RESEARCH ARTICLE

Some Coding Theorem Connected on Generalized Renyi’s Entropy for Incomplete Power Probability Distribution $p^\beta$

Satish Kumar and Arun Choudhary *

Department of Mathematics, Geeta Institute of Management & Technology, Kanipla-136131, Kurukshetra, Haryana, India.

(Received: 9 January 2012, Accepted: 14 May 2012)

A new measure $L_{\beta}^\alpha$, called average codeword length of order $\alpha$ for incomplete power probability distribution $p^\beta$ has been defined and its relationship with a result of generalized Renyi’s entropy has been discussed. Using $L_{\beta}^\alpha$, some coding theorem for discrete noiseless channel has been proved.

Keywords: Renyi’s Entropy; Codeword length; Kraft inequality; Optimal code length and Power probabilities.

AMS Subject Classification: 94A15, 94A17, 94A24, 26D15.

1. Introduction

Throughout the paper $\mathbb{N}$ denotes the set of the natural numbers and for $N \in \mathbb{N}$ we set

$$\Delta_N = \left\{ (p_1, \ldots, p_N) / p_i \geq 0, i = 1, \ldots, N, \sum_{i=1}^{N} p_i = 1 \right\}.$$

In case there is no rise to misunderstanding we write $P \in \Delta_N$ instead of $(p_1, \ldots, p_N) \in \Delta_N$. In case $N \in \mathbb{N}$ the well-known Shannon entropy is defined by

$$H(P) = H(p_1, \ldots, p_N) = -\sum_{i=1}^{N} p_i \log(p_i) \quad ((p_1, \ldots, p_N) \in \Delta_N),$$

where the convention $0 \log(0) = 0$ is adapted, (see Shannon [1]).

Throughout this paper, $\sum$ will stand for $\sum_{i=1}^{N}$ unless otherwise stated and logarithms are taken to the base $D \ (D > 1)$.

Let a finite set of $N$ input symbols

$$X = \{x_1, x_2, \ldots, x_N\}$$

be encoded using alphabet of $D$ symbols, then it has been shown by Feinstien [2] that there is a uniquely decipherable code with lengths $n_1, n_2, \ldots, n_N$ if and only if the Kraft inequality holds, that

* Corresponding author.
Email: arunchoudhary07@gmail.com (Arun Choudhary).
is,
\[ \sum_{i=1}^{N} D^{-n_i} \leq 1. \]  
(1)

Where \( D \) is the size of code alphabet. Furthermore, if

\[ L = \sum_{i=1}^{N} n_i p_i \]  
(2)

is the average codeword length, then for a code satisfying (1), the inequality

\[ L \geq H(P) \]  
(3)

is also fulfilled and equality holds if and only if

\[ n_i = -\log_D(p_i) \quad (i = 1, \ldots, N), \]  
(4)

and that by suitable encoded into words of long sequences, the average length can be made arbitrarily close to \( H(P) \), (see Feinstein [2]). This is Shannon’s noiseless coding theorem.

By considering Renyi’s entropy (see e.g. [3]), a coding theorem and analogous to the above noiseless coding theorem has been established by Campbell [4] and the authors obtained bounds for it in terms of

\[ H_{\alpha}(P) = \frac{1}{1-\alpha} \log_D \sum p_i^\alpha, \alpha > 0(\neq 1). \]

Kieffer [5] defined a class rules and showed \( H_{\alpha}(P) \) is the best decision rule for deciding which of the two sources can be coded with expected cost of sequences of length \( n \) when \( n \to \infty \), where the cost of encoding a sequence is assumed to be a function of length only. Further, in Jelinek [6] it is shown that coding with respect to Campbell’s mean length is useful in minimizing the problem of buffer overflow which occurs when the source symbol is produced at a fixed rate and the code words are stored temporarily in a finite buffer. Concerning Campbell’s mean length the reader can consult [4].

Hooda and Bhaker considered in [7] the following generalization of Campbell’s mean length:

\[ L^\beta(t) = \frac{1}{t} \log_D \left\{ \frac{\sum p_i^\beta D^{-tn_i}}{\sum p_i^\beta} \right\}, \beta \geq 1 \]

and proved

\[ H_{\alpha}^\beta(P) \leq L^\beta(t) < H_{\alpha}^\beta(P) + 1, \quad \alpha > 0, \quad \alpha \neq 1, \quad \beta \geq 1 \]

under the condition

\[ \sum p_i^{\beta-1} D^{-n_i} \leq \sum p_i^\beta \]

where \( H_{\alpha}^\beta(P) \) is generalized entropy of order \( \alpha = \frac{1}{1+t} \) and type \( \beta \) studied by Aczel and Daroczy [8] and Kapur [9]. It may be seen that the mean codeword length (2) had been generalized parametrically and their bounds had been studied in terms of generalized measures of entropies. Here we give another generalization of (2) and study its bounds in terms of generalized entropy of order \( \alpha \) and type \( \beta \).
Generalized coding theorems by considering different information measure under the condition of unique decipherability were investigated by several authors, see for instance the papers [2, 7, 10–14]. An investigation is carried out concerning discrete memoryless sources possessing an additional parameter $\alpha, \beta$ which seems to be significant in problem of storage and transmission (see [6], [5] and [11]).

