Journal of
Applied & Computational Mathematics

Saraswat, J Appl Computat Math 2013, 2:3
DOI: 10.4172/2168-9679.1000131

Research Article Open Access

Upper and Lower Bounds on New F-Divergence in Terms of Relative J-Divergence Measure

Ram Naresh Saraswat*

Research Scholar, Department of Mathematics, Malaviya National Institute of Technology, Jaipur, Rajasthan-302017 India

Abstract

During past years Dragomir, Taneja and Pranesh kumar have contributed a lot of work providing different kinds of bounds on the distance, information and divergence measures. 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 paper, we have established an upper and lower bounds of new f-divergence measure in terms of Relative J-divergence measure. Its particular cases have also considered using a new f-divergence and inequalities.

Keywords: Chi-square divergence; Jenson-Shannon's divergence; Triangular discrimination

Introduction

Let

$$\Gamma_n = \left\{ P = (p_1, p_2, \ldots\ldots p_n) \middle| p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \right\}, n \geq 2 \tag{1.1}$$

be the set of all complete finite discrete probability distributions. There are many information and divergence measures are exist in the literature of information theory and statistics. Csiszar [1,2] introduced a generalized measure of information using f-divergence measure given by

$$I_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \tag{1.2}$$

Where $f:[0,\infty) \to \mathbb{R}$ is a convex function and $P,Q \in \Gamma_n$

The Csiszar's f-divergence is a general class of divergence measures that includes several divergences used in measuring the distance or affinity between two probability distributions. This class is introduced by using a convex function f, defined on (0, ∞). An important property of this divergence is that many known divergences can be obtained from this measure by appropriately defining the convex function f. There are some examples of divergence measures in the category of Csiszar's f divergence measure. Bhattacharya divergence [3], Triangular discrimination [4], Relative J-divergence [5], Hellinger discrimination [6], Chi-square divergence [7], Relative Jensen-Shannon divergence [8], Relative arithmetic-geometric divergence [9], Unified relative Jensen-Shannon and arithmetic-geometric divergence of types [9]. In whole paper, we shall derive some well known divergence measures with help of new f-divergence measure. An inequality between new f divergence and Relative J- divergence measure has established which is shown below. Bounds of well known divergence measure in terms of Relative J- divergence measure have studied below. Numerical bounds of information divergence measure have also studied.

New f-Divergence Measure and Its Particular Cases

Given a function $f:[0,\infty) \to \mathbb{R}$, a new f-divergence measure introduced by Jain and Saraswat [10,11]

$$S_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i + q_i}{2q_i}\right) \tag{2.1}$$

For all $P,Q \in \Gamma_n$

$$S_f(P,Q) \geq f(1) \tag{2.2}$$

Equality holds in (2.2) iff

$$p_i = q_i \forall i = 1,2,..,n \tag{2.3}$$

Corollary 4.1.1: (Non-negativity of new f-divergence measure)

Let $f:[0,\infty) \to \mathbb{R}$

be convex and normalized, i.e..

$$f(1) = 0 \tag{2.4}$$

Then for any, P Q from (2.2) of proposition 2.1 and (2.4), we have the inequality

$$S_f(P,Q) \geq 0 \tag{2.5}$$

If *f* is strictly convex, equality holds in (2.5) iff

$$p_i = q_i \forall i \in [i, 2, \ldots\ldots\ldots, n] \tag{2.6}$$

and

$$S_f(P,Q) \geq 0 \text{ and } S_f(P,Q) = 0 \text{ iff } P = Q \tag{2.7}$$

Proposition 4.2: Let f_1 & f_2 are two convex functions and $g=a f_1+b f_2$ then Sg (P,Q)=a S_{f1}(P,Q)+b S_{f2} (P,Q, where a&b are constants and $P,Q \in \Gamma_n$,

Now we give some examples of well known information divergence measures which are obtain from new f-divergence measure.

