ON A DIFFERENCE OF JENSEN INEQUALITY AND ITS APPLICATIONS TO MEAN DIVERGENCE MEASURES

INDER JEET TANEJA

Abstract. In this paper we have considered a difference of Jensen’s inequality for convex functions and proved some of its properties. In particular, we have obtained results for Csiszar’s f-divergence. A result is established that allow us to compare two measures under certain conditions. By the application of this result we have obtained a new inequality for the well known means such as arithmetic, geometric and harmonic. Some divergence measures based on these means are also defined.

1. Jensen Difference

Let

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

be the set of all complete finite discrete probability distributions.

Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be a differentiable convex function on the interval $I$, $x_i \in \overset{\circ}{I}$ ($\overset{\circ}{I}$ is the interior of $I$). Let $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n) \in \Gamma_n$, then it is well known that

$$f \left( \sum_{i=1}^{n} \lambda_i x_i \right) \leq \sum_{i=1}^{n} \lambda_i f(x_i).$$

The above inequality is famous as Jensen inequality. If $f$ is concave, the inequality sign changes.

Let us consider the following Jensen difference:

$$F_f(\lambda, X) = \sum_{i=1}^{n} \lambda_i f(x_i) - f \left( \sum_{i=1}^{n} \lambda_i x_i \right),$$

Here below we shall give two theorems giving properties of Jensen difference.

Theorem 1. Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be a differentiable convex function on the interval $I$, $x_i \in \overset{\circ}{I}$ ($\overset{\circ}{I}$ is the interior of $I$), $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n) \in \Gamma_n$. If $\eta_1, \eta_2 \in \overset{\circ}{I}$ and $\eta_1 \leq x_i \leq \eta_2$, $\forall i = 1, 2, ..., n$, then we have the inequalities:

$$0 \leq F_f(\lambda, X) \leq L_f(\lambda, X) \leq Z_f(\eta_1, \eta_2),$$

Key words and phrases. Jensen difference; Divergence measures; Csiszar’s f-divergence; Convex function; Mean inequalities.

This paper is a part of author’s chapter to appear in ’Advances in Imaging and Electron Physics’, 2005, Elsevier Publication.
where

\[ L_f(\lambda, X) = \sum_{i=1}^{n} \lambda_i x_i f'(x_i) - \left( \sum_{i=1}^{n} \lambda_i x_i \right) \left( \sum_{i=1}^{n} \lambda_i f'(x_i) \right) \]

and

\[ Z_f(\eta_1, \eta_2) = \frac{1}{4} (\eta_2 - \eta_1) \left[ f'(\eta_2) - f'(\eta_1) \right]. \]

The above theorem is due to Dragomir \[10\]. It has been applied by many authors \[9, 13\]. The measure \( F(\lambda, X) \) has been extensively studied by Burbea and Rao \[3, 4\]. As a consequence of above theorem we have the following corollary.

**Corollary 1.** For all \( a, b, \nu, \omega \in (0, \infty) \), the following inequalities hold:

\[ 0 \leq \frac{vf(a) + \omega f(b)}{v + \omega} - f \left( \frac{va + \omega b}{v + \omega} \right) \leq \frac{va f'(a) + \omega b f'(b)}{v + \omega} - \left( \frac{va + \omega b}{v + \omega} \right) \left( \frac{vf'(a) + \omega f'(b)}{v + \omega} \right) \leq \frac{1}{4} (b - a) \left( f'(b) - f'(a) \right). \]

**Proof.** It follows from Theorem \[\ref{thm1}\] by taking \( \lambda_1 = \frac{\nu}{v + \omega}, \lambda_2 = \frac{\omega}{v + \omega}, \lambda_3 = \ldots = \lambda_n = 0, \)

\( x_1 = a, x_2 = b, x_2 = \ldots = x_n = 0. \)

Now we shall give some examples of Theorem \[\ref{thm1}\].

**Example 1.** For all \( x \in (0, \infty) \), let us consider a function

\[ f_s(x) = \begin{cases} \frac{1-x^s}{s}, & s \neq 0, \\ -\ln x, & s = 0. \end{cases} \]

We can easily check that the function \( f_s(x) \) is convex in \((0, \infty)\) for all \( s \leq 1 \). Let there exist \( \eta_1 \) and \( \eta_2 \) such that \( \eta_1 \leq x_i \leq \eta_2, \forall i = 1, 2, \ldots, n \). Applying Theorem \[\ref{thm1}\] for the function \( f_s(x) \), we have

\[ 0 \leq F_s(\lambda, X) \leq Z_s(\eta_1, \eta_2), \quad s \leq 1, \]

where

\[ F_s(\lambda, X) = \begin{cases} \frac{1}{s} \left[ \left( \sum_{i=1}^{n} \lambda_i x_i \right)^s - \sum_{i=1}^{n} \lambda_i x_i^s \right], & s \neq 0, \\ \ln \left( \frac{A(\lambda, X)}{G(\lambda, X)} \right), & s = 0. \end{cases} \]

\[ A(\lambda, X) = \sum_{i=1}^{n} \lambda_i x_i, \]

\[ G(\lambda, X) = \prod_{i=1}^{n} x_i^{\lambda_i}. \]
and

\[ Z_s(\alpha, \beta) = \frac{1}{4}(\eta_2 - \eta_1)(\eta_1^{s-1} - \eta_2^{s-1}). \]

In particular we have

\[ \frac{A(\lambda, X)}{G(\lambda, X)} \leq \exp \left[ \frac{(\eta_2 - \eta_1)^2}{4\eta_1\eta_2} \right], \quad \eta_1 \leq x_i \leq \eta_2, \quad \forall i = 1, 2, \ldots n. \]

The result (13) is due to Dragomir [10]. The following proposition is a particular case of the inequalities (6) and gives bounds on Burbea and Rao’s [3, 4] Jensen difference divergence measure.

