A New Concept for Deriving the Expected Value of Fuzzy Random Variables

Liang-Hsuan Chen, Chia-Jung Chang

Abstract—Fuzzy random variables have been introduced as an imprecise concept of numeric values for characterizing the imprecise knowledge. The descriptive parameters can be used to describe the primary features of a set of fuzzy random observations. In fuzzy environments, the expected values are usually represented as fuzzy-valued, interval-valued or numeric-valued descriptive parameters using various metrics. Instead of the concept of area metric that is usually adopted in the relevant studies, the numeric expected value is proposed by the concept of distance metric in this study based on two characters (fuzziness and randomness) of FRVs. Comparing with the existing measures, although the results show that the proposed numeric expected value is same with those using the different metric, if only triangular membership functions are used. However, the proposed approach has the advantages of intuitiveness and computational efficiency, when the membership functions are not triangular types. An example with three datasets is provided for verifying the proposed approach.

Keywords—Fuzzy random variables, Distance measure, Expected value.

I. INTRODUCTION

IN practical problems, we often faced a situation in which outcomes of random experiment are not suitable to be described as numeric values, but allowed to be represented as fuzzy values for characterizing the imprecise knowledge. Fuzzy random variables (FRVs) have been introduced as an imprecise concept of numeric values. The concept of FRVs, defined as the extension of classical random variable (CRV), was introduced by Kwakernaak [1], [2]. The various definitions about FRVs have been proposed in the literature. Puri & Ralescu [3] and Diamond & Kloeden [4] viewed an FRV as the extension of a random set. That is, FRVs are collected from a probability space, but are expressed in linguistic terms. Kwakernaak [1] and Kruse & Meyer [5] adopted FRVs to model the imprecise perception of a CRV. Almost all studies followed such definitions to construct various descriptive parameters. Furthermore, Couso & Dubois [6] proposed another view of FRVs, defining the most likely value (the largest membership grade) as the state of a standard randomness and the spread of a fuzzy set as the state of a fuzziness. Thus, FRVs are used to deal with two uncertainties, randomness and fuzziness.

In general, the descriptive parameters are useful to

L. H. Chen is with the Department of Industrial and Information Management, National Cheng Kung University, Tainan 701, Taiwan (e-mail: lhchen@mail.ncku.edu.tw).

C. J. Chang is with the Department of Industrial and Information Management, National Cheng Kung University, Tainan 701, Taiwan (e-mail: r38991028@mail.ncku.edu.tw).

summary the main features of a set of collected data. In classical descriptive statistics, the expected value is well-defined as a measure of center. In fuzzy environments, the expected values are usually developed by the various metrics. To develop the expected value of FRVs, the area metric to measure the difference between two fuzzy numbers is adopted by almost all the existing studies. Three kinds of approaches were used: (1) Hausdorff metric [3] and [7], (2) L_2-metric on the space of Lebesgue integrable [8] and [9], and (3) possibility theory [6] and [10]-[12]. Those approaches addressed the data types of expected value as: (1) fuzzy set [10]-[12], (2) numeric value [3], [5], [8], (3) interval value [9]-[13]. Although various approaches were proposed, they suffer from the cumbersome computations. Instead of the concept of area metric that is usually adopted in the relevant studies, the numeric expected value is proposed by the concept of distance metric in this study based on the two characters (fuzziness and randomness) of FRVs. The proposed approach is based on the concept of distance in the study of Chen & Hsueh [6], [7] to develop the numeric expected value, in which the fuzzy distance is calculated by several α-cuts between a pair of fuzzy numbers. The proposed approach has the advantages of intuitiveness and computational efficiency.

In the following section, we introduce the expected value of FRVs. In Section III, an example with three datasets is used to verify the proposed approach. Finally, some conclusions are provided in Section IV.

II. EXPECTED VALUE

To develop the numeric expected value of FRVs, firstly consider a metric D defined over the class of fuzzy subsets of R, $F(\text{R})$ (i.e., the set of compact fuzzy subsets of the real line). The expected value of an FRV, $\tilde{X} : \Omega \rightarrow F(\text{R})$, with respect to the probability space $(\Omega, \mathcal{A}, \mathcal{P})$ is formulated as $E(\tilde{X}) = \int_{\Omega} D(\tilde{X}, \{0\}) \, d\mathcal{P}$, where $\{0\}$ is a singleton (i.e., $\mu_{\{0\}}(x) = 1$, if $x=0$; otherwise, 0).

