THEORETIC BASED EXPONENTIAL INFORMATION DIVERGENCE MEASURE AND INEQUALITIES

K.C.Jain[1] and Devendra Singh[2]
[1]Department of Mathematics, IIIT KOTA, JLN Marg, MNIT Campus, Jaipur-302017, Rajasthan, India
[2]Department of Mathematics, Malaviya National Institute of Technology, Jaipur-302017, Rajasthan, India
Email: devendra.mnit2013@gmail.com , 2013rma9061@mnit.ac.in

Abstract. A non-parametric theoretic based exponential information divergence measure is proposed. This measure belongs to the category of Csiszár's $f-$divergences. Further for first time, we have derived some inequalities for this exponential information divergence measure in terms of some valuable information divergence measures. Some numerical illustrations are carried out, based on two distinct discrete probability distributions.

Keywords: Non-parametric divergence measure, Symmetric Chi-Square divergence Measure, Csiszár's *f*-divergence, Arithmetic mean divergence, Geometric mean divergence, Harmonic mean divergence.
Mathematics Subject Classification 2010: 94A17, 26D15.

1. INTRODUCTION

Let $\Gamma_n = \{P = (p_1, p_2, p_3, ... p_n);\ p_i > 0, \sum_{i=1}^{n} p_i = 1\}$, n≥2, be the set of all complete finite discrete probability distributions. Csiszár [2, 3] introduced a generalized measure of information using *f*-divergence measure

$$C_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \quad (1.1)$$

Where $f:(0,\infty) \to$R (set of real numbers) is a convex function. Most common choices of function *f* satisfy $f(1) = 0$, so that $C_f(P,Q) = 0$. Convexity of function *f* ensure that divergence measure $C_f(P,Q)$ is nonnegative. An important characteristic of this divergence measure is that many known divergences can be obtained from this measure by appropriately defining the convex function *f*. [Shannon (1958), Renyi (1961), Ali & Silvey (1966), Vajda (1972), Berbea & Rao (1982a, b), Taneja (1995), Kumar & Chhina (2005), Kumar & Johnson (2005) and many more].

Further, these information divergence measures are used to find out distance or affinity between two probability distributions. Non parametric divergence measures give the amount of information supplied by the data for discriminating in favor of a probability distribution P $=\{p_1, p_2, p_3, \dots p_n\}$ against another Q = $\{q_1, q_2, q_3, \dots q_n\}$,where P, Q $\in \Gamma_n$ .The construction of information divergence measure for two distinct probability distributions is not an easy task.

In this research work, we are introducing a new theoretic based non-parametric exponential information divergence measure which fits to the category of Csiszár's $f-$divergences [2, 3].

In unit 2; we discuss some advantageous inequalities. New exponential information divergence measure is achieved in unit 3. In unit 4, we have resulting some information inequalities for the new exponential information divergence measure in terms of some recognized and valued divergence measures. Some numerical illustrations of new exponential information measure are shown in unit 5. Unit 6 concludes the paper.

For shortness, we will denote p_i, q_i and $\sum_{i=1}^{n}$ by p, q and $\sum$ respectively.

During past years P. Kumar and other [4, 5] has contributed a lot of work providing different kinds of information, bounds on the distance and divergence measures. His existing information divergence measures are as follows.

P. Kumar and S. Chhina [5]

$$f(t)=\frac{(t+1)(t-1)^2}{t}\ln\left(\frac{t+1}{2\sqrt{t}}\right) \tag{1.2}$$

Then

$$S(P,Q)=\sum\frac{(p+q)(p-q)^2}{(pq)}\ln\left(\frac{p+q}{2\sqrt{pq}}\right) \tag{1.3}$$

And

$$S_\rho(P,Q)=\sum\left(\frac{(p-q)^2}{2p^2q}\right)\left[2\left(2p^2+pq+q^2\right)\ln\left(\frac{p+q}{2\sqrt{pq}}\right)+(p^2-2pq+q)\right] \tag{1.4}$$

