International Journal on Emerging Technologies **10**(2): 446-449(2019)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Applications of Information Measures to the Theory of Coding

Ruchi Handa[1, 2*], Rajesh Kumar Narula[3] and C.P. Gandhi[4]
[1]Research Scholar, IK Gujral Punjab Technical University, Kapurthala, India.
[2]Department of Applied Sciences, Khalsa College of Engineering and Technology, Amritsar, India.
[3]Department of Mathematical Sciences, IK Gujral Punjab Technical University, Kapurthala, India.
[4]Department of Mathematics, University School of Sciences, Rayat Bahra University, Kharar, Mohali, India.

(Corresponding author: Ruchi Handa)
(Received 01 June 2019, Revised 02 August 2019 Accepted 19 August 2019)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: It has been observed that information measures participate in designing various techniques for the development of mean codeword lengths. The present communication providing the applications of entropy measures for the development of new codeword lengths is a step in this direction. Moreover, our aim is to provide a deeper insight into the problems of correspondence between weighted mean and possible weighted entropy through the possible measures of weighted divergence.

Keywords: Entropy, Weighted entropy, Mean codeword length, Noiseless coding theorem, Kraft inequality, Monotonic increasing function.

I. INTRODUCTION

In the literature of entropy measures, one of many applications will be to the problem of efficient coding of messages to be sent over a noiseless channel, that is, our only concern is to maximize the number of messages that can be sent over the channel in a given time. Let us assume that the messages to be transmitted are generated by a random variable X and each value x_i, $i = 1, 2, \ldots, n$ of X must be represented by a finite sequence of symbols chosen from the set $\{a_1, a_2, ..., a_D\}$. This set is called code alphabet or set of code characters and sequence assigned to each x_i, $i = 1, 2, \ldots, n$ is called code word. While dealing with coding theory, Kraft's (1949) inequality participates with a central role. With D as alphabet size and n_i the length of code word associated with x_i this inequality is given by

$$\sum_{i=1}^{n} D^{-n_i} \le 1 \tag{1}$$

In communication theory, we usually come across those codes which minimize the following code word length:

$$L = \sum_{i=1}^{n} p_i\, n_i \tag{2}$$

Taking into consideration Belis and Guiasu's (1968) entropy, Guiasu and Picard (1971) defined the following quantity as weighted mean codeword length:

$$L(W) = \frac{\sum_{i=1}^{n} n_i w_i p_i}{\sum_{i=1}^{n} w_i p_i} \tag{3}$$

Kapur (1998) introduced implicitly exponentiated mean of order α and type *a* viz.

$$L_{\alpha,a} = \frac{\alpha}{1-\alpha}\log_D\left(\sum_{i=1}^{n} p_i\, D^{a(1-\alpha)n_i/\alpha}\right) + \frac{1}{1-\alpha}\log_D\left(\sum_{i=1}^{n} p_i^{\alpha}\, D^{(\alpha-1)(1-a)n_i} \Big/ \sum_{i=1}^{n} p_i^{\alpha}\right) \tag{4}$$

and showed that its lower bound for uniquely decipherable codes was also between $R_\alpha(P)$ and $R_\alpha(P) + 1$ where $R_\alpha(P)$ is Renyi's (1961) entropy. In fact this gave an infinity of exponentiated means of order α for different values of a between 0 and 1. For a = 1 it gave Campbell's (1965) exponentiated mean and for $a = 0$ it gave another exponentiated mean as

$$\bar{L}_\alpha = \frac{1}{\alpha-1}\log_D\left(\sum_{i=1}^{n} p_i^{\alpha}\, D^{(\alpha-1)n_i} \Big/ \sum_{i=1}^{n} p_i^{\alpha}\right) \tag{5}$$

which was called Kapur's (1998) exponentiated mean of order α. All the infinity of exponentiated means of order α have the same lower bounds $R_\alpha(P)$ and $R_\alpha(P) + 1$. This proves an important result that while a given mean can have only one pair of lower bounds, one pair of lower bounds can correspond to infinity of mean code word lengths.

