Two New Parametric Entropic Models for Discrete Probability Distributions

Om Parkash[a] and Mukesh[b*]

[a]Department of Mathematics, Guru Nanak Dev University, Amritsar (India), omparkash777@yahoo.co.in
[b]Department of Mathematics, Maharaja Ranjit Singh Punjab Technical University, Bathinda (India), mukesh.sarangal@mrsptu.ac.in

_____

Abstract: Information theory fundamentally deals with two types of theoretical models frequently well acknowledged as entropy and divergence. In the literature of entropy models for the discrete probability distributions, there survive numerous standard models but still there is possibility that several innovative models can be constructed so as to provide their applications in a variety of disciplines of mathematical sciences. The present communication is a right step in this direction and participates with the derivation of two new parametric models of entropy. Furthermore, it provides the meticulous study of the most advantageous properties of the projected discrete models to prove their legitimacy.

_____

1. Introduction

The peculiarity that Shannon's (1948) quantitative entropic model for probability spaces has incredibly pleasant properties and provides tremendous applications in various disciplines of mathematical sciences led the researchers to find its drawback that it takes into account only the probabilities associated with the events and not their significance or importance. However, in many realistic situations it becomes inevitable to acquire both the equally important aspects, that is, quantity as well as quality. Inspired by this innovative thought, Guiasu (1971) customized the concept and explained its detailed features. The quantitative expression of the mathematical model measuring uncertainty contained in probabilistic experiment due to Shannon (1948), also well recognized as information entropy is given by

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

Keeping in mind the negative nature of information, it was Burg (1972) who for the first time investigated and introduced the following entropic model:

$$B(P) = \sum_{i=1}^{n} \log p_i \tag{1.2}$$

Kapur (1994) introduced his parametric entropy:

$$H_\alpha(P) = \frac{\sum_{i=1}^{n} p_i^{\alpha} - n^{1-\alpha}}{\alpha(1-\alpha)}; \quad \alpha \geq 0 \tag{1.3}$$

Now

$$\underset{\alpha \to 1}{Lt}\, H_\alpha(P) = H(P) - H(U) \tag{1.4}$$

and

$$\underset{\alpha \to 0}{Lt}\, H_\alpha(P) = B(P) - B(U) \tag{1.5}$$

where $U = \left(\dfrac{1}{n}, \dfrac{1}{n}, \mathrm{K}, \dfrac{1}{n} \right)$ is the uniform distribution.

Parkash and Kakkar (2014a, 2014b) made further study of information theoretic models and enriched the literature of entropy measures by proposing the subsequent parametric entropy models:

$$_\alpha A(P) = \frac{\displaystyle\sum_{i=1}^{n} p_i \alpha^{\log_D p_i} - 1}{1 - \alpha}, \ \alpha > 1 \tag{1.6}$$

$$N_\alpha(P) = \frac{\alpha^{\sum_{i=1}^{n} p_i \ln p_i} - 1}{1 - \alpha}, \ \alpha > 1 \tag{1.7}$$

$$J_\alpha^\beta(p) = \frac{\alpha^{\frac{1}{\beta-1}\ln\left(\sum_{i=1}^{n} p_i^\beta\right)} - 1}{1 - \alpha}, \ \alpha > 1, \ 0 < \beta < 1 \tag{1.8}$$

Parkash and Mukesh (2012, 2014) made furtherance of the above research on entropy models and proposed the subsequent parametric entropy models:

$$H_\alpha(P) = \frac{1}{\alpha} \sum_{i=1}^{n} \left(1 - p_i^{\alpha p_i} \right), \ \alpha \neq 0, \alpha > 0 \tag{1.9}$$

$$H^\alpha(P) = \frac{1}{2(1-\alpha)} \sum_{i=1}^{n} \left[\alpha \left(p_i^{1-\alpha} - 1 \right) + \log p_i^{1-\alpha} \right], \ \alpha \neq 1, \ \alpha > 0 \tag{1.10}$$