**Proposition 1.** Let \( f : (0, \infty) \to \mathbb{R} \) be a differentiable convex function. Then for all \( P, Q \in \Gamma_n \), we have

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

**Proof.** Take \( \omega = \nu = \frac{1}{2} \) in (6), we get

\[ 0 \leq \frac{f(a) + f(b)}{2} - f \left( \frac{a + b}{2} \right) \leq \frac{1}{4} (b - a) \left[ f'(b) - f'(a) \right]. \]

Replace in (13), \( a \) by \( p_i \) and \( b \) by \( q_i \), and sum over all \( i = 1, 2, \ldots, n \), we get the required result. \( \square \)

**Example 2.** Let us consider a convex function

\[ \phi_s(x) = \begin{cases} &\left[ s(s-1) \right]^{-1} [x^s - 1 - s(x - 1)], \quad s \neq 0, 1, \\ &x - 1 - \ln x, \quad s = 0, \\ &1 - x + x \ln x, \quad s = 1, \end{cases} \]

for all \( x \in (0, \infty) \) and \( s \in (-\infty, \infty) \). Then from (14), we get

\[ 0 \leq W_s(P||Q) \leq \frac{1}{4} \varphi_s(P||Q), \]

where

\[ I_s(P||Q) = \left[ s(s-1) \right]^{-1} \sum_{i=1}^{n} \left[ \frac{p_i^s + q_i^s}{2} - \frac{(p_i + q_i)^s}{2} \right], \quad s \neq 0, 1, \]

\[ W_s(P||Q) = \begin{cases} &I_0(P||Q) = \ln \left[ \prod_{i=1}^{n} \left( \frac{p_i + q_i}{2\sqrt{p_i q_i}} \right) \right], \quad s = 0, \\ &I(P||Q) = H \left( \frac{P + Q}{2} \right) - \frac{H(P) + H(Q)}{2}, \quad s = 1, \end{cases} \]
and

\[ J_s(P||Q) = \begin{cases} 
\frac{1}{(s-1)} \sum_{i=1}^{n} (p_i - q_i) (p_i^{s-1} - q_i^{s-1}), & s \neq 0,1, \\
J_0(P||Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i q_i}, & s = 0, \\
J(P||Q) = \sum_{i=1}^{n} (p_i - q_i) \ln \left( \frac{p_i}{q_i} \right), & s = 1.
\]  

The expression \( H(P) = -\sum_{i=1}^{n} p_i \ln p_i \), appearing in (18) is the well known Shannon’s entropy. The expression \( J_s(P||Q) \) appearing in (19) is Jeffreys-Kullback-Leibler’s J-divergence (ref. Jeffreys [16] and Kullback and Leibler [17]). The expression \( J_0(P||Q) \) is due to Burbea and Rao [3]. The measures (18) and (19) has been studied by Burbea and Rao [3] only for positive values of the parameters. Some studies on these generalised measures can be seen in Taneja [18, 20]. Here we have presented them for all \( s \in (-\infty, \infty) \). The function given in (16) is due to Cressie and Read [7].

**Proposition 2.** Let \( f : \mathbb{R}_+ \rightarrow \mathbb{R} \) be differentiable convex and normalized i.e., \( f(1) = 0 \). If \( P, Q \in \Gamma_n \), are such that \( 0 < r \leq \frac{p_i}{q_i} \leq R < \infty \), \( \forall i \in \{1,2,\ldots,n\} \), for some \( r \) and \( R \) with \( 0 < r \leq 1 \leq R < \infty \), then we have

\[ 0 \leq C_f(P||Q) \leq E_{C_f}(P||Q) \leq A_{C_f}(r,R), \tag{20} \]

and

\[ 0 \leq C_f(P||Q) \leq B_{C_f}(r,R) \leq A_{C_f}(r,R), \tag{21} \]

where

\[ C_f(P||Q) = \sum_{i=1}^{n} q_i f \left( \frac{p_i}{q_i} \right), \tag{22} \]

\[ E_{C_f}(P||Q) = \sum_{i=1}^{n} (p_i - q_i) f' \left( \frac{p_i}{q_i} \right), \tag{23} \]

\[ A_{C_f}(r,R) = \frac{1}{4} (R - r) (f'(R) - f'(r)) \tag{24} \]

and

\[ B_{C_f}(r,R) = \frac{(R - 1) f(r) + (1 - r) f(R)}{R - r}. \tag{25} \]

The inequalities (20) follow in view of (18). The inequalities (21) follow in view of (6). For details refer to Taneja [22]. The above proposition is an improvement over the work of Dragomir [11, 12]. The measure (22) is known as Csizsár’s \( f \)-divergence.
Example 3. Under the conditions of Proposition 2, the inequalities (20) and (21) for the function (16) are given by

\[ 0 \leq \Phi_s(P||Q) \leq E_{\Phi_s}(P||Q) \leq A_{\Phi_s}(r, R) \]

and

\[ 0 \leq \Phi_s(P||Q) \leq B_{\Phi_s}(r, R) \leq A_{\Phi_s}(r, R), \]

