ISSN: **1304-7981** Number: **11**, Year:**2016**, Pages: **30-40** **http://jnrs.gop.edu.tr**

Gaziosmanpasa University
Graduate School of Natural and Applied Sciences
Journal of New Results in Science

Received: 17.07.2015 *Editors-in-Chief*: *Bilge Hilal Çadırcı*
Accepted: 31.08.2016 *Area Editor*: *Serkan Demiriz*

New Information Inequalities in Terms of Variational Distance and its Application

K.C. Jain[a,1] (jainkc_2003@yahoo.com)
Praphull Chhabra[b] (prfl24@gmail.com)

[a] *Department of Mathematics, Malaviya National Institute of Technology Jaipur- 302017, India*
[b] *Department of Mathematics, Malaviya National Institute of Technology Jaipur- 302017, India*

Abstract - In this work, new information inequalities are obtained and characterized on new generalized f- divergence (introduced by Jain and Saraswat (2012)) in terms of the Variational distance and these inequalities have been taken for evaluating some new relations among well known divergences. These new relations have been verified numerically by considering two discrete probability distributions: Binomial and Poisson. Asymptotic approximation on new generalized f- divergence is done as well.

Keywords - *Variational distance, new information inequalities, bounded variation, convex and normalized function, bounds of divergences, numerical verification, asymptotic approximation.*

1 Introduction

Divergence measures are basically measures of distance between two probability distributions or compare two probability distributions. It means that any divergence measure must take its minimum value zero when probability distributions are equal and maximum when probability distributions are perpendicular to each other.
Divergence measures have been demonstrated very useful in a variety of disciplines such as economics and political science [22, 23], biology [17], analysis of contingency tables [7], approximation of probability distributions [3, 13], signal processing [11, 12], pattern recognition [1, 2, 10], color image segmentation [15], 3D image segmentation and word alignment [21], cost- sensitive classification for medical diagnosis [18], magnetic resonance image analysis [24] etc.
Also we can use divergences in fuzzy mathematics as fuzzy directed divergences and

[1]K.C. Jain

fuzzy entropies which are very useful to find the amount of average ambiguity or difficulty in making a decision whether an element belongs to a set or not. Fuzzy information measures have recently found applications to fuzzy aircraft control, fuzzy traffic control, engineering, medicines, computer science, management and decision making etc.
Without essential loss of insight, we have restricted ourselves to discrete probability distributions, so let $\Gamma_n = \{P = (p_1, p_2, p_3, ..., p_n) : p_i > 0, \sum_{i=1}^{n} p_i = 1\}$, $n \geq 2$ be the set of all complete finite discrete probability distributions. The restriction here to discrete distributions is only for convenience, similar results hold for continuous distributions. If we take $p_i \geq 0$ for some $i = 1, 2, 3..., n$, then we have to suppose that $0f(0) = 0f\left(\frac{0}{0}\right) = 0$.
Jain and Saraswat [9] introduced new generalized f- divergence measure, which is given by

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

where $f : (0, \infty) \to R$ (set of real no.) is real, continuous, and 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 are probabilities. Many known divergences can be obtained from this generalized measure by suitably defining the convex function f. Some resultant divergences by $S_f(P, Q)$, are as follows.
(a). If we take $f(t) = |t - 1|$ in (1), we obtain

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

where $V(P, Q)$ is called the Variational distance (l_1 distance) [14].
(b). If we take $f(t) = -\log t$ in (1), we obtain

$$S_f(P, Q) = \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right) = F(Q, P). \tag{3}$$

where $F(Q, P)$ is called adjoint of the Relative JS divergence $F(P, Q)$ [19].
(c). If we take $f(t) = (t - 1)\log t$ in (1), we obtain

$$S_f(P, Q) = \frac{1}{2}\sum_{i=1}^{n} (p_i - q_i)\log\left(\frac{p_i + q_i}{2q_i}\right) = \frac{1}{2}J_R(P, Q), \tag{4}$$

