VARIOUS RELATIONS ON NEW INFORMATION
DIVERGENCE MEASURES

K.C. Jain¹, Praphull Chhabra²

Department of Mathematics, Malaviya National Institute of Technology, Jaipur- 302017 (Rajasthan), INDIA.

Abstract

Divergence measures are useful for comparing two probability distributions. Depending on the nature of the problem, different divergence measures are suitable. So it is always desirable to develop a new divergence measure.

Recently, Jain and Chhabra [6] introduced new series \( (\xi^*_{m}(P, Q), \xi_{m}(P, Q) \text{ and } \omega_{m}(P, Q) \text{ for } m \in N ) \) of information divergence measures, defined the properties and characterized, compared with standard divergences and derived the new series \( (\sqrt{\xi^*_{m}(P, Q)} \text{ for } m \in N ) \) of metric spaces.

In this work, various important and interesting relations among divergences of these new series and other well known divergence measures are obtained. Some intra relations among these new divergences are evaluated as well and bounds of new divergence measure \( (\xi^*_1(P, Q)) \) are obtained by using Csiszar’s information inequalities. Numerical illustrations (Verification) regarding bounds are done as well.

Index terms:

Convex function, Divergence measure, Algebraic inequalities, Csiszar’s inequalities, Bounds, Mean divergence measures, Difference of mean divergences, Difference of divergences.

Mathematics subject classification: Primary 94A17, Secondary 26D15.

1. Introduction

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. If we take \( p_i \geq 0 \) for some \( i = 1, 2, 3, \ldots, n \), then we have to suppose that \( 0f(0) = 0f\left(\begin{array}{c} 0 \\ 0 \end{array}\right) = 0 \).

Csiszar’s \( f \)-divergence [2] is a generalized information divergence measure, which is given by (1.1), i.e.,
\[ C_f(P, Q) = \sum_{i=1}^{n} q_i f \left( \frac{p_i}{q_i} \right). \]  

(1.1)

Where \( f: (0, \infty) \to \mathbb{R} \) (set of real no.) is real, continuous and convex function and \( P = (p_1, p_2, p_3, ..., p_n), Q = (q_1, q_2, q_3, ..., q_n) \in \Gamma_n \), where \( p_i \) and \( q_i \) are probability mass functions. Many known divergence measures can be obtained from this generalized measure by suitably defining the convex function \( f \). Some of those are as follows.

\[ \xi_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m}}{(p_i q_i)^{(2m-1)/2} q_i^{2m}}, m = 1, 2, 3 ... \]  

(Jain and Chhabra [6])

(1.2)

\[ \zeta_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m} (p_i^4 - 2p_i^2 q_i^2 + p_i q_i^3 + q_i^4)}{(p_i q_i)^{(2m+1)/2} q_i^{2m+2}}, m = 1, 2, 3 ... \]  

(Jain and Chhabra [6])

(1.3)

\[ \omega_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m} \exp \left( \frac{(p_i^2 - q_i^2)^2}{p_i q_i^3} \right)}{q_i^{2m}}, m = 1, 2, 3 ... \]  

(Jain and Chhabra [6])

(1.4)

\[ E'_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m}}{(p_i q_i)^{(2m-1)/2}}, m = 1, 2, 3, ... \]  

(Jain and Srivastava [11])

(1.5)

\[ J'_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m} \exp \left( \frac{(p_i^2 - q_i^2)^2}{p_i q_i} \right)}{q_i^{2m}}, m = 1, 2, 3, ... \]  

(Jain and Srivastava [11])

(1.6)

\[ N'_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m} \exp \left( \frac{(p_i^2 - q_i^2)^2}{(p_i + q_i)^2} \right)}{(p_i + q_i)^{2m-1}}, m = 1, 2, 3, ... \]  

(Jain and Saraswat [10])

(1.7)

\[ \psi'_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i^2 - q_i^2)^{2m}}{2(p_i q_i)^{3/2}} \]  

(Kumar and Johnson [16])

(1.8)

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

(Kumar and Hunter [15])

(1.9)

Puri and Vineze divergence (Kafka, Osterreicher and Vincze [13])

\[ \Delta_m(P, Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^{2m}}{(p_i + q_i)^{2m-1}}, m = 1, 2, 3, ... \]  

(1.10)

Where \( \Delta(P, Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i + q_i} \) = Triangular discrimination, is a particular case of (1.10) at \( m=1 \).

Relative Arithmetic-Geometric divergence (Taneja [19])
\[ G(P,Q) = \sum_{i=1}^{n} \left( \frac{p_i + q_i}{2} \right) \log \left( \frac{p_i + q_i}{2p_i} \right). \]  

(1.11)

Arithmetic-Geometric mean divergence (Taneja [19])

\[ T(P,Q) = \frac{1}{2} \left[ G(P,Q) + G(Q,P) \right] = \sum_{i=1}^{n} \frac{p_i + q_i}{2} \log \left( \frac{p_i + q_i}{2\sqrt{p_iq_i}} \right). \]  

(1.12)

Where \( G(P,Q) \) is given by (1.11).

d-Divergence (Taneja [20])

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

(1.13)

Symmetric Chi-square divergence (Dragomir, Sunde and Buse [4])

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

(1.14)

Where \( \chi^2(P,Q) \) is given by (1.18).

Relative J- divergence (Dragomir, Gluscevic and Pearce [3])

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

(1.15)

Where \( F(P,Q) \) and \( G(P,Q) \) are given by (1.16) and (1.11) respectively.

