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 Csiszár [5] $f-$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

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

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, $x_i \in \overset{o}{I}$ ($\overset{o}{I}$ is the interior of I). Let $\lambda = (\lambda_1, \lambda_1, ..., \lambda_n) \in \Gamma_n$, then it is well known that

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i). \tag{1.1}$$

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

Let us consider the following *Jensen difference*:

$$F(\lambda, X) = \sum_{i=1}^{n} \lambda_i f(x_i) - f\left(\sum_{i=1}^{n} \lambda_i x_i\right), \tag{1.2}$$

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

Theorem 1.1. *Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be a differentiable convex function on the interval I, $x_i \in \overset{o}{I}$ ($\overset{o}{I}$ is the interior of I), $\lambda = (\lambda_1, \lambda_1, ..., \lambda_n) \in \Gamma_n$. If $\eta_1,\ \eta_2 \in \overset{o}{I}$ and $\eta_1 \leqslant x_i \leqslant \eta_2$, $\forall i = 1, 2, ..., n$, then we have the inequalities:*

$$0 \leqslant F_f(\lambda, X) \leqslant L_f(\lambda, X) \leqslant Z_f(\eta_1, \eta_2), \tag{1.3}$$

2000 *Mathematics Subject Classification.* 94A17; 26D15.

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

where

$$L_f(\lambda, X) = \sum_{i=1}^{n} \lambda_i x_i f'(x_i) - \left(\sum_{i=1}^{n} \lambda_i x_i\right)\left(\sum_{i=1}^{n} \lambda_i f'(x_i)\right) \tag{1.4}$$

and

$$Z_f(\eta_1, \eta_2) = \frac{1}{4}(\eta_2 - \eta_1)\left[f'(\eta_2) - f'(\eta_1)\right]. \tag{1.5}$$

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 1.1. *For all $a, b, \upsilon, \omega \in (0, \infty)$, the following inequality hold:*

$$\begin{aligned} 0 &\leqslant \frac{\upsilon f(a) + \omega f(b)}{\upsilon + \omega} - f\left(\frac{\upsilon a + \omega b}{\upsilon + \omega}\right) \\ &\leqslant \frac{\upsilon a f'(a) + \omega b f'(b)}{\upsilon + \omega} - \left(\frac{\upsilon a + \omega b}{\upsilon + \omega}\right)\left(\frac{\upsilon f'(a) + \omega f'(b)}{\upsilon + \omega}\right) \\ &\leqslant \frac{1}{4}(b - a)\left(f'(b) - f'(a)\right). \end{aligned} \tag{1.6}$$

Proof. It follows from Theorem 1.1, by taking $\lambda_1 = \frac{\upsilon}{\upsilon+\omega}$, $\lambda_2 = \frac{\omega}{\upsilon+\omega}$, $\lambda_3 = ... = \lambda_n = 0$, $x_1 = a$, $x_2 = b$, $x_2 = ... = x_n = 0$. □

Now we shall give some examples of Theorem 1.1.

Example 1.1. *For all $x \in (0, \infty)$, let us consider a function*

$$f_s(x) = \begin{cases} \frac{1-x^s}{s}, & r \neq 0, \\ -\ln x, & r = 0. \end{cases} \tag{1.7}$$

We can easily check that the function $f_s(x)$ is convex in $(0, \infty)$ for all $s \leqslant 1$. Let there exist η_1 and η_2 such that $\eta_1 \leqslant x_i \leqslant \eta_2$, $\forall i = 1, 2, ..., n$. Applying Theorem 1.1 for the function $f_s(x)$, we have

$$0 \leqslant F_s(\lambda, X) \leqslant Z_s(\eta_1, \eta_2),\ \ s \leqslant 1, \tag{1.8}$$

where

$$F_s(\lambda, X) = \begin{cases} \frac{1}{s}\left[\left(\sum\limits_{i=1}^{n} \lambda_i x_i\right)^s - \sum\limits_{i=1}^{n} \lambda_i x_i^s\right], & s \neq 0, \\ \\ \ln\left(\frac{A(\lambda,X)}{G(\lambda,X)}\right), & s = 0. \end{cases} \tag{1.9}$$

$$A(\lambda, X) = \sum_{i=1}^{n} \lambda_i x_i, \tag{1.10}$$

$$G(\lambda, X) = \prod_{i=1}^{n} x_i^{\lambda_i} \tag{1.11}$$

and

$$Z_s(\alpha,\beta)=\frac{1}{4}(\eta_2-\eta_1)\left(\eta_1^{s-1}-\eta_2^{s-1}\right). \tag{1.12}$$