Definition 1. Let $\tilde{X}$ be a triangular FRV, as shown in Fig. 1. A data set consists of n observations denoted as a triple element set $\tilde{X}_i = (a_i, b_i, c_i)$, $i=1,\ldots,n$. Such that, $m = b$, $l = b-a$ and $r = c-b$, represent the most likely value, left spread and right spread, respectively, signifying the most likely values (m) as the state of a standard randomness and the range of spread (l and/or r) as the state of a fuzziness. The α-level sets of $\tilde{X}$, $\tilde{X}_\alpha = \{\omega \in \Omega : \mu_{\tilde{X}}(\omega) \geq \alpha\}$, are closed for all

$\alpha \in (0,1]$.

Firstly, the expected value proposed by Puri & Ralescu [14], [15] is introduced for the subsequent comparisons. They defined the expected value $E(\tilde{X})^{PR}$ as a fuzzy set for $\tilde{X} \in F(R)$, such that $\left(E(\tilde{X})^{PR}\right)_\alpha = \int_\Omega X_\alpha \, d\mathcal{P}$ for $\alpha \in (0,1]$. Hence, the expected value $E(\tilde{X})^{PR}$ is formulated as

$$E(\tilde{X})^{PR} \equiv \left[\frac{1}{n}\sum_{i=1}^n \tilde{X}_i^{\,L}, \ \frac{1}{n}\sum_{i=1}^n \tilde{X}_i^{\,M}, \ \frac{1}{n}\sum_{i=1}^n \tilde{X}_i^{\,U} \right] \qquad (1)$$

where $\tilde{X}^L$, $\tilde{X}^M$ and $\tilde{X}^U$ are the lower bound, the most likely value and the upper bound of the FRV $\tilde{X}$, respectively. In the followings, we will review some approaches in terms of the concept of area metric, and then introduce the proposed expected value.
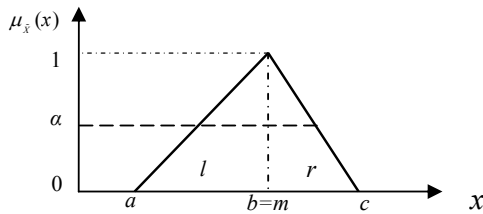

Fig. 1 A generalized triangular FRV

A. The Concept of Area Metric

Gil & Lrpez-Diaz [3] employed the λ-average function (2), suggested by Campos & González [13], to develop the expected value. The λ-average function will produce a numeric value as

$$V_S^\lambda(\tilde{X}) = \int_{(0,1]} \left[\lambda x_{\alpha 2} + (1-\lambda) x_{\alpha 1} \right] dS(\alpha) \qquad (2)$$

where the function V_S^λ is carried out for each $\tilde{X} \in F(R)$ using the α-cuts, $\tilde{X}_\alpha = [x_{\alpha 1}, x_{\alpha 2}]$, for each $\alpha \in (0,1]$. The parameter value of $\lambda \in (0,1]$ is determined as a subjective degree of optimism-pessimism, and S is an additive measure on $(0,1]$ to determine the weight associated with different α-cuts. For the comparison purpose, without loss of generality, we set $\lambda = 1/2$ in Gil &Lrpez-Diaz's approach, and replace $x_{\alpha 1}$ and $x_{\alpha 2}$ by $c_i - \alpha(c_i - b_i)$ and $a_i + \alpha(b_i - a_i)$ in (2), respectively, based on triangular membership functions. Then $V_S^\lambda(\tilde{X}_i) = (1/4)[a_i + c_i + 2b_i]$ can be obtained, and the numeric expected value of FRVs is determined as $E(\tilde{X})^{GL} = E(V_S^\lambda(\tilde{X})) \overset{\lambda=1/2}{=} (1/4)(\overline{a} + \overline{c} + 2\overline{b})$, where $\overline{a}$, $\overline{b}$, and $\overline{c}$ are the corresponding averages.

