Hindawi Publishing Corporation
Journal of Inequalities and Applications
Volume 2009, Article ID 323615, 15 pages
doi:10.1155/2009/323615

Research Article

On Some Improvements of the Jensen Inequality with Some Applications

M. Adil Khan,[1] M. Anwar,[1] J. Jakšetić,[2] and J. Pečarić[1,3]

[1] *Abdus Salam School of Mathematical Sciences, GC University, 5400 Lahore, Pakistan*
[2] *Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, 1000 Zagreb, Croatia*
[3] *Faculty of Textile Technology, University of Zagreb, 1000 Zagreb, Croatia*

Correspondence should be addressed to M. Adil Khan, adilbandai@yahoo.com

Received 23 April 2009; Accepted 10 August 2009

Recommended by Sever Silvestru Dragomir

An improvement of the Jensen inequality for convex and monotone function is given as well as various applications for mean. Similar results for related inequalities of the Jensen type are also obtained. Also some applications of the Cauchy mean and the Jensen inequality are discussed.

1. Introduction

The well-known Jensen's inequality for convex function is given as follows.

Theorem 1.1. *If $(\Omega, \mathbf{A}, \mu)$ is a probability space and if $f \in L^1(\mu)$ is such that $a \le f(t) \le b$ for all $t \in \Omega$, $-\infty \le a < b \le \infty$,*

$$\phi\left(\int_\Omega f(t)d\mu(t)\right) \le \int_\Omega \phi(f(t))d\mu(t) \tag{1.1}$$

is valid for any convex function $\phi : [a,b] \to \mathbb{R}$. In the case when ϕ is strictly convex on $[a,b]$ one has equality in (1.1) *if and only if f is constant almost everywhere on Ω.*

Here and in the whole paper we suppose that all integrals exist. By considering the difference of (1.1) for functional in [1] Anwar and Pečarić proved an interesting result of log-convexity. We can define this result for integrals as follows.

Theorem 1.2. *Let* $(\Omega, \mathbf{A}, \mu)$ *be a probability space and* $f \in L^1(\mu)$ *is such that* $a \leq f(t) \leq b$ *for all* $t \in \Omega$, $-\infty \leq a < b \leq \infty$. *Define*

$$\Lambda_s = \begin{cases} \dfrac{1}{s(s-1)}\left(\displaystyle\int_\Omega (f(t))^s d\mu(t) - \left(\int_\Omega f(t)d\mu(t)\right)^s\right), & s \neq 0, 1, \\ \log\left(\displaystyle\int_\Omega f(t)d\mu(t)\right) - \int_\Omega \log(f(t))d\mu(t), & s = 0, \\ \displaystyle\int_\Omega (f(t)) \log(f(t))d\mu(t) - \left(\int_\Omega f(t)d\mu(t)\right) \log\left(\int_\Omega f(t)d\mu(t)\right), & s = 1, \end{cases} \tag{1.2}$$

and let Λ_s *be positive. Then* Λ_s *is log-convex, that is, for* $-\infty < r < s < u < \infty$, *the following is valid*

$$(\Lambda_s)^{u-r} \leq (\Lambda_r)^{u-s}(\Lambda_u)^{s-r}. \tag{1.3}$$

The following improvement of (1.1) was obtained in [2].

Theorem 1.3. *Let the conditions of Theorem 1.1 be fulfilled. Then*

$$\begin{aligned} &\int_\Omega \phi(f(t))d\mu(t) - \phi\left(\int_\Omega f(t)d\mu(t)\right) \\ &\quad \geq \left| \int_\Omega \left|\phi(f(t)) - \phi\left(\overline{f}\right)\right| d\mu(t) - \left|\phi'_+\left(\overline{f}\right)\right| \int_\Omega \left|f(t) - \overline{f}\right| d\mu(t) \right|, \end{aligned} \tag{1.4}$$

where $\phi'_+(x)$ *represents the right-hand derivative of* ϕ *and*

$$\overline{f} = \int_\Omega f(t)d\mu(t). \tag{1.5}$$

If ϕ *is concave, then left-hand side of* (1.4) *should be* $\phi(\int_\Omega f(t)d\mu(t)) - \int_\Omega \phi f(t)d\mu(t)$.

In this paper, we give another proof and extension of Theorem 1.2 as well as improvements of Theorem 1.3 for monotone convex function with some applications. Also we give applications of the Jensen inequality for divergence measures in information theory and related Cauchy means.

2. Another Proof and Extension of Theorem 1.2

In fact, Theorem 1.2 for $\Omega = [a, b]$ and $0 < r < s < u$, $r, s, u \neq 1$ was first of all initiated by Simić in [3].