P. Kumar and A. Johnson [4]

$$f(t)=\frac{(t^2-1)^2}{2t^{3/2}} \tag{1.5}$$

Then,

$$\Psi M(P,Q)=\sum\frac{(p^2-q^2)^2}{2(pq)^{3/2}} \tag{1.6}$$

$$\Psi M_\rho(P,Q)=\sum\frac{(p-q)(p^2-q^2)(5p^2+3q^2)}{4p^{5/2}q^{3/2}} \tag{1.7}$$

2. WELL KNOW INEQUALITIES

In this section, we give some well-known inequalities which are established in literature of pure and applied mathematics. Using following inequalities we have derived important bounds of well-known divergence measures

$$\frac{x}{1+x}\le \ln(1+x)\le x, \quad x>0 \tag{2.1}$$

$$x-\frac{x^2}{2}\le \ln(1+x)\le x-\frac{x^2}{2(1+x)}, \quad x>0 \tag{2.2}$$

$$1+x\le \exp\{x\}, \quad x>0 \tag{2.3}$$

$$x < \exp\{x\}, \quad x>0 \tag{2.4}$$

3. NEW EXPONENCIAL INFORMATION DIVERGENCE MEASURE

Now, we consider the function $f:(0,\infty)\to\mathbb{R}$ given by

$$f(t)=\frac{(t^2-1)^2}{2t^{3/2}}\exp\{t\} \tag{3.1}$$

And thus the information divergence measure:

$$\Phi D(P,Q)=\sum\frac{(p^2-q^2)^2}{2(pq)^{3/2}}\exp\left\{p/q\right\} \tag{3.2}$$

Next,

$$f'(t)=\frac{1}{4}(t^2-1)\left(\frac{2t^3+5t^2-2t+3}{t^{5/2}}\right)\exp\{t\} \tag{3.3}$$

And

$$f''(t) = \frac{1}{8}\,\frac{\left\lfloor 20t^5 + 7t^4 + 4\left(t^3-1\right)^2 + 6\left(t-1\right)^2 + 5 \right\rfloor}{t^{7/2}}\,\exp\{t\} \qquad (3.4)$$

The function f (t) is convex since $f''(t) > 0$ for all $t > 0$ and normalized also since f (1) = 0.

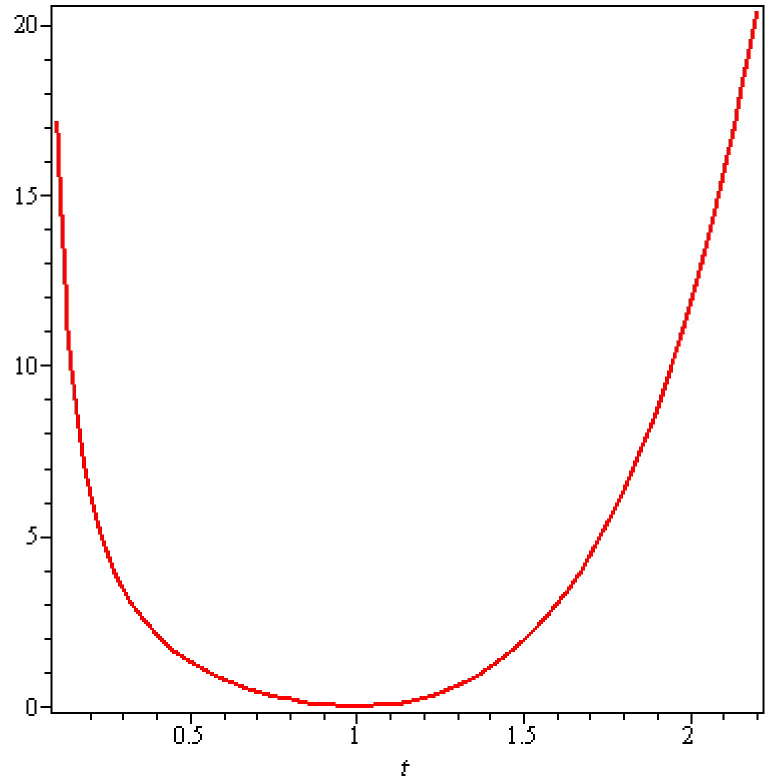

Figure 1: Graph of the convex function f (t).

