International Journal of Mathematical Sciences
Vol. 18, Nos. 1-2, January-June 2019, pp. 163-173

TWO NEW WEIGHTED MODELS OF CROSS ENTROPY

Ruchi Handa, Rajesh Narula & C.P.Gandhi

Abstract: The measure of cross entropy is the key concept in the literature of information theory and has found tremendous applications in different disciplines of science and technology. The various researchers have generalized this cross entropy with different approaches. The object of the present manuscript is to develop new non parametric and parametric models of weighted cross entropy after assigning weights to the probability distribution. The proposed models satisfy all the essential properties of the original measure of cross entropy.

Keywords: Entropy, Cross entropy, Convex function, Hessian matrix.

1. INTRODUCTION

In communication theory, it was Shannon [13] who first introduced the concept of entropy and it was then realized that entropy is a property of any stochastic system and the concept is now widely prevalent in different disciplines. The tendency of the system to become more disordered over time is described by the second law of thermodynamics, which states that the entropy of the system cannot spontaneously decrease. Today, information theory is still principally concerned with communications systems, but there are widespread applications in Mathematical Sciences. Shannon [13] called the probabilistic uncertainty as entropy and developed his measure given by

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \tag{1.1}$$

Hu [8] proposed two new broad classes for measures of uncertainty as the survival exponential and the generalized survival exponential entropies and thus improving Shannon's [13] entropy. Honda and Okazaki [7] generalized Shannon's [13] entropy and proved that their entropy has applicability to the capacity on set systems. Immediately, after Shannon [13] gave his measure, Renyi [12] was the first to derive entropy of order a as follows:

$$_{\alpha}H(P) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{n} p_i^{\alpha} \Big/ \sum_{i=1}^{n} p_i \right), \alpha \neq 1, \alpha > 0 \tag{1.2}$$

The generalized measure of entropy (1.2) includes Shannon's [13] entropy as a limiting case as $\alpha \to 1$. Zyczkowski [15] explored the relationships between the Shannon's [13] and Renyi's [12] entropies of integer order. Havrada and Charvat [5] introduced non-additive entropy, given by:

$$H^{\alpha}(P) = \frac{\left[\sum_{i=1}^{n} p_i^{\alpha}\right] - 1}{2^{1-\alpha} - 1}, \alpha \neq 1, \alpha > 0 \tag{1.3}$$

Many other probabilistic measures of entropy have been discussed and derived by Brissaud [1], Chen [2], Garbaczewski [3], Herremoes [6], Lavenda [10], Nanda and Paul [11], Sharma and Taneja [14] etc. The applications of the results on probabilistic information measures obtained by various authors have been provided to different fields of Linguistics, Biological Sciences, Economics, Social Sciences and Engineering Sciences for developing new entropic models.

The entropy introduced by Shannon takes into account only the probabilities associated with the events and not their importance. But there exist many fields dealing with random events where it is necessary to take into account both these probabilities and some qualitative characteristics of the events. For instance, in two-handed game one should keep in mind both the probabilities of different variants of the game, that is, the random strategies of the players and the wins corresponding to these variants. The concept provides the necessity to associate probability with weight. To explain the concept of weighted entropy, let $E_1, E_2, ..., E_n$ denote n possible outcomes with $p_1, p_2, ..., p_n$ as their probabilities and let $w_1, w_2, ..., w_n$ be non-negative real numbers representing their utilities or weights. Then, Guiasu [4] characterized the following qualitative-quantitative measure of entropy:

$$H(P, U; W) = -\sum_{i=1}^{n} w_i \, p_i \log p_i \tag{1.4}$$

and called it weighted entropy.

A measure $D(P;Q)$ of divergence or cross entropy or directed divergence is found to be very important in various disciplines of Mathematical and Engineering Sciences. This measure is probabilistic in nature and measures the distance of a probability distribution P from Q. The most important and useful measure of divergence is due to Kullback and Leibler [9] and is given by

$$D(P;Q)=\sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \tag{1.5}$$

After attaching weights to different events, the above measure (1.5) becomes

$$D(P,Q;W)=\sum_{i=1}^{n} w_i p_i \log \frac{p_i}{q_i} \tag{1.6}$$

where $w_i \geq 0$ are real numbers.