Various measures of information along with their applications to coding theory have well been discussed by Kapur (1998). In coding theory, generally we don't consider the problem of error correction but our only concern is to maximize the number of messages.

Thus, we find 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 will lie between two bounds and a noiseless coding theorem seeks to find these two lower bounds for a given mean and a given constraint.

Various measures of information along with their applications to coding theory have well been discussed by Joshi and Kumar (2018), Kawan and Yüksel (2018), Lee and Chung (2018), Wondie and Kumar (2017), Reviewed and Ferreira (2019), Frumin, Gelash and Turitsyn (2017), Hayashi (2019) etc.

II. CORRESPONDENCE BETWEEN INFORMATION MEASURES AND CODING THEORY

Below, we demonstrate the connection between entropy measures and the codeword lengths.

(i) Codeword Lengths through Divergence Measures

Here, we develop certain exponentiated mean codeword lengths already existing in the literature of coding theory.

Theorem 2.1: If $n_1, n_2, n_3, ..., n_n$ are the lengths of a uniquely decipherable code, then

$$L_{r,s,k} \geq \left[H_s^r(P)\right]_k - \frac{k(1-s)-(1-r)}{s-r}\log_D \sum_{i=1}^{n} D^{-n_i} \qquad (6)$$

where $L_{r,s,k} = \frac{1}{r-s}\left[(r-1)L^r - k(s-1)L^s\right]$, k is some real constant, r, s are real parameters, $\left[H_s^r(P)\right]_k$ is Kapur's (1967) entropy and

$$L^r = \frac{1}{r-1}\log_D\left(\frac{\sum_{i=1}^{n} p_i^r D^{-n_i(1-r)}}{\sum_{i=1}^{n} p_i^r}\right), L^s = \frac{1}{s-1}\log_D\left(\frac{\sum_{i=1}^{n} p_i^s D^{-n_i(1-s)}}{\sum_{i=1}^{n} p_i^s}\right)$$

are Kapur's (1998) mean codeword lengths.

Proof. The following divergence is due to Kapur (1994):

$$K(P:Q) = \frac{1}{\alpha-\beta}\log_D \frac{\sum_{i=1}^{n} p_i^r q_i^{1-r}}{\left(\sum_{i=1}^{n} p_i^s q_i^{1-s}\right)^k}; r \neq 1, s \neq 1, r, s > 0,\ k > 0$$

Since $K(P:Q) \geq 0$, letting $q_i = \frac{D^{-n_i}}{\sum_{i=1}^{n} D^{-n_i}}$, the above expression gives

$$\frac{1}{r-s}\left[(r-1)\overline{L_r} - k(s-1)\overline{L_s}\right] \geq \frac{1}{s-r}\log_D\left(\frac{\sum_{i=1}^{n} p_i^r}{\left(\sum_{i=1}^{n} p_i^s\right)^k}\right) - \frac{k(1-s)-(1-r)}{r-s}\log_D \sum_{i=1}^{n} D^{-n_i} \qquad (7)$$

The equation (7) further gives

$$L_{r,s,k} \geq \left[H_s^r(P)\right]_k - \frac{k(1-s)-(1-r)}{r-s}\log_D \sum_{i=1}^{n} D^{-n_i}$$

where $\left[H_s^r(P)\right]_k = \frac{1}{s-r}\log_D\left(\frac{\sum_{i=1}^{n} p_i^r}{\left(\sum_{i=1}^{n} p_i^s\right)^k}\right)$ is Kapur's (1986) additive measure of entropy.

Special cases

1. For $k=1$, (7) becomes

$$\frac{1}{r-s}\left[(r-1)L^r - (s-1)L^s\right] \geq \frac{1}{s-r}\log_D\left(\frac{\sum_{i=1}^{n} p_i^r}{\sum_{i=1}^{n} p_i^s}\right) - \log_D \sum_{i=1}^{n} D^{-n_i},$$

that is,

$$L_{r,s} \geq H_s^r(P) - \log_D \sum_{i=1}^{n} D^{-n_i} \qquad (8)$$

where $L_{r,s} = \frac{1}{r-s}\left[(r-1)L^r - (s-1)L^s\right]$ is the exponentiated mean of order r and type S and $H_s^r(P)$ is Kapur's (1986) entropy of order r and type s.

