Journal of Basic and Applied Engineering Research
Print ISSN: 2350-0077; Online ISSN: 2350-0255; Volume 2, Number 13; April-June, 2015 pp. 1135-1138

http://www.krishisanskriti.org/jbaer.html

Bounds on Information Divergence Measures in Terms of Hellinger Discrimination

Ram Naresh Saraswat[1] and Ajay Tak[2]

[1,2]*Department of Mathematics & Statistics, School of Basic Sciences, Manipal University Jaipur, 302026, INDIA*
E-mail: [1]sarswatrn@gmail.com, [2]ajaytak86@gmail.com

Abstract—*There are many information and divergence measures are exist in the literature of Information Theory and statistics. These are very useful and play an important role in many areas like as sensor networks,testing the order in a Markov chain, risk for binary experiments, region segmentation and estimation etc. In this research paper, we shall study bounds on well-known information divergence measures in terms of Hellinger discrimination using information inequalities, convex functions and new f-divergence measures.*

Keywords: *New f-divergence ,Hellinger's discrimination, Relative Jensen-Shannon divergence measure etc.*

1. INTRODUCTION

Let

$$\Gamma_n = \{ p = (p_1, p_2, p_3, \ldots\ldots\ldots p_n) / p_1 \geq 0, \sum_{i=1}^{n} p_i = 1\}, n \geq 2 \quad (1.1)$$

be the set of complete finite discrete probability distributions. There are many information and divergence measures exist in the literature on information theory and statistics. In this section we present some properties of new f-divergence measure introduced in Jain &Saraswat [3] & [4]and its particular cases which are interesting in areas of information theory and statistics is given by

$$S_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i + q_i}{2q_i}\right) \quad (1.2)$$

Where $f : R_+ \to R_+$ is a convex function and P, Q ϵ Γ_n.

Proposition 1.1 Let $f:[0,\infty) \to R$ be convex and $P,Q \in \Gamma_n$ with $P_n = Q_n = 1$ then we have the following inequality

$$S_f(P,Q) \geq f(1) \quad (1.3)$$

Equality holds in (1.3)

$$\text{iff } p_i = q_i, \forall i = 1,2,\ldots,n \quad (1.4)$$

Corollary 1.1.1 (Non-Negativity of New f-divergence measure) Let $f:[0,\infty) \to R$ be convex and normalized i.e.

$$f(1)=0 \quad (1.5)$$

Then for any $P,Q \in \Gamma_n$ from (1.3) of proposition 1.1 and (1.5), we have the inequality

$$S_f(P,Q) \geq 0 \quad (1.6)$$

If *f* is strictly convex, equality holds in (1.6)iff

$$p_i = q_i, \forall i = 1,2,\ldots,n \quad (1.7)$$

and

$$S_f(P,Q) = 0 \text{ iff } P = Q \quad (1.8)$$

Proposition 1.2 Let f_1 and 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 $P,Q \in \Gamma_n$.

It is shown that using new f-divergence measure we have derived some well-known divergence measures like as, Hellinger discrimination [5], Kullback-Leibler divergence [7] & [8], Relative Jensen-Shannon divergence [6], Relative arithmetic-geometric divergence measure [9]. We now give some examples of well-known information divergence measures which are obtained from New f-divergence measure.

- If $f(t) = -\log t$ then relative Jensen-Shannon divergence measure is given by
$$S_f(P,Q) = \sum_{i=1}^{n} q_i \log(2q_i / (p_i + q_i)) = F(Q,P) \quad (1.9)$$
- If $f(t) = t\log t$ then relative arithmetic-geometric divergence measure is given by
$$S_f(P,Q) = \sum_{i=1}^{n} ((p_i + q_i)/2)\log((p_i + q_i)/2q_i) = G(Q,P) \quad (1.10)$$

- If $f(t) = (2t-1)\log(2t-1), t > \frac{1}{2}$ thenKullback-Leibler divergence measure is givenby

$$S_f(P,Q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) = KL(P,Q) \quad (1.11)$$

- If $f(t) = -\log(2t-1), \forall t > \frac{1}{2}$ then Kullback-Leibler divergence measure is given by

$$S_f(P,Q) = \sum_{i=1}^{n} q_i \log\left(\frac{q_i}{p_i}\right) = KL(Q,P) \quad (1.12)$$

