AMO - Advanced Modeling and Optimization, Volume 14, Number 3, 2012

# Some more remarks on a generalized 'useful' Havrda-Charvat and Tsallis's entropy

**Satish Kumar**

Department of Mathematics, G.I.M.T., Kanipla, Kurukshetra, Haryana (India).
E-mail:drsatish74@rediffmail.com

**Rajesh Kumar**

Department of Mathematics, Hindu College, University of Delhi, Delhi-7 (India).
E-mail:rajeshhctm@rediffmail.com

**Abstract.** A parametric mean length is defined as the quantity

$$L_\alpha^\beta(U;P)=\frac{1}{\alpha-1}\left[1-\left(\sum P_i^\beta\left(\frac{u_i}{\sum u_i\, p_i^\beta}\right)^{\frac{1}{\alpha}} D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)}\right)^\alpha\right],$$

where $\alpha>0\ (\neq 1)$, $\beta>0$, $u_i>0$ and $\sum p_i=1$. This being the useful mean length of code words weighted by utilities, $u_i$. Lower and upper bounds for $L_\alpha^\beta(U;P)$ are derived in terms of 'useful' information measure for the incomplete power distribution, $p^\beta$.

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

**Key words:** Tsallis's Entropy, 'Useful' Information measure, Utilities and Power probabilities.

## 1. Introduction

Consider the following model for a random experiment S,

$$S_N=[E;\ P;\ U]$$

where $E=(E_1,\ E_2,\ ....,\ E_N)$ is a finite system of events happening with

respective probabilities $P=(p_1,p_2,...,p_N)$, $p_i\geq 0$, and $\sum p_i=1$ and credited with utilities $U=(u_1,\ u_2,\ ....,\ u_N)$, $u_i>0$, $i=1,\ 2,...,\ N$. Denote the model by *E*,

Corresponding Author: **Satish Kumar** (Email Id: drsatish74@rediffmail.com)

where, $$S_N = \begin{bmatrix} E_1\ E_2\ \ldots.\ E_N \\ p_1\ p_2\ \ldots.\ p_N \\ u_1\ u_2\ \ldots.\ u_N \end{bmatrix} \qquad \text{...(1.1)}$$

We call (1.1) a Utility Information Scheme (UIS). [Belis and Guiasu, 1968] proposed a measure of information called 'useful information' for this scheme, given by

$$H(U;\ P) = -\sum u_i\ p_i \log p_i\ , \qquad \ldots(1.2)$$

where $H(U;\ P)$ reduces to [Shannon's, 1948] entropy when the utility aspect of the scheme is ignored i.e., when $u_i = 1$ for each $i$. Throughout the paper, $\sum$ will stand for $\sum_{i=1}^{N}$ unless otherwise stated and logarithms are taken to base $D(D > 1)$.

[Guiasu and Picard, 1971] considered the problem of encoding the outcomes in (1.1) by means of a prefix code with codewords $w_1, w_2, \ldots., w_N$ having lengths $n_1, n_2, \ldots., n_N$ and satisfying [Kraft's, 1949] inequality.

$$\sum_{i=1}^{N} D^{-n_i} \leq 1. \qquad \text{...(1.3)}$$

where $D$ is the size of the code alphabet. The useful mean length $L(U;P)$ of code was defined as:

$$L(U;P) = \frac{\sum u_i\, n_i\, p_i}{\sum u_i p_i} \qquad \text{...(1.4)}$$

and the authors obtained bounds for it in terms of $H(U;\ P)$. [Longo, 1976, Gurdial and Pessoa, 1977, Autar and Khan, 1989, Singh et al., 2003, Jain and Tuteja, 1989] have studied generalized coding theorems by considering different generalized measures of (1.2) and (1.4) under condition (1.3) of unique decipherability.

In this paper, we study some coding theorems by considering a new function depending on the parameters $\alpha$ and $\beta$ and a utility function. Our motivation for studying this new function is that it generalizes 'useful' information measure already existing in the paper [Tsallis's, 1988 and Arndt, 2001] entropy, which is used in physics.