Now, since $\sum_{i=1}^{n} D^{-n_i}$ always lies between D^{-1} and 1, equation (8) shows that the lower bound for $L_{r,s}$ lies between $H_s^r(P)$ and $H_s^r(P)+1$.

2. For $k=1, s=1$ (7) becomes

$$L^r \geq \frac{1}{1-r}\log_D \sum_{i=1}^{n} p_i^r - \log_D \sum_{i=1}^{n} D^{-n_i}$$

that is, $L^r \geq R_r(P) - \log_D \sum_{i=1}^{n} D^{-n_i}$ where L^r is r order mean and $R_r(P)$ is Renyi's (1961) entropy. This proves that L^r's lower bound lies between $R_r(P)$ and $R_r(P)+1$.

3. For $k=1, s=1$ and $r \to 1$, (7) provides the following expression:

$$L \geq H(P) - \log_D \sum_{i=1}^{n} D^{-n_i}$$

Where $H(P)$ is Shannon's (1948) entropy.

This proves that L's lower bound lies between $H(P)$ and $H(P)+1$.

(ii) Deriving Existing Codeword Lengths

Here, we make available Campbell's (1965) and Shannon's (1948) mean codeword lengths.

Theorem 2.2: If $n_1, n_2, n_3, ..., n_n$ are uniquely decipherable codeword lengths, then

$$-\frac{1+\log_D s}{\log_D s}\log_D \sum_{i=1}^{n}\left(p_i^{\frac{1}{1+\log_D s}} s^{\frac{\log_D p_i}{1+\log_D s}} D^{-n_i\left(\frac{\log_D s}{1+\log_D s}\right)}\right) \geq -\frac{1}{\log_D s}\log_D \sum_{i=1}^{n} p_i s^{\log_D p_i} \qquad (9)$$

where $s > 0, s \neq 1$.

Proof. We know that Holder's inequality is given by

$$\sum_{i=1}^{n} x_i y_i \geq \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} y_i^q\right)^{\frac{1}{q}} \text{ where } \frac{1}{p}+\frac{1}{q}=1, p \text{ or } q<1 \qquad (10)$$

Substituting

$$x_i = p_i^{-\frac{1}{\log_D s}} s^{-\frac{\log_D p_i}{\log_D s}}, y_i = p_i^{\frac{1}{\log_D s}} s^{\frac{\log_D p_i}{\log_D s}} D^{-n_i}, \frac{1}{p} = -\frac{1}{\log_D s}, \frac{1}{q} = \frac{1+\log_D s}{\log_D s}$$

in equation (10), we get

$$\sum_{i=1}^{n} D^{-n_i} \geq \left[\sum_{i=1}^{n}\left(p_i^{\frac{-1}{\log_D s}} s^{\frac{-\log_D p_i}{\log_D s}}\right)^{-\log_D s}\right]^{\frac{-1}{\log_D s}} \left[\sum_{i=1}^{n}\left(p_i^{\frac{1}{\log_D s}} s^{\frac{\log_D p_i}{\log_D s}} D^{-l_i}\right)^{\frac{\log_D s}{1+\log_D s}}\right]^{\frac{1+\log_D s}{\log_D s}}$$

that is,

$$\sum_{i=1}^{n} D^{-n_i} \geq \left[\sum_{i=1}^{n} p_i s^{\log_D p_i}\right]^{\frac{-1}{\log_D s}} \left[\sum_{i=1}^{n}\left(p_i^{\frac{1}{1+\log_D s}} s^{\frac{\log_D p_i}{1+\log_D s}} D^{-n_i\left(\frac{\log_D s}{1+\log_D s}\right)}\right)\right]^{\frac{1+\log_D s}{\log_D s}}$$

or

$$0 \geq -\frac{1}{\log_D s}\log_D \sum_{i=1}^{n} p_i s^{\log_D p_i} + \frac{1+\log_D s}{\log_D s}\log_D \sum_{i=1}^{n}\left(p_i^{\frac{1}{1+\log_D s}} s^{\frac{\log_D p_i}{1+\log_D s}} D^{-l_i\left(\frac{\log_D s}{1+\log_D s}\right)}\right)$$

Reshuffling the terms, we get (9). It is observed that (9) provides relation between entropy and non-mean codeword length.

