

Some More Noiseless Coding Theorem on Generalized R-Norm Entropy

Arun Choudhary & Satish Kumar

Department of Mathematics, Geeta Institute of Management & Technology

Kanipla-136131, Kurukshetra, Haryana, India

E-mail: arunchoudhary07@gmail.com

Abstract

A parametric mean length is defined as the quantity $L_R^\beta = \frac{R}{R-1} \left[1 - \frac{\sum_{i=1}^N p_i^\beta D^{-n_i \left(\frac{R-1}{R}\right)}}{\sum_{j=1}^N p_j^\beta} \right]$

where $R > 0 (\neq 1), \beta > 0, p_i > 0, \sum p_i = 1, i = 1, 2, \dots, N$. This being the mean length of code words. Lower and upper bounds for L_R^β are derived in terms of R-norm information measure for the incomplete power distribution.

AMS Subject classification. 94A15, 94A17, 94A24, 26D15.

Keywords: R-Norm Entropy, Codeword length, Kraft inequality, Holder's inequality, Optimal code length and Power probabilities

1. Introduction

Boekee and Lubbe [1980] studied R-Norm information of probability distribution

$$P = (p_1, p_2, \dots, p_N), p_i \geq 0, i = 1, 2, \dots, N$$

where

$$\sum_{i=1}^N p_i = 1.$$

And

$$R^* = \{R : R > 0, R \neq 1\}$$

given by

$${}_R H(P) = \frac{R}{R-1} \left[1 - \left(\sum_{i=1}^N p_i^R \right)^{\frac{1}{R}} \right]. \quad (1.1)$$

The R-norm information measure (1.1) is a real function $\Delta_N \rightarrow R^+$, defined on Δ_N where $N \geq 2$ and R^+ is the set of positive real numbers. This measure is different from Shannon's [1948], Renyi's [1961], Havrda and Charvat [1967], Daroczy [1970], Tsallis [1988] and Vajda [1961].

The most interesting property of this measure is that when $R \rightarrow 1$, R-Norm information measure (1.1) approaches to Shannon's [1948] entropy and in case $R \rightarrow \infty, {}_R H(P) \rightarrow (1 - \max p_i), i = 1, 2, \dots, N$.

Setting $r = \frac{1}{R}$ in (1.1), we get

$$H^r(P) = \frac{1}{1-r} \left[1 - \left(\sum_{i=1}^N p_i^{\frac{1}{r}} \right)^r \right], r > 0 (\neq 1), \quad (1.2)$$

which is a measure mentioned by Arimoto [1971] as an example of a generalized class of information measure. It may be marked that (1.2) also approaches to Shannon's [1948] entropy as $r \rightarrow 1$.

Let $\Delta_N = \{P = (p_1, p_2, \dots, p_N), p_i \geq 0, \sum p_i = 1\}, N \geq 2$ be the set of all finite discrete probability distributions, for any probability distribution $(p_1, p_2, \dots, p_N) = P \in \Delta_N$,

Shannon [1948] defined entropy as:

$$H(P) = - \sum p_i \log p_i. \quad (1.3)$$

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

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

be encoded using alphabet of D symbols, then it has been shown by Feinstein [1956] that there is uniquely decipherable instantaneous code with length n_1, n_2, \dots, n_N iff

$$\sum D^{-n_i} \leq 1 \quad (1.4)$$

where D is the size of code alphabet.

If

$$L = \sum n_i p_i \quad (1.5)$$

be the average codeword length then for a code which satisfies (1.4) it has also been shown by Feinstein [1956], that

$$L \geq H(P) \quad (1.6)$$

with equality iff

$$n_i = -\log_D p_i \text{ for } i = 1, 2, \dots, N \quad (1.7)$$

and that by suitable encoded into words of long sequences, the average length can be made arbitrary close to $H(P)$. This is Shannon's noiseless coding theorem.

By considering Renyi's [1961] entropy, a coding theorem and analogous to the above noiseless coding theorem has been established by Campbell [1965] 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 [1979] 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 $n \rightarrow \infty$, where the cost of encoding a sequence is assumed to be a function of length only. Further Jelinek [1980] showed that coding with respect to Campbell [1965] mean length is useful in minimizing the problem of buffer overflow which occurs when the source symbol are being produced at a fixed rate and the code words are stored temporarily in a finite buffer.

