

A New Generalized Entropy Measure and its Properties

Mohammad Javid Dar, Saima M. Sofi, Mirza A. K. Baig

Post Graduate Department of Statistics University of Kashmir
Hazratbal, Srinagar, India

*Corresponding author's email: javidstat [AT] gmail.com

ABSTRACT---- *In this research article, a new two parametric measure of Entropy $J_{\alpha\beta}(P)$ and its corresponding code word length $L_{\alpha\beta}(P)$ has been developed. The developed measures are the generalizations to some well known existing measures. Besides, some noiseless coding theorems for discrete noiseless channel have been developed, and the results thus obtained have been verified with the support of an numerical example. Also, at the end of this research article, a comparative study in terms of monotonic behavior among the proposed entropy $J_{\alpha\beta}(P)$, Matahi's entropy $M_{\alpha}(P)$ and Tsallis entropy $T_{\alpha}(P)$ together with their respective average code word length measures have been made and graphically displayed.*

Keywords---- Information, Shannon's Entropy, Coding theory, Prefix Code, Average code-word Length, Kraft's Inequality, Holder's Inequality, Shannon Fanno Codes and Noiseless Coding theorem

1. INTRODUCTION

In the modern world, information is being transmitted through various means and in all such cases the same is being carried from one person/place to another. There is no denying in this fact that we seek information only when we are in doubt and we resort to the concerned quarters to remove our doubts. On the other hand, if an event can occur in just one way, there is no uncertainty/doubt about it and no information is called for either. We get some information by the occurrence of an event only when there was some uncertainty before its occurrence. Naturally, the amount of information received by the occurrence of an event must be equal to the amount of uncertainty prevailing before its occurrence. Thus, uncertainty and information are two sides of the same coin.

The modern information theory can be classified into the following three branches:

- (i) **Shannon Theory**, due to C.E Shannon, [9] deals with mathematical models for communication problems. The concept of 'Entropy' given by Shannon in his mathematical model, has been found useful in many different disciplines and has penetrated into various fields like; linguistics, Psychology, Neurology economics, Business, Accounting, statistics, Biology and Thermodynamics.
- (ii) **Cybernetics**, due to Norbert wiener deals with the communication problems encountered in living beings and social organizations and
- (iii) **Coding Theory**, a recently developed subject that deals with the theory of error correcting codes, finds applications in problems of determining good encoding schemes to combat errors in transmission.

Information theory, the mathematical theory of communication, has two primary goals: The first is the development of the fundamental theoretical limits on the achievable performance when communicating given information source over a given communications channel using coding schemes from within a prescribed class. The second goal is the development of coding schemes that provide performance that is reasonably good in comparison with the optimal performance given by the theory. Information theory was born in a surprisingly rich state in the classic papers of Claude E. Shannon [9] and [10] which contained the basic results for simple memory less sources and channels and introduced more general communication systems models, including finite state sources and channels.

2. INFORMATION CONTENT AND SHANNON'S ENTROPY:

In information theory, the information content, self information or surprisal of a random variable is the amount of information gained when it is sampled.

Definition: Given a random variable $X = (x_1, x_2, \dots, x_n)$ with probability $P = (p_1, p_2, \dots, p_n)$ having probability mass function $P_X(x)$, the self- information of measuring X as outcome x is defined as:

$$I_X(x) = -\log[P_X(x)] = \log\left(\frac{1}{P_X(x)}\right) \quad (2.1)$$

Analogously:

$$I(E) = -\log[p(E)] = -\log(P). \quad (2.1)$$

Entropy is the expectation of the self-information of all the outcomes of the random variable. Let a random variable $X = (x_1, x_2, \dots, x_n)$ with probability $P = (p_1, p_2, \dots, p_n)$, $p_i \geq 0$ and $\sum_{i=1}^n p_i = 1$, Shannon [9] has defined the following measure of information and call it as entropy:

$$H(P) = -\sum_{i=1}^n p_i \log p_i \tag{2.3}$$

Let the probabilities of a transmission of n code-word be $P = (p_1, p_2, \dots, p_n)$ with the corresponding lengths $L = (l_1, l_2, \dots, l_n)$ and these lengths satisfy the Kraft's inequality [6] defined as:

$$\sum_{i=1}^n D^{-l_i} \leq 1 \tag{2.4}$$

Where, D is considered to be the size of code alphabet. Shannon [9] has proved that for all uniquely decipherable codes satisfying inequality (2.4), the lower bound of the average code-word length give as,

$$L = \sum_{i=1}^n p_i l_i \tag{2.5}$$

Lies between $H(P)$ and $H(P) + 1$.