where

\[
\Phi_s(P||Q) = \begin{cases} K_s(P||Q) = [s(s-1)]^{-1} \left[ \sum_{i=1}^{n} p_i^s q_i^{1-s} - 1 \right], & s \neq 0,1, \\ K(Q||P) = \sum_{i=1}^{n} q_i \ln \left( \frac{q_i}{p_i} \right), & s = 0, \\ K(P||Q) = \sum_{i=1}^{n} p_i \ln \left( \frac{p_i}{q_i} \right), & s = 1, \end{cases}
\]

\[
E_{\Phi_s}(P||Q) = \begin{cases} (s-1)^{-1} \sum_{i=1}^{n} (p_i - q_i) \left( \frac{p_i}{q_i} \right)^{s-1}, & s \neq 1, \\ \sum_{i=1}^{n} (p_i - q_i) \ln \left( \frac{p_i}{q_i} \right), & s = 1, \end{cases}
\]

\[
A_{\Phi_s}(r, R) = \frac{1}{4} \begin{cases} \frac{(R-r)(R^{s-1} - r^{s-1})}{4(s-1)}, & s \neq 1, \\ \frac{1}{4} (R - r) \ln \left( \frac{R}{r} \right), & s = 1, \end{cases}
\]

and

\[
B_{\Phi_s}(r, R) = \begin{cases} \frac{(R-1)(r^{s-1} + (1-r)(R^{s-1})}{(R-r) s(s-1)}, & s \neq 0,1, \\ \frac{(R-1) \ln \frac{1}{(R-r)} \ln \frac{1}{R}}{(R-r)}, & s = 0, \\ \frac{(R-1) \ln r + (1-r) \ln R}{(R-r)} \ln R, & s = 1. \end{cases}
\]

The measure \( K(P||Q) \) appearing in (28) is the well known Kullback-Leibler’s relative information. The measure \( \Phi_s(P||Q) \) given in (28) has been extensively studied in [21], [23].

Theorem 2. Let \( f_1, f_2 : [a, b] \subset \mathbb{R}_+ \rightarrow \mathbb{R} \) be twice differentiable functions on \((a, b)\) and there are \( \alpha \) and \( \beta \) such that

\[ \alpha \leq \frac{f_1''(x)}{f_2''(x)} \leq \beta, \quad \forall x \in (a, b), \quad f_2''(x) > 0 \]
If \( x_i \in [a, b] \) and \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \Gamma_n \), then

\[
(33) \quad \alpha F_{f_2}(\lambda, X) \leq F_{f_1}(\lambda, X) \leq \beta F_{f_2}(\lambda, X).
\]

\[
(34) \quad \alpha [L_{f_2}(\lambda, X) - F_{f_2}(\lambda, X)] \leq L_{f_1}(\lambda, X) - F_{f_1}(\lambda, X) \leq \beta [L_{f_2}(\lambda, X) - F_{f_2}(\lambda, X)]
\]

and

\[
(35) \quad \alpha [Z_{f_2}(\eta_1, \eta_2) - F_{f_2}(\lambda, X)] \leq Z_{f_1}(\eta_1, \eta_2) - F_{f_1}(\lambda, X) \leq \beta [Z_{f_2}(\eta_1, \eta_2) - F_{f_2}(\lambda, X)].
\]

Proof. Consider the mapping \( g : [a, b] \to \mathbb{R} \), defined by

\[
(36) \quad g(x) = f_1(x) - \alpha f_2(x), \quad \forall x \in [a, b],
\]

where the functions \( f_1 \) and \( f_2 \) satisfy the condition (32). Then the function \( g \) is twice differentiable on \((a, b)\). This gives

\[
g'(x) = f'_1(x) - \alpha f'_2(x)
\]

and

\[
g''(x) = f''_1(x) - \alpha f''_2(x) = f''_2(x) \left( \frac{f''_1(x)}{f''_2(x)} - \alpha \right) \geq 0, \quad \forall x \in (a, b).
\]

The above expression shows that \( g \) is convex on \([a, b]\). Applying Jensen inequality for the convex function \( g \) one gets

\[
g \left( \sum_{i=1}^{n} \lambda_i x_i \right) \leq \sum_{i=1}^{n} \lambda_i g(x_i),
\]

i.e.,

\[
f_1 \left( \sum_{i=1}^{n} \lambda_i x_i \right) - \alpha f_2 \left( \sum_{i=1}^{n} \lambda_i x_i \right) \leq \sum_{i=1}^{n} \lambda_i [f_1(x_i) - \alpha f_2(x)],
\]

i.e.,

\[
(37) \quad \alpha \left[ \sum_{i=1}^{n} \lambda_i f_2(x_i) - f_2 \left( \sum_{i=1}^{n} \lambda_i x_i \right) \right] \leq \sum_{i=1}^{n} \lambda_i f_1(x_i) - f_1 \left( \sum_{i=1}^{n} \lambda_i x_i \right).
\]

The expression (37) gives the l.h.s. of the inequalities (33).

Again consider the mapping \( k : [a, b] \to \mathbb{R} \) given by

\[
(38) \quad k(x) = \beta f_2(x) - f_1(x),
\]

and proceeding on similar lines as before, we get the proof of the r.h.s. of the inequalities (33).

