International Journal of Mathematical Sciences
Vol. 18, Nos. 1-2, January-June 2019, pp. 175-183

A NEW PROBABILISTIC ENTROPIC MODELING FOR WEIGHTED DISTRIBUTIONS

Ruchi Handa, Rajesh Narula & C.P.Gandhi

Abstract: The measures of weighted information play a significant role in the literature of information theory because of the importance of occurrence of events. In the present communication, we have introduced new generalized model representing weighted information theoretic measure based upon discrete probability distribution and studied its many desirable properties in detail.

Keywords: Uncertainty, Entropy, Probability distribution, Non-negativity, Increasing function, Symmetric function.

1. INTRODUCTION

After the introduction of the concept of entropy by Shannon [8], it was realized that entropy is a property of any stochastic system and the concept is now used widely in many fields. The tendency of the systems 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 communication systems, but there are widespread applications in statistics, information processing and computing.

This entropy measures the amount of uncertainty contained in a probabilistic experiment and is not a single monolithic concept. It can appear in several guises. It can arise in what we normally consider a probabilistic phenomenon. On the other hand, it can also appear in a deterministic phenomenon where we know that the outcome is not a chance event, but we are fuzzy about the possibility of the specific outcome. This type of uncertainty arising out of fuzziness is the subject of investigation of the relatively new discipline of fuzzy set theory. We shall first take up the case of probabilistic uncertainty associated with the probability of outcomes of the random experiments. This uncertainty is called entropy, since this is the terminology that is well entrenched in the literature. Shannon [8] introduced the concept of information theoretic entropy by associating uncertainty with every probability distribution $P = (p_1, p_2, ..., p_n)$ and found that there is a unique function that can measure the uncertainty, is given by

$$H(P) = -\sum_{i=1}^{n} p_i \text{ In } p_i \qquad (1.1)$$

The probabilistic measure of entropy (1.1) possesses a number of interesting properties.

Recently, Yu [9] has presented some entropy comparison results concerning compound distributions on non-negative integers. The main result shows that, under a log-concavity assumption, two compound distributions are ordered in terms of Shannon’s [8] entropy if both the numbers of claims' and the `claim sizes' are ordered accordingly in the convex order. Several maximum/minimum entropy theorems follow as a consequence. Most importantly, two recent results on maximum entropy characterizations of compound poisson and binomial distributions are proved under fewer assumptions and with simpler arguments. Immediately, after Shannon, research workers in many fields saw the potential of application of this entropy and a large number of other information theoretic measures were derived. It was Renyi [7] who for the first time introduced entropy of order α given by the following mathematical expression:

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

Golshani, Pasha and Yari [3] established the definition of the conditional Renyi entropy and showed that the so-called chain rule holds:

$$H_{\alpha}(X_1,...,X_n)=\sum_{i=1}^{n} H_{\alpha}(X_i / X_i,...,X_{i-1}).$$

Here the X_i components of a random variable and $\alpha > 0(\alpha \neq 1)$ is a parameter of the Renyi entropy. The basic definition is based on the earlier work of Renyi which showed that $f^{-1}\left(\sum_i p_i f(-\log p_i)\right)$ leads to the Shannon [6] entropy for $f(x)=x$ and the Renyi [7] entropy for $f(x)=2^{(1-\alpha)x}$ and these two are the only two possible for additive properties. Havrada and Charvat [4] introduced first non-additive entropy, given by

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

Kapur [5] introduced a generalized measure of entropy of order ‘α’ and type ‘β’, given by

$$H_\beta^\alpha(P) = \frac{1}{\alpha+\beta-2}\left[\sum_{n=0}^{\infty} p_n{}^\alpha + \sum_{n=0}^{\infty} p_n{}^\beta - 2\right], \alpha \geq 1, \beta \leq 1 \, or \, \alpha \leq 1, \beta \geq 1 \quad (1.4)$$

Some other pioneer who made the study of entropy theory rigorous include Dehmer and Mowshowitz [2], Lacevic and Amaldi [6], Kapur [5] etc.

In section 2, we have introduced a new generalized information theoretic measure based upon discrete probability distribution $P = \left\{\left(p_1, p_2, ..., p_n\right), p_i \geq 0, \sum_{i=1}^{n} p_i = 1\right\}$ with the help of weighted information introduced by Belis and Guiasu [1].