where $J_R(P, Q)$ is called the Relative J- divergence [6].
(d). If we take $f(t) = t\log t$ in (1), we obtain

$$S_f(P, Q) = \sum_{i=1}^{n} \frac{p_i + q_i}{2}\log\left(\frac{p_i + q_i}{2q_i}\right) = G(Q, P), \tag{5}$$

where $G(Q, P)$ is called adjoint of the Relative AG divergence $G(P, Q)$ [20].
(e). If we take $f(t) = \frac{(t-1)^2}{t}$ in (1), we obtain

$$S_f(P, Q) = \frac{1}{2}\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i + q_i} = \frac{1}{2}\Delta(P, Q), \tag{6}$$

where $\Delta(P,Q)$ is called the Triangular discrimination [4].
(f). If we take $f(t) = (t-1)^2$ in (1), we obtain

$$S_f(P,Q) = \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^2}{q_i} = \frac{1}{4}\chi^2(P,Q), \tag{7}$$

where $\chi^2(P,Q)$ is called the Chi- square divergence or Pearson divergence measure [16]. We can see that

$$J_R(P,Q) = 2[F(Q,P)+G(Q,P)], \Delta(P,Q) = 2[1-W(P,Q)],$$

where $W(P,Q) = \sum_{i=1}^{n}\frac{2p_iq_i}{p_i+q_i}$ is Harmonic mean and $F(P,Q), G(P,Q)$ are given by (3) and (5) respectively. Divergences from (3) to (5) are non- symmetric and (2), (6) are symmetric with respect to probability distribution $P,Q \in \Gamma_n$.
Now, for a differentiable function $f:(0,\infty)\to R$, consider the associated functions $g:(0,\infty)\to R$ and $h:(0,\infty)\to R$, are given by

$$g(t) = (t-1)f'(t) \tag{8}$$

and

$$h(t) = (t-1)f'\left(\frac{t+1}{2}\right). \tag{9}$$

Put (8) and (9) in (1), we get the followings respectively.

$$E_{S_f}(P,Q) = \sum_{i=1}^{n}\left(\frac{p_i-q_i}{2}\right)f'\left(\frac{p_i+q_i}{2q_i}\right) \tag{10}$$

and

$$E^*_{S_f}(P,Q) = \sum_{i=1}^{n}\left(\frac{p_i-q_i}{2}\right)f'\left(\frac{p_i+3q_i}{4q_i}\right). \tag{11}$$

2 New information inequalities

In this section, we introduce new information inequalities on $S_f(P,Q)$ in terms of well known Variational distance. Such inequalities are for instance needed in order to calculate the relative efficiency of two divergences. Now, firstly the following theorem is well known in literature [9].

Theorem 2.1. *If the function f is convex and normalized, i.e., $f''(t)\geq 0 \;\forall\, t>0$ and $f(1)=0$ respectively, then $S_f(P,Q)$ and its adjoint $S_f(Q,P)$ are both non-negative and convex in the pair of probability distribution $(P,Q)\in\Gamma_n\times\Gamma_n$.*

Now, the following lemma 2.2 is very useful for proving the new inequalities (15) and (16). This lemma has been obtained from literature [5].

Lemma 2.2. *Let $\psi:[a,b]\subset R\to R$ be a differentiable function and is of bounded variation on $[a,b]$, i.e., $A_a^b(\psi) = \int_a^b|\psi'(t)|\,dt<\infty$. Then for all $u\in[a,b]$, we have*

$$\left|\int_a^b\psi(t)\,dt - \psi(u)(b-a)\right| \leq \left(\frac{b-a}{2}+\left|u-\frac{a+b}{2}\right|\right)A_a^b(\psi). \tag{12}$$

Now for all $u_1, u_2 \in [a,b]$, if we put $u = u_i$ and summing over i, we get the following inequalities from (12)

$$\left| \int_a^b \psi(t)\,dt - \left(\frac{b-a}{2}\right) \sum_{i=1}^{2} \psi(u_i) \right| \leq \left(\frac{b-a}{2} + \frac{1}{2} \sum_{i=1}^{2} \left| u_i - \frac{a+b}{2} \right| \right) A_a^b(\psi). \tag{13}$$

If we put $u = \frac{a+b}{2}$ in (12), we get the following inequalities

$$\left| \int_a^b \psi(t)\,dt - (b-a)\,\psi\left(\frac{a+b}{2}\right) \right| \leq \left(\frac{b-a}{2}\right) A_a^b(\psi). \tag{14}$$

Now, the following theorem introduces new information inequalities on $S_f(P,Q)$ by using above lemma.

