ON ENTROPY OF FUZZY SETS

CHAO-MING HWANG
Department of Applied Mathematics, Chinese Culture University, HwaKang, Yangminshan, Taipei, Taiwan
hwang@faculty.pccu.edu.tw

MIIN-SHEN YANG
Department of Applied Mathematics, Chung Yuan Christian University, Chung-Li 32023, Taiwan
msyang@math.cycu.edu.tw

Received 8 February 2006
Revised 25 April 2008

In this paper we deal with the entropy of fuzzy sets. We first review several defined entropies of fuzzy sets and then propose a new one. Some comparisons are made with some existing entropies to show the effectiveness of the proposed one.

Keywords: Fuzzy sets; uncertainty; fuzziness; entropy.

1. Introduction

There often exists uncertainty in real systems. Probability has been traditionally used in modelling uncertainty. Since Zadeh\(^1\) originated the idea of fuzzy sets, fuzziness becomes another way to model uncertainty. On the other hand, measuring the fuzziness of fuzzy sets is an important step in fuzzy systems. Measures of fuzziness by contrast to fuzzy measures try to indicate the degree of fuzziness of a fuzzy set. The entropy of fuzzy sets is a measure of fuzziness between fuzzy sets. De Luca and Termini\(^1\) first introduced the axiom construction for entropy of fuzzy sets with reference to Shannon’s probability entropy. Ebanks\(^2\) suggested five properties for an entropy of fuzzy sets to satisfy. Yager\(^17\) defined fuzziness measures of fuzzy sets in terms of a lack of distinction between the fuzzy set and its negation based on \(L_p\) norm. Higashi and Klir\(^5\) extended Yager’s measure to a general class of fuzzy complements. Kosko\(^6\) provided a measure of fuzziness between fuzzy sets using a ratio of distance between the fuzzy set and its nearest set to the distance between the fuzzy set and its farthest set. Liu\(^10\) gave some axiom definitions of entropy and also defined a \(\sigma\)-entropy. Fan \textit{et al.}\(^3,4\) re-organized some fuzzy entropy formulas.
Pal and Pal\cite{13,14} proposed higher order and exponential entropies. More surveys on measuring fuzziness were given by Pal\cite{11} and Pal and Bezdek.\cite{12} Recently, Li and Liu\cite{8,9} consider new defined fuzzy variables and discuss about entropies on fuzzy variables. In Li and Liu,\cite{7} they first give the notion of credibility measure and then consider an entropy of credibility distributions for fuzzy variables. In this paper, we propose a new entropy of fuzzy sets. Some comparisons are made with some existing entropies to show the effectiveness of the proposed one.

2. A New Entropy of Fuzzy Sets

The following notations are used in this section. \( R^+ = [0, +\infty) \); \( X \) is the universal set; \( F(X) \) is the class of all fuzzy sets of \( X \); \( u_A(x) \) is the membership function of \( A \) in \( F(X) \); \( P(X) \) is the class of all crisp sets of \( X \); \( [a] \) is the fuzzy set of \( X \) for which \( u_{[a]}(x) = a, \forall x \in X \); \( F \) is a subclass of \( F(X) \) with (1) \( P(X) \subseteq F \); (2) \( \{x\} \subseteq F \); (3) \( A, B \in F \Rightarrow A \cup B \in F, A^c \in F \), where \( A^c \) is the complement of \( A \) in \( F(X) \) with \( u_{A^c}(x) = 1 - u_A(x), \forall x \in X \). A fuzzy set \( A^* \) is said to be crisper than \( A \), if \( u_{A^*}(x) \geq u_A(x) \), when \( u_A(x) \geq \frac{1}{2} \) and \( u_A(x) \leq u_A(x) \), when \( u_A(x) \leq \frac{1}{2} \). De Luca and Termini\cite{1} first gave axioms for entropy of fuzzy sets as follows.

**Definition 1.** (De Luca and Termini\cite{1}) A real function \( e : F \rightarrow R^+ \) is called an entropy on \( F \), if \( e \) has the following properties:

(E1) \( e(D) = 0 \) if \( D \in P(X) \);

(E2) \( e([\frac{1}{2}]) = \max_{A \in F} e(A) \);

(E3) If \( A^* \) is crisper than \( A \), then \( e(A^*) \leq e(A) \);

(E4) \( e(A^c) = e(A) \) for all \( A \in F \).

Assume that \( X = \{x_1, \ldots, x_n\} \). We review the following entropies of a fuzzy set \( A \) in \( F(X) \).

