Global and Stochastic Analysis
Special Issue: *25th International Conference of Forum for Interdisciplinary Mathematics*

MUK PUBLICATIONS
Open Access Publisher

SYMMETRIC CHI-SQUARE DIVERGENCE, BOUNDS AND RESISTOR-AVERAGE DISTANCE

RAM NARESH SARASWAT AND NEHA KHATOD

ABSTRACT: In this paper we establish symmetric chi-square divergence measure with help of the new convex function and properties of new *f*-divergence measure. Upper and lower bounds of Jenson-Shannon's divergence, in terms of Symmetric chi-square divergence using a new *f*-divergence measure, numerical illustration and inequalities have studied. Relations between Symmetric chi-square divergence and resistor-average distance have also studied.
AMS Classification 62B-10, 94A-17,26D15

1. Introduction

Let

$$\Gamma_n = \left\{ P = (p_1, p_2, \ldots, p_n) \middle| p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \right\}, n \geq 2$$

be the set of all complete finite discrete probability distributions. There are many information and divergence measures exists in the literature on information theory and statistics. Jain and Saraswat [10] have introduced new *f*-divergence measure and its particular cases which is given by

$$S_f(P, Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i + q_i}{2q_i}\right) \tag{1.1}$$

where, $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a convex function and $P, Q \in \Gamma_n$.

The new *f*-divergence is a general class of divergence measures that includes several divergences used in measuring the distance or affinity between two probability distributions. This class is introduced by using a convex function *f*, defined on $(1/2, \infty)$.

Proposition 1.1. Let $f : [1/2, \infty) \to \mathbb{R}$ be convex and $P, Q \in \Gamma_n$ then we have the following inequality

$$S_f(P, Q) \geq f(1) \tag{1.2}$$

Equality holds in (1.2) iff

$$p_i = q_i \; \forall i = 1, 2, \ldots, n \tag{1.3}$$

Corollary 1.1. (Non-negativity of new *f*-divergence measure) Let $f : [0, \infty) \to \mathbb{R}$ be convex and normalized, i.e.

$$f(1) = 0 \tag{1.4}$$

Keyword. Chi-square divergence, Jenson-Shannon's divergence, Triangular discrimination etc.

Then for any P, Q $\in \Gamma_n$ from (1.2) of proposition 1.1 and (1.4), we have the inequality

$$S_f(P, Q) \geq 0 \tag{1.5}$$

If f is strictly convex, equality holds in (1.5) iff

$$p_i = q_i \ \forall i \in [i, 2, ..., n] \tag{1.6}$$

and

$$S_f(P, Q) \geq 0 \text{ and } S_f(P, Q) = 0 \text{ iff } P = Q \tag{1.7}$$

Proposition 1.2. Let f_1 & f_2 are two convex functions and $g = af_1 + bf_2$ then $S_g(P, Q) = aS_{f_1}(P, Q) + bS_{f_2}(P, Q)$, where a & b are constants and P, Q $\in \Gamma_n$.

2. SOME WELL-KNOWN DIVERGENCE MEASURES

It is shown that using new f-divergence measure we derive some well-known divergence measures such as Jenson-Shannon's divergence, symmetric chi-square divergence measure, We now give some examples of well-known information divergence measures which are obtained from new f-divergence measure.

- If $f(t) = \frac{t(t-1)^2}{(2t-1)}, \forall t > \frac{1}{2}$ then symmetric chi-square divergence is given by

$$S_f(P, Q) = \frac{1}{8}\left[\sum_{i=1}^{n} \frac{(p_i + q_i)(p_i - q_i)^2}{p_i q_i}\right] = \frac{1}{8}\Psi(P, Q) \tag{2.1}$$

- If $f(t) = -\log t$ then relative Jensen-Shannon divergence measure is given by

$$S_f(P, Q) = \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right) = F(Q, P) \tag{2.2}$$

