Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

UNIVERSAL WISER PUBLISHER

Research Article

Novel Generalized Entropy Measure's Characteristics Associated with Code-Word Length

Aakanksha Dwivedi[*], **R. N. Saraswat**

Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur, India
E-mail: aakanksha.202505027@muj.manipal.edu

Received: 26 June 2023; **Revised:** 23 October 2023; **Accepted:** 25 December 2023

Abstract: In this article, we introduce a novel, all-purpose entropy measure and derive a new, all-purpose average codeword length in accordance with it. After that, we determined new average generalized codeword length constraints in terms of additional metrics of generalized entropy. In addition, we demonstrate that the metrics employed in this study are generalizations of other metrics typically employed in the coding and information theory communities.

Keywords: mean code-word length, Shannon's entropy, Holder inequality, Kraft's inequality, Huffman and Shannon-Fano codes

MSC: 94A17

1. Introduction

Roads have speed limits; certain movies have age restrictions and the time it takes you to walk to the park are all examples of inequalities. Inequalities do not represent an exact amount but instead represent a limit of what is allowed or possible.

1.1 *Convex function*

A convex function is a function whose inscription is a convex set. Mathematically, a function $f: I \subseteq R \to R$ is called convex function if $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$ holds $\forall x, y \in I$ and all $\lambda \in [0, 1]$.

1.2 *Jensen inequality*

A real valued function f defined on an interval I is convex if $\forall x_1, x_2, ..., x_n \in I$ and all scalars $\forall \lambda_1, \lambda_2, ..., \lambda_n \in [0, 1]$ with $\sum_{i=1}^{n} \lambda_i = 1$, we have

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)$$

DOI: https://doi.org/10.37256/cm.5120243281

In his essay "A Mathematical Theory of Communication", Shannon [1] introduced the concept of entropy. The uncertainty of a random variable is quantified by its entropy, as defined by Wikipedia. Shannon's entropy is provided a numerical measure of the amount of information in a message, often in bits, where "message" refers to a specific realisation of a Random Variable (RV). Similarly, the Shannon entropy quantifies the typical loss of data due to uncertainty in the value of a random variable. The groundwork that entropy provided for our current knowledge of communication theory is immense. In the past fifty years, Shannon's entropy has been one of the most significant contributions to our understanding of probabilistic instability. Statistical thermodynamics, image reconstruction, business, urban and regional planning, money, operation research, queuing theory, biology, spectrum analysis, and even manufacturing have all been linked to the idea of entropy.

Assuming that $y_1, y_2, ..., y_n$ are the discrete random variable which is the subset of discrete random variable set Y and corresponding probabilities $\rho_1, \rho_2, ..., \rho_n, p_i \geq 0$ where i is 1 to n and $\sum_{i=1}^{n} p_i = 1$, we write

$$E = [y_1, y_2, ..., y_n;\ \rho_1, \rho_2, ..., \rho_n] \tag{1}$$

Scheme (1) is referred to as the finite information scheme. Entropy, a measure of uncertainty or information that Shannon [1] developed, is connected to a finite information scheme (1).

$$H(\rho) = E[I(y_i)] = H(\rho_1, \rho_2, ..., \rho_n) = -\sum_{i=1}^{n} \rho_i \log \rho_i \tag{2}$$

Let $\rho_1, \rho_2, ..., \rho_n$ are the set of probabilities and corresponding cord word lengths (to be transmitted) are $\ell_1, \ell_2, ..., \ell_n$ satisfied Kraft inequality [2],

$$\sum_{i=1}^{n} D^{-\ell_i} \leq 1 \tag{3}$$

Where D is the code alphabet's size.

Using equation (3), apply the properties of uniquely decipherable codes Shannon [1], then we the lower bound of mean code-word length (MCWL),

$$L = \sum_{i=1}^{n} \rho_i \ell_i \tag{4}$$

It is lies between $H(\rho)$ and $H(\rho) + 1$. When is specified in (2). The length of the average exponentiated codeword, as determined by Campbell [3], was described as

$$L_\mu = \frac{\mu}{1-\mu} \log_D \left[\sum_{i=1}^{n} \rho_i D^{-\ell_i \left(\frac{\mu-1}{\mu} \right)} \right],\ \mu > 0,\ \mu \neq 1 \tag{5}$$

and using the equation (3), the minimum value of (5) lies between $T_\mu(\rho)$ and $T_\mu(\rho) + 1$.

$$T_\mu = \frac{1}{1-\mu} \log_D \left[\sum_{i=1}^{n} \rho_i^\mu - 1 \right],\ \mu > 0,\ \mu \neq 1 \tag{6}$$

which is Tsallis [4] entropy of order μ.

