journal of inequalities in pure and applied mathematics
http://jipam.vu.edu.au
issn: 1443-5756
Volume 9 (2008), Issue 1, Article 8, 6 pp.

CODING THEOREMS ON GENERALIZED COST MEASURE

RAYEES AHMAD DAR AND M.A.K BAIG

DEPARTMENT OF STATISTICS
UNIVERSITY OF KASHMIR
SRINAGAR-190006, INDIA.
rayees_stats@yahoo.com

Received 6 June, 2007; accepted 13 December, 2007
Communicated by N.S. Barnett

ABSTRACT. In the present communication a generalized cost measure of utilities and lengths of output codewords from a memoryless source are defined. Lower and upper bounds in terms of a non-additive generalized measure of 'useful' information are obtained.

Key words and phrases: Decipherable code, Source alphabet, Codeword, Cost function, Hölders inequality, Codeword length.

2000 *Mathematics Subject Classification.* 94A17, 94A24.

1. INTRODUCTION

Let a finite set of n source symbols $X = (x_1, x_2, \dots, x_n)$ with probabilities $P = (p_1, p_2, \dots, p_n)$ be encoded using D $(D \geq 2)$ code alphabets, then there is a uniquely decipherable/instantaneous code with lengths $l_1, l_2, \dots, l_n$ if and only if

$$\sum_{i=1}^{n} D^{-l_i} \leq 1. \tag{1.1}$$

(1.1) is known as the Kraft inequality [5]. If $L = \sum_{i=1}^{n} l_i p_i$ is the average code word length, then for a code which satisfies (1.1), Shannon's coding theorem for a noiseless channel (Feinstein [5]) gives a lower bound of L in terms of Shannon's entropy [13]

$$L \geq H(P) \tag{1.2}$$

with equality iff $l_i = -\log p_i \;\; \forall \;\; i = 1, 2, \dots, n$. All the logarithms are to base D.

Belis and Guiasu [2] observed that a source is not completely specified by the probability distribution P over the source symbols X, in the absence of its qualitative character. It can also be assumed that the source letters or symbols are assigned weights according to their importance or utilities in the view of the experimenter.

366-07

Let $U = (u_1, u_2, \ldots, u_n)$ be the set of positive real numbers, where u_i is the utility or importance of outcome x_i. The utitlity, in general, is independent of the probability of encoding of source symbol x_i, i.e., p_i. The information source is thus given by

$$S = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \\ u_1 & u_2 & \cdots & u_n \end{bmatrix} \tag{1.3}$$

where $u_i > 0,\ p_i \geq 0,\ \sum_{i=1}^{n} p_i = 1$.

Belis and Guiasu [2] introduced the following qualitative -quantitative measure of information

$$H(P;U) = -\sum_{i=1}^{n} u_i p_i \log p_i \tag{1.4}$$

which is a measure of the average quantity of 'valuable' or 'useful' information provided by the information source (1.3). Guiasu and Picard [6] considered the problem of encoding the letter output of the source (1.3) by means of a single letter prefix code whose code words $w_1, w_2, \ldots, w_n$ are of lengths $l_1, l_2, \ldots, l_n$ respectively satisfying the Kraft inequality (1.1). They introduced the following 'useful' mean length of the code

$$L(P;U) = \frac{\sum_{i=1}^{n} u_i p_i l_i}{\sum_{i=1}^{n} u_i p_i}. \tag{1.5}$$

Longo [11] interpreted (1.5) as the average transmission cost of the letter x_i and obtained the following lower bound for the cost measure (1.5) as

$$L(P;U) \geq H(P;U), \tag{1.6}$$

where

$$H(P;U) = -\frac{\sum_{i=1}^{n} u_i p_i \log p_i}{\sum_{i=1}^{n} u_i p_i}$$

is the 'useful' information measure due to Guiasu and Picard [6], which was also characterized by Bhaker and Hooda [3] by a mean value representation.

2. GENERALIZED MEASURES OF COST

In the derivation of the cost measure (1.5) it is assumed that the cost is a linear function of code length, but this is not always the case. There are occasions when the cost behaves like an exponential function of code word lengths. Such types occur frequently in market equilibrium and growth models in economics. Thus sometimes it might be more appropriate to choose a code which minimizes a monotonic function,

$$C = \sum_{i=1}^{n} u_i^{\beta} p_i^{\beta} D^{\frac{1-\alpha}{\alpha} l_i}, \tag{2.1}$$

where $\alpha > 0\ (\neq 1)\,, \beta > 0$ are the parameters related to the cost.