Campbell [2] considered the more general exponentiated mean code-word length as:

$$L_a = \frac{\alpha}{1-\alpha} \log_D \left[\sum_{i=1}^n p_i D^{-l_i \left(\frac{\alpha-1}{\alpha}\right)} \right]; \alpha > 0, \alpha \neq 1 \tag{2.6}$$

And showed that for all uniquely decodable codes, satisfying constraint (3), the lower bound of (5) lies between $R_\alpha(P)$ and $R_\alpha(P) + 1$, where

$$R_\alpha(P) = \frac{1}{1-\alpha} \log_D \left[\sum_{i=1}^n p_i^\alpha \right], \alpha > 0, \alpha \neq 1 \tag{2.7}$$

Is Renyi's [8] entropy.

3. CODING THEORY

Coding theory is the study of the properties of codes and their fitness for a specific application. Codes are used for data compression, cryptography, error correction and more recently also for network coding. Codes are studied by various scientific disciplines such as information theory, electrical engineering, mathematics and computer sciences – for the purpose of designing efficient and reliable data transmission methods. This typically involves the removal of redundancy and the correction (or detection) of errors in the transmitted data.

4. PREFIX CODE

Prefix code is a code in which no code word forms the prefix of any other code word. Such codes are also called Uniquely Decodable or Instantaneous codes.

5. AVERAGE CODE WORD LENGTH

Consider a set of symbols (alphabets) $S = \{s_1, s_2, \dots, s_M\}$ with their corresponding probabilities

$S = \{p_1, p_2, \dots, p_M\}$; ($p_i > 0, \sum p_i = 1$). If the symbol s_i be assigned a code word of length $l_i, i = 1, 2, \dots, M$, then the definition for the average code word length is $L = \sum_{i=1}^n p_i l_i$, and the basic problem of Noiseless coding is to minimize average code word length. (Noiseless Coding Theorem).

Shannon [10] established the first noiseless coding theorem which states that for all uniquely decipherable codes, the lower bound for the arithmetic mean $L = \sum_{i=1}^n p_i l_i$ lies between $S(P)$ and $S(P) + 1$, where

$$S(P) = -\sum_{i=1}^n p_i \log p_i \text{ is Shannon's measure of entropy.}$$

Here in this communication, we have presented a new two parametric entropy measure $J_{\alpha\beta}(P)$ of order α and type β along with its corresponding average code-word length $L_{\alpha\beta}(P)$ and an attempt has been made to test its genuineness in the line of already existing measures in the literature of information communication.

6. NOISELESS CODING THEOREMS

The proposed measure is given as under:

$$J_{\alpha\beta}(P) = \frac{1}{\alpha\beta-1} \left[\sum_{i=1}^n p_i^{2-\alpha\beta} - 1 \right]; \alpha \neq 1, \beta > 0, -\infty < \alpha < 2 \tag{6.1}$$

Corresponding to this measure, we propose the following average codeword length as:

$$L_{\alpha\beta}(P) = \frac{1}{\alpha\beta-1} \left[\sum_{i=1}^n p_i D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(l_i)} - 1 \right]; \alpha \neq 1, \beta > 0, \tag{6.2}$$

The parameters α and β may be considered as factors affecting the codes used for data compression and data transmission, associated with the source alphabet with corresponding probabilities denotes as, $[x_i, p_i]$.

Particular cases for (6.1) and (6.2):

(I). When $\alpha \rightarrow 1$ and $\beta = 1$, the measure defined in (6.1) tends to Shannon's entropy given as:

$$S(P) = - \sum_i^n p_i \log p_i$$

(II). When $\alpha \rightarrow 1$ and $\beta = 1$, the measure defined in (6.2) becomes $L = \sum_i^n p_i n_i$, which is optimal code-word length due to Shannon [9].

(III). When $\beta = 1$, the measure defined in (6.1) tends to Mathai's entropy [6] communicated by Baig and Javid [1].

(IV). When $\beta = 1$, the measure defined in (6.2) tends to average codeword length corresponding to Mathai's [6] entropy as communicated by Baig and Javid [1].