Many other authors have introduced a diversity of information theoretic models and provided their applications towards neural population coding, generalized rough set models, statistical estimation, correlations in chaotic dynamics, geometric analysis of time series etc. These researchers include Huang and Zhang (2019), Wang, Yue. and Deng (2019), Bulinski and Kozhevin (2019), Cincotta and Giordano (2018), Majumdar and Jayachandran (2018). Some other contributors conscientious for the furtherance of research in the area of entropy models include Sholehkerdar, Tavakoli and Liu (2020), Lenormand (2020), Du, Chen, Guan and Sun (2021), Bajic, Đaji´c and Milovanovi´c (2021), Kumar et.al. (2021) etc.

Further, it was Guiasu (1971) who twisted the subsequent entropic model well acknowledged as

weighted entropy:

$$H(P:W) = -\sum_{i=1}^{n} w_i p_i \log p_i \tag{1.11}$$

Moreover, Kapur (1994) convinced with certain crucial properties of the model (1.11) and wrought a family of proper measures of weighted entropy. Some of these measures are:

$$H_1(P:W) = \sum_{i=1}^{n} w_i \ln p_i \tag{1.12}$$

$$H_2(P:W) = -\sum_{i=1}^{n} w_i\, p_i \ln p_i - \sum_{i=1}^{n} w_i\, (1-p_i) \ln (1-p_i) \tag{1.13}$$

$$H_3(P:W) = -\sum_{i=1}^{n} w_i\, p_i - \sum_{i=1}^{n} w_i\, p_i \ln p_i \tag{1.14}$$

Parkash, Kumar, Mukesh and Kakkar (2019) made further study on weighted measures and consequently introduced the subsequent entropic models for the weighted distributions:

$$H_\alpha^1(P;W) = \frac{1}{2^{1-\alpha}-1}\left[\sum_{i=1}^n w_i p_i^\alpha - \left\{\sum_{i=1}^n w_i p_i^{\frac{1}{\alpha}}\right\}^\alpha\right], \; \alpha > 1 \tag{1.15}$$

$$H_{\alpha,\beta}(P;W) = \frac{1}{\beta-\alpha}\log\left[\frac{\sum_{i=1}^n w_i p_i^\alpha}{\sum_{i=1}^n w_i p_i^\beta}\right], \; \alpha < 1, \beta > 1 \text{ or } \alpha > 1, \beta < 1 \tag{1.16}$$

Some other characterizations of (1.11) have been provided by Longo (1972), Gurdial and Pessoa (1977), Aggarwal and Picard (1975) etc. In the sequel, we outline two entropic models-one each for discrete probability distributions and weighted probability distributions.

2. New Discrete Entropic Models in Probability Spaces

I. We firstly propose the following entropic model:

$$H_{\alpha,\beta}(P) = \frac{\sum_{i=1}^n p_i^\alpha - \sum_{i=1}^n p_i^\beta - n^{1-\alpha} + n^{1-\beta}}{(\beta-\alpha)(\beta+\alpha-1)}; \alpha \geq 1, \beta \leq 1 \text{ or } \alpha \leq 1, \beta \geq 1 \tag{2.1}$$

Obviously, for $\beta = 1$, $H_{\alpha,\beta}(P) = H_{\alpha,}(P)$

and for $\alpha = 1$, $H_{\alpha,\beta}(P) = H_{\beta,}(P)$

Again $\underset{\alpha \to 1}{Lt} H_{\alpha,\beta}(P) = \dfrac{-\sum_{i=1}^n p_i^\beta + n^{1-\beta}}{\beta(\beta-1)}$

Thus $\underset{\substack{\alpha \to 1 \\ \beta \to 1}}{Lt} H_{\alpha,\beta}(P) = -\sum_{i=1}^n p_i \log p_i - \log n = H(P) - H(U)$

Hence, $H_{\alpha,\beta}(P)$ introduced in (2.1) is generalized entropy.

Properties:

(1) $H_{\alpha,\beta}(P)$ is continuous.

(2) $H_{\alpha,\beta}(P)$ is symmetric.

