American. Jr. of Mathematics and Sciences
Vol. 1, No.1 ,(January 2012)

www.journalshub.com

APPLICATIONS OF MEASURES OF FUZZY ENTROPY TO CODING THEORY

R.K.Tuli
Department of Mathematics
S.S.M. College
Dinanagar (India)

Chander Sekhar Sharma
Department of Mathematics
Adesh Institute of Engg. & Tech.
Faridkot (India)

ABSTRACT

In the present communication, we have proved some new coding theorems and consequently, developed some new weighted fuzzy mean codeword lengths corresponding to the well-known measures of weighted fuzzy entropy. A desirable property, that is, monotonicity of the newly developed weighted fuzzy mean codeword lengths has also been studied.

Keywords: Fuzzy sets, Codeword, Code, Uniquely decipherable code, Efficiency, Uncertainty.

INTRODUCTION

The measure of uncertainty introduced by Shannon [11] has tremendous applications in different disciplines. One of its applications is the problem of efficient coding of messages to be sent over a noiseless channel, that is, our only concern is to maximize the number of messages that can be sent over the channel in a given time. Let us assume that the messages to be transmitted are generated by a random variable X and each value x_i , $i = 1, 2, ..., n$ of X must be represented by a finite sequence of symbols chosen from the set $\{a_1, a_2, ..., a_D\}$. Let n_i be the length of code word associated with x_i satisfying Kraft's [7] inequality

$$\sum_{i=1}^{n} D^{-n_i} \le 1 \tag{1.1}$$

where D is the size of alphabet. In calculating the long run efficiency of communications, we choose codes to minimize average code word length, given by

$$L = \sum_{i=1}^{n} p_i n_i \tag{1.2}$$

where p_i is the probability of occurrence of x_i . For uniquely decipherable codes, Shannon's noiseless coding theorem which states that

$$\frac{H(P)}{\log D} \le L < \frac{H(P)}{\log D} + 1 \tag{1.3}$$

determines the lower and upper bounds on L in terms of Shannon's entropy. Campbell [2] for the first time introduced the idea of exponentiated mean code word length for uniquely decodable codes and proved a noiseless coding theorem. He considered an exponentiated mean of order α defined by

$$L_\alpha = \frac{\alpha}{1-\alpha} \log_D \left[\sum_{i=1}^{n} p_i \, D^{(1-\alpha) n_i / \alpha} \right] \tag{1.4}$$

and showed that its lower bound lies between $R_\alpha(P)$ and $R_\alpha(P) + 1$ where

$$R_\alpha(P) = (1-\alpha)^{-1} \log_D \left[\sum_{i=1}^{n} p_i^{\alpha} \right]; \alpha > 0 , \alpha \ne 1 \tag{1.5}$$

Guiasu and Picard [3] defined the weighted average length for a uniquely decipherable code as

$$\bar{L} = \sum_{i=1}^{n} \left(\frac{u_i \; n_i \; p_i}{\sum_{i=1}^{n} u_i \; p_i} \right) \tag{1.6}$$

Longo [8] interpreted (1.6) as the average cost of transmitting letters x_i with probability p_i and utility u_i and gave some practical interpretation of this length. Lower and upper bounds for the cost function (1.6) in terms of weighted entropy have also been derived.

Longo [8] gave lower bound for useful mean codeword length in terms of quantitative-qualitative measure of entropy introduced by Belis and Guiasu [1]. Guiasu and Picard [3] proved a noiseless coding theorem by obtaining lower bounds for similar useful mean codeword length. Gurdial and Pessoa [4] tried to extend the theorem by finding lower bounds for useful mean codeword lengths of order α in terms of useful measures of information of order α. Some other pioneer who extended their results towards the development of coding theory are Korada and Urbanke [6], Szpankowski [12], Merhav [9] etc. Recently, Kapur [5] has established relationships between probabilistic entropy and coding. But there are many situations where probabilistic measures of entropy do not work and to tackle such situations, instead of taking the idea of probability, the idea of fuzziness can be explored.

