World Academy of Science, Engineering and Technology
International Journal of Physical and Mathematical Sciences
Vol:5, No:3, 2011

Normalization and Constrained Optimization of Measures of Fuzzy Entropy

K.C. Deshmukh, P.G. Khot, Nikhil

***Abstract*—**In the literature of information theory, there is necessity for comparing the different measures of fuzzy entropy and this consequently, gives rise to the need for normalizing measures of fuzzy entropy. In this paper, we have discussed this need and hence developed some normalized measures of fuzzy entropy. It is also desirable to maximize entropy and to minimize directed divergence or distance. Keeping in mind this idea, we have explained the method of optimizing different measures of fuzzy entropy.

***Keywords*—**Fuzzy set, Uncertainty, Fuzzy entropy, Normalization, Membership function

I. Introduction

ONE of the most radical and fruitful of the representational capabilities of logic for measuring fuzzy uncertainty was initiated by Zadeh [17] with the publication of his paper "Fuzzy Sets." Starting from the idea of gradual membership, it has been the basis for both logic of gradualness in properties and a new, particularly simple and effective, uncertainty calculus called "Possibility Theory" for handling the notions of possibility and certainty as gradual modalities. When proposing fuzzy sets, Zadeh's [17] concerns were explicitly centered on their potential contribution in the domains of pattern classification, processing and communication of information, abstraction and summarization. Although the claims that fuzzy sets were relevant in these areas appeared unsustained at the time when they were first uttered, namely in the early sixties, the future development of information sciences and engineering proved that these intuitions were right, beyond all expectations. Kapur [6] has well explained the concept of fuzzy uncertainty with the help of examples.

Basically, the Shannon's [16] entropy measures the average uncertainty in bits associated with the prediction of outcomes in a random experiment whereas fuzzy entropy is the quantitative description of fuzziness in fuzzy sets. De Luca and Termini [2] introduced some requirements which capture our intuition for the degree of fuzziness. Fuzzy entropy is one of the important digital features of fuzzy sets and occupies an important place in system model and system design. For example, when generalized fuzzy entropy is used as learning criterion for neural networks, efficient structure parameters are obtained quickly. In other words, generalized fuzzy entropy has better guidance function in neural network system design.

The theory of fuzzy sets which was introduced by Zadeh [17] received a good response from different quarters and after its introduction, many researchers started working around this field. Thus, keeping in view the idea of fuzzy sets, De Luca and Termini [2] introduced a measure of fuzzy entropy corresponding to Shannon's [16] measure. This fuzzy entropy is given by

$$H(A) = -\sum_{i=1}^{n}\begin{bmatrix}\mu_A(x_i)\log\mu_A(x_i) \\ +\left(1-\mu_A(x_i)\right)\log\left(1-\mu_A(x_i)\right)\end{bmatrix} \quad (1)$$

After this development, a large number of measures of fuzzy entropy were discussed, characterized and generalized by various authors. Kapur [6] introduced the following measure of fuzzy entropy:

$$H_{\alpha,\beta}(A) = \frac{1}{\beta-\alpha} \log\frac{\sum_{i=1}^{n}\left\{\mu_A{}^{\alpha}(x_i)+(1-\mu_A(x_i))^{\alpha}\right\}}{\sum_{i=1}^{n}\left\{\mu_A{}^{\beta}(x_i)+(1-\mu_A(x_i))^{\beta}\right\}};\alpha\geq 1,\beta\leq 1 \quad (2)$$

Parkash [10] introduced a new generalized measure of fuzzy entropy involving two real parameters, given by:

$$H_{\alpha}{}^{\beta}(A) = [(1-\alpha)\beta]^{-1}\sum_{i=1}^{n}\left[\left\{\mu_A{}^{\alpha}(x_i)+(1-\mu_A(x_i))^{\alpha}\right\}^{\beta}-1\right] \quad (3)$$
$$;\alpha>0,\alpha\neq 1,\ \beta\neq 0$$

and called it fuzzy entropy which includes some well-known entropies.

Parkash, Sharma and Kumar [13] have provided the characterizations of fuzzy measures by using the concepts of concavity and recursivity. Parkash and Sharma [12] have extended the applications of fuzzy measures to the field of coding theory whereas some desirable applications of weighted measures of fuzzy entropy for the study of maximum entropy principles have been provided by Parkash,

K.C. Deshmukh is with Department of Mathematics, Rashtrasant Tukdoji Maharaj, Nagpur University, Nagpur (kcdeshmukh2000@rediffmail.com).