Following Fréchet's principle, Körner [16] applied the L_2-metric on the space of Lebesgue integrable as (3) to obtain the definition of the expected value of FRVs as

$$E_F(X) = \left\{ B \in F(R) : Ed_2^2(X,B) = \inf_{C \in F(R)} Ed_2^2(X,C) \right\}.$$

$$d_2(X,B) = \|s_X - s_B\|_2$$
$$= \left(n \cdot \int_0^1 \int_{S^{n-1}} \left| s_X(\alpha,u) - s_B(\alpha,u) \right| \mu(du) \, d\alpha \right)^{1/2} (3)$$

where the support function $s_X(\alpha,u) = \sup\{\langle u,a \rangle : a \in X^\alpha\}$, $u \in S^{n-1}$, $\alpha \in [0,1]$, S^{n-1} is the (n-1)-dimensional unit sphere of R, and $\langle \cdot, \cdot \rangle$ is the inner product of the Euclidean space. By using the Steiner point of a fuzzy set, defined as $\sigma_X = n \cdot \int_0^1 \int_{S^{n-1}} u \cdot s_X(u)\mu(du) \, d\alpha$, the approach obtains the numeric expected value of an triangular LR-fuzzy number as

$$E(X_{LR})^{Kr} = Em + \sigma_{X_R} \cdot Er - \sigma_{X_L} \cdot El \qquad (4)$$

where X_L and X_R are the fuzzy set with the membership function $L(x)$ and $R(x)$, respectively (i.e., $L(x)=R(x)=1-x$). And, (4) can be revised as

$$E(X_{LR})^{Kr} = Em + \left(\frac{1}{2}\int_0^1 R^{(-1)}(\alpha)d\alpha \right) \cdot Er - \left(\frac{1}{2}\int_0^1 L^{(-1)}(\alpha)d\alpha \right) \cdot El \,(5)$$

where σ_{X_L} and σ_{X_R} are obtained by $\sigma_{X_L} = \frac{1}{2}\int_0^1 L^{(-1)}(\alpha)d\alpha = \frac{1}{4}$ and $\sigma_{X_R} = \frac{1}{2}\int_0^1 R^{(-1)}(\alpha)d\alpha = \frac{1}{4}$, respectively. Therefore, the numeric expected value of FRVs in Körner's approach is $E(X_{LR})^{Kr} = Em + \frac{1}{4} \cdot Er - \frac{1}{4} \cdot El$, which can then be derived as $\frac{1}{4}(\overline{a} + \overline{c} + 2\overline{b})$ based on triangular membership functions.

Based on possibility theory, a numeric expected value is developed by Liu & Liu [9]. For any closed subset $F(R)$ of R, an FRV $\tilde{X}$ connected with a probability space, $X(\omega) = Pos\{X(\omega) \in F(R)\} = \sup_{x \in R} \mu_{X(\omega)}(x)$ is a measurable function of $\omega \in \Omega$, where $\mu_{X(\omega)}$ is the possibility distribution function of FRV $X(\omega)$. To obtain the well-known uncertain functions, possibility measure (6), necessity measure (7) and the credibility measure (8) of event $\{X \le r\}$, the numeric expected value of $\tilde{X}$ is defined as (9).

$$Pos\{X \le r\} = \sup_{t \le r} \mu_X(t) \qquad (6)$$

$$Nes\left\{X \le r\right\} = 1 - \sup_{t>r} \mu_X(t) \qquad (7)$$

$$Cr\{X \le r\} = \frac{1}{2}\left(Pos(X \le r) + Nes(X \le r)\right) \qquad (8)$$

$$E(\tilde{X})^{LL} = \int_0^\infty Cr\{X \ge r\}\, dr - \int_{-\infty}^0 Cr\{X \le r\}\, dr \qquad (9)$$

where $Cr\{X\} = 1 - Cr\{X^c\}$ is a self-dual set function. Using (9), the expected value of FRVs is $(1/4)\left(a_i + c_i + 2b_i\right)$ based on triangular membership functions, then the numeric expected value of FRVs in Liu & Liu's approach is determined as $E(\tilde{X})^{LL} = (1/4)(\overline{a} + \overline{c} + 2\overline{b})$ [10]-[12].