In particular we have

$$\frac{A(\lambda,X)}{G(\lambda,X)}\leqslant \exp\left[\frac{(\eta_2-\eta_1)^2}{4\eta_1\eta_2}\right],\ \eta_1\leqslant x_i\leqslant \eta_2,\ \forall i=1,2,..n. \tag{1.13}$$

The result (1.13) is due to Dragomir [10]. The following proposition is a particular case of the inequalities (1.6) and gives bounds on Burbea and Rao's [3, 4] *Jensen difference divergence measure.*

Proposition 1.1. *Let $f:(0,\infty)\to\mathbb{R}$ be a differentiable convex function. Then for all $P,Q\in\Gamma_n$, we have*

$$0\leqslant\sum_{i=1}^{n}\left[\frac{f(p_i)+f(q_i)}{2}-f\left(\frac{p_i+q_i}{2}\right)\right]\leqslant\frac{1}{4}\sum_{i=1}^{n}(p_i-q_i)\left[f'(p_i)-f'(q_i)\right]. \tag{1.14}$$

Proof. Take $\omega=\upsilon=\frac{1}{2}$ in (1.6), we get

$$0\leqslant\frac{f(a)+f(b)}{2}-f\left(\frac{a+b}{2}\right)\leqslant\frac{1}{4}(b-a)\left[f'(b)-f'(a)\right]. \tag{1.15}$$

Replace in (1.15), a by p_i and b by q_i, and sum over all $i=1,2,...,n$, we get the required result. □

Example 1.2. *Let us consider a convex function*

$$\phi_s(x)=\begin{cases}\left[s(s-1)\right]^{-1}\left[x^s-1-s(x-1)\right], & s\neq 0,1,\\ x-1-\ln x, & s=0,\\ 1-x+x\ln x, & s=1,\end{cases} \tag{1.16}$$

for all $x\in(0,\infty)$ and $s\in(-\infty,\infty)$. Then from (1.14), we get

$$0\leqslant\mathcal{W}_s(P||Q)\leqslant\mathcal{V}_s(P||Q), \tag{1.17}$$

where

$$\mathcal{W}_s(P||Q)=\begin{cases}I_s(P||Q)=\left[s(s-1)\right]^{-1}\sum_{i=1}^{n}\left[\frac{p_i^s+q_i^s}{2}-\left(\frac{p_i+q_i}{2}\right)^s\right], & s\neq 0,1,\\ I_0(P||Q)=\ln\left[\prod_{i=1}^{n}\left(\frac{p_i+q_i}{2\sqrt{p_iq_i}}\right)\right], & s=0,\\ I(P||Q)=H\left(\frac{P+Q}{2}\right)-\frac{H(P)+H(Q)}{2}, & s=1,\end{cases} \tag{1.18}$$

and

$$\mathcal{V}_s(P||Q) = \begin{cases} J_s(P||Q) = \frac{1}{4(s-1)} \sum_{i=1}^{n} (p_i - q_i)\left(p_i^{s-1} - q_i^{s-1}\right), & s \neq 0, 1, \\ J_0(P||Q) = \frac{1}{4} \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i q_i}, & s = 0, \\ J(P||Q) = \frac{1}{4} \sum_{i=1}^{n} (p_i - q_i) \ln\left(\frac{p_i}{q_i}\right), & s = 1, \end{cases} \tag{1.19}$$

The expression $H(P) = -\sum_{i=1}^{n} p_i \ln p_i$, appearing in (1.18) is the well known *Shannon's entropy*. The expression $J(P||Q)$ appearing in (1.19) is *Jeffreys-Kullback-Leibler's J-divergence* (ref. Jeffreys [16] and Kullback and Leibler [17]). The expression $J_s(P||Q)$ is due to Burbea and Rao [3]. The measures (1.18) and (1.19) 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 (1.16) is due to Cressie and Read [7].

Proposition 1.2. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be differentiable convex and normalized i.e., $f(1) = 0$. If $P, Q \in \Gamma_n$, are such that $0 < r \leqslant \frac{p_i}{q_i} \leqslant R < \infty$, $\forall i \in \{1, 2, ..., n\}$, for some r and R with $0 < r \leqslant 1 \leqslant R < \infty$, then we have*

$$0 \leqslant C_f(P||Q) \leqslant E_{C_f}(P||Q) \leqslant A_{C_f}(r, R), \tag{1.20}$$

and

$$0 \leqslant C_f(P||Q) \leqslant B_{C_f}(r, R) \leqslant A_{C_f}(r, R), \tag{1.21}$$

where

$$C_f(P||Q) = \sum_{i=1}^{n} q_i f(\frac{p_i}{q_i}), \tag{1.22}$$

$$E_{C_f}(P||Q) = \sum_{i=1}^{n} (p_i - q_i) f'(\frac{p_i}{q_i}), \tag{1.23}$$

$$A_{C_f}(r, R) = \frac{1}{4}(R - r)\left(f'(R) - f'(r)\right) \tag{1.24}$$

and

$$B_{C_f}(r, R) = \frac{(R-1)f(r) + (1-r)f(R)}{R - r}. \tag{1.25}$$

The inequalities (1.20) follow in view of (1.3). The inequalities (1.21) follow in view of (1.6). For details refer to Taneja [22]. The above proposition is an improvement over the work of Dragomir [11, 12]. The measure (1.22) is known as *Csiszár's* [5] *f-divergence*.

