Adil Khan et al. *Journal of Inequalities and Applications* (2020) 2020:76
https://doi.org/10.1186/s13660-020-02343-7

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH **Open Access**

New refinement of the Jensen inequality associated to certain functions with applications

Muhammad Adil Khan[1*], Đilda Pečarić[2] and Josip Pečarić[3]

*Correspondence:
adilswati@gmail.com
[1]Department of Mathematics, University of Peshawar, Peshawar, Pakistan
Full list of author information is available at the end of the article

Abstract

This article proposes a new refinement of the celebrated Jensen inequality. Some refinements have been obtained for quasi-arithmetic means, Hölder and Hermite–Hadamard inequalities. Several applications are given in information theory. A more general refinement of Jensen inequality is presented associated to n functions.

MSC: 26D15; 94A17; 94A15

Keywords: Jensen's inequality; Convex functions; Hölder inequality; Means; Csiszár divergence; Shannon entropy

1 Introduction

The celebrated Jensen inequality states that: If I is an interval in $\mathbb{R}$ and $g, p : [a,b] \to \mathbb{R}$ are integrable functions such that $g(\varrho) \in I, p(\varrho) > 0\ \forall \varrho \in [a,b]$. Also, if $\psi : I \to \mathbb{R}$ is convex function and $(\psi \circ g).p$ is integrable on $[a,b]$. Then

$$\psi\left(\frac{\int_a^b g(\varrho)p(\varrho)\,d\varrho}{\int_a^b p(\varrho)\,d\varrho}\right) \leq \frac{\int_a^b p(\varrho)(\psi \circ g)(\varrho)\,d\varrho}{\int_a^b p(\varrho)\,d\varrho}. \tag{1}$$

Jensen's inequality is one of the fundamental inequalities in mathematics and it underlies many vital statistical concepts and proofs. Some important applications involve derivation of the AM-GM inequality, estimations for Zipf–Mandelbrot and Shannon entropies, the convergence property of the expectation maximization algorithm, and positivity of Kullback–Leibler divergence [1–7]. Also, this inequality has been utilized to solve several problems in many areas of science and technology e.g. physics, engineering, financial economics and computer science.

There are several classical important inequalities which may be deduced from (1), for example Hölder, Levinson's, and Ky Fan and Young's inequalities. Due to the great importance of this inequality, several researchers have focused on this inequality and derived many improvements, refinements and extensions of the Jensen inequality. The Jensen inequality also has been given for some other generalized convex functions such as s-convex, preinvex, h-convex and η-convex functions. For some recent results concerning the Jensen inequality see [1–3, 5, 8–20].

In this article first of all we establish an interesting refinement of the Jensen inequality associated to two functions whose sum is equal to unity. Using this refinement, we derive refinements of Hölder, power mean, quasi-arithmetic mean and Hermite–Hadamard inequalities. We also focus on deducing bounds for Csiszár-divergence, Kullback–Leibler divergence, Shannon entropy and variational distance etc. We present a more general refinement of Jensen inequality concerning n functions whose sums are equal to unity.

2 Main results

We start to derive a new refinement of the Jensen inequality associated to two functions whose sum is equal to unity.

Theorem 1 *Let $\psi : I \to \mathbb{R}$ be a convex function defined on the interval I. Let $p, u, v, g : [a, b] \to \mathbb{R}$ be integrable functions such that $g(\varrho) \in I, u(\varrho), v(\varrho), p(\varrho) \in \mathbb{R}^+$ for all $\varrho \in [a, b]$ and $v(\varrho) + u(\varrho) = 1$, $P = \int_a^b p(\varrho)\,d\varrho$. Then*

$$\begin{aligned}
&\frac{1}{P}\int_a^b p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho \\
&\quad\geq \frac{1}{P}\int_a^b u(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b p(\varrho)u(\varrho)g(\varrho)\,d\varrho}{\int_a^b p(\varrho)u(\varrho)\,d\varrho}\right) \\
&\qquad+ \frac{1}{P}\int_a^b p(\varrho)v(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b p(\varrho)v(\varrho)g(\varrho)\,d\varrho}{\int_a^b p(\varrho)v(\varrho)\,d\varrho}\right) \\
&\quad\geq \psi\left(\frac{1}{P}\int_a^b p(\varrho)g(\varrho)\,d\varrho\right).
\end{aligned} \tag{2}$$

If the function ψ is concave then the reverse inequalities hold in (2).