2. A NEW MEASURE OF WEIGHTED ENTROPY BASED UPON DISCRETE PROBABILITY DISTRIBUTION

In this section, we introduce a new measure of weighted entropy given by the following expression:

$$K^\beta(P;W) = -\frac{\sum_{i=1}^{n} w_i p_i^\beta \log\left(\frac{p_i^\beta}{\sum_{i=1}^{n} p_i^\beta}\right)}{\sum_{i=1}^{n} p_i^\beta} + \frac{1}{1-\beta}\left[\sum_{i=1}^{n} w_i p_i^\beta - \sum_{i=1}^{n} w_i p_i\right] - \sum_{i=1}^{n} w_i p_i \log p_i$$

$$; \beta \neq 1, \ \beta > 0 \quad (2.1)$$

$$= H(p_1, p_2, ..., p_n; w_1, w_2, ..., w_n)$$

The measure (2.1) satisfies the following properties:

(i) It is continuous function of $(p_1, w_1); (p_2, w_2); ...; (p_n, w_n)$.

(ii) It is permutationally symmetric function of $(p_1, w_1); (p_2, w_2); ...; (p_n, w_n)$.

(iii) $H(p_1, p_2, ..., p_n; w_1, w_2, ..., w_n) \geq 0$

(iv) $$H^{n+1}(p_1, p_2, ..., p_n, 0; w_1, w_2, ..., w_n, w_{n+1}) = -\frac{\sum_{i=1}^{n+1} w_i p_i^{\beta} \log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n+1} p_i^{\beta}}\right)}{\sum_{i=1}^{n+1} p_i^{\beta}}$$

$$+\frac{1}{1-\beta}\left[\sum_{i=1}^{n+1} w_i p_i^{\beta} - \sum_{i=1}^{n+1} w_i p_i\right] - \sum_{i=1}^{n+1} w_i p_i \log p_i$$

$$= -\frac{\sum_{i=1}^{n} w_i p_i^{\beta} \log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}}\right)}{\sum_{i=1}^{n} p_i^{\beta}} + \frac{1}{1-\beta}\left[\sum_{i=1}^{n} w_i p_i^{\beta} - \sum_{i=1}^{n} w_i p_i\right] - \sum_{i=1}^{n} w_i p_i \log p_i$$

$$= H(p_1, p_2, ..., p_n; w_1, w_2, ..., w_n)$$

This property says that entropy does not change by the inclusion of an impossible event with probability zero.

(v) Since $H(p_1, p_2, ..., p_n; w_1, w_2, ..., w_n)$ is an entropy measure, its maximum value must occur. To find the maximum value, we proceed as follows:

Let $$f(p;w) = -\frac{\sum_{i=1}^{n} w_i p_i^{\beta} \log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}}\right)}{\sum_{i=1}^{n} p_i^{\beta}} + \frac{1}{1-\beta}\left[\sum_{i=1}^{n} w_i p_i^{\beta} - \sum_{i=1}^{n} w_i p_i\right] - \sum_{i=1}^{n} w_i p_i \log p_i - \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$$

Then, we have

$$\frac{\partial f}{\partial p_1} = -\frac{\left(\sum_{i=1}^{n} p_i^{\beta}\right)\frac{\partial}{\partial p_1}\left(\sum_{i=1}^{n} w_i p_i^{\beta}\log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}}\right)\right) - \left(\sum_{i=1}^{n} w_i p_i^{\beta}\log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}}\right)\right)\frac{\partial}{\partial p_1}\left(\sum_{i=1}^{n} p_i^{\beta}\right)}{\left(\sum_{i=1}^{n} p_i^{\beta}\right)^2}$$

$$+\frac{w_1}{1-\beta}\left\{\beta p_1^{\beta-1}-1\right\} - w_1\left(1+\log p_1\right)-\lambda$$

$$= -\frac{\left(\sum_{i=1}^{n} p_i^{\beta}\right)\frac{\partial}{\partial p_1}\left(\sum_{i=1}^{n} w_i p_i^{\beta}\beta\log p_i - \sum_{i=1}^{n} w_i p_i^{\beta}\log\left(\sum_{i=1}^{n} p_i^{\beta}\right)\right) - \left(\sum_{i=1}^{n} w_i p_i^{\beta}\log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}}\right)\right)\left(\beta p_1^{\beta-1}\right)}{\left(\sum_{i=1}^{n} p_i^{\beta}\right)^2}$$

$$+\frac{w_1}{1-\beta}\left\{\beta p_1^{\beta-1}-1\right\} - w_1\left(1+\log p_1\right)-\lambda$$