Moreover, in his proof, he has used convex functions defined on $I = (-\infty, 0) \cup (0, 1) \cup (1, +\infty)$ (see [3, Theorem 1]). In his proof, he has used the following function:

$$\lambda(x) = v^2 \frac{x^s}{s(s-1)} + 2vw \frac{x^r}{r(r-1)} + w^2 \frac{x^u}{u(u-1)}, \tag{2.1}$$

where $r = (s+u)/2$ and v, w, r, s, u are real with $r, s, t \in I$.

In [1] we have given correct proof by using extension of (2.1), so that it is defined on $\mathbb{R}$.

Moreover, we can give another proof so that we use only (2.1) but without using convexity as in [3].

Proof of Theorem 1.2. Consider the function $\lambda(x)$ defined, as in [3], by (2.1).

Now

$$\lambda''(x) = \left(v x^{s/2-1} + w x^{u/2-1}\right)^2 \geq 0, \quad \text{for } x > 0, \tag{2.2}$$

that is, $\lambda(x)$ is convex. By using (1.1) we get

$$v^2 \Lambda_s + 2vw\Lambda_r + w^2 \Lambda_u \geq 0. \tag{2.3}$$

Therefore, (2.3) is valid for all $s, r, u \in I$. Now since left-hand side of (2.3) is quadratic form, by the nonnegativity of it, one has

$$\Lambda^2_{(s+u)/2} = \Lambda^2_r \leq \Lambda_s \Lambda_u. \tag{2.4}$$

Since we have $\lim_{s\to 0} \Lambda_s = \Lambda_0$ and $\lim_{s \to 1} \Lambda_s = \Lambda_1$, we also have that (2.4) is valid for $r, s, u \in \mathbb{R}$. So $s \mapsto \Lambda_s$ is log-convex function in the Jensen sense on $\mathbb{R}$.

Moreover, continuity of Λ_s implies log-convexity, that is, the following is valid for $-\infty < r < s < u < \infty$:

$$(\Lambda_s)^{u-r} \leq (\Lambda_r)^{u-s} (\Lambda_u)^{s-r}. \tag{2.5}$$

□

Let us note that it was used in [4] to get corresponding Cauchy's means. Moreover, we can extend the above result.

Theorem 2.1. *Let the conditions of Theorem 1.2 be fulfilled and let p_i $(i = 1, 2, \ldots, n)$ be real numbers. Then*

$$\left|\Lambda_{p_{ij}}\right|_k \geq 0 \quad (k = 1, 2, \ldots, n), \tag{2.6}$$

where $|a_{ij}|_k$ define the determinant of order k with elements a_{ij} and $p_{ij} = (p_i + p_j)/2$.

Proof. Consider the function

$$f(x) = \sum_{i,j=1}^{n} u_i u_j \frac{x^{p_{ij}}}{p_{ij}(p_{ij}-1)} \tag{2.7}$$

for $x > 0$ and $u_i \in \mathbb{R}$ and $p_{ij} \in I$.

So, it holds that

$$f''(x) = \sum_{i,j=1}^{n} u_i u_j x^{p_{ij}-2} = \left(\sum_{i=1}^{n} u_i x^{p_i/2-1}\right)^2 \geq 0. \tag{2.8}$$

So $f(x)$ is convex function, and as a consequence of (1.1), one has

$$\sum_{i,j=1}^{n} u_i u_j \Lambda_{p_{ij}} \geq 0. \tag{2.9}$$

Therefore, $[\Lambda_{p_{ij}}]$ ($[a_{ij}]$ denote the $n \times n$ matrix with elements a_{ij}) is nonnegative semi definite and (2.6) is valid for $p_{ij} \in I$. Moreover, since we have continuity of $\Lambda_{p_{ij}}$ for all p_{ij}, (2.6) is valid for all $p_i \in \mathbb{R}$ $(i = 1, 2, \ldots, n)$. □

Remark 2.2. In Theorem 2.1, if we set $n = 2$, we get Theorem 1.2.

3. Improvements of the Jensen Inequality for Monotone Convex Function

In this section and in the following section, we denote $\overline{x} = \sum_{i=1}^{n} p_i x_i$ and $P_I = \sum_{i\in I} p_i$.

Theorem 3.1. *If $(\Omega, \mathbf{A}, \mu)$ is a probability space and if $f \in L^1(\mu)$ is such $a \leq f(t) \leq b$ for $t \in \Omega$, and if $f(t) \geq \overline{f}$ for $t \in \Omega' \subset \Omega$ (Ω' is measurable, i.e., $\Omega' \in \mathbf{A}$), $-\infty < a < b \leq \infty$, then*

$$\int_{\Omega} \phi(f(t))d\mu(t) - \phi\left(\int_{\Omega} f(t)d\mu(t)\right) \geq \left| \int_{\Omega} \operatorname{sgn}\left(f(t)-\overline{f}\right)\left[\phi(f(t)) - \phi'_{+}\left(\overline{f}\right)f(t)\right]d\mu(t) + \left[\phi\left(\overline{f}\right) - \overline{f}\phi'_{+}\left(\overline{f}\right)\right]\left[1-2\mu(\Omega')\right]\right|, \tag{3.1}$$

where

$$\overline{f} = \int_{\Omega} f(t)d\mu(t), \tag{3.2}$$

for monotone convex function $\phi : [a,b] \to \mathbb{R}$. If ϕ is monotone concave, then the left-hand side of (3.1) *should be $\phi(\int_{\Omega} f(t)d\mu(t)) - \int_{\Omega}\phi(f(t))d\mu(t)$.*

