ON UNORTHODOX APPROACH TO GENERATING MEASURES OF INFORMATION

Sandeep Singh*

Abstract

A new method for obtaining measures of information is given and used to obtain related measures of information i.e., measure of inaccuracy, measure of directed divergence and measure of entropy.

Key Words

Measure of inaccuracy, Directed Divergence, Entropy

Introduction

Let $P = (p_1, p_2, ..., p_n)$ and $Q = (q_1, q_2, ..., q_n)$ be two non degenerate complete probability distributions such that

$$p_i, q_i > 0 \forall i = 1, 2, ..., n \qquad (1)$$

and

$$\sum_{i=1}^{n} p_i = 1 = \sum_{i=1}^{n} q_i \qquad (2)$$

Then the following measures of inaccuracy [4], measures of directed divergence and measures of entropy [6] respectively are well known in literature respectively:

$$I(P:Q) = -\sum_{i=1}^{n} p_i \ln q_i \qquad (3)$$

$$D(P:Q) = -\sum_{=i=1}^{n} p_{i1} \ln \frac{p_i}{q_i} \qquad (4)$$

$$\text{and } S(P) = -\sum_{i=1}^{n} p_i \ln p_i \qquad (5)$$

Kapur [3] gave a new approach to generate measures of inaccuracy, measures of directed divergence and measures of entrop. He has considered the function.

$$I^*(P:Q) = -\sum_{i=1}^{n} f(p_i) g(q_i) \qquad (6)$$

where (i), $f(p_i)$ is a positive continuous function.

(ii) g(q) is a convex function which is so chosen that

*Department of Mathematics, Institute of Technology, G.G. University, Bilaspur (C.G.), India

(a) $I^*(P:Q)$ *is minimum subject to (2) when* $q_i = p_i$ *for each I so that the minimum value of* $I^*(P:Q)$ *is* $I^*(P:P)$

(b) $I^*(P:Q)$ *is a concave function of p, p .... pn*

The condition (a) will be satisfied if

f(p) (P)=A (a constant) .... (7)

so that $g'(p_i) = \dfrac{A}{f(p_i)}$ .... (8)

So $g(p_i) = \int_c^{p_i} \dfrac{A}{f(p_1)} dp_1$ .... (9)

$g''(p_i) = \dfrac{Af'(p_i)}{f^2(p_i)}$ .... (10)

Since we want $g(p_i)$ to be convex so $Af'(p_i) < 0$. The minimum value of I(P:Q) is

$$\sum_{i=1}^{n} f(p_i) \int_c^{q_i} \frac{A}{f(p_i)} dp_i$$ .... (11)

If $I^*(P:P)$ is a concave function of p_1, p_2, ……….. p_n and vanishes for degenerate distribution it can be used as a measure of entropy since it is clearly a permutationally symmetric function of p_1, p_2, ……..p_n.

Let us consider the following cases (Kapur [2]) :

Case I: When $f(p_i)=p_i$ .... (12)

then $I(P:Q) = -\sum_{i=1}^{n} p_1 \ln q_i$ .... (13)

which is Kerridge's [4] measure of inaccuracy and

$I(P:P) = -\sum_{i=1}^{n} p_i \ln p_i$ .... (14)

is Shannon's [6] measure of entropy.

Case II: If $f(p_i) = p_i^a$, then .... (15)

$$I(P:Q)=\frac{\sum_{i=1}^{n}p_i^a\left(q_i^{1-a}-1\right)}{\alpha-1} \qquad \text{.................................} (16)$$

and $$I(P:P)=\frac{\sum_{i=1}^{n}p_i^a-1}{\alpha-1} \qquad \text{.................................} (17)$$

Which is Havrda-Charvat's [1] measure of entropy.

Case III: If $f(p_i)=p_i+p_i^2$ .... (18)

then $$I(P:Q)=\sum_{i=1}^{n}(p_i+p_i^2)\,ln\left(\frac{q_i+1}{q_i}\right) \qquad \text{.................................} (19)$$

and $$I(P:P)=\sum_{i=1}^{n}(p_i+p_i^2)\,ln\left(\frac{p_i+1}{p_i}\right) \qquad \text{.................................} (20)$$

This also gives a measure of directed divergence

$$D^*(P:Q)=\sum_{i=1}^{n}(p_i+p_i^2)\,ln\left(\frac{q_i+1}{p_i+1}\frac{p_i}{q_i}\right) \qquad \text{.................................} (21)$$

2. Main Result:

Let $f(p_i)=p_i+ap_i^2$ .... (22)

then $$g'(p_i)=\frac{A}{p_i(1+ap_i)}=A\left[\frac{1}{p_i}-\frac{a}{1+ap_i}\right] \qquad \text{.................................} (23)$$