Figure 1, Shows the behavior of the function $f(t)$ and which is continuously convex. Thus the measure is nonnegative and convex in the pair of discrete probability distributions $(P,Q) \in \Gamma_n$.

4. BOUNDS FOR $\Phi D\,(P,Q)$

We now develop information inequalities providing bounds for $\Phi D\,(P,Q)$ in terms of the recognized information divergence measures in the following propositions.

Proposition 4.1: Let $\Phi D\,(P,Q)$ and $\Psi M\,(P,Q)$ be defined as (3.2) and (1.6) respectively and the symmetric χ^2-divergence

$$\Psi\left(P,Q\right)=\chi^{2}\left(P,Q\right)+\chi^{2}\left(Q,P\right)=\sum\frac{\left(p+q\right)\left(p-q\right)^{2}}{pq} \tag{4.1}$$

Then inequality

$$\Psi\left(P,Q\right)\leq\Psi M\left(P,Q\right)\leq\Phi D\left(P,Q\right), \tag{4.2}$$

Holds and equality, *iff* $P = Q$.

Proof: Consider the Harmonic (HM) and Geometric mean (GM) inequality,

$$\text{HM} \leq GM$$

Or,
$$\frac{2pq}{p+q}\leq\sqrt{pq},$$

or,
$$\frac{p+q}{2\sqrt{pq}}\leq\left(\frac{p+q}{2\sqrt{pq}}\right)^{2},$$

or,
$$\frac{\left(p+q\right)\left(p-q\right)^{2}}{pq}\leq\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}, \tag{4.3}$$

Further, for $x > 0$,

$exp\{x\} > 1$,

Thus for, $0 < p \leq 1$ *and* $0 < q \leq 1$, $\frac{p}{q} > 0$,

$$1<\exp\left\{p/q\right\},$$

Or,
$$\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}\leq\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}\exp\left\{p/q\right\} \tag{4.4}$$

Now, from (4.3) and (4.4), we get

$$\frac{\left(p+q\right)\left(p-q\right)^{2}}{pq}\leq\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}\leq\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}\exp\left\{p/q\right\}$$

Summing over all terms we get,

$$\sum\frac{\left(p+q\right)\left(p-q\right)^{2}}{pq}\leq\sum\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}\leq\sum\frac{\left(p^{2}-q^{2}\right)^{2}}{2\left(pq\right)^{3/2}}\exp\left\{p/q\right\}$$

Therefore, we get

$$\Psi\left(P,Q\right)\leq\Psi M\left(P,Q\right)\leq\Phi D\left(P,Q\right),$$

Hence, the inequality.

Proposition 4.2: Let $S\,(P,Q)\,and\,\Phi D\,(P,Q)$ be defined as (1.3) and (3.2), respectively. Then inequality

$$S\,(P,Q)\le \Phi D\,(P,Q) \tag{4.5}$$

And equality holds for $P=Q$

Proof: From inequality (2.1) and (2.4), we get,

$$\ln(1+x)\le x<\exp\{x\},\qquad x>0$$

$$\Rightarrow\qquad \ln\left(\frac{1+x}{2\sqrt{x}}\right)<\exp\{x\},\qquad x>0$$

$$\Rightarrow\qquad \frac{(x^2-1)^2}{2\,x^{3/2}}\ln\left(\frac{1+x}{2\,\sqrt{x}}\right)<\frac{(x^2-1)^2}{2\,x^{3/2}}\exp\{x\} \tag{4.6}$$

Replacing, x by p/q, we get,

$$\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}}\ln\left(\frac{p+q}{2\sqrt{pq}}\right)\le\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}}\exp\left\{p/q\right\} \tag{4.7}$$

Or, $$\left(\frac{p+q}{2\,\sqrt{pq}}\right)\ \frac{(p+q)\,(p-q)^2}{pq}\ln\left(\frac{p+q}{2\sqrt{pq}}\right)\le\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}}\exp\left\{p/q\right\}$$

