International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:5, No:3, 2011

Codes and Formulation of Appropriate Constraints via Entropy Measures

R. K. Tuli

***Abstract*—**In present communication, we have developed the suitable constraints for the given the mean codeword length and the measures of entropy. This development has proved that Renyi's entropy gives the minimum value of the log of the harmonic mean and the log of power mean. We have also developed an important relation between best 1:1 code and the uniquely decipherable code by using different measures of entropy.

***Keywords*—**Codeword, Instantaneous code, Prefix code, Uniquely decipherable code, Best one-one code, Mean codeword length

I. Introduction

IN usual practice, coding theory deals in finding the minimum value of a mean codeword length subject to a given constraint on codeword lengths. However, since the codeword lengths are integers, the minimum value always lies between two bounds and every noiseless coding theorem seeks to find these two lower bounds for a given value of mean and a given constraint. Shannon [11] has shown that the minimal expected length L of a prefix code for a random variable X satisfies the following result:

$$H(X) \le L < H(X)+1 \tag{1}$$

Where H is the entropy of the random variable. Since the set of allowed codeword lengths is the same for the uniquely decipherable and instantaneous codes, the expected codeword length L is the same for both sets of codes. If p_i is the probability of the i^{th} outcome, then Shannon [11] assigned the following codeword length to the outcome of the random variable

$$l_i = \left\lceil \log \frac{1}{p_i} \right\rceil \tag{2}$$

where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. Shannon [11] used Kraft's [6] inequality to prove

his results and this inequality plays an important role in proving a noiseless coding theorem and is uniquely determined by the condition for unique decipherability.

Although the main focus is on the class of uniquely decipherable codes, there has been some interest in the class of one-to-one codes. A one-to-one code is a code that associates a distinct codeword with each source symbol. As such, unlike uniquely decipherable codes, one-to-one codes need not possess unique decipherability. Due to Feinstein [4], one chooses codeword lengths to minimize the average code length and this minimization is done subject to the constraint that the code be uniquely decipherable. Rissanen [9] derived a lower bound for the mean code length of all one-to-one codes for a random variable with n outcomes. Savari and Naheta [10] generalized one-to-one encodings of a discrete random variable to finite code alphabets and established some upper and lower bounds for the best one-to-one codes. Some results regarding the lower bounds of expected lengths of one-one codes have been discussed by Cheng, Huang and Weidmann [3] and Cheng and Huang [2]. Some other work related with best one-one codes has been studied by Leung-Yan-Cheong and Cover [7].

While dealing with coding theory, we usually come across three entities viz means, constraints and measures of entropy and we get the following problems:

(i) Given a specific mean and a specific constraint, find the minimum value of the mean subject to the given constraint.

(ii) Given an entropy measure and a constraint, find the mean codeword length for which the given entropy measure will give the minimum value for the given constraint.

(iii) Given the mean codeword length and the measure of entropy, find a suitable constraint for which the measure of entropy will be the minimum value for the given mean codeword length.

This paper deals with the investigations of the third class of problems dealing with the development of appropriate constraints, the results of which have been shown in section II. In section III, we have developed an important relation between best 1:1 code and the uniquely decipherable code by using different measures of entropy.

II. Appropriate Constraints For The Given Values Of Mean And The Entropy Measures

In such problems, the mean codeword lengths and their minimum values have been given and we have to find the appropriate constraints.

A. Suppose we want the constraint which gives Renyi's [8] entropy as the minimum value of the log of the harmonic mean of codeword length, we proceed as under:

Harmonic mean is given by

$$H = \left(\sum_{i=1}^{n} \frac{p_i}{l_i} \right)^{-1} \tag{3}$$

R.K.Tuli is with the Department of Mathematics, S.S.M.College, Dinanagar (India) r_kumartuli@yahoo.co.in.