In the next section, we have considered the fuzzy distributions and developed some new fuzzy codeword lengths by proving noiseless coding theorems:

2. A CLASS OF FUZZY CODING THEOREMS AND CODEWORD LENGTHS

Theorem 2.1 For all uniquely decipherable codes, we have the following inequality:

$$H_\alpha(A;W) \le L_\alpha(W) \;\; ; \alpha > 1 \tag{2.1}$$

where $H_\alpha(A;W) = \dfrac{1}{\alpha(1-\alpha)} \sum_{i=1}^{n} w_i \left[\mu_A^\alpha(x_i) + (1-\mu_A(x_i))^\alpha - 1 \right]$ is a measure of weighted fuzzy entropy introduced by Parkash and Tuli [10] and

$$L_\alpha(W) = \frac{\sum_{i=1}^{n} w_i \left(\mu_A^{\alpha}(x_i) + \left(1-\mu_A(x_i)\right)^\alpha - 1 \right) D^{n_i\left(\frac{1-\alpha}{\alpha}\right)}}{\alpha(1-\alpha)} \tag{2.2}$$

is a new weighted parametric codeword length.

Proof: By Holder's inequality, we have

$$\sum_{i=1}^{n} x_i \, y_i \ge \left(\sum_{i=1}^{n} x_i^{\,p} \right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} y_i^{\,q} \right)^{\frac{1}{q}}; \; 0 < p < 1, q < 0 \text{ or } 0 < q < 1, p < 0 \tag{2.3}$$

Setting $x_i = \left[f(\mu_A(x_i)) \right]^{-\frac{1}{t}} D^{-n_i}$, $y_i = \left[f(\mu_A(x_i)) \right]^{\frac{1}{t}}$ and $p = -t, q = \dfrac{t}{1+t}$, the inequality (2.3) becomes

$$\left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \right\}^{\frac{1}{1+t}} \right]^{1+t} \le \left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \, D^{n_i t} \right\} \right]$$

Now $\left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \right\} \right] \le \left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \right\}^{\frac{1}{1+t}} \right]^{1+t} \le \left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \, D^{n_i t} \right\} \right]$

Thus, we have

$$\left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \right\} \right] \le \left[\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \, D^{n_i t} \right\} \right]$$

Taking $\alpha = \frac{1}{1+t}$, $\alpha > 1$, $t = \frac{1-\alpha}{\alpha}$, the above equation becomes

$$\sum_{i=1}^{n} \left\{ f(\mu_A(x_i)) \right\} \geq \left[\sum_{i=1}^{n} f(\mu_A(x_i)) D^{n_i \frac{1-\alpha}{\alpha}} \right] \tag{2.4}$$

Again taking

$f(\mu_A(x_i)) = \left[w_i \left\{ \mu_A^{\alpha}(x_i) + (1-\mu_A(x_i))^{\alpha} - 1 \right\} \right]$, and multiplying the denominator by $\alpha(1-\alpha), \alpha > 1$, the inequality (2.4) proves the theorem.

Theorem 2.2 For all uniquely decipherable codes, the following relation holds:

$$H^{\alpha}(A;W) \leq L^{\alpha}(W) \;; \alpha > 1 \tag{2.5}$$

with equality iff $\mu_A(x_i) = \dfrac{D^{-n_i}}{\sum_{j=1}^{n} D^{-n_j}}$

where $$L^{\alpha}(W) = \sum_{i=1}^{n} \log \left[D^{n_i w_i \mu_A^{\alpha}(x_i)} \left\{ 1 - \frac{D^{-n_i}}{\sum_{j=1}^{n} D^{-n_j}} \right\}^{-w_i (1-\mu_A(x_i))^{\alpha}} \right] \tag{2.6}$$

is the new weighted fuzzy codeword length and

$$H^{\alpha}(A;W) = -\sum_{i=1}^{n} w_i \left[\mu_A^{\alpha}(x_i) \log \mu_A(x_i) + (1-\mu_A(x_i))^{\alpha} \log(1-\mu_A(x_i)) \right]; \alpha > 1 \tag{2.7}$$

is a new measure of weighted fuzzy entropy.