P.G. Khot is with Department of Statistics, Rashtrasant Tukdoji Maharaj, Nagpur University, Nagpur

Nikhil is with Department of Mathematics, Rashtrasant Tukdoji Maharaj, Nagpur University. (Nagpur kalia_nikhil@yahoo.com)

Sharma and Mahajan [14,15]. Some other measures of fuzzy entropy and their generalizations have been studied by Zadeh [17], Kapur [6], Klir and Folger [7], Emptoz [3], Kandel [5], Hu, Yu [4], Lowen [8], Zimmermann [18], Pal and Bezdek [9] etc.

In section II, we have discussed the need for normalizing measures of fuzzy entropy and hence developed some normalized measures of fuzzy entropy. In section III, we have explained the method of optimizing different measures of fuzzy entropy.

II. Normalization of Various Measures of Fuzzy Entropy

We first of all discuss the need for normalizing measures of fuzzy entropy. The measure of fuzzy entropy due to De Luca and Termini [2] measures the degree of equality among $\mu_A(x_1), \mu_A(x_2), ..., \mu_A(x_n)$ the fuzzy values, that is, the greater the equality among, , the greater is the value of $H(A)$ and this entropy has its maximum value n log2 when all the fuzzy values are equal, that is, when each $\mu_A(x_i) = \frac{1}{2}$.

Thus, we have

$$H(F) = n\log 2 \tag{4}$$

where $F = \left\{\frac{1}{2}, \frac{1}{2}, ..., \frac{1}{2}\right\}$ is the most fuzzy distribution.

The entropy (1) also measures the uniformity of A or the 'closeness' of A to the most fuzzy distribution F, since according to Bhandari and Pal's [1] measure, the fuzzy directed divergence of A from F is given by

$$D(A,F) = \sum_{i=1}^{n} \left[\begin{array}{l} \mu_A(x_i)\log\dfrac{\mu_A(x_i)}{1/2} \\ +(1-\mu_A(x_i))\log\dfrac{1-\mu_A(x_i)}{1/2} \end{array} \right]$$

$$= n\log 2 - H(A) \tag{5}$$

Thus, greater the value of entropy H (A), the nearer is A to F. This entropy provides a measure of equality or uniformity of the fuzzy values $\mu_A(x_1), \mu_A(x_2), ..., \mu_A(x_n)$ among themselves. Now, let us consider the following fuzzy distributions:

$$\mu_A(x_i) = (0.4, 0.4, 0.4, 0.4)$$

$$\mu_B(x_i) = (0.3, 0.3, 0.4, 0.4, 0.4)$$

Then, we have

$$H(A) = 1.2734 \ , \ H(B) = 1.5825 > H(A)$$

We want to check which fuzzy distribution is more uniform or in which distribution the fuzzy values are more equal? The obvious answer is that fuzzy values of A are more equal than the fuzzy values of B. Thus, A is more uniformly distributed than B but still $H(A) < H(B)$. From the values of the two fuzzy entropies, it appears that B is more uniform than A. This is obviously wrong. The fallacy arises due to the fact that the fuzzy entropy depends not only on the degree of equality among the fuzzy values; it also depends on the value of n. So long as n is the same, entropy can be used to compare the uniformity of the fuzzy distributions. But, if the number of outcomes are different, then fuzzy entropy is not a satisfactory measure of uniformity. In this case, we try to eliminate the effect of n by normalizing the fuzzy entropy, that is, by defining a normalized measure of fuzzy entropy as

$$\overline{H}(A) = \frac{H(A)}{\max H(A)} \tag{6}$$

It is obvious that

$$0 \le \overline{H}(A) \le 1 \tag{7}$$

For De Luca and Termini's [2] measure of fuzzy entropy, we have

$$\overline{H}(A) = -\frac{\displaystyle\sum_{i=1}^{4} \left[\begin{array}{l} \mu_A(x_i) \ \ \log\mu_A(x_i) \\ \\ +\ (1-\mu_A(x_i))\log(1-\mu_A(x_i)) \end{array} \right]}{4\log 2}$$

= 1.05754

and

$\overline{H}(B)$ = 1.05149

Obviously, $\overline{H}(A) > \overline{H}(B)$

Thus, A is more uniform than B. This gives the correct result that B is less uniform than A. Thus, to compare the uniformity, or equality or uncertainty of two fuzzy distributions, we should compare their normalized measures of fuzzy entropy.