In this paper we study some coding theorems by considering a new information measure depending on two parameters. Our motivation is -among others- that this quantity generalizes some information measures already existing in the literature such as the Arndt and Renyi’s entropy, (see [15] and [3]), which is used in physics.

2. Coding Theorem

Definition 2.1 Let $N \in \mathbb{N}$ be arbitrarily fixed, $\alpha, \beta > 0$, $\alpha \neq 1$ be given real numbers. Then the information measure $H^\beta_\alpha : \Delta_N \to \mathbb{R}$ is defined by

$$H^\beta_\alpha (p_1, \ldots, p_N) = \frac{1}{1-\alpha} \log \frac{\sum_{i=1}^N p_i^{\alpha \beta}}{\sum_{j=1}^N p_j^\beta} \quad ((p_1, \ldots, p_N) \in \Delta_N).$$

(5)

It also studied by Roy [14].

Remark 2.2

(i) When $\beta = 1$, then the information measure $H^\beta_\alpha$ reduces to Renyi’s [3] entropy, i.e.,

$$H_\alpha (p_1, \ldots, p_N) = \frac{1}{1-\alpha} \log \sum_{i=1}^N p_i^\alpha \quad ((p_1, \ldots, p_N) \in \Delta_N).$$

(6)

(ii) When $\beta = 1$ and $\alpha \to 1$, then the information measure $H^\beta_\alpha$ reduces to Shannon entropy, i.e.,

$$H (P) = -\sum_{i=1}^N p_i \log p_i.$$  

(7)

(iii) When $\alpha \to 1$, the information measure $H^\beta_\alpha$ is the entropy of the $\beta$-power distribution, i.e.,

$$H^\beta (P) = -\frac{\sum_{i=1}^N p_i^\beta \log(p_i^\beta)}{\sum_{j=1}^N p_j^\beta},$$

(8)

that was considered e.g. in Mathur-Mitter [16].

Definition 2.3 Let $N \in \mathbb{N}$, $\alpha, \beta > 0$, $\alpha \neq 1$ be arbitrarily fixed, then the mean length $L^\beta_\alpha$ corresponding to the generalized information measure $H^\beta_\alpha$ is given by the formula

$$L^\beta_\alpha = \frac{\alpha}{1-\alpha} \log \left( \frac{\sum_{i=1}^N p_i^\beta D^{-n_i \left(1 + \frac{\alpha - 1}{\alpha} \right)}}{\sum_{j=1}^N p_j^\beta} \right),$$

(9)

where $(p_1, \ldots, p_N) \in \Delta_N$ and $D, n_1, n_2, \ldots, n_N$ are positive integers so that

$$\sum_{i=1}^N D^{-n_i} \leq \sum_{j=1}^N p_j^\beta.$$  

(10)
Remark 2.4

(i) When $\beta = 1$, (9) reduces to a mean codeword length defined by Campbell [4]. i.e.,

$$L_\alpha = \frac{\alpha}{1 - \alpha} \log \left\{ \sum_{i=1}^{N} p_i D^{-n_i \left(\frac{\alpha - 1}{\alpha}\right)} \right\}. \tag{11}$$

(ii) When $\beta = 1, \alpha \to 1$, (9) reduces to a mean code length $L = \sum_{i=1}^{N} n_i p_i$, defined in Shannon [1].

Also, we have used the condition (10) to find the bounds. It may be seen that the case $\beta = 1$ inequality (10) reduces to (1).

We establish a result, that in a sense, provides a characterization of $H_\alpha^\beta (P)$ under the condition of (10).

