Applications and Applied Mathematics: An International Journal (AAM)

Volume 7 | Issue 1 Article 23

6-2012

Coding Theorems on a Non-Additive Generalized Entropy of Havrda-Charvat And Tsallis

Satish Kumar
Geeta Institute of Management & Technology

Arun Choudhary
Geeta Institute of Management & Technology

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Vol. 7, Issue 1 (June 2012), pp. 365 – 373

Applications and Applied Mathematics:
An International Journal
(AAM)

Coding Theorems on a Non-Additive Generalized Entropy of Havrda-Charvat And Tsallis

Satish Kumar and Arun Choudhary
Department of Mathematics
Geeta Institute of Management & Technology
Kanipla-136131, Kurukshetra, Haryana (India)
arunchoudhary07@gmail.com

Received: June 28, 2011; Accepted: June 02, 2012

Abstract

A new measure L_α^β, called average code word length of order α and type β is defined and its relationship with a generalized information measure of order α and type β is discussed. Using L_α^β, some coding theorems are proved.

Keywords: Codeword length, Optimal code length, Holder's inequality and Kraft inequality

MSC 2010: 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

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

With no danger of misunderstanding we write $P \in \Delta_N$ instead of $(p_1, ..., p_N) \in \Delta_N$.

In case $P \in \Delta_N$ the well-known Shannon entropy is defined by

$$H(P) = H(p_1, ..., p_N) = -\sum_{i=1}^{N} p_i \log(p_i), \tag{1}$$

where the convention $0\log(0) = 0$ is adopted, [see Shannon (1948)].

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

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

be encoded using an alphabet of D symbols, then it is shown in Feinstein (1956) that there is a uniquely decipherable code with lengths $n_1, n_2, ..., n_N$ if and only if the Kraft inequality holds, that is,

$$\sum_{i=1}^{N} D^{-n_i} \leq 1, \tag{2}$$

where D is the size of the code alphabet.

Furthermore, if

$$L = \sum_{i=1}^{N} n_i p_i \tag{3}$$

is the average codeword length, then for a code satisfying (2), the inequality

$$H(P) \leq L < H(P) + 1 \tag{4}$$

is also fulfilled and the equality, $L = H(P)$, holds if and only if

$$n_i = -\log_D(p_i) \ (i = 1, ..., N), and \sum_{i=1}^{N} D^{-n_i} = 1. \tag{5}$$

If $L < H(P)$, then by suitable encoding of long input sequences, the average number of code letters per input symbol can be made arbitrarily close to $H(P)$ [see Feinstein (1956)]. This is Shannon's noiseless coding theorem.

A coding theorem analogous to Shannon's noiseless coding theorem has been established by Campbell (1965), in terms of Renyi's entropy (1961):

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

Kieffer (1979) defined class rules and showed that $H_\alpha(P)$ is the best decision rule for deciding which of the two sources can be coded with least 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 the length only. Further, in Jelinek (1980) 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 Campbell (1965).

Hooda and Bhaker considered in (1997) 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

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

under the condition

$$\sum p_i^{\beta-1} D^{-n_i} \leq \sum p_i^\beta,$$

where $H_\alpha^\beta(P)$ is the generalized entropy of order $\alpha = \frac{1}{1+t}$ and type β studied by Aczel and Daroczy (1963) and Kapur (1967). It can be seen that the mean codeword length (3) had been generalized parametrically and their bounds studied in terms of generalized measures of entropies. Here we give another generalization of (3) and study its bounds in terms of the generalized entropy of order α and type β.

Generalized coding theorems by considering different information measure under the condition of unique decipherability were investigated by several authors, see for instance the papers [Aczel and Daroczy (1975), Ebanks et al. (1998), Hooda and Bhaker (1997), Khan et al. (2005), Longo (1976), Singh et al. (2003)].

In this paper we obtain two coding theorems by using 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 Havrda-Charvat (1967), Sharma-Mittal (1975) and Tsallis entropy (1988).

2. Coding Theorems

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

$$H_{\alpha}^{\beta}(p_1,...,p_N) = \frac{1}{\beta - 1}\left[1 - \left(\sum_{i=1}^{N} p_i^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}}\right] \quad ((p_1,...,p_N) \in \Delta_N). \tag{6}$$

It is also characterized by Sharma and Mittal (1975).

Remarks:

(i) When $\beta = \alpha$, then the information measure H_{α}^{β} reduces to

$$H^{\alpha}(p_1,...,p_N) = \frac{1}{\alpha - 1}\left[1 - \sum_{i=1}^{N} p_i^{\alpha}\right] \quad ((p_1,...,p_N) \in \Delta_N). \tag{7}$$

The measure (7) was characterized by Havrda-Charvat (1967), Vajda (1968), Daroczy (1970) and Tsallis (1988) through different approaches.

(ii) When $\beta = \alpha$ and $\alpha \to 1$, then the information measure H_{α}^{β} reduces to the Shannon entropy,

$$H(P) = -\sum p_i \log p_i . \tag{8}$$

Here $log = ln$, the natural logarithm.