In order to make the result of the paper more comparable with the usual noiseless coding theorem, instead of minimizing (2.1) we minimize

$$L_{\alpha}^{\beta}(U) = \frac{1}{2^{1-\alpha} - 1}\left[\left(\frac{\sum_{i=1}^{n} (u_i p_i)^{\beta} D^{\frac{1-\alpha}{\alpha} l_i}}{\sum_{i=1}^{n} (u_i p_i)^{\beta}}\right)^{\alpha} - 1\right], \tag{2.2}$$

where $\alpha > 0\ (\neq 1), \beta > 0$, which is a monotonic function of C. We define (2.2) as the 'useful' average code length of order α and type β.

Clearly, if $\alpha \to 1, \beta = 1$, (2.2) reduces to (1.5) which further reduces to the ordinary mean length given by Shannon [13] when $u_i = 1\ \forall\ i = 1, 2, \ldots, n$. We also note that (2.2) is a monotonic non-decreasing function of α and if all the l_i 's are the same, say $l_i = l$ for each i and $\alpha \to 1$, then $L_\alpha^\beta(U) = l$. This is an important property for any measure of length to possess.

In the next section, we derive the lower and upper bounds of the cost function (2.2) in terms of the following 'useful' information measure of order α and type β,

$$H_\alpha^\beta(P;U) = \frac{1}{2^{1-\alpha}-1}\left[\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta} - 1\right], \tag{2.3}$$

where $\alpha > 0\,(\neq 1), \beta > 0,\ p_i \geq 0,\ i = 1, 2, \ldots, n,\ \sum_{i=1}^n p_i \leq 1.$

(i) When $\beta = 1$, (2.3) reduces to the measure of 'useful' information proposed and characterised by Hooda and Ram [9].
(ii) If $\alpha \to 1, \beta = 1$, (2.3) reduces to the measure given by Belis and Guiasu [2].
(iii) If $\alpha \to 1, \beta = 1$ and $u_i = 1\ \forall\ i = 1, 2, \ldots, n$, (2.3) reduces to the well known measure given by Shannon [13].

Also, we have used the condition

$$\sum_{i=1}^n u_i^\beta p_i^{\beta-1} D^{-l_i} \leq \sum_{i=1}^n u_i^\beta p_i^\beta \tag{2.4}$$

to find the bounds. It may be seen that in the case when $\beta = 1$, $u_i = 1\ \forall\, i = 1, 2, \ldots, n$, (2.4) reduces to the Kraft inequality (1.1). $D\ (D \geq 2)$ is the size of the code alphabet.

Longo [12], Gurdial and Pessoa [7], Autar and Khan [1], Jain and Tuteja [10], Taneja et al. [16], Bhatia [4], Singh, Kumar and Tuteja [15] and Hooda and Bhaker [8] considered the problem of a 'useful' information measure in the context of noiseless coding theorems for sources involving utilities.

In this paper, we study upper and lower bounds by considering a new function dependent on the parameter α and type β and a utility function. Our motivation for studying this new function is that it generalizes some entropy functions already existing in the literature. The function under study is closely related to Tsallis entropy which is used in Physics.

3. BOUNDS ON THE GENERALIZED COST MEASURES

Theorem 3.1. *For all integers $D\ (D \geq 2)$, let l_i satisfy (2.4), then the generalized average 'useful' codeword length satisfies*

$$L_\alpha^\beta(U) \geq H_\alpha^\beta(P;U) \tag{3.1}$$

and the equality holds iff

$$l_i = -\log p_i^\alpha + \log \frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}. \tag{3.2}$$

Proof. By Hölder's inequality [14]

$$\sum_{i=1}^n x_i y_i \geq \left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q\right)^{\frac{1}{q}} \tag{3.3}$$

for all $x_i, y_i > 0,\ i = 1, 2, \ldots, n$ and $\frac{1}{p} + \frac{1}{q} = 1$, $p < 1\,(\neq 0)$, $q < 0$,or $q < 1\,(\neq 0)$, $p < 0$.