Tsallis entropy can be applied in many fields like statical physics to measure disorder of a system and rate of a system irreversibility. Several generalized entropy measures have been explored by different researchers, and in accordance with these measures, generalized code-word lengths and theorems under the constraint of uniquely decipherability were developed, e.g., Nath's published article [5] Inaccuracy and coding theory. The weighted entropy discussed by Belis and Guiasu [6], Longo [7] estimated the minimal value of useful mean code-word length. Guiasu and Picard [8] established noiseless coding theorem (NCT) and 60 also discussed smallest value other average code-word length. The average codeword length and bound which is useful, investigated by Gurdial and Pessoa [9]. Jain and Tuteja [10] investigated Different generalized coding theorems and various applications by various well-known authors like Taneja et al [11], Bhatia [12], Hooda [13], Khan et al. [14], Bhat [15-19].

In recent years, some researchers have been focused on the aspects of "Convolutional Neural Network" (CNN) architectures like optimizer [20], various work have been presented to develop CNN models according to loss functions [21], Shannon's entropy used to remove and extract meaningful information from optical patterns. Results showed significantly better accuracy using Shannon's entropy as a segmentation process [22].

2. Bounds on new average code-word length's

Entropy measure is a key measure in information theory. Entropy gauges the unpredictability associated with the value of a random variable or the outcome of a random process. Like Shannon entropy measure (2), here we discuss a new generalized entropy measure which is given by

$$H_{\mu}^{\nu}(\rho)=\frac{1}{\nu(1-\mu)}\left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}-1\right], \tag{7}$$

$$0<\mu<1,\ \nu\geq 1,\ \rho_i\geq 0,\ \forall i=1,2,...,n,\ \sum_{i=1}^{n}\rho_i=1.$$

2.1 *Remarks for (7)*

I - If $\nu = 1$ then the equation (7) gives the following Tsallis entropy [4], i.e.

$$H_{\mu}(\rho)=\frac{1}{(1-\mu)}\left[\sum_{i=1}^{n}\rho_i^{\mu}-1\right],\ 0<\mu<1 \tag{8}$$

II - If $\nu = 1$ and $\mu \to 1$ then the equation (7) gives the following Tsallis entropy [4], i.e.

$$H(\rho)=-\sum_{i=1}^{n}\rho_i\log\rho_i$$

A new generalized average code-word length based on (7)

$$L_{\mu}^{\nu}(\rho)=\frac{\nu}{1-\mu}\left[\left[\sum_{i=1}^{n}\rho_i^{\frac{1}{\nu}}D^{-\ell_i\left(\frac{\mu-1}{\mu}\right)}\right]^{\mu}-1\right],\ 0<\mu<1,\ 0<\nu\leq 1,\ \nu>\mu \tag{9}$$

here D is the size of code alphabet.

2.2 *Remarks for (9)*

I - If $v = 1$ (9) and from equation (9), we get the code word length i.e.

$$L_\mu(\rho) = \frac{1}{1-\mu}\left[\left[\sum_{i=1}^{n} \rho_i D^{-\ell_i\left(\frac{\mu-1}{\mu}\right)}\right]^{\mu} - 1\right],$$

II - If $v = 1$ and $\mu \to 1$ using the equation (9), we get optimal code-word length

$$\sum_{i=1}^{n} \rho_i \ell_i \tag{10}$$

Now we got the bounds of (9) in terms of (7), under the following condition

$$\sum_{i=1}^{n} D^{-\ell_i} \leq 1 \tag{11}$$

The equation number (11) is the Kraft's inequality [2].

Theorem 2.1 Let $\ell_1, \ell_2, ..., \ell_n$ be the code-word lengths and satisfies the equation (11), then the equation (9) satisfies the following inequality

$$L_\mu^v(\rho) \geq H_\mu^v(\rho),\ 0 < \mu < 1,\ v \geq 1 \tag{12}$$

where equality holds well if

$$\ell_i = -\log D\left[\frac{\rho_i^{\frac{\mu}{v}}}{\sum_{i=1}^{n} \rho_i^{\frac{\mu}{v}}}\right], \tag{13}$$

Proof. We know that

$$\sum_{i=1}^{n}(x_i y_i) \geq \left[\sum_{i=1}^{n}(x_i^{\rho})\right]^{\frac{1}{\rho}}\left[\sum_{i=1}^{n}(y_i^{\delta})\right]^{\frac{1}{\delta}} \tag{14}$$

which is the reverse Holder inequality.

