MATHEMATICAL MODELING OF AVERAGE WEIGHTED RENYI'S ENTROPY MEASURE

Savita, Rajeev Kumar

•

Department of Mathematics, Maharishi Dayanand University, Rohtak.
savitanain25@gmail.com, profrajeevmdu@gmail.com

Abstract

A weighted entropy measure of information is provided by a probabilistic experiment whose basic events are described by their objective probabilities and some qualitative (objective or subjective) weights. Weighted entropy has also been applied to equity the amount of information and degree of homogeneity related with a partition of data in classes. These measures have tremendous applications and are found to be quite helpful in many fields. In the present paper, a new weighted Renyi's entropy measure is proposed for the discrete distributions when probabilities are unknown and weights are known. The various characteristics of the measure are investigated. The measure is also studied taking into a particular case. In the last, numerical computation and graphical analysis is also done. Based on the graphical analysis, it is concluded that the proposed measure varies with values of weights and is concave in nature. The developed weighted information measure is useful for the discrete distribution when probabilities are unknown and weights are known.

Keywords: Shannon entropy, Renyi entropy, Weighted entropy, Symmetry, Concavity.

I. Introduction

Shannon's [22] entropy is widely prevalent in the study of probabilistic phenomena pertaining to abroad spectrum of problems. It was given by Shannon as a mathematical function to measure the uncertainty involved in a probabilistic experiment. Shannon entropy for a discrete source is defined as follows:

Let X be a probabilistic experiment with sample space x and probability distribution P, where $p(x_i)$ or p_i is the probability of outcome $x_i \in X$. Then the average amount of information is given by

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

Renyi [19] introduced a flexible extension of Shannon entropy and also, it is one parameter generalization of Shannon entropy. In the analysis of quantum systems and measure of randomness, Renyi entropy is mostly used. Renyi defined entropy as

$$H_\alpha(P) = \frac{1}{1-\alpha}\log\frac{\sum_{i=1}^{n} p_i^{\alpha}}{\sum_{i=1}^{n} p_i}, \quad \alpha \neq 1, \alpha > 0 \tag{2}$$

which is also known as Renyi's entropy measure of order α . Notice that Shannon's entropy measure is the limiting case of Renyi's entropy measure when $\alpha \to 1$.

The probabilistic measure of entropy (1) acquires a number of interesting properties. After the development of this measure, researchers found the potential of the application of this measure as an important quantity in many fields, from probability theory to engineering, ecology, and neuroscience. On the basis of this a large number of other information theoretic measures have also been derived.

Occasionally, standard distributions in statistical modelling are not appropriate for our data and we need to study weighted distribution. This concept has been applied in a variety of statistical fields, including family size analysis, human heredity, world life population study, renewal theory, biomedical and statistical ecology.

Fisher [5] originated the concept of a weighted distribution from his research into the impact of ascertainment methods on frequency estimates. Rao [17,18] expanded on Fisher's core ideas by discussing the need for a unifying idea and identifying several sample scenarios that can be replicated by what he called weighted distributions. The particular case of weighted distribution is the Size-biased distribution. These distributions naturally appear in real-world situations when observations from a sample are recorded with unequal probability. The utility distribution $W = (w_1, w_2, ...w_n)$, where each w_i is a non-negative real number, is proposed by Belis and Guiasu [1] to measure utility aspect of the outcomes.

The applications of the measure to the theory of questionnaires were given by Guiasu and Picard [7]. Longo [12] applied this useful measure to coding theory. Moreover, in many situations, there is need of a measure of uncertainty of a distribution whenever probabilities are unknown and however weights for each value of the random variable are known, i.e. if X be a discrete random variable having values $x_1, x_2, ...x_n$ having weights $w_1, w_2, ...w_n$ respectively but $p_1, p_2, ...p_n$ are unknown. Patsakis et al. [16] gave the applications of the measure in security quantification.