B. The Proposed Concept of Distance Metric

Different from the use of the concept of area metric, the proposed expected value is developed by employing the concept of α-cut between a pair of fuzzy numbers. For the ith fuzzy observation $\tilde{X}_i$, the distance of a fuzzy random observation $\tilde{X}_i$ from the single point $\{0\}$ at the kth α-level is defined as $(1/2)\left[\left((\tilde{X}_i)^L_{\alpha_k} - (\{0\})^L_{\alpha_k}\right) + \left((\tilde{X}_i)^U_{\alpha_k} - (\{0\})^U_{\alpha_k}\right)\right]$, which are illustrated in Fig. 2 as the average distance between two fuzzy numbers at the kth α-level (let $\alpha_k = k/h - 1, k = 0$, 1, …, h-1).Hence, the mean distance between the ith fuzzy random observation and $\{0\}$is determined as the average from $h\alpha$-cuts as

$$D(\tilde{X}_i, \{0\}) \equiv \frac{1}{2h}\sum_{k=0}^{h-1}\left[\left((\tilde{X}_i)^L_{\alpha_k} - (\{0\})^L_{\alpha_k}\right) + \left((\tilde{X}_i)^U_{\alpha_k} - (\{0\})^U_{\alpha_k}\right)\right] (10)$$
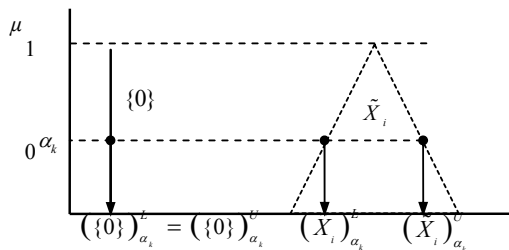

Fig. 2 The distance of a fuzzy random observation $\tilde{X}_i$ from the single point $\{0\}$

The expected value can be formulated as the average of the distance measures $D(\tilde{X}_i, \{0\})$, $i = 1$, …, n, from n fuzzy random observations:

$$E(\tilde{X})^{CC} \equiv \frac{1}{n}\sum_{i=1}^n\left\{\frac{1}{2h}\sum_{k=0}^{h-1}\left[\begin{array}{l}\left((\tilde{X}_i)^L_{\alpha_k} - (\{0\})^L_{\alpha_k}\right) + \\ \left((\tilde{X}_i)^U_{\alpha_k} - (\{0\})^U_{\alpha_k}\right)\end{array}\right]\right\} (11)$$

Thus, we can have

$$E(\tilde{X})^{CC} = \frac{1}{n}\cdot\sum_{i=1}^n\frac{1}{2h}\sum_{k=0}^{h-1}\left[(\tilde{X}_i)^L_{\alpha_k} + (\tilde{X}_i)^U_{\alpha_k}\right]$$

$$= \frac{1}{4}(\overline{a} + \overline{c} + 2\overline{b}) = \overline{m} - \frac{1}{4}\cdot\overline{l} + \frac{1}{4}\cdot\overline{r}, \ \forall\, h \ge 2$$

based on triangular membership functions. It is noted that the proposed expected value contains two parts, i.e., $\overline{m}$ and $-(1/4)\cdot\overline{l} + (1/4)\cdot\overline{r}$, in which $\overline{m}$ is contributed from the randomness, while $-(1/4)\cdot\overline{l} + (1/4)\cdot\overline{r}$ can be attributed from the fuzziness. In addition, obviously the proposed expected value $E(\tilde{X})^{CC}$ is same as those from the other approaches, such as Gil & Lrpez-Diaz [12], [17], Körner [13], and Liu & Liu [9], i.e., $E(\tilde{X})^{GL} = E(X_{LR})^{Kr} = E(\tilde{X})^{LL} = E(\tilde{X})^{CC}$, when triangular membership functions are adopted. However, the proposed approach is intuitive in the formulation, and is efficient in computations, when the membership functions are not triangular type. With the non-triangular membership functions, the existing approaches should carry out the more complicated computations by the area metric, comparing with the proposed approach.

III. ILLUSTRATIVE EXAMPLE

Three datasets are used in this section to exemplify the various approaches mentioned before to determine the expected values of FRVs. The first dataset contains the fuzzy random observations with symmetric triangular membership functions, while the asymmetrical fuzzy random observations are considered in the other two datasets.

