LE MATEMATICHE
Vol. LXX (2015) – Fasc. II, pp. 271–281
doi: 10.4418/2015.70.2.18

NEW INFORMATION INEQUALITIES ON NEW GENERALIZED f-DIVERGENCE AND APPLICATIONS

KAMAL C. JAIN - PRAPHULL CHHABRA

In this work, we introduce new information inequalities on new generalized f-divergence in terms of well known Chi-square divergence. Further we obtain relations of other standard divergences as an application of new inequalities by using Logarithmic power mean and Identric mean, together with numerical verification by taking two discrete probability distributions: Binomial and Poisson.

1. Introduction

Divergence measures are basically measures of distance between two probability distributions or compare two probability distributions.Divergence measure must take its minimum value zero when probability distributions are equal and maximum when probability distributions are perpendicular to each other. So, any divergence measure must increase as probability distributions move apart. Divergence measures have been demonstrated very useful in a variety of disciplines such as economics and political science (1972, 1967) [30, 31], biology (1975) [23], analysis of contingency tables (1978) [12], approximation of probability distributions (1968, 1980) [7, 20], signal processing (1967, 1967)

Entrato in redazione: 14 novembre 2014

AMS 2010 Subject Classification: 94A17, 26D15.
Keywords: New information inequalities, Logarithmic power mean, Identric mean, Binomial and Poisson distributions, Mutual information, Numerical approximation.

[18, 19], pattern recognition (1978, 1973, 1990) [2, 6, 17], color image segmentation (2010) [21], 3D image segmentation and word alignment (2006) [29], cost- sensitive classification for medical diagnosis (2009) [25], magnetic resonance image analysis (2010) [32] etc.

Also we can use divergences in fuzzy mathematics as fuzzy directed divergences and fuzzy entropies (2010, 2004, 2012) [1, 13, 16], 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 $\Gamma_n = \{P = (p_1, p_2, p_3, \ldots, p_n) : p_i > 0, \sum_{i=1}^{n} p_i = 1\}$, $n \geq 2$ 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 $i = 1, 2, 3 \ldots, n$, then we have to suppose that $0f(0) = 0f\left(\frac{0}{0}\right) = 0$.

Some generalized functional information divergence measures had been introduced, characterized and applied in variety of fields, such as: Csiszars f-divergence (1974, 1967) [8, 9], Bregmans f-divergence (1967) [4], Burbea- Raos f-divergence (1982) [5], Renyis like f-divergence (1961) [24] etc. Similarly, Jain and Saraswat (2012) [15] defined 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 $f : (0, \infty) \rightarrow R$ (set of real no.) is real, continuous, and convex function and $P = (p_1, p_2, \ldots, p_n), Q = (q_1, q_2, \ldots, q_n) \in \Gamma_n$, where p_i and q_i are probability mass functions. Many divergence measures can be obtained from the generalized f-divergence (1) by suitably defining the function f.

Some resultant divergences by $S_f(P,Q)$, are as follows.

• If we take $f(t) = -\log t$ in (1), we obtain

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

where $F(Q,P)$ is called adjoint of the Relative JS divergence $F(P,Q)$ (1969) [26].

• If we take $f(t) = \frac{(t-1)^2}{t}$ in (1), we obtain

$$S_f(P,Q) = \frac{1}{2} \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i + q_i} = \frac{1}{2} \Delta(P,Q), \tag{3}$$

where $\Delta(P,Q)$ is called the Triangular discrimination (1978) [10].
• If we take $f(t) = t\log t$ in (1), we obtain

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

where $G(Q,P)$ is called adjoint of the Relative AG divergence $G(P,Q)$ (1995) [28].
• If we take $f(t) = (t-1)\log t$ in (1), we obtain

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

where $J_R(P,Q)$ is called the Relative J-divergence (2001) [11].
• If we take $f(t) = (t-1)^2$ in (1), we obtain

$$S_f(P,Q) = \frac{1}{4}\sum_{i=1}^{n} \frac{(p_i-q_i)^2}{q_i} = \frac{1}{4}\chi^2(P,Q), \tag{6}$$

where $\chi^2(P,Q)$ is called the Chi- square divergence or Pearson divergence measure (1900) [22].
Similarly, we can obtain many divergences by using linear convex functions. Since these divergences are not worthful in practice, therefore we can skip them.
We can see that, divergence (3) is symmetric while (2), (4), (5), and (6) are non-symmetric with respect to probability distribution.
Now, there are two generalized means which are being used in this paper for calculations only. These are as follows.

$$L_p(a,b) = \begin{cases} \frac{b^{p+1}-a^{p+1}}{(p+1)(b-a)} & \text{if } p \neq -1, 0 \\ \frac{\log b - \log a}{b-a} & \text{if } p = -1 \\ 1 & \text{if } p = 0 \end{cases} \qquad a,b > 0,\ a \neq b. \tag{7}$$

$$I(a,b) = \begin{cases} \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-1}} & \text{if } a \neq b \\ b & \text{if } a = b \end{cases} \qquad a,b > 0. \tag{8}$$

Means (7) and (8) are called p- Logarithmic power mean and Identric mean respectively.