(iii) When $\beta \to 1$, then the information measure H_{α}^{β} reduces to a constant times Renyi's (1961) entropy,

$$(ln\, D)^{-1} H_{\alpha}(P) = \frac{1}{1-\alpha} \log_D \sum_{i=1}^{N} p_i^{\alpha} . \tag{9}$$

Definition: Let $N \in \mathbb{N}$, $\alpha > 1$, $\beta > 0(\neq 1)$, $\beta \neq \alpha$ be arbitrarily fixed, then the mean length L_{α}^{β} corresponding to the generalized information measure H_{α}^{β} is given by the formula

$$L_\alpha^\beta = \frac{1}{\beta-1}\left[1-\left(\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)}\right)^{\left(\frac{\beta-1}{\alpha-1}\right)}\right], \tag{10}$$

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

$$\sum_{i=1}^{N} p_i^{(\alpha-1)} D^{-n_i\alpha} \le \sum_{i=1}^{N} p_i^{\alpha}. \tag{11}$$

Since (11) reduces to the Kraft inequality (2) when $\alpha = 1$, it is called the generalized Kraft inequality and codes obtained under this generalized inequality are called personal codes.

Remarks:

(i) If $\alpha \to 1$, then

$$\lim_{\alpha\to 1} L_\alpha^\beta = \frac{1}{\beta-1}\left[1 - D^{(1-\beta)\sum_{i=1}^{N} p_i n_i}\right]; \quad \beta \ne 1$$

(ii) If $\beta = \alpha$ and $\alpha \to 1$, then it becomes an ordinary mean length due to Shannon times $ln\, D$

$$L = \sum_{i=1}^{N} n_i p_i \, lnD.$$

Applications of Holder's Inequality in Coding Theory

In the following theorem, we find the lower bound for L_α^β.

Theorem 2.1. Let $\alpha > 1, \beta > 0(\ne 1), \beta \ne \alpha$ be arbitrarily fixed real numbers, then for all integers $D > 1$ inequality

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

is fulfilled. Furthermore, equality holds if and only if

$$n_i = -\log_D (p_i)^{\frac{1}{\alpha}}. \tag{13}$$

Proof:

The reverse Holder inequality says: If $N \in \mathbb{N}$, $\gamma > 1$ and $x_1,\ldots,x_N$, $y_1,\ldots,y_N$ are positive real numbers then

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

Let $\gamma = \dfrac{\alpha}{\alpha-1}$, $x_i = p_i^{\frac{\alpha^2-\alpha+1}{\alpha-1}} D^{-n_i\alpha}$, $y_i = p_i^{\frac{\alpha}{1-\alpha}}$ $(i = 1,\ldots,N)$.

Putting these values into (14), we get

$$\left(\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)}\right)^{\frac{\alpha}{\alpha-1}} \left(\sum_{i=1}^{N} p_i^{\alpha}\right)^{\frac{1}{1-\alpha}} \le \sum_{i=1}^{N} p_i^{(\alpha-1)} D^{-n_i\alpha} \le \sum_{i=1}^{N} p_i^{\alpha}$$

where we used (11), too. This implies however that

$$\left(\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)}\right)^{\frac{1}{\alpha-1}} \le \left(\sum_{i=1}^{N} p_i^{\alpha}\right)^{\frac{1}{\alpha-1}} . \tag{15}$$

Now consider two cases:

(i) When $0 < \beta < 1$, then raising both sides of (15) to the power $(\beta - 1)$, we have

$$\left(\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)}\right)^{\frac{\beta-1}{\alpha-1}} \ge \left(\sum_{i=1}^{N} p_i^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} . \tag{16}$$

We obtain the result (12) after simplification,

i.e., $L_{\alpha}^{\beta} \ge H_{\alpha}^{\beta}(P)$.

(ii) When $\beta > 1$, then raising power $(\beta - 1)$ to both sides of (15), we have

$$\left(\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)}\right)^{\frac{\beta-1}{\alpha-1}} \le \left(\sum_{i=1}^{N} p_i^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} . \tag{17}$$

We obtain the result (12) after simplification,

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

It is clear that the equality in (12) is valid if $n_i = -\log_D (p_i)^{\frac{1}{\alpha}}$. The necessity of this condition for equality in (12) follows from the condition for equality in Holder's inequality: In the case of the reverse Holder's equality given above, equality holds if and only if for some c,

$$x_i^{(\gamma-1)} = c y_i^{-\gamma}, \; i = 1,2,...,N. \tag{18}$$

Plugging this condition into our situation, with the x_i, y_i, and γ as specified, and using the fact that the $\sum_{i=1}^N p_i = 1$, the necessity is proven.

In the following theorem, we give an upper bound for L_α^β in terms of $H_\alpha^\beta(P)$.

