Asian Journal of Fuzzy and Applied Mathematics (ISSN: 2321 – 564X)
Volume 02 – Issue 04, August 2014

Generalized Fuzzy Inaccuracy Measure and its Applications

M.A.K.Baig and Nusrat Mushtaq
P.G. Department of Statistics
University of Kashmir, Srinagar-190006 (INDIA).
Corresponding author email: baigmak {at} gmail.com

Abstract: - *In this communication, we propose a generalized fuzzy inaccuracy measure and developed some coding theorems. We also study its particular cases.*

Keywords: - *Fuzzy set, Shannon's inequality, Generalized Shannon's inequality, Coding theorem, Kerridge inaccuracy.*

1. Introduction.

Zadeh [12] introduced the concept of Fuzzy sets in which imprecise knowledge can be used to define an event. The importance of fuzzy sets comes from the fact that it can deal with imprecise and inexact information.

Fuzzy sets play a significant role in many deployed system because of their capability to model non-statistical imprecision. Consequently, characterization and quantification of fuzziness are important issues that affect the management of uncertainty in many system models and designs.

A fuzzy set is represented as

$$A = \{x_i/\mu_A(x_i) : i = 1,2,\dots,n\},$$

where $\mu_A(x_i)$ gives the degree of belongingness of the element 'x_i' to the set 'A'. If every element of the set 'A' is '0' or '1', there is no uncertainty about it and a set is said to be a crisp set. On the other hand, a fuzzy set 'A' is defined by a characteristic function

$$\mu_A(x_i) = \{x_1, x_2, .., x_n\} \to [0,1].$$

The function $\mu_A(x_i)$ associates with each $x_i \in R^n$ grade of membership function.

A fuzzy set A^* is called a sharpened version of fuzzy set A if the following conditions are satisfied:

$$\mu_{A^*}(x_i) \le \mu_A(x_i), \quad if \;\; \mu_A(x_i) \le 0.5 \;\; for\ all\ i = 1,2,..,n$$

and $$\mu_{A^*}(x_i) \ge \mu_A(x_i), \quad if \;\; \mu_A(x_i) \ge 0.5 \; for\ all\ i = 1,2,..,n$$

De Luca and Termini [7] formulated a set of properties and these properties are widely accepted as criterion for defining any fuzzy entropy. In fuzzy set theory, the entropy is a measure of fuzziness which expresses the amount of average ambiguity in making a decision whether an element belong to a set or not. So, a measure of average fuzziness is fuzzy set H(A) should have the following properties to be a valid entropy.

i. (Sharpness): $H(A)$ is minimum if and only if A is a crisp set
   i.e. $\mu_A(x_i) = 0 \; or \; 1; \; \forall_i$
ii. (Maximality): $H(A)$ is maximum if and only if A is most fuzzy set
   i.e., $\mu_A(x_i) = \frac{1}{2} \;\; \forall_i$
iii. (Resolution): $H(A^*) \le H(A)$ where A^* is sharpened version of A.
iv. (Symmetry): $H(A) = H(\overline{A})$, where $\overline{A}$ is the complement of set A
   i.e. $\bar{\mu}_A(x_i) = 1 - \mu_A(x_i)$.

The importance of fuzzy set comes from the fact that it can deal with imprecise and inexact information.

2. Basic Concepts

Let X be discrete random variable taking on a finite number of possible values $X = (x_1, x_2, \ldots, x_n)$ with respective membership function $A = \{\mu_A(x_1), \mu_A(x_2), \ldots \mu_A(x_n)\} \to [0,1]$, $\mu_A(x_i)$ gives the elements the degree of belongingness x_i to the set A. The function $\mu_A(x_i)$ associates with each $x_i \in R^n$ a grade of membership to the set A and is known as membership function.

Denote

$$X = \begin{bmatrix} x_1 & x_2 & \ldots & x_2 \\ \mu_A(x_1) & \mu_A(x_2) & \cdots & \mu_A(x_n) \end{bmatrix} \tag{2.1}$$

We call the scheme (2.1) as a finite fuzzy information scheme. Every finite scheme describes a state of uncertainty.

Let a finite source of n source symbols $X = (x_1, x_2, \ldots, x_n)$ be encoded using alphabet of D symbols, then it has been shown by Feinstein [2] that there is a uniquely decipherable/ instantaneous code with lengths $l_1, l_2 \ldots, l_n$ iff the following Kraft [6] inequality is satisfied

$$\sum_i^n D^{-l_i} \leq 1 \tag{2.2}$$

3. Generalization of Fuzzy Shannon's Inequality.

Let $\Gamma_n = \{P = (p_1, p_2, \ldots, p_n);\ p_i \geq 0, \sum_{i=1}^n p_i = 1\}, n \geq 2$ be a set of complete probability distributions.