In this section, we have developed some noiseless coding theorems corresponding to the proposed entropy measure.

Theorem 1:- For all uniquely decipherable codes

$$J_{\alpha\beta}(P) \leq L_{\alpha\beta}(P) \tag{6.3}$$

Where

$$L_{\alpha} = \frac{1}{\alpha\beta-1} \left[\sum_i^n p^{2-\alpha\beta} D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(n_i)} - 1 \right]$$

Proof:-By Holders inequality, we have

$$\sum_i^n x_i y_i \geq \left(\sum_i^n x_i^p \right)^{\frac{1}{p}} \left(\sum_i^n y_i^q \right)^{\frac{1}{q}} \quad ; \quad 0 < p < 1, q < 0 \text{ or } 0 < q < 1, p < 0 \tag{6.4}$$

Setting $x_i = p_i^{-\frac{1}{t}} D^{-n_i}$;

$$y_i = p_i^{\frac{1}{t}} \text{ and } p = -t \Rightarrow 0 < p < 1, q = \frac{t}{t+1} \Rightarrow q < 0$$

Thus equation (6.4) becomes

$$\sum_i^n \left[p_i^{-\frac{1}{t}} D^{-n_i} p_i^{\frac{1}{t}} \right] \geq \left[\left\{ p_i^{-\frac{1}{t}} D^{-n_i} \right\}^{-t} \right]^{\frac{-1}{t}} \left[\left\{ p_i^{\frac{1}{t}} \right\}^{\frac{t}{t+1}} \right]^{\frac{t+1}{t}}$$

Using Kraft's inequality, we have

$$\left[p_i^{\frac{1}{t+1}} \right]^{\frac{t+1}{t}} \leq \left[p_i D^{n_i t} \right]^{\frac{1}{t}}$$

or, $\sum_i^n p_i^{\frac{1}{t}} \leq \sum_i^n \left[p_i D^{n_i t} \right]^{\frac{1}{t}}$

or, $\sum_i^n p_i \leq \sum_i^n p_i D^{n_i t} \tag{6.5}$

Dividing both sides by t, we get:

$$\frac{\sum_i^n p_i}{t} \leq \frac{\sum_i^n p_i D^{n_i t}}{t}$$

Subtracting n from both sides, we get:

$$\sum_i^n \frac{p_i - 1}{t} \leq \sum_i^n \frac{p_i D^{n_i t} - 1}{t} \tag{6.6}$$

Taking $\alpha\beta = \frac{1}{1-t}$, $t = \frac{\alpha\beta-1}{\alpha\beta}$, $\alpha, \beta > 0$, $\alpha\beta \neq 1$ and $p_i = p_i^{2-\alpha\beta}$

Thus equation (6.6) becomes:

$$\frac{\alpha\beta}{\alpha\beta-1} \sum_i^n \left[(p_i^{2-\alpha\beta}) - 1 \right] \leq \frac{\alpha\beta}{\alpha\beta-1} \sum_i^n \left[(p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(n_i)} - 1 \right]$$

Dividing both sides by $\alpha\beta$, we get:

$$\frac{1}{\alpha\beta-1} \sum_i^n [(p_i^{2-\alpha\beta}) - 1] \leq \frac{1}{\alpha\beta-1} \sum_i^n \left[(p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} - 1 \right] \quad (6.7)$$

That is $J_{\alpha\beta}(A) \leq L_{\alpha\beta}$

Which proves the theorem.

Theorem 2:- For all uniquely decipherable codes,

$$J_{\alpha\beta,\beta} \leq L_{\alpha\beta,\beta} \quad (6.8)$$

where,

$$L_{\alpha\beta,\beta} = \frac{1}{\beta-\alpha\beta} \sum_i^n \left[(p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} - (p_i^{2-\beta}) D^{\left(\frac{\beta-1}{\beta}\right)(ni)} \right] \quad (6.9)$$

Where either $\alpha \geq 1, \beta \leq 1$ or $\beta \geq 1, \alpha \leq 1$

Proof: - Since from (6.7), we have

$$\frac{1}{\alpha\beta-1} \sum_i^n [(p_i^{2-\alpha\beta}) - 1] \leq \frac{1}{\alpha\beta-1} \sum_i^n \left[(p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} - 1 \right]$$