Three datasets, as listed in Table I, were analyzed by Sakawa & Yano [10]-[12], Wu [18] and Chen & Hsueh [19], respectively. After performing the approaches in Puri & Ralescu [15], Gil & Lrpez-Diaz [3], Körner [13], Liu & Liu [9] and the proposed $E(\tilde{X})^{CC}$, the obtained expected values are showed in Table II. As described in Section II, all the approaches, excluding Puri & Ralescu [10], [11], results in the same expected value. However, the proposed approach is more efficiently in computational process.

TABLE I
THREE DATASETS OF FUZZY RANDOM OBSERVATIONS

No.	Dataset-1	Dataset-2	Dataset-3
1	(3.5, 4.0, 4.5)	(111, 162, 194)	(19, 30, 39)
2	(5.0, 5.5, 6.0)	(88, 120, 160)	(7, 20, 30)
3	(6.5, 7.5, 8.5)	(161, 222, 288)	(14, 25, 37)
4	(6.0, 6.5, 7.0)	(83, 131, 194)	(33, 45, 55)
5	(8.0, 8.5, 9.0)	(51, 67, 83)	(26, 38, 46)
6	(7.0, 8.0, 9.0)	(124, 169, 213)	(32, 43, 52)
7	(10.0, 10.5, 11.0)	(62, 81, 102)	(23, 40, 51)
8	(9.0, 9.5, 10.0)	(138, 192, 241)	(27, 38, 50)
9		(82, 116, 159)	(25, 37, 49)
10		(41, 55, 71)	(49, 60, 72)
11		(168, 252, 367)	(49, 59, 68)
12		(178, 232, 346)	(43, 54, 62)
13		(111, 144, 198)	(47, 61, 64)
14		(78, 103, 148)	(24, 34, 42)
15		(167, 212, 267)	(29, 38, 47)
16			(48, 64, 73)
17			(43, 56, 63)
18			(52, 63, 72)
19			(50, 66, 71)
20			(37, 49, 58)
21			(45, 55, 67)
22			(56, 67, 81)
23			(43, 53, 62)
24			(45, 54, 64)
25			(57, 70, 77)
26			(59, 68, 78)
27			(55, 65, 74)
28			(70, 75, 89)
29			(74, 84, 91)
30			(68, 80, 86)

TABLE II
EXPECTED VALUES OBTAINED BY DIFFERENT APPROACHES

Approach	Dataset-1	Dataset-2	Dataset-3
$E(\tilde{X})^{PR}$	(6.88,7.50,8.13)	(109.53,150.53,202.07)	(41.63,53.03,62.33)
$E(\tilde{X})^{GL}$	7.50	153.17	52.51
$E(X_{LR})^{Kr}$	7.50	153.17	52.51
$E(\tilde{X})^{LL}$	7.50	153.17	52.51
$E(\tilde{X})^{CC}$	7.50	153.17	52.51

IV. CONCLUSION

Considering the nature of fuzziness and randomness in FRVs, this study proposed the numeric expected value of FRVs based on the distance metric. The two parts of the expected value, $\overline{m}$ and $-(1/4) \cdot \overline{l} + (1/4) \cdot \overline{r}$, can be considered as the measures for randomness and fuzziness, respectively. Through the derivative processes in this study, the expressions of the expected value, presented by Gil & Lrpez-Diaz's approach, Körner's approach, Liu & Liu's approach and our proposed approach, have the same value, although different concepts of metric are adopted, when triangular membership functions are adopted. However, the development of the proposed approach is intuitive, and has the advantage of computational efficiency, if the membership functions are not triangular type.