Hooda and Bhaker [1997] consider the following generalization of Campbell [1965] mean length:

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

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 β studied by Aczel and Daroczy [1963] and Kapur [1967]. It may be seen that the mean codeword length (1.5) had been generalized parametrically and their bounds had been studied in terms of generalized measures of entropies. Here we give another generalization of (1.5) and study its bounds in terms of generalized entropy of order α and type β .

Longo [1976], Gurdial and Pessow [1977], Singh, Kumar and Tuteja [2003], Parkash and Sharma [2004], Hooda and Bhaker [1997], Khan, Bhat and Pirzada [2005] have studied generalized coding theorems by considering different generalized measure of (1.3) and (1.5) under condition (1.4) of unique decipherability.

In this paper we study some coding theorems by considering a new function depending on parameters R and β . Our motivation for studying this new function is that it generalizes some entropy function already existing in literature Boekee and Lubbe [1980] which is used in physics.

2. Coding Theorem

In this section, we define information measure as

$$H_R^\beta(P) = \frac{R}{R-1} \left[1 - \left(\frac{\sum p_i^{R\beta}}{\sum p_i^\beta} \right)^{\frac{1}{R}} \right], \quad (2.1)$$

where $R > 0 (\neq 1)$, $\beta > 0$, $p_i > 0$, $\sum p_i = 1$, $i = 1, 2, \dots, N$.

(i) When $\beta = 1$, (2.1) reduces to Boekee and Lubbe [1980] R-Norm information measure.

$$i.e. H_R(P) = \frac{R}{R-1} \left[1 - \left(\sum p_i^R \right)^{\frac{1}{R}} \right]. \quad (2.2)$$

(ii) When $\beta = 1$, $R \rightarrow 1$, (2.1) reduces to Shannon's [1948] entropy.

$$i.e. H(P) = - \sum p_i \log p_i. \quad (2.3)$$

(iii) When $R \rightarrow 1$, (2.1) reduces to Mathur and Mitter's [1972] entropy for the β - power distribution.

$$i.e. H^\beta(P) = - \frac{\sum p_i^\beta \log p_i^\beta}{\sum p_i^\beta}. \quad (2.4)$$

Definition: The mean length L_R^β with respect to information measure is defined as :

$$L_R^\beta = \frac{R}{R-1} \left[1 - \frac{\sum_{i=1}^N p_i^\beta D^{-n_i \left(\frac{R-1}{R} \right)}}{\sum_{j=1}^N p_j^\beta} \right] \quad (2.5)$$

where $R > 0 (\neq 1)$, $\beta > 0$, $p_i > 0$, $\sum p_i = 1$, $i = 1, 2, \dots, N$.

(i) When $\beta = 1$, (2.5) reduces to Boekee and Lubbe [1980] mean codeword length.

$$i.e. L_R = \frac{R}{R-1} \left[1 - \sum p_i D^{-n_i \left(\frac{R-1}{R} \right)} \right]. \quad (2.6)$$

(ii) When $\beta = 1$, $R \rightarrow 1$, (2.5) reduces to a mean code length defined by Shannon [1948].

$$i.e. L = \sum n_i p_i.$$

Also, we have used the condition

$$\sum_{i=1}^N D^{-n_i} \leq \sum_{j=1}^N p_j^\beta \quad (2.7)$$

to find the bounds. It may be seen that in the case when $\beta = 1$, then (2.7) reduces to Kraft Inequality (1.4).

We establish a result, that in a sense, provides a characterization of $H_R^\beta(P)$ under the condition of (2.7).

Theorem 2.1 For all integers $D > 1$

$$L_R^\beta \geq H_R^\beta(P) \quad (2.8)$$

under the condition (2.7) equality holds iff

$$n_i = -\log_D \left(\frac{p_i^{R\beta}}{\frac{\sum_{i=1}^N p_i^{R\beta}}{\sum_{j=1}^N p_j^\beta}} \right), \quad R > 0, R \neq 1, \beta > 0. \quad (2.9)$$

Proof. By Holder inequality we have

$$\left(\sum_{i=1}^N x_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^N y_i^q \right)^{\frac{1}{q}} \leq \sum_{i=1}^N x_i y_i, \quad (2.10)$$

where $p^{-1} + q^{-1} = 1$; $p (\neq 0) < 1$, $q < 0$ or $q (\neq 0) < 1$, $p < 0$; $x_i, y_i > 0$ for each i .

$$\text{Let } p = \frac{R-1}{R}, \quad x_i = \left(\frac{p_i^\beta}{\sum_{j=1}^N p_j^\beta} \right)^{\left(\frac{R}{R-1} \right)} D^{-n_i},$$

$$q = 1 - R \text{ and } y_i = \left(\frac{p_i^{R\beta}}{\sum_{j=1}^N p_j^\beta} \right)^{\frac{1}{1-R}}.$$