For $P \in \Gamma_n$, Shannon's measure of information [8] is defined as

$$H(P) = -\sum_{i=1}^n p_i log p_i \tag{3.1}$$

The measure (3.1) has been generalized by various authors and has found applications in various disciplines such as economics, accounting, crime, physics, etc.

If P, Q $\epsilon \Gamma_n$,. Then the kerriage inaccuracy measure is defined by

$$H(P, Q) = -\sum_i^n p_i \log q_i \tag{3.3}$$

De Luca and termini [7] introduced a quantity which, in a reasonable way to measures the amount of uncertainty (fuzzy entropy) associated with a given finite scheme. This measure is given by

$$H(A) = -\sum_{i}^{n} \left[\mu_A(x_i) \, log \, \mu_A(x_i) + \left(1 - \mu_A(x_i)\right) log\left(1 - \mu_A(x_i)\right)\right] \qquad (3.4)$$

The measure (3.3) serve as a very suitable measure of fuzzy entropy of the finite information scheme (2.1).

For $A, B \in \Gamma_n$, then Fuzzy Kerridge inaccuracy can be express as

$$H(A,B) = -\sum_{i=1}^{n} \{\mu_A(x_i) + (1 - \mu_A(x_i))\} log\{\mu_B(x_i) + (1 - \mu_B(x_i))\} \qquad (3.4)$$

The Generalized fuzzy entropy corresponding to Sharma and Mittal [9] is given by the following form:

$$H(A;\, \alpha, \beta) = \frac{1}{2^{1-\beta} - 1}\left[\left(\sum_{i=1}^{n}\{\mu_A^{\alpha}(x_i) + (1 - \mu_A(x_i))^{\alpha}\}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right],$$

$$\alpha, \beta > 0,\ \ \alpha \neq \beta, \alpha \neq 1 \neq \beta \qquad (3.5)$$

There is well known relation between $H(A)$ and $H(A, B)$ which is given by

$$H(A) \leq H(A, B) \tag{3.6}$$

The relation (3.6) is known as Fuzzy Shannon inequality and its importance is well known in coding theory.

In the literature of information theory, there are two approaches to extend the relation (3.6) for other measures. Nath and Mittal [8] extended the relation (3.6) in the case of entropy of type β.

Using the method of Nath and Mittal [8], Lubbe [11] generalizes (3.6) in the case of Renyi's entropy. On the other hand, using the method of Campbell, Lubbe [11] generalized (3.6) for the case of entropy of type β. Using these generalizations, coding theorems are proved by these authors for these measures.

The objective of this communication is to generalize (3.6) and give its application in coding theory.

4. Main Result:

For $A, B \in \Gamma_n$, defines a measure of inaccuracy, denoted by $H(A, B; \alpha, \beta)$ as

$$H(A,B;\alpha,\beta) = \frac{1}{2^{1-\beta}-1}\left[\left(\sum_{i=1}^{n}\{\mu_A(x_i) + (1-\mu_A(x_i))\}\{\mu_B(x_i) + (1-\mu_B(x_i))\}^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\beta-1}{\alpha-1}\alpha} - 1\right]$$

$$\alpha, \beta > 0, \;\; \alpha \neq \beta, \alpha \neq 1 \neq \beta \tag{4.1}$$

Since $H(A, B; \alpha, \beta) \neq H(A; \alpha, \beta)$, we will not interpret (4.1) as a measure of inaccuracy. But $H(A, B; \alpha, \beta)$ is a generalization of the measure of inaccuracy defined in (3.5). In spite of the fact that $H(A, B; \alpha, \beta)$ is not a measure of inaccuracy in its usual sense, its study is justified because it leads to meaningful new measures of length. In the following theorem, we will determine a relation between (3.5) and (4.1) of the type (3.6).

Since (4.1) is not a measure of inaccuracy in its usual sense, we will call the generalized relation as Pseudo-generalization of the Fuzzy Shannon inequality.

Theorem 1. If $A, B \in \Gamma_n$, then it holds that

$$H(A;\ \alpha,\beta)\ \leq H(A,B;\alpha,\beta) \tag{4.2}$$

With equality holds if

$$\{\mu_B(x_i) + (1-\mu_B(x_i))\} = \frac{\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}}{\sum_{i=1}^n \{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}}, \quad i = 1,2,\dots,n.$$

Proof. (a) $0 < \alpha < 1 < \beta$.