Next, we have obtained the expression of the normalized measure of fuzzy entropy. For this purpose, we consider Parkash and Sharma's [11] parametric measure of fuzzy entropy of order, given by:

$$H_a(A) = \sum_{i=1}^{n} \left[\begin{array}{c} \log\left(1 + a\mu_A(x_i)\right) \\ +\log\left(1 + a\left(1 - \mu_A(x_i)\right)\right) \\ -\log(a+1) \end{array} \right]; \ a > 0 \tag{8}$$

The maximum value of (8) is given by

$$\left[H_a(A)\right]_{\max} = n\log\frac{\left\{\dfrac{a+2}{2}\right\}^2}{1+a} \tag{9}$$

Thus, the expression for Parkash and Sharma's [11] normalized measure is given by

$$\left[H_a(A)\right]_N = \frac{\sum_{i=1}^{n}\left[\begin{array}{l}\log\left(1+a\mu_A\left(x_i\right)\right)+\\ \log\left(1+a\left(1-\mu_A\left(x_i\right)\right)\right)-\log\left(a+1\right)\end{array}\right]}{n\log\frac{\left\{\frac{a+2}{2}\right\}^2}{1+a}} \quad (10)$$

Proceeding on similar lines, we can obtain maximum values for different fuzzy entropies and consequently develop many other expressions of the normalized measures of fuzzy entropy.

III. OPTIMUM VALUES OF VARIOUS MEASURES OF FUZZY ENTROPY UNDER A SET OF CONSTRAINTS

In this section, we explain the method of optimizing different measures of fuzzy entropy. Firstly, we consider the fuzzy entropy introduced by Parkash and Sharma [11] for measuring its maximum value. Thus, our problem becomes to maximize the following measure:

$$H_a(A) = -\frac{1}{a}\sum_{i=1}^{n}\left\{\begin{array}{l}\left(1+a\mu_A\left(x_i\right)\right)\\ \quad\log\left(1+a\mu_A\left(x_i\right)\right)\\ +\left(1+a\left(1-\mu_A\left(x_i\right)\right)\right)\\ \quad\log\left(1+a\left(1-\mu_A\left(x_i\right)\right)\right)\\ -\left(1+a\right)\log\left(1+a\right)\end{array}\right\} \quad (11)$$

$$;a>0$$

under the following fuzzy constraint

$$\sum_{i=1}^{n}\mu_A\left(x_i\right)=\alpha_0 \quad (12)$$

The corresponding Lagrangian is

$$L = -\frac{1}{a}\sum_{i=1}^{n}\left\{\begin{array}{l}\left(1+a\mu_A\left(x_i\right)\right)\log\left(1+a\mu_A\left(x_i\right)\right)\\ +\left(1+a\left(1-\mu_A\left(x_i\right)\right)\right)\\ \log\left(1+a\left(1-\mu_A\left(x_i\right)\right)\right)\\ -\left(1+a\right)\log\left(1+a\right)\end{array}\right\} +\lambda_0\left\{\sum_{i=1}^{n}\mu_A\left(x_i\right)-\alpha_0\right\}$$

Now

$$\frac{\partial L}{\partial\mu_A\left(x_i\right)} = -\frac{1}{a}\left[\begin{array}{l}a\log\left(1+a\mu_A\left(x_i\right)\right)-\\ a\log\left(1+a\left(1-\mu_A\left(x_i\right)\right)\right)\end{array}\right]+\lambda_0$$

Thus $\frac{\partial L}{\partial\mu_A\left(x_i\right)} = 0$ implies that

$$\mu_A\left(x_i\right) = \frac{(1+a)e^{\lambda_0}-1}{a\left(1+e^{\lambda_0}\right)} \quad (13)$$

Thus, from equation (12), we have

$$\alpha_0 = \sum_{i=1}^{n}\frac{(1+a)e^{\lambda_0}-1}{a\left(1+e^{\lambda_0}\right)} \quad (14)$$

where is to be determined from the following equation:

$$f(\lambda_0) = \sum_{i=1}^{n}\frac{(1+a)e^{\lambda_0}-1}{a\left(1+e^{\lambda_0}\right)}-\alpha_0 \quad (15)$$