Example 1.3. *Under the conditions of Proposition 1.2, the inequalities (1.20) and (1.21) for the function (1.16) are given by*

$$0 \leqslant \Phi_s(P||Q) \leqslant E_{\Phi_s}(P||Q) \leqslant A_{\Phi_s}(r,R) \tag{1.26}$$

and

$$0 \leqslant \Phi_s(P||Q) \leqslant B_{\Phi_s}(r,R) \leqslant A_{\Phi_s}(r,R), \tag{1.27}$$

where

$$\Phi_s(P||Q) = \begin{cases} K_s(P||Q) = [s(s-1)]^{-1}\left[\sum_{i=1}^{n} p_i^s q_i^{1-s} - 1\right], & s \neq 0,1, \\ K(Q||P) = \sum_{i=1}^{n} q_i \ln\left(\frac{q_i}{p_i}\right), & s = 0, \\ K(P||Q) = \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right), & s = 1, \end{cases} \tag{1.28}$$

$$E_{\Phi_s}(P||Q) = \begin{cases} (s-1)^{-1}\sum_{i=1}^{n}(p_i - q_i)\left(\frac{p_i}{q_i}\right)^{s-1}, & s \neq 1, \\ \sum_{i=1}^{n}(p_i - q_i)\ln\left(\frac{p_i}{q_i}\right), & s = 1, \end{cases} \tag{1.29}$$

$$A_{\Phi_s}(r,R) = \frac{1}{4}\begin{cases} \frac{(R-r)\left(R^{s-1} - r^{s-1}\right)}{4(s-1)}, & s \neq 1, \\ \frac{1}{4}(R-r)\ln\left(\frac{R}{r}\right), & s = 1, \end{cases} \tag{1.30}$$

and

$$B_{\Phi_s}(r,R) = \begin{cases} \frac{(R-1)(r^s-1)+(1-r)(R^s-1)}{(R-r)s(s-1)}, & s \neq 0,1, \\ \frac{(R-1)\ln\frac{1}{r}+(1-r)\ln\frac{1}{R}}{(R-r)}, & s = 0, \\ \frac{(R-1)r\ln r+(1-r)R\ln R}{(R-r)}, & s = 1. \end{cases} \tag{1.31}$$

The measure $K(P||Q)$ appearing in (1.28) is the well known *Kullback-Leibler's* [17] *relative information*. The measure $\Phi_s(P||Q)$ given in (1.28) has been extensively studied in [21], [23].

Theorem 1.2. *Let $f_1, f_2 : [a,b] \subset \mathbb{R}_+ \to \mathbb{R}$ be twice differentiable functions on (a,b) and there are α and β such that*

$$\alpha \leqslant \frac{f_1''(x)}{f_2''(x)} \leqslant \beta,\ \forall x \in (a,b),\ f_2''(x) > 0 \tag{1.32}$$

If $x_i \in [a,b]$ *and* $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n) \in \Gamma_n$, *then*

$$\alpha\, F_{f_2}(\lambda, X) \leqslant F_{f_1}(\lambda, X) \leqslant \beta\, F_{f_2}(\lambda, X). \tag{1.33}$$

$$\begin{aligned} \alpha\left[L_{f_2}(\lambda, X) - F_{f_2}(\lambda, X)\right] &\leqslant L_{f_1}(\lambda, X) - F_{f_1}(\lambda, X) \\ &\leqslant \beta\left[L_{f_2}(\lambda, X) - F_{f_2}(\lambda, X)\right] \end{aligned} \tag{1.34}$$

and

$$\begin{aligned} \alpha\left[Z_{f_2}(\eta_1, \eta_2) - F_{f_2}(\lambda, X)\right] &\leqslant Z_{f_1}(\eta_1, \eta_2) - F_{f_1}(\lambda, X) \\ &\leqslant \beta\left[Z_{f_2}(\eta_1, \eta_2) - F_{f_2}(\lambda, X)\right]. \end{aligned} \tag{1.35}$$