We want to minimize

$$\log_D H = -\log_D\left(\sum_{i=1}^{n}\frac{p_i}{l_i}\right) \quad (4)$$

subject to the following constraint

$$f(l_1,l_2,...,l_n) = k \quad (5)$$

The corresponding Lagrangian is given by

$$L = -\log_D\left(\sum_{i=1}^{n}\frac{p_i}{l_i}\right) - \lambda\left\{f-k\right\}$$

Differentiating the above equation both sides with respect to l_i and equating to zero, we get

$$\frac{\partial f}{\partial l_i} = K'\frac{p_i}{l_i^2} \quad (6)$$

where

$$K' = \frac{1}{\lambda \log_D \sum_{i=1}^{n}\frac{p_i}{l_i}}$$

Equation (6) to have the solution, we must have

$$l_i = \frac{1}{p_i} \quad (7)$$

that is,

$$p_i = \frac{1}{l_i} \quad (8)$$

Substituting equation (8) in (6), we get

$$\frac{\partial f}{\partial l_i} = K' l_i^{-3} \quad (9)$$

Integrating equation (9), we get

$$f(l_1,l_2,...,l_n) = \sum_{i=1}^{n}\frac{K''}{l_i^2} + A \quad (10)$$

where $K'' = -\frac{K'}{2}$ and A is some arbitrary constant.

From equations (5) and (10), we have

$$\sum_{i=1}^{n}\frac{K''}{l_i^2} = K'''$$

where $K''' = k - A$ is a constant.

or

$$\sum_{i=1}^{n}\frac{1}{l_i^2} = C$$

where C is another constant given by

$$C = \frac{K'''}{K''}$$

Thus, the required constraint to be found is

$$. \sum_{i=1}^{n}\frac{1}{l_i^2} = C$$

If, instead of (9), we want the solution $l_i = \frac{a}{p_i}$, then the constraint will be $\sum_{i=1}^{n}\frac{a}{l_i^2} = C$ and the minimum value of (6) is $-\log_D\left(\sum_{i=1}^{n} p_i.\frac{p_i}{a}\right) = \log_D a - \log_D\left(\sum_{i=1}^{n} p_i^{\,2}\right)$

$$= \log_D a + R_2(P)$$

Where $R_2(P)$ is Renyi's [8] entropy of order 2. Thus, the minimum value of log of harmonic mean lies between $R_2(P)$ and $R_2(P)+1$ if a lies between 1 and D.

B. Suppose we want the constraint which gives Renyi's [8] entropy as the minimum value of the log of the power mean of codeword length, we proceed as under:

Power mean is given by

$$P = \left(\sum_{i=1}^{n} p_i l_i^{\,r}\right)^{\frac{1}{r}} \quad (11)$$

We want to minimize

$$\log_D P = \frac{1}{r}\log_D\left(\sum_{i=1}^{n} p_i l_i^{\,r}\right)$$

that is,

$$\log_D P = \frac{1}{1-(1-r)}\log_D\left(\sum_{i=1}^{n} p_i l_i^{\,r}\right) \quad (12)$$

subject to the following constraint

$$h(l_1,l_2,...,l_n) = k \quad (13)$$

To solve the problem, we apply Lagrange's method of maximum multipliers. The corresponding Lagrangian is given by

$$L = \frac{1}{r}\log_D\left(\sum_{i=1}^{n} p_i l_i^{\,r}\right) - K\left\{h-k\right\}$$

Differentiating the above equation both sides with respect to l_i and equating to zero, we get:

$$\frac{\partial h}{\partial l_i} = K' p_i l_i^{\,r-1} \quad (14)$$

where

$$K' = \frac{1}{K \log_D \sum_{i=1}^{n} p_i l_i^{\,r}}$$

Equation (14) to have the solution, we must have

$$l_i = \frac{1}{p_i} \quad (15)$$

that is,

$$p_i = \frac{1}{l_i} \tag{16}$$

Substituting equation (16) in (14), we get

$$\frac{\partial h}{\partial l_i} = K' l_i^{r-2} \tag{17}$$

Integrating equation (17), we get

$$h(l_1, l_2, ..., l_n) = \sum_{i=1}^{n} K'' l_i^{r-1} + A \tag{18}$$

where

$$K'' = \frac{K'}{r-1}$$

From equations (13) and (18), we have

$$\sum_{i=1}^{n} K'' l_i^{r-1} = K'''$$

where $K''' = k - A$ is some constant.

or $\sum_{i=1}^{n} l_i^{r-1} = C$ where C is another constant

Thus the required constraint to be found is $\sum_{i=1}^{n} l_i^{r-1} = C$.

