International Journal of Advance Research in Science and Engineering
Volume No.06, Special Issue No.(01), Nov 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

A Non-Additive Generalized Inaccuracy Measure and Its Application in Coding Theory

Sunit kumar[1], Satish Kumar[2]

[1]Department of Mathematics

Guru Nanak Institute of Technology, Mullana

[2]Department of Mathematics,

College of Natural and Computational Science,

University of Gondar, Ethiopia.

ABSTRACT

A relation between Shannon entropy and Kerridge inaccuracy, which is known as Shannon inequality, is well known in information theory. In this communication, first we generalized Shannon inequality and then given its application in coding theory.

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

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

I. INTRODUCTION

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

$$\Delta_n = \left\{ P = (p_1, p_2, ..., p_n);\ 0 < p_k \le 1,\ 0 < \sum_{k=1}^{n} p_k \le 1 \right\},$$

$$\Delta_n^* = \left\{ P = (p_1, p_2, ..., p_n);\ 0 < p_k \le 1,\ \sum_{k=1}^{n} p_k = 1 \right\},$$

denote the sets of n-components, $n \ge 2$, generalized probability distributions and complete probability distributions respectively.

For $(p_1, p_2, ..., p_n) = P \in \Delta_n, (q_1, q_2, ..., q_n) = Q \in \Delta_n$, we define a non-additive measure of inaccuracy, denoted by $H(P,Q;\alpha)$ as

$$H(P,Q;\alpha) = \frac{1}{1-\alpha}\left[\left(\frac{\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} q_k^{\left(\frac{\alpha-1}{\alpha}\right)}}{\sum_{k=1}^{n} p_k}\right) - 1\right] \qquad \alpha > 1. \qquad (1.1)$$

$$= -\frac{\sum_{k=1}^{n} p_k \log_2 q_k}{\sum_{k=1}^{n} p_k}; \quad \alpha \to 1.$$

If $P = Q$, then $H(P,Q;\alpha)$ reduces to non-additive entropy.

$$i.e., \; H(P;\alpha) = \frac{1}{1-\alpha}\left[\left(\frac{\sum_{k=1}^{n} p_k^{\alpha}}{\sum_{k=1}^{n} p_k}\right) - 1\right] \; \alpha > 0(\neq 1). \qquad (1.2)$$

$$= -\frac{\sum_{k=1}^{n} p_k \log_2 p_k}{\sum_{k=1}^{n} p_k}; \quad \alpha \to 1.$$

The entropy (1.2) was first of all characterized by Havrda and Charvat [7]. Later on, Daroczy [5] and M. Behara and P. Nath [2,3] studied this entropy. Vajda [20] also characterized this entropy for finite discrete generalized probability distributions. Tsallis's [19] gave its applications in physics for $P \in \Delta_n^*$, and $\alpha \to 1$, $H(P;\alpha)$ reduces to Shannon [17] entropy.

$$i.e., \; H(P) = -\sum_{k=1}^{n} p_k \log_D p_k. \qquad (1.3)$$

II. FORMULATION OF THE PROBLEM

For $\alpha \to 1$ and $P \in \Delta_n^*$, $Q \in \Delta_n^*$, then an important property of Kerridge's inaccuracy [9] is that

$$\mathrm{H}(P) \leq H(P,Q). \qquad (2.1)$$

equality if and only if $P = Q$. In other words, Shannon's entropy is the minimum value of Kerridge's inaccuracy. If $P \in \Delta_n$, $Q \in \Delta_n$, then (2.1) is no longer necessarily true. Also, the corresponding inequality

$$H(P;\alpha) \leq H(P,Q;\alpha) \qquad (2.2)$$

are not necessarily true even for generalized probability distributions. Hence, it is natural to ask the following question:"For generalized probability distributions, what are the quantity the minimum values of which are $H(P;\alpha)$?" We give below an answer to the above question separately for $H(P;\alpha)$ by dividing the discussion into two parts (i) $\alpha \to 1$ and (ii) $1 \neq \alpha$. Also we shall assume that $n \geq 2$, because the problem is trivial for $n = 1$.

