Asian Journal of Fuzzy and Applied Mathematics (ISSN: 2321 – 564X)
Volume 04 – Issue 06, December 2016

Noiseless Coding Theorems on New Generalized Useful Information Measure of order α and type β

Ashiq Hussain Bhat[1,*], Mirza Abdul Khaliq Baig[2]

[1]Research Scholar (Ph.D) in Department Of Statistics University Of Kashmir Srinagar
(Srinagar, India)

[2]Associate professor in Department Of Statistics University Of Kashmir Srinagar
(Srinagar, India)

[*]*Corresponding author's email: ashiqhb14 [AT] gmail.com*

ABSTRACT— ***In this communication we define a new generalized useful average code-word length $L_\alpha^\beta(P;U)$ of order α and type β and its relationship with new generalized useful information measure $H_\alpha^\beta(P;U)$ of order α and type β has been discussed. The lower and upper bound of $L_\alpha^\beta(P;U)$, in terms of $H_\alpha^\beta(P;U)$ are derived for a discrete noiseless channel. The measures defined in this communication are not only new but some well known measures are the particular cases of our proposed measures that already exist in the literature of useful information and coding theory. The noiseless coding theorems for discrete channel proved in this paper are verified by considering Huffman and Shannon-Fano coding schemes on taking empirical data. The important properties of $H_\alpha^\beta(P;U)$ have also been studied.***

Keywords— Shannon's entropy, codeword length, useful information measure, Kraft inequality, Holder's inequality, Huffman codes, Shannon-Fano codes, Noiseless coding theorem

***AMS* Classification—** 94A17, 94A24

---

1. INTRODUCTION

The growth of telecommunication in the early twentieth century led several researchers to study the information control of signals, the seminal work of Shannon [17], based on papers by Nyquists [14], [15] and Hartley [7] rationalized these early efforts into a coherent mathematical theory of communication and initiated the area of research now known as information theory. The central paradigm of classical information theory is the engineering problem of the transmission of information over a noisy channel. The most fundamental results of this theory are Shannon's source coding theorem which establishes that on average the number of bits needed to represent the result of an uncertain event is given by its entropy; and Shannon's noisy-channel coding theorem which states that reliable communication is possible over noisy channels provided that the rate of communication is below a certain threshold called the channel capacity. Information theory is a broad and deep mathematical theory with equally broad and deep applications, amongst which is the vital field of coding theory. Information theory is a new branch of probability and statistics with extensive potential application to communication system. The term information theory does not possess a unique definition. Broadly speaking, information theory deals with the study of problems concerning any system. This includes information processing, information storage and decision making. In a narrow sense, information theory studies all theoretical problems connected with the transmission of information over communication channels. This includes the study of uncertainty (information) measure and various practical and economical methods of coding information for transmission.

Let X be a finite discrete random variable or finite source taking values $x_1, x_2, \dots, x_n$ with respective probabilities $P = (p_1, p_2, \dots, p_n)$, $p_i \geq 0 \;\forall\; i = 1, 2, \dots, n$ and $\sum_{i=1}^{n} p_i = 1$. Shannon [17] gives the following measure of information and call it entropy.

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \tag{1.1}$$

The measure (1.1) serves as a suitable measure of entropy. Let $p_1, p_2, p_{3,\dots,}p_n$ be the probabilities of n code-words to be transmitted and let their lengths $l_1, l_2, \dots, l_n$ satisfy Kraft [11] inequality,

$$\sum_{i=1}^{n} D^{-l_i} \leq 1 \tag{1.2}$$

For uniquely decipherable codes, Shannon [17] showed that for all codes satisfying (1.2), the lower bound of the mean codeword length,

$$L(P) = \sum_{i=1}^{n} p_i l_i \tag{1.3}$$

lies between $H(P)$ and $H(P) + 1$.Where D is the size of code alphabet. Shannon's entropy (1.1) is indeed a measure of uncertainty and is treated as information supplied by a probabilistic experiment. This formula gives us the measure of information as a function of the probabilities only in which various events occur without considering the effectiveness or importance of the events. Belis and Guiasu [1] remarked that a source is not completely specified by the probability distribution P over the source alphabet X in the absence of quality character. They enriched the usual description of the information source (i.e., a finite source alphabet and finite probability distribution) by introducing an additional parameter measuring the utility associated with an event according to their importance or utilities in view of the experimenter.