We see that equality holds if and only if there exists a positive constant c such that

$$x_i^p = c y_i^q. \tag{3.4}$$

Making the substitutions,

$$p=\frac{\alpha-1}{\alpha},\qquad q=1-\alpha,$$

$$x_i=\frac{(u_ip_i)^{\frac{\beta\alpha}{\alpha-1}}D^{-l_i}}{\sum_{i=1}^n(u_ip_i)^{\frac{\beta\alpha}{\alpha-1}}},\qquad y_i=\frac{(u_i)^{\frac{\beta}{1-\alpha}}p_i^{\frac{\alpha+\beta-1}{1-\alpha}}}{\sum_{i=1}^n(u_ip_i)^{\frac{\beta}{1-\alpha}}}$$

in (3.3), we get

$$\frac{\sum_{i=1}^n u_i^\beta p_i^{\beta-1}D^{-l_i}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\geq\left[\frac{\sum_{i=1}^n u_i^\beta p_i^\beta D^{\frac{1-\alpha}{\alpha}l_i}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\right]^{\frac{\alpha}{\alpha-1}}\left[\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\right]^{\frac{1}{1-\alpha}}.$$

Using the condition (2.4), we get

$$\left[\frac{\sum_{i=1}^n u_i^\beta p_i^\beta D^{\frac{1-\alpha}{\alpha}l_i}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\right]^{\frac{\alpha}{1-\alpha}}\geq\left[\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\right]^{\frac{1}{1-\alpha}}.$$

Taking $0<\alpha<1$ and raising both sides to the power $(1-\alpha)$,

$$\left[\frac{\sum_{i=1}^n u_i^\beta p_i^\beta D^{\frac{1-\alpha}{\alpha}l_i}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\right]^{\alpha}\geq\left[\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\right].$$

Multiplying both sides by $\frac{1}{2^{1-\alpha}-1}>0$ for $0<\alpha<1$ and simplifying, we obtain

$$L_\alpha^\beta(U)\geq H_\alpha^\beta(P;U).$$

For $\alpha>1$, the proof follows along similar lines. □

Theorem 3.2. *For every code with lengths* $\{l_i\}$, $i=1,2,\ldots,n$ *of Theorem 3.1,* $L_\alpha^\beta(U)$ *can be made to satisfy the inequality,*

$$L_\alpha^\beta(U)<H_\alpha^\beta(P;U)D^{1-\alpha}+\frac{D^{1-\alpha}-1}{2^{1-\alpha}-1}. \tag{3.5}$$

Proof. Let l_i be the positive integer satisfying the inequality,

$$-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\leq l_i<-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}+1. \tag{3.6}$$

Consider the interval,

$$\delta_i=\left[-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta},-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}+1\right] \tag{3.7}$$

of length 1. In every δ_i, there lies exactly one positive integer l_i such that

$$0<-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\leq l_i<-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}+1. \tag{3.8}$$

We will first show that the sequence $l_1,l_2,\ldots,l_n$, thus defined satisfies (2.4). From (3.8), we have,

$$-\log p_i^\alpha+\log\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}\leq l_i,$$

$$\frac{p_i^\alpha}{\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}}\geq D^{-l_i}.$$

Multiplying both sides by $u_i^\beta p_i^{\beta-1}$ and summing over $i = 1, 2, \ldots, n,$ we get (2.4). The last inequality of (3.8) gives,

$$l_i < -\log p_i^\alpha + \log \frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta} + 1$$

or

$$D^{l_i} < \left(\frac{p_i^\alpha}{\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}} \right)^{-1} D.$$

For $0 < \alpha < 1$, raising both sides to the power $\frac{1-\alpha}{\alpha}$ we obtain,

$$D^{\frac{1-\alpha}{\alpha} l_i} < \left(\frac{p_i^\alpha}{\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n u_i^\beta p_i^\beta}} \right)^{\frac{\alpha-1}{\alpha}} D^{\frac{1-\alpha}{\alpha}}.$$

Multiplying both sides by $\frac{u_i^\beta p_i^\beta}{\sum_{i=1}^n u_i^\beta p_i^\beta}$ and summing over $i = 1, 2, \ldots, n,$ gives

$$\frac{\sum_{i=1}^n (u_i p_i)^\beta D^{\frac{1-\alpha}{\alpha} l_i}}{\sum_{i=1}^n (u_i p_i)^\beta} < \left(\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n (u_i p_i)^\beta} \right)^{\frac{1}{\alpha}} D^{\frac{1-\alpha}{\alpha}},$$

$$\left(\frac{\sum_{i=1}^n (u_i p_i)^\beta D^{\frac{1-\alpha}{\alpha} l_i}}{\sum_{i=1}^n (u_i p_i)^\beta} \right)^\alpha < \left(\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n (u_i p_i)^\beta} \right) D^{1-\alpha}.$$