Putting these values in (2.10) and using (2.9), we get

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

it implies

$$\left(\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right)^{\frac{1}{1-R}} \leq \left(\frac{\sum_{i=1}^N P_i^\beta D^{-n_i(\frac{R-1}{R})}}{\sum_{j=1}^N P_j^\beta} \right)^{\frac{R}{1-R}} \tag{2.11}$$

Here two cases arise

Case 1. When $0 < R < 1$, then raising power $\left(\frac{1-R}{R}\right)$ to both sides of (2.11), we have

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

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

i.e.

$$L_R^\beta(P) \geq H_R^\beta(P), \text{ when } 0 < R < 1. \tag{2.12}$$

Case 2. Similarly we can prove (2.8) for $R > 1$.

Theorem 2.2 On properly choosing the lengths n_1, n_2, \dots, n_N in the code of theorem 2.1, $L_R^\beta(P)$ can be made to satisfy the following inequality:

$$H_R^\beta(P) \leq L_R^\beta(P) < D^{\frac{1-R}{R}} H_R^\beta(P) + \frac{R}{R-1} (1 - D^{\frac{1-R}{R}}), \tag{2.13}$$

where $H_R^\beta(P)$ and $L_R^\beta(P)$ are given by (2.1) and (2.5) respectively.

Proof: It can be proved that there is equality in (2.8) if and only if

$$D^{-n_i} = \frac{P_i^{R\beta}}{\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta}}, \quad R > 0, R \neq 1, \beta > 0$$

$$\text{or } n_i = -\log_D P_i^{R\beta} + \log_D \left[\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right].$$

We choose the codeword lengths n'_i 's as integers satisfying

$$-\log_D P_i^{R\beta} + \log_D \left[\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right] \leq n_i < -\log_D P_i^{R\beta} + \log_D \left[\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right] + 1. \tag{2.14}$$

From the left inequality of (2.14), we have

$$D^{-n_i} \leq \frac{P_i^{R\beta}}{\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta}}, \tag{2.15}$$

taking sum over i, we get the generalized inequality (2.7). So there exists a generalized personal code with length n'_i 's.

Let $0 < R < 1$, then (2.14) can be written as

$$P_i^{\beta(R-1)} \left(\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right)^{\frac{1-R}{R}} \geq D^{-n_i(\frac{R-1}{R})} > P_i^{\beta(R-1)} \left(\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right)^{\frac{1-R}{R}} D^{\frac{1-R}{R}}. \tag{2.16}$$

Multiplying (2.16) by $\frac{P_i^\beta}{\sum_{j=1}^N P_j^\beta}$ throughout, summing over i and we obtain the result (2.13) after simplification for $\frac{R}{R-1} < 0$ as $0 < R < 1$.

$$\begin{aligned} \text{i.e. } \frac{R}{R-1} \left(1 - \left(\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right)^{\frac{1}{R}} \right) &\leq \frac{R}{R-1} \left(1 - \frac{\sum_{i=1}^N P_i^\beta D^{-n_i(\frac{R-1}{R})}}{\sum_{j=1}^N P_j^\beta} \right) \\ &< D^{\frac{1-R}{R}} \left\{ \frac{R}{R-1} \left(1 - \left(\frac{\sum_{i=1}^N P_i^{R\beta}}{\sum_{j=1}^N P_j^\beta} \right)^{\frac{1}{R}} \right) \right\} + \frac{R}{R-1} (1 - D^{\frac{1-R}{R}}) \end{aligned}$$

$H_R^\beta(P) \leq L_R^\beta(P) < D^{\frac{1-R}{R}} H_R^\beta(P) + \frac{R}{R-1}(1 - D^{\frac{1-R}{R}})$, which gives (2.13).

Similarly we can prove (2.13) for $R > 1$.

Theorem 2.3 For every code with length $\{n_i\}$, $i = 1, 2, \dots, N$ of theorem 2.1, L_R^β can be made to satisfy,

$$L_R^\beta \geq H_R^\beta(P) > H_R^\beta(P)D + \frac{R}{R-1}(1 - D). \quad (2.17)$$

Proof: Suppose

$$\bar{n}_i = -\log_D \left(\frac{P_i^{R\beta}}{\sum_{i=1}^N P_i^{R\beta}} \right), \quad R > 0, R \neq 1, \beta > 0. \quad (2.18)$$

Clearly \bar{n}_i and $\bar{n}_i + 1$ satisfy 'equality' in Holder's inequality (2.10). Moreover, \bar{n}_i satisfies (2.7). Suppose n_i is the unique integer between \bar{n}_i and $\bar{n}_i + 1$, then obviously, n_i satisfies (2.7).