Multiplying both sides by $(\alpha\beta - 1)$, we have

$$\sum_i^n [(p_i^{2-\alpha\beta}) - 1] \leq \sum_i^n \left[(p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} - 1 \right] \quad (6.10)$$

Substituting $\alpha=1$, we have

$$\sum_i^n [(p_i^{2-\beta}) - 1] \leq \sum_i^n \left[(p_i^{2-\beta}) D^{\left(\frac{\beta-1}{\beta}\right)(ni)} - 1 \right] \quad (6.11)$$

Subtract (6.11) to (6.10) and divide by $(\beta - \alpha\beta)$, we get

$$\begin{aligned} \frac{1}{\beta-\alpha\beta} \sum_i^n [(p_i^{2-\alpha\beta}) - (p_i^{2-\beta})] \\ \leq \frac{1}{\beta-\alpha\beta} \sum_i^n \left[(p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} - (p_i^{2-\beta}) D^{\left(\frac{\beta-1}{\beta}\right)(ni)} \right] \end{aligned} \quad (6.12)$$

That is

$J_{\alpha\beta,\beta} \leq L_{\alpha\beta,\beta}$. This proves the theorem.

Theorem 3:- For all uniquely decipherable codes

$$J'_{\alpha\beta,\beta} \leq L'_{\alpha\beta,\beta} \quad (6.13)$$

Where $L'_{\alpha\beta,\beta} = \frac{1}{\alpha\beta+\beta+2} \sum_i^n \left[\left\{ (p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} - (p_i^{2-\alpha\beta}) D^{\left(\frac{\beta-1}{\beta}\right)(ni)} \right\} - 2 \right] \quad (6.14)$

Proof:- The result can be easily proved by adding (6.10) and (6.11) and then dividing by $(\alpha\beta + \beta + 2)$.

Theorem 4:- For all uniquely decipherable codes

$$J'_{\alpha\beta,\beta} \leq L'_{\alpha\beta,\beta} \tag{6.15}$$

Where

$$J'_{\alpha\beta,\beta} = \frac{1}{\beta - \alpha\beta} \left[\frac{\sum_i^n \{ (p_i^{2-\alpha\beta}) - 1 \}}{\sum_i^n \{ (p_i^{2-\beta}) - 1 \}} \right] \tag{6.16}$$

And

$$L'_{\alpha\beta,\beta} = \frac{1}{\beta - \alpha\beta} \left[\frac{\sum_i^n \left[\left\{ (p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} \right\} - 1 \right]}{\sum_i^n \left[\left\{ (p_i^{2-\beta}) D^{\left(\frac{\beta-1}{\beta}\right)(ni)} \right\} - 1 \right]} \right] \tag{6.17}$$

To prove this theorem, we first prove the following lemma

Lemma 1: For all uniquely decipherable codes

$$\sum_{i=1}^n \{ (p_i^{2-\alpha\beta}) - 1 \} \leq \sum_{i=1}^n \left[\left\{ (p_i^{2-\alpha\beta}) D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} \right\} - 1 \right]$$

Proof of the Lemma. From equation (6.5) we have

$$\sum_i^n (p_i) \leq \sum_i^n (p_i) D^{ni t}$$

Subtracting ‘n’ from both sides, we get

$$\sum_i^n \{ (p_i) - 1 \} \leq \sum_i^n \{ (p_i) D^{ni t} - 1 \}$$

Taking $\alpha\beta = \frac{1}{1-t}$, $t = \frac{\alpha\beta-1}{\alpha\beta}$, and $p_i = p_i^{2-\alpha\beta}$

we have

$$\sum_i^n \left[\{ p_i^{2-\alpha\beta} \} - 1 \right] \leq \sum_i^n \left[\left\{ p_i^{2-\alpha\beta} D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} \right\} - 1 \right] \tag{6.18}$$

Which proves the lemma

Proof of the theorem 4.