Relative Jensen- Shannon divergence (Sibson [18])

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

(1.16)

Hellinger discrimination (Hellinger [5])

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

(1.17)

Where \( G^*(P,Q) \) is given by (1.29).

Chi-square divergence (Pearson [17])

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

(1.18)

Relative information (Kullback and Leibler [14])

\[ K(P,Q) = \sum_{i=1}^{n} p_i \log \left( \frac{p_i}{q_i} \right). \]  

(1.19)

J- Divergence (Jeffreys, Kullback and Leibler [12, 14])

\[ J(P,Q) = K(P,Q) + K(Q,P) = J_R(P,Q) + J_R(Q,P) = \sum_{i=1}^{n} (p_i - q_i) \log \left( \frac{p_i}{q_i} \right). \]  

(1.20)

Where \( J_R(P,Q) \) and \( K(P,Q) \) are given by (1.15) and (1.19) respectively.

Jensen- Shannon divergence (Burbea, Rao and Sibson [1, 18])
\[
I(P,Q) = \frac{1}{2} \left[ F(P,Q) + F(Q,P) \right] = \frac{1}{2} \left[ \sum_{i=1}^{n} p_i \log \left( \frac{2p_i}{p_i + q_i} \right) + \sum_{i=1}^{n} q_i \log \left( \frac{2q_i}{p_i + q_i} \right) \right].
\]

(1.21)

Where \( F(P,Q) \) is given by (1.16).

Some mean divergence measures and difference of particular mean divergences can be seen in literature (Taneja [21]), these are as follows.

Root mean square divergence \( = S(P,Q) = \sum_{i=1}^{n} \sqrt{\frac{p_i^2 + q_i^2}{2}}. \) (1.22)

Harmonic mean divergence \( = H(P,Q) = \sum_{i=1}^{n} \frac{2p_i q_i}{p_i + q_i} \) (1.23)

Arithmetic mean divergence \( = A(P,Q) = \sum_{i=1}^{n} \frac{p_i + q_i}{2} = 1. \) (1.24)

Square root mean divergence \( = N_1(P,Q) = \sum_{i=1}^{n} \left( \sqrt{\frac{p_i^2 + q_i^2}{2}} \right)^2. \) (1.25)

\[ N_2(P,Q) = \sum_{i=1}^{n} \left( \sqrt{\frac{p_i^2 + q_i^2}{2}} \right) \sqrt{\frac{p_i + q_i}{2}}. \] (1.26)

Heronian mean divergence \( = N_3(P,Q) = \sum_{i=1}^{n} p_i + \sqrt{p_i q_i} + q_i \) (1.27)

Logarithmic mean divergence \( = L(P,Q) = \sum_{i=1}^{n} \frac{p_i - q_i}{\log p_i - \log q_i}, p_i \neq q_i \forall i. \) (1.28)

Geometric mean divergence \( = G(P,Q) = \sum_{i=1}^{n} \sqrt{p_i q_i}. \) (1.29)

Square root- arithmetic mean divergence

\[ M_{SA}(P,Q) = S(P,Q) - A(P,Q) = \sum_{i=1}^{n} \left( \sqrt{\frac{p_i^2 + q_i^2}{2}} - 1 \right). \] (1.30)

Where \( S(P,Q) \) and \( A(P,Q) \) are given by (1.22) and (1.24) respectively.

Square root- geometric mean divergence

\[ M_{SG^*}(P,Q) = S(P,Q) - G^*(P,Q) = \sum_{i=1}^{n} \left( \sqrt{\frac{p_i^2 + q_i^2}{2}} - \sqrt{p_i q_i} \right). \] (1.31)

Where \( S(P,Q) \) and \( G^*(P,Q) \) are given by (1.22) and (1.29) respectively.

Square root- harmonic mean divergence

\[ M_{SH}(P,Q) = S(P,Q) - H(P,Q) = \sum_{i=1}^{n} \left( \sqrt{\frac{p_i^2 + q_i^2}{2}} - \frac{2p_i q_i}{p_i + q_i} \right). \] (1.32)
Where \( S(P, Q) \) and \( H(P, Q) \) are given by (1.22) and (1.23) respectively.

Some difference of particular divergences can be seen in literature (Taneja [20]), these are as follows.

\[
D_{\psi T}(P, Q) = \frac{1}{16} \psi(P, Q) - T(P, Q) = \frac{1}{2} \sum_{i=1}^{n} \left( p_i + q_i \right) \left( \frac{1}{8} \left( \frac{p_i - q_i}{p_i q_i} \right)^2 - \log \left( \frac{p_i + q_i}{2 \sqrt{p_i q_i}} \right) \right). 
\tag{1.33}
\]

\[
D_{\psi J}(P, Q) = \frac{1}{2} \psi(P, Q) - J(P, Q) = \sum_{i=1}^{n} \left( p_i - q_i \right) \left( \frac{p_i^2 - q_i^2}{2 p_i q_i} - \log \left( \frac{p_i}{q_i} \right) \right). 
\tag{1.34}
\]

Where \( \psi(P, Q), T(P, Q) \) and \( J(P, Q) \) are given by (1.14), (1.12) and (1.20).

Divergences from (1.2) to (1.7) and (1.10) are series of divergence measures corresponding to series of convex functions respectively. Out of them, (1.2), (1.3) and (1.4) are recently introduced by Jain and Chhabra and these series have been taken in this paper for deriving various important and interesting relations. (1.4), (1.6) and (1.7) are exponential while (1.9), (1.11), (1.12), (1.15), (1.16), (1.19), (1.20), (1.21), (1.33), (1.34) are logarithmic divergences respectively. Mean divergences from (1.22) to (1.32) are not members of Csiszar’s class, as corresponding function is need not be convex but divergences (1.33) and (1.34) are members as their corresponding function is convex.

