Jordan Journal of Mathematics and Statistics (JJMS) 11(4), 2018, pp 309 - 323

OSTROWSKI INEQUALITY AND APPLICATIONS IN INFORMATION THEORY

RAM NARESH SARASWAT [(1)] AND AJAY TAK [(2)]

Abstract. Here we provide, an approximation for new f-divergence between two probability distributions defined on finite set in terms of integral means using Ostrowskis integral inequality. Some particular cases are derived. Consider some numerical illustration with the case when each pair p_i, q_i is very close and deal with applications to mutual information.

1. Introduction

Let

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

be the set of all complete finite discrete probability distributions. The new f-divergence between two probability distributions P and Q is defined 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 $f : (0, \infty) \to \Re_+$ is a convex function and $P = (p_1, p_2...p_n), Q = (q_1, q_2...q_n) \in \Gamma_n$, where p_i and q_i and for some $i = 1, 2, 3...n$, are probabilities (see Jain & Saraswat ([10]-[11]). All the below divergences are particular instances of new f-divergence. Some options of f satisfy $f(1) = 0$, so that $S_f(P, P) = 0$. Convexity ensures that

2000 *Mathematics Subject Classification.* 40H05, 46A45.

Key words and phrases. New f- divergence, Ostrowski inequality, Divergence measure, Mutual information.

Received: Nov. 19, 2017 Accepted: Feb. 8, 2018 .

$S_f(P,Q)$ is non-negative.

The following wellknown information divergence measures, Bhattacharya divergence[1], Kullback-Leibler divergence measure[13], Variational distance[14], Triangular discrimination[21], Relative J-divergence[7],Hellinger discrimination[9], Chi-square divergence[16], Relative Jensen-Shannon divergence[19], Relative arithmetic-geometric divergence[20] will be used in this paper.

- If $f(t) = (t-1)^2$, for $t > 0$, then Chi-square divergence measure is given by

$$S_f(P,Q) = \frac{1}{4}\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i} = \frac{1}{4}\sum_{i=1}^{n} \frac{p_i^2}{q_i} - 1 = \frac{1}{4}\chi^2(P,Q) \tag{1.2}$$

- If $f(t) = t \ln t$, for $t > 0$, then Relative Arithmetic-Geometric divergence measure is given by

$$S_f(P,Q) = \sum_{i=1}^{n} \left(\frac{p_i + q_i}{2}\right) \log\left(\frac{p_i + q_i}{2q_i}\right) = G(Q,P) \tag{1.3}$$

- If $f(t) = -\ln t$, for $t > 0$, then Relative Jensen-Shannon divergence measure is given by

$$S_f(P,Q) = \sum_{i=1}^{n} q_i \ln\left(\frac{2q_i}{p_i + q_i}\right) = F(Q,P) \tag{1.4}$$

- If $f(t) = \frac{(t-1)^2}{t}$, for $t > 0$, then Triangular discrimination is given by

$$S_f(P,Q) = \frac{1}{2}\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{(p_i + q_i)} = \frac{1}{2}\Delta(P,Q) \tag{1.5}$$

- If $f(t) = (t-1)\ln t$, for $t > 0$, then Relative J-divergence measure is given by

$$S_f(P,Q) = \sum_{i=1}^{n} \left(\frac{p_i - q_i}{2}\right) \ln\left(\frac{p_i + q_i}{2q_i}\right) = \frac{1}{2}J_R(P,Q) \tag{1.6}$$

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

$$S_f(P,Q) = \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right) = K(P,Q) \tag{1.7}$$

- If $f(t) = 1 - \sqrt{t}$, for $t > 0$, then Hellinger discrimination is given by

$$S_f(P,Q) = \sum_{i=1}^{n} \left[1 - B\left(\frac{P+Q}{2}, Q\right)\right] = h\left(\frac{P+Q}{2}, Q\right) \tag{1.8}$$

- If $f(t) = |t-1|$, for $t > 0$, then Variational distance is given by

$$S_f(P,Q) = \frac{1}{2}\sum_{i=1}^{n} |p_i - q_i| = \frac{1}{2}V(P,Q) \tag{1.9}$$

The following theorem 1.1 of Ostrowskis integral inequalityis given in [2],[4]-[6] and [7].Some applications to numerical integration, special means and short proof of inequality are defined in Dragomir [8].