Now differentiating equation (15) w.r.t. λ_0, we get

$$f'(\lambda_0) = a\sum_{i=1}^{n}\frac{(1+a)e^{\lambda_0}+e^{\lambda_0}}{\left(a\left(1+e^{\lambda_0}\right)^2\right)}$$

$$= a\sum_{i=1}^{n}\frac{(2+a)e^{\lambda_0}}{\left(a\left(1+e^{\lambda_0}\right)^2\right)} > 0 \quad (16)$$

Equation (16) shows that $f(\lambda_0)$ is an increasing function of λ_0.

Now $f(-\infty) = -\frac{n}{a}-\alpha_0$

$$f(0) = \frac{n}{2}-\alpha_0$$

$$f(\infty) = \frac{n(1+a)}{a}-\alpha_0$$

Hence, we see that $f(\lambda_0)$ has a unique real root λ_0.

Now $\lambda_0 < 0$ if $f(0)>0$, that is, if $\alpha_0 < \frac{n}{2}$ then $\mu_A(x_i) \le \frac{1}{2}$

$\lambda_0 > 0$ if $f(0)<0$, that is, if $\alpha_0 > \frac{n}{2}$ then $\mu_A(x_i) \ge \frac{1}{2}$

Thus, we see that the maximizing $\mu_A(x_i)$ are either $\le \frac{1}{2}$ or $\ge \frac{1}{2}$.

We, now consider the following two cases:

Case-I: When $\mu_A(x_i) \le \frac{1}{2}$ for each i.

Then the maximum value of the entropy (11) is given by

$$[H_a(A)]_{\max} = -\frac{1}{a}\sum_{i=1}^{n}\left\{\begin{array}{l}\left(1+a\mu_A(x_i)\right)\\ \quad\log\left(1+a\mu_A(x_i)\right)\\ +\left(1+a\left(1-\mu_A(x_i)\right)\right)\\ \quad\log\left(1+a\left(1-\mu_A(x_i)\right)\right)\\ -(1+a)\log(1+a)\end{array}\right\} \quad ; a>0$$

where

$$\mu_A(x_i) = \frac{(1+a)e^{\lambda_0}-1}{a\left(1+e^{\lambda_0}\right)}$$

and

$$\alpha_0 = \sum_{i=1}^{n}\frac{(1+a)e^{\lambda_0}-1}{a\left(1+e^{\lambda_0}\right)}$$

Now

$$\frac{d}{d\mu_A(x_i)}[H_a(A)]_{\max} = -\log\frac{1+a\mu_A(x_i)}{1+a\left(1-\mu_A(x_i)\right)}$$

Now, for fixed a, since $\mu_A(x_i) \le \frac{1}{2}$, we have

$$\frac{1+a\mu_A(x_i)}{1+a\left(1-\mu_A(x_i)\right)} \le 1$$

Thus, we must have

$$\frac{d}{d\mu_A(x_i)}[H_a(A)]_{\max} \ge 0$$

Thus, we notice that the maximum fuzzy entropy is an increasing function of $\mu_A(x_i)$.

Also, we have

$$\frac{d}{d\lambda_0}\mu_A(x_i) = \frac{(2+a)a\,e^{\lambda_0}}{\left(a+e^{\lambda_0}\right)^2} > 0$$

This shows that $\mu_A(x_i)$ increase as λ_0 increases.

Again, we have

$$\frac{d\alpha_0}{d\lambda_0} = \sum_{i=1}^{n}\frac{(2+a)a\,e^{\lambda_0}}{\left(a+e^{\lambda_0}\right)^2} > 0$$

This shows that as α_0 increases, λ_0 increases, that is, $\sum_{i=1}^{n}\mu_A(x_i)$ increases.

Thus, we see that the maximum fuzzy entropy is an increasing function of α_0 and is maximum at $\alpha_0 = \frac{n}{2}$, that is, when $\mu_A(x_i) = \frac{1}{2}$ for each i.