Case 1. Let $\alpha \to 1$. If $P \in \Delta_n^*, Q \in \Delta_n^*$, then as remarked earlier (2.1) is true. For $P \in \Delta_n, Q \in \Delta_n$, it can be easily seen by using Jenson's inequality that (2.1) is true if $\sum_{k=1}^n p_k \geq \sum_{k=1}^n q_k$, equality in (2.1) holding if and only if

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \ldots = \frac{p_n}{q_n} = \frac{\sum_{k=1}^n p_k}{\sum_{k=1}^n q_k}.$$

Case 2. Let $1 \neq \alpha$. Since (2.2) is not necessarily true, we need an inequality

$$\sum_{k=1}^n p_k^{\alpha-1} q_k \leq \sum_{k=1}^n p_k^{\alpha}; \ \alpha > 1 \ . \tag{2.3}$$

such that $H(P;\alpha) \leq H(P,Q;\alpha)$ and equality holds if and only if $P = Q$.

Remark: If $\alpha = 1$, the inequality (2.3) reduces to case 1.

Since $\alpha > 1$, by reverse Hölder inequality, that is, if $n = 2,3,\ldots,$ $\gamma > 1$ and $x_1,\ldots,x_n,y_1,\ldots,y_n$ are positive real numbers then

$$\left(\sum_{k=1}^n x_k^{\frac{1}{\gamma}}\right)^{\gamma} \cdot \left(\sum_{k=1}^n y_k^{-\frac{1}{(\gamma-1)}}\right)^{-(\gamma-1)} \leq \sum_{k=1}^n x_k y_k. \tag{2.4}$$

Let $\gamma = \dfrac{\alpha}{\alpha-1}$, $x_k = p_k^{\frac{\alpha^2-\alpha+1}{\alpha-1}} q_k$, $y_k = p_k^{\frac{\alpha}{1-\alpha}}$ $(k = 1,2,3,\ldots n)$.

Putting these values into (2.4), we get

$$\left(\sum_{k=1}^n p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} q_k^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}} \left(\sum_{k=1}^n p_k^{\alpha}\right)^{\frac{1}{1-\alpha}} \leq \sum_{k=1}^n p_k^{\alpha-1} q_k \leq \sum_{k=1}^n p_k^{\alpha},$$

where we used (2.3), too. This implies however that

$$\left(\sum_{k=1}^n p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} q_k^{\frac{\alpha-1}{\alpha}}\right) \leq \left(\sum_{k=1}^n p_k^{\alpha}\right). \tag{2.5}$$

Or

$$\left(\frac{\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} q_k^{\frac{\alpha-1}{\alpha}}}{\sum_{k=1}^{n} p_k}\right) \le \left(\frac{\sum_{k=1}^{n} p_k^{\alpha}}{\sum_{k=1}^{n} p_k}\right) \qquad (2.6)$$

using (2.6) and the fact that $\alpha > 1$, , we get (2.2) .

Particular's case: If $\alpha = 1$, then (2.2) becomes

$$H(P) \le H(P,Q) \quad ,$$

which is Kerridge's inaccuracy [9] .

3. Mean Codeword Length and their bounds.

We will now give an application of inequality (2.2) in coding theory for

$$\Delta_n^* = \left\{ P = (p_1, p_2, ..., p_n);\ 0 < p_k \le 1,\ \sum_{k=1}^{n} p_k = 1 \right\} \quad .$$

Let a finite set of n input symbols

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

be encoded using alphabet of D symbols, then it has been shown by Feinstien [6] that there is a uniquely decipherable code with lengths $N_1, N_2, ..., N_n$ if and only if the Kraft inequality holds that is,

$$\sum_{k=1}^{n} D^{-N_k} \le 1. \qquad (3.1)$$

Where D is the size of code alphabet.

Furthermore, if

a code satisfying (3.1), the inequality

$$L \ge H(P) \qquad (3.3)$$

$$L = \sum_{k=1}^{n} N_k p_k \qquad (3.2)$$

is the average codeword length, then for

is also fulfilled and equality holds if and only if

$$N_k = -\log_D(p_k) \ (k = 1, ..., n), \qquad (3.4)$$

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. [15]), 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

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

Kieffer [11] 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 [8] 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].

It may be seen that the mean codeword length (3.2) had been generalized parametrically by Campbell [4] and their bounds had been studied in terms of generalized measures of entropies. Here we give another generalization of (3.2) and study its bounds in terms of 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 [6, 10, 12, 13, 14, 16].

An investigation is carried out concerning discrete memoryless sources possessing an additional parameter α, which seems to be significant in problem of storage and transmission (see [8], [11] and [12]).

In this section we study a coding theorem 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 [1] entropy, which is used in physics.