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

$$g(x) = f_1(x) - \alpha f_2(x), \quad \forall x \in [a,b], \tag{1.36}$$

where the functions f_1 and f_2 satisfy the condition (1.32). Then the function g is twice differentiable on (a,b). This gives

$$g'(x) = f_1'(x) - \alpha f_2'(x)$$

and

$$g''(x) = f_1''(x) - \alpha f_2''(x) = f_2''(x)\left(\frac{f_1''(x)}{f_2''(x)} - \alpha\right) \geqslant 0, \ \forall x \in (a,b).$$

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

$$g\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i g(x_i),$$

i.e.,

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

i.e.,

$$\alpha\left[\sum_{i=1}^{n} \lambda_i f_2(x_i) - f_2\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right] \leqslant \sum_{i=1}^{n} \lambda_i f_1(x_i) - f_1\left(\sum_{i=1}^{n} \lambda_i x_i\right). \tag{1.37}$$

The expression (1.37) gives the *l.h.s.* of the inequalities (1.33).

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

$$k(x) = \beta f_2(x) - f_1(x), \tag{1.38}$$

and proceeding on similar lines as before, we get the proof of the *r.h.s.* of the inequalities (1.33).

Now we shall prove the inequalities (1.34). Applying the inequalities (1.3) for the convex function g given by (1.36), we get

$$F_g(\lambda, X) \leqslant L_g(\lambda, X) \leqslant Z_g(\eta_1, \eta_2).$$

i.e.,

$$F_{f_1}(\lambda, X) - \alpha F_{f_2}(\lambda, X) \leqslant L_{f_1}(\lambda, X) - \alpha L_{f_2}(\lambda, X) \leqslant Z_{f_1}(\lambda, X) - \alpha F_{f_2}(\eta_1, \eta_2). \tag{1.39}$$

Simplifying the first inequality of (1.39) we get

$$\alpha \left[L_{f_2}(\lambda, X) - F_{f_2}(\lambda, X)\right] \leqslant L_{f_1}(\lambda, X) - F_{f_1}(\lambda, X). \tag{1.40}$$

Again simplifying the last inequality of (1.39) we get

$$\alpha \left[Z_{f_2}(\eta_1, \eta_2) - F_{f_2}(\lambda, X)\right] \leqslant Z_{f_1}(\eta_1, \eta_2) - F_{f_1}(\lambda, X). \tag{1.41}$$

The expressions (1.40) and (1.41) complete the first part of the inequalities (1.34) and (1.35) respectively. The last part of the inequalities (1.34) and (1.35) follows by considering the function $k(x)$ given by (1.38) over the inequalities (1.3). □

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 1.3. *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 $0 < r \leqslant 1 \leqslant R < \infty$;

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

$$\alpha \leqslant \frac{f_1''(x)}{f_2''(x)} \leqslant \beta, \;\; f_2''(x) > 0, \;\; \forall x \in (r, R). \tag{1.42}$$

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

$$0 < r \leqslant \frac{p_i}{q_i} \leqslant R < \infty,$$

then we have the inequalities:

$$\alpha \; C_{f_2}(P||Q) \leqslant C_{f_1}(P||Q) \leqslant \beta \; C_{f_2}(P||Q), \tag{1.43}$$

$$\alpha \left[E_{f_2}(P||Q) - C_{f_2}(P||Q)\right] \leqslant E_{f_1}(P||Q) - C_{f_1}(P||Q) \leqslant \beta \left[E_{f_2}(P||Q) - C_{f_2}(P||Q)\right], \tag{1.44}$$

$$\alpha \left[A_{f_2}(r, R) - C_{f_2}(P||Q)\right] \leqslant A_{f_1}(r, R) - C_{f_1}(P||Q) \leqslant \beta \left[A_{f_2}(r, R) - C_{f_2}(P||Q)\right] \tag{1.45}$$

and

$$\alpha \left[B_{f_2}(r, R) - C_{f_2}(P||Q)\right] \leqslant B_{f_1}(r, R) - C_{f_1}(P||Q) \leqslant \beta \left[B_{f_2}(r, R) - C_{f_2}(P||Q)\right]. \tag{1.46}$$

Proof. It is an immediate consequence of the Theorem 1.2. □

2. Applications to Mean Divergence Measures

Let us consider the following mean of order t:

$$D_t(a,b) = \begin{cases} \left(\frac{a^t+b^t}{2}\right)^{1/t}, & t \neq 0, \\ \sqrt{ab}, & t = 0, \\ \max\{a,b\}, & t = \infty, \\ \min\{a,b\}, & t = -\infty, \end{cases} \tag{2.1}$$

