SOME CODING THEOREMS ON GENERALIZED
HAVRDA-CHARVAT AND TSALLIS’S ENTROPY

SATISH KUMAR AND ARUN CHOUDHARY

Abstract. A new measure \( L_{\alpha}^\beta \), called average code word length of order \( \alpha \) and type \( \beta \) has been defined and its relationship with a result of generalized Havrda-Charvat and Tsallis’s entropy has been discussed. Using \( L_{\alpha}^\beta \), some coding theorem for discrete noiseless channel has been proved.

1. Introduction

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

\[
\Delta_N = \left\{ (p_1, \ldots, p_N) \in [0, 1], \sum_{i=1}^{N} p_i = 1 \right\}.
\]

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 [19]).

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 [6] 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 is,

\[
\sum_{i=1}^{N} D^{-n_i} \leq 1. \tag{1.2}
\]

Corresponding author: Arun Choudhary.

2010 Mathematics Subject Classification. 94A15, 94A17, 94A24, 26D15.

Key words and phrases. Tsallis’s entropy, codeword length, Kraft inequality, optimal code length and power probabilities.
Where $D$ is the size of code alphabet.

Furthermore, if

$$L = \sum_{i=1}^{N} n_i p_i \quad (1.3)$$

is the average codeword length, then for a code satisfying (1.2), the inequality

$$L \geq H(P) \quad (1.4)$$

is also fulfilled and equality holds if and only if

$$n_i = -\log_D(p_i) \quad (i = 1, \ldots, N), \quad (1.5)$$

and that by suitable encoded into words of long sequences, the average length can be made arbitrarily close to $H(P)$, (see Feinstein [6]). This is Shannon’s noiseless coding theorem.

By considering Renyi’s entropy (see e.g. [18]), a coding theorem and analogous to the above noiseless coding theorem has been established by Campbell [3] 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 [13] 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 [10] 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 [3].

Hooda and Bhaker considered in [9] the following generalization of Campbell’s mean length:

$$L^\beta(t) = \frac{1}{t} \log_D \left\{ \frac{\sum p_i^\beta D^{-t n_i}}{\sum p_i^\beta} \right\}, \quad \beta \geq 1$$

and proved

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

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 [1] and Kapur [11]. It may be seen that the mean codeword length (1.3) had been general-
ized parametrically and their bounds had been studied in terms of generalized measures of entropies. Here we give another generalization of (1.3) and study its bounds in terms of gen-
eralized 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 [7, 9, 12, 14, 17, 20].

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 Tsallis (or Havrda-Charvat) entropy, (see [8] and [21]).

2. Coding theorem

**Definition.** 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

$$H^\beta_\alpha(p_1, \ldots, p_N) = \frac{1}{\alpha - 1} \left( 1 - \frac{\sum_{i=1}^{N} p_i^\beta}{\sum_{j=1}^{N} p_j^\beta} \right)^{\alpha - 1} \left( (p_1, \ldots, p_N) \in \Delta_N \right).$$  \hspace{1cm} (2.1)

**Remark (i)** When $\beta = 1$, then the information measure $H^\beta_\alpha$ reduces to entropy,

$$i.e., \quad H_\alpha(p_1, \ldots, p_N) = \frac{1}{\alpha - 1} \left[ 1 - \sum_{i=1}^{N} p_i^\alpha \right] \left( (p_1, \ldots, p_N) \in \Delta_N \right).$$  \hspace{1cm} (2.2)

The measure (2.2) was characterized by Havrda-Charvat [8], Vajda [22], Tsallis [21] and Daroczy [4] by different approaches.

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

$$i.e., \quad H(P) = -\sum p_i \log p_i.$$  \hspace{1cm} (2.3)

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

$$i.e., \quad H^\beta(P) = -\frac{\sum_{i=1}^{N} p_i^\beta \log(p_i^\beta)}{\sum_{j=1}^{N} p_j^\beta},$$  \hspace{1cm} (2.4)

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

**Definition.** 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

$$L^\beta_\alpha = \frac{1}{\alpha - 1} \left( 1 - \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i(a-1)}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha \right).$$  \hspace{1cm} (2.5)
where \((p_1, \ldots, p_N) \in \Delta_N\) and \(D, \ n_1, n_2, \ldots, n_N\) are positive integers so that
\[
\sum_{i=1}^{N} D^{-n_i} \leq \sum_{j=1}^{N} p_j^\beta.
\] (2.6)

**Remark (i)** When \(\beta = 1\), (2.5) reduces to a mean codeword length,
\[
i.e., \quad L_a = \frac{1}{\alpha - 1} \left[ 1 - \left\{ \sum p_i D^{-n_i (\frac{\alpha - 1}{\alpha})} \right\}^\alpha \right].
\] (2.7)

(ii) When \(\beta = 1, \ \alpha \to 1\), (2.5) reduces to a mean code length \(L = \sum_{i=1}^{N} n_i p_i\), defined in Shannon [19].