2. New information inequalities

In this section, we introduce new information inequalities on $S_f(P,Q)$. Such inequalities are for instance needed in order to calculate the relative efficiency of two divergences.

Definition 2.1. Convex function: A function $f(x)$ is said to be convex over an interval (a,b) if for every $x_1, x_2 \in (a,b)$ and $0 \leq \lambda \leq 1$, we have

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda) f(x_2),$$

and said to be strictly convex if equality does not hold only if $\lambda \neq 0$ or $\lambda \neq 1$.

Theorem 2.2. *Let $f:(\alpha,\beta) \subset R_+ \to R$ be a real, convex differentiable function on (α,β) with $0 < \alpha \leq 1 \leq \beta < \infty$, $\alpha \neq \beta$.*
If $P,Q \in \Gamma_n$ and satisfying the assumption $0 < \alpha < \frac{1}{2} \leq \frac{p_i+q_i}{2q_i} \leq \beta < \infty \forall i = 1,2,3,\ldots,n$, then we have the following inequalities

$$\begin{aligned} 0 &\leq B_{S_f}(\alpha,\beta) - S_f(P,Q) \\ &\leq \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\left[(\beta-1)(1-\alpha) - \frac{1}{4}\chi^2(P,Q)\right] \leq A_{S_f}(\alpha,\beta), \end{aligned} \tag{9}$$

where $S_f(P,Q)$ and $\chi^2(P,Q)$ are given by (1) and (6) respectively and

$$B_{S_f}(\alpha,\beta) = \frac{(\beta-1) f(\alpha) + (1-\alpha) f(\beta)}{\beta-\alpha}, \tag{10}$$

$$A_{S_f}(\alpha,\beta) = \frac{1}{4}(\beta-\alpha)\left[f'(\beta) - f'(\alpha)\right]. \tag{11}$$

Proof. Since f is differentiable and convex, therefore we can write the following for $a,b \in (\alpha,\beta)$ by mean value theorem

$$f'(a)(b-a) \leq f(b) - f(a). \tag{12}$$

Now put $a = c$ and $b = \frac{\delta c+\gamma d}{\delta+\gamma}$ (by assuming $\delta,\gamma \geq 0$ with $\delta+\gamma > 0$) in (12), we get

$$f'(c)\left(\frac{\delta c+\gamma d}{\delta+\gamma} - c\right) \leq f\left(\frac{\delta c+\gamma d}{\delta+\gamma}\right) - f(c),$$

i.e.,

$$(d-c)\frac{\gamma}{\delta+\gamma} f'(c) \leq f\left(\frac{\delta c+\gamma d}{\delta+\gamma}\right) - f(c). \tag{13}$$

In a similar manner, by putting $a = d$ and $b = \frac{\delta c+\gamma d}{\delta+\gamma}$ in (12), we get

$$(c-d)\frac{\delta}{\delta+\gamma} f'(d) \leq f\left(\frac{\delta c+\gamma d}{\delta+\gamma}\right) - f(d). \tag{14}$$

Now multiply (13) by δ and (14) by γ and add the resultant inequalities, we get

$$(d-c)\frac{\delta\gamma}{\delta+\gamma}\left[f'(c)) - f'(d))\right] \leq (\delta+\gamma) f\left(\frac{\delta c+\gamma d}{\delta+\gamma}\right) - \delta f(c) - \gamma f(d),$$

which is nothing but the following

$$0 \leq \frac{\delta f(c)+\gamma f(d)}{\delta+\gamma} - f\left(\frac{\delta c+\gamma d}{\delta+\gamma}\right) \leq (d-c)\frac{\delta\gamma}{(\delta+\gamma)^2}\left[f'(d))-f'(c)\right], \tag{15}$$

where first inequality of (15) is the definition of convex function itself.
If we choose $c=\alpha$, $d=\beta$, $\delta=\beta-x$, and $\gamma=x-\alpha$ in (15), we get

$$0 \leq \frac{(\beta-x)f(\alpha)+(x-\alpha)f(\beta)}{\beta-\alpha} - f(x) \leq \frac{(\beta-x)(x-\alpha)}{\beta-\alpha}\left[f'(\beta))-f'(\alpha)\right]. \tag{16}$$