Proof Since $u(\varrho) + v(\varrho) = 1$, so we have

$$\int_a^b p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho = \int_a^b u(\varrho)p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho + \int_a^b v(\varrho)p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho. \tag{3}$$

Applying the integral Jensen inequality on both terms on the right side of (3) we obtain

$$\begin{aligned}
&\frac{1}{P}\int_a^b p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho \\
&\quad\geq \frac{1}{P}\int_a^b u(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b u(\varrho)p(\varrho)g(\varrho)\,d\varrho}{\int_a^b u(\varrho)p(\varrho)\,d\varrho}\right) \\
&\qquad+ \frac{1}{P}\int_a^b v(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b v(\varrho)p(\varrho)g(\varrho)\,d\varrho}{\int_a^b v(\varrho)p(\varrho)\,d\varrho}\right) \\
&\quad\geq \psi\left(\frac{1}{P}\int_a^b u(\varrho)p(\varrho)g(\varrho)\,d\varrho + \frac{1}{P}\int_a^b v(\varrho)p(\varrho)g(\varrho)\,d\varrho\right) \\
&\qquad\text{(by the convexity of } \psi)
\end{aligned}$$

$$= \psi\left(\frac{1}{P}\int_a^b p(\varrho)g(\varrho)\,d\varrho\right). \tag{4}$$

□

As a consequence of the above theorem we deduce the following refinement of the Hölder inequality.

Corollary 1 *Let $r_1, r_2 > 1$ be such that $\frac{1}{r_1} + \frac{1}{r_2} = 1$. If u, v, τ, g_1 and g_2 are non-negative functions defined on $[a,b]$ such that $\tau g_1^{r_1}, \tau g_2^{r_2}, u\tau g_2^{r_2}, v\tau g_2^{r_2}, u\tau g_1 g_2, v\tau g_1 g_2, \tau g_1 g_2 \in L^1([a,b])$ and $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$, then*

$$\begin{aligned}
&\left(\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}}\left(\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\quad\geq \left(\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\qquad\times\left\{\left(\int_a^b u(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{1-r_1}\left(\int_a^b u(\varrho)\tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho\right)^{r_1}\right. \\
&\qquad\left.+\left(\int_a^b v(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{1-r_1}\left(\int_a^b v(\varrho)\tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho\right)^{r_1}\right\}^{\frac{1}{r_1}} \\
&\quad\geq \int_a^b \tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho.
\end{aligned} \tag{5}$$

In the case when $0 < r_1 < 1$ and $r_2 = \frac{r_1}{r_1-1}$ with $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho > 0$ or $r_1 < 0$ and $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho > 0$, then we have

$$\begin{aligned}
&\int_a^b \tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho \\
&\quad\geq \left(\int_a^b u(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}}\left(\int_a^b u(\varrho)\tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}} \\
&\qquad+\left(\int_a^b v(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}}\left(\int_a^b v(\varrho)\tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}} \\
&\quad\geq \left(\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}}\left(\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}}.
\end{aligned} \tag{6}$$

Proof If $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho > 0$, then by using Theorem 1 for $\psi(\varrho) = \varrho^{r_1}, \varrho > 0, r_1 > 1, p(\varrho) = \tau(\varrho)g_2^{r_2}(\varrho), g(\varrho) = g_1(\varrho)g_2^{\frac{-r_2}{r_1}}(\varrho)$, we obtain (5). If $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho > 0$, then applying the same procedure but taking r_1, r_2, g_1, g_2 instead of r_2, r_1, g_2, g_1, we obtain (5).

Set $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho = 0$ and $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho = 0$. We know that

$$0 \leq \tau(\varrho)g_1(\varrho)g_2(\varrho) \leq \frac{1}{r_1}\tau(\varrho)g_1^{r_1}(\varrho) + \frac{1}{r_2}\tau(\varrho)g_2^{r_2}(\varrho). \tag{7}$$

Therefore taking the integral and then using the given conditions we have $\int_a^b \tau(\varrho)g_1(\varrho) \times g_2(\varrho)\,d\varrho = 0$.

For the case $r_1 > 1$, the proof is completed.

For the case when $0 < r_1 < 1$, $M = \frac{1}{r_1} > 1$ and applying (5) for M and $N = (1-r_1)^{-1}, \overline{g}_1 = (g_1g_2)^{r_1}, \overline{g}_2 = g_2^{-r_1}$ instead of r_1, r_2, g_1, g_2.