In this communication, we have introduced two new parametric and non-parametric probabilistic measure of cross entropy and studied their important properties.

2. A NEW NON PARAMETRIC WEIGHTED MODEL OF CROSS ENTROPY

Consider the following set of all complete finite discrete probability distributions

$$\Omega_n=\left\{P=\left(p_1,p_2,\ldots,p_n\right): p_i>0, \sum_{i=1}^{n} p_i=1\right\}, n\geq 2, \tag{2.1}$$

Let $P,Q\in\Omega_n$ be any two probability distributions. Then, we propose a new weighted measure of cross entropy given by

$$D\left(P,Q;W\right)=\sum_{i=1}^{n} w_i p_i \ln \frac{p_i}{\left(\frac{\sqrt{p_i}+\sqrt{q_i}}{2}\right)^2} \tag{2.2}$$

Next, to prove that the measure (2.2) is a valid measure of weighted cross entropy, we study its essential properties as follows:

(i) $D\left(P,Q;W\right)$ is convex.

Let $D\left(P,Q;W\right)=f\left(p_1,p_2,\ldots p_n,q_1,q_2,\ldots,q_n;w_1,w_2,\ldots,w_n\right)$. Then, we have

$$\frac{\partial f}{\partial p_i}=w_i\left[\frac{\sqrt{q_i}}{\sqrt{p_i}+\sqrt{q_i}}+\ln\frac{4p_i}{\left(\sqrt{p_i}+\sqrt{q_i}\right)^2}\right]$$

$$\text{Also, } \frac{\partial^2 f}{\partial p_i^2} = w_i \left[\frac{\sqrt{q_i}\left(p_i + 2\sqrt{p_i q_i}\right)}{2 p_i^{\frac{3}{2}} \left(\sqrt{p_i} + \sqrt{q_i}\right)^2} \right] > 0$$

and $\dfrac{\partial^2 f}{\partial p_i \partial p_j} = 0$ for $i, j = 1, 2, \ldots, n; i \neq j$.

The Hessian matrix of second order partial derivatives of f with respect to $p_1, p_2, \ldots, p_n$ is

$$\begin{bmatrix} \dfrac{w_1\sqrt{q_1}\left(p_1 + 2\sqrt{p_1 q_1}\right)}{2 p_1^{\frac{3}{2}} \left(\sqrt{p_1} + \sqrt{q_1}\right)^2} & 0 & \cdots & 0 \\ 0 & \dfrac{w_2\sqrt{q_2}\left(p_2 + 2\sqrt{p_2 q_2}\right)}{2 p_2^{\frac{3}{2}} \left(\sqrt{p_2} + \sqrt{q_2}\right)^2} & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & \dfrac{w_n\sqrt{q_n}\left(p_n + 2\sqrt{p_n q_n}\right)}{2 p_n^{\frac{3}{2}} \left(\sqrt{p_n} + \sqrt{q_n}\right)^2} \end{bmatrix}$$

which is positive definite matrix. Thus, $D(P,Q;W)$ is a convex function of $p_1, p_2, \ldots, p_n$.

A similar result is also true with respect to $q_1, q_2, \ldots, q_n$

that is, $\dfrac{\partial f}{\partial q_i} = -\dfrac{w_i p_i}{\sqrt{p_i q_i} + q_i}$

Also, $\dfrac{\partial^2 f}{\partial q_i^2} = \dfrac{w_i p_i\left(p_i + 2\sqrt{p_i q_i}\right)}{2\sqrt{p_i q_i}\left(\sqrt{p_i q_i} + q_i\right)^2}$

and $\dfrac{\partial^2 f}{\partial p_i \partial p_j} = 0$ for $i, j = 1, 2, \ldots, n; i \neq j$.

