Data Analysis with Fuzzy Measure on Intuitionistic Fuzzy Sets

Sanghyuk Lee*, Ka Lok Man, Eng Gee Lim, Mark Leach

Abstract- Study of fuzzy entropy on intuitionistic fuzzy sets (IFSs) were proposed, and analyzed. Built in uncertainty in IFSs, it is named as hesitation. It is contained in fuzzy membership function in itself by definition. Hence, designing fuzzy entropy is not easy because there is no general entropy definition about IFSs. By considering existing fuzzy entropy definitions, fuzzy entropy on IFSs is designed and proved its usefulness. Similarity measure was also extended with the information of entropy, and the similarity measure is verified its usefulness by the proof.

I. INTRODUCTION

Generally, data uncertainty is calculated with the help of entropy [1]. Then, the difference between fuzzy set and intuitionistic fuzzy sets is clear, because IFSs include hesitation – uncertainty - by definition. Hence, there are difference and difficulty if we design fuzzy entropy on IFSs. In previous results, fuzzy entropy was designed as explicitly [2–4]. By the dual concepts, similarity measure was also proposed to compute the degree of similarity between fuzzy sets [5–7]. With the obtained fuzzy entropy and similarity measure, result represents the complementary information for each other. Hence, the relation was clear that the similarity measure (entropy) was derived from entropy (similarity measure) in the previous literature [6]. Previous research on fuzzy entropy has been derived by many researchers [2–4]. It was shown that entropy design on fuzzy sets is enough than the research on IFSs, it means that entropy design on fuzzy sets is rather easy and convenient. They derived rather explicit formula, hence it is easy to apply calculation. Specially, Liu proposed axiomatic definitions of entropy, distance measures, and similarity measures and discussed the relationships among these three concepts.

By the way, similarity measure design is simplified than fuzzy entropy, it can be designed easily with the help of fuzzy numbers [5]. Even it has restriction, it has only application to triangular or trapezoidal membership functions [5]. For the general formulation of similarity measure even for IFSs, the distance measures are applicable to general fuzzy membership functions, including non-convex fuzzy membership functions [8].

As an extension of fuzzy sets, IFSs and vague sets were introduced by Atanassov and Gau and Buehrer, and other authors [9–12]. Basically, IFSs include hesitancy, that is not belonging to membership and non-membership. Furthermore, comparison between IFSs and vague sets was pointed out by Bustince and Burillo, two sets are the same [13]. Data distribution with IFSs shows more realistic than fuzzy sets, and it is more practical and accurate. Hence, we do want to point out that design of fuzzy entropy and similarity measure on IFSs are very important to obtain more reliable results.

In this paper, fuzzy entropy on IFSs has been derived by the definition. Proposed fuzzy entropy was considered for the general fuzzy membership functions, it is applicable even it is non-convex structure. Unfortunately, still there is no unified fuzzy entropy definition on IFSs yet. Hence, we apply two fuzzy entropy definitions, and the fuzzy entropy was considered for both cases. By the dual concept, similarity measure was also derived through fuzzy entropy results.

II. PRELIMINARIES

IFS was proposed as general formulation of fuzzy sets with hesitancy by Atanassov [9].

A. Intuitionistic fuzzy sets

Atanassov gave the definition of IFSs, in which the uncertainty of data, hesitancy is included more practically [9].

Definition 2.1 IFSs $V$ in the universe of discourse $X = \{x_1, x_2, \cdots, x_n\}$ defined as follows:

$$V = \{(x, t_r(x), f_r(x)) \mid x \in X, t_r(x) \in [0,1], f_r(x) \in [0,1], 0 \leq t_r(x) + f_r(x) \leq 1\}.$$  

Where, $t_r(x)$ and $f_r(x)$ denote the membership function and the non-membership function of $x$ in $X$, respectively. And $t_r(x)$ is the lowest bound of membership degree of $x$, $f_r(x)$ is the lowest bound of non-membership degree of $x$, respectively. From the definition, it is clear that membership degree in IFS $V$, in between $[t_r, 1-f_r]$ is not belonging in anywhere. Uncertainty degree can be calculated by $1-t_r - f_r$ for all $x$ in $X$. Furthermore, if $t_r + f_r = 1$, then $V$ is expressed as the same as fuzzy set definition. In order to evaluate the uncertainty or entropy on IFSs, analysis on relation among the hesitation information, membership and
non-membership degree is needed to be considered.

By the meaning of fuzzy set, null set has meaning that there is no any other information about data themselves. By graphical representation, analysis between the membership degree and non-membership degree is illustrated in Fig. 1. It shows the projection with Szmidt and Kacprzyk’s, it represents the coordination of \( t_v(x) \cdot f_v(x) \) plane.