Theorem 2.5 Let $\alpha, \beta > 0$, $\alpha \neq 1$ be arbitrarily fixed real numbers, then for all integers $D > 1$ inequality

$$L_\alpha^\beta \geq H_\alpha^\beta (P) \tag{12}$$

is fulfilled. Furthermore, equality holds if and only if

$$n_i = - \log_D \left( \frac{\sum_{j=1}^{N} P_i^{\alpha \beta}}{\sum_{j=1}^{N} P_j^{\beta}} \right) \tag{13}$$

Proof By reverse Hölder inequality, that is, if $N \in \mathbb{N}$, $\gamma > 1$ and $x_1, \ldots, x_N, y_1, \ldots, y_N$ are positive real numbers then

$$\left( \sum_{i=1}^{N} x_i^{\frac{1}{\gamma}} \right) \gamma \left( \sum_{i=1}^{N} y_i^{\frac{1}{(\gamma - 1)}} \right)^{-(\gamma - 1)} \leq \sum_{i=1}^{N} x_i y_i. \tag{14}$$

Let

$$\gamma = \frac{\alpha}{\alpha - 1}, \quad x_i = \left( \frac{P_i^{\beta}}{\sum_{j=1}^{N} P_j^{\beta}} \right)^{\left(\frac{\alpha - 1}{\alpha}\right)} D^{-n_i}, \quad y_i = \left( \frac{P_i^{\alpha \beta}}{\sum_{j=1}^{N} P_j^{\beta}} \right)^{\frac{1}{1 - \alpha}} (i = 1, \ldots, N).$$

Putting these values into (14), we get

$$\left( \sum_{i=1}^{N} P_i^{\beta} D^{-n_i \left(\frac{\alpha - 1}{\alpha}\right)} \right)^{\frac{\alpha}{\alpha - 1}} \left( \sum_{j=1}^{N} P_j^{\beta} \right)^{\frac{1}{1 - \alpha}} \leq \sum_{i=1}^{N} D^{-n_i} \sum_{j=1}^{N} P_j^{\beta} \leq 1,$$

where we used (10) too. This implies however that

$$\left( \frac{\sum_{i=1}^{N} P_i^{\alpha \beta}}{\sum_{j=1}^{N} P_j^{\beta}} \right)^{\frac{1}{1 - \alpha}} \leq \left( \frac{\sum_{i=1}^{N} P_i^{\beta} D^{-n_i \left(\frac{\alpha - 1}{\alpha}\right)}}{\sum_{j=1}^{N} P_j^{\beta}} \right)^{\frac{\alpha}{\alpha - 1}}. \tag{15}$$

We obtain the result (12) after taking logarithm on both sides of (15). i.e., $L_\alpha^\beta \geq H_\alpha^\beta (P)$. From (13)
and after simplification, we get
\[
D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)} = p_i^{\beta(\alpha-1)} \left( \frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right)^{\frac{1-\alpha}{\alpha}}.
\]

This implies
\[
\left( \frac{\sum_{i=1}^{N} p_i^{\beta} D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{j=1}^{N} p_j^{\beta}} \right)^{\alpha} = \left( \frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right),
\]

which gives \( L_\alpha^\beta = H_\alpha^\beta (P) \). \[\blacksquare\]

**Theorem 2.6** Let \( N \in \mathbb{N}, \alpha, \beta > 0, \alpha \neq 1 \) be fixed. Then there exist code length \( n_1, \ldots, n_N \) so that
\[
H_\alpha^\beta (P) \leq L_\alpha^\beta < H_\alpha^\beta (P) + \log D
\]
holds. Where \( H_\alpha^\beta (P) \) and \( L_\alpha^\beta \) are given by (5) and (9), respectively.

**Proof** Due to the previous theorem, \( L_\alpha^\beta = H_\alpha^\beta (P) \) holds if and only if
\[
D^{-n_i} = \frac{p_i^{\alpha\beta}}{\sum_{i=1}^{N} p_i^{\alpha\beta}}, \quad \alpha > 0, \quad \alpha \neq 1, \quad \beta > 0,
\]
i.e.,
\[
n_i = -\log_D p_i^{\alpha\beta} + \log_D \left[ \frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right].
\]

We choose the codeword lengths \( n_i, i = 1, \ldots, N \) in such a way that
\[
-\log_D p_i^{\alpha\beta} + \log_D \left[ \frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right] \leq n_i < -\log_D p_i^{\alpha\beta} + \log_D \left[ \frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right] + 1 \tag{17}
\]
is fulfilled for all \( i = 1, \ldots, N \).

From the left inequality of (17), we have
\[
D^{-n_i} = \frac{p_i^{\alpha\beta}}{\sum_{i=1}^{N} p_i^{\alpha\beta}}, \tag{18}
\]
taking sum over \( i \), we get the generalized inequality (10). So there exists a generalized code with code lengths \( n_i, i = 1, \ldots, N \).

**Case 1.** Let \( 0 < \alpha < 1 \), then (17) can be written as
\[
\frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \leq D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)} < p_i^{\beta(\alpha-1)} \left( \frac{\sum_{i=1}^{N} p_i^{\alpha\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right)^{\frac{1-\alpha}{\alpha}} D^{1-\alpha}. \tag{19}
\]
Since $n$...Clearly...of Theorem 2.5, $L$.

Suppose

Theorem 2.7 For arbitrary $N$.