Theorem 2.2. For α and β as in Theorem 2.1, there exist positive integers n_i satisfying (11) such that

$$L_\alpha^\beta < D^{1-\beta} \mathrm{H}_\alpha^\beta(P) + \frac{1-D^{1-\beta}}{\beta-1}. \tag{19}$$

Proof:

Let n_i be the positive integer satisfying the inequalities

$$-\log_D (p_i)^{\frac{1}{\alpha}} \leq n_i < -\log_D (p_i)^{\frac{1}{\alpha}} + 1. \tag{20}$$

It is easy to see that the sequence $\{n_i\}, \; i = 1,2,...,N$ thus defined, satisfies (11).
Now, from the right of inequality (20), we have

$$n_i < -\log_D (p_i)^{\frac{1}{\alpha}} + 1$$

$$\Rightarrow D^{-n_i(\alpha-1)} > D^{(1-\alpha)} p_i^{\frac{\alpha-1}{\alpha}}. \tag{21}$$

Multiplying (21) throughout by $p_i^{\frac{\alpha^2-\alpha+1}{\alpha}}$ and then summing up from $i = 1$ to $i = N$,

$$\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)} > D^{(1-\alpha)} \sum_{i=1}^{N} p_i^{\alpha} . \qquad (22)$$

Now consider two cases:

(i) Let $0<\beta<1$, we obtain inequality (19) after simplification:

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

$$i.e., L_\alpha^\beta < D^{1-\beta} \mathrm{H}_\alpha^\beta(P) + \frac{1-D^{1-\beta}}{\beta-1} .$$

(ii) Let $\beta>1$, we obtain inequality (19) after simplification:

$$\frac{1}{\beta-1}\left[1-\left(\sum_{i=1}^{N} p_i^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{n_i(1-\alpha)}\right)^{\left(\frac{\beta-1}{\alpha-1}\right)}\right] < D^{1-\beta}\left\{\frac{1}{\beta-1}\left[1-\left(\sum_{i=1}^{N} p_i^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}}\right]\right\} + \frac{1-D^{1-\beta}}{\beta-1} ,$$

$$i.e., L_\alpha^\beta < D^{1-\beta} \mathrm{H}_\alpha^\beta(P) + \frac{1-D^{1-\beta}}{\beta-1} .$$

Acknowledgement

We are thankful to the anonymous referee for their valuable comments and suggestions.

REFERENCES

Aczel, J. and Daroczy, Z. (1963). Uber Verallgemeinerte quasilineare mittelwerte, die mit Gewichtsfunktionen gebildet sind, Publ. Math. Debrecen, Vol. **10**, pp. 171-190.

Aczel, J. and Daroczy, Z. (1975). On measures of information and their characterizations. Mathematics in Science and Engineering, Vol. **115**, Academic Press, New York, London.

Campbell, L.L. (1965). A coding theorem and Renyi's entropy, Information and Control, Vol. **8**, pp. 423-429.

Daroczy, Z. (1970). Generalized Information Functions, Information and Control, Vol. **16**, pp. 36-51.

Ebanks, B., Sahoo, P. and Sander, W. (1998). Characterizations of information measures, World Scientific Publishing Co., Inc., River Edge, NJ.

Feinstein, A. (1956). Foundation of Information Theory, McGraw Hill, New York.

Havrda, J.F. and Charvat, F. (1967). Qualification Method of Classification Process, the concept of structural α -entropy, Kybernetika, Vol. **3**, pp. 30-35.

Hooda, D.S. and Bhaker, U.S. (1997). A generalized `useful' information measure and coding theorems, Soochow J. Math., Vol. **23**, pp. 53-62.

Jelinek, F. (1980). Buffer overflow in variable lengths coding of fixed rate sources, IEEE, Vol. **3**, pp. 490-501.

Kapur, J.N. (1967). Generalized entropy of order α and type β. Maths Seminar, Delhi, Vol. **4**, pp. 78-94.

Khan, A.B., Bhat, B.A. and Pirzada, S. (2005). Some Results on a Generalized Useful Information Measure, Journal of Inequalities in Pure and Applied Mathematics, Vol. **6,** No. **4**, Art. 117.

Kieffer, J.C. (1979). Variable lengths source coding with a cost depending only on the codeword length, Information and Control, Vol. **41**, pp. 136-146.

Longo, G. (1976). A Noiseless Coding Theorem for Sources Having Utilities, Siam J. Appl. Math., Vol. **30**, pp. 739-748.

Renyi, A. (1961). On Measure of entropy and information, Proc. 4th Berkeley Symp. Maths. Stat. Prob., Vol. **1**, pp. 547-561.

Shannon, C.E. (1948). A Mathematical Theory of Communication, Bell System Tech. J., Vol. **27**, pp. 379-423, 623-656.

Sharma, B.D. and Mittal, D.P. (1975). New non-additive measures of entropy for discrete Probability distributions, J. Math. Sci., Vol. **10**, pp. 28-40.

Singh, R.P., Kumar, R. and Tuteja, R.K. (2003). Application of Hölder's Inequality in Information Theory, Information Sciences, Vol. **152**, pp. 145-154.

Tsallis, C. (1988). Possible Generalization of Boltzmann Gibbs Statistics, J. Stat. Phy., Vol. **52**, pp. 479.

Vajda, I. (1968). Axioms for α -entropy of a generalized probability scheme, Kybernetika, (Prague), Vol. **4**, pp. 105-112.