The suitable generalization of classical entropy is the weighted entropy, which has been proposed by Belis and Guiasu [1], Guiasu [6]. A detailed discussion on weighted entropies have been made by Suhov and sekeh [24]. Mahdy [13] studied the weighted entropy measures and its application in Reliability theory and stochastic. Using two weighted entropy measures, Singh et al. [23] provided the applications of Holder's inequality to coding theory. Some other substantial measures of weighted entropy introduced by Kapur [9] are as under:

- $$H_\alpha(P:W) = -\sum_{i=1}^{n} w_i p_i^\alpha \ln p_i, \quad 1/2 \le \alpha \le 1 \tag{3}$$

- $$H_a(P:W) = -\sum_{i=1}^{n} w_i p_i \ln p_i + 1/a \sum_{i=1}^{n} w_i (1+ap_i)\ln(1+ap_i) - w_{min.} \sum_{i=1}^{n} (1+a)\ln(1+a)p_i, \tag{4}$$

where $w_{min.} = \min(w_1, w_2, ...w_3)$

Some characterizations and generalizations of the weighted measure have been provided by Longo [12], Hooda and Tuteja [8], Taneja and Hooda [25], Parkash and Taneja [15], Taneja and Tuteja [26], Kapur [9], Kumar et.al. [10, 11], Endo and Kudo [4], Mohammadi [14], Savita and Kumar [21], Bhat and Pundir [2], Sahni and Kumar [20] etc. Thus, weighted measures of information find tremendous applications and are quite helpful to the researchers in many fields.

II. Average Weighted Entropy Measure

Average weighted Renyi's entropy measure is proposed as under:

$$H_1(w) = \frac{1}{1-\alpha} \ln\left(\sum_{i=1}^{n} (w_i)^\alpha / \sum_{i=1}^{n} w_i \right) \tag{5}$$

where $w_i > 0$ and $\sum_{i=1}^{n} w_i = w$ (constant).

The important characteristics of the measure (5) are investigated as under:

- It is a continuous and non -increasing function of α.
- It is permutationally symmetric function of $w_1, w_2, ...w_n$, i.e., it does not change when $w_1, w_2, ...w_n$ are permuted among themselves.
- $H_1(w)$ is non-negative for $\alpha < 0$ and negative for $\alpha > 0$.
- Expansible property: This property is satisfied for the measure (5) which states that entropy does not change by the addition of weight with zero value.

  $H_1(w_1, w_2, ..., w_n, 0) = H_1(w_1, w_2, ..., w_n)$

  Here we use convention $0^\alpha = 0$ for all real values of α.
- The maximum value of the function can be obtained by considering following Lagrangian:

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

  Now $\frac{\partial L}{\partial w_1} = \frac{1}{(1-\alpha)} \left[\frac{\alpha w_1^{\alpha-1}}{\sum_{i=1}^{n} w_i^\alpha} \right] - \lambda$

  $$\frac{\partial L}{\partial w_2} = \frac{1}{(1-\alpha)} \left[\frac{\alpha w_2^{\alpha-1}}{\sum_{i=1}^{n} w_i^\alpha} \right] - \lambda$$

  Continuing like this

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

  Now $\frac{\partial L}{\partial w_i} = 0 \Rightarrow \frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial w_3} = ... = \frac{\partial L}{\partial w_n} = 0$

  $$\frac{1}{(1-\alpha)} \left[\frac{\alpha w_1^{\alpha-1}}{\sum_{i=1}^{n} w_i^\alpha} \right] - \lambda = \frac{1}{(1-\alpha)} \left[\frac{\alpha w_2^{\alpha-1}}{\sum_{i=1}^{n} w_i^\alpha} \right] - \lambda = ... = \frac{1}{(1-\alpha)} \left[\frac{\alpha w_n^{\alpha-1}}{\sum_{i=1}^{n} w_i^\alpha} \right] - \lambda$$

  $$w_1^{\alpha-1} = w_2^{\alpha-1} = w_3^{\alpha-1} = ... = w_n^{\alpha-1}$$

which is possible only if $w_1 = w_2 = w_3 = ... = w_n$

Also, $\sum_{i=1}^{n} w_i = w \Rightarrow nw_i = w \Rightarrow w_i = \frac{w}{n}$