Divergences from (1.5) to (1.10), (1.12) to (1.14), (1.17) and (1.20) to (1.34) are symmetric while (1.2) to (1.4), (1.11), (1.15), (1.16), (1.18) and (1.19) are non-symmetric, with respect to probability distributions \( P, Q \in \Gamma_n \).  

Divergences (1.17), (1.18), (1.19), (2.20) and (1.21) are also known as Kolmogorov’s measure, Pearson divergence, Directed divergence, Jeffreys- Kullback- Leibler divergence and Information radius, respectively.

Now, the whole paper is organized as follows. In section 2, we have evaluated some new basic relations on new divergence measures (1.2) to (1.4) by using algebraic inequalities. In section 3, we have obtained various new important relations on new divergences (1.2) to (1.4) with other standard divergences. Further, in section 4, we have discussed the bounds of new divergence measure \( \xi(P, Q) \) by using well known Csiszar’s information inequalities. In next section 5, we have done numerical illustrations for verifying the bounds of \( \xi(P, Q) \). Section 6 concludes the paper. Section 7 gives the references.

2. Some basic new relations

Now, firstly the following theorem is well known in literature [2].

**Theorem 2.1** If the function \( f \) is convex and normalized, i.e., \( f^{(n)}(t) \geq 0 \) and \( f(1) = 0, t > 0 \) then \( C_f(P, Q) \) and its adjoint \( C_f(Q, P) \) are both non-negative and convex in the pair of probability distribution \( (P, Q) \in \Gamma_n \times \Gamma_n \).
There are some new algebraic inequalities, which are important tool to derive some interesting and important new relations in this paper. These inequalities are as follows.

Proposition 2.1 Let \( t \in (0, \infty) \), then we have the following algebraic inequalities.

\[
\frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} > \frac{(t-1)^{2m}}{t^{(2m-1)/2}}, \quad m = 1, 2, 3, \ldots
\]  

(2.1)

\[
\frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} > \frac{(t-1)^{2m}}{(t+1)^{2m-1}}, \quad m = 1, 2, 3, \ldots
\]  

(2.2)

\[
\frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} \exp \left\{ \frac{(t^2 - 1)^2}{t} \right\} > \frac{(t-1)^{2m}}{t^{(2m-1)/2}} \exp \left\{ \frac{(t-1)^2}{t} \right\}, \quad m = 1, 2, 3, \ldots
\]  

(2.3)

\[
\frac{(t^2 - 1)^2}{\sqrt{t}} > (t-1)^2.
\]  

(2.4)

\[
\frac{(t^2 - 1)^2}{\sqrt{t}} \geq 1 - \sqrt{t}.
\]  

(2.5)

All functions involve in (2.1) to (2.5) are convex and normalized, since \( f''(t) \geq 0 \) \( \forall t > 0 \) and \( f(1) = 0 \) respectively.

Proof:

From (2.1) at \( m=1 \), we have to prove that

\[
\frac{(t^2 - 1)^2}{\sqrt{t}} > \frac{(t-1)^2}{\sqrt{t}}.
\]

Or \( (t+1)^2 > 1 \).

Or \( t^2 + 2t > 0 \).

Which is obvious for \( t > 0 \). Hence proved (2.1).

Now, from (2.1) at \( m=2 \), we have to prove that

\[
\frac{(t^2 - 1)^4}{t^{3/2}} > \frac{(t-1)^4}{t^{3/2}}.
\]

Or \( (t+1)^4 > 1 \).

Which is obvious for \( t > 0 \). Hence proved (2.1).

Similarly, proofs are obvious for \( m=3, 4, 5 \ldots \).

Further, from (2.2) at \( m=1 \), we have to prove that

\[
\frac{(t^2 - 1)^2}{\sqrt{t}} > \frac{(t-1)^2}{t+1}.
\]

Or \( (t+1)^3 > \sqrt{t} \).
Or \((t+1)^3 - \sqrt{t} > 0\).

Which is true (figure 1) for \(t > 0\). Hence proved (2.2).

Similarly, proofs are obvious for \(m=2, 3, 4, \ldots\).

Now, from (2.3) at \(m=1\), we have to prove that

\[
\frac{(t^2 - 1)^2}{\sqrt{t}} e^{\frac{(t^2-1)^2}{t}} > \frac{(t-1)^2}{\sqrt{t}} e^{\frac{(t-1)^2}{t}}.
\]

Or \((t+1)^2 e^{\frac{(t^2-1)^2}{t}} > e^{\frac{(t-1)^2}{t}}\).

Or \((t+1)^2 e^{\frac{(t^2-1)^2}{t} - \frac{(t-1)^2}{t}} > 1\).

Or \((t+1)^2 e^{(t+2)(t-1)^2} > 1\).

Or \((t+1)^2 e^{(t+2)(t-1)^2} - 1 > 0\).

Which is true (figure 2) for \(t > 0\). Hence proved (2.3).

Similarly, proofs are true for \(m=2, 3, 4, \ldots\).

Second lastly, from (2.4), we have to prove that
\[
\frac{(t^2 - 1)^2}{\sqrt{t}} > (t - 1)^2.
\]
Or \( (t + 1)^2 - \sqrt{t} > 0. \)

![Figure 3: Graph of \((t + 1)^2 - \sqrt{t}\)](image)

Which is true (figure 3) for \( t > 0 \). Hence proved (2.4).