Theorem 1.1. *Assume that* $g : [a,b] \to \Re$ *is absolutely continuous with* $g' \in L_\infty[a,b]$ *that is,* $||g'||_\infty := ess\sup_{t\in[a,b]} |g'(t)| < \infty$ *then*

$$|g(x) - \frac{1}{b-a}\int_a^b g(t)dt| \le \left[\frac{1}{4} + \left(\frac{x - \frac{a+b}{2}}{b-a}\right)^2\right](b-a)||g'||_\infty \tag{1.10}$$

$\forall x \in [a,b]$. *The following important result is given by Diaz and Metcalf (see[*[15], *p.61]).*

Theorem 1.2. *Suppose* $a_k(\neq 0)$ *and* $b_k(k = 1,2...n)$ *are real numbers satisfying* $m \le \frac{b_k}{a_k} \le M$. *Then*

$$\sum_{k=1}^{n} b_k^2 + mM\sum_{k=1}^{n} a_k^2 \le (M+m)\sum_{k=1}^{n} a_k b_k \tag{1.11}$$

Equality holds if and only if for each k *either* $b_k = ma_k$ *or* $b_k = Ma_k$. *The following extension of this will be used.*

Proposition 1.1. *Suppose the conditions of the Diaz-Metcalf result hold and* $t_k > 0$ *for* $k = 1,2...n$. *Then*

$$\sum_{k=1}^{n} t_k b_k^2 + mM\sum_{k=1}^{n} t_k a_k^2 \le (M+m)\sum_{k=1}^{n} t_k a_k b_k \tag{1.12}$$

Equality holds if and only if for each k either $b_k = ma_k$ or $b_k = Ma_k$. This follows for $k = 1, 2...n$ that

$$(\frac{b_k}{a_k} - m)(M - \frac{b_k}{a_k})t_k a_k^2 \geq 0$$

Then desired result follows on summation over k.

In Section 2. We provide, by the use of Ostrowskis integral inequality for absolutely continuous mappings with essentially bounded of first derivative,an approximation for the new f-divergence in terms of an integral mean.Here we use proposition 1.1 with slight modification according as new f-divergence.Section 3 considers some particular cases of Theorem 2.1 and section 4 considers some of the examples with the case when each pair p_i, q_i is very close. Finally in Section 5, we deal with applications to mutual information. This research work is similar to paper [6].

2. INFORMATION INEQUALITY ON NEW f-DIVERGENCE

In our main result, Theorem 2.1 below does not assume convexity.In this way we assume that there exist real numbers r, R with $0 < r \leq \frac{p_i+q_i}{2q_i} \leq R < \infty$ for all $i \in (1, 2...n)$. We will establish in Theorem 2.1 below that if p and q are close in the sense that $R-r$ is small, then the integral mean $\frac{1}{R-r}\int_r^R f(t)dt$ approximates the new f-divergence to first order.

Theorem 2.1. *Assume that $f : [r, R] \to \Re$ is absolutely continuous on $[r, R]$ and $f' \in L_\infty[r, R]$ Then*

$$|S_f(P,Q) - \frac{1}{R-r}\int_r^R f(t)dt| \tag{2.1}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(Q,P) + \left(\frac{R+r}{2} - 1\right)^2\right]\right](R-r)||f'||_\infty \leq \frac{1}{2}(R-r)||f'||_\infty$$