Substituting $\alpha = 1$ in (2.16), we have:

$$\sum_i^n \left[\{ p_i^{2-\beta} \} - 1 \right] \leq \sum_i^n \left[\left\{ p_i^{2-\beta} D^{\left(\frac{\beta-1}{\beta}\right)(ni)} \right\} - 1 \right] \tag{6.19}$$

Dividing (2.17) to (2.16), we get

$$\frac{\sum_i^n \left[\{ p_i^{2-\alpha\beta} \} - 1 \right]}{\sum_i^n \left[\{ p_i^{2-\beta} \} - 1 \right]} \leq \frac{\sum_i^n \left[\left\{ p_i^{2-\alpha\beta} D^{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} \right\} - 1 \right]}{\sum_i^n \left[\left\{ p_i^{2-\beta} D^{\left(\frac{\beta-1}{\beta}\right)(ni)} \right\} - 1 \right]}$$

Dividing both sides by $\beta - \alpha\beta$, we have

$$\frac{1}{\beta-\alpha\beta} \frac{\sum_i^n \left[\left\{ p_i^{2-\alpha\beta} \right\}_{-1} \right]}{\sum_i^n \left[\left\{ p_i^{2-\beta} \right\}_{-1} \right]} \leq \frac{1}{\beta-\alpha\beta} \frac{\sum_i^n \left[\left\{ p_i^{2-\alpha\beta} D_{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} \right\}_{-1} \right]}{\sum_i^n \left[\left\{ p_i^{2-\beta} D_{\left(\frac{\beta-1}{\beta}\right)(ni)} \right\}_{-1} \right]} \quad (6.20)$$

$$\Rightarrow J'_{\alpha\beta,\beta} \leq L'_{\alpha\beta,\beta}.$$

The R.H.S. is a new exponentiated mean codeword length of order α and type β and is defined as:

$$\frac{1}{\beta-\alpha\beta} \frac{\sum_i^n \left[\left\{ p_i^{2-\alpha\beta} \right\}_{-1} \right]}{\sum_i^n \left[\left\{ p_i^{2-\beta} \right\}_{-1} \right]} \leq \frac{1}{\beta-\alpha\beta} \frac{\sum_i^n \left[\left\{ p_i^{2-\alpha\beta} D_{\left(\frac{\alpha\beta-1}{\alpha\beta}\right)(ni)} \right\}_{-1} \right]}{\sum_i^n \left[\left\{ p_i^{2-\beta} D_{\left(\frac{\beta-1}{\beta}\right)(ni)} \right\}_{-1} \right]}$$

7. ILLUSTRATION

In this section, we shall verify the above proved theorems by taking into consideration numerical example. Let a source memory has six characters with the following probabilities of transmission:

X	A	B	C	D	E	F
P(X)	1/3	1/4	1/8	1/8	1/12	1/12

Moreover, by using above depicted data, we shall use Shannon Fanno encoding procedure to obtain uniquely decodable code to the above message ensemble, and a comparative study among the Proposed Entropy $J_{\alpha\beta}(P)$, Mathai’s Entropy $M_{\alpha}(P)$, Tsalli’s Entropy $T_{\alpha}(P)$ and their corresponding Average Code word lengths shall also be carried out.

The outcome of $J_{\alpha\beta}(P)$, $M_{\alpha}(P)$ and $T_{\alpha}(P)$ together with their corresponding average code word lengths for different values of α & fixed $\beta=1$ using Shannon encoding are displayed in the table 1 as under:

Table: 1
Values of $J_{\alpha\beta}(P)$, $M_{\alpha}(P)$, $T_{\alpha}(P)$ and their corresponding code word lengths for different values of α at fixed $\beta=1$, using Shannon encoding.

P_i	Shannon Fano Codeword	l_i	α	β	Proposed Entropy		Mathai’s Entropy		Tsalli’s Entropy	
					$J_{\alpha\beta}(P)$	$L_{\alpha\beta}(P)$	$M_{\alpha}(P)$	$L_{\alpha}(P)$	$T_{\alpha}(P)$	$L_{\alpha}(P)$
1/4	00	2	0.20	1	0.95	0.95	0.95	1.25	5.16	5.16
1/4	01	2	0.30	1	1.02	1.02	1.02	1.40	4.47	4.47
1/8	100	3	0.40	1	1.10	1.10	1.10	1.54	3.90	3.90
1/8	101	3	0.50	1	1.20	1.20	1.20	1.66	3.41	3.41
1/16	1100	4	0.60	1	1.30	1.30	1.30	1.75	3.00	3.00
1/16	1101	4	0.70	1	1.42	1.42	1.42	1.82	2.66	2.66
1/16	1110	4	0.80	1	1.56	1.56	1.56	1.86	2.37	2.37
1/16	1111	4	0.90	1	1.72	1.72	1.72	1.89	2.12	2.12