(a) De Luca and Termini\cite{1} defined

\[
e_{DT}(A) = -k \sum_{i=1}^{n} (u_A(x_i) \log(u_A(x_i)) + (1-u_A(x_i)) \log(1-u_A(x_i))), \quad \forall A \in F(X).
\]

(b) In Yager,\cite{17} he defined

\[
e_{Y_p}(A) = 1 - \frac{d_p(A, A^c)}{n^p}, \quad \forall A \in F(X).
\]

where \( d_p \) is defined as \( d_p(A, B) = (\sum_{i=1}^{n} |u_A(x_i) - u_B(x_i)|^p)^{\frac{1}{p}}, \forall A, B \in F(X) \).

(c) For a fuzzy set \( A \), Kosko\cite{6} defined

\[
e_{K}(A) = \frac{d_p(A, A_{\text{near}})}{d_p(A, A_{\text{far}})}, \quad \forall A \in F(X),
\]
where \( A_{\text{near}}, A_{\text{far}} \in P(X) \) are defined as

\[
u_{A_{\text{near}}}(x) = \begin{cases} 
1 & \text{if } u_{A}(x) \geq \frac{1}{2} \\
0 & \text{if } u_{A}(x) < \frac{1}{2},
\end{cases}
\]

\[
u_{A_{\text{far}}}(x) = \begin{cases} 
1 & \text{if } u_{A}(x) \leq \frac{1}{2} \\
0 & \text{if } u_{A}(x) > \frac{1}{2}.
\end{cases}
\]

(d) Pal and Pal\(^{14}\) defined

\[
p_{PP}(A) = \frac{1}{n} \sum_{i=1}^{n} (u_{A}(x_{i})e^{(1-u_{A}(x_{i}))} + (1 - u_{A}(x_{i}))e^{u_{A}(x_{i})}), \quad \forall A \in F(X).
\]

(e) Li and Liu\(^{7}\) defined

\[
p_{LL}(A) = \sum_{i=1}^{n} S(\text{Cr}(\xi_{A} = x_{i})) \quad \text{with} \quad S(t) = -(t)\ln(t) - (1 - t)\ln(1 - t)
\]

where \( \text{Cr}(\xi_{A} = x_{i}) = \frac{1}{2}(u_{A}(x_{i}) + 1 - \sup_{x \neq x_{i}} u_{A}(x)), \quad \forall A \in F(X). \)

It is obviously that \( p_{DT}, p_{Y}, p_{K}, p_{PP} \) and \( p_{LL} \) are all entropies on \( F(X) \) that satisfy the conditions of Definition 1.

Because an entropy for a fuzzy set \( A \) is a measure of fuzziness of \( A \), it is reasonable to have the fuzziness of \( A \) only focus on \( A \cap D \) and \( A \cup D^{c} \) for any \( D \in P(X) \). Therefore, Liu\(^{10}\) gave a definition of a \( \sigma \)-entropy on \( F \) as follows.

**Definition 2.** (Liu\(^{10}\)) Let \( e \) be an entropy on \( F \). If for any \( A \in F \), we have \( e(A) = e(A \cap D) + e(A \cap D^{c}) \forall D \in P(X) \), then \( e \) is called a \( \sigma \)-entropy on \( F \).

Based on Definition 2, we have the following \( \sigma \)-entropies on \( F(X) \):

(a) The entropy \( p_{DT} \) is a \( \sigma \)-entropy on \( F(X) \).

(b) When \( p = 1 \), the entropy \( p_{Y} \) is a \( \sigma \)-entropy on \( F(X) \).

(c) The entropy \( p_{PP} \) is a \( \sigma \)-entropy on \( F(X) \).

Next, we derive a new entropy of fuzzy sets. Let us first consider \( p_{Y} \) defined by Yager\(^{17}\) with \( p = 1 \). That is,

\[
p_{Y}(A) = 1 - \frac{1}{n} \sum_{i=1}^{n} |u_{A}(x_{i}) - u_{A^{c}}(x_{i})|
\]