Lastly, from (2.5), we have to prove that
\[
\frac{(t^2 - 1)^2}{\sqrt{t}} \geq 1 - \sqrt{t}.
\]
Or \( (t^2 - 1)^2 + \sqrt{t} \geq 1. \)
Or \( (t^2 - 1)^2 + t - \sqrt{t} \geq 0. \)

![Figure 4: Graph of \((t^2 - 1)^2 + t - \sqrt{t}\)](image)

Which is true (figure 4) for \( t > 0 \). Hence proved (2.5).

**Proposition 2.2** Let \((P, Q) \in \Gamma_n \times \Gamma_m\), then we have the followings new inter relations.

\[
\xi_m (P, Q) > E_m^* (P, Q). \quad (2.6)
\]
\[
\xi_m (P, Q) > \Delta_m (P, Q). \quad (2.7)
\]
\[
\omega_m (P, Q) > J_m^* (P, Q). \quad (2.8)
\]
\[ \xi_1(P, Q) > \chi^2(P, Q). \]  
\[ \xi_1(P, Q) \geq h(P, Q). \]  
(2.9)  
(2.10)  

Where \( \xi_m(P, Q), E_m^r(P, Q), \Delta_m(P, Q), \omega_m(P, Q), J_m^r(P, Q), \chi^2(P, Q) \) and \( h(P, Q) \) are given by (1.2), (1.5), (1.10), (1.4), (1.6), (1.18) and (1.17) respectively.

**Proof:**

If we put \( t = \frac{D_i}{q}, i = 1, 2, 3, ..., n \) in (2.1) to (2.5) and multiply by \( q_i \), and then sum over all \( i = 1, 2, 3, ..., n \), we get the desired relations (2.6) to (2.10) respectively.

Now we can easily say from (2.6), (2.7) and (2.8) that

\[ \xi_1(P, Q) > E_1^r(P, Q), \xi_2(P, Q) > E_2^r(P, Q), \xi_3(P, Q) > E_3^r(P, Q), ..., \]  
\[ \xi_1(P, Q) > \Delta_1(P, Q) = \Delta(P, Q), \xi_2(P, Q) > \Delta_2(P, Q), \xi_3(P, Q) > \Delta_3(P, Q), ... \]  
(2.11)  
(2.12)  

and

\[ \omega_1(P, Q) > J_1^r(P, Q), \omega_2(P, Q) > J_2^r(P, Q), \omega_3(P, Q) > J_3^r(P, Q), ... \]  
(2.13)  

respectively.

**Proposition 2.3** Let \((P, Q) \in \Gamma_n \times \Gamma_n\), then we have the following new intra relations.

\[ \xi_m(P, Q) \leq \xi_m(P, Q). \]  
\[ \xi_m(P, Q) \leq \omega_m(P, Q). \]  
(2.14)  
(2.15)  

Where \( m = 1, 2, 3, ... \) and \( \xi_m(P, Q), \xi_m(P, Q) \) and \( \omega_m(P, Q) \) are given by (1.2), (1.3) and (1.4) respectively.

**Proof:**

Since
\[
\frac{(t^2 - 1)^{2m} (t^4 - 2t^2 + t + 1)}{t^{(2m+1)/2}} = \frac{(t^2 - 1)^{2m}}{t^{(2m+1)/2}} + \frac{(t^2 - 1)^{2m+2}}{t^{(2m+1)/2}}
\]

Therefore
\[
\frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} \leq \frac{(t^2 - 1)^{2m} (t^4 - 2t^2 + t + 1)}{t^{(2m+1)/2}}.
\]  
(2.16)  

Similarly
\[
\frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} \leq \exp \left\{ \frac{(t^2 - 1)^2}{t} \right\} = \frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} \left[ 1 + \frac{(t^2 - 1)^2}{t} + \frac{(t^2 - 1)^4}{t^2} + \frac{(t^2 - 1)^6}{t^3} + ... \right]
\]

Therefore
\[
\frac{(t^2 - 1)^{2m}}{t^{(2m-1)/2}} \leq \exp \left\{ \frac{(t^2 - 1)^2}{t} \right\}.
\]  
(2.17)
\( \forall t > 0 \text{ and } m = 1, 2, 3, \ldots \). Here all functions involve in (2.16) to (2.17) are convex and normalized, since \( f''(t) \geq 0 \forall t > 0 \text{ and } f(1) = 0 \) respectively.

Now put \( t = \frac{p_i}{q_i}, i = 1, 2, 3, \ldots, n \) in (2.16) and (2.17) and multiply by \( q_i \), and then sum over all \( i = 1, 2, 3, \ldots, n \), we get the relations (2.14) and (2.15) respectively. Particularly from (2.14) and (2.15), we get the followings.

\[
\xi_1(P, Q) \leq \zeta_1(P, Q), \xi_2(P, Q) \leq \zeta_2(P, Q), \xi_3(P, Q) \leq \zeta_3(P, Q), \ldots \tag{2.18}
\]

and

\[
\xi_1(P, Q) \leq \omega_1(P, Q), \xi_2(P, Q) \leq \omega_2(P, Q), \xi_3(P, Q) \leq \omega_3(P, Q), \ldots \tag{2.19}
\]

respectively.

3. Various new relations on new divergences

In this section, we obtain various new important relations on new divergence measures (1.2) to (1.4) with other standard divergences by taking help of section 2 relations.