(3) $H_{\alpha,\beta}(P)$ is expansible.

(4) **Concavity:** We have

$$\frac{\partial}{\partial p_i} H_{\alpha,\beta}(P) = \frac{\alpha p_i^{\alpha-1} - \beta p_i^{\beta-1}}{(\beta-\alpha)(\beta+\alpha-1)}$$

and

$$\frac{\partial^2}{\partial p_i^2} H_{\alpha,\beta}(P) = \frac{\alpha(\alpha-1)p_i^{\alpha-2} - \beta(\beta-1)p_i^{\beta-2}}{(\beta-\alpha)(\beta+\alpha-1)} < 0 \quad \forall \alpha, \beta \text{ which proves that } H_{\alpha,\beta}(P) \text{ is concave.}$$

(5) For degenerate distributions $(1,0,0K,0),(0,1,0K,0),K,(0,0,0K,1)$, we have

$$H_{\alpha,\beta}(P) = \frac{n^{1-\beta} - n^{1-\alpha}}{(\beta-\alpha)(\beta+\alpha-1)}$$

(6) **Entropy Maximization**

For entropy maximization, we differentiate the following Lagrangian

$$L = \frac{\sum_{i=1}^n p_i^\alpha - \sum_{i=1}^n p_i^\beta - n^{1-\alpha} + n^{1-\beta}}{(\beta-\alpha)(\beta+\alpha-1)} - \lambda\left(\sum_{i=1}^n p_i - 1\right)$$

and putting the derivatives equal to zero, we get

$$\frac{\alpha p_1^{\alpha-1} - \beta p_1^{\beta-1}}{(\beta-\alpha)(\beta+\alpha-1)} = \frac{\alpha p_2^{\alpha-1} - \beta p_2^{\beta-1}}{(\beta-\alpha)(\beta+\alpha-1)} = L = \frac{\alpha p_n^{\alpha-1} - \beta p_n^{\beta-1}}{(\beta-\alpha)(\beta+\alpha-1)}$$

which is possible only if $p_1 = p_2 = \Lambda = p_n$

Using the property $\sum_{i=1}^{n} p_i = 1$,

we get $p_i = \dfrac{1}{n} \ \forall i$.

(7) **Maximum Value**

$$\left[H_{\alpha,\beta}(P) \right]_{\max} = \frac{\sum_{i=1}^{n}\left(\dfrac{1}{n}\right)^{\alpha} - \sum_{i=1}^{n}\left(\dfrac{1}{n}\right)^{\beta} - n^{1-\alpha} + n^{1-\beta}}{(\beta - \alpha)(\beta + \alpha - 1)} = 0$$

With the above mentioned properties, we claim that the entropic model introduced in (2.1) is an appropriate model of information entropy. Next, for $\alpha = 2$ and $\beta = 0.5$, we have calculated various values for the entropy measure $H_{\alpha,\beta}(P)$ as shown in the following Table 1 and wrought Figure 1 screening the concavity of $H_{\alpha,\beta}(P)$.

Table 1. $H_{\alpha,\beta}(P)$ **against** p_i **for** $\alpha = 2$, $\beta = 0.5$ **and** $n = 2$

p_i	$H_{\alpha,\beta}(P)$
0.1	-0.20858
0.2	-0.11225
0.3	-0.04881
0.4	-0.01207
0.5	0.00000
0.6	-0.01207
0.7	-0.04881
0.8	-0.11225
0.9	-0.20858
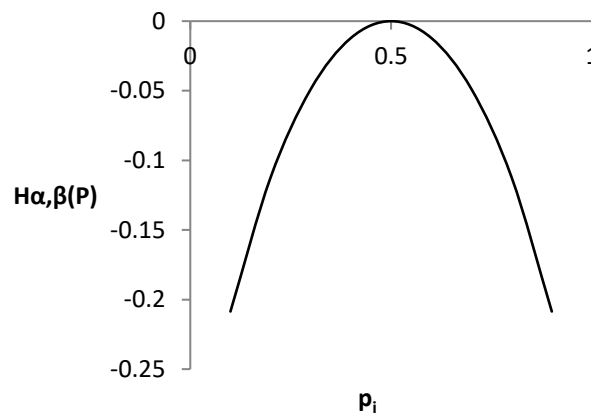
Figure1. Concavity of $H_{\alpha,\beta}(P)$ **with respect to *p*.**