$$= -\frac{\left(\sum_{i=1}^{n} p_i^{\beta}\right)\frac{\partial}{\partial p_1}\left(\sum_{i=1}^{n} w_i p_i^{\beta}\beta\log p_i\right) - \left(\sum_{i=1}^{n} p_i^{\beta}\right)\frac{\partial}{\partial p_1}\left(\sum_{i=1}^{n} w_i p_i^{\beta}\log\left(\sum_{i=1}^{n} p_i^{\beta}\right)\right) - \left(\sum_{i=1}^{n} w_i p_i^{\beta}\log\left(\frac{p_i^{\beta}}{\sum_{i=1}^{n} p_i^{\beta}}\right)\right)\left(\beta p_1^{\beta-1}\right)}{\left(\sum_{i=1}^{n} p_i^{\beta}\right)^2}$$

$$+\frac{w_1}{1-\beta}\left\{\beta p_1^{\beta-1}-1\right\}-w_1\left(1+\log p_1\right)-\lambda$$

Thus, we have

$$\frac{\partial f}{\partial p_1}=-\frac{\left(\sum_{i=1}^{n}p_i^\beta\right)w_1p_1^{\beta-1}\beta(1+\beta\log p_1)-\left(\sum_{i=1}^{n}w_ip_i^\beta\log\left(\frac{p_i^\beta}{\sum_{i=1}^{n}p_i^\beta}\right)\right)\left(\beta p_1^{\beta-1}\right)-\left(\sum_{i=1}^{n}p_i^\beta\right)\left[w_1\beta\,p_1^{\beta-1}\log\left(\sum_{i=1}^{n}p_i^\beta\right)+\left(\sum_{i=1}^{n}w_ip_i^\beta\right)\frac{\beta\,p_1^{\beta-1}}{\sum_{i=1}^{n}p_i^\beta}\right]}{\left(\sum_{i=1}^{n}p_i^\beta\right)^2}$$

$$+\frac{w_1}{1-\beta}\left\{\beta p_1^{\beta-1}-1\right\}-w_1\left(1+\log p_1\right)-\lambda$$

Similarly, we have

$$\frac{\partial f}{\partial p_2}=-\frac{\left(\sum_{i=1}^{n}p_i^\beta\right)w_2p_2^{\beta-1}\beta(1+\beta\log p_2)-\left(\sum_{i=1}^{n}w_ip_i^\beta\log\left(\frac{p_i^\beta}{\sum_{i=1}^{n}p_i^\beta}\right)\right)\left(\beta p_2^{\beta-1}\right)-\left(\sum_{i=1}^{n}p_i^\beta\right)\left[w_2\beta\,p_2^{\beta-1}\log\left(\sum_{i=1}^{n}p_i^\beta\right)+\left(\sum_{i=1}^{n}w_ip_i^\beta\right)\frac{\beta\,p_2^{\beta-1}}{\sum_{i=1}^{n}p_i^\beta}\right]}{\left(\sum_{i=1}^{n}p_i^\beta\right)^2}$$

$$+\frac{w_2}{1-\beta}\left\{\beta p_2^{\beta-1}-1\right\}-w_2\left(1+\log p_2\right)-\lambda$$

and so on

$$\frac{\partial f}{\partial p_n} = -\frac{\left(\sum_{i=1}^{n} p_i^\beta\right) w_n p_n^{\beta-1} \beta (1+\beta \log p_n) - \left(\sum_{i=1}^{n} w_i p_i^\beta \log\left(\frac{p_i^\beta}{\sum_{i=1}^{n} p_i^\beta}\right)\right)\left(\beta p_n^{\beta-1}\right) - \left(\sum_{i=1}^{n} p_i^\beta\right)\left[w_n \beta\, p_n^{\beta-1} \log\left(\sum_{i=1}^{n} p_i^\beta\right) + \left(\sum_{i=1}^{n} w_i p_i^\beta\right)\frac{\beta\, p_n^{\beta-1}}{\sum_{i=1}^{n} p_i^\beta}\right]}{\left(\sum_{i=1}^{n} p_i^\beta\right)^2}$$

$$+\frac{w_n}{1-\beta}\left\{\beta p_n^{\beta-1} - 1\right\} - w_n\left(1+\log p_n\right) - \lambda$$

For maximum value, we put

$$\frac{\partial f}{\partial p_1} = \frac{\partial f}{\partial p_2} = \ldots = \frac{\partial f}{\partial p_n} = 0$$

which gives that each p_i is a function of w_i. In particular, when weights are ignored, then $p_1 = p_2 = \ldots = p_n$ and applying the condition that $\sum_{i=1}^{n} p_i = 1$, we get $p_i = \frac{1}{n}\ \forall\ i.$