Proof. Consider the case when ϕ is nondecreasing on $[a,b]$. Then

$$
\begin{aligned}
\int_{\Omega}\left|\phi(f(t))-\phi\left(\overline{f}\right)\right|d\mu(t)
&=\int_{\Omega'}\left(\phi(f(t))-\phi\left(\overline{f}\right)\right)d\mu(t)+\int_{\Omega\setminus\Omega'}\left(\phi\left(\overline{f}\right)-\phi(f(t))\right)d\mu(t)\\
&=\int_{\Omega'}\phi(f(t))d\mu(t)-\int_{\Omega\setminus\Omega'}\phi(f(t))d\mu(t)-\phi\left(\overline{f}\right)\mu(\Omega')+\phi\left(\overline{f}\right)\mu(\Omega\setminus\Omega')\\
&=\int_{\Omega}\operatorname{sgn}\left(f(t)-\overline{f}\right)\phi(f(t))d\mu(t)+\phi\left(\overline{f}\right)(\mu(\Omega\setminus\Omega')-\mu(\Omega')).
\end{aligned}
\tag{3.3}
$$

Similarly,

$$
\int_{\Omega}\left|f(t)-\overline{f}\right|d\mu(t)=\int_{\Omega}\operatorname{sgn}\left(f(t)-\overline{f}\right)f(t)d\mu(t)+\overline{f}(\mu(\Omega\setminus\Omega')-\mu(\Omega')).
\tag{3.4}
$$

Now from (1.4), (3.3), and (3.4) we get (3.1).

The case when ϕ is nonincreasing can be treated in a similar way. □

Of course a discrete inequality is a simple consequence of Theorem 3.1.

Theorem 3.2. *Let $\phi:[a,b]\to\mathbb{R}$ be a monotone convex function, $x_i\in[a,b]$, $p_i>0$, $\sum_{i=1}^{n}p_i=1$. If $x_i\geq\overline{x}$ for $i\in I\subset\{1,2,\ldots,n\}(=I_n)$, then*

$$
\begin{aligned}
&\sum_{i=1}^{n}p_i\phi(x_i)-\phi\left(\sum_{i=1}^{n}p_ix_i\right)\\
&\qquad\geq\left|\sum_{i=1}^{n}p_i\operatorname{sgn}(x_i-\overline{x})\left[\phi(x_i)-x_i\phi'_+(\overline{x})\right]+\left[\phi(\overline{x})-\overline{x}\phi'_+(\overline{x})\right][1-2P_I]\right|.
\end{aligned}
\tag{3.5}
$$

If ϕ is monotone concave, then the left-hand side of (3.5) *should be*

$$
\phi\left(\sum_{i=1}^{n}p_ix_i\right)-\sum_{i=1}^{n}p_i\phi(x_i).
\tag{3.6}
$$

The following improvement of the Hermite-Hadamard inequality is valid [5].

Corollary 3.3. *Let $\phi : [a,b] \to \mathbb{R}$ be a differentiable convex. Then*

(i) *the inequality*

$$\frac{1}{b-a}\int_a^b \phi(t)dt - \phi\left(\frac{a+b}{2}\right) \geq \left|\frac{1}{b-a}\int_a^b \left|\phi(t) - \phi\left(\frac{a+b}{2}\right)\right| dt - \left(\frac{b-a}{4}\right)\left|\phi'\left(\frac{a+b}{2}\right)\right|\right| \tag{3.7}$$

holds.

If ϕ is differentiable concave, then the left-hand side of (3.7) *should be $\phi((a+b)/2) - (1/(b-a))\int_a^b \phi(t)dt$;*

(ii) *if ϕ is monotone, then the inequality*

$$\frac{1}{b-a}\int_a^b \phi(t)dt - \phi\left(\frac{a+b}{2}\right) \geq \left|\frac{1}{b-a}\int_a^b \operatorname{sgn}\left(t - \frac{a+b}{2}\right)\left(\phi(t) - t\phi'\left(\frac{a+b}{2}\right)\right)dt\right| \tag{3.8}$$

holds. If ϕ is differentiable and monotone concave then the left-hand side of (3.8) *should be $\phi((a+b)/2) - (1/(b-a))\int_a^b \phi(t)dt$.*

Proof. (i) Setting $\Omega = [a,b]$, $f(t) = t$, $d\mu(t) = dt/(b-a)$ in (1.4), we get (3.7).