For all$(x_i\, y_i) > 0$, $i = 1, 2, 3, ...$ and $\frac{1}{\rho} + \frac{1}{\delta} = 1$, $\rho < 1\ (\neq 0)$, $\delta < 0$, or $\rho > 0\ (\neq 0)$.

The equality holds if a positive constant k occurs, we get

$$(x_i^{\rho}) = k(y_i^{\delta}) \tag{15}$$

and

$$x_i = \rho_i^{\frac{\mu}{\nu(\mu-1)}} D^{-l_i},\ y_i = \rho_i^{\frac{\mu}{\nu(1-\mu)}},\ \rho = \frac{\mu-1}{\mu} \text{ and } \delta = 1-\mu$$

from the equation (14) in (13), we obtain

$$\sum_{i=1}^{n} D^{-l_i} \geq \left[\sum_{i=1}^{n} \rho_i^{\frac{1}{\nu}} D^{-l_i \frac{\mu-1}{\mu}}\right]^{\frac{\mu}{\mu-1}} \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}}\right]^{\frac{1}{1-\mu}} \tag{16}$$

Now using the inequality (11) we get

$$\left[\sum_{i=1}^{n} \rho_i^{\frac{1}{\nu}} D^{-l_i \frac{\mu-1}{\mu}}\right]^{\frac{\mu}{\mu-1}} \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}}\right]^{\frac{1}{1-\mu}} \leq 1 \tag{17}$$

Or equivalently (17), can be written as

$$\left[\sum_{i=1}^{n} \rho_i^{\frac{1}{\nu}} D^{-l_i \frac{\mu-1}{\mu}}\right]^{\frac{\mu}{\mu-1}} \leq \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}}\right]^{\frac{1}{1-\mu}} \tag{18}$$

Here the following new particular cases.

Case 1

If $0 < \mu < 1$, $\nu \geq 1$ then $(\mu - 1) < 0$, taking power both sides of $(\mu - 1)$ then using the inequality (18), we get

$$\left[\sum_i^n \rho_i^{\frac{1}{\nu}} D^{-l_i \frac{\mu-1}{\mu}}\right]^{\mu} \geq \left[\sum_i^n \rho_i^{\frac{\mu}{\nu}}\right]$$

$$\left[\sum_{i}^{n} \rho_i^{\frac{1}{\nu}} D^{-l_i \frac{\mu-1}{\mu}}\right]^{\mu} - 1 \leq \left[\sum_{i}^{n} \rho_i^{\frac{\mu}{\nu}}\right] - 1 \tag{19}$$

As $0 < \mu < 1$, $\nu \geq 1$ then $\nu(1 - \mu) > 0$ and $\frac{1}{\nu(1-\mu)} > 0$ multiply inequality (19), throughout by $\frac{1}{\nu(1-\mu)} > 0$, we get

$$\frac{1}{\nu(1-\mu)}\left[\left[\sum_{i}^{n} \rho_i^{\frac{1}{\nu}} D^{-l_i \frac{\mu-1}{\mu}}\right]^{\mu} - 1\right] \geq \frac{1}{\nu(1-\mu)}\left[\left[\sum_{i}^{n} \rho_i^{\frac{\mu}{\nu}}\right] - 1\right] \tag{20}$$

and $L_\mu^\nu \rho \geq H_\mu^\nu \rho$, $0 < \mu < 1$, $\nu \geq 1$ Hence proved the result.