The Hessian matrix of second order partial derivatives of f with respect to $q_1, q_2, ..., q_n$ is

$$\begin{bmatrix} \dfrac{w_1 p_1\left(p_1 + 2\sqrt{p_1 q_1}\right)}{2\sqrt{p_1 q_1}\left(\sqrt{p_1 q_1} + \sqrt{q_i}\right)^2} & 0 & \cdots & 0 \\ 0 & \dfrac{w_2 p_2\left(p_2 + 2\sqrt{p_2 q_2}\right)}{2\sqrt{p_2 q_2}\left(\sqrt{p_2 q_2} + \sqrt{q_2}\right)^2} & \cdots & 0 \\ . & & & \\ . & & & \\ . & & & \\ 0 & 0 & \cdots & \dfrac{w_n p_n\left(p_n + 2\sqrt{p_n q_n}\right)}{2\sqrt{p_n q_n}\left(\sqrt{p_n q_n} + \sqrt{q_n}\right)^2} \end{bmatrix}$$

which is positive definite matrix. Thus, $D(P,Q;W)$ is a convex function of both $p_1, p_2, ..., p_n$ and $q_1, q_2, \ldots, q_n$. Thus, we have

(i) $D(P,Q;W)$ is convex function of both $p_1, p_2, ..., p_n$ and $q_1, q_2, ..., q_n$.

(ii) $D(P,Q;W) = 0$ iff $P = Q$

(iii) $D(P,Q;W)\geq 0$ as its minimum value is zero and it is convex.

Under the above properties, the mathematical model introduced in (2.2) proves its validity to be a correct measure of weighted cross entropy.

Moreover, with the help of numerical data shown in Table-2.1, we have presented $D(P,Q;W)$ as shown in the following Fig.-2.1.

Table-2.1

p_i	q_i	w_i	$D(P,Q;W)$
0.1	0.5	0.10	0.021481
0.2	0.5	0.15	0.016219
0.3	0.5	0.20	0.009036
0.4	0.5	0.25	0.002733
0.5	0.5	0.30	0
0.6	0.5	0.35	0.003826
0.7	0.5	0.40	0.018071
0.8	0.5	0.45	0.048657
0.9	0.5	0.50	0.107407

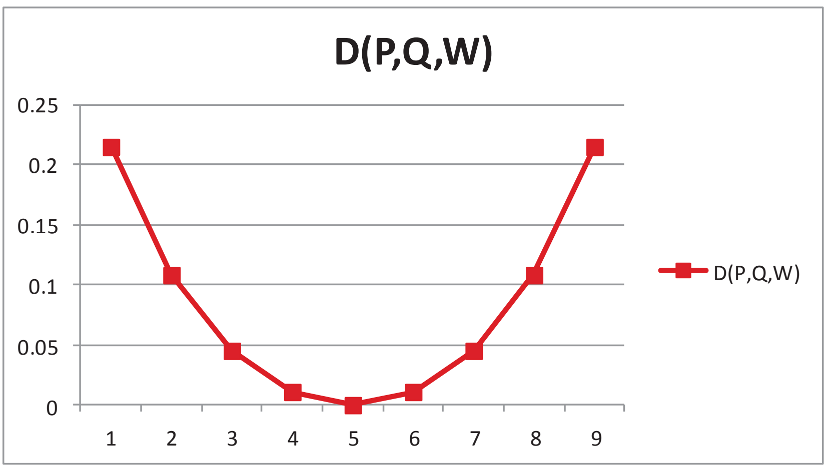

Fig. 2.1

3. A NEW GENERALIZED WEIGHTED PARAMETRIC MODEL OF CROSS ENTROPY

Now, we propose a new generalized parametric measure of weighted cross entropy given by the following mathematical model:

$$D_\alpha\left(P,Q;W\right)=\frac{\ln\left(\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{(\alpha-1)w_i p_i}\right)}{(\alpha-1)},\ \alpha>0,\ \alpha\neq 1 \tag{3.1}$$

We have, $\underset{\alpha\to 1}{lt}\left[D_\alpha\left(P,Q;W\right)\right]=\underset{\alpha\to 1}{lt}\left(\frac{\ln\left(\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{(\alpha-1)w_i p_i}\right)}{\alpha-1}\right)$

$$=\sum_{i=1}^{n} w_i p_i \ln\frac{p_i}{q_i}$$

Thus $D_\alpha\left(P,Q;W\right)$ is a generalization of Kullback-Leibler's [9] measure of cross entropy after attaching weights to the probability distribution.