Now we shall prove the inequalities (34). Applying the inequalities (3) for the convex function \( g \) given by (36), we get

\[
F_g(\lambda, X) \leq L_g(\lambda, X) \leq Z_g(\eta_1, \eta_2).
\]
i.e.,

\[(39) \quad F_1(\lambda, X) - \alpha F_2(\lambda, X) \leq L_1(\lambda, X) - \alpha L_2(\lambda, X) \leq Z_1(\lambda, X) - \alpha F_2(\eta_1, \eta_2).\]

Simplifying the first inequality of (39) we get

\[(40) \quad \alpha [L_2(\lambda, X) - F_2(\lambda, X)] \leq L_1(\lambda, X) - F_1(\lambda, X).\]

Again simplifying the last inequality of (39) we get

\[(41) \quad \alpha [Z_2(\eta_1, \eta_2) - F_2(\lambda, X)] \leq Z_1(\eta_1, \eta_2) - F_1(\lambda, X).\]

The expressions (40) and (41) complete the first part of the inequalities (34) and (35) respectively. The last part of the inequalities (34) and (35) follows by considering the function \(k(x)\) given by (38) over the inequalities (3).

\[\square\]

Particular cases of above theorem can be seen in \[1\], \[8\], \[9\], \[14\]. Applications of the above theorem for the Csiszár’s \(f\)-divergence are given in the following proposition.

**Proposition 3.** Let \(f_1, f_2 : I \subset \mathbb{R}_+ \to \mathbb{R}\) be two normalized convex mappings, i.e., \(f_1(1) = f_2(1) = 0\) and suppose the assumptions:

(i) \(f_1\) and \(f_2\) are twice differentiable on \((r, R)\), where \(0 < r \leq 1 \leq R < \infty\);

(ii) there exists the real constants \(\alpha, \beta\) such that \(\alpha < \beta\) and

\[(42) \quad \alpha \leq \frac{f''_1(x)}{f''_2(x)} \leq \beta, \quad f''_2(x) > 0, \quad \forall x \in (r, R).\]

If \(P, Q \in \Gamma_n\) are discrete probability distributions satisfying the assumption

\[0 < r \leq \frac{p_i}{q_i} \leq R < \infty,\]

then we have the inequalities:

\[(43) \quad \alpha C_{f_1}(P || Q) \leq C_{f_1}(P || Q) \leq \beta C_{f_2}(P || Q),\]

\[(44) \quad \alpha [E_{f_2}(P || Q) - C_{f_2}(P || Q)] \leq E_{f_1}(P || Q) - C_{f_1}(P || Q) \leq \beta [E_{f_2}(P || Q) - C_{f_2}(P || Q)],\]

\[(45) \quad \alpha [A_{f_2}(r, R) - C_{f_2}(P || Q)] \leq A_{f_1}(r, R) - C_{f_1}(P || Q) \leq \beta [A_{f_2}(r, R) - C_{f_2}(P || Q)],\]

and

\[(46) \quad \alpha [B_{f_2}(r, R) - C_{f_2}(P || Q)] \leq B_{f_1}(r, R) - C_{f_1}(P || Q) \leq \beta [B_{f_2}(r, R) - C_{f_2}(P || Q)].\]

**Proof.** It is an immediate consequence of the Theorem \(\square\)
2. Applications to Mean Divergence Measures

Let us consider the following mean of order \( t \):

\[
D_t(a, b) = \begin{cases} 
\left( \frac{a^t + b^t}{2} \right)^{1/t}, & t \neq 0, \\
\sqrt{ab}, & t = 0, \\
\max\{a, b\}, & t = \infty, \\
\min\{a, b\}, & t = -\infty,
\end{cases}
\]

for all \( a, b > 0 \) and \( t \in \mathbb{R} \). In particular, we have

\[
D_{-1}(a, b) = H(a, b) = \frac{2}{\frac{1}{a} + \frac{1}{b}} = \frac{2ab}{a + b} = A(a^{-1}, b^{-1})^{-1},
\]

\[
D_0(a, b) = G(a, b) = \sqrt{ab} = \sqrt{A(a, b)H(a, b)},
\]

\[
D_{1/2}(a, b) = N_1(a, b) = \left( \frac{\sqrt{a} + \sqrt{b}}{2} \right)^2 = A\left( \frac{\sqrt{a} + \sqrt{b}}{2} \right)^2,
\]

and

\[
D_1(a, b) = A(a, b) = \frac{a + b}{2},
\]

where \( H(a, b) \), \( G(a, b) \) and \( A(a, b) \) are the well known harmonic, geometric and arithmetic means respectively. It is well known [2] that the mean of order \( t \) given in (47) is monotonically increasing in \( t \), then we can write

\[
D_{-1}(a, b) \leq D_0(a, b) \leq D_{1/2}(a, b) \leq D_1(a, b),
\]

or equivalently,

\[
H(a, b) \leq G(a, b) \leq N_1(a, b) \leq A(a, b).
\]

We can easily check that the function \( f(x) = -x^{1/2} \) is convex in \((0, \infty)\). This allows us to conclude the following inequality:

\[
\frac{\sqrt{a} + \sqrt{b}}{2} \leq \sqrt{\frac{a + b}{2}}.
\]

From (49), we can easily derive that

\[
\left( \frac{\sqrt{a} + \sqrt{b}}{2} \right)^2 \leq \left( \frac{\sqrt{a} + \sqrt{b}}{2} \right) \left( \sqrt{\frac{a + b}{2}} \right) \leq \frac{a + b}{2}.
\]

Finally, the expressions (48) and (50) lead us to following inequalities:

\[
H(a, b) \leq G(a, b) \leq N_1(a, b) \leq N_2(a, b) \leq A(a, b),
\]

where

\[
N_2(a, b) = \left( \frac{\sqrt{a} + \sqrt{b}}{2} \right) \left( \sqrt{\frac{a + b}{2}} \right).
\]
Let $P, Q \in \Gamma_n$. In (51), replace $a$ by $p_i$ and $b$ by $q_i$ sum over all $i = 1, 2, ...n$ we get
\begin{equation}
H(P||Q) \leq G(P||Q) \leq N_1(P||Q) \leq N_2(P||Q) \leq 1.
\end{equation}