From the table: 1, as depicted above, we have summarized the following results:

- 1) That due to the increase in the value of alpha there is corresponding increase the in the value of proposed entropy $J_{\alpha\beta}(P)$ and Mathai’s entropy $M_{\alpha}(P)$ together with their respective average code word lengths. However, in case of Tsalli’s entropy and its corresponding average code word length, its value decreases due the increase in the value of alpha under Shannon Fanno encoding scheme.
- 2) That the average code word length in case of proposed entropy $J_{\alpha\beta}(P)$ is less than the average code word length under both Mathai’s entropy $M_{\alpha}(P)$ and Tsalli’s entropy $T_{\alpha}(P)$, very considerably, though the values of both proposed entropy $J_{\alpha\beta}(P)$ and Mathai’s entropy $M_{\alpha}(P)$ are same under Shannon Fanno Coding scheme.
- 3) That the proposed entropy $J_{\alpha\beta}(P)$ and its corresponding average code word length $L_{\alpha\beta}(P)$ is monotonic increasing with the increase in the value of alpha.

Next, we plot the Table 1 and see from the Fig.1 and Fig. 2, that both the proposed entropy $J_{\alpha\beta}(\mathbf{P})$ and Mathai's entropy $M_{\alpha}(\mathbf{P})$ are almost similar and Monotonic increasing with the increase in the value of α , but in case of Tsallis entropy $T_{\alpha}(\mathbf{P})$, the same is monotonic decreasing with the increase in the value of α . However, the average code word length in case of proposed entropy measure is less than that of Mathai's entropy and Tsallis entropy very considerably as is depicted in Fig. 2 below.