Chi-square divergence measure: If $f(t) = (t-1)^2$ then Chi-square

***Corresponding author:** Ram Naresh Saraswat, Research Scholar, Department of Mathematics, Malaviya National Institute of Technology, Jaipur, Rajasthan-302017 India, E-mail: sarswatrn@gmail.com

Received January 02, 2013; **Accepted** June 20, 2013; **Published** July 05, 2013

Citation: Saraswat RN (2013) Upper and Lower Bounds on New F-Divergence in Terms of Relative J-Divergence Measure. J Appl Computat Math 2: 131. doi:10.4172/2168-9679.1000131

Citation: Saraswat RN (2013) Upper and Lower Bounds on New F-Divergence in Terms of Relative J-Divergence Measure. J Appl Computat Math 2: 131. doi:10.4172/2168-9679.1000131

divergence measure is given by

$$S_f(P,Q)=\frac{1}{4}\left[\sum_{i=1}^{n}\frac{p_i^2}{q_i}-1\right]=\frac{1}{4}\chi^2(P,Q) \tag{2.8}$$

Relative Jensen-Shannon divergence measure: If $f(t)=\log t$ then relative Jensen-Shannon divergence measure is given by

$$S_f(P,Q)=\sum_{i=1}^{n}q_i\log\left(\frac{2q_i}{p_i+q_i}\right)=F(Q,P) \tag{2.9}$$

Relative arithmetic-geometric divergence measure: If $f(t)=t\log t$ then relative arithmetic-geometric divergence measure is given by

$$S_f(P,Q)\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2q_i}\right)=G(Q,P) \tag{2.10}$$

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

$$S_f(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^2}{2(p_i+q_i)}=\frac{1}{2}\Delta(P,Q) \tag{2.11}$$

Relative J-divergence measure: If $f(t)=(t-1)\log t$ then Relative J-divergence measure is given by

$$S_f(P,Q)=\sum_{i=1}^{n}\left(\frac{p_i-q_i}{2}\right)\log\left(\frac{p_i+q_i}{2q_i}\right)=\frac{1}{2}J_R(P,Q) \tag{2.12}$$

Hellinger discrimination: If $f(t)=(1-\sqrt{t})$ then Hellinger discrimination is given by

$$S_f(P,Q)=\left[1-B\left(\frac{P+Q}{2},Q\right)\right]=h\left(\frac{P+Q}{2},Q\right) \tag{2.13}$$

Unified relative Jensen-Shannon and arithmetic-geometric divergence of type α

$$f(t)=\begin{cases}[\alpha(\alpha-1)]^{-1}[t^\alpha-1], & \alpha\neq 0,1\\ -\log t & \text{if } \alpha=0\\ t\log t & \text{if } \alpha=1\end{cases} \tag{2.14}$$

Then Unified relative Jensen-Shannon and Arithmetic-Geometric divergence measure of type α is given by

$$f(t)=\begin{cases}[FG_\alpha(Q,P)=[\alpha(\alpha-1]^{-1}\left[\sum_{i=1}^{n}q_i\left(p_i+q_i\right)^\alpha-1\right], & \alpha\neq 0,1\\ F(Q,P)=\sum_{i=1}^{n}q_i\log\left(\frac{2q_i}{p_{i+}q_i}\right), & \alpha=0\ =\Omega_\alpha(Q,P)\\ G(Q,P)=\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2q_i}\right) & \alpha=1\end{cases} \tag{2.15}$$

Inequality among New F-Divergence and Relative J-Divergence Measure

In the following theorem we have obtained an inequality between a new f-divergence measure and Relative J-divergence measure. The results are on similar lines to the result presented by Dragomir [12] and Jain and Saraswat [13,11,14].

Theorem 3.1: Assume that generating mapping $f:[0,\infty)\to\mathbb{R}$ is normalized i.e. f (1)=0 and satisfies the assumptions.

(i) f is twice differentiable on (r,R), where $0\le r\le 1\le R\le\infty$ (ii) there exist constants m, M such that

$$m\le\frac{t^2}{(1+t)}f''(t)\le M \tag{3.1}$$

If P, Q are discrete probability distributions satisfying the assumptions

$$r<\frac{1}{2}\le r_i=\frac{p_i+q_i}{2q_i}\le R,\forall i\in\{1,2,..........n\} \tag{3.2}$$

Then we have the inequality

$$\frac{m}{2}J_R(P,Q)\le S_f(P,Q)\le\frac{M}{2}J_f(P,Q) \tag{3.3}$$

Proof: Define a mapping: $f_m:[0,\infty)\to\mathbb{R},F_m(t)=f(t)-m(t-1)\log t,\forall t>0$

Then $F\,m$ (.) is normalized and twice differentiable, since

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

For all $t\in(r,R)$, it follows that $F\,m$ (.) is convex on (r,R). Applying non-negativity property of new f-divergence measure for $F\,m$ (.) and the linearity property, we may state that

$$\begin{aligned}&0\le S_F\ (P,Q)=S_f(P,Q)-mS_{(t\ \ 1)\log t}(P,Q)=S_f(P,Q)-—J_R(P,Q)\\&\Rightarrow 0\le S_f(P,Q)-—J_R(P,Q)\end{aligned} \tag{3.5}$$