3. SYMMETRIC CHI-SQUARE DIVERGENCE MEASURE

Now we consider the function $f : (1/2, \infty) \to \mathbb{R}$ given by

$$F(t) = \frac{t(t-1)^2}{(2t-1)} \tag{3.1}$$

$$F'(t) = t - \frac{1}{4(2t-1)^2} - \frac{3}{4} \tag{3.2}$$

$$F''(t) = 1 + \frac{1}{(2t-1)^3} > 0 \tag{3.3}$$

Function $f_k(t)$ is always convex, $\forall t > \frac{1}{2}$

Now applying new f-divergence measure property (2.1) on (3.1)

$$S_f(P, Q) = \frac{1}{8}\sum_{i=1}^{n} \frac{(p_i - q_i)(p_i + q_i)^2}{p_i q_i} = \frac{1}{8}\Psi(P, Q) \tag{3.4}$$

where, $\Psi(P, Q)$ is symmetric chi-square divergence measure.

4. NEW INFORMATION INEQUALITIES

The following theorem concerning inequalities among new f-divergence measure and symmetric chi-square divergence measure holds. Its particular cases are given in Section 6.

Theorem 4. Let $f:(0,\infty)\to\mathbb{R}$ is normalized mapping i.e. $f(1)=0$ and satisfy the assumptions.

(i) f is twice differentiable on (r, R), where $0\le r\le 1\le \mathrm{R}\le\infty$

(ii) there exist constants m, M such that

$$m \le \frac{2t(4t^2-6t+3)}{(2t-1)^3} f''(t) \le \mathrm{M} \tag{4.1}$$

If P, Q are discrete probability distributions satisfying the assumptions

$$r < \frac{1}{2} \le r_i = \frac{p_i+q_i}{2q_i} \le \mathrm{R},\ \forall i\in\{1, 2, \ldots, n\} \tag{4.2}$$

Then we have the inequality

$$m\,\Psi(\mathrm{P},\mathrm{Q}) \le \mathrm{S}_f(\mathrm{P},\mathrm{Q}) \le \mathrm{M}\,\Psi(\mathrm{P},\mathrm{Q}) \tag{4.3}$$

Proof: Define a mapping $\mathrm{F}_m:(0,\infty)\to\mathbb{R},\ \mathrm{F}_m(t)=f(t)-m\,\dfrac{t(t-1)^2}{(2t-1)},\ \forall t>\dfrac{1}{2}$.

Then $\mathrm{F}_m(.)$ is normalized, twice differentiable and since

$$\begin{aligned}\mathrm{F}''_m(t) &= f''(t) - \frac{2t(4t^2-6t+3)m}{(2t-1)^3}\\ &= \frac{2t(4t^2-6t+3)}{(2t-1)^3}\left[\frac{(2t-1)^3}{2t(4t^2-6t+3)} f''(t) - m\right] \ge 0\end{aligned} \tag{4.4}$$

For all $r\in(r,\mathrm{R})$, it follows that $\mathrm{F}_m(.)$ is convex on (r,R). Applying non-negativity property of new f-divergence functional for $\mathrm{F}_m(.)$ and the linearity property, we may state that

$$0\le \mathrm{S}_{\mathrm{F}_m}(\mathrm{P},\mathrm{Q}) = \mathrm{S}_f(\mathrm{P},\mathrm{Q}) - m\,\mathrm{S}_{\frac{t(t-1)^2}{(2t-1)}}(\mathrm{P},\mathrm{Q}) = \mathrm{S}_f(\mathrm{P},\mathrm{Q}) - m\Psi(\mathrm{P},\mathrm{Q})$$

$$\Rightarrow \qquad 0\le \mathrm{S}_f(\mathrm{P},\mathrm{Q}) - m\Psi(\mathrm{P},\mathrm{Q}) \tag{4.5}$$

from where the first inequality of (4.3) results.