(ii) Setting $f(t) = t$, $d\mu(t) = dt/(b-a)$, and $\Omega = [a,b]$ in (3.1), we get (3.8). □

4. Improvements of the Levinson Inequality

Theorem 4.1. *If the third derivative of f exist and is nonnegative, then for $0 < x_i < a$, $p_i > 0$ $(1 \leq i \leq n)$, $\sum_{i=1}^n p_i = 1$ and $P_k = \sum_{i=1}^k p_i$ $(2 \leq k \leq n-1)$ one has*

(i)

$$\sum_{i=1}^n p_i f(2a - x_i) - f(2a - \overline{x}) - \sum_{i=1}^n p_i f(x_i) + f(\overline{x}) \geq \left|\sum_{i=1}^n p_i \left|f(2a - x_i) - f(x_i) - f(2a - \overline{x}) + f(\overline{x})\right| - \left|f'(2a - \overline{x}) + f'(\overline{x})\right| \sum_{i=1}^n p_i |x_i - \overline{x}|\right|, \tag{4.1}$$

(ii) *if* $\phi(x) = f(2a - x) - f(x)$ *is monotone and* $x_i \geq \overline{x}$ *for* $i \in I \subset \{1, 2, \ldots, n\} = I_n$, *then*

$$\begin{aligned}
&\sum_{i=1}^{n} p_i f(2a - x_i) - f(2a - \overline{x}) - \sum_{i=1}^{n} p_i f(x_i) + f(\overline{x}) \\
&\quad \geq \Bigg| \sum_{i=1}^{n} p_i \operatorname{sgn}(x_i - \overline{x}) \left[f(2a - x_i) - f(x_i) + x_i \left(f'(2a - \overline{x}) + f'(\overline{x}) \right) \right] \\
&\qquad + \left[f(2a - \overline{x}) - f(\overline{x}) + \overline{x} \left(f'(2a - \overline{x}) + f'(\overline{x}) \right) \right] [1 - 2P_I] \Bigg|.
\end{aligned} \tag{4.2}$$

Proof. (i) As for 3-convex function $f : [0, 2a] \to \mathbb{R}$ the function $\phi(x) = f(2a - x) - f(x)$ is convex on $[0, a]$, so by setting $\phi = f(2a - x) - f(x)$ in the discrete case of [2, Theorem 2], we get (4.1).

(ii) As $f(2a - x) - f(x)$ is monotone convex, so by setting $\phi = f(2a - x) - f(x)$ in (3.5), we get (5.16). □

Ky Fan Inequality

Let $x_i \in (0, 1/2]$ be such that $x_1 \geq x_2 \geq \cdots \geq x_k \geq \overline{x} \geq x_{k+1} \cdots \geq x_n$. We denote G_k and A_k, the weighted geometric and arithmetic means, respectively, that is,

$$A_k = \frac{1}{P_k} \left(\sum_{i=1}^{k} p_i x_i \right) (= \overline{x}), \qquad G_k = \left(\prod_{i=1}^{k} x_i^{p_i} \right)^{1/P_k}, \tag{4.3}$$

and also by A'_k and G'_k, the arithmetic and geometric means of $1 - x_i$, respectively, that is,

$$A'_k = \frac{1}{P_k} \sum_{i=1}^{k} p_i (1 - x_i) (= 1 - A_k), \qquad G'_k = \left(\prod_{i=1}^{k} (1 - x_i)^{p_i} \right)^{1/P_k}. \tag{4.4}$$

The following remarkable inequality, due to Ky Fan, is valid [6, page 5],

$$\frac{G_n}{G'_n} \leq \frac{A_n}{A'_n}, \tag{4.5}$$

with equality sign if and only if $x_1 = x_2 = \cdots = x_n$.

Inequality (4.5) has evoked the interest of several mathematicians and in numerous articles new proofs, extensions, refinements and various related results have been published [7].

The following improvement of Ky Fan inequality is valid [2].

Corollary 4.2. *Let A_n, G_n and A'_n, G'_n be as defined earlier. Then, the following inequalities are valid*

(i)

$$\frac{A_n/A'_n}{G_n/G'_n} \geq \exp\left(\left|\sum_{i=1}^{n} p_i \left|\ln\left(\frac{(1-x_i)A_n}{x_i A'_n}\right)\right| - \frac{1}{A_n A'_n}\sum_{i=1}^{n} p_i |x_i - A_n|\right|\right), \tag{4.6}$$

(ii)

$$\frac{A_n/A'_n}{G_n/G'_n} \geq \exp\left[\left|2P_k\left\{\ln\left(\frac{G'_k A_n}{G_k A'_n}\right) + \frac{A_k - A_n}{A_n A'_n}\right\} + \ln\left(\frac{G_n A'_n}{A_n G'_n}\right)\right|\right]. \tag{4.7}$$

Proof. (i) Setting $a = 1/2$, $f(x) = \ln x$ in (4.1), we get (4.6).