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 utility, in general, is independent of p_i, i.e., the probability of encoding of source symbol x_i. The information source is thus given by

$$S = \begin{bmatrix} X_1 & X_2 \ldots & X_n \\ p_1 & p_2 \ldots & p_n \\ u_1 & u_2 \ldots & u_n \end{bmatrix}, u_i > 0\ p_i \geq 0, \sum_i^n p_i = 1 \tag{1.4}$$

We call (1.4) a Utility Information Scheme. Belis and Guiasu [1] introduced the following quantitative-qualitative measure of information for this scheme.

$$H(P, U) = -\sum_{i=1}^{n} u_i p_i \log p_i \tag{1.5}$$

and call it as 'useful' entropy. The measure (1.5) can be taken as satisfactory measure for the average quantity of 'valuable' or 'useful' information provided by the information source (1.4). Guiasu and Picard [5] considered the problem of encoding the letter output by the source (1.4) by means of a single letter prefix code whose codeword's $c_1, c_2, \ldots, c_n$ have lengths $l_1, l_2, \ldots, l_n$ respectively and satisfy the Kraft's inequality (1.2), they introduced the following quantity

$$L(P; U) = \frac{\sum_{i=1}^{n} u_i p_i l_i}{\sum_{i=1}^{n} u_i p_i} \tag{1.6}$$

and call it as 'useful' mean length of the code. Further they derived a lower bound for (1.6). However, Longo [12] interpreted (1.6) as the average transmission cost of the letters x_i with probabilities p_i and utility u_i and gave some practical interpretations of this length; also bounds for the cost function (1.6) in terms of (1.5) are also derived by him.

In coding theory, usually we come across the problem of efficient coding of messages to be sent over a noiseless channel where our concern is to maximize the number of messages that can be sent through a channel in a given time. Therefore, we find the minimum value of a mean codeword length subject to a given constraint on codeword lengths. As the code-word lengths are integers, the minimum value lies between two bounds, so a noiseless coding theorem seeks to find these bounds which are in terms of some measure of entropy for a given mean and a given constraint. Shannon [17] found the lower bounds for the arithmetic mean by using his own entropy. Campbell [4] defined his own exponentiated mean and by applying Kraft's [11] inequality, found lower bounds for his mean in terms of Renyi's [16] measure of entropy. Longo [12] developed lower bound for useful mean codeword length in terms of weighted entropy introduced by Belis and Guiasu [1]. Guiasu and Picard [5] proved a noiseless coding theorem by obtaining lower bounds for another useful mean codeword length. Gurdial and Pessoa [6] extended the theorem by finding lower bounds for useful mean codeword length of order α also the various authors like Jain and Taneja [9], Taneja et al [19], Hooda and Bhaker [8], Khan et al [10], have studied generalized coding theorems by considering different generalized 'useful' information measures under the condition of uniquely decipherability.

In this paper we define a new generalized useful average code-word length $L_\alpha^\beta(P;U)$ of order α and type β and its relationship with new generalized useful information measure of order α and type β has been discussed. The lower and upper bound of $L_\alpha^\beta(P;U)$, in terms of $H_\alpha^\beta(P;U)$ have been obtained for a discrete noiseless channel in section 2. The measures defined in this communication are not only new but some well known measures are the particular cases of our proposed measures that already exist in the literature of useful information and coding theory. In section 3, the noiseless coding theorems for discrete channel proved in this paper are verified by considering Huffman and Shannon-Fano coding schemes on taking empirical data. The important properties of $H_\alpha^\beta(P;U)$ have also been studied in section 4.