2. HELINGER DISCRIMINATION

Let us considera function

$$f(t) = \frac{1}{2}\left(\sqrt{2t-1} - 1\right)^2 \quad (2.1)$$

$$f'(t) = \frac{\sqrt{2t-1}-1}{\sqrt{2t-1}}, \; f''(t) = \frac{1}{(2t-1)^{3/2}} > 0, \forall t > \frac{1}{2}.$$

Hence function is convex and normalized i.e. f(1)=0, It's second derivative is positive. So function is convex

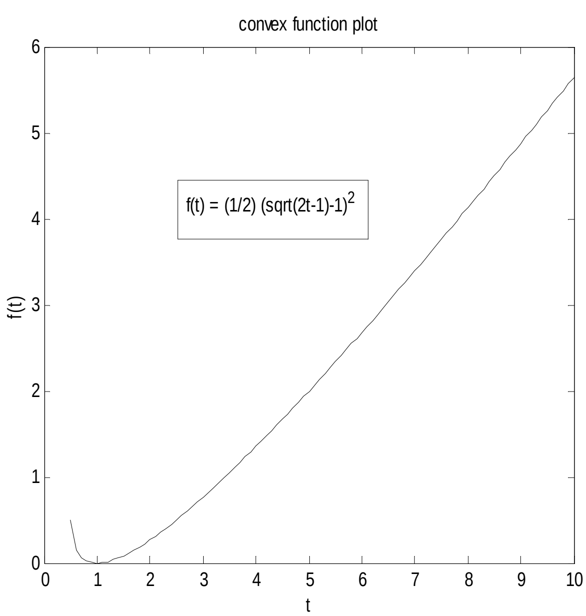

Fig. 2.1: (Graph of f(t))

Consider a function $f(t) = \frac{1}{2}\left(\sqrt{2t-1} - 1\right)^2$ then Hellinger discrimination is given by

$$\sum_{i=1}^{n}\left[q_i \frac{1}{2}\left\{\sqrt{\left(2\left(\frac{p_i+q_i}{2q_i}\right)-1\right)} - 1\right\}^2\right] = \frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{p_i} - \sqrt{q_i}\right)^2 = h(P,Q)$$

Then we can say

$$S_f(P,Q) = h(P,Q) \quad (2.2)$$

3. NEW INFORMATION INEQUALITY

The following theorem concerning an upper and lower bound for a new f-divergence measure in terms of the Hellinger discrimination holds.

Results are similar to presented by Dragomir [1] and Jain and Saraswat[2& 4].

Theorem 3.1: Let suppose that generating mapping $f : \left(\frac{1}{2}, \infty\right) \to R$ is normalized that is *f*(1)=0 and satisfies the assumptions.

(i) f is twice differentiable on (r, R), where $0.5 \le r \le 1 \le R \le \infty$

(ii) there exist constants m, M such that

$m \le [(2t-1)^{\frac{3}{2}} f''(t)] \le M$ for all $t \in (r, R)$. (3.1)

If P, Q are discrete probability distributions satisfying the assumptions $r \le r_i := \frac{(p_i+q_i)}{2q_i} \le R, \forall i \in \{1,2,3.........n\},$ (3.2)

Then we have the Inequality

$$mh(P,Q) \le S_f(P,Q) \le Mh(P,Q) \quad (3.3)$$

Proof: Define a mapping

$$F_m : (0.5, \infty) \to R, F_m(t) = f(t) - \frac{m}{2}\left(\sqrt{2t-1} - 1\right)^2.$$

Then $F_m(.)$ is normalized and twice differentiable and since

$$F_m''(t) = f''(t) - \frac{m}{(2t-1)^{\frac{3}{2}}} = \frac{1}{(2t-1)^{\frac{3}{2}}}\left[(2t-1)^{\frac{3}{2}} f''(t) - m\right] \ge 0 \quad (3.4)$$

For all t∈(a, b), implied by the first inequality in (3.1). It follows that the mapping $F_m(.)$ is convex on (r, R). Applying the non-negativity property of New *f*-divergence functional for

$F_m(.)$ and the linearity,(by proposition 1.2), we may state that

$$0 \le S_{F_m}(P,Q) = S_f(P,Q) - mS_{\frac{1}{2}\left(\sqrt{2t-1}-1\right)^2}(P,Q)$$
$$= S_f(P,Q) - mh(P,Q) \qquad (3.5)$$

From where we get the first inequality in (3.3).