(ii) Consider $a = 1/2$ and $f(x) = \ln x$, then $\phi(x) = \ln(1-x) - \ln x$ is strictly monotone convex on the interval $(0, 1/2)$ and has derivative

$$\phi'(x) = -\frac{1}{x(x-1)}. \tag{4.8}$$

Then the application of inequality (4.2) to this function is given by

$$\begin{aligned} &\sum_{i=1}^{n} p_i \ln\frac{1-x_i}{x_i} - \ln\frac{1-\overline{x}}{\overline{x}} \\ &\quad \geq \left|\sum_{i=1}^{n} p_i \operatorname{sgn}(x_i - \overline{x})\left[\ln\frac{1-x_i}{x_i} + \frac{x_i}{\overline{x}(1-\overline{x})}\right] + \left[\ln\frac{1-\overline{x}}{\overline{x}} + \frac{1}{1-\overline{x}}\right](1-2P_k)\right|. \end{aligned} \tag{4.9}$$

From (4.9) we get (4.7). □

5. On Some Inequalities for Csiszár Divergence Measures

Let $(\Omega, \mathbf{A}, \mu)$ be a measure space satisfying $|\mathbf{A}| > 2$ and μ a σ-finite measure on Ω with values in $\mathbb{R} \cup \{\infty\}$. Let $\mathbf{P}$ be the set of all probability measures on the measurable space $(\Omega, \mathbf{A})$ which are absolutely continuous with respect to μ. For $P, Q \in \mathbf{P}$, let $p = dP/d\mu$ and $q = dQ/d\mu$ denote the Radon-Nikodym derivatives of P and Q with respect to μ, respectively.

Csiszár introduced the concept of f-divergence for a convex function, $f : [0, \infty) \to (-\infty, \infty)$ that is continuous at 0 as follows (cf. [8], see also [9]).

Definition 5.1. Let $P, Q \in \mathbf{P}$. Then

$$I_f(Q, P) = \int_{\Omega} p(s) f\left(\frac{q(s)}{p(s)}\right) d\mu(s), \tag{5.1}$$

is called the f-divergence of the probability distributions Q and P.

We give some important f-divergences, playing a significant role in Information Theory and statistics.

(i) The class of χ-divergences: the f-divergences, in this class, are generated by the family of functions:

$$f_\alpha(u) = |u-1|^\alpha \quad u \geq 0,\ \alpha \geq 1,$$
$$I_{f_\alpha}(Q,P) = \int_\Omega p^{1-\alpha}(s)|q(s)-p(s)|^\alpha d\mu(s). \tag{5.2}$$

For $\alpha = 1$, it gives the total variation distance:

$$V(Q,P) = \int_\Omega \left|q(s)-p(s)\right| d\mu(s). \tag{5.3}$$

For $\alpha = 2$, it gives the Karl pearson χ^2-divergence:

$$I_{\chi^2}(Q,P) = \int_\Omega \frac{[q(s)-p(s)]^2}{p(s)} d\mu(s). \tag{5.4}$$

(ii)The α-order Renyi entropy: for $\alpha \in \mathbb{R} \setminus \{0,1\}$, let

$$f(t) = t^\alpha, \quad t > 0. \tag{5.5}$$

Then I_f gives α-order entropy

$$D_\alpha(Q,P) = \int_\Omega q^\alpha(s) p^{1-\alpha}(s) d\mu(s). \tag{5.6}$$

(iii) Harmonic distance: let

$$f(t) = \frac{2t}{1+t}, \quad t > 0. \tag{5.7}$$

Then I_f gives Harmonic distance

$$D_H(Q,P) = \int_\Omega \frac{2p(s)q(s)}{p(s)+q(s)} d\mu(s). \tag{5.8}$$

(iv) Kullback-Leibler: let

$$f(t) = t\log t, \quad t > 0. \tag{5.9}$$

Then f-divergence functional gives rise to Kullback-Leibler distance [10]

$$D_{\mathrm{KL}}(Q,P)=\int_{\Omega}q(s)\log\left(\frac{q(s)}{p(s)}\right)d\mu(s). \tag{5.10}$$

The one parametric generalization of the Kullback-Leibler [10] relative information studied in a different way by Cressie and Read [11].