Case 2. Let $\alpha > 1$, then (17) can be written as

\[ p_i^{(a-1)} \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right)^{1-\alpha} \geq D^{-n_i(\frac{a-1}{\alpha})} > p_i^{\beta(\alpha-1)} \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right)^{1-\alpha} D^{1-\alpha}. \]  

(20)

Multiplying (20) throughout by $\frac{\beta^\alpha}{\sum_{j=1}^{N} p_j^\beta}$ and then summing up from $i = 1$ to $i = N$, we obtain inequality (16) after simplification with $\frac{\alpha}{1-\alpha}$, i.e.,

\[ \frac{1}{1-\alpha} \log \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right) \leq \frac{1}{1-\alpha} \log \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i(\frac{a-1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right) \]

\[ < \frac{1}{1-\alpha} \log \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right) + \log D \]

$H_1^\alpha(P) \leq L_1^\alpha < H_1^\beta(P) + \log D$, which gives (16).

\[ \mathbf{\text{Case 2.}} \]

\[ \mathbf{Let \ \alpha > 1, \ then \ (17) \ \can \ \be \ \written \ \as} \]

\[ p_i^{(a-1)} \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right)^{1-\alpha} \geq D^{-n_i(\frac{a-1}{\alpha})} > p_i^{\beta(\alpha-1)} \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right)^{1-\alpha} D^{1-\alpha}. \]

(20)

Multiplying (20) throughout by $\frac{\beta^\alpha}{\sum_{j=1}^{N} p_j^\beta}$ and then summing up from $i = 1$ to $i = N$, we obtain inequality (16) after simplification with $\frac{\alpha}{1-\alpha}$, i.e.,

\[ \frac{1}{1-\alpha} \log \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right) \leq \frac{1}{1-\alpha} \log \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i(\frac{a-1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right) \]

\[ < \frac{1}{1-\alpha} \log \left( \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{N} p_j^\beta} \right) + \log D \]

$H_1^\alpha(P) \leq L_1^\alpha < H_1^\beta(P) + \log D$, which gives (16).

\[ \mathbf{\text{Theorem 2.7}} \]

\[ \mathbf{For \ arbitrary \ N, \ \alpha, \ \beta > 0, \ \alpha \neq 1 \ \and \ \for \ \every \ \code \ \word \ \lengths \ \n_i, \ \i = 1, \ldots, N \ \of \ \Theorem 2.5, \ \L_\alpha^\beta \ \can \ \be \ \made \ \to \ \satisfy,} \]

\[ L_\alpha^\beta \geq H_\alpha^\beta (P) > H_\alpha^\beta (P) + \frac{\log D}{1-\alpha}. \]

(21)

\[ \mathbf{\text{Proof}} \]

\[ \mathbf{Suppose} \]

\[ \mathbf{n_i} = -\log D \left( \frac{p_i^\alpha}{\sum_{i=1}^{N} p_i^\alpha} \right), \ \ \ \alpha > 0, \ \alpha \neq 1, \ \beta > 0. \]

(22)

Clearly $\mathbf{n_i}$ and $\mathbf{n_i} + 1$ satisfy ‘equality’ in Holder’s inequality (14). Moreover, $\mathbf{n_i}$ satisfies (10). Suppose $\mathbf{n_i}$ is the unique integer between $\mathbf{n_i}$ and $\mathbf{n_i} + 1$, then obviously, $\mathbf{n_i}$ satisfies (10).

Since $\alpha > 0 (\neq 1)$, we have

\[ \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i(\frac{a-1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha \leq \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-\mathbf{n_i}(\frac{a-1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha \]

\[ < D \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-\mathbf{n_i}(\frac{a-1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha. \]

(23)
Since
\[
\left( \frac{\sum_{i=1}^{N} p_i^{\beta} D^{-n_i \left( \frac{\alpha-1}{\alpha} \right)}}{\sum_{j=1}^{N} p_j^{\beta}} \right) = \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}}.
\]
Hence (23) becomes
\[
\left( \frac{\sum_{i=1}^{N} p_i^{\beta} D^{-n_i \left( \frac{\alpha-1}{\alpha} \right)}}{\sum_{j=1}^{N} p_j^{\beta}} \right) \leq \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} < D \left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right)
\]
which gives (21).

3. Conclusion

It is clear that the optimal code is that code for which the value of \( L_{\alpha}^{\beta} \) is equal to its lower bound. Considering Theorem 2.5, we remark that the optimal code lengths are dependent of \( \alpha, \beta \) in contrast with the optimal code lengths of Shannon which do not depend of a parameter. So it can be reduced significantly by taking suitable value of parameters. However it is also possible to prove coding theorems with respect to \( H_{\alpha}^{\beta}(P) \) such that the optimal code lengths are identical to those of Shannon.

References