Properties of Cross entropy measure:

(i) $D_\alpha\left(P,Q;W\right)\geq 0$

that is, $D_\alpha\left(P,Q;W\right)=\frac{\ln\left(\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{(\alpha-1)w_i p_i}\right)}{(\alpha-1)}=\frac{\sum_{i=1}^{n}\ln\left(\left(\frac{p_i}{q_i}\right)^{(\alpha-1)w_i p_i}\right)}{(\alpha-1)}$

$$=\sum_{i=1}^{n} w_i p_i \ln\frac{p_i}{q_i}\geq 0.$$

(ii) $D_\alpha\left(P,Q;W\right)=0$ iff $P=Q$

(iii) To prove $D_\alpha\left(P,Q;W\right)$ is convex.

We have, $\dfrac{\partial\left(D_\alpha\left(P,Q;W\right)\right)}{\partial p_i}=w_i\left[1+\ln\dfrac{p_i}{q_i}\right]$

Also, $\dfrac{\partial^2\left(D_\alpha\left(P,Q;W\right)\right)}{\partial p_i^2}=\dfrac{w_i}{p_i}>0$

and $\dfrac{\partial^2 D_\alpha\left(P,Q;W\right)}{\partial p_i\partial p_j}=0$ for $i,j=1,2,\ldots,n;i\neq j$.

Therefore, Hessian matrix is given by

$$\begin{bmatrix} \frac{w_1}{p_1} & 0 & \ldots & 0 \\ 0 & \frac{w_2}{p_2} & \ldots & 0 \\ . & & & \\ . & & & \\ . & & & \\ 0 & 0 & \ldots & \frac{w_n}{p_n} \end{bmatrix}$$

which is positive definite matrix. So, it is convex function of $p_1, p_2,..., p_n$

Similarly, $\dfrac{\partial\left(D_\alpha\left(P,Q;W\right)\right)}{\partial q_i}=-\dfrac{w_i p_i}{q_i}$

Also, $\dfrac{\partial^2\left(D_\alpha\left(P,Q;W\right)\right)}{\partial q_i^2}=\dfrac{w_i p_i}{q_i^2}>0$

and $\dfrac{\partial\left(D_\alpha\left(P,Q;W\right)\right)}{\partial q_i\partial q_j}=0$ for $i,j=1,2,\ldots,n;i\neq j$.

Therefore, Hessian matrix is given by

$$\begin{bmatrix} \frac{w_1 p_1}{q_1^2} & 0 & \cdots & 0 \\ 0 & \frac{w_2 p_2}{q_2^2} & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & \frac{w_n p_n}{q_n^2} \end{bmatrix}$$

which is also positive definite matrix. So, it is convex function of $q_1, q_2, ..., q_n$.

Under the above properties, the mathematical model introduced in (3.1) proves its validity to be a correct parametric measure of weighted cross entropy.

Moreover, with the help of numerical data shown in Table-2.2, we have presented $D_\alpha(P,Q;W)$ as shown in the Fig.-2.2.

Table-2.2

p_i	q_i	w_i	$D_2(P,Q;W)$	$D_{2.2}(P,Q;W)$	$D_{2.3}(P,Q;W)$	$D_{2.4}(P,Q;W)$	$D_{2.5}(P,Q;W)$
0.1	0.5	1	0.159848254	0.368064207	0.36806421	0.368064	0.368064207
0.2	0.5	1.025	0.085800684	0.197563376	0.19756338	0.197563	0.197563376
0.3	0.5	1.075	0.038415125	0.088454094	0.08845409	0.088454	0.088454094
0.4	0.5	1.1	0.009619217	0.022149065	0.02214906	0.022149	0.022149065
0.5	0.5	1.125	0	0	0	0	0
0.6	0.5	1.15	0.010056454	0.023155841	0.02315584	0.023156	0.023155841
0.7	0.5	1.175	0.041988625	0.096682382	0.09668238	0.096682	0.096682382
0.8	0.5	1.2	0.100449581	0.231293708	0.23129371	0.231294	0.231293708
0.9	0.5	1.225	0.195814111	0.450878654	0.45087865	0.450879	0.450878654