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

$$\frac{\sum_{i=1}^N P_i^\beta D^{-n_i(\frac{R-1}{R})}}{\sum_{j=1}^N P_j^\beta} \leq \frac{\sum_{i=1}^N P_i^\beta D^{-\bar{n}_i(\frac{R-1}{R})}}{\sum_{j=1}^N P_j^\beta} < D \left(\frac{\sum_{i=1}^N P_i^\beta D^{-\bar{n}_i(\frac{R-1}{R})}}{\sum_{j=1}^N P_j^\beta} \right). \quad (2.19)$$

Since, $\frac{\sum_{i=1}^N P_i^\beta D^{-\bar{n}_i(\frac{R-1}{R})}}{\sum_{j=1}^N P_j^\beta} = \left(\frac{\sum P_i^{R\beta}}{\sum P_j^\beta} \right)^{\frac{1}{R}}$.

Hence (2.19) becomes

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

which gives (2.17).

References

- A. Feinstein. (1956). *Foundation of Information Theory*, McGraw Hill, New York.
- A. Renyi. (1961). On Measure of entropy and information, Proc. 4th Berkeley Symp. *Maths. Stat. Prob.*, Vol.1 547-561.
- A.B. Khan, B.A. Bhat and S. Pirzada. (2005). Some Results on a Generalized Useful Information Measure. *Journal of Inequalities in Pure and Applied Mathematics*, Vol. 6, Issue 4, Article 117.
- C. Tsallis. (1988). Possible Generalization of Boltzmann Gibbs Statistics. *J. Stat. Phys.*, 52 479.
- C.E. Shannon. (1948). A Mathematical Theory of Communication. *Bell System Tech.J.*, 27 379-423 623-656.
- D.S. Hooda and U.S. Bhaker. (1997). A generalized 'useful' information measure and coding theorems. *Soochow J. Math.*, 23 53-62.
- E. Boekee and J.C.A. Vander Lubbe. (1980). The R-norm Information Measure. *Information and Control*, 45 136-155.
- F. Jelinek. (1980). Buffer overflow in variable lengths coding of fixed rate sources, *IEEE*, 3 490-501.
- G. Longo. (1976). A Noiseless Coding Theorem for Sources Having Utilities. *Siam J. Appl. Math.*, 30 739-748.
- Gurdial and F. Pessow. (1977). On Useful Information of order α . *J. Comb. Information and Syst. Sci.*, 2 30-35.
- I. Vajda. (1961). On measure of entropy and information, Proc. Fourth Berk. *Symp. in Math, Stat. and Prob.*, 1 547-561.
- J. Aczel and Z. Daroczy. (1963). Uber Verallgemeineste quasilineare mittelveste die mit grewinebts functionen gebildet Sind. *Pub. Math. Debrecan*, 10 171-190.
- J. Mitter and Y.D Mathur. (1972). Comparison of entropies of power distribution. *ZAMM*, 52 239-240.
- J.C. Kieffer. (1979). Variable lengths source coding with a cost depending only on the codeword length. *Information and Control*, Vol. 41 136-146.
- J.F. Havrda and F. Charvat. (1967). Qualification Method of Classification Process, the concept of structural α -entropy. *Kybernetika*, 3 30-35.
- J.N. Kapur. (1967). Generalized entropy of order α and type β . *Maths Seminar*, Delhi, Vol. 4.
- L.L. Campbell. (1965). A coding theorem and Renyi's entropy. *Information and Control*, Vol. 8 423-429.

- Om Parkash and P.K Sharma. (2004). Noiseless Coding Theorems Corresponding to Fuzzy Entropies. *Southeast Asian Bulletin of Mathematics*, 27 1073-1080.
- R.P. Singh, R. Kumar and R.K. Tuteja. (2003). Application of Holder's Inequality in Information Theory. *Information Sciences*, 152 145-154.
- S. Arimoto. (1971). Information Theoretical Considerations on Estimation Problems. *Information and Control*, 1 181-199.
- Satish Kumar. (2009). Some More Results on R-Norm Information Measure. *Tamkang Journal of Mathematics*, Vol. 40 No. 1 41-58.
- Z. Daroczy. (1970). Generalized Information Functions. *Information and Control*, 16 36-51.