Proof. We have introduced the following generalized weighted measure of fuzzy divergence:

$$D_{\alpha}(A,B;W) = \sum_{i=1}^{n} w_i \left[\mu_A^{\alpha}(x_i) \log \frac{\mu_A(x_i)}{\mu_B(x_i)} + (1-\mu_A(x_i))^{\alpha} \log \frac{(1-\mu_A(x_i))}{(1-\mu_B(x_i))} \right] \tag{2.8}$$

Now, since equation (2.8) represents a measure of distance, it must satisfy the following inequality:

$$\sum_{i=1}^{n} w_i \left[\mu_A^{\alpha}(x_i) \log \frac{\mu_A(x_i)}{\mu_B(x_i)} + (1-\mu_A(x_i))^{\alpha} \log \frac{(1-\mu_A(x_i))}{(1-\mu_B(x_i))} \right] \geq 0 \tag{2.9}$$

Taking $\mu_B(x_i) = \dfrac{D^{-n_i}}{\sum_{j=1}^{n} D^{-n_j}}$, $1 \leq i \leq n$, inequality (2.9) becomes

$$H^{\alpha}(A,W) \leq -\sum_{i=1}^{n} w_i \mu_A^{\alpha}(x_i) \left[\log D^{-n_i} - \log\left(\sum_{j=1}^{n} D^{-n_j} \right) \right] - \sum_{i=1}^{n} w_i (1-\mu_A(x_i))^{\alpha} \log \left(1 - \frac{D^{-n_i}}{\sum_{j=1}^{n} D^{-n_i}} \right) \tag{2.10}$$

where $H^{\alpha}(A;W) = -\sum_{i=1}^{n} w_i \left[\mu_A^{\alpha}(x_i) \log \mu_A(x_i) + (1-\mu_A(x_i))^{\alpha} \log(1-\mu_A(x_i)) \right]; \alpha > 1$

Using Kraft's (1949) inequality, inequality (2.10) proves the theorem.

In the next section, we have studied the monotonicity of the newly constructed codeword lengths.

3. MONOTONICITY OF THE CODEWORD LENGTHS

Here, we discuss the monotonicity of the following codeword lengths:

I. Monotonicity of $L^{\alpha}(W)$

We have developed the codeword length which can be rewritten as

$$L^{\alpha}(W)=\sum_{i=1}^{n} n_i\ w_i\mu_A^{\alpha}(x_i)\log D-\sum_{i=1}^{n} w_i(1-\mu_{\mathbf{A}}(x_i))^{\alpha}\log\left\{1-\frac{D^{-n_i}}{\sum_{j=1}^{n}D^{-n_j}}\right\} \qquad (3.1)$$

Differentiating equation (3.1) w.r.t. α , we get

$$\frac{\partial}{\partial\alpha}L^{\alpha}(W)=\sum_{i=1}^{n} n_i\ w_i\log D\mu_A^{\alpha}(x_i)\log\mu_{\mathbf{A}}(x_i)+\sum_{i=1}^{n} w_i(1-\mu_{\mathbf{A}}(x_i))^{\alpha}\log(1-\mu_{\mathbf{A}}(x_i))\log\left\{1-\frac{D^{-n_i}}{\sum_{j=1}^{n}D^{-n_j}}\right\} \qquad (3.2)$$

Clearly, the R.H.S. of (3.2) consists of two terms and both terms are –ve. Hence, we must have $\frac{\partial}{\partial\alpha}L^{\alpha}(W)<0$ which shows that $L^{\alpha}(W)$ is a monotonically decreasing function of α .

Next, with the help of the data, we have presented the weighted codeword length $L^{\alpha}(W)$ graphically. For this purpose, we have computed different values of $L^{\alpha}(W)$ for different values of the parameter α , corresponding to different fuzzy values $\mu_A(x_i)$ under the weighted distribution $W=\{2,3,4,5\}$. Next, we have presented $L^{\alpha}(W)$ graphically and obtained the Fig.-3.1 which clearly shows that the codeword length $L^{\alpha}(W)$ is monotonically decreasing function of α .

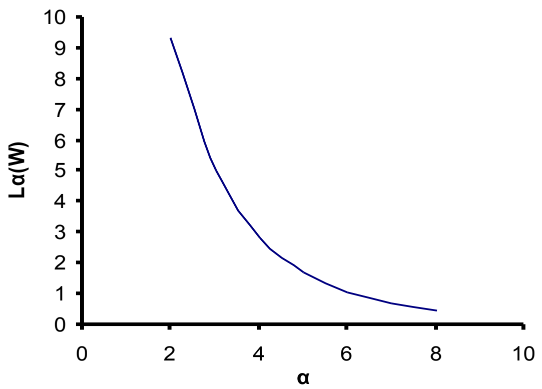

Fig.-3.1

II. Monotonicity of $L_{\alpha}(W)$

We have developed the weighted codeword length in equation (2.2) which gives

$$\alpha^2(1-\alpha)^2\frac{dL_{\alpha}(W)}{d\alpha}$$

$$=\alpha\left[\sum_{i=1}^{n}w_i\left\{\mu_A^{\alpha}(x_i)\log\mu_A^{1-\alpha}(x_i)+\left(1-\mu_A(x_i)\right)^{\alpha}\log\left(1-\mu_A(x_i)\right)^{1-\alpha}+\mu_A^{\alpha}(x_i)+\left(1-\mu_A(x_i)\right)^{\alpha}-1\right\}\right]$$