2. NOISELESS CODING THEOREMS ON 'USEFUL' CODES

Define a new generalized useful information measure of order α and type β for incomplete probability distribution as:

$$H_\alpha^\beta(P;U) = \frac{\beta}{\beta-\alpha}\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}\right], \tag{2.1}$$

Where $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha, p_i \geq 0 \ \forall\ i = 1,2,\dots,n\ , \sum_{i=1}^n p_i \leq 1$

Remarks for (2.1)

I. When $\beta = 1$ and $\alpha \to 1$, (2.1) reduces to 'useful' information measure for the incomplete distribution due to Bhakar and Hooda [2]. i.e.,

$$H(P,U) = -\frac{\sum_{i=1}^n u_i p_i \log p_i}{\sum_{i=1}^n u_i p_i} \tag{2.2}$$

II. When $\beta = 1, u_i = 1,\ \forall\ i = 1,2,\dots,n,$ i.e., when the utility aspect is ignored, $\sum_{i=1}^n p_i = 1$ and $\alpha \to 1$, the measure (2.1) reduces Shannon's [17] entropy. i.e.,

$$H(P) = -\sum_{i=1}^{n} p_i log p_i \tag{2.3}$$

III. When $\beta = 1,\ u_i = 1,\ \forall\ i = 1,2, \dots, n,$ i.e., when the utility aspect is ignored, $\sum_{i=1}^{n} p_i = 1$, $\alpha \to 1,$ and $p_i = \frac{1}{n} \forall\ i = 1,2, \dots, n$, the measure (2.1) reduces to maximum entropy. i.e.,

$$H\left(\frac{1}{n}\right) = log_D\, n \tag{2.4}$$

IV. When $\alpha \to 1,$ the measure (2.1) reduces to useful information measure for the incomplete power distribution p^{β} due to Sharma, Man and Mitter [18]. i.e.,

$$H^{\beta}(P;U) = -\frac{\beta \sum_{i=1}^{n} u_i p_i^{\beta} log p_i^{\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \tag{2.5}$$

V. When $\alpha \to 1, u_i = 1,\ \forall\ i = 1,2, \dots, n,$ i.e., when the utility aspect is ignored, (2.1) reduces to a measure of incomplete power probability distribution due to Mitter and Mathur [13]. i.e.,

$$H^{\beta}(P) = -\frac{\beta \sum_{i=1}^{n} p_i^{\beta} log p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}} \tag{2.6}$$

Further we define a new generalized useful average code-word length of order α and type β corresponding to (2.1) and is given by

$$L_{\alpha}^{\beta}(P;U) = \frac{\beta}{\beta-\alpha}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\alpha}, \tag{2.7}$$

Where $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha, p_i \geq 0\ \forall\ i = 1,2, \dots, n\ , \sum_{i=1}^{n} p_i \leq 1$ and *D* is the size of code alphabet.

Remarks for (2.7)

I. When $\beta = 1$ and $\alpha \to 1,$ (2.7) reduces to 'useful' codeword length due to Guiasu and Picard [5]. i.e.,

$$L(P;U) = \frac{\sum_{i=1}^{n} u_i p_i l_i}{\sum_{i=1}^{n} u_i p_i} \tag{2.8}$$

II. When $\beta = 1, u_i = 1,\ \forall\ i = 1,2, \dots, n,$ i.e., when the utility aspect is ignored, $\sum_{i=1}^{n} p_i = 1$ and $\alpha \to 1,$ (2.7) reduces to optimal codeword length defined by Shannon [17]. i.e.,

$$L(P) = \sum_{i=1}^{n} p_i l_i \tag{2.9}$$

III. When $\beta = 1,\ u_i = 1,\ \forall\ i = 1,2,\dots,n,$ i.e., when the utility aspect is ignored, $\sum_{i=1}^{n} p_i = 1\ , \alpha \to 1,$ and $l_1 = l_2 = \cdots = l_n = 1,$ then (2.7) reduces to 1.