Now put $x=\frac{p_i+q_i}{2q_i}$ in inequalities (16), multiply by q_i and then sum over all $i=1,2,3,\dots,n$, we get

$$0 \leq \frac{(\beta-1)f(\alpha)+(1-\alpha)f(\beta)}{\beta-\alpha} - \sum_{i=1}^{n} q_i f\left(\frac{p_i+q_i}{2q_i}\right) \leq \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}$$
$$\left[(\beta-1)(1-\alpha)-\frac{1}{4}\sum_{i=1}^{n}\left(\frac{{p_i}^2}{q_i}-1\right)\right]$$

or

$$0 \leq B_{S_f}(\alpha,\beta)-S_f(P,Q) \leq \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\left[(\beta-1)(1-\alpha)-\frac{1}{4}\chi^2(P,Q)\right]. \tag{17}$$

Since it is well known that $\chi^2(P,Q)\geq 0$ and $(\beta-1)(1-\alpha)\leq\frac{1}{4}(\beta-\alpha)^2$, therefore (17) can be written as

$$0 \leq B_{S_f}(\alpha,\beta)-S_f(P,Q) \leq \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\left[(\beta-1)(1-\alpha)-\frac{1}{4}\chi^2(P,Q)\right] \leq$$
$$\frac{1}{4}(\beta-\alpha)^2\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}, \text{ i.e.,}$$

$$\begin{aligned} 0 &\leq B_{S_f}(\alpha,\beta)-S_f(P,Q) \\ &\leq \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\left[(\beta-1)(1-\alpha)-\frac{1}{4}\chi^2(P,Q)\right] \leq A_{S_f}(\alpha,\beta). \end{aligned}$$

Hence the result is proven. □

3. Application of new information inequalities

In this section, we obtain bounds of different divergences in terms of the Chi-square divergence by using new inequalities (9).

Proposition 3.1. *Let $F(P,Q)$ and $\chi^2(P,Q)$ be defined as in (2) and (6) respectively. For $P,Q \in \Gamma_n$, we have*

$$0 \le \log I\left(\frac{1}{\alpha},\frac{1}{\beta}\right) - L_{-1}(\alpha,\beta) + L_0(\alpha,\beta) - F(Q,P)$$
$$\le L_{-2}(\alpha,\beta)\left[(\beta-1)(1-\alpha) - \frac{1}{4}\chi^2(P,Q)\right] \le \frac{1}{4}(\beta-\alpha)^2 L_{-2}(\alpha,\beta). \quad (18)$$

Proof. Let us consider

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

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 (10) and put $f'(t)$ in (11), we get the followings by considering means (7) and (8).

$$S_f(P,Q) = \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i+q_i}\right) = F(Q,P). \quad (19)$$

$$\begin{aligned} B_{S_f}(\alpha,\beta) &= \frac{(\beta-1)(-\log\alpha) + (1-\alpha)(-\log\beta)}{\beta-\alpha} \\ &= \frac{\log\beta^{\alpha} - \log\alpha^{\beta}}{\beta-\alpha} - \frac{\log\beta - \log\alpha}{\beta-\alpha} \\ &= \log I\left(\frac{1}{\alpha},\frac{1}{\beta}\right) - L_{-1}(\alpha,\beta) + L_0(\alpha,\beta). \end{aligned} \quad (20)$$

$$A_{S_f}(\alpha,\beta) = \frac{1}{4}(\beta-\alpha)\left[\frac{1}{\alpha} - \frac{1}{\beta}\right] = \frac{1}{4}\frac{(\beta-\alpha)^2}{\alpha\beta} = \frac{1}{4}(\beta-\alpha)^2 L_{-2}(\alpha,\beta). \quad (21)$$

The result (18) is obtained by using (19), (20), and (21) in (9). □

By using the similar approach, we obtain the relation of other divergences with chi-square divergence; these results are as follows (which can be proved on the same lines).