## 2. Coding Theorems

In this section, we define generalized 'useful' information measure as:

$$H_\alpha^\beta(U;\ P) = \frac{1}{\alpha - 1}\left[1 - \left(\frac{\sum u_i\ p_i^{\alpha\beta}}{\sum u_i\ p_i^{\beta}}\right)\right] \qquad \ldots(2.1)$$

where $\alpha > 0(\neq 1), \beta > 0$ , $\sum p_i = 1$

(i) When $\beta$ = 1 then (2.1) reduces to 'useful' information measure studied by [Hooda and Ram, 2002].

i.e., $$H_\alpha(U;\,P) = \frac{1}{\alpha-1}\left[1-\left(\frac{\sum u_i\ p_i^\alpha}{\sum u_i\ p_i}\right)\right] \qquad ...(2.2)$$

(ii) When $u_i = 1$ then (2.1) reduces to new generalized information measure of order $\alpha$ and type $\beta$.

i.e., $$H_\alpha^\beta(P) = \frac{1}{\alpha-1}\left[1-\left(\frac{\sum p_i^{\alpha\beta}}{\sum p_i^\beta}\right)\right] \qquad ...(2.3)$$

(iii) When $u_i = 1$ and $\beta$ = 1, (2.1) reduces entropy as considered by [Tsallis, 1988].

i.e., $$H_\alpha(P) = \frac{1}{\alpha-1}\left[1-\left(\sum p_i^\alpha\right)\right] \qquad ...(2.4)$$

The measure (2.4) was characterized by many authors by different approaches. [Harvda and Charvat, 1967] characterized (2.4) by an axiomatic approach. [Darcozy, 1970] studied by a functional equation.

(iv) When $\beta$ = 1 and $\alpha \to 1$, (2.1) reduces to a measure of ‘useful’ information for the incomplete distribution due to [Belis and Guiasu, 1968].

i.e., $$H(U;P) = -\frac{\sum u_i p_i \log p_i}{\sum u_i p_i} \qquad \ldots(2.5)$$

(v) When $u_i = 1$ for each $i$, i.e. when the utility aspect is ignored and $\alpha \to 1$, the measure (2.1) reduces to entropy considered by [Mathur and Mitter, 1972] for $\beta$ - power distribution.

i.e., $$H^\beta(P) = -\frac{\sum p_i^\beta \log p_i^\beta}{\sum p_i^\beta} \qquad ...(2.6)$$

(vi) When $u_i = 1, \beta$ = 1 and $\alpha \to 1$, then (2.1) reduces to [Shannon’s, 1948] entropy.

$$H(P) = -\sum p_i \log p_i \qquad \ldots(2.7)$$

(vii) When $\alpha \to 1$,then (2.1) becomes generalized ‘useful’ information mesure of $\beta$-power distribution.

i.e., $${}_\beta H(U;P) = -\frac{\sum u_i p_i^\beta \log p_i^\beta}{\sum u_i p_i^\beta} \qquad \ldots(2.8)$$

**Further consider**

**Definition:** The generalized ‘useful’ mean length $L_\alpha^\beta(U;P)$ with respect to ‘useful’ information measure is defined as :

$$L_\alpha^\beta(U;P) = \frac{1}{\alpha-1}\left[1-\left(\sum P_i^\beta\left(\frac{u_i}{\sum u_i\ p_i^\beta}\right)^{\frac{1}{\alpha}} D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)}\right)^\alpha\right] \qquad ...(2.9)$$

where $\alpha > 0(\neq 1), \beta > 0, p_i > 0, \sum p_i = 1,\ i = 1, 2, \ldots\ldots\ldots\ldots, N$

(i) For $\beta$ = 1 ,then (2.9) reduces to the new useful mean length.

i.e., $$L_\alpha(U;P)=\frac{1}{\alpha-1}\left[1-\left(\sum p_i\left(\frac{u_i}{\sum u_i\, p_i}\right)^{\frac{1}{\alpha}} D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)}\right)^{\alpha}\right] \qquad ...(2.10)$$