Then,

\[
p_{Y}(A) = 1 - \frac{1}{n} \sum_{i=1}^{n} |u_{A}(x_{i}) - u_{A^{c}}(x_{i})|
\]

\[
= 1 - \frac{1}{n} \sum_{i=1}^{n} (1 - 2u_{A^{c}}(x_{i}))
\]

\[
= 1 - \frac{1}{n} \sum_{i=1}^{n} (1 - 2u_{A^{c}}(x_{i}))I_{[u_{A}(x_{i}) \geq \frac{1}{2}]} + (2u_{A^{c}}(x_{i}) - 1)I_{[u_{A}(x_{i}) < \frac{1}{2}]}
\]

\[
= 1 - \frac{1}{n} \sum_{i=1}^{n} (\frac{1}{2} - u_{A^{c}}(x_{i}))I_{[u_{A}(x_{i}) \geq \frac{1}{2}]} + (u_{A^{c}}(x_{i}) - \frac{1}{2})I_{[u_{A}(x_{i}) < \frac{1}{2}]}
\]

\[
= \frac{1}{2n} \sum_{i=1}^{n} u_{A^{c}}(x_{i})I_{[u_{A}(x_{i}) \geq \frac{1}{2}]} + u_{A}(x_{i})I_{[u_{A}(x_{i}) < \frac{1}{2}]}
\]
Based on the above derivation, we proposed a new entropy \( e(A) \) of a fuzzy set \( A \) as follows:

For a discrete universal set \( X = \{x_1, \cdots, x_n\} \),

\[
e(A) = \frac{\sum_{i=1}^{n} \left[ (1 - e^{-u_A(x_i)})I_{[u_A(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_A(x_i) < \frac{1}{2}]} \right]}{n(1 - e^{-\frac{1}{2}})}
\]

Similarly, for a continuous universal set \( X \),

\[
e(A) = \frac{\int_{x \in X} \left[ (1 - e^{-u_A(x)})I_{[u_A(x) \geq \frac{1}{2}]} + (1 - e^{-u_A(x)})I_{[u_A(x) < \frac{1}{2}]} \right]dx}{(1 - e^{-\frac{1}{2}})}
\]

It is obviously that \( e(A) \) satisfies the conditions of Definition 1 so that it is an entropy of fuzzy sets. We may say that the proposed entropy \( e(A) \) has combined the concepts of Yager\(^\text{17}\) and Pal and Pal\(^\text{14}\) by replacing the memberships \( u_A(x) \) and \( u_A(x) \) with the exponential distance-types \( (1 - e^{-u_A(x)}) \) and \( (1 - e^{-u_A(x)}) \). In fact, these exponential distance-types had been successfully used in clustering algorithms by Wu and Yang\(^\text{15}\) and Yang and Wu\(^\text{16}\) where they had claimed its robustness for clustering. We next claim that \( e(A) \) is a \( \sigma \)-entropy on \( F \).

**Property 1.** The proposed entropy \( e \) is a \( \sigma \)-entropy on \( F \).

**Proof.** By Definition 2, we need to prove that

\[
e(A) = e(A \cap D) + e(A \cap D^c), \quad \forall D \in P(X)
\]

Assume that the universal set \( X \) is discrete with \( X = \{x_1, \cdots, x_n\} \). We mention that \( D \) is a crisp set with \( u_D(x_i) \in \{0, 1\} \). Thus, we have

\[
e(A \cap D) = \frac{\sum_{i=1}^{n} \left[ (1 - e^{-u_A(x_i)})I_{[u_D(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_D(x_i) < \frac{1}{2}]} \right]}{n(1 - e^{-\frac{1}{2}})}
\]

\[
= \frac{\sum_{i=1}^{n} \left[ (1 - e^{u_A(x_i) - 1})I_{[u_A(x_i) \land u_D(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_A(x_i) \land u_D(x_i) < \frac{1}{2}]} \right]}{n(1 - e^{-\frac{1}{2}})}
\]

\[
= \frac{\sum_{i=1}^{n} \left[ (1 - e^{u_A(x_i) - 1})I_{[u_A(x_i) \geq \frac{1}{2}, u_D(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_A(x_i) \leq \frac{1}{2}, u_D(x_i) \geq \frac{1}{2}]} + 0 \right]}{n(1 - e^{-\frac{1}{2}})}
\]

\[
= \frac{\sum_{i=1}^{n} \left[ (1 - e^{-u_A(x_i)})I_{[u_A(x_i) \geq \frac{1}{2}, u_D(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_A(x_i) < \frac{1}{2}, u_D(x_i) \geq \frac{1}{2}]} \right]}{n(1 - e^{-\frac{1}{2}})}
\]
Similarly,
\[
e(A \cap D^c) = \frac{\sum_{i=1}^{n} (1 - e^{-u_{A^c}(x_i)})I_{[u_{A^c}(x_i) \geq \frac{1}{2}, u_{D^c}(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_A(x_i) < \frac{1}{2}, u_{D^c}(x_i) \geq \frac{1}{2}]} )}{n(1 - e^{-\frac{1}{2}})}.
\]