Theorem 2.3. *Let $f : [\alpha, \beta] \subset (0,\infty) \to R$ be a twice differentiable function which is normalized, i.e., $f(1) = 0$ and f' is of bounded variation on $[\alpha,\beta]$, i.e., $A_\alpha^\beta(f') = \int_\alpha^\beta |f''(t)|\,dt < \infty$.*
If $P, Q \in \Gamma_n$ is such that $0 < \alpha < \frac{1}{2} \leq \frac{p_i+q_i}{2q_i} \leq \beta < \infty\ \forall\ i = 1,2,3...,n$ for some α and β with $0 < \alpha \leq 1 \leq \beta < \infty, \alpha \neq \beta$, then we have the following inequalities

$$\left| S_f(P,Q) - \frac{1}{2} E_{S_f}(P,Q) \right| \leq \frac{1}{2} V(P,Q)\, A_\alpha^\beta(f') \tag{15}$$

and

$$\left| S_f(P,Q) - E_{S_f}^*(P,Q) \right| \leq \frac{1}{4} V(P,Q)\, A_\alpha^\beta(f'), \tag{16}$$

where $S_f(P,Q), E_{S_f}(P,Q), E_{S_f}^*(P,Q)$ and $V(P,Q)$ are given by (1), (10) (11) and (2) respectively.

Proof:
Case I (for $1 \leq u$): Put $\psi = f', u_1 = a = 1$, and $u_2 = b = u \in [\alpha,\beta]$ in (13) and put $\psi = f', a = 1$, and $b = u \in [\alpha, \beta]$ in (14), we get respectively

$$\left| \int_1^u f'(t)\,dt - \left(\frac{u-1}{2}\right)(f'(1) + f'(u)) \right| \leq \left[\frac{u-1}{2} + \frac{1}{2}\left(\left|1 - \frac{u+1}{2}\right| + \left|u - \frac{u+1}{2}\right| \right) \right] A_1^u(f')$$

and

$$\left| \int_1^u f'(t)\,dt - (u-1)\, f'\left(\frac{u+1}{2}\right) \right| \leq \left(\frac{u-1}{2}\right) A_1^u(f').$$

Or

$$\left| f(u) - f(1) - \left(\frac{u-1}{2}\right)(f'(1) + f'(u)) \right| \leq \left(\frac{u-1}{2} + \left|\frac{u-1}{2}\right| \right) A_1^u(f')$$

and

$$\left| f(u) - f(1) - (u-1)\, f'\left(\frac{u+1}{2}\right) \right| \leq \left(\frac{u-1}{2}\right) A_1^u(f').$$

Or

$$\left| f(u) - \left(\frac{u-1}{2}\right)(f'(1) + f'(u)) \right| \leq (u-1)\, A_1^u(f') \leq (u-1)\, A_\alpha^\beta(f') \tag{17}$$

and

$$\left|f\left(u\right)-\left(u-1\right)f'\left(\frac{u+1}{2}\right)\right|\le\left(\frac{u-1}{2}\right)A_1^u\left(f'\right)\le\left(\frac{u-1}{2}\right)A_\alpha^\beta\left(f'\right). \tag{18}$$