(ii) For $u_i = 1$ for each *i*, then (2.9) becomes new optimal code length

i.e., $$L_\alpha^\beta(P)=\frac{1}{\alpha-1}\left[1-\left(\sum p_i^\beta\left(\frac{1}{\sum p_i^\beta}\right)^{\frac{1}{\alpha}} D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)}\right)^{\alpha}\right] \qquad ...(2.11)$$

(iii) For $\beta$ = 1, $u_i = 1$ and $\alpha \to 1$ then (2.9) reduced to $L$ considered by [Shannon, 1948].

i.e., $$L=\sum n_i p_i \qquad ...(2.12)$$

(iv) For $\beta$ = 1 and $u_i = 1$ for each *i*, then (2.9) becomes new optimal code length

i.e., $$L_\alpha(P)=\frac{1}{\alpha-1}\left[1-\left(\sum p_i\; D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)}\right)^{\alpha}\right] \qquad ...(2.13)$$

We establish a result, that in a sense, provides a characterization of $H_\alpha^\beta(U;\, P)$ under the condition of unique decipherability.

**Theorem 2.1** For all integers $D > 1$

$$L_\alpha^\beta(U;P) \geq H_\alpha^\beta(U;\, P) \qquad ...(2.14)$$

under the condition (1.3). Equality holds if and only if

$$n_i = -\log_D\left(\frac{u_i\, p_i^{\alpha\beta}}{\sum u_i\, p_i^{\alpha\beta}}\right) \qquad ...(2.15)$$

**Proof:** We use [Holder's, 1967] inequality

$$\sum x_i y_i \geq \left(\sum x_i^p\right)^{\frac{1}{p}} \left(\sum y_i^q\right)^{\frac{1}{q}} \qquad ...(2.16)$$

for all $x_i \geq 0$, $y_i \geq 0, i = 1, 2,$ ......, $N$ when $P < 1$ $(\neq 1)$ and $p^{-1}+q^{-1}=1$, with equality if and only if there exists a positive number *c* such that

$$x_i^p = c y_i^q\,. \qquad ...(2.17)$$

Setting

$$x_i = p_i^{\frac{\alpha\beta}{\alpha-1}}\left(\frac{u_i}{\sum u_i\, p_i^\beta}\right)^{\frac{1}{\alpha-1}} D^{-n_i},$$

$$y_i = p_i^{\frac{\alpha\beta}{1-\alpha}}\left(\frac{u_i}{\sum u_i\, p_i^\beta}\right)^{\frac{1}{1-\alpha}},$$

$p = 1 - \frac{1}{\alpha}$ and $q = 1 - \alpha$ in (2.16) and using (1.3) we obtain the result (2.14) after simplification for $\frac{1}{\alpha - 1} > 0$ as $\alpha > 1$.

**Theorem 2.2** For every code with lengths $\{n_i\}$, $i = 1, 2, ..., N$, $L_\alpha^\beta(U;P)$ can be made to satisfy,

$$L_\alpha^\beta(U;P) < H_\alpha^\beta(U;P) D^{(1-\alpha)} + \frac{1}{\alpha - 1}\left[1 - D^{(1-\alpha)}\right] \qquad ...(2.18)$$

**Proof:** Let $n_i$ be the positive integer satisfying, the inequality

$$-\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) \le n_i < -\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) + 1 \qquad ...(2.19)$$

Consider the intervals

$$\delta_i = \left[-\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right), -\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) + 1\right] \qquad ...(2.20)$$

of length 1. In every $\delta_i$, there lies exactly one positive number $n_i$ such that

$$0 < -\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) \le n_i < -\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) + 1 \qquad ...(2.21)$$

It can be shown that the sequence $\{n_i\}$, $i = 1, 2, ..., N$ thus defined, satisfies (1.3). From (2.21) we have

$$n_i < -\log_D\left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) + 1$$

$$\Rightarrow D^{-n_i} > \left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right) D^{-1}$$

$$\Rightarrow D^{-n_i\left(\frac{\alpha-1}{\alpha}\right)} > \left(\frac{u_i p_i^{\alpha\beta}}{\sum u_i p_i^{\alpha\beta}}\right)^{\frac{\alpha-1}{\alpha}} D^{\frac{1-\alpha}{\alpha}} \qquad ...(2.22)$$

Multiplying both sides of (2.22) by $p_i^\beta\left(\frac{u_i}{\sum u_i p_i^\beta}\right)^{\frac{1}{\alpha}}$, summing over $i = 1, 2, ...., N$ and simplification for $\frac{1}{\alpha - 1} > 0$ as $\alpha > 1$, gives (2.18).