Based on inequalities (52), we shall build some mean divergence measures. Let us consider the following differences:
\begin{align}
M_{AG}(P||Q) &= 1 - G(P||Q), \\
M_{AH}(P||Q) &= 1 - H(P||Q), \\
M_{AN_2}(P||Q) &= 1 - N_2(P||Q), \\
M_{N_2G}(P||Q) &= N_2(P||Q) - G(P||Q),
\end{align}
and
\begin{equation}
M_{N_2N_1}(P||Q) = N_2(P||Q) - N_1(P||Q).
\end{equation}

We can easily verify that
\begin{align}
M_{AG}(P||Q) &= 1 - G(P||Q) \\
&= 2 [N_1(P||Q) - G(P||Q)] := 2M_{N_1G}(P||Q) \\
&= 2 [1 - N_1(P||Q)] := 2M_{AN_1}(P||Q).
\end{align}

We can also write
\begin{equation}
M_{AG}(P||Q) = 1 - G(P||Q) := h(P||Q)
\end{equation}
and
\begin{equation}
M_{AH}(P||Q) = 1 - H(P||Q) := \frac{1}{2} \Delta(P||Q),
\end{equation}
where $h(P||Q)$ is the well known Hellinger’s discrimination and $\Delta(P||Q)$ is known by triangular discrimination. These two measures are well known in the literature of statistics. The measure $M_{AN_2}(P||Q)$ is new and has been recently studied by Taneja [22].

Now we shall prove the convexity of these measures. This is based on the well known result due to Csiszár [5, 6].

**Result 1.** If the function $f : \mathbb{R}_+ \to \mathbb{R}$ is convex and normalized, i.e., $f(1) = 0$, then the $f$-divergence, $C_f(P||Q)$ is nonnegative and convex in the pair of probability distribution $(P, Q) \in \Gamma_n \times \Gamma_n$.

**Example 4.** Let us consider
\begin{equation}
f_{AH}(x) = \frac{(x - 1)^2}{2(x + 1)}, \quad x \in (0, \infty),
\end{equation}
in (15), then $C_f(P||Q) = M_{AH}(P||Q)$, where $M_{AH}(P||Q)$ is as given by (54).
Moreover,
\begin{equation}
_f'(x) = \frac{(x - 1)(x + 3)}{2(x + 1)^2}
\end{equation}
and
\begin{equation}
 f''_{AH}(x) = \frac{4}{(x+1)^3} > 0, \ x \in (0, \infty).
\end{equation}

**Example 5.** Let us consider
\begin{equation}
 f_{AG}(x) = \frac{1}{2}(\sqrt{x} - 1)^2, \ x \in (0, \infty),
\end{equation}
in (15), then \( C_f(P||Q) = M_{AG}(P||Q) \), where \( M_{AG}(P||Q) \) is as given by (63).
Moreover,
\begin{equation}
 f'_{AG}(x) = \frac{\sqrt{x} - 1}{2\sqrt{x}}
\end{equation}
and
\begin{equation}
 f''_{AG}(x) = \frac{1}{4x\sqrt{x}} > 0, \ x \in (0, \infty).
\end{equation}

**Example 6.** Let us consider
\begin{equation}
 f_{N2N1}(x) = \frac{(x+1)\sqrt{2(x+1)} - 1 - x - 2\sqrt{x}}{4}, \ x \in (0, \infty)
\end{equation}
in (13), then we have \( C_f(P||Q) = M_{N2N1}(P||Q) \), where \( M_{N2N1}(P||Q) \) is as given by (57).
Moreover,
\begin{equation}
 f'_{N2N1}(x) = \frac{2x + 1 + \sqrt{x} - (\sqrt{x} + 1)\sqrt{2(x+1)}}{6\sqrt{x}(x+1)^2}
\end{equation}
and
\begin{equation}
 f''_{N2N1}(x) = \frac{-2x - 2x^{5/2} + x(2x + 2)^{3/2}}{8x^{5/2}(2x + 2)^{3/2}}
 = \frac{x[(2x + 2)^{3/2} - 2(x^{3/2} + 1)]}{8x^{5/2}(2x + 2)^{3/2}}.
\end{equation}
Since \((x + 1)^{3/2} \geq x^{3/2} + 1, \ \forall \ x \in (0, \infty) \) and \( 2^{3/2} \geq 2 \), then obviously, \( f''_{N2N1}(x) \geq 0, \ \forall \ x \in (0, \infty) \).

**Example 7.** Let us consider
\begin{equation}
 f_{N2G}(x) = \frac{(\sqrt{x} + 1)\sqrt{2(x+1)} - 4x}{4}, \ x \in (0, \infty),
\end{equation}
in (13), then \( C_f(P||Q) = M_{N2G}(P||Q) \), where \( M_{N2G}(P||Q) \) is as given by (56).
Moreover,
\begin{equation}
 f'_{N2G}(x) = \frac{2x + 1 + \sqrt{x} - 2\sqrt{2(x+1)}}{4\sqrt{2x(x+1)}}
\end{equation}
and
\begin{equation}
 f''_{N2G}(x) = \frac{(2x + 2)^{3/2} - x^{3/2} - 1}{4x^{3/2}(2x + 2)^{3/2}}.
\end{equation}
Since \((x + 1)^{3/2} \geq x^{3/2} + 1, \ \forall \ x \in (0, \infty) \) and \( 2^{3/2} \geq 1 \), then obviously, \( f''_{N2G}(x) \geq 0, \ \forall \ x \in (0, \infty) \).
Example 8. Let us consider
\begin{equation}
    f_{AN^2}(x) = \frac{2(x+1) - (\sqrt{x} + 1)\sqrt{2(x+1)}}{4}, \quad x \in (0, \infty),
\end{equation}
in (3.1), then \( C_f(P||Q) = M_{AN^2}(P||Q) \), where \( M_{AN^2}(P||Q) \) is as given by (53).
Moreover,
\begin{equation}
    f'_{AN^2}(x) = -\frac{2x + 1 + \sqrt{x} - 2\sqrt{2x(x+1)}}{4\sqrt{2(x+1)}},
\end{equation}
and
\begin{equation}
    f''_{AN^2}(x) = \frac{1 + x^{3/2}}{8x^{3/2}(x+1)\sqrt{2x+2}} > 0, \quad x \in (0, \infty).
\end{equation}