Hence,
\[
e(A \cap D) + e(A \cap D^c)
\]
\[
= \frac{\sum_{i=1}^{n} (1 - e^{-u_{A^c}(x_i)})I_{[u_{A^c}(x_i) \geq \frac{1}{2}, u_{D^c}(x_i) \geq \frac{1}{2}]} + I_{[u_A(x_i) < \frac{1}{2}, u_{D^c}(x_i) \geq \frac{1}{2}]} )}{n(1 - e^{-\frac{1}{2}})}
\]
\[
+ \sum_{i=1}^{n} (1 - e^{-u_A(x_i)})I_{[u_A(x_i) < \frac{1}{2}, u_{D^c}(x_i) \geq \frac{1}{2}]} + I_{[u_A(x_i) < \frac{1}{2}, u_{D^c}(x_i) \geq \frac{1}{2}]} )}{n(1 - e^{-\frac{1}{2}})}
\]
\[
= \frac{\sum_{i=1}^{n} ((1 - e^{-u_{A^c}(x_i)})I_{[u_{A^c}(x_i) \geq \frac{1}{2}]} + (1 - e^{-u_A(x_i)})I_{[u_A(x_i) < \frac{1}{2}]})}{n(1 - e^{-\frac{1}{2}})}
\]
\[
= e(A).
\]

Similarly, it can be proved for a continuous universal \(X\).

3. Comparisons and Results

The linguistic hedges, like “very”, “more or less”, “slightly”, are used to represent the modifiers of linguistic variables. Fuzzy sets are usually used as linguistic variables. Thus, the hedges may be viewed as operations on fuzzy sets (see Zimmerman\(^{19}\)). In this section, we will consider these operations on fuzzy sets and then make comparisons of the proposed entropy of fuzzy sets with others.

For a given fuzzy set \(A = \{(x, \mu_A(x))|x \in X\}\), the modifier \(A^n\) for the fuzzy set \(A\) is defined as follows:

\[
A^n = \{(x, (\mu_A(x))^n)|x \in X\}.
\]

We then define the concentration and dilation of \(A\) with concentration: \(CON(A) = A^2\) and dilation: \(DIL(A) = A^{1/2}\). The concentration and dilation are mathematical models frequently to be used for modifiers. Thus, we can use these mathematical operators to define the linguistic hedges on a fuzzy set \(A\) as follows.

Very \(A = CON(A) = A^2\)

more or less \(A = DIL(A) = A^{1/2}\)

Quite Very \(A = A^3\)

Very Very \(A = A^4\)
Let us consider a fuzzy set $A_1$ of $X = \{6, 7, 8, 9, 10\}$ with $A_1 = \{(6, 0.1), (7, 0.3), (8, 0.4), (9, 0.9), (10, 1)\}$. By taking into account the characterization of linguistic variables, we regarded $A_1$ as “Large” on $X$. We can generate the following fuzzy sets:

$$A_1^2 = \{(6, 0.316), (7, 0.548), (8, 0.632), (9, 0.949), (10, 1)\}$$
$$A_1^3 = \{(6, 0.01), (7, 0.09), (8, 0.16), (9, 0.81), (10, 1)\}$$
$$A_1^4 = \{(6, 0.001), (7, 0.027), (8, 0.64), (9, 0.729), (10, 1)\}$$
$$A_1^5 = \{(6, 0), (7, 0.008), (8, 0.026), (9, 0.656), (10, 1)\}.$$

The hedges represented by the above fuzzy sets are described as follows.

- $A_1^2$ may be treated as “More or less Large”,
- $A_1^3$ may be treated as “Very Large”,
- $A_1^4$ may be treated as “Quite Very Large”,
- $A_1^5$ may be treated as “Very Very Large”.