**Proposition 3.1** Let \((P, Q) \in \Gamma_n \times \Gamma_m\), then we have the following new inter relations.

\[
2\left[N_i^*(P, Q) - N_i^+(P, Q)\right] \leq 2\Delta(P, Q) \leq 8I(P, Q) \leq 8h(P, Q) \leq J(P, Q) \leq 8T(P, Q) \leq E_i^+(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.1}
\]

\[
H(P, Q) \leq G^+(P, Q) \leq L^+(P, Q) \leq N_1(P, Q) \leq N_3(P, Q) \leq N_2(P, Q) \leq A(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.2}
\]

\[
\frac{1}{4} J^\ast_i(P, Q) \leq K(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.3}
\]

\[
2\Delta(P, Q) - \frac{1}{2} \psi(P, Q) \leq \chi^+(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.4}
\]

\[
\frac{1}{2} \left[\psi^+(P, Q) - \frac{1}{2} J_i^+(P, Q)\right] \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.5}
\]

\[
\left[J_i^+(P, Q) - J_i^+(P, Q)\right] \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.6}
\]

\[
2\left[\Delta(P, Q) + L(P, Q)\right] \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.7}
\]

\[
4M_{SA}(P, Q) \leq \frac{4}{3} M_{SH}(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.8}
\]

\[
\frac{1}{2} M_{SG^+}(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.9}
\]

\[
32d^\ast(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.10}
\]

\[
2F(P, Q) \leq \xi_1(P, Q) \leq \zeta_1(P, Q), \omega_1(P, Q) \right]. \tag{3.11}
\]

\[
6D_{w^\ast}(P, Q) \leq 64D_{w^+}(P, Q) \leq E_i^+(P, Q) \leq \xi_2(P, Q) \leq \zeta_2(P, Q), \omega_2(P, Q) \right]. \tag{3.12}
\]
**Proof:** We know that

\[
\frac{1}{4} \Delta(P, Q) \leq I(P, Q) \leq h(P, Q) \leq \frac{1}{8} J(P, Q) \leq T(P, Q) \leq \frac{1}{8} E_1^+(P, Q) \tag{3.13}
\]

\[
N_1^+(P, Q) - N_2^+(P, Q) \leq \Delta(P, Q) \tag{3.14}
\]

\[
H(P, Q) \leq G^+(P, Q) \leq L(P, Q) \leq N_1(P, Q) \leq N_3(P, Q) \leq N_2(P, Q) \leq A(P, Q) \tag{3.15}
\]

\[
A(P, Q) \leq h(P, Q) \tag{3.16}
\]

\[
\frac{1}{4} J_x(P, Q) \leq K(P, Q) \leq J(P, Q) \tag{3.17}
\]

\[
\Delta(P, Q) \leq \frac{1}{2} \left[ \frac{1}{2} \psi(P, Q) + \chi^+(P, Q) \right] \tag{3.18}
\]

\[
\frac{1}{2} \psi_m(P, Q) \leq E_{1+}^+(P, Q) + \frac{1}{4} J_{1+}^+(P, Q) \tag{3.19}
\]

\[
J_1^+(P, Q) - J_2^+(P, Q) \leq E_{1+}^+(P, Q) \tag{3.20}
\]

\[
\Delta(P, Q) \leq \frac{1}{2} E_{1+}^+(P, Q) - L(P, Q) \tag{3.21}
\]

\[
M_{sh}(P, Q) \leq \frac{1}{3} M_{sh}(P, Q) \leq \frac{1}{4} \Delta(P, Q) \tag{3.22}
\]

\[
\frac{1}{2} M_{sys}^+(P, Q) \leq h(P, Q) \tag{3.23}
\]

\[
4d(P, Q) \leq \frac{1}{8} J(P, Q) \tag{3.24}
\]

\[
F(P, Q) \leq \frac{1}{2} \Delta(P, Q) \tag{3.25}
\]

\[
\frac{1}{4} D_{\psi f}(P, Q) \leq \frac{8}{3} D_{\psi f}(P, Q) \leq \frac{1}{24} E_{2+}^+(P, Q) \tag{3.26}
\]

By taking (3.13), (3.14) and first part of the relations (2.11), (2.18) and (2.19) together, we get the relation (3.1).

By taking (2.10), (3.15), (3.16) and first part of the relations (2.18) and (2.19) together, we get the relation (3.2).

By taking (3.17) and fifth, eighth, ninth elements of the proved relation (3.1) together, we get the relation (3.3).

By taking (2.9), (3.18) and first part of the relations (2.18) and (2.19) together, we get the relation (3.4).

By taking (3.19) and first part of the relations (2.11), (2.18) and (2.19) together, we get the relation (3.5).

By taking (3.20) and first part of the relations (2.11), (2.18) and (2.19) together, we get the relation (3.6).

By taking (3.21) and first part of the relations (2.11), (2.18) and (2.19) together, we get the relation (3.7).
By taking (3.22) and first part of the relations (2.12), (2.18) and (2.19) together, we get the relation (3.8).
By taking (2.10), (3.23) and first part of the relations (2.18) and (2.19) together, we get the relation (3.9).
By taking (3.24) and fifth, eighth, ninth elements of the proved relation (3.1) together, we get the relation (3.10).
By taking (3.25) and first part of the relations (2.12), (2.18) and (2.19) together, we get the relation (3.11).
By taking (3.26) and second part of the relations (2.11), (2.18) and (2.19) together, we get the relation (3.12).