Thus, the maximum value of function will exist at $w_i = \frac{w}{n}$ and is given by

$$H_1(w) = \frac{1}{(1-\alpha)} \ln \frac{\sum_{i=1}^{n} \left(\frac{w}{n}\right)^{\alpha}}{w} = \frac{\ln w^{\alpha-1} - \ln\left(n^{\alpha} - 1\right)}{(1-\alpha)}$$

- Additive property: The measure (5) is additive in nature since when

$$H_{n1}(w) = \frac{1}{1-\alpha} \ln\left(\sum_{i=1}^{n} (w_i)^{\alpha} / \sum_{i=1}^{n} w_i\right) \quad \text{and}\ H_{m1}(w') = \frac{1}{1-\alpha} \ln\left(\sum_{j=1}^{m} (w_j')^{\alpha} / \sum_{j=1}^{m} w_j'\right)$$

Then $$H_{n1+m1}(w \cup w') = \frac{1}{(1-\alpha)} \ln\left(\frac{\left(\sum_{i=1}^{n}\sum_{j=1}^{m} (w_i w_j')^{\alpha}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_i w_j'}\right) = H_{n1}(w) + H_{m1}(w')$$

- Concave property: Since for the measure (5),

$$H_1(w) = \frac{1}{1-\alpha} \ln\left(\frac{\sum_{i=1}^{n} (w_i)^{\alpha}}{\sum_{i=1}^{n} w_i}\right)$$

$$H_1'(w) = \frac{1}{(1-\alpha)} \frac{\alpha\left(\sum_{i=1}^{n} (w_i)^{\alpha-1}\right)}{\sum_{i=1}^{n} w_i^{\alpha}}$$

$$H_1''(w) = -\frac{\alpha\left(\sum_{i=1}^{n} (w_i)^{\alpha-2}\right)}{\left(\sum_{i=1}^{n} w_i^{\alpha}\right)} - \frac{\alpha^2\left(\sum_{i=1}^{n} (w_i)^{\alpha-1}\right)^2}{(1-\alpha)\left(\sum_{i=1}^{n} w_i^{\alpha}\right)^2}$$

Thus, the measure (5) is concave upward for $\alpha > 1$.

I. Particular Cases

- when $w_i = p_{i,} \quad 0 \le p_i \le 1$

then measure (5) becomes Renyi type entropy measure.

- when $w_i = p_{i,} \quad 0 \le p_i \le 1$ and $\alpha \to 1$

then measure (5) becomes Shannon type entropy measure.

- when $\alpha \to 1$

then measure (5) reduces to average weighted Shannon's entropy measure.

III. Numerical Computation and Graphical Analysis

For computation of various values of the measure $H_1(w)$ given in (5), three values of weights w_1, w_2 and w_3 such that $w = w_1 + w_2 + w_3$ are considered. Total weights w are assumed to take values 50, 100 and 150 and $\alpha = 3$. The computed values of these cases are presented in table 1. Various graphs are plotted for the obtained values of $H_1(w)$ w.r.t. different weights to study the behavior of the information measure (5).

Table 1: *Average weighted entropy measure*

W=150			
w1	w2	w3	$H_1(w)$
5	50	95	-4.3936
10	50	90	-4.3241
15	50	85	-4.2536
20	50	80	-4.1832
25	50	75	-4.1148
30	50	70	-4.0508
35	50	65	-3.9948
40	50	60	-3.9505
45	50	55	-3.9219
50	50	50	-3.9120
55	50	45	-3.9219
60	50	40	-3.9505
65	50	35	-3.9948
70	50	30	-4.0508
75	50	25	-4.1148
80	50	20	-4.1832
85	50	15	-4.2536
90	50	10	-4.3241
95	50	5	-4.3936

W=100			
w1	w2	w3	$H_1(w)$
5	30	65	-4.0061
10	30	70	-3.8999
15	30	55	-3.7923
20	30	50	-3.6889
25	30	45	-3.5993
30	30	40	-3.5366
35	30	35	-3.5139
40	30	30	-3.5366
45	30	25	-3.5993
50	30	20	-3.6889
55	30	15	-3.7923
60	30	10	-3.8999
65	30	5	-4.0061