Particular Cases:

Case-I: Taking $s = D$, (9) becomes

$$-2\log_D \sum_{i=1}^{n}\left(p_i^{\frac{1}{2}} D^{\frac{\log_D p_i}{2}} D^{-n_i\left(\frac{1}{2}\right)}\right) \geq -\log_D \sum_{i=1}^{n} p_i D^{\log_D p_i}$$

that is,

$$-2\log_D \sum_{i=1}^{n}\left(p_i D^{-n_i\left(\frac{1}{2}\right)}\right) \geq -\log_D \sum_{i=1}^{n} p_i^2 \qquad (11)$$

This inequality provides relation between Campbell's (1965) codeword length and Renyi's (1961) entropy.

Case-II: Letting $s \to 1$ in (9), we get

$$\sum_{i=1}^{n} p_i n_i \geq -\sum_{i=1}^{n} p_i \log_D p_i \qquad (12)$$

The inequality (12) provides relation between Shannon's (1948) entropy and the standard codeword length.

(iii) Relation between Weighted Mean and Possible Weighted Entropy

We first of all define the following weighted mean:

$$L^r(W) = \frac{1}{r-1}\log_D\left[\frac{\sum_{i=1}^{n} w_i p_i^r D^{-n_i(1-r)}}{\sum_{i=1}^{n} w_i p_i^r}\right]; r > 1 \qquad (13)$$

We observe that

$$\underset{r\to 1}{Lt}\, L^r(W) = \frac{\sum_{i=1}^{n} w_i\{p_i n_i + p_i \log_D p_i\} - \sum_{i=1}^{n} w_i p_i \log_D p_i}{\sum_{i=1}^{n} w_i p_i} = \frac{\sum_{i=1}^{n} w_i p_i n_i}{\sum_{i=1}^{n} w_i p_i}$$

which is Guiasu and picard's (1971) weighted mean. Thus, we see that the weighted mean introduced in (13) is a generalized weighted mean.

Next, we provide the correspondence between weighted mean and possible weighted entropy through the possible measures of weighted divergence.

Theorem 2.3: If $n_1, n_2, ..., n_n$ be the lengths of uniquely decipherable codes, then

$$L^r(W) \geq H_r(P;W) - \log_D \sum_{i=1}^{n} D^{-n_i}$$

where $H_r(P;W) = \frac{1}{1-r}\log_D(\sum_{i=1}^{n} w_i p_i^r)$ is possible measure of weighted entropy and $L^r(W)$ is weighted mean defined above.

Proof: To prove the above theorem, we make use of the possible weighted divergence given by

$$K_r(P:Q;W) = \frac{1}{r-1}\left[\tan^{-1}\left\{\sum_{i=1}^{n} w_i p_i^r q_i^{1-r}\right\} - \frac{\pi}{4}\right]; r > 1 \qquad (14)$$

This is to be noted that upon ignoring weights, measure (14) reduces to

$$K_r(P;r) = \frac{1}{r-1}\left[\tan^{-1}\left\{\sum_{i=1}^{n} p_i^r q_i^{1-r}\right\} - \frac{\pi}{4}\right]; r > 1 \qquad (15)$$

which is Kapur's (1994) divergence.