Definition: Let $n \in \mathbb{N}, \alpha > 0(\neq 1)$ be arbitrarily fixed, then the mean length $L(\alpha)$ corresponding to the generalized information measure $H(P;\alpha)$ is given by the formula

$$L(\alpha,\beta) = \frac{1}{1-\alpha}\left[\left(\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{N_k\left(\frac{1-\alpha}{\alpha}\right)}\right) - 1\right], \tag{3.5}$$

where $P = (p_1,..., p_n) \in \Delta_n^*$ and $D, N_1, N_2,..., N_n$ are positive integers so that

$$\sum_{k=1}^{n} p_k^{\alpha-1} D^{-N_k} \leq \sum_{k=1}^{n} p_k^\alpha. \tag{3.6}$$

Since (3.6) reduces to Kraft inequality when $\alpha = 1$, therefore it is called generalized Kraft inequality and codes

International Journal of Advance Research in Science and Engineering
Volume No.06, Special Issue No.(01), Nov 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

obtained under this generalized inequality are called personal codes.

Theorem 1. Let $n \in \mathbb{N}, \alpha > 1.$ be arbitrarily fixed. Then there exist code length $N_1,..., N_n$ so that

$$H(P;\alpha) \le L(\alpha) < D^{\frac{1-\alpha}{\alpha}} H(P;\alpha) + \frac{1 - D^{\frac{1-\alpha}{\alpha}}}{1-\alpha} \tag{3.7}$$

holds under the condition (3.6) and equality holds if and only if

$$N_k = -\log_D p_k \quad ; \quad k = 1,2,..., n. \tag{3.8}$$

Where $H(P;\alpha)$ and $L(\alpha)$ are given by (1.2) and (3.5) respectively.

Proof: First of all we shall prove the lower bound of $L(\alpha)$.

By reverse Hölder inequality, that is, if $n = 2,3,...,$ $\gamma > 1$ and $x_1,..., x_n, y_1,..., y_n$ are positive real numbers then

$$\left(\sum_{k=1}^{n} x_k^{\frac{1}{\gamma}}\right)^{\gamma} \cdot \left(\sum_{k=1}^{n} y_k^{-\frac{1}{(\gamma-1)}}\right)^{-(\gamma-1)} \le \sum_{k=1}^{n} x_k y_k . \tag{3.9}$$

Let $\gamma = \frac{\alpha}{\alpha - 1}$, $x_k = p_k^{\frac{\alpha^2-\alpha+1}{\alpha-1}} D^{-N_k}$, $y_k = p_k^{\frac{\alpha}{1-\alpha}}$ $(k = 1,2,3,...n)$.

Putting these values into (3.9), we get

$$\left(\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{-N_k\left(\frac{\alpha-1}{\alpha}\right)}\right)^{\frac{\alpha}{\alpha-1}} \left(\sum_{k=1}^{n} p_k^{\alpha}\right)^{\frac{1}{1-\alpha}} \le \sum_{k=1}^{n} p_k^{\alpha-1} D^{-N_k} \le \sum_{k=1}^{n} p_k^{\alpha} \; ; \alpha > 1.$$

where we used (3.6), too. This implies however that

$$\left(\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{-N_k\left(\frac{\alpha-1}{\alpha}\right)}\right)^{\frac{\alpha}{\alpha-1}} \le \left(\sum_{k=1}^{n} p_k^{\alpha}\right)^{\frac{\alpha}{\alpha-1}} \tag{3.10}$$

For $\alpha > 1$, (3.10) becomes

$$\left(\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{-N_k\left(\frac{\alpha-1}{\alpha}\right)}\right) \le \left(\sum_{k=1}^{n} p_k^{\alpha}\right) \tag{3.11}$$

using (3.11) and the fact that $\alpha > 1$, we get

$$H(P;\alpha) \le L(\alpha) \quad . \tag{3.12}$$

From (3.8) and after simplification, we get

$$p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{-N_k\left(\frac{\alpha-1}{\alpha}\right)} = p_k^{\alpha}$$

This implies

$$\left(\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{-N_k\left(\frac{\alpha-1}{\alpha}\right)}\right) = \left(\sum_{k=1}^{n} p_k^{\alpha}\right), \qquad (3.13)$$

which gives $L(\alpha) = H(P;\alpha)$. Then equality sign holds in (3.12).

Now we will prove the inequality (3.7) for upper bound of $L(\alpha)$.

We choose the codeword lengths N_k, $k = 1,...,n$ in such a way that

$$-\log_D p_k \le N_k < -\log_D p_k + 1. \qquad (3.14)$$

is fulfilled for all $k = 1,...,n$.

