International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

Measures of Entropy and Their Equivalence

H. R. Trivedi

Associate Professor, Department of Mathematics, D. B. Science College Gondia (441614), India

Abstract: ***With the emergence of new measures of entropy there is a need to find relation between entropies. There are various methods to deal with these. The present paper deals with the the equivalence of Kapur's [4] and Behra Chawla's [1] measures, using three different methods.***

Keywords: Measures of Entropy, Variance,Measure of uncertainty,MEPD, ,proof of equivalence

1. Introduction

Let $P = (p_1, p_2, p_3 \ldots \ldots \ldots . p_n)$ be a probability distribution then Shannon[10] gave the measure of entropy

$$S(P) = -\sum_{i=1}^{n} p_i ln p_i \qquad (1)$$

to measure the uncertainty, diversity or equality represented by P.

Later *Behra Chawla[1]* and *Kapur[4]* gave following measures of entropy:

$$B(P) = \frac{1 - (\sum_1^n {p_i}^{1/\gamma})^\gamma}{1 - 2^{\gamma-1}}, \quad \gamma > 0, \quad \gamma \neq 1 \qquad (2)$$

$$and K(P) = \frac{1 - (\sum_1^n {p_i}^{\frac{1}{\gamma}})^\gamma}{1-\gamma} \gamma > 0, \quad \gamma \neq 1 \qquad (3)$$

Both these measures are non-additive and obtained from quite different consideration. It is also obvious that $S(P)$ is the limiting care of $B(P)$ and $K(P)$

$$\lim_{\propto \to 1} \mathrm{B(P)} = S(P) and \lim_{\propto \to 1} \mathrm{K(P)} = S(P) \qquad (4)$$

Now suppose only partial information about $p_1, p_2, p_3 \ldots \ldots \ldots . p_n$ in terms of moments are available in the following form.

$$p_1, p_2, p_3 \ldots \ldots \ldots . p_n \geq 0, \sum_1^n p_i = 1 \; and \sum_1^n p_i \, g_{ri} = a_r \, (5)$$

$$r = 1,2, \ldots \ldots \ldots m, \quad m + 1 < n$$

These are in general not sufficient to determine p_1, p_2, p_n uniquely. However with the help of *Jaynes [3]* maximum entropy principle "we should choose $p_1, p_2, \ldots \ldots \ldots . p_n$ which maximizes $S(P)$ subject to (5)" we can find $p_1, p_2, \ldots \ldots \ldots . p_n$.

Later *Kapur*and *Kesavan [6]* generalized Jayne's principle and stated that we shall choose $p_1, p_2, \ldots \ldots \ldots . p_n$ which maximizes any other measure of entropy subject to *(5)Nayak [7]*describes that in general different measures of entropy arrange a given set of probability distribution in different order for their uncertainty or diversity. Thus one measure may give the result that P is more uncertain then Q, while another measure may give the result that P is less uncertain then Q. In spite of this different measures of entropy have been used in Economics, Genetics, Sociology, Ecology and so many other fields *[11]*, because of the fitness of different measures for different situations.

However some measures in spite of being different, may lead to same arrangements and inparticular they may lead to the same probability distribution as MEPD we call these measures as equivalent from the generalized maximum entropy point of view or from the point of view of arranging probability distributions according to their entropies or equalities or diversities. *Kapur [5]*in his famous treatise has shown *Havrda-Charvat [2],Renyi [8] Behra Chawla [1] Sharma mittal [9]* entropies as equivalent entropies. Using the same methodology we will prove in section 2 equivalence of $B(P)$ $and K(P)$

2. Proof of Equivalence

From (2) and (3) it is obvious that

$$(\sum {p_i}^\gamma)^\gamma = 1 - (1-\gamma)K(P) = \; 1 - (1 - 2^{\gamma-1})B(P) \quad (6)$$

Now take the case $\gamma > 1$

$$K(P) \gtrless K(Q) \; \Leftrightarrow (1-\gamma)K(P) \gtrless (1-\gamma)K(Q)$$
$$\Leftrightarrow 1 - (1-\gamma)K(P) \gtrless 1 - (1-\gamma)K(Q)$$
$$\Leftrightarrow 1 - (1 - 2^{\gamma-1})B(P) \gtrless 1 - (1 - 2^{\gamma-1})B(Q)$$
$$B(P) \Leftrightarrow B(Q) \qquad (7)$$

This proves that $K(P)$ and $B(P)$ increases or decreases together when $\gamma < 1$

$$K(P) \gtrless K(Q) \Leftrightarrow (1-\gamma)K(P) \gtrless (1-\gamma)K(Q)$$
$$\Leftrightarrow 1 - (1-\gamma)K(P) \gtrless 1 - (1-\gamma)K(Q)$$
$$\Leftrightarrow 1 - (1 - 2^{\gamma-1})B(P) \gtrless 1 - (1 - 2^{\gamma-1})B(Q)$$
$$\Leftrightarrow -(1 - 2^{\gamma-1})B(P)$$
$$\gtrless 1 - (1 - 2^{\gamma-1})B(Q)$$
$$\Leftrightarrow B(P) \gtrless B(Q) \qquad (8)$$

So it is established that $K(P)$ and $B(P)$ increases or decreases together when $\gamma < 1$. In case$\gamma = 1$, $K(P)$ and $B(P)$ both becomes $S(P)$ so equivalent.