Now we again Define a mapping $\mathrm{F}_\mathrm{M}:(0,\infty)\to\mathbb{R},\ \mathrm{F}_\mathrm{M}(t)=\mathrm{M}\,\dfrac{t(t-1)^2}{(2t-1)}-f(t)$,

which is obviously normalized, twice differentiable and by (4.1), convex on (r,R). Applying non-negativity property of f-divergence functional for $\mathrm{F}_\mathrm{M}(.)$ and the linearity property, we obtain the second part of (3.3) i.e.

$$0\le \mathrm{M}\,\Psi(\mathrm{P},\mathrm{Q}) - \mathrm{S}_f(\mathrm{P},\mathrm{Q}) \tag{4.6}$$

from (4.5) and (4.6) give the result (4.3)

Remark 1. If we have strict inequality ">" in (4.3) for any $t \in (r, R)$then the mapping $F_m(.)$ and $F_M(.)$ are strictly convex and equality holds in (4.3) iff P = Q.

Remark 2. It is important note that f is twice differentiable on $(0, \infty)$ and $m \le \frac{2t(4t^2 - 6t + 3)}{(2t-1)^3} f''(t) \le M < \infty, \forall t \in (0, \infty),$ then inequality (4.1) holds for any probability distributions P, Q.

5. RESISTOR-AVERAGE DISTANCE

Here we use the Resistor-Average distance as a measure of dissimilarity between two probability densities on new f-divergence measure which is defined as

$$D_{RAD}(S, D) = [D_\psi(S, D)^{-1} + D_\psi(S, D)^{-1}]^{-1}$$

Symmetric divergence measure from which is derived, it is non-negative and equal to zero iff $p(x) \equiv q(x)$, but unlike it, it is symmetric. Another important property of the Resistor-Average distance is that when two classes of patterns C_p and C_q are distributed according $p(x)$ and $q(x)$, To see in what manner RAD differs from the symmetric Chi-square divergence, it is instructive to consider two special cases: when divergences in both directions between two pdfs are approximately equal and when one of them is much greater than the other:

$$*K_\psi(S, D) \approx K_\psi(S, D) \approx K$$
$$K_{RAD}(S, D) \approx D$$
$$*K_\psi(S, D) \approx K_\psi(S, D) \approx K$$
$$K_\psi(S, D) \approx K_\psi(D, S) \text{ or } K_\psi(S, D) \approx K_\psi(D, S)$$
$$K_{RAD}(S, D) \approx \min K_\psi(S, D) \text{ or } K_\psi(D, S)$$

6. Some Particular Cases

Using equation (4.3) of Theorem 4, we shall be able to point out the following particular cases which are may be interested in Information Theory.

Proposition 6.1: Let P, Q $\in \Gamma_n$ be two probability distribution with the property that

$$r < \frac{1}{2} \le r_i = \frac{p_i + q_i}{2q_i} \le R, \forall i \in \{1, 2, ..., n\}$$

Then we have the following inequalities

$$\frac{2(4R^2 - 6R + 3)}{R(2R-1)^3} \psi(P, Q) \le F(Q, P) \le \frac{2(4r^2 - 6r + 3)}{r(2r-1)^3} \psi(P, Q) \quad (6.1)$$

Proof: Consider the mapping $f : (r, R) \to \mathbb{R}$.

$$f(t) = -\log t, f'(t) = -\frac{1}{t}, f''(t) = \frac{1}{t^2} > 0, \forall t > 0$$

$f''(t) \ge 0$ and $f(1) = 0$, So function f is convex and normalized.

Define, $g(t) = \frac{2t(4t^2 - 6t + 3)}{(2t-1)^3} f''(t) = \frac{2t(4t^2 - 6t + 3)}{(2t-1)^3}\left(\frac{1}{t^2}\right) = \frac{2(4t^2 - 6t + 3)}{t(2t-1)^3}$

$$g(t) = \frac{2(4t^2 - 6t + 3)}{t(2t-1)^3}$$

Then obviously

$$m = \sup_{t \in [r, R]} g(t) = \frac{2(4R^2 - 6R + 3)}{R(2R-1)^3}, \ M = \inf_{t \in [r, R]} g(t) = \frac{2(4r^2 - 6r + 3)}{r(2r-1)^3} \quad (6.2)$$

Also then

$$S_f(P, Q) = \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right) = F(Q, P)$$

Prove of the result (6.1)

7. Numerical Illustration

Let P be the binomial probability distribution for the random valuable X with parameter ($n = 8$, $p = 0.5$) and Q its approximated normal probability distribution. The following table has also discussed [8].