Hence, the maximum value is given by

$$[H_a(A)]_{\max} = -\frac{1}{a}\sum_{i=1}^{n}\left[\begin{array}{l}\left(1+\frac{a}{2}\right)\log\left(1+\frac{a}{2}\right)\\ +\left(1+\frac{a}{2}\right)\log\left(1+\frac{a}{2}\right)\\ -(1+a)\log(1+a)\end{array}\right]$$

$$= \sum_{i=1}^{n}\left[\frac{(1+a)\log(1+a)}{a} - \frac{(2+a)}{a}\log\frac{(2+a)}{2}\right] \quad (17)$$

Thus, we see that the maximum fuzzy entropy is an increasing function of α_0 and it increases from 0 to $\sum_{i=1}^{n}\left[\frac{(1+a)\log(1+a)}{a} - \frac{(2+a)}{a}\log\frac{(2+a)}{2}\right]$

Case-II: When $\mu_A(x_i) \ge \frac{1}{2}$ for each i.

In this case, we see that for fixed a, we have

$$\frac{d}{d\mu_A(x_i)}[H_a(A)]_{\max} = -\log\frac{1+a\mu_A(x_i)}{1+a\left(1-\mu_A(x_i)\right)} < 0$$

Thus, we notice that the maximum fuzzy entropy is a decreasing function of α_0 and assumes its maximum value at $\alpha_0 = \frac{1}{2}$ and then starts decreasing and it decreases from $\sum_{i=1}^{n}\left[\frac{(1+a)\log(1+a)}{a} - \frac{(2+a)}{a}\log\frac{(2+a)}{2}\right]$ to 0.

Hence, the maximum value is given by equation (17).

With similar arguments, we can find the maximum values of the other existing well known measures of fuzzy entropy.

References

[1] Bhandari, D. and Pal, N.R. (1993). Some new information measures for fuzzy sets. Information Sciences 67: 209-228.

[2] De Luca, A. and Termini, S. (1972). A definition of non-probabilistic entropy in setting of fuzzy set theory. Information and Control 20: 301-312.
[3] Emptoz, H. (1981). Non-probabilistic entropies and indetermination process in the setting of fuzzy set theory. Fuzzy Sets and Systems 5: 307-317.
[4] Hu,Q.and Yu, D. (2004). Entropies of fuzzy indiscernibility relation and its operations. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12: 575-589.
[5] Kandel, A. (1986). Fuzzy Mathematical Techniques with Applications. Addison-wesley.
[6] Kapur, J.N. (1997). Measures of Fuzzy Information. Mathematical Sciences Trust Society, New Delhi.
[7] Klir, G.J. and Folger, T.A. (1988). Fuzzy Sets, Uncertainty and Indetermination. Prentice Hall, New York.
[8] Lowen, R. (1996). Fuzzy Set Theory-Basic Concepts, Techniques and Bibliography. Kluwer Academic Publishers, Boston.
[9] Pal, N.R. and Bezdek, J.C. (1994). Measuring fuzzy uncertainty. IEEE Transaction on Fuzzy Systems 2: 107-118.
[10] Parkash, O. (1998). A new parametric measure of fuzzy entropy. Information Processing and Management of Uncertainty 2:1732-1737.
[11] Parkash, O. and Sharma, P.K. (2004). Measures of fuzzy entropy and their relations. Inernationa. Journal of Management & Systems 20 : 65-72.
[12] Parkash, O. and Sharma, P. K. (2004). Noiseless coding theorems corresponding to fuzzy entropies. Southeast Asian Bulletin of Mathematics 27: 1073-1080.
[13] Parkash, O., Sharma, P. K. and Kumar, J. (2008). Characterization of fuzzy measures via concavity and recursivity. Oriental Journal of Mathematical Sciences 1:107-117.
[14] Parkash, O, Sharma, P. K. and Mahajan, R (2008). New measures of weighted fuzzy entropy and their applications for the study of maximum weighted fuzzy entropy principle. Information Sciences 178: 2389-2395.
[15] Parkash, O., Sharma, P. K. and Mahajan, R (2010). Optimization principle for weighted fuzzy entropy using unequal constraints. Southeast Asian Bulletin of Mathematics 34: 155-161.
[16] Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal 27: 379-423, 623-659.
[17] Zadeh, L. A. (1965). Fuzzy sets. Information and Control 8: 338-353.
[18] Zimmermann, H. J. (2001). Fuzzy Set Theory and its Applications. Kluwer Academic Publishers, Boston.