Case 2

From equation (13),

$$\ell_i = -\log_D\left[\frac{\rho_i^{\frac{\mu}{\nu}}}{\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}}\right]$$

Alternatively, the above-mentioned equation can be expressed as

$$D^{-l_i} = \frac{\rho_i^{\frac{\mu}{\nu}}}{\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}} \tag{21}$$

After properly simplifying equation (21) and taking $\frac{\mu-1}{\mu}$, we obtain.

$$D^{-l_i\frac{\mu-1}{\mu}} = \rho_i^{\frac{\mu-1}{\nu}}\left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}\right]^{\frac{1-\mu}{\nu}} \tag{22}$$

By multiplying both sides of equation (22) by $\rho_i^{\frac{1}{\nu}}$ then taking summation from 1 to n and finally performing certain simplifications, we obtain

$$\sum_{i=1}^{n}\rho_i^{\frac{1}{\nu}}D^{-l_i\frac{\mu-1}{\mu}} = \left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}\right]^{\frac{1}{\mu}} \tag{23}$$

Taking both the sides to the power μ in the equation (23), we obtain

$$\left[\sum_{i=1}^{n}\rho_i^{\frac{1}{\nu}}D^{-l_i\frac{\mu-1}{\mu}}\right]^{\mu} = \left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}\right]$$

$$\left[\left(\sum_{i=1}^{n}\rho_i^{\frac{1}{\nu}}D^{-l_i\frac{\mu-1}{\mu}}\right)^{\mu} - 1\right] = \left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}\right] - 1$$

Now multiply both sides by $\frac{1}{\nu(1-\mu)}$, we obtain

$$L_\mu^\nu(\rho) = H_\mu^\nu(\rho)$$

Hence the result.

Theorem 2.2 Let ℓ_1, ℓ_2, ..., ℓ_n be the series of code-word lengths and $L_\mu^\nu(\rho)$ be the Kraft inequality, then the inequality holds:

$$L_\mu^\nu(\rho) < (H_\mu^\nu(\rho) + \xi) D^{(1-\mu)}, \text{ where } 0 < \mu < 1,\ \nu \geq 1,\ \xi = \frac{1}{\nu(1-\mu)} > 0, \tag{24}$$

Proof. From (7)

$$L_\mu^\nu(\rho) = H_\mu^\nu(\rho)$$

If

$$D^{-\ell_i} = \frac{\rho_i^{\frac{\mu}{\nu}}}{\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}},\ 0 < \mu < 1,\ \nu \geq 1$$

Alternatively, the abovementioned equation can be expressed as

$$\ell_i = -\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right]$$

In order to ensure that the inequality holds, we select the code-word lengths as ℓ_i and the value of i is equal to 1 to n

$$-\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right] \leq \ell_i < -\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right] + 1 \tag{25}$$

Consider the interval

$$\xi_i = \left[-\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right],\ -\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right] + 1\right]$$

having length 1. So, $\forall \xi$, $\exists$ a positive integer ℓ_i s.t.

$$0 < -\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right] \leq \ell_i - \log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right] + 1 \tag{26}$$

We will first show that $\ell_1, \ell_2, ..., \ell_n$ as specified, satisfies the Kraft [2] inequality. From the left part of the inequality (26), we get

$$-\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^n \rho_i^{\frac{\mu}{\nu}}\right] \leq \ell_i$$

Alternatively, the abovementioned equation can be expressed as

$$D^{-l_i} \leq \frac{\rho_i^{\frac{\mu}{\nu}}}{\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}}} \tag{27}$$

Taking summation over i is from 1 to n on both sides to the inequality (27), we get $D^{-l_i} \leq 1$, which is the Kraft's [2] inequality equation (26) provides

$$\ell_i \leq -\log_D \rho_i^{\frac{\mu}{\nu}} + \log_D \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} \right] + 1$$

Or

$$D^{-l_i} < \left[\frac{\rho_i^{\frac{\mu}{\nu}}}{\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}}} \right]^{-1} D \tag{28}$$

As $0 < \mu < 1$ then $(1 - \mu) > 0$ and $\frac{1-\mu}{\mu} > 0$ multiply inequality (28), throughout by $\frac{1-\mu}{\mu} > 0$, we get

$$D^{l_i} \frac{1-\mu}{\mu} < \left[\frac{\rho_i^{\frac{\mu}{\nu}}}{\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}}} \right]^{\frac{\mu-1}{\mu}} D^{\frac{1-\mu}{\mu}}$$

Alternatively, the above mentioned equation can be expressed as

$$D^{-l_i\left(\frac{\mu-1}{\mu}\right)} < \rho_i^{\frac{\mu-1}{\nu}} \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} \right]^{\frac{1-\mu}{\mu}} D^{\frac{1-\mu}{\mu}} \tag{29}$$

Multiply inequality (29), both sides by $\rho_i^{\frac{\mu}{\nu}}$ then taking summation i, 1 to n, we obtain

$$\sum_{i=1}^{n} \rho_i^{\frac{1}{\beta}} D^{-l_i\left(\frac{\mu-1}{\mu}\right)} < \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} \right]^{\frac{1}{\mu}} D^{\frac{1-\mu}{\mu}} \tag{30}$$