**Theorem 2.3** For every code with lengths $\{n_i\}$, $i = 1, 2, ...., N$, of Theorem 2.1, $L_\alpha^\beta(U;P)$ can be made to satisfy,

$$L_\alpha^\beta(U;P) \ge H_\alpha^\beta(U;P) > H_\alpha^\beta(U;P) D + \frac{\alpha}{\alpha - 1}(1 - D) \qquad ...(2.23)$$

**Proof:** Suppose

$$\bar{n}_i = -\log_D\left(\frac{u_i\, p_i^{\alpha\beta}}{\sum u_i\, p_i^{\alpha\beta}}\right) \qquad ...(2.24)$$

Clearly $\bar{n}_i$ and $\bar{n}_i$ + 1 satisfy ‘equality’ in Holder’s inequality (2.16). Moreover, $\bar{n}_i$ satisfies Kraft’s inequality (1.3).

Suppose $n_i$ is the unique integer between $\bar{n}_i$ and $\bar{n}_i$ + 1, then obviously, $n_i$ satisfied (1.3).

Since $\alpha > 0\ (\neq 1)$, we have

$$\left(\sum p^{\beta}{}_i\left(\frac{u_i}{\sum u_i\, p^{\beta}{}_i}\right)^{\frac{1}{\alpha}} D^{-n_i\frac{(\alpha-1)}{\alpha}}\right)^{\alpha}$$

$$\leq\left(\sum p_i^{\beta}\left(\frac{u_i}{\sum u_i\, p_i^{\beta}}\right)^{\frac{1}{\alpha}} D^{-\bar{n}_i\frac{(\alpha-1)}{\alpha}}\right)^{\alpha}$$

$$< D\left(\sum p^{\beta}{}_i\left(\frac{u_i}{\sum u_i\, p^{\beta}{}_i}\right)^{\frac{1}{\alpha}} D^{-\bar{n}_i\frac{(\alpha-1)}{\alpha}}\right)^{\alpha} \qquad \ldots(2.25)$$

Since, $$\left(\sum p_i^{\beta}\left(\frac{u_i}{\sum u_i\, p_i^{\beta}}\right)^{\frac{1}{\alpha}} D^{-\bar{n}_i\frac{(\alpha-1)}{\alpha}}\right)^{\alpha} = \left(\frac{\sum u_i\, p_i^{\alpha\beta}}{\sum u_i\, p_i^{\beta}}\right)$$

Hence, (2.25) becomes

$$\left(\sum p_i^{\beta}\left(\frac{u_i}{\sum u_i\, p_i^{\beta}}\right)^{\frac{1}{\alpha}} D^{-n_i\frac{(\alpha-1)}{\alpha}}\right)^{\alpha} \leq \left(\frac{\sum u_i\, p_i^{\alpha\beta}}{\sum u_i\, p_i^{\beta}}\right) < D\left(\frac{\sum u_i\, p_i^{\alpha\beta}}{\sum u_i\, p_i^{\beta}}\right)$$

which gives the result (2.23).