Case II (for $u<1$): Put $\psi=f', u_1=a=u\in[\alpha,\beta]$, and $u_2=b=1$ in (13) and put $\psi=f', a=u\in[\alpha,\beta]$, and $b=1$ in (14), we get similarly respectively

$$\left|-f\left(u\right)-\left(\frac{1-u}{2}\right)\left(f'\left(1\right)+f'\left(u\right)\right)\right|\le\left(1-u\right)A_u^1\left(f'\right)\le\left(1-u\right)A_\alpha^\beta\left(f'\right)$$

and

$$\left|-f\left(u\right)-\left(1-u\right)f'\left(\frac{u+1}{2}\right)\right|\le\left(\frac{1-u}{2}\right)A_u^1\left(f'\right).$$

Or

$$\left|f\left(u\right)-\left(\frac{u-1}{2}\right)\left(f'\left(1\right)+f'\left(u\right)\right)\right|\le\left(1-u\right)A_u^1\left(f'\right)\le\left(1-u\right)A_\alpha^\beta\left(f'\right) \tag{19}$$

and

$$\left|f\left(u\right)-\left(u-1\right)f'\left(\frac{u+1}{2}\right)\right|\le\left(\frac{1-u}{2}\right)A_u^1\left(f'\right)\le\left(\frac{1-u}{2}\right)A_\alpha^\beta\left(f'\right). \tag{20}$$

From (17), (19) and from (18), (20), we get respectively

$$\left|f\left(u\right)-\left(\frac{u-1}{2}\right)\left(f'\left(1\right)+f'\left(u\right)\right)\right|\le\left|u-1\right|A_\alpha^\beta\left(f'\right) \tag{21}$$

and

$$\left|f\left(u\right)-\left(u-1\right)f'\left(\frac{u+1}{2}\right)\right|\le\left|\frac{u-1}{2}\right|A_\alpha^\beta\left(f'\right). \tag{22}$$

Now put $u=\frac{p_i+q_i}{2q_i}, i=1,2,3...,n$ in (21) and (22), we get respectively

$$\left|f\left(\frac{p_i+q_i}{2q_i}\right)-\left(\frac{p_i-q_i}{4q_i}\right)\left[f'\left(\frac{p_i+q_i}{2q_i}\right)+f'\left(1\right)\right]\right|\le\left|\frac{p_i-q_i}{2q_i}\right|A_\alpha^\beta\left(f'\right)$$

and

$$\left|f\left(\frac{p_i+q_i}{2q_i}\right)-\left(\frac{p_i-q_i}{2q_i}\right)f'\left(\frac{p_i+3q_i}{4q_i}\right)\right|\le\left|\frac{p_i-q_i}{4q_i}\right|A_\alpha^\beta\left(f'\right).$$

Now multiply the above expressions by q_i and sum over all $i=1,2,3...,n$ by taking into account $\sum_{i=1}^n p_i=\sum_{i=1}^n q_i=1$, we get the desire results (15) and (16) respectively.

3 Application of new information inequalities

In this section, we obtain bounds of standard divergence measures by using new inequalities defined in (15), in terms of Variational (taking only convex functions here).

Proposition 3.1. *Let $F\left(P,Q\right), V\left(P,Q\right)$, and $G\left(P,Q\right)$ be defined as in (3), (2), and (5) respectively. For $P,Q\in\Gamma_n$, we have*

$$\left|G\left(P,Q\right)-F\left(P,Q\right)\right|\le\log\left(\frac{\beta}{\alpha}\right)V\left(P,Q\right). \tag{23}$$

Proof: Let us consider

$$f(t) = t\log t, t > 0, f(1) = 0, f'(t) = 1 + \log t \text{ and } f''(t) = \frac{1}{t}.$$

Since $f''(t) > 0 \ \forall \ t > 0$ and $f(1) = 0$, so $f(t)$ is convex and normalized function respectively.
Now put $f(t)$ in (1) and put $f'(t)$ in (10), we get the followings respectively.

$$S_f(P,Q) = \sum_{i=1}^{n} \left(\frac{p_i + q_i}{2}\right) \log \frac{p_i + q_i}{2q_i} = G(Q,P). \tag{24}$$