As $0 < \mu < 1$ using equation (30)

$$\left[\sum_{i=1}^{n} \rho_i^{\frac{1}{\nu}} D^{-l_i\left(\frac{\mu-1}{\mu}\right)} \right]^{\mu} < \left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} \right] D^{(1-\mu)}$$

$$\left[\sum_{i=1}^{n}\rho_i^{\frac{1}{\nu}}D^{-l_i\left(\frac{\mu-1}{\mu}\right)}\right]^{\mu}-1<\left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}\right]D^{(1-\mu)}+D^{(1-\mu)}-D^{(1-\mu)}$$

$$\left[\sum_{i=1}^{n}\rho_i^{\frac{1}{\nu}}D^{-l_i\left(\frac{\mu-1}{\mu}\right)}\right]^{\mu}-1<\left[\sum_{i=1}^{n}\rho_i^{\frac{\mu}{\nu}}-1\right]D^{(1-\mu)}+D^{(1-\mu)}$$

As $0<\mu<1$, $\nu\geq 1$ then $\nu(1-\mu)<0$ and $\frac{1}{\nu(1-\mu)}>0$ multiply in the above inequality, we get

$$L_{\mu}^{\nu}(\rho)<(H_{\mu}^{\nu}(\rho)+\omega)D^{(1-\mu)}$$

and for $0<\mu<1$, $\nu\geq 1$ and $\omega=\frac{1}{\nu(1-\mu)}>0$.

Thus, we have demonstrated from the two coding theorems above that

$$H_{\mu}^{\nu}(\rho)\leq L_{\mu}^{\nu}(\rho)<(H_{\mu}^{\nu}(\rho)+\omega)D^{(1-\mu)}$$

Where $0<\mu<1$, $\nu\geq 1$ and $\omega>0$.

In the next part, using the Huffman and Shannon-Fano coding can be verified to discrete channel of noiseless coding theorems.

3. Illustration

In this part, using empirical data from Tables 1 and 2 and following the methodology of Hooda et al. [19], we demonstrate the validity of theorems 2.1 and 2.2. The value of $H_{\mu}^{\nu}(\rho)$, $L_{\mu}^{\nu}(\rho)$ $(H_{\mu}^{\nu}(\rho)+\omega)D^{(1-\mu)}$ with particular values of μ and ν for ψ, $D=2$ given in Table 1. It is based on Huffman coding scheme.

Table 1. Using Huffman coding scheme the value of $H_{\mu}^{\nu}(\rho)$, $L_{\mu}^{\nu}(\rho)$ $(H_{\mu}^{\nu}(\rho)+\omega)D^{(1-\mu)}$ and ψ for different values of μ and ν

Probabilities	Huffman Code words	ℓ_i	μ	ν	$H_{\mu}^{\nu}(\rho)$	$L_{\mu}^{\nu}(\rho)$	$\psi=\frac{H_{\mu}^{\nu}(\rho)}{L_{\mu}^{\nu}(\rho)}\times 100$	$(H_{\mu}^{\nu}(\rho)+\omega)D^{(1-\mu)}$
ρ_i								
0.3846	0	1	0.9	1	1.6466	1.7003	96.8417	12.4825
0.1795	100	3	0.5	1	2.3725	2.7714	85.6066	6.1836
0.1538	101	3						
0.1538	110	3						
0.1282	111	3						

Table 2. Using Shannon-Fano coding scheme the value of $H_\mu^v(\rho)$, $L_\mu^v(\rho)$ $(H_\mu^v(\rho)+\omega)D^{(1-\mu)}$ and ψ for different values of μ and v

Probabilities	Huffman Code words	ℓ_i	μ	v	$H_\mu^v(\rho)$	$L_\mu^v(\rho)$	$\psi = \frac{H_\mu^v(\rho)}{L_\mu^v(\rho)} \times 100$	$(H_\mu^v(\rho)+\omega)D^{(1-\mu)}$
ρ_i								
0.3846	0	1	0.9	1	1.6466	1.8035	91.3002	12.4825
0.1795	10	2	0.5	1	2.3725	3.3776	70.2442	6.1836
0.1538	110	3						
0.1538	1110	4						
0.1282	1111	4						

The values of $H_\mu^v(\rho)$, $L_\mu^v(\rho)$ $(H_\mu^v(\rho)+\omega)D^{(1-\mu)}$ and ψ for particular values of μ and v calculated by using "Shannon-Fano coding scheme" are discussed in below table.