From where the first inequality of (3.3) result

Now we again Define a mapping: $F_M:(0,\infty)\to\mathbb{R},F_M(t)=M(t-1)\log t-f(t)$, which is obviously normalized, twice differentiable and by (3.1), convex on (r,R). Applying non-negativity property of f-divergence functional for $F\,M$ (.) and the linearity property, we obtain the second part of (3.3) i.e.

$$0\le\frac{M}{2}J_R(P,Q)-S_f(P,Q) \tag{3.6}$$

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

Some Particular Cases

Using the result (3.3) of Theorem (3.1) we able to point out the following particular cases which may be interest in Information Theory and Statistics.

Result 6.1: Let $P,Q\in\Gamma n$ be two probability distribution with the property that

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

Then we have the following relation

$$\frac{1}{2(1+R)}J_f(P,Q)\le F(Q,P)\le R\frac{1}{2(1+r)}J_R(P,Q) \tag{4.1}$$

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

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

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

Define $g(t)=\frac{t^2}{1+t}f''(t)=\frac{t^2}{1+t}\left(\frac{1}{t^2}\right)=\frac{1}{1+t}\Rightarrow g'(t)=-\frac{1}{(1+t)^2}<0$

Hence function f is decreasing, then we get

$$M=\sup_{t\in[r,R]} g(t)=\frac{1}{(1+r)}, m=\inf_{t\in[r,R]} g(t)=\frac{1}{(1+R)} \tag{4.2}$$

Also $S_f(P,Q)=F(Q,P)$ from (2.9)

From equation (2.9), (3.3), (4.2) then prove of the result (4.1).

Result 6.2: Let, $P\ Q\in\Gamma n$ be two probability distribution satisfying (3.2), then we have the following relation

$$\frac{r}{2(1+r)}J_R(P,Q)\le G(P,Q)\le\frac{R}{2(1+R)}J_R(P,Q) \tag{4.3}$$

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

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

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

Define $g(t)=\frac{t^2}{(1+t)}f''(t)=\frac{t^2}{(1+t)}\left(\frac{1}{t}\right)=\frac{t}{1+t}\Rightarrow g'(t)=\frac{1}{(1+t^2)}>0$

Hence function f is increasing, then we get

$$M=\sup_{t\in[r,R]} g(t)=\frac{R}{(1+R)}, m=\inf_{t\in[r,R]} g(t)=\frac{r}{(1+r)} \tag{4.4}$$

Also $S_f(P,Q)=G(Q,P)$ from (2.10)

From equation (2.10), (3.3) & (4.4) prove of the result (4.3).

Result 6.3: Let P Q Î G n be two probability distributions satisfying (3.2), then we have the following relation

$$\frac{4r^2}{(1+r)}J_R(P,Q)\le\chi^2(P,Q)\le\frac{4R^2}{(1+R)}J_R(P,Q) \tag{4.5}$$

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

$f(t)=t^2-1, f'(t)=2t, f''(t)=2>0,\forall t>0$

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

Define $g(t)=\frac{t^2}{(1+t)}f''(t)=\frac{2t^2}{(1+t)}, g'(t)=\frac{2t(t+2)}{(1+t)^2}>0$

Hence function f is increasing, then obviously

$$M+\sup_{t\in[r,R]} g(t)=\frac{2}{(1+R)}, m=\inf_{t\in[r,R]} g(t)=\frac{2r^2}{(1+r)} \tag{4.6}$$

Also $S_f(P,Q)=\frac{1}{4}\chi^2(P,Q)$ (P,Q) from (2.8)

From equation (2.8), (3.3) & (4.6) prove of the result (4.5).

Result 6.4: Let $P\ Q\in\Gamma n$ be two probability distributions satisfying (3.2), then we have the following relation

$$\frac{4}{R(1+R)}J_R(P,Q)\le\Delta(P,Q)\le\frac{4}{r(1+r)}J_R(P,Q) \tag{4.7}$$

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

$f(t)=\frac{(t-1)^2}{t}=\left(t+\frac{1}{t}-2\right), f'(t)=\left(1-\frac{1}{t^2}\right), f''(t)=\frac{2}{t^3}$

$f''(t)\ge 0$ and $f(1)=0$, so function f is convex and normalized.