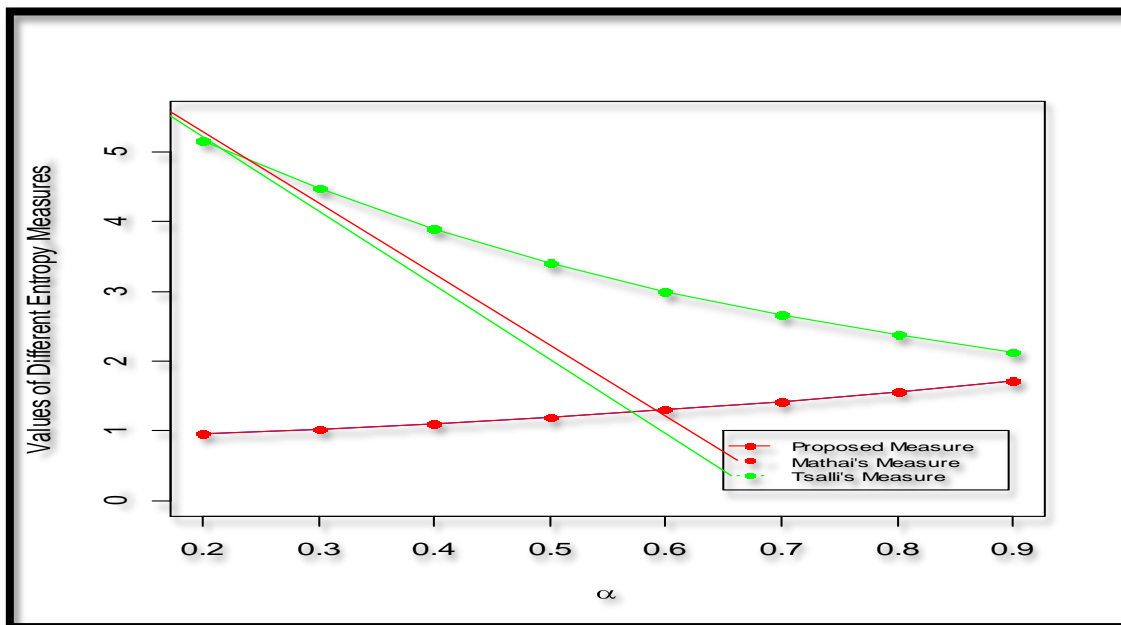


Figure 1 Graphic overview of values of different Entropy measures

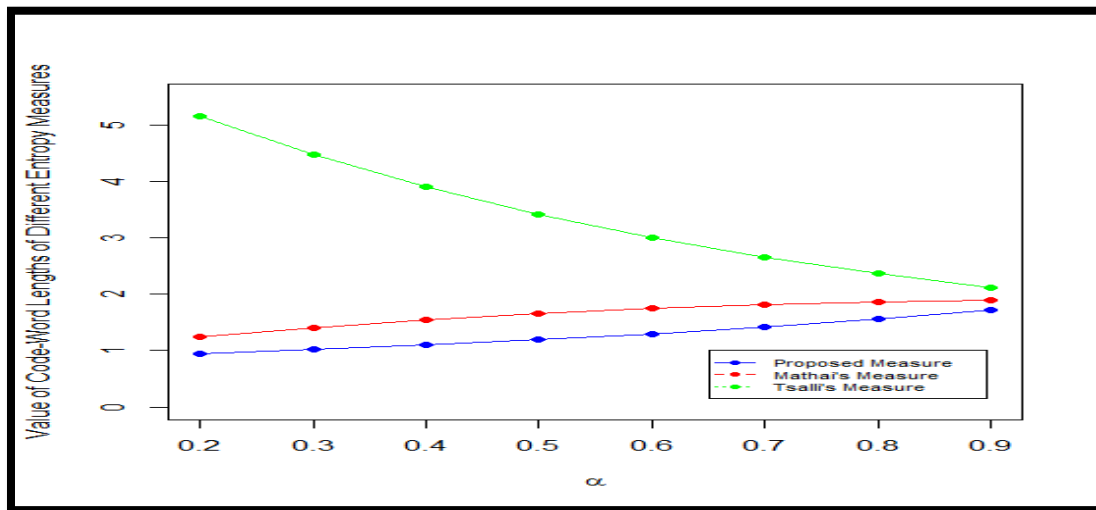


Figure 2. Graphical Overview of Code-Word Lengths of Different Entropy Measures

8. CONCLUSION

In this research article, we proposed a new two parametric measure of Entropy $J_{\alpha\beta}(P)$ and its corresponding code word length $L_{\alpha\beta}(P)$ and we analyzed its main properties in the graphic-theoretic setting. We also compare the proposed entropy with some already existing entropy measure observed that average code word length in case of proposed two parametric entropy measure is less than that of compared entropy measures under Shannon Fanno encoding scheme which makes it more efficient in terms of efficiency and redundancy of transmission/communication system.

9. REFERENCES

- [1] Baig M.A.K. and Javid M., [2013]: Some New Generalizations of Fuzzy Average Codeword Length and Their Bounds, American Journal of Applied Mathematics and Statistics, Volume 2, No. 2, pp. 73 – 76.
- [2] Campbell, L.L. [1965]: A coding theorem and Renyi's entropy, Information and Control, Vol. 8, pp. 423-429.
- [3] Havrada, J. H. and Charvat, F. [1967]: Quantification methods of classificatory processes, the concepts of structural α entropy, Kybernetika, Vol.3, pp. 30-35..
- [4] Kapur, J. N. [1986]: A generalization of Campbell's noiseless coding theorem, Jour. Bihar Math, Society, Vol.10, pp.1-10.
- [5] Kapur, J. N. [1998]: Entropy and Coding, Mathematical Science Trust Society, New Delhi.
- [6] Kraft, L.J. [1949]: A device for quantizing grouping and coding amplitude modulates pulses. M.S. Thesis, Department of of Electrical Engineering, MIT, Cambridge.
- [7] Mathai, A.M. and Rathie, P.N. [1975]: Basic Concept in Information Theory and Statistics. Wiley Eastern Limited, New Delhi.
- [8] Renyi, A. [1961]: On measures of entropy and information. Proceedings 4th Berkeley Symposium on Mathematical Statistics and Probability, Vol.1, pp.541-561.
- [9] Shannon, C. E. [1948]: A mathematical theory of communication. Bell System Technical Journal, Vol.27, pp.379-423, 623-659.
- [10] Shannon, C. E. [1959]: Coding theorems for a discrete source with a fidelity criterion. In IRE National Convention Record, Part 4, pp142-163,
- [11] Sharma, B.D. and Taneja, I. J.: Entropies of type α, β and other generalized measures of information theory, Mathematika, Vol.22, pp. 205-215.
- [12] Tsalli's C. [1988]: Possible Generalization of Boltzmann-Gibbs statistics, Vol. 52, pp. 479-487.