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

$$-\alpha(1-\alpha)^2\frac{\sum_{i=1}^{n} w_i\left\{\mu_A^{\alpha}(x_i)+\left(1-\mu_A(x_i)\right)^{\alpha}-1\right\}D^{n_i\left(\frac{1-\alpha}{\alpha}\right)}}{\alpha(1-\alpha)} \quad (3.3)$$

The equation (3.3) consists of three terms. Obviously, the IInd and IIIrd terms are negative. Now, we discuss the sign of Ist term:

Ist term can be written as $\alpha\left[\sum_{i=1}^{n} w_i F(x)\right]$

where $F(x)=x(x^{\alpha-1}-x^{\alpha-1}\log x^{\alpha-1})+(1-x)\left[(1-x)^{\alpha-1}-(1-x)^{\alpha-1}\log(1-x)^{\alpha-1}\right]-1\leq 0$

as $u - u\log u \leq 1$

Thus, we have $F(\mu_A(x_i))\leq 0$

Since, $w_i \geq 0$ and $\alpha > 1$, we see that the first term is ≤ 0

Thus, equation (3.3) gives $\frac{dL_\alpha(W)}{d\alpha}\leq 0, \alpha > 1$.

So $L_\alpha(W)$ is monotonically decreasing function of α .

Next, we have presented $L_\alpha(W)$ graphically and obtained the Fig.-3.2 which clearly shows that the codeword length $L_\alpha(W)$ is monotonically decreasing function of α .

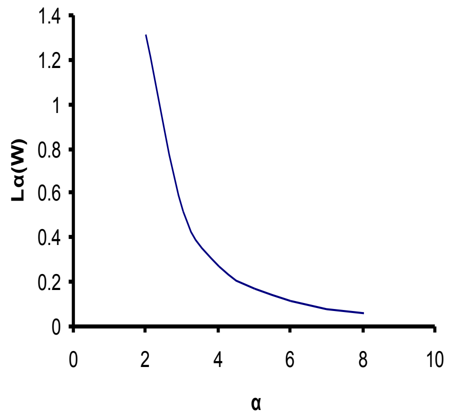

Fig.-3.2

REFERENCES

[1] Belis, M. and Guiasu, S. (1968). A quantitative-qualitative measure of information in cybernetic systems. IEEE Transactions on Information Theory **14**: 593-594.

[2] Campbell, L.L. (1965). A Coding theorem and Renyi's entropy. Information and control, **8:** 423-429.

[3] Guiasu, S. and Picard, C.F. (1971). Borne in ferictur de la longuerur utile de certains codes. Comptes Rendus Mathematique Academic des Sciences Paris **273**: 248-251.

[4] Gurdial and Pessoa, F. (1977). On useful information of order α. Journal of Combanatrics Information and System Sciences **2**: 158-162.

[5] Kapur, J.N. (1998). Entropy and Coding. Mathematical Sciences Trust Society, New Delhi.

[6] Korada, S. B. and Urbanke, R. L. (2010). Polar codes are optimal for lossy source coding. IEEE Trans. Inform. Theory **56**(4): 1751–1768.

[7] Kraft, L.G. (1949): "A Device for Quantizing Grouping and Coding Amplitude Modulated Pulses", M.S. Thesis, Electrical Engineering Department, MIT.

[8] Longo, G. (1972). Quantitative-Qualitative Measures of Information. Springer-Verlag, New York.

[9] Merhav, N. (2008). Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels. IEEE Transactions on Information Theory **54**: 2723-2734.

[10] Parkash, O. and Tuli, R.K. (2007). On generalized Havrada-Charvat fuzzy entropy. Quality, Reliability and Infocom Tech. (Ed.P.K.Kapur &A.K.Verma), MacMillan India Ltd., 509-513.

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

[12] Szpankowski, W. (2008). A one-to-one code and its anti-redundancy. IEEE Trans. Inform. Theory **54**(10): 4762–4766.