(v) The Dichotomy class: this class is generated by the family of functions $g_\alpha : [0,\infty) \to \mathbb{R}$,

$$g_\alpha(u)=\begin{cases} u-1-\log u, & \alpha=0,\\ \dfrac{1}{\alpha(1-\alpha)}\left[\alpha u+1-\alpha-u^{\alpha}\right], & \alpha\in\mathbb{R}\setminus\{0,1\},\\ 1-u+u\log u, & \alpha=1. \end{cases} \tag{5.11}$$

This class gives, for particular values of α, some important divergences. For instance, for $\alpha = 1/2$, we have Hellinger distance and some other divergences for this class are given by

$$I_{g_\alpha}(Q,P)=\begin{cases} Q-P+D_{\mathrm{KL}}(P,Q), & \alpha=0,\\ \dfrac{\alpha(Q-P)+P-D_\alpha(Q,P)}{\alpha(1-\alpha)}, & \alpha\in\mathbb{R}\setminus\{0,1\},\\ D_{\mathrm{KL}}(Q,P)+P-Q, & \alpha=1, \end{cases} \tag{5.12}$$

where $p(x)$ and $q(x)$ are positive integrable functions with $\int_\Omega p(s)d\mu(s) = P$, $\int_\Omega q(s)d\mu(s) = Q$.

There are various other divergences in Information Theory and statistics such as Arimoto-type divergences, Matushita's divergence, Puri-Vincze divergences (cf. [12–14]) used in various problems in Information Theory and statistics. An application of Theorem 1.1 is the following result given by Csiszár and Körner (cf. [15]).

Theorem 5.2. *Let $f : [0,\infty) \to R$ be convex, and let p and q be positive integrable function with $\int_\Omega p(s)d\mu(s) = P$, $\int_\Omega q(s)d\mu(s) = Q$. Then the following inequality is valid:*

$$I_f(P,Q)\geq Qf\left(\frac{P}{Q}\right), \tag{5.13}$$

where $I_f(P,Q) = \int_\Omega q(s)f(p(s)/q(s))d\mu(s)$.

Proof. By substituting $\phi(s) = f(s)$, $f(s) = p(s)/q(s)$ and $d\mu(s) = q(s)d\mu(s)$ in Theorem 1.1 we get (5.13). □

Similar consequence of Theorems 1.2 and 2.1 in information theory for divergence measures discussed above is the following result.

Theorem 5.3. *Let p and q be positive integrable functions with $\int_\Omega p(s)d\mu(s) = P$, $\int_\Omega q(s)d\mu(s) = Q$. Define the function*

$$\Phi_t = \begin{cases} \dfrac{1}{t(1-t)}\left(P^t Q^{1-t} - D_t(P,Q)\right), & t \neq 0,1, \\ D_{\mathrm{KL}}(Q,P) + Q\log\dfrac{P}{Q}, & t = 0, \\ D_{\mathrm{KL}}(P,Q) + P\log\dfrac{Q}{P}, & t = 1, \end{cases} \tag{5.14}$$

and let Φ_t be positive. Then

(i) *it holds that*

$$\left|\Phi_{p_{ij}}\right|_k \geq 0 \quad (k = 1,2,\ldots,n), \tag{5.15}$$

where $|a_{ij}|_k$ define the determinant of order n with elements a_{ij} and $p_{ij} = (p_i + p_j)/2$,

(ii) *Φ_t is log-convex.*

As we said in [4] we define new means of the Cauchy type, here we define an application of these means for divergence measures in the following definition.

Definition 5.4. Let p and q be positive integrable functions with $\int_\Omega p(s)d\mu(s) = P$, $\int_\Omega q(s)d\mu(s) = Q$. The mean $\mathrm{M}_{s,t}$ is defined as

$$\begin{aligned} \mathrm{M}_{s,t} &= \left(\frac{\Phi_s}{\Phi_t}\right)^{1/(s-t)}, \quad s \neq t \neq 0,1, \\ \mathrm{M}_{s,s} &= \exp\left(\frac{P^s Q^{1-s}\log(P/Q) - D_s'(P,Q)}{P^s Q^{1-s} - D_s(P,Q)} - \frac{1-2s}{s(1-s)}\right), \quad s \neq 0,1, \\ \mathrm{M}_{0,0} &= \exp\left(\frac{Q\left(\log(P/Q)\right)^2 - D_0''(P,Q)}{2\left(Q\log(P/Q) - D_0'(P,Q)\right)} + 1\right), \end{aligned} \tag{5.16}$$

where $D_0'(P,Q) = \int_\Omega q(s)\log(p(s)/q(s))d\mu(s)$ and $D_0''(P,Q) = \int_\Omega q(s)\log\left(p(s)/q(s)\right)^2 d\mu(s)$,

$$\mathrm{M}_{1,1} = \exp\left(\frac{Q\left(\log(P/Q)\right)^2 - D_1''(P,Q)}{2\left(P\log(P/Q) - D_1'(P,Q)\right)} - 1\right), \tag{5.17}$$

where $D_1'(P,Q) = \int_\Omega p(s)\log(q(s)/p(s))d\mu(s)$ and $D_1''(P,Q) = \int_\Omega p(s)\log\left(q(s)/p(s)\right)^2 d\mu(s)$.