Proposition 3.2. *If we take $f(t) = \frac{(t-1)^2}{t}$, then we have*

$$\begin{aligned} 0 &\le (\beta-1)(1-\alpha)L_{-2}(\alpha,\beta) - \frac{1}{2}\Delta(P,Q) \\ &\le \frac{\alpha+\beta}{\alpha^2\beta^2}\left[(\beta-1)(1-\alpha) - \frac{1}{4}\chi^2(P,Q)\right] \\ &\le \frac{1}{2}(\beta-\alpha)^2 L_{-3}(\alpha,\beta). \end{aligned} \quad (22)$$

Proposition 3.3. *If we take $f(t) = t\log t$, then we have*

$$\begin{aligned} 0 &\leq \log I(\alpha,\beta) + L_0(\alpha,\beta) - G(Q,P) \\ &\leq L_{-1}(\alpha,\beta)\left[(\beta-1)(1-\alpha) - \frac{1}{4}\chi^2(P,Q)\right] \\ &\leq \frac{1}{4}(\beta-\alpha)^2 L_{-1}(\alpha,\beta). \end{aligned} \tag{23}$$

Proposition 3.4. *If we take $f(t) = (t-1)\log t$, then we have*

$$\begin{aligned} 0 &\leq (\beta-1)(1-\alpha)L_{-1}(\alpha,\beta) - \frac{1}{2}J_R(P,Q) \leq \\ &[L_{-1}(\alpha,\beta) + L_{-2}(\alpha,\beta)] \\ &\cdot\left[(\beta-1)(1-\alpha) - \frac{1}{4}\chi^2(P,Q)\right] \leq \frac{1}{4}(\beta-\alpha)^2[L_{-1}(\alpha,\beta) + L_{-2}(\alpha,\beta)]. \end{aligned} \tag{24}$$

4. Numerical verification of inequalities

In this section, we give two examples for calculating the divergences $\Delta(P,Q)$, $J_R(P,Q)$, and $\chi^2(P,Q)$ and verify the inequalities (22) and (24).

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
$p_i \approx$	.000976	.00976	.043	.117	.205	.246	.205
$q_i \approx$	.00673	.033	.084	.140	.175	.175	.146
$\frac{p_i+q_i}{2q_i} \approx$	.573	.648	.757	.918	1.086	1.203	1.202

x_i	7	8	9	10
$p_i \approx$	.117	.043	.00976	.000976
$q_i \approx$	.104	.065	.036	.018
$\frac{p_i+q_i}{2q_i} \approx$	1.063	.831	.636	.527

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

By using Table 1, we get the followings.

$$\alpha(= .527) \leq \frac{p_i+q_i}{2q_i} \leq \beta(= 1.203). \tag{25}$$

$$\Delta(P,Q) = \sum_{i=1}^{11} \frac{(p_i-q_i)^2}{p_i+q_i} \approx .0917. \tag{26}$$

$$J_R(P,Q)=\sum_{i=1}^{11}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)\approx .0808. \tag{27}$$

$$\chi^2(P,Q)=\sum_{i=1}^{11}\frac{(p_i-q_i)^2}{q_i}\approx .1471. \tag{28}$$

Put the approximated numerical values from (25) to (28) in (22) and (24) respectively and verify them.

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 By using

x_i	0	1	2	3	4	5	6
$p_i \approx$	.0000059	.000137	.00144	.009	.036	.102	.200
$q_i \approx$	.000911	.00638	.022	.052	.091	.177	.199
$\frac{p_i+q_i}{2q_i}\approx$	.503	.510	.532	.586	.697	.788	1.002

x_i	7	8	9	10
$p_i \approx$	.266	.233	.121	.0282
$q_i \approx$	.149	.130	.101	.0709
$\frac{p_i+q_i}{2q_i}\approx$	1.392	1.396	1.099	.698

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

Table 2, we get the followings.

$$\alpha\,(=.503)\le\frac{p_i+q_i}{2q_i}\le\beta\,(=1.396). \tag{29}$$

$$\Delta(P,Q)=\sum_{i=1}^{11}\frac{(p_i-q_i)^2}{p_i+q_i}\approx .1812. \tag{30}$$

$$J_R(P,Q)=\sum_{i=1}^{11}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)\approx .1686. \tag{31}$$

$$\chi^2(P,Q)=\sum_{i=1}^{11}\frac{(p_i-q_i)^2}{q_i}\approx .3298. \tag{32}$$

Put the approximated numerical values from (29) to (32) in (22) and (24) respectively and verify them.