If, instead of (15), we want the solution $l_i = \frac{a}{p_i}$, then the constraint will be $\sum_{i=1}^{n} a l_i^{r-1} = C$ and the minimum value of (12) is given by

$$\log_D\left(\sum_{i=1}^{n} p_i \cdot \left(\frac{a}{p_i}\right)^r\right)$$

$$= \frac{1}{1-(1-r)} \log_D a^r + \frac{1}{1-(1-r)} \log_D \left(\sum_{i=1}^{n} p_i^{1-r}\right)$$

$$= \log_D a + R_{(1-r)}(P)$$

where $R_{(1-r)}(P)$ is Renyi's [8] entropy of order $1-r$.

Thus, the minimum value of log of power mean lies between $R_{(1-r)}(P)$ and $R_{(1-r)}(P)+1$ if a lies between 1 and D.

III. Relation Between Best 1:1 Code And The Uniquely Decipherable Code

In this section, we develop an important relation between best 1:1 code and the uniquely decipherable code by using different measures of entropy.

A. A mean codeword length of order t for the best 1:1 code and Renyi's [8] entropy of type α

Campbell [1] introduced a generalized mean codeword length of order t given by

$$L_{UD}(t) = \frac{1}{t} \log_D \left(\sum_{i=1}^{n} p_i D^{t l_i}\right), \ (0 < t < \infty) \tag{19}$$

Where, D represents the size of the code alphabet.

Also,

$$t = \frac{1-\alpha}{\alpha}$$

and

l_i , $i = 1, 2, .., n$ are the lengths of the codewords associated with the value of X. He showed that for uniquely decipherable codes, lower bound for $L_{UD}(t)$ lies $R_\alpha(P)$ between $R_\alpha(P)+1$ and

that is,

$$R_\alpha(P) \le L_{UD}(t) < R_\alpha(P) + 1 \tag{20}$$

where

$$R_\alpha(P) = \frac{1}{1-\alpha} \log_D \sum_{i=1}^{n} p_i^\alpha \ , \alpha > 0, \alpha \ne 1 \tag{21}$$

is the Renyi's [8] entropy of type α .

Let the probability distribution of the random variable X taking finite number of values be

$$X = \begin{pmatrix} x_1 & x_2 & .... & x_n \\ p_1 & p_2 & .... & p_n \end{pmatrix}$$

Without any loss of generality, we can assume that $p_1 \ge p_2 \ge ..... \ge p_n$. Let l_i , $i = 1, 2, .., n$ be the lengths of the codewords in the best 1:1 code for encoding the random variable X , where l_i is the length of the codeword assigned to x_i .

Remark: The set of available codewords is $\{0, 1, 00, 01, 10, 11, 000, 001, ...\}$. It is clear that the best 1:1 code must have $l_1 \le l_2 \le l_3 \le ....$ Thus, by Inspection, wehaveprecisely

$$l_1 = 1, \ l_2 = 1, \ l_3 = 2, ..., \ l_i = \left\lceil \log_D \left(\frac{i}{2} + 1\right) \right\rceil$$

Where, $\lceil x \rceil$ denotes the smallest integer greater than or equal to x .

We consider the generalized mean codeword length of order t for the best 1:1 code, that is

$$L_{1:1}(t) = \frac{1}{t} \log_D \left(\sum_{i=1}^{n} p_i D^{t l_i}\right)$$

$$= \frac{1}{t} \log_D \left(\sum_{i=1}^{n} p_i D^{t \left\lceil \log_D \left(\frac{i}{2}+1\right)\right\rceil}\right) \tag{22}$$

and find a lower bound to this. In fact, we prove the following result:

Theorem-I: For $R_\alpha(P)$ as given in equation (21), the following estimates hold

$$L_{1:1}(t) \geq R_\alpha(P) - \log_D\left(\sum_{i=1}^{n} \frac{2}{i+2}\right) \tag{23}$$

and moreover,

$$L_{1:1}(t) \geq L_{UD}(t) - 2 - \log_D\left(\sum_{i=1}^{n} \frac{1}{i+2}\right) \tag{24}$$

Proof: From equation (22), we can have

$$L_{1:1}(t) \geq \frac{1}{t}\log_D\left(\sum_{i=1}^{n} p_i\left(\frac{i}{2}+1\right)^t\right)$$