Theorem 5.5. *Let r, s, t, u be nonnegative reals such that $r \leq t$, $s \leq u$, then*

$$\mathrm{M}_{r,t} \leq \mathrm{M}_{s,u}. \tag{5.18}$$

Proof. By using log convexity of Φ_t, we get the following result for $r, s, t, u \in \mathbb{R}$ such that $r \leq t,\ s \leq u$ and $r \neq s,\ t \neq u$

$$\left(\frac{\Phi_s}{\Phi_r}\right)^{1/(s-r)} \leq \left(\frac{\Phi_u}{\Phi_t}\right)^{1/(u-t)}. \tag{5.19}$$

Also for $r = s,\ t = u$, we consider limiting case and the result follows from continuity of $\mathrm{M}_{s,u}$. □

An application of Theorem 1.3 in divergence measure is the following result given in [16].

Theorem 5.6. *Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be differentiable convex function on $\overset{o}{I}$, then*

$$I_f(P,Q) - Qf\left(\frac{P}{Q}\right) \geq \left|\left|I_f(P,Q) - Qf\left(\frac{P}{Q}\right)\right| - \frac{|f'(P/Q)|}{Q}\overline{Q}\right|, \tag{5.20}$$

where

$$\overline{Q} = \int_{\Omega} |Qp(s) - Pq(s)| d\mu(s). \tag{5.21}$$

Proof. By substituting $\phi(s) = f(s),\ f(s) = p(s)/q(s)$, and $d\mu(s) \mapsto q(s)d\mu(s)$ in Theorem 1.3, we get (5.20). □

Theorem 5.7. *Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be differentiable monotone convex function on $\overset{o}{I}$ and let $p(s)\ /q(s) > P/Q$ for $s \in \Omega' \subset \Omega$*

$$\begin{aligned} I_f(P,Q) - Qf\left(\frac{P}{Q}\right) \geq & \int_{\Omega} \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right) f\left(\frac{p(s)}{q(s)}\right) q(s) d\mu(s) \\ & - f'\left(\frac{P}{Q}\right) \int_{\Omega} \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right) p(s) d\mu(s) \\ & + Q\left[f\left(\frac{P}{Q}\right) - \frac{P}{Q} f'\left(\frac{P}{Q}\right)\right]\left[1 - 2\frac{Q'}{Q}\right], \end{aligned} \tag{5.22}$$

where

$$Q' = \int_{\Omega'} q(s) d\mu(s), \tag{5.23}$$

and Ω' as in Theorem 5.7.

Proof. By substituting $\phi(s) = f(s),\ f(s) = p(s)/q(s)$ and $d\mu(s) \mapsto q(s)d\mu(s)$ in Theorem 3.1(ii) we get (5.22). □

Corollary 5.8. *It holds that*

$$D_{H_\alpha}(P,Q) - \frac{2PQ}{P+Q} \geq \int_\Omega \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)\frac{2p(s)q(s)}{p(s)+q(s)}d\mu(s)$$
$$- \frac{2Q^2}{(P+Q)^2}\int_\Omega \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)p(s)d\mu(s) \tag{5.24}$$
$$+ \left[\frac{2PQ}{P+Q} - \frac{2PQ^2}{(P+Q)^2}\right]\left[1 - \frac{2Q'}{Q}\right],$$

where

$$Q' = \int_{\Omega'} q(s)d\mu(s), \tag{5.25}$$

and Ω' *as in Theorem 5.7.*

Proof. The proof follows by setting $f(t) = 2t/(1+t),\ t > 0$ in Theorem 5.7. □

Corollary 5.9. *Let* $g_\alpha : \mathbb{R}_+ \to \mathbb{R}$ *be as given in* (5.11), *then*

(i) *for* $\alpha = 0$ *one has*

$$D_{\mathrm{KL}}(Q,P) + Q\log\left(\frac{P}{Q}\right)$$
$$\geq \int_\Omega \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)\left(\frac{p(s)}{q(s)} - 1 - \log\left(\frac{p(s)}{q(s)}\right)\right)q(s)d\mu(s) \tag{5.26}$$
$$- \frac{P-Q}{P}\int_\Omega \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)p(s)d\mu(s) + Q\ \log\left(\frac{P}{Q}\right)\left(1 - \frac{2Q'}{Q}\right),$$