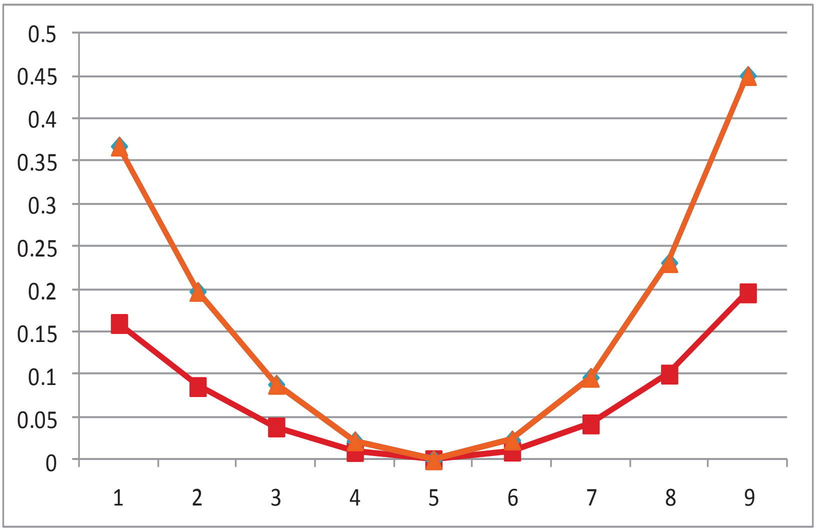

Fig. 2.2

REFERENCES

1. Brissaud, J. B. (2005). The meaning of entropy. *Entropy* **7**: 68-96.

2. Chen, Y. (2006). Properties of quasi-entropy and their applications. *Journal of Southeast University Natural Sciences* **36**: 222-225.

3. Garbaczewski, P. (2006). Differential entropy and dynamics of uncertainty. *Journal of Statistical Physics* **123**: 315-355.

4. Guiasu, S. (1971). Weighted entropy. *Reports on Mathematical Physics* **2**: 165-171.

5. Havrada, J. H. and Charvat, F. (1967). Quantification methods of classification process: Concept of structural α-entropy. *Kybernetika* **3**: 30-35.

6. Herremoes, P. (2006). Interpretations of Renyi entropies and divergences. *Journal of Physics A* **365**: 57-62.

7. Honda, A. and Okazaki, Y. (2007). Generalization of Shannon entropy of capacities on set systems. *Applied Functional Analysis*: 89-98.

8. Hu, H. (2007). Survival exponential entropies and their analytic properties. *Journal of Huazhong Normal University, Natural Sciences* **41**: 493-496.

9. Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics* **22**: 79-86.

10. Lavenda, B. H. (2005). Mean Entropies. *Open Systems & Information Dyn*amics **12**: 289-302.

11. Nanda, A. K. and Paul, P. (2006). Some results on generalized residual entropy. *Information Sciences* **176**: 27-47.

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

13. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* **27**: 379-423, 623-659.

14. Sharma, B. D. and Taneja, I. J. (1975). Entropies of type (α,β) and other generalized measures of information theory. *Metrica* **22**: 202-215.

15. Zyczkowski, K. (2003). Renyi extrapolation of Shannon entropy. *Open Systems and Information Dynamics* **10**: 297-310.

Ruchi Handa
Department of Applied Sciences,
Khalsa College of Engg.& Tech.,
Amritsar, India
ruchihanda175@gmail.com

Rajesh Narula
Department of Applied Sciences,
I.K.Gujral Punjab Tech. University,
Jalandhar, Kapurthala, India
dr.rknarula@gmail.com

C.P.Gandhi
Department of Mathematics,
University School of Sciences,
Rayat Bahra University,
Kharar, Mohali, India.
cchanderr@gmail.com