Now

$$R_\alpha(P) - L_{1:1}(t)$$

$$\leq \frac{1+t}{t}\log_D\left(\sum_{i=1}^{n} p_i^{\frac{1}{t+1}}\right) - \frac{1}{t}\log_D\left(\sum_{i=1}^{n} p_i\left(\frac{i}{2}+1\right)^t\right)$$

$$= \log_D\left[\left(\sum_{i=1}^{n}\left(p_i^{\frac{1}{t}}\right)^{\frac{t}{t+1}}\right)^{\frac{1+t}{t}} \cdot \left(\sum_{i=1}^{n}\left(p_i^{-\frac{1}{t}}\left(\frac{i}{2}+1\right)^{-1}\right)^{-t}\right)^{-\frac{1}{t}}\right]$$

$$\leq \log_D\left[\sum_{i=1}^{n}\left(p_i^{\frac{1}{t}} p_i^{-\frac{1}{t}}\left(\frac{i}{2}+1\right)^{-1}\right)\right]$$

(Using Holder's inequality)

$$= \log_D\left[\sum_{i=1}^{n} \frac{2}{i+2}\right]$$

Thus, we have

$$L_{1:1}(t) \geq R_\alpha(P) - \log_D\left(\sum_{i=1}^{n} \frac{2}{i+2}\right)$$

Now, from equation (20), we have

$$R_\alpha(P) \leq L_{UD}(t) < R_\alpha(P) + 1$$

Thus, we have

$$L_{UD}(t) - L_{1:1}(t) < 1 + R_\alpha(P) - L_{1:1}(t)$$

$$\leq 1 + \log\left(\sum_{i=1}^{n} \frac{2}{i+2}\right)$$

$$= 2 + \log\left(\sum_{i=1}^{n} \frac{1}{i+2}\right)$$

Which gives equation (24).

B. A 2-parameter exponentiated mean codeword length of order α and type β for the best 1:1 code and Kapur's [5] additive measure of entropy

Kapur [5] introduced a mean codeword length of order α and type β, given by:

$$L_{\alpha,\beta} = \frac{1}{\alpha-1}\log_D\left(\frac{\sum_{i=1}^{n} p_i^{\beta} D^{(\alpha-1)l_i}}{\sum_{i=1}^{n} p_i^{\beta}}\right)$$

Where, D represents the size of the code alphabet, and l_i , $i = 1,2,..,n$ are the lengths of the codewords associated with the value of X. He showed that for uniquely decipherable codes, lower bound for $L_{\alpha,\beta}$ given by

$$L_{UD}(t) = \frac{1}{t}\log_D\left(\frac{\sum_{i=1}^{n} p_i^{\beta} D^{tl_i}}{\sum_{i=1}^{n} p_i^{\beta}}\right) \tag{25}$$

Where, $t = \alpha - 1$ and this lower bound lies between $E_{\alpha,\beta}(P)$ and $E_{\alpha,\beta}(P)+1$, that is,

$$E_{\alpha,\beta}(P) \leq L_{UD}(t) < E_{\alpha,\beta}(P) + 1 \tag{26}$$

Where

$$E_{\alpha,\beta}(P) = \frac{1}{\alpha-1}\log_D\left(\frac{\left(\sum_{i=1}^{n} p_i^{\frac{\beta}{\alpha}}\right)^{\alpha}}{\sum_{i=1}^{n} p_i^{\beta}}\right) \tag{27}$$

is a two parameter additive measure of entropy given by Kapur [5].

We consider 2-parameter exponentiated mean codeword length of order t and type β for the best 1:1 code, that is

$$L_{1:1}(t) = \frac{1}{t}\log_D\left(\frac{\sum_{i=1}^{n} p_i^{\beta} D^{t\left[\log_D\left(\frac{i}{2}+1\right)\right]}}{\sum_{i=1}^{n} p_i^{\beta}}\right) \tag{28}$$

and find a lower bound to this. In fact, we prove the following result:

Theorem-II: For $E_{\alpha,\beta}(P)$ as given in equation (27), the following estimates hold

$$L_{1:1}(t) \geq E_{\alpha,\beta}(P) - \log_D\left(\sum_{i=1}^{n} \frac{2}{i+2}\right) \tag{29}$$

and moreover,

$$L_{1:1}(t) \geq L_{UD}(t) - 2 - \log_D\left(\sum_{i=1}^{n} \frac{1}{i+2}\right) \tag{30}$$