Proof: By Ostrowskis integral inequality, we have

$$|f\left(\frac{p_i+q_i}{2q_i}\right) - \frac{1}{R-r}\int_r^R f(t)dt| \leq \left[\frac{1}{4} + \left(\frac{\frac{p_i+q_i}{2q_i} - \frac{R+r}{2}}{(R-r)}\right)^2\right](R-r)||f'||_\infty$$

for each $i \in 1, 2...n$. We may multiply by q_i sum the resultant inequalities and use the generalized triangle inequality to obtain

$$\begin{aligned} &|S_f(P,Q) - \frac{1}{R-r}\int_r^R f(t)dt| \\ &\leq \sum_{i=1}^n q_i |f\left(\frac{p_i+q_i}{2q_i}\right) - \frac{1}{R-r}\int_r^R f(t)dt| \\ &\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\sum_{i=1}^n q_i\left(\frac{p_i+q_i}{2q_i} - \frac{R+r}{2}\right)^2\right](R-r)||f'||_\infty \end{aligned}$$

Since

$$\begin{aligned} &\sum_{i=1}^n q_i\left(\frac{p_i+q_i}{2q_i} - \frac{R+r}{2}\right)^2 \\ &= \sum_{i=1}^n \frac{q_i}{4}\left[\frac{(p_i+q_i)^2}{q_i^2} - \frac{2(p_i+q_i)(R+r)}{q_i} + (R+r)^2\right] \\ &= \frac{1}{4}\sum_{i=1}^n \left[\frac{(p_i+q_i)^2}{q_i} - 2(p_i+q_i)(R+r) + (R+r)^2\right] \\ &= \frac{1}{4}\sum_{i=1}^n \left[\frac{(p_i+q_i)^2}{q_i} + (R+r)(R+r-2)(p_i+q_i))\right] \\ &= \frac{1}{4}\sum_{i=1}^n \left[\left(\frac{p_i^2}{q_i} + 3\right) + (R+r)(R+r-4)\right] \end{aligned}$$

$$= \frac{1}{4}\sum_{i=1}^{n}\left[\left(\frac{p_i^2}{q_i}+3\right)+(R+r)^2-4(R+r)\right]$$

$$= \frac{1}{4}\sum_{i=1}^{n}\left[\left(\frac{p_i^2}{q_i}+3-4\right)+(R+r-2)^2\right]$$

$$= \frac{1}{4}\sum_{i=1}^{n}\left[\left(\frac{p_i^2}{q_i}-1\right)+(R+r-2)^2\right]$$

$$= \left[\frac{1}{4}\sum_{i=1}^{n}\chi^2(P,Q)+\frac{(R+r-2)^2}{2}\right] = \frac{1}{4}\sum_{i=1}^{n}\chi^2(P,Q)+\left(\frac{R+r}{2}-1\right)^2,$$

this yields the first inequality in (2.1).
For the second, set $b_k = \sqrt{\frac{p_k+q_k}{2q_k}}$ and $a_k = \sqrt{\frac{2q_k}{p_k+q_k}}, (k = 1, 2...n)$. Then $\frac{b_k}{a_k} = \frac{p_k+q_k}{2q_k} \in [r, R], (k = 1, 2...n)$. On applying Proposition 1.1 for

$$t_k = (p_k + q_k), (k = 1, 2...n),$$

we get

$$\left(\sum_{k=1}^{n}(p_k+q_k)\left(\sqrt{\frac{p_k+q_k}{2q_k}}\right)^2 + rR\sum_{k=1}^{n}(p_k+q_k)\left(\sqrt{\frac{2q_k}{p_k+q_k}}\right)^2\right)$$
$$\leq (r+R)\sum_{k=1}^{n}(p_k+q_k)\sqrt{\frac{p_k+q_k}{2q_k}}\sqrt{\frac{2q_k}{p_k+q_k}},$$