Now we derive the lower and upper bound of (2.7) in terms of (2.1) under the condition

$$\frac{\sum_{i=1}^{n} u_i D^{-l_i}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \leq 1. \tag{2.10}$$

This is generalization of Kraft's inequality (1.2). It is easy to see that when $\beta = 1, u_i = 1,\ \forall\ i = 1,2,\dots,n,$ i.e., when the utility aspect is ignored and $\sum_{i=1}^{n} p_i = 1,$ then the inequality (2.10) reduces to Kraft's [11] inequality (1.2). A code satisfying (2.10) would be termed as a 'useful' personal probability code.

Theorem 2.1: Let $\{u_i\}_{i=1}^{n}$, $\{p_i\}_{i=1}^{n}$ and $\{l_i\}_{i=1}^{n}$, satisfies the inequality (2.10), then the two parametric generalized 'useful' code-word length (2.7) satisfies the inequality

$$L_{\alpha}^{\beta}(P;U) \geq H_{\alpha}^{\beta}(P;U), 0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha \tag{2.11}$$

is fulfilled .Where $H_{\alpha}^{\beta}(P;U)$ and $L_{\alpha}^{\beta}(P;U)$ are defined in (2.1) and (2.7) respectively. Furthermore, equality holds good iff

$$l_i = -\log_D \left[\frac{\frac{p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right] \tag{2.12}$$

Proof: By Holder's inequality, we have

$$\left(\sum_{i=1}^{n} x_i^{p}\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} y_i^{q}\right)^{\frac{1}{q}} \leq \sum_{i=1}^{n} x_i y_i \tag{2.13}$$

For all $x_i, y_i > 0,\ i = 1,2,3,\dots,n$ and $\frac{1}{p} + \frac{1}{q} = 1, p < 1(\neq 0), q < 0$ or $q < 1(\neq 0), p < 0$. We see the equality holds iff there exists a positive constant c such that

$$x_i^{p} = c y_i^{q} \tag{2.14}$$

Making the substitution

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

$$y_i = \frac{u_i^{\frac{1}{1-\alpha}} p_i^{\frac{\alpha\beta}{1-\alpha}}}{\left(\sum_{i=1}^{n} u_i p_i^{\beta}\right)^{\frac{1}{1-\alpha}}} \qquad \text{and} \qquad q = 1-\alpha.$$

Using these values in (2.13), and after suitable simplification we get

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{\alpha}{\alpha-1}} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{1}{1-\alpha}} \leq \frac{\sum_{i=1}^{n} u_i D^{-l_i}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \tag{2.15}$$

Now using the inequality (2.10), we get

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{\alpha}{\alpha-1}} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{1}{1-\alpha}} \leq 1 \tag{2.16}$$

Or equation (2.16), can be written as

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{\alpha}{\alpha-1}} \leq \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{1}{\alpha-1}} \tag{2.17}$$

As $0 < \alpha < 1$, then $(\alpha - 1) < 0$, raising both sides to the power $(\alpha - 1) < 0$, to equation (2.17), we get

$$\left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\alpha} \geq \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right] \tag{2.18}$$

As $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha$ then $(\beta - \alpha) > 0$ and $\frac{\beta}{\beta-\alpha} > 0$, multiply equation (2.18) both sides by $\frac{\beta}{\beta-\alpha} > 0$, we get

$$\frac{\beta}{\beta-\alpha}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\alpha} \geq \frac{\beta}{\beta-\alpha}\left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right] \tag{2.19}$$

This implies

$L_{\alpha}^{\beta}(P;U) \geq H_{\alpha}^{\beta}(P;U)$. Hence the result for $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha.$

Now we will show that the equality in (2.11) holds if and only if

$$l_i = -\log_D\left[\frac{p_i^{\alpha\beta}}{\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}}\right], 0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha.$$

Or equivalently, we can write

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

Or we can write

$$D^{-l_i} = p_i^{\alpha\beta} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right]^{-1} \tag{2.20}$$

Raising both sides to the power $\frac{\alpha-1}{\alpha}$, to equation (2.20), and after simplification we get

$$D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)} = p_i^{\beta(\alpha-1)} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right]^{\frac{1-\alpha}{\alpha}} \tag{2.21}$$