**Conclusion and Discussion:**

The problem of coding is that of associating the messages that have to be sent with the sequences of symbols in a one to one fashion. In coding theory, generally we come across the problem of efficient coding of messages to be sent over a noiseless channel. We do not consider the problem of error correction, but attempt to maximize the number of messages that can be sent through a channel in a given time. Therefore, we find the minimum value of a mean codeword length subject to a given constraint on codeword lengths. As the codeword lengths are integers, the minimum value lies between two bounds, so a noiseless coding theorem seeks to find these bounds for a given mean and a given constraint. For uniquely decipherable codes, [Shannon, 1948] found the lower bounds for the arithmetic mean by using his entropy. [Longo, 1976] obtained the lower bound for useful mean codeword length in terms of quantitative-qualitative measure of entropy, introduced by [Belis and Guiasa, 1968]. [Guiasa and Picard, 1971] proved a noiseless coding theorem and obtained the lower bounds for similar useful mean codeword length. [Gurdial and Pessoa, 1977] extended this by finding lower bounds for useful mean codeword length of order $\alpha$ in terms of useful measures of information of order $\alpha$. It is important to note that for other standard means like the geometric mean, the harmonic mean and the power mean, the lower bounds can be found in principle, but except for the arithmetic mean, no closed expressions for the lower bounds can be obtained. [Kraft’s, 1949] inequality plays an important role in proving the noiseless coding theorem. It is uniquely determined by the condition for unique decipherability.

We know that optimal code is that code for which the value $L_\alpha^\beta(U;P)$ is equal to its lower bound. From the result of the theorem 2.1, it can be seen that the mean codeword length of the optimal code is dependent on two parameters $\alpha, \beta$ and a utility function, while in the case of Shannon's theorem it does not depend on any parameter . So it can be reduced significantly by taking suitable values of parameters.

## References


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

[2] Autar, R. and Khan, A.B. (1989), On Generalized`Useful’Information for Incomplete distribution. J. of Comb. Information and Syst. Sci., Vol. 14, No. 4, 187-191.

[3] Belis, M. and Guiasu, S. (1968), A Qualitative-Quantitative Measure of Information in Cybernetics Systems. IEEE Transation in Information Theory, IT 14, 593-594.

[4] Daroczy, Z. (1970), Generalized Information function. Information and Control, Vol. 16, 36-51.

[5] Feinstein, A. (1956), Foundation of Information Theory. McGraw Hill, New York.

[6] Guiasu, S. and Picard, C.F. (1971), Borne Inferieture de la Longuerur de utile Certain Codes. C.R. Acad. Sci, Paris, Ser A-B Vol. 273, 248-251.

[7] Gurdial and Pessoa, F. (1977), On Useful Information of order $\alpha$. J. Comb. Information and Syst. Sci., Vol. 2, 30-35.

[8] Havrda, J.F. and Charvat, F. (1967), Qualification Method of Classification Process, the concept of structural $\alpha$-entropy. Kybernetika, Vol. 3, 30-35.

[9] Hooda, D.S. and Ram, A. (2002), On ‘useful’ Relative information and ‘useful’ J. Divergence measure. Tamkang Journals of Mathematics, Vol. 33, No. 2, 149-160.

[10] Jain, P. and Tuteja, R.K. (1989), On coding theorems connected with ‘useful’ entropy of order $\beta$. International Journal of Mathematics and Mathematical Sciences, Vol. 12, No. 1, 193-198.

[11] Kraft, L.G. (1949), A device for quantizing, grouping and coding amplitude modulated pulses. MS Thesis, Electrical Engineering Department, MIT.
[12] Longo, G. (1976), A Noiseless Coding Theorem for Sources Having Utilities. SIAM J. Appl. Math., Vol. 30, No. 4, 739-748.
[13] Mc-Millan, B. (1956), Two inequalities implied by using decipherability. IEEE Trans Inform. Theory, IT-2, 115-116.
[14] Mitter, J. and Mathur, Y.D. (1972), Comparison of entropies of power distribution. ZAMM, Vol. 52, 239-240.
[15] Singh, R.P., Kumar, R. and Tuteja, R.K. (2003), Application of Holder's Inequality in Information Theory. Information Sciences, Vol. 152, 145-154.
[16] Shannon, C.E. (1948), A Mathematical Theory of Communication. Bell System Tech. J., Vol. 27, 379-423, 623-636.
[17] Shisha, O. (1967), Inequalities. Academic Press, New York.
[18] Tsallis, C. (1988), Possible Generalization of Boltzmann Gibbs Statistics. J. Stat. Phy., Vol. 52, 479.