We have, $\frac{2\,\sqrt{pq}}{p+q}\le 1\Rightarrow\frac{p+q}{2\,\sqrt{pq}}\ge 1$ and thus,

$$\frac{(p+q)\,(p-q)^2}{pq}\ln\left(\frac{p+q}{2\sqrt{pq}}\right)\le\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}}\exp\left\{p/q\right\}$$

Summing over both sides, we get,

$$\sum\frac{(p+q)\,(p-q)^2}{pq}\ln\left(\frac{p+q}{2\sqrt{pq}}\right)\le\sum\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}}\exp\left\{p/q\right\}$$

Hence, the required inequality,

$$S(P,Q) \leq \Phi D(P,Q).$$

Proposition 4.3: Let $\Phi D(P,Q)$, $\Psi M(P,Q)$ *and* $\Psi(P,Q)$ be defined as (3.2), (1.6) and (4.1), respectively. Then inequality

$$2\,\Psi(P,Q) - \frac{3}{2}\,\Psi M(P,Q) - \sum \frac{(p-q)^2}{\sqrt{pq}} \leq \Phi D(P,Q) \tag{4.8}$$

and equality holds for $P = Q$.

Proof: From inequality (4.6), we have

$$\frac{(x^2-1)^2}{2\,x^{3/2}} \ln\left(\frac{1+x}{2\sqrt{x}}\right) < \frac{(x^2-1)^2}{2\,x^{3/2}} \exp\{x\}$$

Or,
$$\frac{(x^2-1)^2}{2\,x^{3/2}} \ln\left(\frac{2\sqrt{x}}{1+x}\right) < \frac{(x^2-1)^2}{2\,x^{3/2}} \exp\{x\}$$

Replacing, x by p/q, we get,

$$\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \ln\left(\frac{2\sqrt{pq}}{p+q}\right) \leq \frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \exp\left\{p/q\right\} \tag{4.9}$$

Further, we know that

$$\ln\left(\frac{2\sqrt{pq}}{p+q}\right) \approx \frac{4\sqrt{pq}}{p+q} - \frac{2pq}{(p+q)^2} - \frac{3}{2} \tag{4.10}$$

Now, from (4.9) and (4.10), we get

$$\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \left(\frac{4\sqrt{pq}}{p+q} - \frac{2pq}{(p+q)^2} - \frac{3}{2}\right) \leq \frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \exp\left\{p/q\right\}$$

Arranging in appropriate forms and summing over all terms, we get

$$2\sum\frac{(p+q)(p-q)^2}{pq}-\frac{3}{2}\sum\frac{(p^2-q^2)^2}{2(pq)^{3/2}}-\sum\frac{(p-q)^2}{\sqrt{pq}}\le\sum\frac{(p^2-q^2)^2}{2(pq)^{3/2}}\exp\left\{p/q\right\}$$

Using (4.1), (1.6), and (3.2), we get

$$2\,\Psi\,(P,Q)-\frac{3}{2}\,\Psi M\,(P,Q)-\sum\frac{(p-q)^2}{\sqrt{pq}}\le\Phi D\,(P,Q)$$

Proposition 4.4: Let $\Phi D\,(P,Q)$, $\Psi M\,(P,Q)$ *and* $\Psi M_\rho(P,Q)$ be defined as (3.2), (1.6) and (1.7), respectively. Then inequality

$$\Psi M_\rho(P,Q)-\Psi M\,(P,Q)\le\Phi D_\rho(P,Q)-\Phi D\,(P,Q) \tag{4.11}$$

Holds and equality i*ff* $P=Q$.