Now, we know that $K_r(P,Q;W) \geq 0$

$$\Rightarrow \tan^{-1}\left\{\sum_{i=1}^{n} p_i^r q_i^{1-r}\right\} \geq \frac{\pi}{4} \qquad (16)$$

Letting $q_i = \frac{D^{-n_i}}{\sum_{i=1}^{n} D^{-n_i}}$ in equation (16), we get

$$\tan^{-1}\left\{\sum_{i=1}^{n} w_i p_i^r \left(\frac{D^{-n_i}}{\sum_{i=1}^{n} D^{-n_i}}\right)^{1-r}\right\} \geq \frac{\pi}{4}$$

$$\Rightarrow \sum_{i=1}^{n} w_i p_i^r \; D^{-n_i(1-r)} \geq \left(\sum_{i=1}^{n} D^{-n_i}\right)^{1-r} \qquad (17)$$

(or $-\log_D\left[\sum_{i=1}^{n} w_i p_i^r \; D^{-n_i(1-r)}\right] \leq (r-1)\log_D\left(\sum_{i=1}^{n} D^{-n_i}\right)$ (18)

Adding, $\log_D\left(\sum_{i=1}^{n} w_i p_i^r\right)$ (18) provides

$$L^r(W) = H_r(P;W) - \log_D\left(\sum_{i=1}^{n} D^{-n_i}\right)$$

$L^r(W) = H_r(P;W) - \log_D\left(\sum_{i=1}^{n} D^{-n_i}\right)$ which proves the theorem.

Note: The possible measure of entropy $H_r(P;W)$ reduces to Renyi's (1961) entropy after ignoring the weights.

Next, we provide another interesting correspondence.

Theorem 2.4: If $n_1, n_2, ..., n_n$ be the lengths of uniquely decipherable codes, then

$$H^r(P;W) \leq L(W)$$

where $H^r(P;W) = \sum_{i=1}^{n} w_i p_i^r\left[1 - \log_D p_i^r\right]; r > 1$ is weighted entropy and $L(W)$ is some weighted function.

Proof: To prove the above theorem, we employ Gurdial and Pessoa's (1977) fundamental theorem which states that

$$\frac{r}{1-r}\log_D\left[\sum_{i=1}^{n} p_i\left\{\frac{w_i}{\sum_{i=1}^{n} w_i p_i}\right\}^{\frac{1}{r}} D^{-n_i\frac{r-1}{r}}\right] \ge \frac{1}{1-r}\log_D\left(\frac{\sum_{i=1}^{n} w_i p_i^r}{\sum_{i=1}^{n} w_i p_i}\right) \qquad (19)$$

where $H_r(P;W) = \frac{1}{1-r}\log_D\left(\frac{\sum_{i=1}^{n} w_i p_i^r}{\sum_{i=1}^{n} w_i p_i}\right); r \ne 1, r > 0$

is Gurdial and Pessoa's (1977) weighted entropy and

$$L_r(W) = \frac{r}{1-r}\log_D\left[\sum_{i=1}^{n} p_i\left\{\frac{w_i}{\sum_{i=1}^{n} w_i p_i}\right\}^{\frac{1}{r}} D^{-n_i\frac{r-1}{r}}\right]$$

is parametric weighted code word length again introduced by Gurdial and Pessoa (1977).

From equation (19), we have

$$\sum_{i=1}^{n} w_i p_i^r \le \left\{\sum_{i=1}^{n} w_i p_i\right\}\left[\sum_{i=1}^{n} p_i\left\{\frac{w_i}{\sum_{i=1}^{n} w_i p_i}\right\}^{\frac{1}{r}} D^{-n_i\frac{r-1}{r}}\right]^r \qquad (20)$$

Substituting $x_i = -w_i^{\frac{r}{1-r}} p_i^{\frac{r^2}{1-r}}\left\{\log_D p_i^r\right\}^{\frac{r}{1-r}} D^{-n_i}$