II. We, now propose the following weighted entropic model:

$$H_{\alpha}(P;W) = \frac{\sum_{i=1}^{n} w_i p_i^{\alpha} - n^{1-\alpha}}{\alpha(1-\alpha)}; \ \ \alpha > 0, \alpha \neq 1 \tag{2.2}$$

Obviously, ignoring weights (2.2) reduces to Kapur's (1994) entropy introduced in (1.3). Thus, we observe that $H_{\alpha}(P)$ introduced in (2.2) is generalized entropy.

Properties:

(1) $H_\alpha(P;W)$ is continuous.

(2) $H_\alpha(P;W)$ is symmetric.

(3) $H_\alpha(P;W)$ is expansible.

(4) **Concavity:**

The second derivative of the model (2.2) is

$$\frac{\partial^2}{\partial p_i^2} H_\alpha(P;W) = -w_i p_i^{\alpha-2} < 0$$

which proves that $H_\alpha(P;W)$ is concave.

(5) **Entropy Maximization**

For entropy maximization, we differentiate the following Lagrangian

$$L = \frac{\sum_{i=1}^{n} w_i p_i^\alpha - n^{1-\alpha}}{\alpha(1-\alpha)} - \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$$

and putting the derivatives equal to zero, we get

$$\frac{p_1^{\alpha-1}}{(1-\alpha)} = \frac{\lambda}{w_1}, \frac{p_2^{\alpha-1}}{(1-\alpha)} = \frac{\lambda}{w_2}, ..., \frac{p_n^{\alpha-1}}{(1-\alpha)} = \frac{\lambda}{w_n}$$

which shows that each p_i is a function of w_i. Ignoring weights, we get

$$\frac{p_1^{\alpha-1}}{(1-\alpha)} = \frac{p_2^{\alpha-1}}{(1-\alpha)} = ... = \frac{p_n^{\alpha-1}}{(1-\alpha)}$$

which is possible only if $p_1 = p_2 = \Lambda = p_n$. Using the property $\sum_{i=1}^{n} p_i = 1$, we get $p_i = \frac{1}{n} \, \forall i$.

(6) **Maximum Value**

$$\left[H_\alpha(P;W)\right]_{max} = \frac{n^{1-\alpha} - n^{1-\alpha}}{\alpha(1-\alpha)} = 0$$

(7) For degenerate distributions $(1,0,0K,0),(0,1,0K,0),K,(0,0,0K,1)$, we have

$$H_\alpha(P;W) = \frac{1 - n^{1-\alpha}}{\alpha(1-\alpha)} < 0; \, \forall \alpha > 0$$

(8) **Negativity:** $H_\alpha(P;W) < 0$

The negativity of the weighted entropy model clearly indicates that it has resemblance with Burg's (1972) entropy.

With the above mentioned properties, we claim that the weighted entropic model introduced in (2.2) is an appropriate model of information entropy for weighted distributions.

Concluding Remarks:

It is well accredited reality that information theory plays an important role in designing various techniques for data compression and deals with an assortment of parametric, non-parametric, weighted and non-weighted entropic models for probability distributions but still there is inevitability to develop more parametric models for furtherance of research and broaden the applications areas for assortment of disciplines of mathematical sciences. The proposal of two weighted and non-weighted entropic models is a step in this direction for the furtherance of research. Many new such entropic models can be created from application point of analysis to a choice of an assortment of mathematical disciplines.

References

[1]. Aggarwal, N. L., & Picard, C. F. (1975). Functional equations and information measures with preference. *Kybernetika*, 14, 174-181.