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

$$D_{-1}(a,b) = H(a,b) = \frac{2}{\frac{1}{a}+\frac{1}{b}} = \frac{2ab}{a+b} = A(a^{-1},b^{-1})^{-1},$$
$$D_0(a,b) = G(a,b) = \sqrt{ab} = \sqrt{A(a,b)H(a,b)},$$
$$D_{1/2}(a,b) = N_1(a,b) = \left(\frac{\sqrt{a}+\sqrt{b}}{2}\right)^2 = A\left(\sqrt{a},\sqrt{b}\right)^2$$

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 know [2] that the *mean of order* t given in (2.1) is monotonically increasing in t, then we can write

$$D_{-1}(a,b) \leqslant D_0(a,b) \leqslant D_{1/2}(a,b) \leqslant D_1(a,b),$$

or equivalently,

$$H(a,b) \leqslant G(a,b) \leqslant N_1(a,b) \leqslant A(a,b). \tag{2.2}$$

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:

$$\frac{\sqrt{a}+\sqrt{b}}{2} \leqslant \sqrt{\frac{a+b}{2}}. \tag{2.3}$$

From (2.3), we can easily derive that

$$\left(\frac{\sqrt{a}+\sqrt{b}}{2}\right)^2 \leqslant \left(\frac{\sqrt{a}+\sqrt{b}}{2}\right)\left(\sqrt{\frac{a+b}{2}}\right) \leqslant \frac{a+b}{2}. \tag{2.4}$$

Finally, the expressions (2.2) and (2.4) lead us to following inequalities:

$$H(a,b) \leqslant G(a,b) \leqslant N_1(a,b) \leqslant N_2(a,b) \leqslant A(a,b), \tag{2.5}$$

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 (2.5), replace a by p_i and b by q_i sum over all $i = 1, 2, ...n$ we get

$$H(P||Q) \leqslant G(P||Q) \leqslant N_1(P||Q) \leqslant N_2(P||Q) \leqslant 1. \tag{2.6}$$

Based on inequalities (2.6), we shall build some *mean divergence measures*. Let us consider the following differences:

$$M_{AG}(P||Q) = 1 - G(P||Q), \tag{2.7}$$

$$M_{AH}(P||Q) = 1 - H(P||Q), \tag{2.8}$$

$$M_{AN_2}(P||Q) = 1 - N_2(P||Q), \tag{2.9}$$

$$M_{N_2G}(P||Q) = N_2(P||Q) - G(P||Q), \tag{2.10}$$

and

$$M_{N_2N_1}(P||Q) = N_2(P||Q) - N_1(P||Q). \tag{2.11}$$

We can easily verify that

$$\begin{aligned} M_{AG}(P||Q) &= 1 - G(P||Q) \\ &= 2\left[N_1(P||Q) - G(P||Q)\right] := 2M_{N_1G}(P||Q) \\ &= 2\left[1 - N_1(P||Q)\right] := 2M_{AN_1}(P||Q). \end{aligned} \tag{2.12}$$

(2.13)

We can also write

$$M_{AG}(P||Q) = 1 - G(P||Q) := h(P||Q) \tag{2.14}$$

and

$$M_{AH}(P||Q) = 1 - H(P||Q) := \frac{1}{2}\Delta(P||Q), \tag{2.15}$$

where $h(P||Q)$ is the well known *Hellinger's* [15] *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 2.1. *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 2.1. *Let us consider*

$$f_{AH}(x) = \frac{(x-1)^2}{2(x+1)},\ x \in (0, \infty), \tag{2.16}$$

in (1.15), then $C_f(P||Q) = M_{AH}(P||Q)$, where $M_{AH}(P||Q)$ is as given by (2.8).

Moreover,

$$f'_{AH}(x) = \frac{(x-1)(x+3)}{2(x+1)^2}$$

and

$$f''_{AH}(x) = \frac{4}{(x+1)^3} > 0,\ x \in (0,\infty). \tag{2.17}$$

Example 2.2. *Let us consider*

$$f_{AG}(x) = \frac{1}{2}(\sqrt{x}-1)^2,\ x \in (0,\infty), \tag{2.18}$$

in (1.15), then $C_f(P||Q) = M_{AG}(P||Q)$*, where* $M_{AG}(P||Q)$ *is as given by (2.7).*
Moreover,

$$f'_{AG}(x) = \frac{\sqrt{x}-1}{2\sqrt{x}}$$

and

$$f''_{AG}(x) = \frac{1}{4x\sqrt{x}} > 0,\ x \in (0,\infty). \tag{2.19}$$

Example 2.3. *Let us consider*

$$f_{N_2N_1}(x) = \frac{(x+1)\sqrt{2(x+1)} - 1 - x - 2\sqrt{x}}{4},\ x \in (0,\infty) \tag{2.20}$$