Define

$$g(t)=\frac{t^2}{(1+t)}f''(t)\frac{t^2}{(1+t)}\left(\frac{2}{t^3}\right)=\frac{2}{t(1+t)}, g'(t)=-\frac{(2t+1)}{t^2(1+t)^2}<0$$

Then obviously

$$M=\sup_{t\in[r,R]} g(t)=\frac{2}{r(1+r)}, m=\inf_{t\in[r,R]} g(t)=\frac{2}{R(1+R)} \tag{4.8}$$

Since $S_f(P,Q)=\frac{1}{2}\Delta(P,Q)$ from equation (2.11), from equation (2.11), (3.3) & (4.8) prove of the result (4.7).

Result 6.5: Let, $P\ Q\in\Gamma n$ be two probability distributions satisfying (3.2), and then we have the following relations

$$\frac{r^\alpha}{2(1+r)}J_R(P,Q)\le\Omega_\alpha(Q,P)\le\frac{R^\alpha}{2(1+R)}J_R(P,Q) \tag{4.9}$$

$$\frac{R^\alpha}{2(1+R)}J_R(P,Q)\le\Omega_\alpha(Q,P)\le\frac{r^\alpha}{2(1+r)}J_R(P,Q) \tag{4.10}$$

Proof: Consider the mapping $f:(r,R)\to\mathbb{R}$ and from equation (2.14), then we get

$f(t)=[\alpha(\alpha-1)]^{-1}[t^\alpha-1], f'(t)=[\alpha-1]^{-1}t^{\alpha-1}, f''(t)=t^{\alpha-2}>0,\forall t>0$

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

Define $g(t)=\frac{t^2}{(1+t)}f''(t)=\frac{t^2}{(1+t)}(t^{\alpha-2})=\frac{t^\alpha}{(1+t)}\Rightarrow g(t)'\frac{\alpha t^{\alpha-1}+t^\alpha(\alpha-1)}{(1+t^2)}$

If $g(t)'=0\Rightarrow t=\frac{\alpha}{1-\alpha}$

If α<1 then function *g(t)* is increasing and if α>1 then function g(t) is decreasing, we get

$$M=\sup_{t\in[r,R]} g(t)=\frac{R^\alpha}{(1+R)}, m=\inf_{t\in[r,R]} g(t)=\frac{r^\alpha}{(1+r)} \tag{4.11}$$

$$M=\sup_{t\in[r,R]} g(t)=\frac{r^\alpha}{(1+r)}, m=\inf_{t\in[r,R]} g(t)=\frac{R^\alpha}{(1+R)} \tag{4.12}$$

Also $S_f(P,Q)=\Omega_\alpha(Q,P)$ from (2.15)

From equation (3.3), (2.14), (2.15), (4.11) & (4.10) prove of the results (4.9) & (4.10).

Corollary 6.5.1: For α=1/2 of equations (4.9) and (4.10) and Let, $P\ Q\in\Gamma n$ be two probability distributions satisfying (3.2), then we have the following inequality

$$\frac{\sqrt{r}}{2(1+r)}J_R(P,Q)\le 4\left[1-B\left(\frac{P+Q}{2},Q\right)\right]\le\frac{\sqrt{R}}{2(1+R)}J_R(P,Q)\ \text{if}\ \alpha<1 \tag{4.13}$$

$$\frac{\sqrt{r}}{2(1+r)}J_R(P,Q)\le 4h\left[\left(\frac{P+Q}{2},Q\right)\right]\le\frac{\sqrt{R}}{2(1+R)}J_R(P,Q)\ \ \alpha<1 \tag{4.14}$$

$$\frac{\sqrt{R}}{2(1+R)}J_R(P,Q)\le 4\left[1-B\left(\frac{P+Q}{2},Q\right)\right]\le\frac{\sqrt{r}}{2(1+r)}J_R(P,Q)\ \ \alpha<1 \tag{4.15}$$

$$\frac{\sqrt{R}}{2(1+R)}J_R(P,Q)\le 4h\left[\left(\frac{P+Q}{2},Q\right)\right]\le\frac{\sqrt{r}}{2(1+r)}J_R(P,Q)\ \text{if}\ \alpha>1 \tag{4.16}$$

Where $\left[1-B\left(\frac{P+Q}{2},Q\right)\right]=h\left(\frac{P+Q}{2},Q\right) and\ B\left(\frac{P+Q}{2},Q\right)$ is Bhattacharya divergence measure [1].