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

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

$$= \sum_{i=1}^{n} \left[q_i \log \frac{2q_i}{p_i + q_i} + \left(\frac{p_i + q_i}{2}\right) \log \frac{p_i + q_i}{2q_i}\right] = F(Q,P) + G(Q,P). \tag{25}$$

$$A_\alpha^\beta(f') = \int_\alpha^\beta |f''(t)|\, dt = \int_\alpha^\beta \left|\frac{1}{t}\right| dt = \int_\alpha^\beta \frac{1}{t} dt = \log\left(\frac{\beta}{\alpha}\right). \tag{26}$$

The result (23) is obtained by using (24), (25), and (26) in (15), after interchanging P and Q.

Proposition 3.2. *Let $J_R(P,Q), V(P,Q)$, and $\Delta(P,Q)$ be defined as in (4), (2), and (6) respectively. For $P, Q \in \Gamma_n$, we have*

$$|J_R(P,Q) - \Delta(P,Q)| \leq 2\left(\frac{\beta - \alpha}{\alpha\beta} + \log \frac{\beta}{\alpha}\right) V(P,Q). \tag{27}$$

Proof: Let us consider

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

Since $f''(t) > 0 \ \forall \ t > 0$ and $f(1) = 0$, so $f(t)$ is convex and normalized function respectively.
Now put $f(t)$ in (1) and put $f'(t)$ in (10), we get the followings respectively.

$$S_f(P,Q) = \frac{1}{2} \sum_{i=1}^{n} (p_i - q_i) \log \left(\frac{p_i + q_i}{2q_i}\right) = \frac{1}{2} J_R(P,Q). \tag{28}$$

$$E_{S_f}(P,Q) = \frac{1}{2} \sum_{i=1}^{n} (p_i - q_i) \log \left(\frac{p_i + q_i}{2q_i}\right) + \frac{1}{2} \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i + q_i} = \frac{1}{2}\left[J_R(P,Q) + \Delta(P,Q)\right]. \tag{29}$$

$$A_\alpha^\beta(f') = \int_\alpha^\beta |f''(t)|\, dt = \int_\alpha^\beta \left|\frac{1+t}{t^2}\right| dt = \int_\alpha^\beta \frac{1+t}{t^2} dt = \frac{\beta - \alpha}{\alpha\beta} + \log \frac{\beta}{\alpha}. \tag{30}$$

The result (27) is obtained by using (28), (29) and (30) in (15).

Proposition 3.3. *Let $F(P,Q), V(P,Q)$, and $\Delta(P,Q)$ be defined as in (3), (2), and (6) respectively. For $P,Q \in \Gamma_n$, we have*

$$|4F(P,Q) - \Delta(P,Q)| \leq 2\left(\frac{\beta-\alpha}{\alpha\beta}\right) V(P,Q). \tag{31}$$

Proof: Let us consider

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

Since $f''(t) > 0 \ \forall\ t > 0$ and $f(1) = 0$, so $f(t)$ is convex and normalized function respectively.
Now put $f(t)$ in (1) and put $f'(t)$ in (10), we get the followings respectively.

$$S_f(P,Q) = \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i+q_i}\right) = F(Q,P). \tag{32}$$

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

$$= \sum_{i=1}^{n}\left[\frac{q_i^2+p_iq_i}{p_i+q_i} - \frac{2p_iq_i}{p_i+q_i}\right] = \sum_{i=1}^{n}\left[q_i - \frac{2p_iq_i}{p_i+q_i}\right] = 1 - W(P,Q) = \frac{1}{2}\Delta(P,Q). \tag{33}$$

$$A_\alpha^\beta(f') = \int_\alpha^\beta |f''(t)|\, dt = \int_\alpha^\beta \left|\frac{1}{t^2}\right| dt = \int_\alpha^\beta \frac{1}{t^2} dt = \frac{\beta-\alpha}{\alpha\beta}. \tag{34}$$

The result (31) is obtained by using (32), (33), and (34) in (15), after interchanging P and Q.