in (1.15), then we have $C_f(P||Q) = M_{N_2N_1}(P||Q)$*, where* $M_{N_2N_1}(P||Q)$ *is as given by (2.11).*
Moreover,

$$f'_{N_2N_1}(x) = \frac{2x+1+\sqrt{x} - (\sqrt{x}+1)\sqrt{2(x+1)}}{6\sqrt{x}(x+1)^2}$$

and

$$\begin{aligned} f''_{N_2N_1}(x) &= \frac{-2x - 2x^{5/2} + x(2x+2)^{3/2}}{8x^{5/2}(2x+2)^{3/2}} \\ &= \frac{x\left[(2x+2)^{3/2} - 2(x^{3/2}+1)\right]}{8x^{5/2}(2x+2)^{3/2}}. \end{aligned} \tag{2.21}$$

Since $(x+1)^{3/2} \geqslant x^{3/2}+1$, $\forall x \in (0,\infty)$ *and* $2^{3/2} \geqslant 2$*, then obviously,* $f''_{N_2N_1}(x) \geqslant 0$, $\forall x \in (0,\infty)$.

Example 2.4. *Let us consider*

$$f_{N_2G}(x) = \frac{(\sqrt{x}+1)\sqrt{2(x+1)} - 4x}{4},\ x \in (0,\infty), \tag{2.22}$$

in (1.15), then $C_f(P||Q) = M_{N_2G}(P||Q)$*, where* $M_{N_2G}(P||Q)$ *is as given by (2.10).*
Moreover,

$$f'_{N_2G}(x) = \frac{2x+1+\sqrt{x} - 2\sqrt{2(x+1)}}{4\sqrt{2x(x+1)}}$$

and

$$f''_{N_2G}(x) = \frac{(2x+2)^{3/2} - x^{3/2} - 1}{4x^{3/2}(2x+2)^{3/2}}. \tag{2.23}$$

Since $(x+1)^{3/2} \geqslant x^{3/2}+1$, $\forall x \in (0,\infty)$ *and* $2^{3/2} \geqslant 1$, *then obviously,* $f''_{N_2G}(x) \geqslant 0$, $\forall x \in (0,\infty)$.

Example 2.5. *Let us consider*

$$f_{AN_2}(x) = \frac{2(x+1) - (\sqrt{x}+1)\sqrt{2(x+1)}}{4},\ x \in (0,\infty), \tag{2.24}$$

in (3.1), then $C_f(P||Q) = M_{AN_2}(P||Q)$, *where* $M_{AN_2}(P||Q)$ *is as given by (2.9).*

Moreover,

$$f'_{AN_2}(x) = -\frac{2x+1+\sqrt{x}-2\sqrt{2x(x+1)}}{4\sqrt{2(x+1)}},$$

and

$$f''_{AN_2}(x) = \frac{1+x^{3/2}}{8x^{3/2}(x+1)\sqrt{2x+2}} > 0,\ x \in (0,\infty). \tag{2.25}$$

In the above examples 2.1-2.5 the generating function $f_{(\cdot)}(1) = 0$ and the second derivative is positive for all $x \in (0,\infty)$. This proves the *nonegativity* and *convexity* of the measures (2.7)-(2.11) in the pair of probability distributions $(P,Q) \in \Gamma_n \times \Gamma_n$.

The inequality (2.6) also admits more nonnegative differences, but here we have considered only the convex ones.

Based on the Proposition 1.2, we can obtain bounds on the *mean divergence measures*, but we omit these details here. Now we shall apply the inequalities (1.34) given in Proposition 1.3 to obtain inequalities among the measures (2.7)-(2.11).

Theorem 2.1. *The following inequalities among the six mean divergences hold:*

$$\frac{1}{8}M_{AH}(P||Q) \leqslant M_{N_2N_1}(P||Q) \leqslant \frac{1}{3}M_{N_2G}(P||Q) \tag{2.26}$$

$$\leqslant \frac{1}{4}M_{AG}(P||Q) \leqslant M_{AN_2}(P||Q).$$

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

Proposition 2.1. *The following inequality hold:*

$$\frac{1}{8}M_{AH}(P||Q) \leqslant M_{N_2N_1}(P||Q). \tag{2.27}$$

Proof. Let us consider

$$\begin{aligned} g_{AH\_N_2N_1}(x) &= \frac{f''_{AH}(x)}{f''_{N_2N_1}(x)} \\ &= \frac{32x^{5/2}(2x+2)^{3/2}}{(x+1)^3\left[-2x-2x^{5/2}+x(2x+2)^{3/2}\right]},\ x \in (0,\infty), \end{aligned} \tag{2.28}$$

where $f''_{AH}(x)$ and $f''_{N_2N_1}(x)$ are as given by (2.17) and (2.21) respectively.