From the Fig. 1, under diagonal area is the point that hesitancy is included. Hence A-F-B, and lines under the fuzzy line satisfies \( t_v + f_v + \text{hesitancy} = 1 \) under the variation of hesitancy. By the inspection, it is clear that hesitance satisfies one when \( t_v + f_v = 0 \), that is origin.

Fig. 1 F membership function and non-membership function.

Analysis indicates that the fuzzy line means that it has no hesitance. IFSs satisfies lines under the fuzzy line, hence the relation between membership degree and non-membership degree are defined by \( 0 \leq t_v + f_v \leq 1 \). Especially the line of \( t_v + f_v = 0.5 \) is specified in dotted line. As the point approach to the origin \((0, 0)\), degree of hesitance also approaches one. Graphical information provide rather geometrical information about membership, non-membership degree and hesitancy information.

B. Uncertainty on IFSs

Typical IFSs membership function is illustrated in Fig. 2. Research on uncertainty of IFSs, fuzzy entropy on IFSs has been considered by some researchers [13, 17]. However, similarity measure design on IFSs reported more than fuzzy entropy design on IFSs. Burillo and Bustince [13] and Szmidt and Kacprzyk [17] have proposed the fuzzy entropy on IFSs [13, 17]. The knowledge provides us methodology on how to measure the degree of intuitionism on IFS, and non-probabilistic type of entropy measure with a geometric interpretation on IFSs.

Burillo and Bustince proposed an axiomatic definition of IFSs, which was considered by taking into account fuzzy set consideration. Gaussian type IFS membership function is illustrated in Fig. 2.

C. Entropy and Similarity Measure Definition on IFSs

Fuzzy entropy was proposed by De Luca and Termini at first time [4], and the axiomatic definition was referred to Shannon’s probability entropy.

Definition 2.3 [4] A real function \( e: \text{FS}(X) \rightarrow R^+ \) is called an entropy on fuzzy set FS(X) if \( e \) has the following properties:

\[
\begin{align*}
(\text{E1}) & \quad e(\tilde{A}) = 0, \text{ if } \tilde{A} \text{ is a crisp set.} \\
(\text{E2}) & \quad e(\tilde{A}) \text{ assumes a unique maximum if } \mu_{\tilde{A}} = 1/2. \\
(\text{E3}) & \quad e(\tilde{A}) \leq e(\tilde{B}) \text{ if } \tilde{A} \text{ is crisper than } \tilde{B}, \text{ that is, if } \mu_{\tilde{A}} \leq \mu_{\tilde{B}} \text{ for } \mu_{\tilde{B}} \leq 1/2 \text{ and } \mu_{\tilde{A}} \geq \mu_{\tilde{B}} \text{ for } \mu_{\tilde{B}} \geq 1/2.
\end{align*}
\]

Fig. 2 Gaussian type IFS Membership Function.
Szmidt and Kacprzyk. They proposed entropy on IFSs as follows

\[ E(F) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\max Count(F_i) \cap C_i}{\max Count(F_i)} \right) \]

Where \( F_i \cap C_i \leq \min(\mu_{F_i}, \mu_{C_i}) \), max\((\nu_{F_i}, \nu_{C_i})\) > 
and \( F_i \cup C_i \geq \max(\mu_{F_i}, \mu_{C_i}) \), min\((\nu_{F_i}, \nu_{C_i})\) > .

Here, (2) and (3) have quite different meaning, that is, the difference could be found Example 1 of Szmidt and Kacprzyk [17].

Another fuzzy entropy on IFSs has been defined by Hung and Yang [16]. In their definition, IP2 and IP4 of Definition 2.2 are different from those of their properties. Difference of two definitions has their own characteristics. However, it is not matched to the actual consideration, because of the property (E1) in Definition 2.3.

Now we introduce entropy on fuzzy set briefly with our previous result. Fig. 3 shows Gaussian fuzzy membership function and crisp set \( A_{\text{near}} \), especially “near” is 0.5.

![Fig. 3 Membership functions of fuzzy set](image)

Then, it could be sufficient to calculate the entropy representation for “AB” line of Fig. 1, that is the case of \( l_x + f_y = 1 \).

B. Entropy on IFSs
Now we propose fuzzy entropy on IFSs with the definition of Szmidt and Kacprzyk that was generalized to IFSs case [17]. Now fuzzy entropy on IFSs satisfying definition is obtained as follows.

**Theorem 3.1** Following equation satisfies a fuzzy entropy on IFSs.

\[ E_t(A, A_{\text{near}}) = \frac{1}{N} \sum_{i=1}^{N} \left( \pi_i(x_i) + d(\mu_i(x_i), \mu_{A_{\text{near}}}) \right) \]

Where subscript “near” means close crisp set to set \( A \), in this theorem near = 0.5 is considered, then \( \mu_{A_{\text{near}}} = 1 \) when \( \mu_A \geq 0.5 \) and \( \mu_{A_{\text{near}}} = 0 \) when \( \mu_A \leq 0.5 \).

