezatti R. 2010. Defuzzification Through a Bi Symmetrical weighted Function. Aust. J. Basic Appl. Sci. 10: 4976-4984.

[14] R. Saneifard, A method for defuzzification by weighted distance, Int. J. Industrial Mathematics. 3 (2009) 209 - 217.

[15] Saneifard R. 2009. Ranking L-R fuzzy numbers with weighted averaging based on levels. Int. J. Industrial Mathematics. 2:163-173.

[16] Saneifard Rahim, Saneifard Rasoul. 2011. A modified method for defuzzification by probability density function. Journal of Applied Sciences Research. 2:102-110.

[17] Saneifard R, Asgari A. 2011. A Method For Defuzzification Based on Probability Density Function (II). Applied Mathematical Sciences. 5:1357-1365.

[18] Saneifard Rahim, Saneifard Rasoul. 2011. Evaluation of Fuzzy Linear Regression Models by Parametric Distance. Australian Journal of Basic and Applied Sciences. 5:261-267.

[19] Saneifard R, Noori E. 2011. A novel approach for defuzzification based on ambiguity-preserving. International Journal of Natural and Engineering Sciences. 5 (3): 27-32.

[20] Bodjanova S. Median value and median interval of a fuzzy number. Information Sciences. 172: 73-89.

[21] Bustince H, Burillo P. 1995. Correlation of interval-valued intuitionistic fuzzy sets. Fuzzy Sets and Systems. 74:237-244.

[22] Hung W, Wu J. 2002. Correlation of intuitionistic fuzzy sets by centroid method. Information Sciences. 144:219-225.

[23] Yu C. 1993. Correlation of fuzzy numbers. Fuzzy Sets and Systems. 55:303-307.

[24] Yen CL. 2006. Using fuzzy sets in developing mathematical learning-progress indicator. Proc. Natl. Sci. Coune. ROC.

[25] Hong DH, Hwang SY. 1995. Correlation of intuitionistic fuzzy sets in probability spaces. Fuzzy Sets and Systems. 75:77-81.

[26] Murty C.A, Pal SK, Majumder DD. 1985. Correlation between two fuzzy membership functions. Fuzzy Sets and Systems. 17:23-38.

[27] Sahnoun Z, Dicesare F. 1991. Efficient methods for computing linguistic consistency. Fuzzy Sets and Systems. 39:15-26.