Since $2^{1-\alpha} - 1 > 0$ for $0 < \alpha < 1$,after suitable operations, we obtain

$$\frac{1}{2^{1-\alpha}-1} \left[\left(\frac{\sum_{i=1}^n (u_i p_i)^\beta D^{\frac{1-\alpha}{\alpha} l_i}}{\sum_{i=1}^n (u_i p_i)^\beta} \right)^\alpha - 1 \right] < \frac{1}{2^{1-\alpha}-1} \left[\frac{\sum_{i=1}^n u_i^\beta p_i^{\alpha+\beta-1}}{\sum_{i=1}^n (u_i p_i)^\beta} - 1 \right] D^{1-\alpha} + \frac{D^{1-\alpha}-1}{2^{1-\alpha}-1}.$$

We can write

$$L_\alpha^\beta (U) < H_\alpha^\beta (P; U) D^{1-\alpha} + \frac{D^{1-\alpha}-1}{2^{1-\alpha}-1}.$$

As $D \geq 2,$ we have $\frac{D^{1-\alpha}-1}{2^{1-\alpha}-1} > 1$ from which it follows that the upper bound, $L_\alpha^\beta (U)$ in (3.5), is greater than unity.

Also, for $\alpha > 1,$ the proof follows similarly. □

REFERENCES

[1] R. AUTAR AND A.B. KHAN, On generalized 'useful' information for incomplete distribution, *J. Combinatorics, Information and System Sciences*, **14**(4) (1989), 187–191.

[2] M. BELIS AND S. GUIASU, A quantitative-qualitative measure of information in Cybernetics system, *IEEE Transactions in Information Theory*, **14** (1968), 593–594.

[3] U.S. BHAKER AND D.S. HOODA, Mean value characterization of 'useful' information measures, *Tamkang J. Math.*, **24** (1993), 283–294.

[4] P.K. BHATIA, On a generalized 'useful' inaccuracy for incomplete probability distribution, *Soochow J. Math.*, **25**(2) (1999), 131–135.

[5] A. FEINSTEIN, *Foundation of Information Theory*, Mc Graw Hill, New York (1956).

[6] S. GUIASU AND C.F. PICARD, Borne inferieure dela langueur de certian codes, *C.R Academic Sciences, Paris*, **27 C** (1971), 248–251.

[7] GURDIA AND F. PESSOA, On useful information of order α, *J. Combinatorics, Information and System Sciences*, **2** (1977), 30–35.

[8] D.S HOODA AND U.S. BHAKER, A generalized 'useful' information measure and coding theorems, *Soochow J. Math.*, **23** (1997), 53–62.

[9] D.S. HOODA AND A. RAM, Characterization of non-additive 'useful' information measure, *Recent Advances in Information Theory, Statistics and Computer Applications*, CCS Haryana Agricultural University, Hisar, 64–77 (1998).

[10] P. JAIN AND R.K. TUTEJA, On coding theorms connected with 'useful' entropy of order β, *International Journal of Mathematics and Mathematical Sciences*, **12**(1) (1989), 193–198.

[11] G. LONGO, A noiseless coding theorems for source having utilities, *SIAM J. Appl. Math.*, **30** (1976), 739-748.

[12] G. LONGO, *Quantitative-Qualitative Measure of Information*, Springer Verlag, New York (1972).

[13] C.E. SHANNON, A mathematical theory of communication, *Bell System Technical Journal*, **27** (1948), 394–423, 623–656.

[14] O. SHISHA, *Inequalities*, Academic Press, New York (1967).

[15] R.P. SINGH, R. KUMAR AND R.K. TUTEJA, Applications of Hölder's inequality in information theory, *Information Sciences*, **152** (2003), 145–154.

[16] H.C. TENAJA, D.S. HOODA AND R.K. TUTEJA, Coding theorems on a generalized 'useful' information, *Soochow J. Math.*, **11** (1985), 123–131.