$y_i = -w_i^{\frac{r}{1-r}} p_i^{\frac{r^2}{1-r}}\left\{\log_D p_i^r\right\}^{\frac{r}{1-r}}$, $p_i = 1-r, q = \frac{r-1}{r}$ in (10)

and applying Kraft's (1949) equality $\sum_{i=1}^{n} D^{-n_i} = 1$, we get

$$0 \ge \frac{1}{1-r}\log_D\left[-\sum_{i=1}^{n} w_i^r p_i^{r^2}\left\{\log_D p_i^r\right\}^r D^{-n_i(1-r)}\right] + \frac{r}{r-1}\log_D\left[-\sum_{i=1}^{n} w_i p_i^r \log_D p_i^r\right]$$

or $$-\sum_{i=1}^{n} w_i p_i^r \log_D p_i^r \le \left[-\sum_{i=1}^{n} w_i^r p_i^{r^2}\left\{\log_D p_i^r\right\}^r D^{-n_i(1-r)}\right]^{\frac{1}{r}} \qquad (21)$$

Adding (10) and (21), we have

$H^r(P;W) \le L(W)$

where

$$L(W) = \left\{\sum_{i=1}^{n} w_i p_i\right\}\left[\sum_{i=1}^{n} p_i\left\{\frac{w_i}{\sum_{i=1}^{n} w_i p_i}\right\}^{\frac{1}{r}} D^{-n_i\frac{r-1}{r}}\right]^r - \left[\sum_{i=1}^{n} w_i^r p_i^{r^2}\left\{\log_D p_i^r\right\}^r D^{-n_i(1-r)}\right]^{\frac{1}{r}}$$

is neither any weighted mean codeword length nor its monotonic increasing function. Hence the theorem.

III. CONCLUSION

It can be shown that taking into consideration the existing as well as new entropy measures, many new coding theorems can be proved and consequently, any new codeword lengths can be developed. The advantage of this technique is that many new measures of entropy via coding theorems can be developed. The work can further be extended for other entropy measures.

Conflict of Interest: Author has no any conflict of interest.

REFERENCES

[1]. Belis, M. and Guiasu, S. (1968). A quantitative-qualitative measure of information in cybernetic systems. *IEEE Transactions on Information Theory,* **14**: 593-594.

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

[3]. Frumin, L. L., Gelash, A. A. and Turitsyn, S. K. (2017). New approaches to coding information using inverse scattering transform. *Phys. Rev. Lett.*, **118**(22): 5 pp.

[4]. Guiasu, S. and Picard, C.F. (1971). Borne in ferictur de la longuerur utile de certains codes. *Comptes Rendus Mathematique Academic des Sciences Paris,* **273**: 248-251.

[5]. Gurdial and Pessoa, F. (1977). On useful information of order α. *Journal of Combinatorics Information and System Sciences,* **2**: 158-162.

[6]. Hayashi, M. (2019). Universal channel coding for general output alphabet. *IEEE Trans. Inform. Theory*, **65** (1): 302–321.

[7]. Joshi, R. and Kumar, S. (2018). A new weighted (α,β)-norm information measure with application in coding theory. *Physica A: Statistical Mechanics and its Applications*, **510**(C): 538-551.

[8]. Kapur, J.N. (1967). Generalized entropy of order α and type β. *Mathematics Seminar,* **4**: 79-84.

[9]. Kapur, J.N. (1986). Four families of measures of entropy. *Indian Journal of Pure and Applied Mathematics,* **17**: 429-449.

[10]. Kapur, J.N. (1994). *Measures of Information and their Applications*. Wiley Eastern, New York.

[11]. Kapur, J.N. (1998). *Entropy and Coding*. Mathematical Sciences Trust Society, New Delhi.

[12]. Kawan, C. and Yüksel, S. (2018). On optimal coding of non-linear dynamical systems. *IEEE Trans. Inform. Theory*, **64**(10): 6816–6829.

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

[14]. Lee, Si-Hyeon and Chung, Sae-Young (2018). A unified random coding bound. *IEEE Trans. Inform. Theory*, **64** (10): 6779–6802.

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

[16]. Ouahada, K., & Ferreira, H.C. (2019). New Distance Concept and Graph Theory Approach for Certain Coding Techniques Design and Analysis. *Communications in Applied and Industrial Mathematics*, ***10***(1), 53-70.

How to cite this article: Handa, R., Narula, R.K. and Gandhi, C.P. (2019). Applications of Information Measures to the Theory of Coding. *International Journal on Emerging Technologies*, **10**(2): 446–449.