Numerical Illustrations

Example 7.1: Let P be the binomial probability distribution for the random valuable X with parameter (n=8 p=0.5) and Q its approximated normal probability distribution

Here $\chi^2(P,Q)$ =0.00145837 and $J_R(P,Q)$ =0.00151848

It is noted that r=(0.774179933) $\leq \frac{p}{2} \leq R(1.050330018)$

We shall consider the upper and lower bounds of F (P,Q) , Δ(P,Q), G(P,Q) & $\chi^2(P,Q)$ are following.

(i) $\frac{1}{2(1+R)}J_R(P,Q) \leq F(Q,P) \leq \frac{1}{2(1+r)}J_R(P,Q)$ $\frac{1}{2(1+r)}J_R(P,Q)=0.00042791$

Lower bound: 0.00042791 ≤ $F(Q,P)$ ≤ 0.24386318

$\frac{1}{2(1+R)}J_R(P,Q)=0.24386318$ The width of the interval is 0.243435

Upper bound:

$$\frac{r}{2(1+r)}J_R(P,Q) \leq G(P,Q) \leq \frac{R}{2(1+R)}J_R(P,Q)$$

Lower bound:

$$\frac{r}{2(1+r)}J_R(P,Q)=0.000331$$

Upper bound:

$$\frac{R}{2(1+R)}J_R(P,Q)=0.256137$$

0.000331 ≤ G(P,Q) ≤ 0.256137

The width of the interval is 0.255806

iii $\frac{2}{R(1+R)}J_R(P,Q) \leq \Delta(P,Q) \leq \frac{2}{r(1+r)}J_R(P,Q)$

Lower bound:

$$\frac{2}{R(1+R)}J_R(P,Q)=0.001410$$

Upper bound: $\frac{2}{R(1+R)}J_R(P,Q)=2.15222533$

0.001410 ≤ $\Delta(P,Q) \leq 1.456097$

The width of the interval is 1.454687

iv $\frac{4r^2}{(1+r^2)}J_R(P,Q) \leq \chi^2(P,Q) \leq \frac{4R^2}{(1+R^2)}J_R(P,Q)$

Lower bound:

$$\frac{4r^2}{(1+r)}J_R(P,Q)=.002052$$

Upper bound:

$$\frac{4R^2}{(1+R)}J_R(P,Q)=2.15222533$$

0.002052 ≤ χ^2(P,Q) ≤ 2.15222533

The width of the interval is 2.150173.

References

1. Csiszar I (1978) Information measure: A critical survey. Trans 7th prague conf on info Th Statist Decius Funct, Random Processes and 8th European meeting of statist B.
2. Csiszar I (1961) Information-type measures of difference of probability functions and indirect observations. studia Sci Math hunger 2: 299-318.
3. Bhattacharya A (1947) Some analogues to amount of information and their uses in statistical estimation. Sankhya 8: 1-14.
4. Dacunha Castella D (1978) Ecole d'Ete de probabilities de Saint-Flour VII-1977 Berline, Heidelberg. Springer, USA.
5. Dragomir SS, Gluscevic V, Pearce CEM (2001) Approximation for the Csiszar f –divergence via mid-point inequalities, in inequality theory and applications. Nova science publishers, USA.
6. Hellinger E (2009) Neue Begrundung der Theorie der quadratischen Formen Von unendlichen vielen veranderlichen. J Rein Aug Math 1909: 210-271.
7. Pearson K (1900) 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: 157-172.
8. Sibson R (1969) Information Radius, Z, Wahrs Undverw geb. 14: 149-160.
9. Taneja IJ (1995) New Developments in generalized information measures. Advances in imaging and Electron Physics 91: 27-135.
10. Jain KC, Saraswat RN (2012) A New Information Inequality and its Application in Establishing Relation among various f-Divergence Measures. Journal of Applied Mathematics, Statistics and Informatics (JAMSI) 8: 17-32.
11. Jain KC, Saraswat RN (2013) Some bounds of information divergence measure in term of Relative arithmetic-geometric divergence measure. International Journal of Applied Mathematics and Statistics 32: 48-58.
12. Dragomir SS (2001) Some inequalities for (m, M)-convex mappings and applications for the Csiszar's F-divergence in information theory. Math J Ibaraki Univ 33: 35-50.
13. Jain KC, Saraswat RN (2011) 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.
14. Jain KC, Saraswat RN (2012) Inequalities among well known divergence measures and triangular discrimination. International Journal of Computational Science and Mathematics 4: 215-224.