Multiply equation (2.21) both sides by $\frac{u_i p_i^{\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}$, and then summing over $i = 1,2,\dots,n,$ both sides to the resulted expression and after suitable simplification, we get

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

Or equivalently,

$$\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}} = \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right]^{\frac{1}{\alpha}} \tag{2.22}$$

Raising both sides to the power α to equation (2.22), then multiply both sides by $\frac{\beta}{\beta-\alpha}$, we get

$$\frac{\beta}{\beta-\alpha} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right]^{\alpha} = \frac{\beta}{\beta-\alpha} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right] \tag{2.23}$$

This implies

$L_{\alpha}^{\beta}(P;U) = H_{\alpha}^{\beta}(P;U)$. Hence the result

***Theorem 2.2*:** For every code with lengths $l_1, l_2, \dots, l_n$ satisfies the condition (2.10), $L_{\beta}^{\alpha}(P;U)$ can be made to satisfy the inequality

$$L_\beta^\alpha(P;U) < H_\alpha^\beta(P;U)D^{(1-\alpha)}, \text{where } 0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha. \quad (2.24)$$

Proof: From the theorem (2.1) we have

$L_\alpha^\beta = H_\alpha^\beta(P)$, holds if and only if

$$l_i = -\log_D\left[\frac{p_i^{\alpha\beta}}{\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}}\right], 0 < \alpha < 1, 0 < \beta \leq 1.$$

Or equivalently we can write

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

Now we choose the code-word lengths $l_1, l_2, \ldots, l_n$, in such a way that they satisfy the inequality,

$$-\log_D p_i^{\alpha\beta} + \log_D\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}\right] \leq l_i < -\log_D p_i^{\alpha\beta} + \log_D\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}\right] + 1$$

Consider the interval

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

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

$$0 < -\log_D p_i^{\alpha\beta} + \log_D\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}\right] \leq l_i < -\log_D p_i^{\alpha\beta} + \log_D\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}\right] + 1 \quad (2.25)$$

Now we will first show that the sequence $l_1, l_2, \ldots, l_n$, thus defined satisfies the inequality (2.10), which is generalization of Kraft inequality.

From the left inequality of (2.25), we have

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

Or equivalently, we can write

$$D^{-l_i} \leq \left[\frac{p_i^{\alpha\beta}}{\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}} \right] \tag{2.26}$$

Multiply equation (2.26) both sides by u_i, then summing over $i = 1,2,\dots,n$, both sides to the resulted expression, and after suitable operations, we get the required result (2.10), i.e.,

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

Now the last inequality of (2.25), gives

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

Or equivalently, we can write

$$D^{l_i} < p_i^{-\alpha\beta} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right] D \tag{2.27}$$

As $0 < \alpha < 1$, then $(1-\alpha) > 0$, and $\left(\frac{1-\alpha}{\alpha}\right) > 0$, raising both sides to the power $\left(\frac{1-\alpha}{\alpha}\right) > 0$, to equation (2.27), and after suitable operations, we get

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

Or equivalently, we can write

$$D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)} < p_i^{\beta(\alpha-1)} \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}} \right]^{\frac{1-\alpha}{\alpha}} D^{\frac{1-\alpha}{\alpha}} \tag{2.28}$$

Multiply equation (2.28) both sides by $\frac{u_i p_i^{\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}$, and then summing over $i = 1,2,\dots,n$, both sides to the resulted expression and after suitable simplification, we get

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

Or equivalently, we can write

$$\frac{\sum_{i=1}^{n} u_i p_i^{\beta} D^{-l_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{i=1}^{n} u_i p_i^{\beta}} < \left[\frac{\sum_{i=1}^{n} u_i p_i^{\alpha\beta}}{\sum_{i=1}^{n} u_i p_i^{\beta}}\right]^{\frac{1}{\alpha}} D^{\frac{1-\alpha}{\alpha}} \tag{2.29}$$