Using Holder's inequality, we get

$$\left(\sum_{i=1}^n \{\mu_A(x_i) + (1-\mu_A(x_i))\}\{\mu_B(x_i) + (1-\mu_B(x_i))\}^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}} \left(\sum_{i=1}^n \{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}\right)^{\frac{1}{1-\alpha}}$$

$$\leq 1, \quad \alpha > 0, \alpha \neq 1 \tag{4.3}$$

Since $\alpha < 1$, (4.3) becomes

$$\sum_{i=1}^n \{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\} \leq \left(\sum_{i=1}^n \{\mu_A(x_i) + (1-\mu_A(x_i))\}\{\mu_B(x_i) + (1-\mu_B(x_i))\}^{\frac{\alpha-1}{\alpha}}\right)^\alpha \tag{4.4}$$

Raising both sides of (4.4) with $(\beta-1)/\alpha-1\ (<0)$ we get

$$\left(\sum_{i=1}^n \{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}\right)^{\frac{\beta-1}{\alpha-1}}$$

$$\geq \left(\sum_{i=1}^{n}\{\mu_A(x_i)+(1-\mu_A(x_i))\}\{\mu_B(x_i)+(1-\mu_B(x_i))\}^{\frac{\alpha-1}{\alpha}}\right)^{\alpha\frac{\beta-1}{\alpha-1}} \tag{4.5}$$

Using (4.5) and the fact that $\beta > 1$, we get (4.2).

(b) $\alpha > 1, \beta > 1; 0 < \alpha < 1, \beta > 1$ $(\alpha < \beta \ or \ \beta < \alpha); 0 < \beta < 1 < \alpha$.

The proof follows on the similar lines.

We will now give an application of Theorem 1 in coding theory. Let a finite set of n-input symbols with probabilities $p_1, p_2, \dots, p_n$ be encoded in terms of symbols taken from the alphabet $\{a_1, a_2, \dots, a_n\}$, there always exists a uniquely decipherable code with length $N_1, N_2, \dots, N_n$ iff

$$\sum_{i=1}^{n} D^{-N_i} \leq 1, \tag{4.6}$$

If $L = \sum_{i=1}^{n}\{\mu_A(x_i)+(1-\mu_A(x_i))\}\, N_i$ is the average codeword length, then for a code which satisfies (4.6) it has been shown that [2],

$$L \geq H(P) \tag{4.7}$$

With equality iff $N_i = -log\{\mu_A(x_i)+(1-\mu_A(x_i))\}, i = 1,2,\dots,n$.

We define the measure of length $L(\alpha,\beta)$ by

$$L(\alpha,\beta) = \frac{1}{2^{1-\beta}-1}\left[\left(\sum_{i=1}^{n}\{\mu_A(x_i)+(1-\mu_A(x_i))\}D^{N_i\frac{\alpha-1}{\alpha}}\right)^{\alpha\frac{\beta-1}{\alpha-1}} - 1\right],$$

$$\alpha,\beta > 0, \qquad \alpha \neq \beta, \qquad \alpha \neq 1 \neq \beta. \tag{4.8}$$

We prove the following theorem in respect to the relation between $L(\alpha,\beta)$ and $H(A;\ \alpha,\beta)$.

Theorem 2. If $N_i, i = 1,2,\dots,n$ are the lengths of codewords satisfying (4.6), then

$$H(A;\ \alpha,\beta) \leq\ L(\alpha,\beta) < D^{1-\beta} H(A;\ \alpha,\beta) + \frac{1-D^{1-\beta}}{1-2^{1-\beta}}. \tag{4.9}$$

Proof. In (4.2), choose $B = (b_1, b_2, \dots, b_n)$ where

$$\mu_B = \frac{D^{-N_i}}{\sum_{i=1}^{n} D^{-N_i}} \tag{4.10}$$

With choice of B, (4.2) becomes

$$H(A;\ \alpha,\beta) \leq \frac{1}{2^{1-\beta}-1}\left[\left(\sum_{i=1}^{n}\{\mu_A(x_i) + (1-\mu_A(x_i))\}\left(\frac{D^{-N_i}}{\sum_{i=1}^{n} D^{-N_i}}\right)^{\frac{\alpha-1}{\alpha}}\right)^{\alpha\frac{\beta-1}{\alpha-1}} - 1\right]$$

$$= \frac{1}{2^{1-\beta}-1}\left[\left(\sum_{i=1}^{n}\{\mu_A(x_i) + (1-\mu_A(x_i))\}\frac{D^{N_i\frac{\alpha-1}{\alpha}}}{\left(\sum_{i=1}^{n} D^{-N_i}\right)^{\frac{\alpha-1}{\alpha}}}\right)^{\alpha\frac{\beta-1}{\alpha-1}} - 1\right]$$

$$= \frac{1}{2^{1-\beta}-1}\left[\left(\sum_{i=1}^{n}\{\mu_A(x_i) + (1-\mu_A(x_i))\}D^{N_i\frac{\alpha-1}{\alpha}}\right)^{\alpha\frac{\beta-1}{\alpha-1}}\left(\sum_{i=1}^{n} D^{-N_i}\right)^{\beta-1} - 1\right]$$

Using the relation (4.6), we get

$H(A;\ \alpha,\beta) \leq\ L(A,\alpha,\beta)$ which proves the first part of (4.9).