or equivalently

$$\sum_{k=1}^{n}\frac{(p_k+q_k)^2}{2q_k} + 2rR \leq 2(R+r),$$

$$\sum_{k=1}^{n}\frac{(p_k+q_k)^2}{q_k} \leq 4((R+r)-rR),$$

$$\sum_{k=1}^{n}\frac{p_k^2}{q_k} + 3 \leq 4((R+r)-rR),$$

$$\sum_{k=1}^{n}\frac{p_k^2}{q_k} - 1 \leq 4((R+r)-rR-1),$$

$$\chi^2(P,Q) \leq 4((R+r)-rR-1),$$

Thus

$$\frac{1}{4}\chi^2(P,Q) \leq R + r - rR - 1 = (1-r)(R-1)$$

and so

$$\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \frac{1}{4}((R+r)-2)^2\right] \leq \frac{1}{2}$$

and the theorem is proved.

Corollary 2.1. *Let f satisfy the conditions of Theorem 2.1. If $\epsilon > 0$ and $0 \leq R - r \leq 2\epsilon/||f'||_\infty$, then*

$$|S_f(P,Q) - \frac{1}{R-r}\int_r^R f(t)dt| \leq \epsilon.$$

For the approximation aspect Theorem 2.1 can be reformulated.

Corollary 2.2. *Let $f : [0,2] \to \Re$ be absolutely continuous with $f' \in L_\infty[0,2]$. If $\eta \in (0,1)$ and $p_i(\eta), q_i(\eta)$ are such that*

$$|\frac{p_i(\eta) + q_i(\eta)}{2q_i(\eta)} - 1| \leq \eta$$

for all $i \in \{1, 2...n\}$, then

$$S_f(P(\eta), Q(\eta)) = \frac{1}{2\eta}\int_{1-\eta}^{1+\eta} f(t)dt + R_f(P,Q,\eta)$$

and the remainder $R_f(P,Q,\eta)$ satisfies

$$R_f(P,Q,\eta) \leq \frac{\eta}{2}\left[1 + \frac{1}{\eta^2}\frac{1}{4}\chi^2(P(\eta),Q(\eta))\right]||f'||_\infty \leq \eta||f'||_\infty.$$

This follows by Theorem 2.1 with the options $R = 1 + \eta$ and $r = 1 - \eta(\eta \in (0,1))$.

3. SOME PARTICULAR CASES

In some applications, we can use of definition 1.1 for functions $f : [0, \infty] \to \Re$ which are continuous but not necessarily convex.

3.1 For Relative Arithmetic Geometric divergence, we take $f(t) = t \ln t$. With this option we have $||f'||_\infty = \ln(eR)$ and

$$\int_r^R f(t)dt = \frac{1}{4}\left[R^2 \ln R^2 - r^2 \ln r^2 - (R^2 - r^2)\right]$$
$$= \frac{R^2 - r^2}{4} \ln\left[\left(\frac{(R^2)^{R^2}}{(r^2)^{r^2}}\right)^{\frac{1}{R^2 - r^2}} \cdot \frac{1}{e}\right]$$
$$= \frac{R^2 - r^2}{4} I\left[R^2, r^2\right],$$

where the identic mean $I(a, b)$ for positive argument is given by

$$I(a,b) := \begin{cases} a & \text{if b=a} \\ \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}} & \text{if b}\neq\text{a} \end{cases}$$

The conclusion of Theorem 2.1 reads

$$|G(Q,P) - \frac{R+r}{4} \ln\left[I(R^2, r^2)\right]| \tag{3.1}$$
$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2} - 1\right)^2\right]\right](R-r)\ln(eR) \leq \frac{1}{2}(R-r)\ln(eR)$$

3.2 If we take the convex map $f : [0, \infty] \to \Re$ given by $f(t) = -\ln t$, then Relative Jensen-Shannon divergence measure we have