From the Tables 1 and 2 following results have been inferred:

1. In the instances of Shannon-Fano codes and Huffman codes, theorems 2.1 and 2.2 are true i.e.,

$$H_\mu^v(\rho) \le L_\mu^v(\rho) < (H_\mu^v(\rho)+\omega)D^{(1-\mu)}, \text{ where } 0<\mu<1, v \ge 1$$

2. The mean code-word length $L_\mu^v(\rho)$ of the Huffman coding scheme is less than that of the Shannon-Fano coding scheme.

3. The Huffman code efficiency coefficient is greater than the Shannon-Fano code efficiency coefficient. In other words, it has been determined that the scheme of Huffman coding is more effective as compared to the scheme of Shannon-Fano coding.

4. Some properties of new generalized entropy measure

Using $H_\mu^v(\rho)$ we will examine a few properties of measure given in (7).

(i) $H_\mu^v(\rho)$ is non-negative.

Proof. From equation (7), we get

$$H_\mu^v(\rho) = \frac{1}{v(1-\mu)}\left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{v}} - 1\right]$$

where

$$0<\mu<1,\ v \ge 1,\ \rho_i \ge 0,\ \forall i = 1, 2, \ldots, n,\ \sum_{i=1}^{n} \rho_i = 1$$

For any values of μ and v, it is clear that,

$$\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} \geq 1$$

also we have $0 < \mu < 1$, $\nu \geq 1$ and $\dfrac{1}{\nu(1-\mu)}$ therefore, we conclude that

$$H_\mu^\nu(\rho) = \frac{1}{\nu(1-\mu)}\left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} - 1\right] \geq 0$$

(ii): $H_\mu^\nu(\rho)$ is a symmetric function ρ_i, i is from 1 to n.

Proof.

$$H_\mu^\nu(\rho_1, \rho_2, ..., \rho_{n-1}, \rho_n) = H_\mu^\nu(\rho_n, \rho_1, ..., \rho_{n-1})$$

(iii): $H_\mu^\nu(\rho)$ is maximum when $\nu = 1$, $\mu \to 1$ with equal probabilities.

Proof. When $\rho_i = \dfrac{1}{n}$, i is from 1 to n, and $\nu = 1$, $\mu \to 1$. Then $H_\mu^\nu(\rho) = \log n$ which is maximum entropy.

(iv): For $\nu = 1$, $\mu \to 1$ $H_\mu^\nu(\rho)$ is convex function for $\rho_1, \rho_2, ..., \rho_{n-1}, \rho_n$.

Proof. From (7),

$$H_\mu^\nu(\rho) = \frac{1}{\nu(1-\mu)}\left[\sum_{i=1}^{n} \rho_i^{\frac{\mu}{\nu}} - 1\right]$$

where

$$0 < \mu < 1,\ \ \nu \geq 1,\ \ \rho_i \geq 0,\ \ \forall i = 1, 2, ..., n,\ \ \sum_{i=1}^{n} \rho_i = 1$$

If $\nu = 1$, $\mu \to 1$, then we can calculate its first derivative regarding ρ_i as

$$\left[\frac{d}{d\rho} H_\mu^\nu(\rho)\right]_{\nu=1, \mu \to 1} = -1 - \log(\rho_i)$$

And the second derivative is given by

$$\left[\frac{d^2}{d\rho^2} H_\mu^\nu(\rho)\right]_{\nu=1, \mu \to 1} = -\left(\frac{1}{\rho_i}\right) < 0,\ \ \forall \rho_i \in [0, 1]$$

and i from 1 to n.

But the 2nd derivativities of $H_\mu^\nu(\rho)$, $\forall \rho_i \in [0, 1]$ and i from 1 to n. as $\nu = 1$ and $\mu \to 1$.

Therefore $H_\mu^\nu(\rho)$ is concave downward function for $\rho_1, \rho_2, ..., \rho_n$.

5. Conclusion

In this study, a new generalized entropy measure has been constructed and by using this measure, we created a new generalized average code-word length and determined its boundaries in terms of the new generalized entropy measure. In this paper, established the bounds for discrete channel. These bounds have been substantiated by considering Huffman and Shannon-Fano coding schemes by using an empirical data. The key characteristics of the new entropy measure have also been discussed.