Proof. From equation (28), we can have

$$L_{1:1}(t) \geq \frac{1}{t}\log_D\left(\frac{\sum_{i=1}^{n} p_i^{\beta}\left(\frac{i}{2}+1\right)^t}{\sum_{i=1}^{n} p_i^{\beta}}\right)$$

Now

$$E_{\alpha,\beta}(P) - L_{1:1}(t)$$

$$\leq \frac{1}{t}\log_D\left(\frac{\left(\sum_{i=1}^{n} p_i^{\frac{\beta}{t+1}}\right)^{t+1}}{\sum_{i=1}^{n} p_i^{\beta}}\right) - \frac{1}{t}\log_D\left(\frac{\sum_{i=1}^{n} p_i^{\beta}\left(\frac{i}{2}+1\right)^t}{\sum_{i=1}^{n} p_i^{\beta}}\right)$$

$$= \log_D\left(\left(\sum_{i=1}^{n} p_i^{\frac{\beta}{t+1}}\right)^{\frac{t+1}{t}}\left(\sum_{i=1}^{n} p_i^{\beta}\left(\frac{i}{2}+1\right)^t\right)^{\frac{-1}{t}}\right)$$

$$= \log_D\left(\left(\sum_{i=1}^{n}\left(p_i^{\frac{\beta}{t}}\right)^{\frac{t}{t+1}}\right)^{\frac{t+1}{t}}\left(\sum_{i=1}^{n}\left(p_i^{-\frac{\beta}{t}}\left(\frac{i}{2}+1\right)^{-1}\right)^{-t}\right)^{\frac{-1}{t}}\right)$$

$$\leq \log_D\left(\sum_{i=1}^{n} p_i^{\frac{\beta}{t}} p_i^{-\frac{\beta}{t}}\left(\frac{i}{2}+1\right)^{-1}\right)$$

(Using Holder's inequality)

$$= \log_D\left[\sum_{i=1}^{n}\frac{2}{i+2}\right]$$

Thus, we have

$$L_{1:1}(t) \geq R_{\alpha}(P) - \log_D\left(\sum_{i=1}^{n}\frac{2}{i+2}\right)$$

Now, from equation (26), we get the following result:

$$E_{\alpha,\beta}(P) \leq L_{UD}(t) < E_{\alpha,\beta}(P) + 1$$

Thus, we have

$$L_{UD}(t) - L_{1:1}(t) < 1 + E_{\alpha,\beta}(P) - L_{1:1}(t)$$

$$\leq 1 + \log\left(\sum_{i=1}^{n}\frac{2}{i+2}\right)$$

$$= 2 + \log\left(\sum_{i=1}^{n}\frac{1}{i+2}\right)$$

which gives equation (30).

REFERENCES

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

[2] Cheng, J. and Huang, T. K. (2006): “New lower and upper bounds on the expected length of optimal one-to-one codes”, Proceedings of the Data Compression Conference, 43-52.

[3] Cheng, J., Huang, T. K. and Weidmann, C. (2007): “New bounds on the expected length of optimal one-to-one codes”, IEEE Trans. Inform. Theory, 53(5), 1884-1895.

[4] Feinstein, A. (1958): “Foundations of Information Theory”, McGraw-Hill, New York.

[5] Kapur, J.N. (1995): “Measures of Information and Their Applications”, Wiley Eastern, New York.

[6] Kraft, L. G. (1949): “A Device for Quantizing Grouping and Coding Amplitude Modulated Pulses”, M.S. Thesis, Electrical Engineering Department, MIT.

[7] Leung-Yan-Cheong, S. K. and Cover, T. M. (1978): “Some equivalences between Shannon entropy and Kolmogrov complexity”, IEEE Trans. Inform. Theory, 24, 331-338.

[8] Renyi, A. (1961): “On measures of entropy and information”, Proceedings 4th Berkeley Symposium on Mathematical Statistics and Probability, 1, 547-561.

[9] Rissanen, J. (1992): “Tight lower bounds for optimal code length”, IEEE Trans. Inform. Theory, 28(2), 348-349.

[10] Savari, S. A. and Naheta, A. (2004): “Bounds on the expected cost of one-to-one codes”, Proc. IEEE Int. Symp. Information Theory, 94.

[11] Shannon, C. E. (1948): “A mathematical theory of communication”, Bell System Tech J., 27, 379-423.