5. Conclusion and discussion

Jain and Saraswat (2012) [15] introduced and characterized the new f-divergence $S_f(P,Q)$. Properties and relation of $S_f(P,Q)$ with well known Csiszars f-divergence can be seen in the same literature. In this work, we presented new information inequalities on convex functions for $S_f(P,Q)$. Further, bounds of various well known divergences have been obtained in terms of the chi-square divergence in an interval (α,β), $0<\alpha\leq 1\leq\beta<\infty$ with $\alpha\neq\beta$ as an application of new inequalities. These bounds have been verified numerically by taking two discrete distributions: Binomial and Poisson.

We found in our previous article (2014) [14] that square root of some particular divergences of Csiszars class is a metric space but $C_f(P,Q)$ is not a metric 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 and such investigations are actually in progress because this is also an area worth being investigated.

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. Such types of divergences are also very useful to find utility (2010, 1986) [3, 27] of an event, i.e., an event is how much useful compare to other event.

REFERENCES

[1] R. K. Bajaj - D. S. Hooda, *Generalized measures of fuzzy directed divergence, total ambiguity and information improvement*, Journal of Applied Mathematics, Statistics and Informatics 6 (2) (2010), 31–44.

[2] M. B. Bassat, *f- Entropies, probability of error and feature selection*, Inform. Control 39 (1978), 227–242.

[3] J S. Bhullar - O P. Vinocha - M. Gupta, *Generalized measure for two utility distributions*, Proceedings of the World Congress on Engineering 3, June 30- July 2, 2010.

[4] L. M. Bregman, *The relaxation method to find the common point of convex sets and its applications to the solution of problems in convex programming*, USSR Comput. Math. Phys. 7 (1967), 200–217.

[5] J. Burbea - C. R. Rao, *On the convexity of some divergence measures based on entropy functions*, IEEE Trans. on Inform. Theory IT-28 (1982), 489–495.

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

[7] C. K. Chow - C. N. Lin, *Approximating discrete probability distributions with dependence trees*, IEEE Trans. Inform. Theory 14 (3) (1968), 462–467.

[8] I. Csiszar, *Information measures: A Critical survey*, in Trans. In: Seventh Prague Conf. on Information Theory, Academia, Prague, 1974, 73–86.

[9] I. Csiszar, *Information type measures of differences of probability distribution and indirect observations*, Studia Math. Hungarica 2 (1967), 299–318.

[10] D. Dacunha-Castelle, *École d'Été de Probabilités de Saint-Flour VII-1977*, Springer,Berlin, Heidelberg, New York, 1978.

[11] S. S. Dragomir - V. Gluscevic - C. E. M. Pearce, *Approximation for the Csiszars 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, 139–154.

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

[13] D. S. Hooda, *On generalized measures of fuzzy entropy*, Mathematica Slovaca 54 (3) (2004), 315–325.

[14] K. C. Jain - P. Chhabra, *Series of new information divergences, properties and corresponding series of metric spaces*, International Journal of Innovative Research in Science, Engineering and Technology 3 (5) (2014), 12124–12132.

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

[16] P. Jha - V. K. Mishra, *Some new trigonometric, hyperbolic and exponential measures of fuzzy entropy and fuzzy directed divergence*, International Journal of Scientific and Engineering Research 3 (4) (2012), 1–5.

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

[18] T. T. Kadota - L. A. Shepp, *On the best finite set of linear observables for discriminating two Gaussian signals*, IEEE Trans. Inform. Theory 13 (1967), 288–294.

[19] T. Kailath, *The divergence and Bhattacharyya distance measures in signal selection*, IEEE Trans. Comm. Technology COM-15 (1967), 52–60.

[20] D. Kazakos - T. Cotsidas, *A decision theory approach to the approximation of discrete probability densities*, IEEE Trans. Perform. Anal. Machine Intell. 1 (1980), 61–67.

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

[22] 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.

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

[24] A. Renyi, *On measures of entropy and information*, Proc. 4th Berkeley Symposium on Math. Statist. and Prob. 1 (1961), 547–561.

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

[26] R. Sibson, *Information radius*, Z. Wahrs. Undverw. Geb. 14 (1969), 149–160.

[27] H. C. Taneja - R. K. Tuteja, *Characterization of a quantitative- qualitative measure of inaccuracy*, Kybernetika 22 (5) (1986), 393–402.

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

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

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

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

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

KAMAL C. JAIN
Department of Mathematics
Malaviya National Institute of Technology
Jaipur- 302017 (Rajasthan), India
e-mail: `jainkc_2003@yahoo.com`

PRAPHULL CHHABRA
Department of Mathematics
Malaviya National Institute of Technology
Jaipur- 302017 (Rajasthan), India
e-mail: `prfl24@gmail.com`