**Proof:** (4) stands for the sum of conventional fuzzy entropy
concept and hesitancy area on IFSs. For the conventional area in (4), $d(\mu_A(x), \mu_B(x))$ represent easy of proof. Hence, it is the same result of [16]. For the non-specific distribution, (4) is clear because $A_{\text{near}}$ satisfies $A$ itself. Therefore, (P1) is satisfied. For (IP2), $d(\mu_A(x), \mu_{A_{\text{near}}}) = 0$ is proved easily for the case of $\mu_A(x) = 0$ and $v_A(x) = 0$. And it is also quite natural, $I(A) = \text{Cardinal}(X) = N$, so (IP3) is clear from the definition. Finally, $A \preceq B$ guarantee two properties
\[
\frac{1}{N} \sum_{i=1}^{N} \pi_A(x_i) \geq \frac{1}{N} \sum_{i=1}^{N} \pi_A(x_i) \\
\text{and} \quad d(\mu_A(x), \mu_{A_{\text{near}}}) \geq d(\mu_A(x), \mu_{A_{\text{near}}}).
\]

Hence, Theorem 3.1 considers for more flexible fuzzy membership function including Gaussian fuzzy membership function. Hence, (4) is useful as a fuzzy entropy on IFSs. Fuzzy entropy has to be carried out between IFS and corresponding numeric data or set. Whereas, the similarity measure can be designed through comparison between IFSs.

C. Similarity Measure Design on IFSs with Distance Measure

It is well known fact that the similarity measure has counter-intuitive case of fuzzy entropy measure [15]. However, condition P2 is very strict to overcome, because similarity measure based on the difference calculation between two IFSs. For example, Hong and Kim showed similarity measure as follows;

\[
S_H(A, B) = 1 - \frac{\sum_{i=1}^{n} (|t_A(x_i) - t_B(x_i)| + |f_A(x_i) - f_B(x_i)|)}{2n}
\]

Equation has counter-intuitive case for P2, and the result can be shown in the related references. Other conventional similarity measure showed almost same results. Here novel similarity measure between IFSs is considered as follows.

Theorem 3.2 Following equation satisfies a similarity measure on IFS(X).

\[
S_L(A, B) = 1 - E_L(A, B)
\]

Where $E_L(A, B)$ is the same formulation of (1), and $A_{\text{near}}$ is replaced into $B$.

Proof: From (P1) to (P3), it is clear from (4) itself. And $d(\mu_A(x_i), \mu_{A_{\text{near}}}(x_i))$ and $d(\mu_A(x_i), \mu_{B_{\text{near}}}(x_i))$ are greater than $d(\mu_B(x_i), \mu_{A_{\text{near}}}(x_i))$ and $d(\mu_B(x_i), \mu_{B_{\text{near}}}(x_i))$, respectively. Hence, (P4) is satisfied. Finally, $A = \Phi$ and $B = \overline{A}$, or, $A = \overline{B}$ and $B = \Phi$, IFSs $A$ and $B$ are complementary each other., therefore (P5) is also satisfied.

IV. DISCUSSION AND CONCLUSIONS

Since the fuzzy set was introduced by Zadeh, many new approaches and theories treating imprecision and uncertainty have been proposed, such as the interval-valued fuzzy sets. Among these theories, an extension of the classic fuzzy set is intuitionistic fuzzy set theory, which was introduced by Atanassov [9,10]. Since then, many researchers have investigated this topic and obtained some meaningful conclusions such as implication of intuitionistic fuzzy sets, generalization of intuitionistic fuzzy rough approximation operators, and multi-criteria decision-making methods based on intuitionistic fuzzy sets. Fuzzy entropy and similarity measure represent dual meaning each other, that is, dissimilarity and similarity between considering two sets. In the previous result, their relation was showed and discussed about each characteristics [6]. However, fuzzy entropy contains some debate from definition itself. In Definition 2.2, fuzzy entropy was defined by the hesitance area, there it was not necessary corresponding numerical data. Basically, fuzzy set contains the uncertainty in itself, which is evaluated by the entropy measure. Hence, additional term between $\mu_A(x_i)$ and $\mu_{A_{\text{near}}}$ is considered, which make possible to consider abnormal intuitionistic fuzzy membership function. Furthermore, relation between fuzzy entropy and similarity measure shows similar result compared to our previous result. With the help of fuzzy entropy and similarity measure of fuzzy set, analysis on IFSs also carried out. Usefulness of proposed measures is proved by evaluating the definitions. Relation between fuzzy entropy and similarity measure are also discussed. By discussing the relation, it was verified that one measure is derived from another measure, hence two measures satisfy dual concept for data analysis.

ACKNOWLEDGMENT

This work was supported by the Centre for Smart Grid and Information Convergence (CeSGIC) at Xi’an Jiaotong-Liverpool University.

REFERENCES