Also, we have used the condition (2.6) to find the bounds. It may be seen that the case \(\beta = 1\) inequality (2.6) reduces to (1.2).

We establish a result, that in a sense, provides a characterization of \(H_{\alpha}^\beta (P)\) under the condition of (2.6).

**Theorem 2.1.** Let \(\alpha, \beta > 0, \ \alpha \neq 1\) be arbitrarily fixed real numbers, then for all integers \(D > 1\) inequality
\[
I_{\alpha}^\beta \geq H_{\alpha}^\beta (P)
\] (2.8)
is fulfilled. Furthermore, equality holds if and only if
\[
n_i = -\log D \left( \frac{p_i^{\alpha \beta}}{\sum_{i=1}^{N} p_i^{\alpha \beta}} \right).\] (2.9)

**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
\[
\left( \sum_{i=1}^{N} x_i^{\gamma} \right)^{\frac{1}{\gamma}} \cdot \left( \sum_{i=1}^{N} y_i^{-(\gamma - 1)} \right)^{-(\gamma - 1)} \leq \sum_{i=1}^{N} x_i y_i.
\] (2.10)

Let \(\gamma = \frac{\alpha}{\alpha - 1}\), \(x_i = \left( \frac{p_i^{\beta}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\frac{\alpha - 1}{\alpha}} D^{-n_i}, \ y_i = \left( \frac{p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\alpha \beta}} \right)^{\frac{1}{\alpha}} (i = 1, \ldots, N).

Putting these values into (2.10), we get
\[
\left( \frac{\sum_{i=1}^{N} p_i^{\beta} D^{-n_i (\frac{\alpha - 1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\frac{\alpha - 1}{\alpha}} \leq \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\alpha \beta}} \leq 1,
\]
where we used (2.6), too. This implies however that
\[
\left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\frac{1}{\alpha}} \leq \left( \frac{\sum_{i=1}^{N} p_i^{\beta} D^{-n_i (\frac{\alpha - 1}{\alpha})}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\frac{\alpha}{1 - \alpha}}.
\] (2.11)
Here two cases arise

**Case 1.** When $0 < \alpha < 1$, then raising power $(1 - \alpha)$ to both sides of (2.11), we have

$$
\left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta_i}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\alpha} \leq \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i (\frac{\alpha}{\alpha - 1})}}{\sum_{j=1}^{N} p_j^\beta} \right)
$$

we obtain the result (2.8) after simplification for $\frac{1}{\alpha - 1} < 0$ as $0 < \alpha < 1$.

i.e.,

$$
L_\alpha^\beta \geq H_\alpha^\beta (P), \text{ when } 0 < \alpha < 1. \quad (2.12)
$$

**Case 2.** When $\alpha > 1$, then raising power $(1 - \alpha)$ to both sides of (2.11), we have

$$
\left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta_i}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\alpha} \geq \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i (\frac{\alpha}{\alpha - 1})}}{\sum_{j=1}^{N} p_j^\beta} \right)
$$

we obtain the result (2.8) after simplification for $\frac{1}{\alpha - 1} > 0$ as $\alpha > 1$.

i.e., $L_\alpha^\beta \geq H_\alpha^\beta (P)$, when $\alpha > 1$.

**Case 3.** From (2.9) and after simplification, we get

$$
D^{-n_i (\frac{\alpha - 1}{\alpha})} = p_i^{\beta (\alpha - 1)} \left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta_i}}{\sum_{j=1}^{N} p_j^\beta} \right)^{\left( \frac{\alpha - 1}{\alpha} \right)}.
$$

This implies

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

which gives $L_\alpha^\beta = H_\alpha^\beta (P)$. \qed

**Theorem 2.2.** 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

$$
H_\alpha^\beta (P) \leq L_\alpha^\beta < D^{1-\alpha} H_\alpha^\beta (P) + \frac{1}{\alpha - 1} (1 - D^{1-\alpha})
$$

holds. Where $H_\alpha^\beta (P)$ and $L_\alpha^\beta$ are given by (2.1) and (2.5) respectively.

**Proof.** Due to the previous theorem,

$$
L_\alpha^\beta = H_\alpha^\beta (P)
$$

holds if and only if

$$
D^{-n_i} = \frac{p_i^{\alpha \beta_i}}{\sum_{i=1}^{N} p_i^{\alpha \beta_i}}, \quad \alpha > 0, \quad \alpha \neq 1, \quad \beta > 0,
$$
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

\[
-\log_D p_i^{\alpha \beta} + \log_D \left[ \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right] \leq n_i < -\log_D p_i^{\alpha \beta} + \log_D \left[ \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right] + 1 \tag{2.14}
\]

is fulfilled for all \( i = 1, \ldots, N \).