Conflict of interest

Authors declare that there is no conflict of interest.

References

[1] Shannon CE. A mathematical theory of communication. *Bell System Technical Journal*. 1948; 27: 379-423. Available from: doi: 10.1002/j.1538-7305.1948.tb01338.x.
[2] Kraft LG. *A Device for Quantizing, Grouping and Coding Amplitude-Modulated Pulses*. Doctoral dissertation, Massachusetts Institute of Technology. Cambridge; 1949.
[3] Campbell LL. A coding theorem and Renyi's entropy. *Information and Control*. 1965; 8: 423-429.
[4] Constantino T. The nonadditive entropy sq and its applications in physics and elsewhere: Some remarks. *Entropy.* 2011; 13(10): 1765-1804. Available from: doi: 10.3390/e13101765.
[5] Nath P. On a coding theorem connected with Renyi's entropy. *Information and Control*. 1975; 29(3): 234-242.
[6] Belis M, Guiasu S. A quantitative-qualitative measure of information in cybernetic systems. *IEEE Transactions on Information Theory*. 1968; 14(4): 593-594.
[7] Longo G. A noiseless coding theorem for sources having utilities. *SIAM Journal on Applied Mathematics*. 1976; 30(4): 739-748. Available from: doi: 10.1137/0130067.
[8] Guiasu S, Picard CF. Borne inferieure de la longueur de certain codes [Lower bound of the length of certain codes]. *Comptes Rendus, Academy of Sciences, Paris*. 1971; 273(A): 248-251.
[9] Gurdial PF. On 'useful' information of order. *Journal of Combinatorics Information and System Sciences*. 1977; 2: 158-162.
[10] Jain P, Tuteja RK. On coding theorem connected with 'useful' entropy of order-β. *International Journal of Mathematics and Mathematical Sciences*. 1989; 12: 193-198. Available from: doi: 10.1155/S0161171289000232.
[11] Tenaja H, Hooda DS, Tuteja RK. Coding theorems on a generalized 'useful' information. *Soochow Journal of Mathematics*. 1985; 11: 123-131.
[12] Bhatia PK. Useful inaccuracy of order and 1.1 coding. *Soochow Journal of Mathematics*. 1995; 21(1): 81-87.
[13] Hooda DS, Bhaker US. A generalized 'useful' information measure and coding theorems. *Soochow Journal of Mathematics*. 1997; 23: 53-62.
[14] Khan AB, Bhat BA, Pirzada S. Some result on a generalized 'useful' information measure. *Journal of Inequalities in Pure Applied Mathematics*. 2005; 6(4): 117.
[15] Bhat AH, Baig MAK. Characterization of new two parametric generalized useful information measure. *Journal of Information Science Theory and Practice*. 2016; 4(4): 64-74.
[16] Bhat AH, Baig MAK. Noiseless coding theorem on new generalized useful information measure of order and type. *Asian Journal of Fuzzy and Applied Mathematics*. 2016; 4(6): 73-85.
[17] Bhat AH, Baig MAK. New generalized measure of entropy of order and type and its coding theorems. *International Journal of Information Science and System*. 2016; 5: 1-7.
[18] Bhat AH, Baig MAK. Some coding theorems on generalized reyni's entropy of order and type. *International Journal of Applied Mathematics and Information Science Letters*. 2017; 5(1): 13-19.
[19] Hooda DS, Upadhyay K, Saxena S. Characterization of a generalized information measure by optimization technique. *International Journal of Engineering Research and Management Technology*. 2014; 1: 302-311.
[20] Rajesh G, Sunoj S. Some properties of cumulative Tsallis entropy of order alpha. *Statistical Papers*. 2019; 60(3): 583-593.

[21] Mohamed MS, Barakat HM, Alyami SA, Abd Elgawad MA. Cumulative residual tsallis entropy-based test of uniformity and some new findings. *Mathematics*. 2022; 10(5): 771. Available from: doi: 10.3390/math10050771.

[22] Garcia-González W, Flores-Fuentes W, Sergiyenko O, Rodríguez-Quiñonez JC, Miranda-Vega JE, Hernández-Balbuena D. Shannon entropy used for feature extractions of optical patterns in the context of structural health monitoring. *Entropy*. 2023; 25(8): 1207. Available from: doi: 10.3390/e25081207.