4 Numerical verification of obtained bounds

In this section, we give two examples for calculating the divergences

$$G(P,Q), F(P,Q), \Delta(P,Q), J_R(P,Q), V(P,Q)$$

together with verification of the inequalities (23), (27), and (31).
Example 4.1 Let P be the binomial probability distribution with parameters $(n = 10, p = 0.5)$ and Q its approximated Poisson probability distribution with parameter $(\lambda = np = 5)$ for the random variable X, then we have

x_i	0	1	2	3	4	5	6	7	8	9	10
$p_i \approx$	.000976	.00976	.043	.117	.205	.246	.205	.117	.043	.00976	.000976
$q_i \approx$	.00673	.033	.084	.140	.175	.175	.146	.104	.065	.036	.018
$\frac{p_i+q_i}{2q_i} \approx$	.573	.648	.757	.918	1.086	1.203	1.202	1.063	.831	.636	.527

Table 1: $(n = 10, p = 0.5, q = 0.5)$

By using Table 1, we get the followings.

$$\alpha\,(= .527) \leq \frac{p_i+q_i}{2q_i} \leq \beta\,(= 1.203). \tag{35}$$

$$G(P,Q)=\sum_{i=1}^{11}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2p_i}\right)\approx .031. \tag{36}$$

$$F(P,Q)=\sum_{i=1}^{11}p_i\log\left(\frac{2p_i}{p_i+q_i}\right)\approx .036. \tag{37}$$

$$\Delta(P,Q)=\sum_{i=1}^{11}\frac{(p_i-q_i)^2}{p_i+q_i}\approx .0917. \tag{38}$$

$$J_R(P,Q)=\sum_{i=1}^{11}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)\approx .0808. \tag{39}$$

$$V(P,Q)=\sum_{i=1}^{11}|p_i-q_i|\approx .3312. \tag{40}$$

Put the approximated numerical values from (35) to (40) in (23), (27), and (31), we get the followings respectively

$$5\times10^{-3}\le .2733,\ .0109\le 1.253,\ \text{and } .0523\le .7063.$$

Hence verified the inequalities (23), (27), and (31) for $p=0.5$.
Example 4.2 Let P be the binomial probability distribution with parameters $(n=10, p=0.7)$ and Q its approximated Poisson probability distribution with parameter $(\lambda=np=7)$ for the random variable X , then we have

x_i	0	1	2	3	4	5	6	7	8	9	10
$p_i\approx$	.0000059	.000137	.00144	.009	.036	.102	.200	.266	.233	.121	.0282
$q_i\approx$	.000911	.00638	.022	.052	.091	.177	.199	.149	.130	.101	.0709
$\frac{p_i+q_i}{2q_i}\approx$	.503	.510	.532	.586	.697	.788	1.002	1.392	1.396	1.099	.698

Table 2: $(n=10, p=0.7, q=0.3)$

By using Table 2, we get the followings.

$$\alpha\,(=.503)\le\frac{p_i+q_i}{2q_i}\le\beta\,(=1.396)\,. \tag{41}$$

$$G(P,Q)=\sum_{i=1}^{11}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2p_i}\right)\approx .0746. \tag{42}$$

$$F(P,Q)=\sum_{i=1}^{11}p_i\log\left(\frac{2p_i}{p_i+q_i}\right)\approx .0842. \tag{43}$$

$$\Delta(P,Q)=\sum_{i=1}^{11}\frac{(p_i-q_i)^2}{p_i+q_i}\approx .1812. \tag{44}$$

$$J_R(P,Q)=\sum_{i=1}^{11}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)\approx .1686. \tag{45}$$

$$V(P,Q)=\sum_{i=1}^{11}|p_i-q_i|\approx .4844. \tag{46}$$

Put the approximated numerical values from (41) to (46) in (23), (27), and (31), we get the followings respectively

$$9.6\times10^{-3}\le .4944, .0126\le 2.22098, \text{ and } .1556\le 1.2320.$$

Hence verified the inequalities (23), (27), and (31) for $p=0.7$.