W=50			
w1	w2	w3	$H_1(w)$
5	10	35	-3.3900
10	10	30	-3.1815
15	10	25	-2.9957
20	10	20	-2.9145
25	10	15	-2.9957
30	10	10	-3.1815
35	10	5	-3.3900

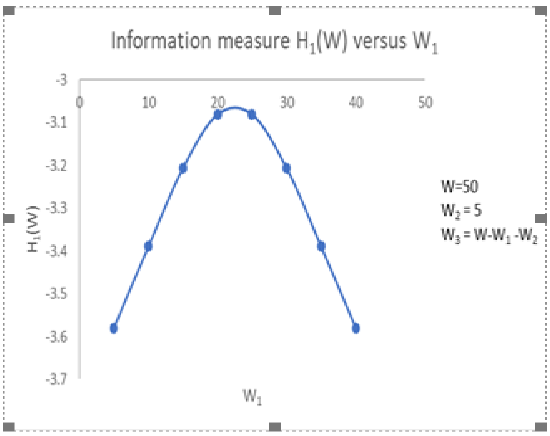

Figure 1: *Information measure* $H_1(w)$ *w.r.t.* w_1 *for* $w = 50$ *and* $w_2 = 5$

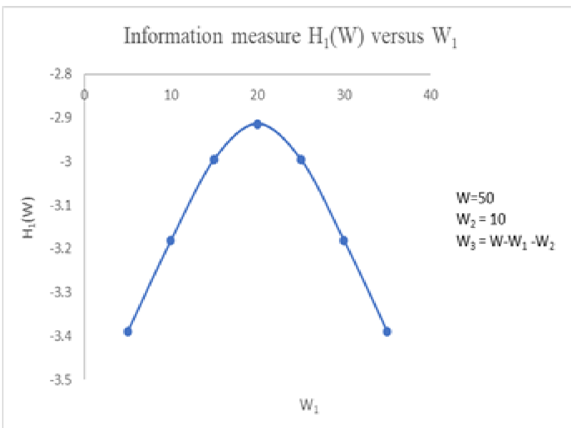

Figure 2: *Information measure* $H_1(w)$ *w.r.t.* w_1 *for* $w = 50$ *and* $w_2 = 10$

Figure 1 and Figure 2 depict the graph between $H_1(w)$ and w_1 for different values of w_i, total weight w = 50 and we have fixed w_2 = 5,10. From the graphs, it can be seen that $H_1(w)$ increases as weights increases up to $w_i = \frac{w}{n}$, and thereafter it decreases.

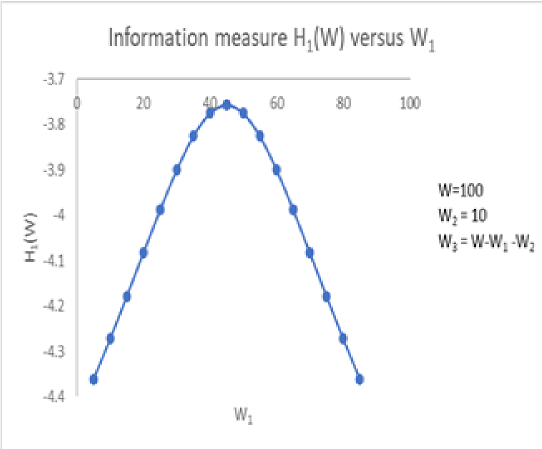

Figure 3: *Information measure* $H_1(w)$ *w.r.t.* w_1 *for* $w = 100$ *and* $w_2 = 10$

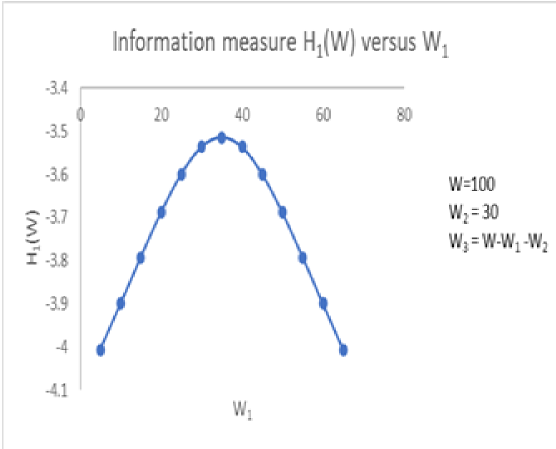

Figure 4: *Information measure* $H_1(w)$ *w.r.t.* w_1 *for* $w = 100$ *and* $w_2 = 30$