Hence $$g(p_i)=A\,ln\left(\frac{p_i}{1+ap_i}\right)+B \qquad \text{.................................} (24)$$

Also $$g''(p_i)=-\frac{A(1+2ap_i)}{p_i^2(1+ap_i)^2} \qquad \text{.................................} (25)$$

Thus for g (p_i) to be convex $A<0$ and $a\geq-\frac{1}{2}$

Now $$g(p_i)=A\,ln\left(\frac{1+ap_i}{p_i}\right)+B, A>0 \qquad \text{.................................} (26)$$

So $$I(P:Q)\sum_{i=1}^{n}(p_1+ap_1^2)\left\{A\,ln\left(\frac{1+aq_1}{q_1}\right)+B\right\} \qquad \text{.................................} (27)$$

and $$I(P:P)\sum_{i=1}^{n}A(p_i+ap_i^2)\,ln\left(\frac{1+ap_i}{p_i}\right)+\sum_{i=1}^{n}BP_i+\sum_{i=1}^{n}Bap_i^2 \qquad \text{.................................} (28)$$

I(P:P) will be a measure of entropy if it vanishes for any degenerate distribution

$\Delta_r = (0,0,......,1,......,0)$

i.e. $A(1+a)ln(1+a) + B(1+a) = 0$ .... (29)

i.e. $B = -A\,ln(1+a)$ .... (30)

Thus $I(P:P) = A\left[\sum_{i=1}^{n}\left(p_i + ap_i^2\right)ln\left(\frac{1+ap_i}{p_i}\right) - ln(1+a)\right]$ .... (31)

is a valid measure of entropy when A > 0 and $a \geq \frac{1}{2}$

Particular Cases:

(i) When a = 0 and A = 1, then

$$I(P:P) = \sum_{i=1}^{n} p_i\, ln\, p_i$$ .... (32)

which is Shannon's [6] measure of entropy,

I(P:P) a measure of entropy suffers from an infirmity that it does not vanish for any degenerate probability distribution = (0, 0, …., 1, …, 0)

(ii) When $A = I, 1(P:P)$

$$= \sum_{i=1}^{n}\left(p_i + ap_i^2\right)\left\{ln\left(\frac{1+ap_i}{p_i}\right) - ln(1+a)\right\}$$ .... (33)

is a valid measure of entropy,

Measure of inaccuracy will be

$$I(P:Q) = \sum_{i=1}^{n} p_i(1+ap_i)\left\{ln\left(\frac{1+aq_i}{q_i}\right) - ln(1+a)\right\}$$ .... (34)

and measure of directed divergence will be

$D_a(P:Q) = I(P:Q) - I(P:P)$

$$= \sum_{i=1}^{n} p_i(1+ap_i)\,ln\left\{\frac{1+aq_i}{1+ap_i}\,\frac{q_i}{p_i}\right\}$$

When $a = 0,\ D_a(P:Q) = D(P:Q)$ given by Kullback and Leibler [5]

When $a = 1, D_a(P:Q) = D(P:Q)$ given by (21)

When $a = \frac{1}{\lambda}, I(P:Q) = \sum_{i=1}^{n} p_i\left(\frac{\lambda + p_i}{\lambda}\right)ln\left\{\frac{\lambda + q_i}{p_i(\lambda+1)}\right\}$

Which is Ferrari's measure of entropy and corresponding measure of inaccuracy is

$$I(P:Q)=\sum_{i=1}^{n} p_i\left(\frac{\lambda+p_i}{\lambda}\right) ln\left\{\frac{\lambda+q_i}{q_i(\lambda+1)}\right\}$$

and directed divergence is

$$D(P:Q)=\sum_{i=1}^{n} p_1\left(\frac{\lambda+p_i}{\lambda}\right) ln\left\{\frac{\lambda+q_1}{\lambda+p_i}\frac{p_i}{q_i}\right\}$$

References

1. J.H. Havrda and F. Charvat (1967): "Quantification methods of classification processes, concept of structural entropy", Kybernetika 3, 30-35.
2. J.N. Kapur (1994) : "Measures of information and their application", Wiley Eastern and Wiley.
3. J.N. Kapur (2000): "Geometry of probability spaces", MSTS, New Delhi.
4. D.E.F. Kerridge (1961): "Inaccuracy and inference", J. Roy, Soc. 23A, 184-194.
5. S. Kullback and R.A. Leibler (1951): "On information and Sufficiency", An. Math. State, 22, 77-89.
6. C.E. Shannon (1948): "The Mathematical Theory of Communication: Bell. Syst. Tech. Journ. 27, 423-467.