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

With this option $||f'||_\infty = \frac{1}{r}$ and the identic mean reappears through

$$\frac{1}{R-r}\int_r^R f(t)dt| = \ln\left[I(R,r)\right]$$

Theorem 2.1 provides

$$|F(Q,P) - \ln\left[\frac{1}{I(r,R)}\right]| \tag{3.2}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2} - 1\right)^2\right]\right]\left(\frac{R}{r} - 1\right) \leq \frac{1}{2}\left(\frac{R}{r} - 1\right)$$

3.3 For Variational distance, $f(t) = |t-1|$, this is absolutely continuous on $[r,R]$. We have $||f'||_\infty = sup_{t\in[r,R]}|f'(t)| = 1$ Further,

$$\frac{1}{R-r}\int_r^R f(t)dt = \frac{1}{R-r}\left[\int_r^1 (1-u)du + \int_1^R (u-1)du\right]$$
$$= \frac{1}{R-r}\left[\frac{(r-1)^2}{2} + \frac{(R-1)^2}{2}\right]$$
$$= \frac{1}{R-r}\left[\frac{(R-r)^2}{4} + \left(\frac{r+R}{2} - 1\right)^2\right]$$

Theorem 2.1 provides

$$|\frac{1}{2}V(P,Q) - \frac{1}{R-r}\left[\frac{(R-r)^2}{4} + \left(\frac{r+R}{2} - 1\right)^2\right]| \tag{3.3}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2} - 1\right)^2\right]\right](R-r) \leq \frac{1}{2}(R-r).$$

3.4 For Relative J-divergence, $f(t) = (t-1)\ln t$, this is absolutely continuous on [r,R]. We have

$$||f'||_\infty = sup_{u\in[r,R]}|f'(t)| = \ln(eR) - \frac{1}{R}$$

Further,

$$\int_r^R f(t)dt = \left[\ln R\left(\frac{R^2}{2} - R\right) - \frac{R^2}{4} + R\right] - \left[\ln r\left(\frac{r^2}{2} - r\right) - \frac{r^2}{4} + r\right]$$
$$= \left[\ln\frac{R^{(\frac{R^2}{2}-R)}}{r^{(\frac{r^2}{2}-r)}} - \frac{1}{4}(R^2 - r^2) + (R-r)\right].$$

Theorem 2.1 provides

$$|\frac{1}{2}J_R(P,Q) - \left[\frac{1}{R-r}\ln\frac{R^{(\frac{R^2}{2}-R)}}{r^{(\frac{r^2}{2}-r)}} - \frac{1}{4}(R+r)+1\right]| \tag{3.4}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2}-1\right)^2\right]\right](R-r)\left(\ln(eR) - \frac{1}{R}\right)$$

$$\leq \frac{1}{2}(R-r)\left(\ln(eR) - \frac{1}{R}\right).$$

3.5 For Triangular discrimination, $f(t) = \frac{(t-1)^2}{t}$, this is absolutely continuous on [r,R]. We have

$$||f'||_\infty = sup_{t\in[r,R]}|f'(t)| = 1 - \frac{1}{R^2}$$

Further,

$$\int_r^R f(t)dt = \left[\left(\frac{R^2}{2} - 2R + \ln R\right) - \left(\frac{r^2}{2} - 2r + \ln r\right)\right]$$

$$= \left[\ln\frac{R}{r} + \frac{1}{2}(R^2 - r^2) - 2(R-r)\right].$$

Theorem 2.1 provides

$$|\frac{1}{2}\Delta(P,Q) - \left[\frac{1}{R-r}\ln\frac{R}{r} + \frac{1}{2}(R+r) - 2\right]| \tag{3.5}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2}-1\right)^2\right]\right](R-r)\left(1 - \frac{1}{R^2}\right) \leq \frac{1}{2}(R-r)\left(1 - \frac{1}{R^2}\right)$$

3.6 Triangular discrimination, which arises with $f(t) = \frac{(t-1)^2}{(t+1)}$, We have

$$||f'||_\infty = sup_{u\in[r,R]}|f'(u)| = |f'(R)| = \frac{(R-1)(R+1)}{(R+1)^2}$$