5 Asymptotic Approximation

In this section, we introduce asymptotic approximation on $S_f(P,Q)$ in terms of well known Chi- square divergence.

Theorem 5.1. *If $f:(0,\infty)\to R$ is twice differentiable, convex, and normalized function, i.e., $f''(t)>0$ and $f(1)=0$ respectively, then we have*

$$S_f(P,Q)\approx\frac{f''(1)}{8}\chi^2(P,Q). \tag{47}$$

Equivalently

$$\left|\frac{S_f(P,Q)}{\chi^2(P,Q)}-\frac{f''(1)}{8}\right|<\epsilon \text{ ,when} |P-Q|<\delta,$$

where $\epsilon,\delta\to0$, i.e., ϵ,δ are very small and $S_f(P,Q),\chi^2(P,Q)$ are given by (1) and (7) respectively.

Proof: We know by Taylor's series expansion of function $f(t)$ at $t=1$, that

$$f(t)=f(1)+(t-1)f'(1)+\frac{(t-1)^2}{2!}f''(1)+(t-1)^2g(t), \tag{48}$$

where $g(t)=\frac{(t-1)}{3!}f'''(1)+\frac{(t-1)^2}{4!}f''''(1)+...$ and we can see that $g(t)\to0$ as $t\to1$, $f(1)=0$ because $f(t)$ is normalized, therefore from (48) we get

$$f(t)\approx(t-1)f'(1)+\frac{(t-1)^2}{2!}f''(1). \tag{49}$$

Now Put $t=\frac{p_i+q_i}{2q_i}$ in (49), multiply with q_i and then sum over all $i=1,2,3...,n$, we get the desire result (47).

Remark 5.2. *Particularly if we take $f(t)=\frac{(t-1)^2}{t},(t-1)\log t,-\log t,$ and $t\log t$ in (47), we get $\Delta(P,Q)\approx\frac{1}{2}\chi^2(P,Q), J_R(P,Q)\approx\frac{1}{2}\chi^2(P,Q), F(Q,P)\approx\frac{1}{8}\chi^2(P,Q)$, and $G(Q,P)\approx\frac{1}{8}\chi^2(P,Q)$ respectively, where $\Delta(P,Q), J_R(P,Q), \chi^2(P,Q), F(Q,P)$, and $G(Q,P)$ are given by (6), (4), (7), (3), and (5) respectively.*

Figure 1 shows the behavior of $F(P,Q), G(P,Q), \Delta(P,Q), J_R(P,Q), \chi^2(P,Q)$, and $V(P,Q)$. We have considered $p_i=(a,1-a), q_i=(1-a,a)$, where $a\in(0,1)$. It is clear from figure that the $V(P,Q)$ has a steeper slope than all others, $\chi^2(P,Q)$ has a steeper slope than remaining except $V(P,Q)$ and so on.

6 Conclusion and discussion

In this work, we presented new information inequalities on functions that have derivatives are of bounded variation for for $S_f(P,Q)$. Further, bounds of various well known divergences have been obtained in terms of the Variational distance in an interval (α,β), $0<\alpha\le 1\le\beta<\infty$ with $\alpha\ne\beta$ as an application of new inequalities. These bounds have been verified numerically by taking two discrete distributions: Binomial and Poisson. An approximation on $S_f(P,Q)$ has been done, which relates $S_f(P,Q)$ to $\chi^2(P,Q)$ approximately.

We found in our previous article [8] that square root of some particular divergences of Csiszar's class is a metric space but not each because of violation of triangle inequality, so we strongly believe that divergence measures can be extended to other significant problems of functional analysis and its applications, such investigations are actually in progress because this is also an area worth being investigated. Such types of divergences are also very useful to find the utility of an event, i.e., an event is how much useful compare to other event.

We hope that this work will motivate the reader to consider the extensions of divergence measures in information theory, other problems of functional analysis and fuzzy mathematics.