As $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha$ then, $(\beta - \alpha) > 0$ and $\frac{\beta}{\beta-\alpha} > 0$, raising both sides to the power α to equation (2.29), then multiply the resulted expression both sides by $\frac{\beta}{\beta-\alpha} > 0$, we get

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

This implies

$L_{\alpha}^{\beta}(P;U) < H_{\alpha}^{\beta}(P;U)D^{1-\alpha}$. Hence the result for $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha$

Thus from above two coding theorems, we have shown that

$H_{\alpha}^{\beta}(P;U) \leq L_{\alpha}^{\beta}(P;U) < H_{\alpha}^{\beta}(P;U)D^{1-\alpha}$ .where $0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha.$

In the next section we verify the noiseless coding theorems by considering Shannon-Fano coding scheme and Huffman coding scheme on taking an empirical data.

3. ILLUSTRATION

In this section we illustrate the veracity of the theorems 2.1 and 2.2 by taking empirical data as given in table (3.1) and (3.2) on the lines of Ashiq and Baig [3].

Using Huffman coding scheme the values of $H_{\alpha}^{\beta}(P;U)$, $H_{\alpha}^{\beta}(P;U)D^{1-\alpha}$, $L_{\alpha}^{\beta}(P;U)$ and η for different values of α and β are shown in the following table:

Table (3.1): Here *D*=2 in this case, as we use here binary code

Probabilities P_i	***Huffman Code words***	l_i	u_i	α	β	$H_{\alpha}^{\beta}(P;U)$	$L_{\alpha}^{\beta}(P;U)$	$\eta = \frac{H_{\alpha}^{\beta}(P;U)}{L_{\alpha}^{\beta}(P;U)} \times 100$	$H_{\alpha}^{\beta}(P;U)D^{(1-\alpha)}$
0.3846	0	1	1	0.9	1	11.894	12.032	98.853%	12.748
0.1795	100	3	2	0.8	0.9	12.338	13.117	94.061%	14.173
0.1538	101	3	3	0.5	1	4.820	5.269	91.478%	6.817
0.1538	110	3	3						
01282	111	3	4						

Now using Shannon-Fano coding scheme the values of $H_\alpha^\beta(P;U)$, $H_\alpha^\beta(P;U)D^{1-\alpha}$, $L_\alpha^\beta(P;U)$ and η for different values of α and β are shown in the following table:

Table (3.2): Here *D*=2 in this case, as we use here binary code

Probabilities P_i	***Shannon Fano Code words***	l_i	u_i	α	β	$H_\alpha^\beta(P;U)$	$L_\alpha^\beta(P;U)$	$\eta = \frac{H_\alpha^\beta(P;U)}{L_\alpha^\beta(P;U)} \times 100$	$H_\alpha^\beta(P;U)D^{(1-\alpha)}$
0.3846	0	1	1	0.9	1	11.894	12.294	96.746%	12.748
0.1795	100	2	2	0.8	0.9	12.338	13.779	89.542%	14.173
0.1538	101	3	3	0.5	1	4.820	6.280	76.751%	6.817
0.1538	110	4	3						
0.1282	111	4	4						

From table (3.1) and (3.2) we infer the following:

I. Theorems 2.1 and 2.2 hold both the cases of Shannon-Fano codes and Huffman codes. i.e.

$$H_\alpha^\beta(P;U) \le L_\alpha^\beta(P;U) < H_\alpha^\beta(P;U)D^{1-\alpha} \text{ .where } 0<\alpha<1, 0<\beta\le 1, \beta>\alpha.$$

II. Huffman mean code-word length is less than Shannon-Fano mean code-word length.

III. Coefficient of efficiency of Huffman codes is greater than coefficient of efficiency of Shannon-Fano codes i.e. it is concluded that Huffman coding scheme is more efficient than Shannon-Fano coding scheme.

4. PROPERTIES OF NEW GENERALIZED 'USEFUL' INFORMATION MEASURE OF ORDER α AND TYPE β $H_\alpha^\beta(P;U)$

In this section we will discuss some properties of new generalized 'useful' information measure of order α and type β $H_\alpha^\beta(P;U)$ given in (2.1)

Property 4.1: $H_\alpha^\beta(P;U)$ is non-negative.