Also

$$\frac{1}{R-r}\int_r^R f(u)du = \frac{R+r}{2} + \ln\left(\frac{R+1}{r+1}\right)^{(\frac{4}{R-r})} - 3$$

Theorem 2.1 provides

$$|\Delta\left(\frac{P+Q}{2}, Q\right) - \left[\frac{R+r}{2} + \ln\left(\frac{R+1}{r+1}\right)^{(\frac{4}{R-r})} - 3\right]| \tag{3.6}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2} - 1\right)^2\right]\right]\frac{(R-r)(R-1)(R+3)}{(R+1)^2}$$
$$\leq \frac{1}{2}\frac{(R-r)(R-1)(R+3)}{(R+1)^2}.$$

3.7 For Kullback-Leibler divergence measures,$f(t) = (2t-1)\ln(2t-1), t > \frac{1}{2}$, this is absolutely continuous on [r,R]. We have

$$||f'||_\infty = sup_{t\in[r,R]}|f'(t)| = \ln(e(2R-1))^2$$

Further,

$$\int_r^R f(t)dt = \left[\frac{(2R-1)^2}{8}(\ln(2R-1)^2 - 1) - \frac{(2r-1)^2}{8}(\ln(2r-1)^2 - 1)\right]$$
$$= \left[\ln\frac{(2R-1)^{(\frac{(2R-1)^2}{4})}}{(2r-1)^{(\frac{(2r-1)^2}{4})}} - \frac{1}{2}(R^2 - r^2) + \frac{1}{2}(R-r)\right].$$

Theorem 2.1 provides

$$|K(P,Q) - \left[\frac{1}{R-r}\ln\frac{(2R-1)^{(\frac{(2R-1)^2}{4})}}{(2r-1)^{(\frac{(2r-1)^2}{4})}} - \frac{1}{2}(R+r) + \frac{1}{2}\right]| \tag{3.7}$$

$$\leq \left[\frac{1}{4} + \frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q) + \left(\frac{R+r}{2} - 1\right)^2\right]\right](R-r)\ln(e(2R-1))^2$$
$$\leq \frac{1}{2}(R-r)\ln(e(2R-1))^2.$$

3.8 For Hellinger discrimination, $f(t) = (1-\sqrt{t}), t > 0$, this is absolutely continuous on [r,R]. We have

$$||f'||_\infty = sup_{t\in[r,R]}|f'(t)| = \frac{1}{2\sqrt{r}}$$

Further,

$$\int_r^R f(t)dt = \left[\left(R - \frac{2}{3}R^{\frac{3}{2}}\right) - \left(r - \frac{2}{3}r^{\frac{3}{2}}\right)\right]$$
$$= \left[(R-r) - \frac{2}{3}(R\sqrt{R} - r\sqrt{r})\right].$$

Theorem 2.1 provides

$$|h\left(\frac{P+Q}{2},Q\right)-\left[1-\frac{1}{R-r}\frac{2}{3}(R\sqrt{R}-r\sqrt{r})\right]| \tag{3.8}$$

$$\leq\left[\frac{1}{4}+\frac{1}{(R-r)^2}\left[\frac{1}{4}\chi^2(P,Q)+\left(\frac{R+r}{2}-1\right)^2\right]\right](R-r)\frac{1}{2\sqrt{r}}\leq\frac{1}{2}(R-r)\frac{1}{2\sqrt{r}}.$$

4. Numerical Illustration

In practically, where p_i and q_i are close, so that we have $p_i = p_i(\epsilon), q_i = q_i(\epsilon)$ and

$$|\frac{p_i(\epsilon)+q_i(\epsilon)}{2q_i(\epsilon)}-1|\leq\epsilon,\epsilon\in(0,1) \tag{4.1}$$

for all $i \in (1, 2...n)$, With $R = 1+\epsilon$ and $r = 1-\epsilon$, we obtain from 3.1 that

$$|G(q(\epsilon),p(\epsilon))-\frac{1}{2}\ln\left[I\left((1+\epsilon)^2,(1-\epsilon)^2\right)\right]|$$
$$\leq\frac{\epsilon}{2}\left[1+\frac{1}{\epsilon^2}\frac{1}{4}\chi^2(p(\epsilon),q(\epsilon))\right]\ln\left[e(1+\epsilon)\right]\leq\epsilon\ln\left[e(1+\epsilon)\right]$$