References

[1] M.B. Bassat, *f- Entropies, probability of error and feature selection*, Inform. Control, vol. 39, 1978, pp: 227-242.

[2] H.C. Chen, *Statistical pattern recognition*, Hoyderc Book Co., Rocelle Park, New York, 1973.

[3] C.K. Chow and C.N. Lin, *Approximating discrete probability distributions with dependence trees*, IEEE Trans. Inform. Theory, vol. 14, 1968, no 3, pp: 462-467.

[4] D. Dacunha- Castelle, *Ecole d'Ete de Probabilites de*, Saint-Flour VII-1977, Berlin, Heidelberg, New York: Springer, 1978.

[5] S.S. Dragomir, *On the Ostrowski's integral inequality for mappings of bounded variation and applications*, RGMIA Res. Rep. Coll. 2 (1999) No.1, 73?80.

[6] S.S. Dragomir, V. Gluscevic, and C.E.M. Pearce, *Approximation for the Csiszar f-divergence via midpoint inequalities*, in inequality theory and applications - Y.J. Cho, J.K. Kim, and S.S. Dragomir (Eds.), Nova Science Publishers, Inc., Huntington, New York, Vol. 1, 2001, pp. 139-154.

[7] D.V. Gokhale and S. Kullback, *Information in contingency Tables*, New York, Marcel Dekker, 1978.

[8] K.C. Jain and P. Chhabra, *Series of new information divergences, properties and corresponding series of metric spaces*, International Journal of Innovative Research in Science, Engineering and Technology, vol. 3- no.5 (2014), pp: 12124- 12132.

[9] K.C. Jain and R.N. Saraswat, *Some new information inequalities and its applications in information theory*, International Journal of Mathematics Research, vol. 4, Number 3 (2012), pp. 295-307.

[10] L. Jones and C. Byrne, *General entropy criteria for inverse problems with applications to data compression, pattern classification and cluster analysis*, IEEE Trans. Inform. Theory, vol. 36, 1990, pp: 23-30.

[11] T. Kailath, *The divergence and Bhattacharyya distance measures in signal selection*, IEEE Trans. Comm. Technology, vol. COM-15, 1967, pp: 52-60.

[12] T.T. Kadota and L.A. Shepp, *On the best finite set of linear observables for discriminating two Gaussian signals*, IEEE Trans. Inform. Theory, vol. 13, 1967, pp: 288-294.

[13] D. Kazakos and T. Cotsidas, *A decision theory approach to the approximation of discrete probability densities*, IEEE Trans. Perform. Anal. Machine Intell., vol. 1, 1980, pp: 61-67.

[14] A.N. Kolmogorov, *On the approximation of distributions of sums of independent summands by infinitely divisible distributions*, Sankhya, 25, 159-174, 1963.

[15] F. Nielsen and S. Boltz, *The Burbea-Rao and Bhattacharyya centroids*, Apr. 2010, arxiv.

[16] 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(1900), 157-172.

[17] E.C. Pielou, *Ecological diversity*, New York, Wiley, 1975.

[18] R. Santos-Rodriguez, D. Garcia-Garcia, and J. Cid-Sueiro, *Cost-sensitive classification based on Bregman divergences for medical diagnosis*, In M. A. Wani, editor, Proceedings of the 8th International Conference on Machine Learning and Applications (ICMLA'09), Miami Beach, Fl., USA, December 13-15, 2009, pp: 551- 556, 2009.

[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] B. Taskar, S. Lacoste-Julien, and M.I. Jordan, *Structured prediction, dual extra gradient and Bregman projections*, Journal of Machine Learning Research, 7, pp: 1627- 1653, July 2006.

[22] H. Theil, *Economics and information theory*, Amsterdam, North-Holland, 1967.

[23] H. Theil, *Statistical decomposition analysis*, Amsterdam, North-Holland, 1972.

[24] B. Vemuri, M. Liu, S. Amari, and F. Nielsen, *Total Bregman divergence and its applications to DTI analysis*, IEEE Transactions on Medical Imaging, 2010.