4. Csiszar’s information inequalities and its application

In this section, we are taking well known information inequalities on \( C_f (P, Q) \), such inequalities are for instance needed in order to calculate the relative efficiency of two divergences. This theorem is due to literature [22], which relates two generalized f- divergence measures.

**Theorem 4.1** Let \( f_1, f_2 : I \subset R_+ \rightarrow R \) be two convex and normalized functions, i.e. \( f_1''(t), f_2''(t) \geq 0, t > 0 \) and \( f_1(1) = f_2(1) = 0 \) respectively and suppose the assumptions:

a. \( f_1 \) and \( f_2 \) are twice differentiable on \((\alpha, \beta)\) where \(0 < \alpha \leq 1 \leq \beta < \infty, \alpha \neq \beta\).

b. There exists the real constants \( m, M \) such that \( m < M \) and

\[
m \leq \frac{f_1''(t)}{f_2''(t)} \leq M, \quad f_2''(t) > 0 \forall t \in (\alpha, \beta).
\] (4.1)

If \( P, Q \in \Gamma_n \) and satisfying the assumption \( 0 < \alpha \leq \frac{P_i}{Q_i} \leq \beta < \infty \), then we have the inequalities,

\[
mC_{f_1} (P, Q) \leq C_{f_1} (P, Q) \leq MC_{f_2} (P, Q).
\] (4.2)

Where \( C_f (P, Q) \) is given by (1.1).

Now by using theorem 4.1 or inequalities (4.2), we will get the bounds of \( \xi_1 (P, Q) \) in terms of \( \chi^2 (Q, P) \) and \( K (Q, P) \), where \( \chi^2 (Q, P) \) and \( K (Q, P) \) are adjoint of \( \chi^2 (P, Q) \) and \( K (P, Q) \) respectively.

Firstly, let us consider

\[
f_1(t) = \frac{(t^2 - 1)^2}{\sqrt{t}}, t > 0, f_1(1) = 0 \quad \text{and} \quad f_1'(t) = \frac{(t^2 - 1)(7t^2 + 1)}{2t^{3/2}} \quad \text{and}
\]

\[
f_1''(t) = \frac{(35t^4 - 6t^2 + 3)}{4t^{5/2}}.
\] (4.3)

Put \( f_1(t) \) in (1.1), we get

\[
C_{f_1} (P, Q) = \sum_{i=1}^{n} \left( \frac{p_i^2 - q_i^2}{p_i q_i} \right)^2 = \xi_1 (P, Q).
\] (4.4)
(Properties of \( \xi_m(P, Q) \) have been discussed in literature [6], in detail. So we are skipping that.)

**Proposition 4.1** Let \( \xi_1(P, Q) \) and \( \chi(P, Q) \) be defined as in (4.4) and (1.18) respectively. For \( P, Q \in \Gamma_n \), we have
\[
\frac{35\alpha^{9/2} - 6\alpha^{5/2} + 3\alpha^{1/2}}{8} \leq \xi_1(P, Q) \leq \frac{35\beta^{9/2} - 6\beta^{5/2} + 3\beta^{1/2}}{8} \chi(Q, P).
\]  
(4.5)

**Proof:** Let us consider
\[
f_2(t) = \left(\frac{t-1}{t}\right)^2, t \in (0, \infty), f_2(1) = 0, f_2'(t) = \frac{(t^2 - 1)}{t^2}
\]
and
\[
f_2''(t) = \frac{2}{t^3}.
\]
(4.6)

Since \( f_2''(t) > 0 \) \( \forall \; t > 0 \) and \( f_2(1) = 0 \), so \( f_2(t) \) is convex and normalized function respectively. Now put \( f_2(t) \) in (1.1), we get,
\[
C_{f_2}(P, Q) = \sum_{i=1}^{n} \left( \frac{p_i - q_i}{p_i} \right)^2 = \chi^2(Q, P).
\]  
(4.7)

Now, let \( g(t) = \frac{f_2'(t)}{f_2(t)} = \frac{35t^{9/2} - 6t^{5/2} + 3t^{1/2}}{8} \), where \( f_2'(t) \) and \( f_2''(t) \) are given by (4.3) and (4.6) respectively.
And \( g'(t) = \frac{3(105t^4 - 10t^2 + 1)}{16t^{12}} > 0 \).

![Figure 5: Graph of \( g'(t) \)](image)

It is clear by figure 5 of \( g'(t) \) that \( g'(t) > 0, t > 0 \) or \( g(t) \) is always increasing in \( (0, \infty) \), so
\[
m = \inf_{\alpha \in (a, \beta)} g(t) = g(\alpha) = \frac{35\alpha^{9/2} - 6\alpha^{5/2} + 3\alpha^{1/2}}{8}.
\]  
(4.8)

\[
M = \sup_{\beta \in (a, \beta)} g(t) = g(\beta) = \frac{35\beta^{9/2} - 6\beta^{5/2} + 3\beta^{1/2}}{8}.
\]  
(4.9)

The result (4.5) is obtained by using (4.4), (4.7), (4.8) and (4.9) in (4.2).
Proposition 4.2 Let $\xi_1(P, Q)$ and $K(P, Q)$ be defined as in (4.4) and (1.19) respectively. For $P, Q \in \Gamma_n$, we have

a. If $0 < \alpha \leq 0.3916$, then
\[
1.158 K(Q, P) \leq \xi_1(P, Q) \leq \max \left[ \frac{35\alpha^{7/2} - 6\alpha^{3/2} + 3\alpha^{-1/2}}{4}, \frac{35\beta^{7/2} - 6\beta^{3/2} + 3\beta^{-1/2}}{4} \right] K(Q, P). \tag{4.10}
\]

b. If $0.3916 < \alpha \leq 1$, then
\[
\frac{35\alpha^{7/2} - 6\alpha^{3/2} + 3\alpha^{-1/2}}{4} K(Q, P) \leq \xi_1(P, Q) \leq \frac{35\beta^{7/2} - 6\beta^{3/2} + 3\beta^{-1/2}}{4} K(Q, P). \tag{4.11}
\]