(ii) *for* $\alpha \in \mathbb{R}\setminus\{0,1\}$ *one has*

$$\frac{P^\alpha Q^{1-\alpha} - D_\alpha(P,Q)}{\alpha(1-\alpha)} \geq \int_\Omega \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)\left(\alpha\frac{p(s)}{q(s)} + 1 - \alpha - p(s)^\alpha q(s)^{-\alpha}\right)q(s)d\mu(s)$$
$$- \alpha\left(1 - P^{\alpha-1}Q^{1-\alpha}\right)\int_\Omega \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)p(s)d\mu(s)$$
$$+ \frac{\left[\alpha\ P\ +\ Q - \alpha\ Q + P\ Q^{(1-\alpha)} - \alpha\ P/Q\left(1 - P^{(1-\alpha)}\ Q^{(1-\alpha)}\right)\right](1 - 2Q'/Q)}{\alpha(1-\alpha)}, \tag{5.27}$$

(iii) *for* $\alpha = 1$ *one has*

$$D_{\mathrm{KL}}(P,Q) + \frac{P}{Q}\log\left(\frac{P}{Q}\right) \geq \int_{\Omega} \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)\left(1 - \frac{p(s)}{q(s)} + \frac{p(s)}{q(s)}\log\frac{p(s)}{q(s)}\right)q(s)d\mu(s)$$
$$- \log\left(\frac{P}{Q}\right)\int_{\Omega} \operatorname{sgn}\left(\frac{p(s)}{q(s)} - \frac{P}{Q}\right)p(s)d\mu(s) \quad (5.28)$$
$$+ [Q - P]\left[1 - \frac{2Q'}{Q}\right],$$

where

$$Q' = \int_{\Omega'} q(s)d\mu(s), \quad (5.29)$$

and Ω' *as in Theorem 5.7.*

Proof. The proof follows be setting $f = g_\alpha$ to be as given in (5.11), in Theorem 3.1. □

Acknowledgments

This research work is funded by the Higher Education Commission Pakistan. The research of the fourth author is supported by the Croatian Ministry of Science, Education and Sports under the Research Grants 117-1170889-0888.

References

[1] M. Anwar and J. Pečarić, "On logarithmic convexity for differences of power means and related results," *Mathematical Inequalities & Applications*, vol. 12, no. 1, pp. 81–90, 2009.
[2] S. Hussain and J. Pečarić, "An improvement of Jensen's inequality with some applications," *Asian-European Journal of Mathematics*, vol. 2, no. 1, pp. 85–94, 2009.
[3] S. Simić, "On logarithmic convexity for differences of power means," *Journal of Inequalities and Applications*, vol. 2007, Article ID 37359, 8 pages, 2007.
[4] M. Anwar and J. Pečarić, "New means of Cauchy's type," *Journal of Inequalities and Applications*, vol. 2008, Article ID 163202, p. 10, 2008.
[5] S. S. Dragomir and A. McAndrew, "Refinements of the Hermite-Hadamard inequality for convex functions," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 6, no. 2, article 140, 6 pages, 2005.
[6] H. Alzer, "The inequality of Ky Fan and related results," *Acta Applicandae Mathematicae*, vol. 38, no. 3, pp. 305–354, 1995.
[7] E. F. Beckenbach and R. Bellman, *Inequalities*, vol. 30 of *Ergebnisse der Mathematik und ihrer Grenzgebiete, N. F.*, Springer, Berlin, Germany, 1961.
[8] I. Csiszár, "Information measures: a critical survey," in *Transactions of the 7th Prague Conference on Information Theory, Statistical Decision Functions and the 8th European Meeting of Statisticians*, pp. 73–86, Academia, Prague, Czech Republic, 1978.
[9] M. C. Pardo and I. Vajda, "On asymptotic properties of information-theoretic divergences," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1860–1868, 2003.
[10] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79–86, 1951.

[11] P. Cressie and T. R. C. Read, "Multinomial goodness-of-fit tests," *Journal of the Royal Statistical Society. Series B*, vol. 46, no. 3, pp. 440–464, 1984.

[12] P. Kafka, F. Österreicher, and I. Vincze, "On powers of f-divergences defining a distance," *Studia Scientiarum Mathematicarum Hungarica*, vol. 26, no. 4, pp. 415–422, 1991.

[13] F. Liese and I. Vajda, *Convex Statistical Distances*, vol. 95 of *Teubner Texts in Mathematics*, BSB B. G. Teubner Verlagsgesellschaft, Leipzig, Germany, 1987.

[14] F. Österreicher and I. Vajda, "A new class of metric divergences on probability spaces and its applicability in statistics," *Annals of the Institute of Statistical Mathematics*, vol. 55, no. 3, pp. 639–653, 2003.

[15] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless System*, Probability and Mathematical Statistics, Academic Press, New York, NY, USA, 1981.

[16] M. Anwar, S. Hussain, and J. Pečarić, "Some inequalities for Csiszár-divergence measures," *International Journal of Mathematical Analysis*, vol. 3, no. 26, pp. 1295–1304, 2009.