[2]. Bajic D., Đaji´c V., & Milovanovi´c B. (2021). Entropy analysis of COVID-19 cardiovascular signals. *Entropy*, 23(1), 87 (1-15).

[3]. Bulinski, A., & Kozhevin, A. (2019). Statistical estimation of conditional Shannon entropy. *ESAIM: Probability and Statistics*, **23,** 350–386.

[4]. Burg, J. P. (1972). The relationship between maximum entropy spectra and maximum likelihood spectra. In: Childrers, D.G. (ed). *Modern Spectral Analysis*. Pp.130-131.

[5]. Cincotta, P. M., & Giordano, C. M. (2018). Phase correlations in chaotic dynamics: a Shannon entropy measure. *Celestial Mechanics and Dynamical Astronomy*, 130(11), 130:74.

[6]. Du Y. M., Chen J. F., Guan X., & Sun C. P. (2021). Maximum entropy approach to reliability of multi-component systems with non-repairable or repairable components. *Entropy*, 23(3), 348(1-14).

[7]. Guiasu, S. (1971). Weighted entropy. *Reports on Mathematical Physics*, 2, 165-171.

[8]. Gurdial, & Pessoa, F. (1977). On useful information of order α. *Journal of Combinatorics Information and System Sciences*, 2, 158-162.

[9]. Huang, W., & Zhang, K. (2019). Approximations of Shannon mutual information for discrete variables with applications to neural population coding. *Entropy*, 21(3), Paper No. 243, 21 pp.

[10]. Kapur, J.N. (1994). *Measures of Information and their Applications*. Wiley Eastern, New York.

[11]. Kumar, R., Singh, S., Bilga, P.S., Jatin, Singh, J., Singh, S., Scutaru, M.L., & Pruncu, C.I. (2021). Revealing the benefits of entropy weights method for multi-objective optimization in matching operations: A critical review. *Journal of Materials Research and Technology*, 10, 1471-1492.

[12]. Lenormand, M., Samaniego, H., Chaves, J.C., Vieira, V.D.F., Barbosa da Silva, M.A.H., & Evsukoff, A.G. (2020). Entropy as a measure of attractiveness and socioeconomic complexity in Rio de Janeiro metropolitan area. *Entropy*, 22, 368(1-18).

[13]. Longo, G. (1972). *Quantitative-Qualitative Measures of Information*. Springer Verlag, New York.

[14]. Majumdar, K., & Jayachandran, S. (2018). A geometric analysis of time series leading to information encoding and a new entropy measure. *Journal of Computational and Applied Mathematics*, 328, 469–484.

[15]. Parkash, O., & Kakkar, P. (2014a). New information theoretic models, their detailed properties and new inequalities. *Canadian Journal of Pure and Applied Sciences*, 8(3), 3115-3123.

[16]. Parkash, O., & Kakkar, P. (2014b). Generating entropy measures through quasilinear entropy and its generalized forms, In: Parkash. O. (ed). Mathematical Modeling and applications, Lambert Academic Publishers, pp. 42-53.

[17]. Parkash, O., Kumar, R., Mukesh, & Kakkar, P. (2019). Weighted measures and their correspondence with coding theory. *A Journal of Composition Theory*, 12(9), 1670-1680.

[18]. Parkash, O., & Mukesh (2012). New generalized parametric measure of entropy and cross entropy. *American Journal of Mathematics and Sciences*, 1(1), 91-96.

[19]. Parkash, O., & Mukesh (2014). Portfolio optimization using measures of cross entropy. *Canadian Journal of Pure and Applied Sciences*, 8(2), 2963-2967.

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

[21]. Sholehkerdar, A., Tavakoli, J., & Liu, Z. (2020). Theoretical analysis of Tsallis entropy-based quality measure for weighted averaging image fusion. *Information Fusion*, 58, 69-81.

[22]. Wang, Z., Yue, H., & Deng, J. (2019). An uncertainty measure based on lower and upper approximations for generalized rough set models. *Fundamenta Informatica*, 166(3), 273–296.