Now we again Define a mapping,

$$F_M : (0.5,\infty) \to R, F_M(t) = f(t) - \frac{M}{2}\left(\sqrt{2t-1}-1\right)^2$$

which is obviously normalized, twice differentiable and, by (3.1), convex on (r, R). Applying non-negativity property of New *f*-divergence for $F_M(.)$, and Proposition 1.2, we obtain the second part of (3.3) i.e.

$$0 \le Mh(P,Q) - S_f(P,Q) \qquad (3.6)$$

From (3.5) and (3.6) give the result (3.3)

$$mh(P,Q) \le S_f(P,Q) \le Mh(P,Q)$$

4. SOME PARTICULAR CASES

In this section we established bounds of particular well known divergence measures in terms of Hellinger discrimination using inequality of (3.3) of Theorem 3.1 which may be interested in Information Theory and statistics.

The result is on similar lines to the result presented by Dragomir [1] and Jain &Saraswat [2]& [3].

Proposition 4.1: Let $P, Q \in \Gamma_n$ be two probability distributions with the property that

$$r \le r_i := \frac{(p_i + q_i)}{2q_i} \le R, \forall i \in \{1,2,3.........n\}$$

Then we have the following inequality

$$4\sqrt{(2r-1)}h(P,Q) \le KL(P,Q) \le 4\sqrt{(2R-1)}h(P,Q) \quad (4.1)$$

Proof: consider the mapping $f : (0.5,\infty) \to R$, $f : (r,R) \to R$.

$f(t) = (2t-1)\log(2t-1), f'(t) = 2[1+\log(2t-1)], f''(t) = \dfrac{4}{(2t-1)} > 0,$ So function is convex and normalized i.e. $f(1) = 0$.

$$\text{Define } g(t) = \left[(2t-1)^{\frac{3}{2}} f''(t)\right]$$

$$= \left[(2t-1)^{\frac{3}{2}} \frac{4}{(2t-1)}\right] = \left[4(2t-1)^{\frac{1}{2}}\right]$$

Then obviously

$$M = \sup_{t\in[r,R]} g(t) = 4\sqrt{(2R-1)}, m = \inf_{t\in[r,R]} g(t) = 4\sqrt{(2r-1)} \quad (4.2)$$

Also $S_f(P,Q) = KL(P,Q)$ from equation (1.11)

From equation (1.11), (3.3) & (4.2) give the result (4.1).

Proposition 4.2: Let $P, Q \in \Gamma_n$ be two probability distributions satisfying (3.1)

Then we have the following inequality

$$\frac{4}{\sqrt{(2R-1)}}h(P,Q) \le KL(Q,P) \le \frac{4}{\sqrt{(2r-1)}}h(P,Q) \qquad (4.3)$$

Proof: Consider the mapping $f : (0.5,\infty) \to R$, $f : (r,R) \to R$,

$$f(t) = -\log(2t-1), f'(t) = \left(-\frac{2}{2t-1}\right), f''(t) = \left(\frac{4}{(2t-1)^2}\right) > 0,$$

So function is convex and normalized i.e. $f(1) = 0$

$$\text{Define } g(t) = \left[(2t-1)^{\frac{3}{2}} f''(t)\right] = \left[(2t-1)^{\frac{3}{2}} \frac{4}{(2t-1)^2}\right]$$

$$g(t) = \frac{4}{\sqrt{(2t-1)}} > 0, \forall t > \frac{1}{2}$$

Then obviously

$$M = \sup_{t\in[r,R]} g(t) = \frac{4}{\sqrt{(2r-1)}},$$
$$m = \inf_{t\in[r,R]} g(t) = \frac{4}{\sqrt{(2R-1)}} \qquad (4.4)$$

Also $S_f(P,Q) = KL(Q,P)$ from equation (1.12)