3. Another Proof of Equivalence

From (6) $\; 1 - \gamma K(P) = \; (1 - 2^{\gamma-1})B(P)$

$$\because \frac{dK(P)}{dB(P)} = \frac{1 - 2^{\gamma-1}}{1-\gamma} > 0 \; for all values of \gamma \gtrless_1^1 \quad (9)$$

So,$K(P)$ and $B(P)$ increases or decreases together. Since $\sum_1^n p_i = \sum_1^n q_i = 1$. So, MEPD given by $K(P)$ and $B(P)$ are same.

4. Alternative Proof of Equivalence

Using Lagrange's method maximizing $B(P)$and $K(P)$ subject to constraints (5) we get

$Lagrangian L \equiv B(P) + \lambda_0(\sum p_i - 1) + \lambda_1 (\sum p_i g_{1i} -$ *a1)+λ2(pig2i−a2)+ ……… . λm(pigmi−am) . ..(10)*

$$\frac{\partial L}{\partial p_i} = \Longrightarrow \frac{(\sum_1^n {p_i}^{1/\gamma})^{\gamma-1} {p_i}^{1/\gamma-1}}{1-2^{\gamma-1}} = \lambda_0 + \lambda_1 g_{1i} + \lambda_2 g_{2i} + \ldots\ldots\ldots + \lambda_m g_{mi} \quad \ldots(11)$$

$LagrangianL \equiv K(P) + \lambda_0{}'(\sum p_i - 1) + \lambda_1{}'(\sum p_i g_{1i} - a1) + \ldots\ldots\ldots\ldots + \lambda m'(pigmi - am)$

$$\frac{\partial L}{\partial p_i} = \Longrightarrow \frac{(\sum_1^n {p_i}^{1/\gamma})^{\gamma-1} {p_i}^{1/\gamma-1}}{1-\gamma} = \lambda_0{}' + \lambda_1{}' g_{li} + \lambda_2{}' g_{zi} + \ldots\ldots\ldots\ldots\ldots\ldots + \lambda_m{}' g_{mi} \quad \ldots(12)$$

where $(\lambda_0, \lambda_1, \lambda_2 \ldots\ldots\ldots \lambda_m) and\ (\lambda_0{}', \lambda_1{}', \ldots\ldots\ldots\ldots \lambda_m{}')$ are obtained by using the constraint (5)

Equations (11) and (12) can be written as

$$p_i = (\mu_0 + \mu_1 g_{1i} + \mu_2 g_{2i} + \ldots\ldots\ldots\ldots + \mu_m g_{mi})^{\frac{\gamma}{1-\gamma}} \quad (13)$$

where $\mu_0, \mu_1, \mu_2 \ldots\ldots\ldots\ldots \mu_m$ are obtained by using the constraints (5) and as such have the same values in case of $B(P)$and $K(P)$. Thus we have established MEPD under constraint (5) remains the same whether we use $B(P)$or $K(P)$

5. Acknowledgement

Author is very much thankful to Dr. S.K. Varma Ex. Vice Chancellor, Sarguja University, Chhattisgarh for their kind support.

References

[1] Behra M. and Chawla J. S. (1974) "Generalised Gamma Entropy" Selecta Statistica Canadian Vol. 2, p. 15-38
[2] Havrda J. H. and Charvat F.C. (1967) "Quantification Methods of Classification Processes Concept of Structural & Entropy" Kybernetika Vol. 3, P. 30-35.
[3] Jaynes E.T. (1957) Information Theory and Statistical Mechanics Physical Review. Vol. 106 P. 620-630.
[4] Kapur J. N. (1993) "Maximum Entropy Models in Science & Engineering Wiley Eastern New Delhi P. 7
[5] Kapur J. N. (1994) "Measures of Information and Their Applications Wiley Eastern Limited New Delhi Chapter 15
[6] Kapur J. N. and Kesavan (1987) Generalised Maximum Entropy principle (with applications) Sanford Educational Press University of Waterloo Canada
[7] Nayak T.K. (1985) "On Diversity Measures based on Entropy function" Communications in Statistics 14(11) 223
[8] Renyi A (1961) "On Measures of Entropy and Information's" Proc First Berkeley Symp Stat. 1. 547-561
[9] Sharma B.D. and Mittal D. P. (1975) "New Non additive Measures of entropy for Discrete Probatrility Distributions J. Maths Vol. 10 p. 28-40
[10] Shannon C. E. (1948) "A Mathematical Theory of Communications" Bell System Tech Journ, 27. p. 379-423.

Author Profile

Haresh. R. Trivedi, is an Associate Professor in Department of Mathematics completed his M.Sc and Ph.D in Mathematics, the work area is Maximum Entropy Principle, published and presented number of papers in Conferences and seminar working at D B Science College since last 31 years, Life member ISITA.