From (2.28), we have

$$\begin{aligned} g'_{AH\_N_2N_1}(x) &= -\frac{48\sqrt{2x(x+1)}}{(x+1)^4\left[-2x-2x^{5/2}+x(2x+2)^{3/2}\right]^2}\times \\ &\quad \times\left[4x^2(1-x^{5/2})+x^2(x-1)(2x+2)^{5/2}\right] \\ &= \frac{48x^2(x+1)\left(1-\sqrt{x}\right)\sqrt{2x(x+1)}}{(x+1)^4\left[-2x-2x^{5/2}+x(2x+2)^{3/2}\right]^2}\times \\ &\quad \times\left[\sqrt{2}\left(\sqrt{x}+1\right)(x+1)^{3/2}-\left(x^2+x^{3/2}+x+\sqrt{x}+1\right)\right]. \end{aligned}$$

Since $\sqrt{2(x+1)} \geqslant \sqrt{x}+1$, $\forall x \in (0,\infty)$, then

$$\begin{aligned} \sqrt{2}(x+1)^{3/2}\left(\sqrt{x}+1\right) &\geqslant \left(\sqrt{x}+1\right)^2(x+1) \\ &\geqslant x^2+x^{3/2}+x+\sqrt{x}+1. \end{aligned}$$

Thus we conclude that

$$g'_{AH\_N_2N_1}(x)\begin{cases} <0, & x>1, \\ >0, & x<1. \end{cases} \tag{2.29}$$

In view of (2.29), 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

$$M = \sup_{x\in(0,\infty)} g_{AH\_N_2N_1}(x) = g_{AH\_N_2N_1}(1) = 8. \tag{2.30}$$

Applying the inequalities (1.34) for the measures $M_{AH}(P||Q)$ and $M_{N_2N_1}(P||Q)$ along with (2.30) we get the required result. □

Proposition 2.2. *The following inequality hold:*

$$M_{N_2N_1}(P||Q) \leqslant \frac{1}{3}M_{N_2G}(P||Q). \tag{2.31}$$

Proof. Let us consider

$$\begin{aligned} g_{N_2N_1\_N_2G}(x) &= \frac{f''_{N_2N_1}(x)}{f''_{N_2G}(x)} \\ &= \frac{-2x-2x^{5/2}+x(2x+2)^{3/2}}{2x\left[1+x^{3/2}-(2x+2)^{3/2}\right]},\ x\in(0,\infty), \end{aligned} \tag{2.32}$$

where $f''_{N_2N_1}(x)$ and $f''_{N_2G}(x)$ are as given by (2.21) and (2.23) respectively.

From (2.32), we have

$$g'_{N_2N_1\_N_2G_1}(x) = \frac{3x^2\sqrt{2x+2}\left(1-\sqrt{x}\right)}{2x^2\left[-1-x^{3/2}+(2x+2)^{3/2}\right]^2}\begin{cases} <0, & x>1, \\ >0, & x<1. \end{cases} \tag{2.33}$$

In view of (2.33), 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

$$M = \sup_{x\in(0,\infty)} g_{N_2N_1\_N_2G}(x) = g_{N_2N_1\_N_2G}(1) = \frac{1}{3}. \tag{2.34}$$

Applying the inequalities (1.34) for the measures $M_{N_2N_1}(P||Q)$ and $M_{N_2G}(P||Q)$ along with (2.34) we get the required result. □

Proposition 2.3. *The following inequality hold:*

$$M_{N_2G}(P||Q) \leqslant \frac{3}{4} M_{AG}(P||Q). \tag{2.35}$$

Proof. Let us consider

$$g_{N_2G\_AG}(x) = \frac{f''_{N_2G}(x)}{f''_{AG}(x)} = -\frac{1 + x^{3/2} - (2x+2)^{3/2}}{(2x+2)^{3/2}}, \ x \in (0, \infty), \tag{2.36}$$

where $f''_{N_2G}(x)$ and $f''_{AG}(x)$ are as given by (2.23) and (2.19) respectively.