Proof: From (2.1), we have

$$H_\alpha^\beta(P;U) = H_\alpha^\beta(P;U) = \frac{\beta}{\beta-\alpha}\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^\beta}\right], 0<\alpha<1, 0<\beta\le 1, \beta>\alpha.$$

From table (3.1) and (3.2) it is observed that $H_\alpha^\beta(P;U)$ is non-negative for given values of α and β.

Property 4.2: $H_\alpha^\beta(P;U)$ is a symmetric function on every $p_i, i = 1, 2, 3, \dots, n.$

Proof: It is obvious that $H_\alpha^\beta(P;U)$ is a symmetric function on every $p_i, i = 1, 2, 3, \dots, n.$ i.e., $H_\alpha^\beta(p_1u_1, p_2u_2, \dots, p_{n-1}u_{n-1}, p_nu_n) = H_\alpha^\beta(p_nu_n, p_1u_1, p_2u_2, \dots, p_{n-1}u_{n-1})$

Property 4.3: $H_\alpha^\beta(P;U)$ is maximum when all the events have equal probabilities.

Proof: When $p_i = \frac{1}{n} \forall\, i = 1,2,\dots,n$ and $\beta = 1, \alpha \to 1.$ and $u_i = 1,\ \forall\, i = 1,2,\dots,n,$ i.e., when the utility aspect is ignored, and $\sum_{i=1}^n p_i = 1.$ Then $H_\alpha^\beta(P;U) = \log_D n,$ which is maximum entropy.

Property 4.4: $H_\alpha^\beta(P)$ is concave function for $p_1, p_2, \dots, p_n.$

Proof: From (2.1), we have

$$H_\alpha^\beta(P;U) = \frac{\beta}{\beta-\alpha}\left[\frac{\sum_{i=1}^n u_i p_i^{\alpha\beta}}{\sum_{i=1}^n u_i p_i^{\beta}}\right], 0 < \alpha < 1, 0 < \beta \le 1, \beta > \alpha.$$

If $\beta = 1, \alpha \to 1, u_i = 1,\ \forall\, i = 1,2,\dots,n,$ i.e., when the utility aspect is ignored, and $\sum_{i=1}^n p_i = 1.$ then the first derivative of (2.1) with respect p_i is given by

$$\left[\frac{d}{dp_i} H_\alpha^\beta(P;U)\right]_{\substack{\beta=1\\ \alpha\to 1\\ u_i=1}} = -n - \sum_{i=1}^n \log_D p_i$$

And the second derivative is given by

$$\left[\frac{d^2}{dp_i^2} H_\alpha^\beta(P;U)\right]_{\substack{\beta=1\\ \alpha\to 1\\ u_i=1}} = -\sum_{i=1}^n \left(\frac{1}{p_i}\right) < 0.$$ For all $p_i \epsilon\, [0,1]$ and $i = 1, 2, \dots, n.$

Since the second derivative of $H_\alpha^\beta(P;U)$ with respect to p_i is negative on given interval $p_i \epsilon [0,1]$ $i = 1, 2, \dots, n.$ as $\beta = 1, \alpha \to 1, u_i = 1,\ \forall\, i = 1,2,\dots,n,$ i.e., when the utility aspect is ignored, and $\sum_{i=1}^n p_i = 1,$ therefore,

$H_\alpha^\beta(P;U)$ is concave function for $p_1, p_2, \dots, p_n.$

5. CONCLUSION

In this paper we define a new generalized 'useful' entropy measure of order α and type β i.e., $H_\alpha^\beta(P;U)$.This measure also generalizes some well-known information measures already existing in the literature of 'useful' information theory. Also we define a new generalized 'useful' code-word mean length i.e., $L_\alpha^\beta(P;U)$ corresponding to $H_\alpha^\beta(P;U)$, and then we characterize $L_\alpha^\beta(P;U)$ in term of $H_\alpha^\beta(P;U)$ and showed that

$$H_\alpha^\beta(P;U) \leq L_\alpha^\beta < H_\alpha^\beta(P;U)D^{1-\alpha}.\text{where } 0 < \alpha < 1, 0 < \beta \leq 1, \beta > \alpha.$$

Further we have established the noiseless coding theorems proved in this paper with the help of two different techniques by taking experimental data and show that Huffman coding scheme is more efficient than Shannon-Fano coding scheme. The important properties of $H_\alpha^\beta(P;U)$ have also been studied.