Table 1. Binomial Probability Distribution ($n = 8$, $p = 0.5$)

x	0	1	2	3	4	5
$p(x)$	0.004	0.031	0.109	0.219	0.274	0.219
$q(x)$	0.005	0.030	0.104	0.220	0.282	0.220
$p(x)/q(x)$	0.774	1.042	1.0503	0.997	0.968	0.997

It is noted that $r = 0.77$ and $R = 1.05$. Here we shall discuss the numerical bounds of new information divergence measure in terms of symmetric chi-square divergence measure. From equation (6.1) and using the table of Binomial distribution where R and r are the lower and upper bounds then we get

$$m = \sup_{t \in [r, R]} g(t) = \frac{2(4(1.05)^2 - 6(1.05) + 3))}{(1.05)((2.10 - 1)^3)},$$

$$M = \inf_{t \in [r, R]} g(t) = \frac{2(4(0.77)^2 - 6(0.77) + 3))}{(0.77)(2(0.77 - 1)^3)}$$

$$m = \sup_{t \in [r, R]} g(t) = \frac{2(4(1.05)^2 - 6(1.05) + 3))}{(1.05)((2.10 - 1)^3)},$$

$$M = \inf_{t \in [r, R]} g(t) = \frac{2(4(0.77)^2 - 6(0.77) + 3))}{(0.77)(2(0.77) - 1)^3}$$

$$m = \frac{2.22}{1.39755} = 1.58, \ M = \frac{1.5032}{0.1212} = 12.40$$

$(1.58)\ \psi(P, Q) \leq F(Q, P) \leq (12.40)\ \psi(P, Q)$

References

1. Arandjelovi´cOgnjen, Gregory Shakhnarovich, John Fisher, Roberto Cipolla, Trevor Darrell, "Face Recognition with Image Sets Using Manifold Density Divergence, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).
2. Dragomir S.S., V. Gluscevic and C.E.M. Pearce, "Approximation for the Csiszar f –divergence via mid-point inequalities, in inequality theory and applications- Y. J. Cho, J. K. Kim and S.S. dragomir (Eds), nova science publishers, inc., Huntington, new York, vol1, 2001, pp.139-154.
3. Jain K.C. and R.N. Saraswat "Some New Information Inequalities and its Applications in Information Theory" International Journal of Mathematics Research, Volume 4, Number 3 (2012), pp. 295-307.
4. Jain K.C. and R.N. Saraswat "Inequalities among Generalized Information Measures of Type s and its Applications in Information Theory, Journal of Combinatorics, Information & System Sciences 38,(1-2), 69-88, 2013.
5. Jain K.C. and R.N. Saraswat "A New Information Inequality and its Application in Establishing Relation among various f-Divergence Measures", Journal of Applied Mathematics, Statistics and Informatics, 8(1) (2012), 17-32.
6. Jain K.C. and R.N. Saraswat "Some bounds of information divergence measure in term of Relative arithmetic-geometric divergence measure" International Journal of Applied Mathematics and Statistics, Volume 32, Number 2 (2013), pp. 48-58.
7. Pearson K., "On the criterion that a give system of deviations from the probable in the case of correlated system of variables in such that it can be reasonable supposed to have arisen from random sampling", Phil. Mag., 50(1900),157-172.
8. Sibson R., "Information Radius", Z, Wahrs. undverw.geb. (14) (1969), 149-160.
9. Taneja I.J. and Pranesh Kumar, "On symmetric and non-symmetric divergence measure and their generalizations", Chapter in: Advances in Imaging and Electron Physics, Ed. P.W. Hawkes, 91 (1995), 37-136.

Ram Naresh Saraswat: Department of Mathematics and Statistics, Manipal Univeristy, Jaipur

E-mail address: **sarswatrn@gmail.com**

Neha Khatod: VPO-Dehmikala Jaipur-Ajmer Express Highway, Jaipur, Raj-303007, India