Figure 3 and Figure 4 depict the graph between $H_1(w)$ and w_1 for different values of w_i, total weight w = 100 and we have fixed w_2 = 10, 20, 30. From the graphs, it can be seen that $H_1(w)$ increases as weights increases up to $w_i = \frac{w}{n}$, and thereafter it decreases.

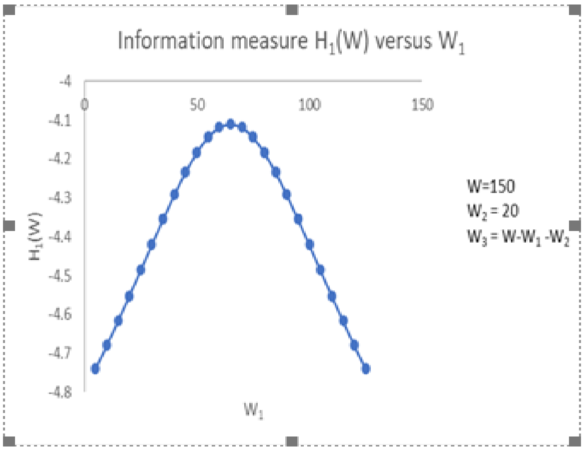

Figure 5: *Information measure* $H_1(w)$ *w.r.t.* w_1 *for* $w = 150$ *and* $w_2 = 20$

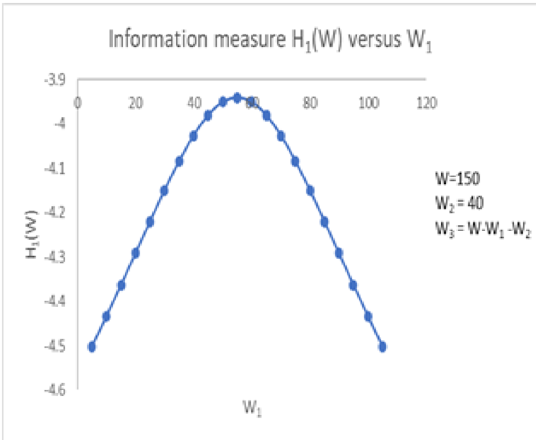

Figure 6: *Information measure* $H_1(w)$ *w.r.t.* w_1 *for* $w = 150$ *and* $w_2 = 40$

Figure 5 and Figure 6 depict the graph between $H_1(w)$ and w_1 for different values of w_i, total weight w = 150 and we have fixed w_2 = 10, 20, 30, 40, 50. From the graphs, it can be seen that $H_1(w)$ increases as weights increases up to $w_i = \frac{w}{n}$, and thereafter it decreases.

IV. Conclusion

The proposed weighted Renyi's measure varies with values of weights and is concave in nature. In case, weights of the distribution are their probabilities, the average weighted Renyi entropy measure reduces to Renyi 's entropy. In the weighted sense, this is in fact generalizations of the Shannon [22] entropy and Burg [3] entropy. The developed weighted information measure is useful for the discrete distributions when probabilities are unknown and weights are known.

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. https://doi.org/10.1109/TIT.1968.1054185

[2] Bhat, V.A. and Pundir, S. (2023). Weighted intervened exponential distribution as a lifetime distribution. *Reliability: Theory & Applications.* 2(73): 24-38. https://doi.org/10.24412/1932-2321-2023-273-253-266

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

[4] Endo, T. and Kudo, M. (2013). Weighted naïve Bayes classifiers by Renyi entropy. *J. Ruiz-Shulcloper and G. Sanniti di Baja (Eds.), Springer-Verlag Berlin Heidelberg, CIARP, Part I, LNCS 8258,* 149–156.