From the left inequality of (3.14), we have

$$D^{-N_k} \le p_k, \qquad (3.15)$$

multiplying both sides by $p_k^{\alpha-1}$ and then taking sum over k, we get the generalized inequality (3.6). So there exists a generalized code with code lengths N_k, $k = 1,...,n$.

Since $\alpha > 1$, then (3.14) can be written as

$$p_k^{\left(\frac{\alpha-1}{\alpha}\right)} \ge D^{-N_k\left(\frac{\alpha-1}{\alpha}\right)} > p_k^{\left(\frac{\alpha-1}{\alpha}\right)} D^{-\left(\frac{\alpha-1}{\alpha}\right)} \qquad (3.16)$$

Multiplying (3.16) throughout by $p_k^{\frac{\alpha^2-\alpha+1}{\alpha}}$ and then summing up from $k = 1$ to n, we obtain inequality

$$\left(\sum_{k=1}^{n} p_k^{\alpha}\right) \ge \left(\sum_{k=1}^{n} p_k^{\frac{\alpha^2-\alpha+1}{\alpha}} D^{N_k\left(\frac{1-\alpha}{\alpha}\right)}\right) > \left(\sum_{k=1}^{n} p_k^{\alpha}\right) D^{\left(\frac{1-\alpha}{\alpha}\right)}. \qquad (3.17)$$

Since $1-\alpha < 0$ for $\alpha > 1$, we get from (3.17) the inequality (3.7).

Particular's cases:

For $\alpha \to 1$, then (3.7) becomes

$$\frac{H(P)}{\log D} \leq L < \frac{H(P)}{\log D} + 1.$$

Which is the Shannon [17] classical noiseless coding theorem.

REFERENCES

[1] C. Arndt, Information Measure-Information and its description in Science and Engineering, Springer, Berlin, 2001.

[2] M. Behara and P. Nath. Additive and Non-Additive Entropies of Finite Measurable Partitions, Probability and Information Theory II, Springer-Verlag, 102-138 (1970).

[3] M. Behara and P. Nath. An Axiomatic Characterization of Entropy of a Finite Discrete Generalized Probability Distribution, McMaster Univ. Math, report no. 40.

[4] L.L. Campbell. A coding theorem and Renyi's entropy, Information and Control 8:423-429 (1965).

[5] Z. Daroczy. Generalized Information Functions, Information and Control 16:36-51 (1970).

[6] A. Feinstein, Foundation of Information Theory, McGraw Hill, New York, 1956.

[7] J.F. Havrda and F. Charvat. Qualification Method of Classification Process, the concept of structural α-entropy, Kybernetika 3:30-35 (1967).

[8] F. Jelinek. Buffer overflow in variable lengths coding of fixed rate sources, IEEE 3:490-501 (1980).

[9] D.F. Kerridge. Inaccuracy and inference, Jour. Roy. Statist. Soc. Ser B 23:184-196 (1961).

[10] A.B. Khan, B.A. Bhat and S. Pirzada. Some Results on a Generalized Useful Information Measure, Journal of Inequalities in Pure and Applied Mathematics 6(4):Art.117 (2005).

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

[12] G. Longo. A Noiseless Coding Theorem for Sources Having Utilities, Siam J. Appl. Math. 30:739-748 (1976).

[13] Om Parkash and P.K. Sharma. Noiseless Coding Theorems Corresponding to Fuzzy Entropies, Southeast Asian Bulletin of Mathematics 27:1073-1080 (2004).

[14] R. Pereira and Gur Dial. Pseudo-generalization of Shannon inequality for Mittal's entropy and its application in coding theory, Kybernetika 20(1):73-77 (1984).

[15] A. Renyi. On Measure of entropy and information, Proc. 4th Berkeley Symp. Maths. Stat. Prob. 1:547-561 (1961).

[16] L.K. Roy. Comparison of Renyi's entropy of power distribution, ZAMM 56:217-218 (1976).

[17] C.E. Shannon. A Mathematical Theory of Communication, Bell System Tech. J. 27:379-423, 623-656 (1948).

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

[19] C. Tsallis. Possible Generalization of Boltzmann Gibbs Statistics, J. Stat. Phy. 52:479 (1988).

[20] I. Vajda. Axioms for α-entropy of a generalized probability scheme, Kybernetika (Prague) 4:105-112 (1968).