**Proof:** Let us consider
\[
f_2(t) = -\log t, t \in (0, \infty), f_2(1) = 0, f_2'(t) = -\frac{1}{t^2}, \quad f_2''(t) = \frac{1}{t^2}. \tag{4.12}
\]

Since $f_2''(t) > 0 \forall t > 0$ and $f_2(1) = 0$, so $f_2(t)$ is convex and normalized function respectively. Now put $f_2(t)$ in (1.1), we get,
\[
\xi_1(P, Q) = \frac{1}{n} \sum_{i=1}^{n} q_i \log \frac{q_i}{p_i} = K(Q, P). \tag{4.13}
\]

Now, let $g(t) = \frac{f_2''(t)}{f_2'(t)} = \frac{35t^{7/2} - 6t^{3/2} + 3t^{-1/2}}{4}$, where $f_2''(t)$ and $f_2'(t)$ are given by (4.3) and (4.12) respectively and $g'(t) = \frac{245t^2 - 18t^2 - 3}{8t^{3/2}}$.

![Figure 6: Graph of $g'(t)$](image-url)

It is clear by figure 6 of $g'(t)$ that $g'(t) < 0$ in $(0, 0.3916)$ and $g'(t) \geq 0$ in $[0.3916, \infty)$ i.e. $g(t)$ is decreasing in $(0, 0.3916)$ and increasing in $[0.3916, \infty)$. So $g(t)$ has a minimum value at $t=0.3916$ because $g''(0.3916) > 0$. Now

a. If $0 < \alpha \leq 0.3916$, then
\[
m = \inf_{\alpha, \beta} g(t) = g(0.3916) = 1.158. \quad (4.14)
\]
\[
M = \sup_{\alpha, \beta} g(t) = \max\{g(\alpha), g(\beta)\}
\]
\[
= \max\left[ \frac{35\alpha^{\frac{3}{2}} - 6\alpha^{\frac{3}{2}} + 3\alpha^{-\frac{1}{2}}}{4}, \frac{35\beta^{\frac{3}{2}} - 6\beta^{\frac{3}{2}} + 3\beta^{-\frac{1}{2}}}{4} \right]. \quad (4.15)
\]

b. If \(0.3916 < \alpha \leq 1\), then
\[
m = \inf_{\alpha, \beta} g(t) = g(\alpha) = \frac{35\alpha^{\frac{3}{2}} - 6\alpha^{\frac{3}{2}} + 3\alpha^{-\frac{1}{2}}}{4}. \quad (4.16)
\]
\[
M = \sup_{\alpha, \beta} g(t) = g(\beta) = \frac{35\beta^{\frac{3}{2}} - 6\beta^{\frac{3}{2}} + 3\beta^{-\frac{1}{2}}}{4}. \quad (4.17)
\]

The results (4.10) and (4.11) are obtained by using (4.4), (4.13), (4.14), (4.15), (4.16) and (4.17) in (4.2).

5. **Numerical illustration**

In this section, we give two examples for calculating the divergences \(\xi_1(P,Q), \chi^2(Q,P)\) and \(K(Q,P)\) and verify the inequalities (4.5) and (4.10) or verify bounds of \(\xi_1(P,Q)\).

**Example 5.1** Let \(P\) be the binomial probability distribution with parameters \((n=10, p=0.5)\) and \(Q\) its approximated Poisson probability distribution with parameter \((\lambda = np = 5)\), then for the discrete random variable \(X\), we have

<table>
<thead>
<tr>
<th>(x_i)</th>
<th>(0)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p(x_i) = p_i)</td>
<td>.009976</td>
<td>.00976</td>
<td>.043</td>
<td>.117</td>
<td>.205</td>
<td>.246</td>
<td>.205</td>
<td>.117</td>
<td>.043</td>
<td>.00976</td>
<td>.00976</td>
</tr>
<tr>
<td>(q(x_i) = q_i)</td>
<td>.00673</td>
<td>.033</td>
<td>.084</td>
<td>.140</td>
<td>.175</td>
<td>.175</td>
<td>.146</td>
<td>.104</td>
<td>.065</td>
<td>.036</td>
<td>.018</td>
</tr>
<tr>
<td>(\frac{p_i}{q_i})</td>
<td>.1450</td>
<td>.2957</td>
<td>.5119</td>
<td>.8357</td>
<td>1.171</td>
<td>1.405</td>
<td>1.404</td>
<td>1.125</td>
<td>.6615</td>
<td>.2711</td>
<td>.0542</td>
</tr>
</tbody>
</table>

By using the above table 1, we get the followings.
\[
\alpha \left(= 0.0542 \right) \leq \frac{p_i}{q_i} \leq \beta \left(= 1.405 \right). \quad (5.1)
\]
\[
\xi_1(P,Q) = \sum_{i=0}^{10} \left(\frac{p_i^2 - q_i^2}{p_i q_i} \right)^2 = 0.5928. \quad (5.2)
\]
\[
K(Q,P) = \sum_{i=1}^{11} q_i \log \left( \frac{q_i}{p_i} \right) = 0.11177. \quad (5.3)
\]
\[ \chi^2(Q, P) = \sum_{i=1}^{11} \left( \frac{p_i - q_i}{p_i} \right)^2 \approx 0.5548. \quad (5.4) \]

Put the approximated numerical values from (5.1) to (5.4) in (4.5) and (4.10) and get the followings respectively.

\[ 0.04815 \leq \xi_1(Q, P) = 0.5928 \leq 10.4843. \text{ From (4.5)} \]
\[ 0.1294 \leq \xi_2(Q, P) = 0.5928 \leq 3.0066. \text{ From (4.10)} \]
Hence verified the inequalities (4.5) and (4.10) for \( p=0.5 \).