REFERENCES

[1] H. Kwakernaak, "Fuzzy Random Variabhs-I. Definitions and Theorems," *Information Sciences*, vol. 15, pp. 1-29, 1978.
[2] H. Kwakernaak, "Fuzzy Random Variables-II. Algorithms and Examples for the Discrete Case," *Information Sciences*, vol. 17, 1979.
[3] M. L. Puri and D. A. Ralescu, "Fuzzy random variables," *Journal of Mathematical Analysis and Applications*, vol. 114, pp. 409-422, 1986.
[4] P. Diamond and P. Kloeden, *Metric Spaces of Fuzzy Sets: Theory and Applications. Singapore*: World Scientific, 1994.
[5] R. Kruse and K. D. Meyer, *Statistics with Vague Data*: Reidel, Dordrecht, The Netherlands, 1987.
[6] I. Couso and D. Dubois, "On the Variability of the Concept of Variance for Fuzzy Random Variables," *IEEE Transactions on Fuzzy Systems*, vol. 17, pp. 1070-1080, 2009.
[7] R. Kruse, "On the Variance of Random Sets," *Journal of Mathematical Analysis and Applications*, vol. 122, pp. 469-473, 1987.
[8] Y. Feng, L. Hu, and H. Shu, "The variance and covariance of fuzzy random variables and their application," *Fuzzy Sets and Systems*, vol. 120, pp. 487-497, 2001.
[9] R. Körner, "On the variance of fuzzy random variables," *Fuzzy Sets and Systems*, vol. 92, pp. 83-93, 1997.
[10] B. Liu and Y. K. Liu, "Expected value of fuzzy variable and fuzzy expected value models," *IEEE Transactions on Fuzzy Systems*, vol. 10, pp. 445-450, 2002.
[11] Liu and Y. K. Liu, "A class of fuzzy random optimization: expected value models," *Information Sciences*, vol. 155, pp. 89-102, 2003.
[12] B. Liu and Y. K. Liu, "Fuzzy random variables: a scalar expected value operator," *Fuzzy Optimization and Decision Making*, vol. 2, pp. 143-160, 2003.
[13] M. A. Gil and M. Lrpez-Diaz, "The λ-average value and the fuzzy expectation of a fuzzy random variable," *Fuzzy Sets and Systems*, vol. 99, pp. 347-352, 1998.
[14] L. H. Chen and C. C. Hsueh, "A mathematical programming method for formulating a fuzzy regression model based on distance criterion," *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, vol. 37, pp. 705-712, Jun 2007.
[15] L. H. Chen and C. C. Hsueh, "Fuzzy Regression Models Using the Least-Squares Method Based on the Concept of Distance," *IEEE Transactions on Fuzzy Systems*, vol. 17, pp. 1259-1272, Dec 2009.
[16] L. M. d. Campos and A. González, "A subjective approach for ranking fuzzy numbers," *Fuzzy Sets and Systems*, vol. 29, pp. 145-153, 1989.
[17] J. Watada, S. Wang, and W. Pedrycz, "Building confidence-interval-based fuzzy random regression models," *IEEE Transcations on Fuzzy Systems*, vol. 17, pp. 1273-1283, 2009.
[18] M. Sakawa and H. Yano, "Multiobjective fuzzy linear regression analysis for fuzzy input-output data," *Fuzzy Sets and Systems*, vol. 47, pp. 173-181, 1992.
[19] H. C. Wu, "Linear regression analysis for fuzzy input and output data using the extension principle," *Computers and Mathematics with Applications*, vol. 45, pp. 1849-1859, 2003.

Liang-Hsuan Chen received the B.S. and M.S. degrees in industrial management from the National Cheng Kung University, Tainan, Taiwan, in 1980 and1982, respectively, and the Ph.D. degree in industrial engineering from the University of Missouri, Columbia, in 1991.

He is currently a Professor of industrial and information management with the National Cheng Kung University. His current research interests include intelligent computations, fuzzy systems, and fuzzy set theory and its applications in decision-making problems, decision analysis, and robust system design. He has authored or coauthored a number of research papers published in reputed international journals, such as the IEEE Transactions on Systems, Man, and Cybernetics—Part B, the European Journal of Operational Research, Fuzzy Sets and Systems, etc. He was the Editor-in-Chief of the Journal of the Chinese Institute of Industrial Engineers in 2005 and 2006.

Prof. Chen received the Outstanding Research Award in the field of management science from the National Science Council of Taiwan in 2000.

Chia-Jung Chang received the M.S. degree in Statistics in 2002 from the National Central University, Taoyuan County, Taiwan, and in industrial and information management in 2010 from the National Cheng Kung University, Tainan, Taiwan. She is currently working toward the Ph.D. degree with the Department of Industrial and Information Management in National Cheng Kung University. Her current research interests include: operations research, mathematical programming, fuzzy regression, fuzzy statistical process control and fuzzy set theory and its applications in traditional statistics problems, etc.