Where

$$\Phi D_\rho(P,Q)=\frac{1}{4}\frac{(p-q)(p^2-q^2)(2p^3+5p^2q-2pq^2+3q^3)}{(pq)^{5/2}}\exp\left\{p/q\right\} \tag{4.12}$$

Proof: From (3.3), we have

$$f'(t)=\frac{1}{4}(t^2-1)\left(\frac{2t^3+5t^2-2t+3}{t^{5/2}}\right)\exp\{t\}$$

And, thus

$$\Phi D_\rho(P,Q)=\sum(p-q)f'\left(\frac{p}{q}\right)$$

$$=\frac{1}{4}\frac{(p-q)(p^2-q^2)(2p^3+5p^2q-2pq^2+3q^3)}{(pq)^{5/2}}\exp\left\{p/q\right\} \tag{4.13}$$

Further, from (3.2) and (4.6), we get

$$\Phi D_\rho(P,Q)-\Phi D\,(P,Q)=\frac{1}{4}\frac{(p-q)(p^2-q^2)(2p^3+5p^2q-2pq^2+3q^3)}{(pq)^{5/2}}\exp\left\{p/q\right\} \tag{4.14}$$

From (1.6) and (1.7), we get

$$\Psi M_{\rho}(P,Q)-\Psi M\,(P,Q)=\frac{1}{4}\frac{(p-q)\left(p^{2}-q^{2}\right)q\left(3p^{2}-2pq+3q^{2}\right)}{(pq)^{5/2}} \tag{4.15}$$

From (4.14) and (4.15), we get inequality (4.11).

5. NUMERICAL ILLUSTRATION

In this section, we consider an example of symmetrical probability distributions. We will numerically verify the bounds achieved in the earlier section. For this, we calculate measures $\Phi D\,(P,Q)$, $\Psi M\,(P,Q)$, $S(P,Q)$, *and* $\Psi\,(P,Q)$.

Let P be the binomial probability distribution for the random variable X with parameter (n = 8, p = 0.5) and Q its approximated normal probability distribution.

Table 5.1 Binomial Probability Distribution (n = 8, p = 0.5)

X	0	1	2	3	4	5	6	7	8
$p(x)$	0.0040	0.0310	0.1090	0.2190	0.2740	0.2190	0.1090	0.0310	0.0040
$q(x)$	0.0050	0.0300	0.1040	0.2200	0.2820	0.2200	0.1040	0.0300	0.0050
$\frac{p(x)}{q(x)}$	0.8000	1.0333	1.0481	0.9955	0.9716	0.9955	1.0481	1.0333	0.8000

The divergence measures $\Phi D\,(P,Q)$, $\Psi M\,(P,Q)$, $S(P,Q)$, *and* $\Psi\,(P,Q)$ *are*:

$$\Phi D\,(P,Q)=0.08103214, \qquad \Psi M\,(P,Q)=0.00306097$$

$$S(P,Q)=0.00001030, \qquad \Psi\,(P,Q)=0.00305063$$

It is noted that

$$0.8000 \le \left(\frac{p}{q}\right) \le 1.0481$$

These numerical values of measures verified the propositions of previous section.

6. CONCLUSION

During past years Dragomir [20-25], Teneja [11-13 and 19], Kumar and others [4, 5] gave the idea of divergence measures, their properties, bounds and relations with other measures. Kumar and other did a lot of work especially in the field of information theory. In [4, 5], he derived new bounds in terms of different symmetric and nonsymmetric divergence measures. We have introduced a new exponential nonparametric divergence measure, in the Csiszár's *f*-divergence category [2, 3], by considering a convex function *f*, defined on $(0, \infty)$. This paper also defines the bounds and properties of new exponential information measure with the work of Kumar's divergence measures and some other well-known measures.