6. REFERENCES

[1] Belis M, Guiasu S, "A quantitative-qualitative Measure of Information in cybernetic System", IEEE Transaction on information theory, vol. 14, pp. 593-594, 1968.
[2] Bhaker U. S, Hooda D. S, "Mean value characterization of 'useful' information measure", Tamkang Journal of Mathematics, vol. 24, pp. 283-294, 1993.
[3] Bhat A. H, Baig M. A. K, "Some coding theorems on generalized Reyni's entropy of order α and type β", International Journal of Applied Mathematics and Information Sciences Letters, vol. 5, pp. 1-5, 2016.
[4] Campbell L. L, "A Coding Theorem and Renyi's Entropy", Information and Control, Vol. 8, pp. 423-429, 1965.
[5] Guiasu S and Picard C. F, "Borne Inferieure De la Longer De Certain Codes", C. R Academic Sciences, Paris, Vol. 273, pp. 248-251, 1971.
[6] Gurdial and Pessoa F, "On Useful Information of Order α", Journal of Combinatorics, Information and System Sciences, Vol. 2, pp. 158 – 162, 1977.
[7] Hartley R. V. L, "Transmission of Information", Bell System Technical Journal, Vol. 7, pp. 535-563, 1928.
[8] Hooda D. S, Bhaker U. S, "A Generalized Useful Information Measure and Coding Theorems", Soochow Journal of Mathematics, Vol. 23, pp. 53-62, 1997.
[9] Jain P, Tuteja R. K, "On Coding Theorem Connected with Useful Entropy of Order β", International Journal of Mathematics and Mathematical Sciences, Vol. 12, pp. 193-198, 1989.
[10] Khan A. B, Bhat B. A, Pirzada S, "Some Results on a Generalized Useful Information Measure", Journal of Inequalities in Pure and Applied Mathematics, Vol. 6, Article 117, 2005.
[11] Kraft L. J, "A Device for Quantizing Grouping and Coding Amplitude Modulates Pulses", M.S Thesis, Department of Electrical Engineering, MIT, Cambridge, 1949.
[12] Longo G, "A Noiseless Coding Theorem for Source Having Utilities", SIAM Journal of Applied Mathematics, Vol. 30, pp. 739-748, 1976.
[13] Mitter J, Mathur Y. D, "Comparison of Entropies of Power Distributions", ZAMM, Vol. 52, pp. 239-240, 1972.
[14] Nyquist H, "Certain Factors Affecting Telegraph Speed", Bell System Technical Journal, Vol. 3, pp. 324-346, 1924.
[15] Nyquist H, "Certain Topics in Telegraphy Transmission Theory", Journal of the American Institute of Electrical Engineers, Vol. 47, pp. 617, 1928.
[16] Renyi A, "On measure of entropy and information", Proceeding Fourth Berkely Symposium on Mathematical Statistics and probability, University of California Press, Vol. 1, pp. 547-561, 1961.
[17] Shannon C. E, "A Mathematical Theory of Communication", Bell System Technical Journal, Vol. 27, pp. 379-423, 623-659, 1948.
[18] Sharma B. D, Man Mohan, Mitter J, "On Measure of Useful Information", Information and Control, Vol. 39, pp. 323-336, 1978.
[19] Taneja H. C, Hooda D. S, Tuteja R. K, "Coding Theorems on a Generalized Useful Information", Soochow Journal of Mathematics, Vol. 11, pp. 123-131, 1985.