Therefore if $p(\epsilon), q(\epsilon)$ are in the sense of 4.1, we can approximate the Relative Arithmetic Geometric divergence $G(q(\epsilon), p(\epsilon))$ by $\frac{1}{2}\ln\left[I\left((1+\epsilon)^2,(1-\epsilon)^2\right)\right]$ and the error of the approximation is less than

$$E(\epsilon) := \epsilon\ln\left[e(1+\epsilon)\right].$$

From 3.2, we derive

$$|F(q(\epsilon),p(\epsilon))-\ln\left[\frac{1}{I(1-\epsilon,1+\epsilon)}\right]|$$
$$\leq\frac{1}{2}\frac{\epsilon}{1-\epsilon}\left[1+\frac{1}{\epsilon^2}\frac{1}{4}\chi^2(p(\epsilon),q(\epsilon))\right]\leq\frac{\epsilon}{1-\epsilon},$$

for $\epsilon \in (0, 1)$. Therefore for $p(\epsilon), q(\epsilon)$ satisfying 4.1, we can approximate the Relative Jensen-Shannon divergence measure $F(q(\epsilon), p(\epsilon))$ by $\ln\left[\frac{1}{I(1-\epsilon,1+\epsilon)}\right]$ and the error of the approximation is less than $\frac{\epsilon}{1-\epsilon}$ for $\epsilon \in (0, 1)$

5. APPLICATION TO MUTUAL INFORMATION

Mutual information [18] is a measure of the amount of information that one random variable provides about another. The reduction of uncertainty about one variable due to knowledge of the other is considered.

It is defined for two discrete- valued random variables X and Y with a joint probability mass function $t(x, y)$ and marginal probability mass functions $p(x)(x \in X)$ and $q(y)(y \in Y)$. It is the relative entropy between the joint distribution and the product distribution, that is,

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} t(x,y)\ln\left[\frac{t(x,y)}{p(x)q(x)}\right] = K(t(x,y), p(x)q(x)),$$

Whereas before $K(.,.)$ denotes Kullback-Leibler distance. We assume that

$$s \leq \frac{t(x,y)}{p(x)q(x)} \leq S \tag{5.1}$$

for all $(x, y) \in (XxY)$. Much as with r, R we have $s \leq 1 \leq S$ We also may consider mutual information in a chi-squared sense, that is,

$$I_{\chi^2}(X,Y) := \sum_{x\in X}\sum_{y\in Y} \frac{t^2(x,y)}{p(x)q(x)} - 1.$$

Inequality 3.7 yields the following proposition.

Proposition 5.1. *If t, p and q satisfy 5.1, then*

$$|I(X,Y) - \left[\frac{1}{S-s}\ln\left(\frac{(2S-1)^{\frac{(2S-1)^2}{4}}}{(2s-1)^{\frac{(2s-1)^2}{4}}}\right) - \frac{1}{2}(S+s) + \frac{1}{2}\right]|$$

$$\leq \left[\frac{1}{4} + \frac{1}{(S-s)^2}\{\frac{1}{4}I_{\chi^2}(X,Y) + \left(\frac{S+s}{2} - 1\right)^2\}\right](S-s)\ln(e(2S-1))^2$$

$$\leq \frac{1}{2}(S-s)\ln(e(2S-1))^2$$

6. **Acknowledgement**

Authors are grateful to the Editor-in-Chief and referees for their comments and valuable suggestions which helped a lot in improving the quality of this paper.