**Example 5.2** Let \( P \) be the binomial probability distribution with parameters (\( n=10, p=0.7 \)) and \( Q \) its approximated Poisson probability distribution with parameter (\( \lambda = np = 7 \)), then for the discrete random variable \( X \), we have

<table>
<thead>
<tr>
<th>( x_i = p_i )</th>
<th>( q(x_i) = q_i )</th>
<th>( \frac{p_i}{q_i} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.0000059</td>
<td>0.00647</td>
</tr>
<tr>
<td>9</td>
<td>0.000137</td>
<td>0.0214</td>
</tr>
<tr>
<td>8</td>
<td>0.00144</td>
<td>0.0654</td>
</tr>
<tr>
<td>7</td>
<td>0.009</td>
<td>0.173</td>
</tr>
<tr>
<td>6</td>
<td>0.036</td>
<td>0.395</td>
</tr>
<tr>
<td>5</td>
<td>0.102</td>
<td>0.871</td>
</tr>
<tr>
<td>4</td>
<td>0.266</td>
<td>1.005</td>
</tr>
<tr>
<td>3</td>
<td>0.233</td>
<td>1.785</td>
</tr>
<tr>
<td>2</td>
<td>0.121</td>
<td>1.792</td>
</tr>
<tr>
<td>1</td>
<td>0.0282</td>
<td>1.198</td>
</tr>
<tr>
<td>0</td>
<td>0.397</td>
<td></td>
</tr>
</tbody>
</table>

By using the above table 2, we get the followings.

\[ \alpha = 0.00647 \leq \frac{P}{Q} \leq \beta (= 1.792). \quad (5.5) \]
\[ \xi_1(Q, P) = \sum_{i=1}^{11} \left( \frac{p_i^2 - q_i^2}{p_i q_i} \right) = 1.5703. \quad (5.6) \]
\[ K(Q, P) = \sum_{i=1}^{11} q_i \log \left( \frac{q_i}{p_i} \right) = 0.2461. \quad (5.7) \]
\[ \chi^2(Q, P) = \sum_{i=1}^{11} \left( \frac{p_i - q_i}{p_i} \right)^2 \approx 1.2259. \quad (5.8) \]

Put the approximated numerical values from (5.5) to (5.8) in (4.5) and (4.10) and get the followings respectively.

\[ 0.03697 \leq \xi_1(Q, P) = 1.5703 \leq 70.700. \text{ From (4.5)} \]
\[ 0.2849 \leq \xi_2(Q, P) = 1.5703 \leq 15.8406. \text{ From (4.10)} \]
Hence verified the inequalities (4.5) and (4.10) for \( p=0.7 \).

6. **Conclusion**
In this work, we derived some new intra relations and new inter relations of divergence measures \( \xi_m(P, Q), \zeta_m(P, Q) \) and \( \omega_m(P, Q) \) for \( m = 1, 2, 3 \ldots \) by using some algebraic inequalities and some standard relations of several standard divergences.

Related works done so far and future scope with limitations:

Many research papers have been studied by I.J. Taneja, P. Kumar, S.S. Dragomir, K.C. Jain and others, who gave the idea of divergence measures, their properties, their bounds and relations with other measures. Especially I. J. Taneja and K. C. Jain did a lot of quality work in this field. Such instance, Taneja derived bounds on different non-symmetric divergences in terms of different symmetric divergences and vice versa. He also introduced new generalized divergences, new divergences as a result of difference of means and characterized their properties and so on. Similarly Jain introduced a new generalized f-divergence measure by which many standard and some new divergences have been obtained. He defined and characterized its properties, derived many new information inequalities and obtained several new relations on well known divergence measures.

We also found in our previous article [6] that square root of some particular divergences of Csiszar’s class is a metric space, 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. Also we can use divergences in fuzzy mathematics as fuzzy directed divergences and fuzzy entropies which are very useful to find amount of average ambiguity or difficulty in making a decision whether an element belongs to a set or not. Such types of divergences are also very useful to find 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.

For getting the information divergence measures from Csiszar’s f-divergence and Jain’s f-divergence, the function must be convex in the interval \((0, \infty)\) because the function must satisfy certain conditions and probability is always positive. Therefore, we cannot take the concave functions. This is the limitation of this area.

Note: In view of theorem 2.1, all divergence measures from (1.2) to (1.21), (1.33) and (1.34) are both non-negative and convex in the pair of probability distribution \((P, Q) \in \Gamma_n \times \Gamma_n\).

7. References


[17] Pearson K., “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), pp: 157-172.


Authors

K. C. Jain, Professor in Department of Mathematics, Malaviya National Institute of Technology, Jaipur- 302017 (Rajasthan), India. Ph: +91- 9314682159

Praphull Chhabra ,Research Scholar in Department of Mathematics ,Malaviya National Institute of Technology Jaipur- 302017 (Rajasthan), India. Ph: +91- 9783372873, 8952826855