REFERENCES

1. S.M. Ali and S.D. Silvey, A general class of coefficients of divergence of one distribution from another, Jour. Roy. Statist. Soc. B (28), 131-142, (1966).
2. I. Csiszár's, Information-type measure of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar. 2, 299-318, (1967).
3. I. Csiszár's, Information measures: A critical survey, In Trans. Seventh Prague Conf. On Information Theory, A, pp.73-86, Academia, Prague (1974).
4. P. Kumar and A. Johnson, on a symmetric divergence measure and information inequalities, J. Inequalities in Pure Appl.Math. 6(3) 1-13, (2005).
5. P. Kumar and S. Chhina, A symmetric information divergence measure of the Csiszár's *f*-divergence class and its bounds. An I.J. computers & mathematics with applications, 49, 575-588(2005).
6. C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. Jour. 27, 623-659, (1958).
7. R. Sibson, Information radius, Z. Wahrs. undverw Geb. 14,149-160, (1969).
8. K. Pearson ,On the Criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonable supposed to have arisen from random sampling, Phil. Mag., 50, 157-172, (1900) .
9. H. Jeffreys, An invariant form for the prior probability in estimation problems, Proc. Roy. Soc. London Ser. A 186, 453-461, (1946).
10. P.Kafka , F. Osterreicher and I.Vincze, On powers of f-divergence defining a distance, Studia Sci.Math.Hungar.26, 415-422 (1991).
11. I.J. Taneja,Inequalities among logarithmic mean measures,2011. Available online: http://arxiv.org/abs/1103.2580v1.
12. I.J. Taneja, New developments in generalized information measures, Chapter in: Advances in Imaging and Electron Physics (Edited by P.W. Hawkes) 91, 37-135, (1995).

13. I.J. Taneja, Generalized symmetric divergence measures and Inequalities, RGMIA Research Report Collection, http://rgmia.vu.edu.au, 7(4) (2004), Art. 9. Available online: arxiv:math.ST/0501301 v1 19 Jan 2005.
14. L. Le Cam, Asymptotic Methods in Statistical Decision Theory, Springer-Verlag, New York, (1986).
15. F. *Topsφe* , 85-98, Res. Rep. Collection: RGMIA 2(1), (1999).
16. A. Bhattacharyya, On some analogues to the amount of information and their uses in statistical estimation, *Sankhya* 8, 1-14, (1946).
17. F. Liese and I. Vajda, *Teubner−Texte zur Mathematik* 95, Leipzig, (1987).
18. E. Hellinger, Neue Berudung der Theorie der quadratischen Formen von Un endlichen Vielen Veran derlichen, J.Reine Aug.Math. 136, 210-271 (1909).
19. I.J. Taneja and P. Kumar, Relative information of $type-s$, Csiszár's $f-$divergence and information inequalities, *Information Sciences*, (2003).
20. S.S. Dragomir, Some inequalities for the Csiszar's ϕ- divergence, Inequalities for the Csiszar's *f*-divergence in Information Theory, (Edited by S.S. Dragomir), http://rgmia.vu.edu.au/monographs/csiszar.html, (2000).
21. S.S. Dragomir, A converse inequality for the Csiszar's ϕ- divergence, Inequalities for the Csiszar's *f*-divergence in Information Theory, (Edited by S.S. Dragomir), http://rgmia.vu.edu.au/monographs/csiszar.html, (2ooo).
22. S.S. Dragomir, Some inequalities for (m, M)-convex mappings and applications for the Csiszar's ϕ-divergence in information theory, Inequalities for the Csiszar's *f*-divergence in Information Theory, (Edited by S.S. Dragomir), http://rgmia.vu.edu.au/monographs/csiszar.html, (2000).
23. S.S. Dragomir, Upper and lower bounds for Csiszar's *f*-divergence in terms of the Kullback-Leibler distance and applications, Inequalities for the Csiszar's *f*-divergence in Information Theory, (Editedby S.S. Dragomir), http://rgmia.vu.edu.au/monographs/csiszar.html, (2000).
24. S.S. Dragomir,Upper and lower bounds for Csiszar's *f*-divergence in terms of the Hellinger discrimination and applications, Inequalities for the Csiszar's *f*-divergence in Information Theory, (Edited by S.S. Dragomir), http://rgmia.vu.edu.au/monographs/csiszar.html, (2000).
25. S.S.Dragomir, V.Gluscevic, and C.E.M.Pearce, Inequality theory and Applications (Edited by Y.J.Cho, J.K. Kim and S.S. Dragomir), Nova Science Publishers, Huntington, New York, (2001).