References

[1] A. Bhattacharya, Some analogues to amount of information and their uses in statistical estimation, *Sankhya* **8** (1946), 1–14

[2] P. Cerone, S.S. Dragomir, E. Kikianty, Jensen-Ostrowski type inequalities and applications for f-divergence measures, *Appl. Math. Comput.* **266** (2015), 304–315

[3] I. Csiszar and J. Fischer, Information sentfernungen in raum der wahrscheinlichkeistwahrscheinlichkeistverteilungen, *Magyar Tud. Akad. Mat. Kutat Int. Ksl* **7** (1962), 159-180

[4] P. Cerone, S.S. Dragomir, E. Kikianty, Ostrowski and Jensen type inequalities for higher derivatives with applications, *Journal of Inequalities and Special Functions* **7(1)** (2016), 61–77

[5] S.S. Dragomir, New Jensen and Ostrowski type Inequalities for General Lebesgue Integral with Applications, *Iranian Journal of Mathematical Sciences and Informatics* **11(2)** (2016), 1–22

[6] S.S. Dragomir, V. Gluscevic and C.E.M. Pearce, Csiszrs fdivergence, Ostrowskis Inequality and Mutual Information, *Nonlinear Analysis* **47** (2001), 2375–2386

[7] S.S. Dragomir, V. Gluscevic and C.E.M. Pearce, Approximation of Csiszrs fdivergence via midpoint Inequalities, in inequality theory and applications, Y. J. Cho, J.K. Kim and S.S. dragomir (Eds),*nova science publishers inc., Hutington New York*, **1** (2001), 139–154

[8] S.S. Dragomir and S. Wang, Application of Ostrowskis inequality to the estimation of error bounds for some special means and for some numerical quadrature rules, *Appl. Math. Lett.* **11** (1998), 105–109

[9] E. Hellinger, NeueBegrundung der Theorieder quadratischenFormen Von unendlichenvielenveranderlichen,*J.Rein.Aug.Math*,**136** (1909), 210–271

[10] K. C. Jain 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

[11] K. C. Jain and R. N. Saraswat, Some Bounds of Information Divergence Measures in Terms of Relative-Arithmetic-Geometric Divergence,*International Journal of Applied Mathematics and Statistics*, **32(2)** (2013), 48–58

[12] K. C. Jain and R. N. Saraswat, Some well-known inequalities and its applications in information theory, *Jordan Journal of Mathematics and Statistics* **6(2)** (2013), 157–167

[13] S. Kullback and A. Leibler, On information and sufficiency, *Ann. Math. Statist.* **22** (1951), 79-86

[14] J. N. Kapur, A comparative assessment of various measures of directed divergence, *Advances in Management Studies*, **3(1)** (1984), 1–16

[15] D. S. Mitrinovic, *Analytic inequalities*, Springer Verlag, 1970

[16] K. Pearson, 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

[17] R. N. Saraswat and A. Tak, Bounds on Information Divergence Measures inTerms of Hellinger Discrimination, *Journal of Basic and Applied Engineering Research*, **2(13)** (2015), 1135–1138

[18] C.E. Shannon, A mathematical theory of communication, *Bell. Syst. Tech. J.*, **27** (1948),370–423 and 623–656

[19] R. Sibson , Information Radius, *Z, Wahrs. Undverw.geb.* **(14)** (1969), 149–160

[20] I. J. Taneja, New developments in generalized information measures, *Chapter in Advances in imaging and Electron Physics, Ed. P. W. Hawkes* **91** (1995), 37–135

[21] Topsoe F., Some inequalities for information divergence and related measures ofdiscrimination, *RGMIA Res. Rep. Coll.*, **2(1)** (1999), 85-98

Department of Mathematics & Statistics, School of Basic Science, Manipal University Jaipur, 303007, INDIA

E-mail address: (1) `sarswatrn@gmail.com`

E-mail address: (2) `ajaytak86@gmail.com`