(vi) The maximum value is an increasing function of n. Let $f(n)$ be maximum value. Then, we have

$$f(n) = \frac{n^{-\beta} \log\left(\frac{n^{-\beta}}{n^{1-\beta}}\right)}{n^{1-\beta}} + \frac{1}{1-\beta}\left[n^{1-\beta} - 1\right] + \left\{\log n - 1\right\}$$

$$= -\frac{1}{n}\log n + \frac{1}{1-\beta}\left[n^{1-\beta} - 1\right] + \left\{\log n - 1\right\}$$

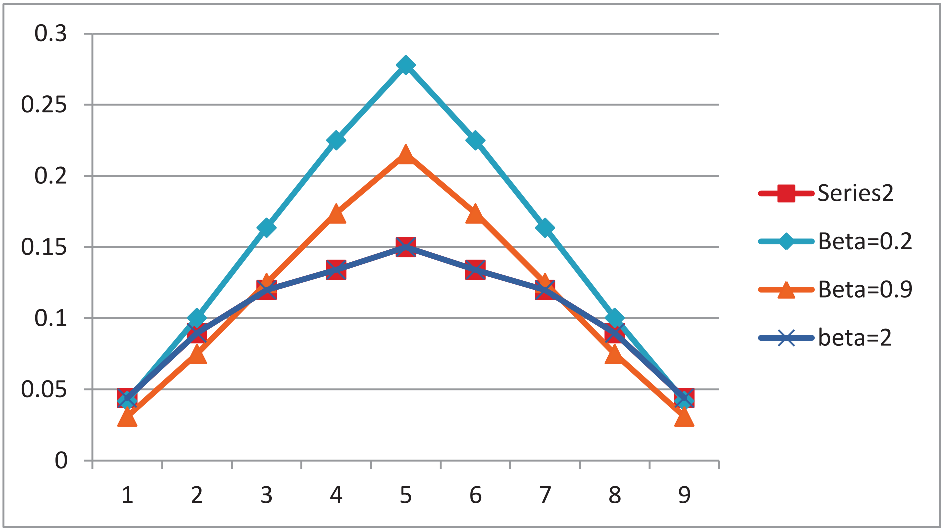
Thus

$$f^{'}(n) = -\left(\frac{1}{n^2} - \frac{1}{n^2}\log n\right) + \frac{1}{n^\beta} + \frac{1}{n}$$

$$= \frac{1}{n^2}(\log n - 1) + \frac{1}{n^\beta} + \frac{1}{n} \quad > 0$$

Hence, maximum value is an increasing function of n and this property is most desirable. Under the above properties, we observe that the proposed measure (2.1) is a valid measure of weighted entropy.

χ	p_i	$0\cdot1$	$0\cdot2$	$0\cdot3$	$0\cdot4$	$0\cdot5$	$0\cdot6$	$0\cdot7$	$0\cdot8$	$0\cdot9$
$0\cdot1$	$K^\chi(P,W)$	0.043922	0.089503	0.119857	0.133942	0.15	0.133942	0.119857	0.089503	0.043922\
$0\cdot2$	$K^\chi(P,W)$	0.04181365	0.10022819	0.16354702	0.22499447	0.27791292	0.22499447	0.16354702	0.10022819	0.04181365
$0\cdot9$	$K^\chi(P,W)$	0.030926	0.074903	0.124691	0.173632	0.21532	0.173632	0.124691	0.074903	0.030926
2	$K^\chi(P,W)$	0.043922	0.089503	0.119857	0.133942	0.15	0.133942	0.119857	0.089503	0.043922



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. Dehmer, M. and Mowshowitz, A. (2011). A history of graph entropy measures. *Inform. Sci.* 181(1): 57-78.

3. Golshani, L., Pasha, E. and Yari, G. (2009). Some properties of Renyi entropy and Renyi entropy rate. *Information Sciences* 179(14): 2426-2433.

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

5. Kapur, J.N. (1995). *Measures of Information and Their Applications*. Wiley Eastern, New York.

6. Lacevic, B. and Amaldi, E. (2011). Entropy of diversity measures for populations in Euclidean space. *Inform. Sci.* **181**(11): 2316-2339.

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

8. Shannon, C.E. (1948). A mathematical theory of communication. *Bell. Sys. Tech. Jr.* **27**: 379-423, 623-659.

9. Yu, Y. (2009). On the entropy of compound distributions on nonnegative integers, *IEEE Trans. Inform. Theory* **55** (8): 3645–3650.

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