The equality holds iff $D^{-N_i} = \frac{\{\mu_A^{\alpha}(x_i)+(1-\mu_A(x_i))^{\alpha}\}}{\sum_{i=1}^{n}\{\mu_A^{\alpha}(x_i)+(1-\mu_A(x_i))^{\alpha}\}}, i = 1,2,\dots,n$ which is equivalent to

$$N_i = -log\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\} + log\left[\sum_{i=1}^{n}\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}\right], i = 1,2,.,n \quad (4.11)$$

Choose all N_i such that

$$-log\frac{\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}}{\sum_{i=1}^{n}\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}} \le N_i < -log\frac{\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}}{\sum_{i=1}^{n}\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}} + 1$$

Using the above relation, it follows that

$$D^{-N_i} \ge \frac{\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}}{\sum_{i=1}^{n}\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}D} \quad (4.12)$$

We now have two possibilities:

1) If $\alpha > 1$, (4.12) gives us

$$\left[\sum_{i=1}^{n}\{\mu_A(x_i) + \left(1-\mu_A(x_i)\right)\}D^{N_i\frac{\alpha-1}{\alpha}}\right]^\alpha \ge \sum_{i=1}^{n}\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}D^{1-\alpha} \quad (4.13)$$

Now consider two cases:

(a) Let $0 < \beta < 1$

Raising both sides of (4.13) with $\beta - 1/\alpha - 1$ we get

$$\left[\sum_{i=1}^{n}\{\mu_A(x_i) + (1-\mu_A(x_i))\}D^{N_i\frac{\alpha-1}{\alpha}}\right]^{\alpha(\frac{\beta-1}{\alpha-1})} \le \left[\sum_{i=1}^{n}\{\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha\}D^{1-\alpha}\right]^{\frac{\beta-1}{\alpha-1}} \quad (4.14)$$

Since $2^{1-\beta} - 1 > 0$ for $\beta < 1$, we get from (4.14) the right hand side in (4.9).

(b) Let $\beta > 1$. The proof follows similarly.

(c) $0 < \alpha < 1$, the proof follows on the same lines.

Remarks.

1) Since $D \ge 2$, we have

$$\frac{1 - D^{1-\beta}}{1 - 2^{1-\beta}} \geq 1$$

It follows then the upper bound of $L(\alpha, \beta)$ in (4.9) is greater than unity.

2) If $\beta = \alpha$ (4.9) reduces to the Nath and Mittal [8] entropy of type β.

References

[1] Belis and S.Guiasu: Fuzzy sets with inner interdependence, Information Sciences, (1968) 79(3-4), pp. 315-338.

[2] A. Feinstein: Foundations of information theory, McGraw-Hill, New York (1958).

[3] S. Guiasu, C.F. Picard: Borne Inferieure de la longueur utile de certain codes, C.R. Acad. Sci. Paris 273 (1971) 248-251

[4] D.F.Kerridge: Inaccuracy and inference. J. Roy. Statist. Soc. Sec. B 23 (1961), 184-194.

[5] G.J.Klir and B. Parviz: A note on the measure of discord, Proc. Eighth conference on Artificial Intelligence, San Mateo, California, (1992) pp.138-141.

[6] D.H. Kraft: Fuzzy sets and generalized retrieval systems, Intern. J. of Man-Machine studies, (1949) Vol. 19 (1), pp. 45-56.

[7] A. De Luca and S.Termini: A definition of non- probabilistic entropy in setting of fuzzy set theory. Information and Control (1972) Vol. 20, pp. 301-312.

[8] P. Nath and D.P.Mittal: A generalization of Shannon's inequality and its application in coding theory. Inform. and Control (1973), Vol. 23, pp. 439-445.

[9] C.E. Shannon: A mathematical theory of communication, Bell System Technical Journal, (1948), Vol. 27, pp. 379-423, 623-659.

[10] B.D.Sharma and D.P.Mittal: New non-additive measures of entropy for discrete probability distribution, J. Math. Sci. (1975), Vol. 10, pp.28-40.

[11] J.C.A. van der Lubbe: On certain coding theorems for the information of order α and of type β. In: Trans. Eight Prague Conf. Inform. Theory, Statist. Dee. Function, Random Processes, Vol. C. Academia, Prague (1978), pp.253-266.

[12] L.A. Zadeh: Fuzzy sets, Information and control, (1965) Vol. 8, pp, 338-353.