[5] Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *The Annals of Eugenics*, 6: 13-25.

[6] Guiasu, S. (1986). Grouping data by using the weighted entropy. *Journal of Statistical Planning and Inference*, 15:63-69. https://doi.org/10.1016/0378-3758(86)90085-6

[7] Guiasu, S. and Picard, C.F. (1971). Borne in ferictur de la longuerur utile de certains code. *Comptes Rendus Mathematique Academic Des Sciences Paris*, 273: 248-251.

[8] Hooda, D.S. and Tuteja, R.K. (1981). Two generalized measures of useful information. *Information Sciences*, 23: 11-24.

[9] Kapur. J.N. (1995). Measures of information and their applications. *Wiley Eastern*, New York.

[10] Kumar, R., Kumar, S. and Kumar, A. (2010). A new measure of probabilistic entropy and its properties. *Applied Mathematical Sciences*, 4(28): 1387-1394**.**

[11] Kumar, R., Kumar, S. and Kumar, A. (2010). Some new information theoretic measures based upon statistical constants. *Aryabhatta Journal of Mathematics and Informatics*, 2(2): 329-338.

[12] Longo, G. (1972). Quantitative-qualitative measure of information. *Springer-Verlag*, New York.

[13] Mahdy, M. (2018). Weighted entropy measure: a new measure of information with its properties in reliability theory and stochastic orders. Journal of Statistical Theory and Applications. 17(4):703-718. DOI:10.2991/jsta.2018.4.17.11

[14] Mohammadi, U. (2015). Weighted entropy function as an extension of the Kolmogorov-Sinai entropy. *U.P.B. Sci. Bull.*, Series A, 77(4): 117-122.

[15] Parkash, O. and Taneja, H.C. (1986). Characterization of quantitative-qualitative measure of inaccuracy for discrete generalized probability distributions. *Comm. Statistics theory and Methods*, 15(12): 3763-3772.

[16] Patsakis, C., Mermigas, D., Pirounias, S. and Chondrokoukis, G. (2013). The role of weighted entropy in security quantification. *International Journal of Information and Electronics Engineering,* 3(2):156-159.

[17] Rao, C.R. (1965). On discrete distributions arising out of methods of ascertainment, in Classical and Contagious Discrete Distributions. *G.P. Patil, ed., Pergamon Press and Statistical Publishing Society, Calcutta*,320–332.

[18] Rao, C.R. (1985). Weighted distributions arising out of methods of ascertainment, in *A* Celebration of Statistics, *A.C. Atkinson & S.E. Fienberg, eds, Springer-Verlag,* New York, Chapter 24:543–569.

[19] Renyi, A. (1961). On measures of entropy and information. *Proceedings IV Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, 20-30 June 1961,*1: 547-561.

[20] Sahni, P. and Kumar, R. (2023). Huntsberger type shrinkage entropy estimator for variance of normal distribution under linex loss function. *Reliability: Theory & Applications.* 1(72): 491-499. https://doi.org/10.24412/1932-2321-2023-172-491-499

[21] Savita and Kumar, R. (2019). On A New Weighted Entropy Measure. *Journal of Advanced research in dynamical and control systems.* 11(01)-special issue:1264-1271.

[22] Shannon, C.E (1948). A mathematical theory of communication, *Bell System Technical Journal*, 27:379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

[23] Singh, R.P., Kumar, R. and Tuteja, R.K. (2003). Applications of holder's inequality in information theory. *Information Sciences*, 152: 145-154.

[24] Suhov, Y., Sekeh, S.Y. (2015). Weighted cumulative entropies: An extension of CRE and CE. *arXiv: 1507.0705*[v][cs.IT]. https://doi.org/10.48550/arXiv.1507.07051

[25] Taneja, H.C. and Hooda, D.S. (1983). On characterization of generalized measure of useful information, *Sooch. Jr. Math*, 9: 221-230.

[26] Taneja, H.C. and Tuteja, R.K. (1986). Characterization of quantitative-qualitative measure of inaccuracy, *Kybernetika,* 22: 393-402.