In the above examples the generating function \( f(1) = 0 \) and the second derivative is positive for all \( x \in (0, \infty) \). This proves the nonnegativity and convexity of the measures (53)-(57) in the pair of probability distributions \((P, Q) \in \Gamma_n \times \Gamma_n\).

The inequality (52) also admits more nonnegative differences, but here we have considered only the convex ones.

Based on the Proposition 2, we can obtain bounds on the mean divergence measures, but we omit these details here. Now we shall apply the inequalities (54) given in Proposition 1.3 to obtain inequalities among the measures (53)-(57).

**Theorem 3.** The following inequalities among the six mean divergences hold:
\begin{equation}
    \frac{1}{8} M_{AH}(P||Q) \leq M_{N_2N_1}(P||Q) \leq \frac{1}{3} M_{N_2G}(P||Q)
\end{equation}
\begin{equation}
    \leq \frac{1}{4} M_{AG}(P||Q) \leq M_{AN^2}(P||Q).
\end{equation}

The proof of the above theorem is based on the following propositions, where we have proved each part separately.

**Proposition 4.** The following inequality hold:
\begin{equation}
    \frac{1}{8} M_{AH}(P||Q) \leq M_{N_2N_1}(P||Q).
\end{equation}

**Proof.** Let us consider
\begin{equation}
    g_{AH*N_2N_1}(x) = \frac{f''_{AH}(x)}{f''_{N_2N_1}(x)}
\end{equation}
\begin{equation}
    = \frac{32x^{5/2}(2x + 2)^{3/2}}{(x+1)^3 \left[-2x - 2x^{5/2} + x(2x+2)\right]}, \quad x \in (0, \infty),
\end{equation}
where \( f''_{AH}(x) \) and \( f''_{N_2N_1}(x) \) are as given by (63) and (67) respectively.
From (74), we have
\[ g'_{AH,N_2N_1}(x) = \frac{-48\sqrt{2x(x+1)}}{(x+1)^4[-2x - 2x^{5/2} + x(2x + 2)^{3/2}]^2} \times \]
\[ \times \left[4x^2(1 - x^{5/2}) + x^2(x-1)(2x + 2)^{5/2}\right] \]
\[ = \frac{48x^2(x+1)(1 - \sqrt{x})\sqrt{2x(x+1)}}{(x+1)^4[-2x - 2x^{5/2} + x(2x + 2)^{3/2}]^2} \times \]
\[ \times \left[\sqrt{2} (\sqrt{x} + 1)(x + 1)^{3/2} - (x^2 + x^{3/2} + x + \sqrt{x} + 1)\right]. \]

Since \(\sqrt{2}(x+1) \geq \sqrt{x} + 1, \forall x \in (0, \infty)\), then
\[ \sqrt{2}(x+1)^{3/2} (\sqrt{x} + 1) \geq (\sqrt{x} + 1)^2 (x + 1) \]
\[ \geq x^2 + x^{3/2} + x + \sqrt{x} + 1. \]

Thus we conclude that
\[
(75) \quad g'_{AH,N_2N_1}(x) = \begin{cases} < 0, & x > 1, \\ > 0, & x < 1. \end{cases}
\]

In view of (75), we conclude that the function \(g_{AH,N_2N_1}(x)\) is increasing in \(x \in (0, 1)\) and decreasing in \(x \in (1, \infty)\), and hence
\[
(76) \quad M = \sup_{x \in (0, \infty)} g_{AH,N_2N_1}(x) = g_{AH,N_2N_1}(1) = 8.
\]

Applying the inequalities (34) for the measures \(M_{AH}(P\|Q)\) and \(M_{N_2N_1}(P\|Q)\) along with (76) we get the required result.

**Proposition 5.** The following inequality hold:

\[
(77) \quad M_{N_2N_1}(P\|Q) \leq \frac{1}{3} M_{N_2G}(P\|Q).
\]

**Proof.** Let us consider
\[
(78) \quad g_{N_2N_1,N_2G}(x) = \frac{f_{N_2N_1}(x)}{f_{N_2G}(x)} = -\frac{2x - 2x^{5/2} + x(2x + 2)^{3/2}}{2x[1 + x^{3/2} - (2x + 2)^{3/2}], \ x \in (0, \infty),}
\]
where \(f_{N_2N_1}(x)\) and \(f_{N_2G}(x)\) are as given by (67) and (69) respectively.