From equation (1.12), (3.3) & (4.4) give the result (4.3)

Proposition 4.3: Let $P, Q \in \Gamma_n$ be two probability distributions satisfying (3.1),

Then we have the following inequality

$$\left(2-\frac{1}{r}\right)\sqrt{(2r-1)}h(P,Q) \le G(Q,P) \le \left(2-\frac{1}{R}\right)\sqrt{(2R-1)}h(P,Q) \quad (4.5)$$

Proof: Consider the mapping $f : (r,R) \to R$

$f(t)=t\log t,\ f'(t)=(1+\log t),\ f''(t)=\frac{1}{t}>0,$ So function is convex and normalized i.e. $f(1)=0$

Define $g(t)=\left[(2t-1)^{\frac{3}{2}} f''(t)\right]=\left[(2t-1)^{\frac{3}{2}}\frac{1}{t}\right]$

$$g(t)=\left[\left(2-\frac{1}{t}\right)\sqrt{(2t-1)}\right]>0, \forall t>\frac{1}{2}$$

Then obviously

$$M=\sup_{t\in[r,R]} g(t)=\left(2-\frac{1}{R}\right)\sqrt{(2R-1)}, m=\inf_{t\in[r,R]} g(t)=\left(2-\frac{1}{r}\right)\sqrt{(2r-1)} \quad (4.6)$$

Also $S_f(P,Q)=G(Q,P)$ from equation (1.10)

From equation (1.10), (3.3) & (4.6) give the result (4.5)

Proposition 4.4: Let $P,Q\in\Gamma_n$ be two probability distributions satisfying (3.1),

Then we have the following inequality

$$\left(2-\frac{1}{R}\right)^{\frac{3}{2}}\frac{1}{\sqrt{R}}h(P,Q)\le F(Q,P)\le\left(2-\frac{1}{r}\right)^{\frac{3}{2}}\frac{1}{\sqrt{r}}h(P,Q) \quad (4.7)$$

Proof: Consider the mapping $f:(r,R)\to R$

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

So function is convex and normalized i.e. $f(1)=0$

Define $g(t)=\left[(2t-1)^{\frac{3}{2}} f''(t)\right]=\left[(2t-1)^{\frac{3}{2}}\frac{1}{t^2}\right]$

$$g(t)=\left[\left(2-\frac{1}{t}\right)^{\frac{3}{2}}\frac{1}{\sqrt{t}}\right]>0, \forall t>\frac{1}{2}$$

Then obviously

$$M=\sup_{t\in[r,R]} g(t)=\left(2-\frac{1}{r}\right)^{\frac{3}{2}}\frac{1}{\sqrt{r}}, m=\inf_{t\in[r,R]} g(t)=\left(2-\frac{1}{R}\right)^{\frac{3}{2}}\frac{1}{\sqrt{R}} \quad (4.8)$$

Also $S_f(P,Q)=F(Q,P)$ from equation (1.9)

From equation (1.9), (3.3)& (4.8) give the result (4.7)

REFERENCES

[1] Dragomir S.S., "Some inequalities for (m, M)-convex mappings and applications for the Csiszar's Φ-divergence in information theory". Math. J. Ibaraki Univ. (Japan) 33 (2001), 35-50.

[2] Jain K. C. and R. N. Saraswat , "Some bounds of Csiszar's f-divergence measure in terms of the well-known divergence measures of Information Theory" International Journal of Mathematical Sciences and Engineering Applications,5 (5), 2011.

[3] 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.

[4] 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, 32,(2) (2013), 48-58.

[5] Hellinger E., "NeueBegrundung derTheorie der quadratischenFormen Vonunendlichenvielenveranderlichen", J.Rein.Aug.Math, 136(1909), 210-271.

[6] Sibson R., Information Radius, Z, Wahrs. undverw.geb. (14) (1969), 149-160.

[7] Kullback S. , R. A. Leibler: On information and sufficiency. Ann. Math. Statistics 22(1951), 79-86.

[8] Kullback S.: Information Theory and Statistics. J. Wiley, New York, 1959.

[9] Taneja I.J., "New Developments in generalized information measures", Chapter in: Advances in imaging and Electron Physics, Ed. P. W. Hawkes 91 (1995),37-135.