The comparison results are shown in Table 1 for those entropies of fuzzy sets:

1. $e_{Y_1}(A) = 1 - \frac{d_p(A, A^c)}{n^2}$ (Yager\(^{17}\));
2. $e_K(A) = \frac{d_p(A, A_{near})}{d_p(A, A_{far})}$ (Kosko\(^6\));
3. $e_{PP}(A) = \frac{1}{n} \sum_{i=1}^{n} u_{A}(x_i) e(1 - u_{A}(x_i)) + (1 - u_{A}(x_i)) e(u_{A}(x_i))$ (Pal and Pal\(^{14}\));
4. $e_{LL}(A) = \sum_{i=1}^{n} S(Cr(\xi_A = x_i))$ (Li and Liu\(^7\));
5. $e(A) = \frac{\sum_{i=1}^{n} (1 - e^{-u_{A}(x_i)}) I_{u_{A}(x_i) \geq \frac{1}{2}} + (1 - e^{-u_{A}(x_i)}) I_{u_{A}(x_i) < \frac{1}{2}}}{n(1 - e^{-\frac{1}{2}})}$ (our proposed one).

We use the following notions for different entropies.

- $E(Y_1)$ = the entropy measure with $p = 1$ by Yager.
- $E(K)$ = the entropy measure by Kosko.
- $E(P)$ = the entropy measure by Pal and Pal.
- $E(L)$ = the entropy measure by Li and Liu.
- $E(O)$ = the entropy measure by our proposed one.
Table 1. Results of the measures of fuzziness with different entropies.

<table>
<thead>
<tr>
<th>Fuzzy set</th>
<th>$E(Y_1)$</th>
<th>$E(K)$</th>
<th>$E(P)$</th>
<th>$E(L)$</th>
<th>$E(O)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>0.397</td>
<td>0.220</td>
<td>1.389</td>
<td>0.810</td>
<td>0.505</td>
</tr>
<tr>
<td>$A_1$</td>
<td>0.360</td>
<td>0.311</td>
<td>1.331</td>
<td>0.723</td>
<td>0.397</td>
</tr>
<tr>
<td>$A_2$</td>
<td>0.167</td>
<td>0.099</td>
<td>1.202</td>
<td>0.378</td>
<td>0.212</td>
</tr>
<tr>
<td>$A_3$</td>
<td>0.145</td>
<td>0.078</td>
<td>1.151</td>
<td>0.870</td>
<td>0.167</td>
</tr>
<tr>
<td>$A_4$</td>
<td>0.151</td>
<td>0.082</td>
<td>1.136</td>
<td>0.692</td>
<td>0.165</td>
</tr>
</tbody>
</table>

Table 2. Results of the measures of fuzziness with $E(P)$ and $E(O)$.

<table>
<thead>
<tr>
<th>Fuzzy set</th>
<th>$E(P)$</th>
<th>$E(O)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_2$</td>
<td>1.501</td>
<td>0.653</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.513</td>
<td>0.616</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.386</td>
<td>0.577</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.094</td>
<td>0.393</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1.241</td>
<td>0.298</td>
</tr>
</tbody>
</table>

According to the results of Table 1, we see that

\[ E(Y_1)(A_1^2) > E(Y_1)(A_1) > E(Y_1)(A_2^2) > E(Y_1)(A_3^2) < E(Y_1)(A_4^2). \]
\[ E(K)(A_1^2) < E(K)(A_1) > E(K)(A_2^2) > E(K)(A_3^2) < E(K)(A_4^2). \]
\[ E(O)(A_1^2) > E(O)(A_1) > E(O)(A_2^2) > E(O)(A_3^2) > E(O)(A_4^2). \]

Hence, the order of the entropy measures $E(P)$ and $E(O)$ presents better than others. Next, we compare only for $E(P)$ and $E(O)$.

We consider another fuzzy set $A_2$ of $X = \{6, 7, 8, 9, 10\}$ defined by

\[ A_2 = \{(6, 0.2), (7, 0.3), (8, 0.4), (9, 0.7), (10, 0.8)\}. \]

We generate $A_1^2, A_2^2, A_3^2$ and $A_4^2$ based on the previous operators. The results for the entropies $E(P)$ and $E(O)$ are shown in Table 2. According to the results of Table 2, we see that

Obviously, the order of the proposed entropy $E(O)$ presents better than $E(P)$. From the previous comparisons, the proposed entropy is actually better for presenting the measure of fuzziness for a fuzzy set.

4. Conclusions

Based on the Yager’s entropy formation,\textsuperscript{17} we proposed a new entropy of fuzzy sets. The hedges, like “very”, “highly”, “more or less”, with linguistic variables are used for the comparisons of our proposed one with several existing entropies. The results show the effectiveness of our proposed entropy.

Acknowledgments

The authors are grateful to the anonymous reviewers for their comments to improve the presentation of the paper. This work was supported in part by the National Science Council of Taiwan, under Miin-Shen Yang’s Grant: NSC 94-2118-M-033-001.

References