From (78), we have
\[
(79) \quad g'_{N_2N_1,N_2G}(x) = \frac{3x^2\sqrt{2x+2}(1 - \sqrt{x})}{2x^2[-1 - x^{3/2} + (2x + 2)^{3/2}]^2} \begin{cases} < 0, & x > 1, \\ > 0, & x < 1. \end{cases}
\]

In view of (79), we conclude that the function \(g_{N_2N_1,N_2G}(x)\) is increasing in \(x \in (0, 1)\) and decreasing in \(x \in (1, \infty)\), and hence
\[
(80) \quad M = \sup_{x \in (0, \infty)} g_{N_2N_1,N_2G}(x) = g_{N_2N_1,N_2G}(1) = \frac{1}{3}.
\]
Applying the inequalities (34) for the measures $M_{N_2 N_1}(P \| Q)$ and $M_{N_2 G}(P \| Q)$ along with (80) we get the required result.

**Proposition 6.** The following inequality hold:

(81) \[ M_{N_2 G}(P \| Q) \leq \frac{3}{4} M_{AG}(P \| Q). \]

**Proof.** Let us consider

(82) \[ g_{N_2 G \cdot AG}(x) = \frac{f''_{N_2 G}(x)}{f''_{AG}(x)} = \frac{1 + x^{3/2} - (2x + 2)^{3/2}}{(2x + 2)^{3/2}}, \quad x \in (0, \infty), \]

where $f''_{N_2 G}(x)$ and $f''_{AG}(x)$ are as given by (69) and (65) respectively.

From (82), we have

(83) \[ g'_{N_2 G \cdot AG}(x) = \frac{3(1 - \sqrt{x})}{(2x + 2)^{5/2}} \begin{cases} \leq 0, & x \geq 1, \\ \geq 0, & x \leq 1. \end{cases} \]

In view of (83), we conclude that the function $g_{AH \cdot N_2 N_1}(x)$ is increasing in $x \in (0, 1)$ and decreasing in $x \in (1, \infty)$, and hence

(84) \[ M = \sup_{x \in (0, \infty)} g_{N_2 G \cdot AG}(x) = g_{N_2 G \cdot AG}(1) = \frac{3}{4}. \]

Applying the inequalities (34) for the measures $M_{N_2 G}(P \| Q)$ and $M_{AG}(P \| Q)$ along with (84) we get the required result.

**Proposition 7.** The following inequality hold:

(85) \[ \frac{1}{4} M_{AG}(P \| Q) \leq M_{N_2}(P \| Q). \]

**Proof.** Let us consider

(86) \[ g_{AG \cdot AN_2}(x) = \frac{f''_{AG}(x)}{f''_{AN_2}(x)} = \frac{(2x + 2)^{3/2}}{(\sqrt{x} + 1)(x - \sqrt{x} + 1)}, \quad x \in (0, \infty), \]

where $f''_{AG}(x)$ and $f''_{AN_2}(x)$ are as given by (65) and (71) respectively.

From (86), we have

(87) \[ g'_{AG \cdot AN_2}(x) = \frac{3(1 - \sqrt{x}) \sqrt{2x + 2}}{(\sqrt{x} + 1)^2(x - \sqrt{x} + 1)^2} \begin{cases} \leq 0, & x \geq 1, \\ \geq 0, & x \leq 1. \end{cases} \]

In view of (87), we conclude that the function $g_{AG \cdot AN_2}(x)$ is increasing in $x \in (0, 1)$ and decreasing in $x \in (1, \infty)$, and hence

(88) \[ M = \sup_{x \in (0, \infty)} g_{AG \cdot AN_2}(x) = g_{AG \cdot AN_2}(1) = 4. \]

Applying the inequalities (34) for the measures $M_{AG}(P \| Q)$ and $M_{AN_2}(P \| Q)$ along with (88) we get the required result.
Combining the results given in the Propositions 4-7, we get the proof of the theorem. The expression (73) can also be written as

\[
\frac{1}{16} \Delta(P\|Q) \leq M_{N_2N_1}(P\|Q) \leq \frac{1}{3} M_{N_2G}(P\|Q) \\
\leq \frac{1}{4} h(P\|Q) \leq M_{AN_2}(P\|Q).
\]

**Remark 1.**

(i) The classical divergence measures \(I(P\|Q)\) and \(J(P\|Q)\) appearing in the Section 1 can be written in terms of Kullback-Leibler’s relative information as follows:

\[
I(P\|Q) = \frac{1}{2} \left[ K \left( P\| \frac{P + Q}{2} \right) + K \left( Q\| \frac{P + Q}{2} \right) \right]
\]

and

\[
J(P\|Q) = K(P\|Q) + K(Q\|P).
\]

Also we can write

\[
J(P\|Q) = 4 \left[ I(P\|Q) + T(P\|Q) \right],
\]

where

\[
T(P\|Q) = \frac{1}{2} \left[ K \left( \frac{P + Q}{2} \| P \right) + K \left( \frac{P + Q}{2} \| Q \right) \right] = \sum_{i=1}^{n} A(p_i, q_i) \ln \left( \frac{A(p_i, q_i)}{G(p_i, q_i)} \right),
\]

is the arithmetic and geometric mean divergence measure due to Taneja [19].

(ii) Recently, Taneja [22] proved the following inequality:

\[
\frac{1}{4} \Delta(P\|Q) \leq I(P\|Q) \leq h(P\|Q) \leq 4 M_{AN_2}(P\|Q) \leq \frac{1}{8} J(P\|Q) \leq T(P\|Q).
\]

Following the lines of the Propositions 4-7, we can also show that

\[
\frac{1}{4} I(P\|Q) \leq M_{N_2N_1}(P\|Q).
\]

Thus combining [33] with (92), (93) and (94), we get the following inequalities among the classical and mean divergence measures:

\[
\frac{1}{16} \Delta(P\|Q) \leq \frac{1}{4} I(P\|Q) \leq M_{N_2N_1}(P\|Q) \\
\leq \frac{1}{3} M_{N_2G}(P\|Q) \leq \frac{1}{4} h(P\|Q) \leq M_{AN_2}(P\|Q) \\
\leq \frac{1}{32} J(P\|Q) \leq \frac{1}{4} T(P\|Q) \leq \frac{1}{16} J(P\|Q).
\]
References

[1] D. ANRICA and I. RAŞA, The Jensen Inequality: Refinement and Applications, Anal. Num. Theory Approx., 14(1985), 105-108.