From the left inequality of (2.14), we have

\[
D^{-n_i} \leq \frac{p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}}, \tag{2.15}
\]

taking sum over \( i \), we get the generalized inequality (2.6). So there exists a generalized code with code lengths \( n_i, i = 1, \ldots, N \).

**Case 1.** Let \( 0 < \alpha < 1 \), then (2.14) can be written as

\[
p_i^{\beta(\alpha-1)} \left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right)^{\frac{1-\alpha}{\alpha}} \leq D^{-n_i \left( \frac{\alpha-1}{\alpha} \right)} < p_i^{\beta(\alpha-1)} \left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right)^{\frac{1-\alpha}{\alpha}} D^{\frac{1-\alpha}{\alpha}}. \tag{2.16}
\]

Multiplying (2.16) throughout by \( \frac{p_i^{\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \) and then summing up from \( i = 1 \) to \( i = N \), we obtain inequality (2.13) after simplification with \( \frac{\alpha}{\alpha-1} \),

i.e.,

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

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

\( H_\alpha(P) \leq L_\alpha < D^{1-\alpha} H_\alpha(P) + \frac{1}{\alpha-1} (1 - D^{1-\alpha}) \), which gives (2.13).

**Case 2.** Let \( \alpha > 1 \), then (2.14) can be written as

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

Multiplying (2.16) throughout by \( \frac{p_i^{\beta}}{\sum_{j=1}^{N} p_j^{\beta}} \) and then summing up from \( i = 1 \) to \( i = N \), we obtain inequality (2.13) after simplification with \( \frac{\alpha}{\alpha-1} \),

i.e.,

\[
\frac{1}{\alpha-1} \left( 1 - \left( \frac{\sum_{i=1}^{N} p_i^{\alpha \beta}}{\sum_{j=1}^{N} p_j^{\beta}} \right) \right) \leq \frac{1}{\alpha-1} \left( 1 - \frac{\sum_{i=1}^{N} p_i^{\beta} D^{-n_i \left( \frac{\alpha-1}{\alpha} \right)}}{\sum_{j=1}^{N} p_j^{\beta}} \right) ^{\alpha}
\]
\[ \alpha > 0 \quad \frac{1}{\alpha - 1} \left( 1 - \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{\beta} p_j^\beta} \right) < D^{1-\alpha} \left\{ \frac{1}{\alpha - 1} \left( 1 - \frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{\beta} p_j^\beta} \right) + \frac{1}{\alpha - 1} (1 - D^{1-\alpha}) \right\} \]

\[ H_\alpha^\beta (P) \leq L_\alpha^\beta < D^{1-\alpha} H_\alpha^\beta (P) + \frac{1}{\alpha - 1} (1 - D^{1-\alpha}), \text{ which gives (2.13).} \]

**Theorem 2.3.** For arbitrary \( N \in \mathbb{N}, \alpha, \beta > 0, \alpha \neq 1 \) and for every code word lengths \( n_i, i = 1, \ldots, N \) of Theorem 2.1, \( L_\alpha^\beta \) can be made to satisfy,

\[ L_\alpha^\beta \geq H_\alpha^\beta (P) > H_\alpha^\beta (P) D + \frac{1}{\alpha - 1} (1 - D). \quad (2.17) \]

**Proof.** Suppose

\[ \tilde{n}_i = -\log_D \left( \frac{p_i^\alpha}{\frac{\sum_{i=1}^{N} p_i^\alpha}{\sum_{j=1}^{\beta} p_j^\beta}} \right), \quad \alpha > 0, \alpha \neq 1, \beta > 0. \quad (2.18) \]

Clearly \( \tilde{n}_i \) and \( \tilde{n}_i + 1 \) satisfy ‘equality’ in Holder’s inequality (2.10). Moreover, \( \tilde{n}_i \) satisfies (2.6). Suppose \( n_i \) is the unique integer between \( \tilde{n}_i \) and \( \tilde{n}_i + 1 \), then obviously, \( n_i \) satisfies (2.6).

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

\[ \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-n_i (\frac{a-1}{a})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha \leq \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-\tilde{n}_i (\frac{a-1}{a})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha < D \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-\tilde{n}_i (\frac{a-1}{a})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha. \quad (2.19) \]

Since

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

Hence (2.19) becomes

\[ \left( \frac{\sum_{i=1}^{N} p_i^\beta D^{-\tilde{n}_i (\frac{a-1}{a})}}{\sum_{j=1}^{N} p_j^\beta} \right)^\alpha \leq \frac{\sum p_i^\alpha}{\sum p_j^\beta} < D \left( \frac{\sum p_i^\alpha}{\sum p_j^\beta} \right)^\alpha \]

which gives (2.17).

**Acknowledgement**

We thank the referee for his valuable suggestions.

**References**


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

E-mail: arunchoudhary07@gmail.com