From (2.36), we have

$$g'_{N_2G\_AG}(x) = \frac{3\,(1-\sqrt{x})}{(2x+2)^{5/2}} \begin{cases} \leqslant 0, & x \geqslant 1, \\ \geqslant 0, & x \leqslant 1. \end{cases} \tag{2.37}$$

In view of (2.37), 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

$$M = \sup_{x \in (0,\infty)} g_{N_2G\_AG}(x) = g_{N_2G\_AG}(1) = \frac{3}{4}. \tag{2.38}$$

Applying the inequalities (1.34) for the measures $M_{N_2G}(P||Q)$ and $M_{AG}(P||Q)$ along with (2.38) we get the required result. □

Proposition 2.4. *The following inequality hold:*

$$\frac{1}{4} M_{AG}(P||Q) \leqslant M_{AN_2}(P||Q). \tag{2.39}$$

Proof. Let us consider

$$g_{AG\_AN_2}(x) = \frac{f''_{AG}(x)}{f''_{AN_2}(x)} = \frac{(2x+2)^{3/2}}{(\sqrt{x}+1)\,(x-\sqrt{x}+1)}, \ x \in (0, \infty), \tag{2.40}$$

where $f''_{AG}(x)$ and $f''_{AN_2}(x)$ are as given by (2.19) and (2.25) respectively.

From (2.40), we have

$$g'_{AG\_AN_2}(x) = \frac{3\,(1-\sqrt{x})\,\sqrt{2x+2}}{\left(\sqrt{x}+1\right)^2 \left(x-\sqrt{x}+1\right)^2} \begin{cases} \leqslant 0, & x \geqslant 1 \\ \geqslant 0, & x \leqslant 1 \end{cases}. \tag{2.41}$$

In view of (2.41), we conclude that the function $g_{AG\_AN_2}(x)$ is increasing in $x \in (0, 1)$ and decreasing in $x \in (1, \infty)$, and hence

$$M = \sup_{x \in (0,\infty)} g_{AG\_AN_2}(x) = g_{AG\_AN_2}(1) = 4. \tag{2.42}$$

Applying the inequalities (1.34) for the measures $M_{AG}(P||Q)$ and $M_{AN_2}(P||Q)$ along with (2.42) we get the required result. □

Combining the results given in the Propositions 2.1-2.4, we get the proof of the theorem. The expression (2.27) can also be written as

$$\frac{1}{16}\Delta(P||Q) \leqslant M_{N_2N_1}(P||Q) \leqslant \frac{1}{3}M_{N_2G}(P||Q) \leqslant \frac{1}{4}h(P||Q) \leqslant M_{AN_2}(P||Q). \tag{2.43}$$

Remark 2.1. (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:*

$$I(P||Q) = \frac{1}{2}\left[K\left(P||\frac{P+Q}{2}\right) + K\left(Q||\frac{P+Q}{2}\right)\right] \tag{2.44}$$

and

$$J(P||Q) = K(P||Q) + K(Q||P). \tag{2.45}$$

Also we can write

$$J(P||Q) = 4\left[I(P||Q) + T(P||Q)\right], \tag{2.46}$$

where

$$T(P||Q) = \frac{1}{2}\left[K\left(\frac{P+Q}{2}||P\right) + K\left(\frac{P+Q}{2}||Q\right)\right] = \sum_{i=1}^{n} A(p_i,q_i)\ln\left(\frac{A(p_i,q_i)}{G(p_i,q_i)}\right), \tag{2.47}$$

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

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

$$\frac{1}{4}\Delta(P||Q) \leqslant I(P||Q) \leqslant h(P||Q) \leqslant 4\,M_{AN_2}(P||Q) \leqslant \frac{1}{8}J(P||Q) \leqslant T(P||Q). \tag{2.48}$$

Following the lines of the Propositions 2.1-2.4, we can also show that

$$\frac{1}{4}I(P||Q) \leqslant M_{N_2N_1}(P||Q). \tag{2.49}$$

Thus combining (2.43) with (2.46), (2.48) and (2.49), we get the following inequalities among the classical and mean divergence measures:

$$\begin{aligned} \frac{1}{16}\Delta(P||Q) &\leqslant \frac{1}{4}I(P||Q) \leqslant M_{N_2N_1}(P||Q) \\ &\leqslant \frac{1}{3}M_{N_2G}(P||Q) \leqslant \frac{1}{4}h(P||Q) \leqslant M_{AN_2}(P||Q) \\ &\leqslant \frac{1}{32}J(P||Q) \leqslant \frac{1}{4}T(P||Q) \leqslant \frac{1}{16}J(P||Q). \end{aligned} \tag{2.50}$$

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. CSISZÁR, Information Type Measures of Differences of Probability Distribution and Indirect Observations, *Studia Math. Hungarica*, **2**(1967), 299-318.

[6] I. CSISZÁR, 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 Differetiable Convex Mappings and Applications for the Shannon and Rnyi's Entropies, *RGMIA Research Report Collection*, *http://rgmia.vu.edu.au*, **..**(1999).

[9] S.S. DRAGOMIR, On An Inequality for Logrithmic 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, 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*.

[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*.

[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*, 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**(2004).

[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.

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`