[2] E. F. BECKENBACH and R. BELLMAN, Inequalities, Springer-Verlag, New York, 1971.

[3] J. BURBEA, J. and C.R. RAO, Entropy Differential Metric, Distance and Divergence Measures in Probability Spaces: A Unified Approach, J. Multi. Analysis, 12(1982), 575-596.

[4] J. BURBEA, J. and C.R. RAO, On the Convexity of Some Divergence Measures Based on Entropy Functions, IEEE Trans. on Inform. Theory, IT-28(1982), 489-495.

[5] I. CSIZÁR, Information Type Measures of Differences of Probability Distribution and Indirect Observations, Studia Math. Hungarica, 2(1967), 299-318.

[6] I. CSIZÁR, On Topological Properties of $f$–Divergences, Studia Math. Hungarica, 2(1967), 329-339.

[7] P. CRESSIE and T.R.C. READ, Multinomial Goodness-of-fit Tests, J. Royal Statist. Soc., Ser. B, 46(1984), 440-464.

[8] S.S. DRAGOMIR, An Inequality for Twice Differentiable Convex Mappings and Applications for the Shannon and Rényi’s Entropies, RGMIA Research Report Collection, http://rgmia.vu.edu.au, ..(1999).

[9] S.S. DRAGOMIR, On An Inequality for Logarithmic Mapping and Applications for the Shannon Entropy, RGMIA Research Report Collection, http://rgmia.vu.edu.au, ..(1999).

[10] S. S. DRAGOMIR, A Converse Result for Jensen’s Discrete Inequality via Gruss’ Inequality and Applications in Information Theory, available on line: [http://rgmia.vu.edu.au/authors/SSDragomir.htm](http://rgmia.vu.edu.au/authors/SSDragomir.htm) 1999.

[11] S. S. DRAGOMIR, Some Inequalities for the Csiszár’s $f$–Divergence - Inequalities for Csiszár’s $f$–Divergence in Information Theory - Monograph - Chapter I - Article 1 - [http://rgmia.vu.edu.au/monographs/csiszar.htm](http://rgmia.vu.edu.au/monographs/csiszar.htm).

[12] S. S. DRAGOMIR, Other Inequalities for Csiszár’s Divergence and Applications - Inequalities for Csiszár’s f-Divergence in Information Theory - Monograph - Chapter I - Article 4 - [http://rgmia.vu.edu.au/monographs/csiszar.htm](http://rgmia.vu.edu.au/monographs/csiszar.htm).

[13] S. S. DRAGOMIR, N.M. Dragomir and K. PRANESH, An Inequality for Logarithms and Applications in Information Theory, Computers and Math. with Appl., 38(1999), 11-17.

[14] S.S. DRAGOMIR and N.M. IONESCU, Some Converse of Jensen’s Inequality Anal. Num. Theory Approx., 23(1994), 71-78.

[15] E. HELLINGER, Neue Begründung der Theorie der quadratischen Formen von unendlichen vielen Veränderlichen, J. Reine Aug. Math., 136(1909), 210-271.

[16] H. JEFFREYS, An Invariant Form for the Prior Probability in Estimation Problems, Proc. Roy. Soc. Lon., Ser. A, 186(1946), 453-461.

[17] S. KULLBACK and R.A. LEIBLER, On Information and Sufficiency, Ann. Math. Statist., 22(1951), 79-86.

[18] I. J. TANEJA, On Generalized Information Measures and Their Applications, Chapter in: Advances in Electronics and Electron Physics, Ed. P.W. Hawkes, Academic Press, 76(1989), 327-413.

[19] I. J. TANEJA, New Developments in Generalized Information Measures, Chapter in: Advances in Imaging and Electron Physics, Ed. P.W. Hawkes, 91(1995), 37-136.

[20] I. J. TANEJA, Generalized Information Measures and their Applications - on line book: [http://www.mtm.ufsc.br/~taneja/book/book.html](http://www.mtm.ufsc.br/~taneja/book/book.html), 2001.
[21] I. J. TANEJA, Generalized Relative Information and Information Inequalities, *Journal of Inequalities in Pure and Applied Mathematics*. Vol. 5, No.1, 2004, Article 21, 1-19. Also in: *RGMIA Research Report Collection*, http://rgmia.vu.edu.au, 6(3)(2003), Article 10.

[22] I.J. TANEJA, Generalized Symmetric Divergence Measures and Inequalities – *RGMIA Research Report Collection*, http://rgmia.vu.edu.au, 7(4)(2004), Art. 9

[23] I. J. TANEJA and P. KUMAR, Relative Information of Type s, Csiszár f–Divergence, and Information Inequalities, *Information Sciences*, 166(1-4),2004, 105-125. Also in: http://rgmia.vu.edu.au, *RGMIA Research Report Collection*, 6(3)(2003), Article 12.

Inder Jeet Taneja, Departamento de Matemática, Universidade Federal de Santa Catarina, 88.040-900 Florianópolis, SC, Brazil

*E-mail address*: taneja@mtm.ufsc.br

*URL*: http://www.mtm.ufsc.br/~taneja