Finally, if $r_1 < 0$ then $0 < r_2 < 1$ and we may apply similar arguments with r_1, r_2, g_1, g_2 replaced by r_2, r_1, g_2, g_1 provided that $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho > 0$. □

Another refinement of the Hölder inequality presented in the following corollary.

Corollary 2 *Let $r_1 > 1, r_2 = \frac{r_1}{r_1-1}$. If u, v, τ, g_1 and g_2 are non-negative functions defined on $[a,b]$ such that $\tau g_1^{r_1}, \tau g_2^{r_2}, u\tau g_2^{r_2}, v\tau g_2^{r_2}, \tau g_1 g_2 \in L^1([a,b])$ and $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$, also assuming that $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho) > 0$, then*

$$\begin{aligned}
&\left(\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}}\left(\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\quad\geq \left(\int_a^b u(\varrho)\tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}}\left(\int_a^b u(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\qquad+ \left(\int_a^b v(\varrho)\tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}}\left(\int_a^b v(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\quad\geq \int_a^b \tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho. \qquad (8)
\end{aligned}$$

In the case when $0 < r_1 < 1$ and $r_2 = \frac{r_1}{r_1-1}$ with $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho > 0$ or $r_1 < 0$ and $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho > 0$, then we have

$$\begin{aligned}
&\left(\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho\right)^{\frac{1}{r_1}}\left(\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\quad\leq \left(\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{\frac{1}{r_2}} \\
&\qquad\times \left\{\left(\int_a^b u(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{1-r_1}\left(\int_a^b u(\varrho)\tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho\right)^{r_1}\right. \\
&\qquad\left.+ \left(\int_a^b v(\varrho)\tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho\right)^{1-r_1}\left(\int_a^b v(\varrho)\tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho\right)^{r_1}\right\}^{\frac{1}{r_1}} \\
&\quad\leq \int_a^b \tau(\varrho)g_1(\varrho)g_2(\varrho)\,d\varrho. \qquad (9)
\end{aligned}$$

Proof Assume that $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho > 0$. Let $\psi(\varrho) = \varrho^{\frac{1}{r_1}}$, $\varrho > 0, r_1 > 1$. Then clearly the function ψ is concave. Therefore applying Theorem 1 for $\psi(\varrho) = \varrho^{\frac{1}{r_1}}, p = \tau g_2^{r_2}, g = g_1^{r_1}g_2^{-r_2}$, we obtain (8). If $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho > 0$, then applying the same procedure but taking r_1, r_2, g_1, g_2 instead of r_2, r_1, g_2, g_1, we obtain (8).

If $\int_a^b \tau(\varrho)g_2^{r_2}(\varrho)\,d\varrho = 0$ and $\int_a^b \tau(\varrho)g_1^{r_1}(\varrho)\,d\varrho = 0$, then since as we know that

$$0 \leq \tau(\varrho)g_1(\varrho)g_2(\varrho) \leq \frac{1}{r_1}\tau(\varrho)g_1^{r_1}(\varrho) + \frac{1}{r_2}\tau(\varrho)g_2^{r_2}(\varrho). \qquad (10)$$

Therefore taking the integral and then using the given conditions we have $\int_a^b \tau(\varrho) \times g_1(\varrho)g_2(\varrho)\,d\varrho = 0$.

In the case when $0 < r_1 < 1$, $M = \frac{1}{r_1} > 1$ and applying (8) for M and $N = (1-r_1)^{-1}, \overline{g}_1 = (g_1 g_2)^{r_1}, \overline{g}_2 = g_2^{-r_1}$ instead of r_1, r_2, g_1, g_2, we get (9).

Finally, if $r_1 < 0$ then $0 < r_2 < 1$ and we may apply similar arguments with r_1, r_2, g_1, g_2 replaced by r_2, r_1, g_2, g_1 provided that $\int_a^b \tau(\varrho) g_1^{r_1}(\varrho)\, d\varrho > 0$. □

Remark 1 If we put $u(\varrho) = \frac{b-\varrho}{b-a}, v(\varrho) = \frac{\varrho-a}{b-a}$ in (8), then we deduce the inequalities which have been obtained by Işcan in [21].

Let p and g be positive integrable functions defined on $[a,b]$. Then the integral power means of order $r \in \mathbb{R}$ are defined as follows:

$$M_r(p;g) = \begin{cases} \left(\frac{1}{\int_a^b p(\varrho)\, d\varrho} \int_a^b p(\varrho) g^r(\varrho)\, d\varrho\right)^{\frac{1}{r}}, & \text{if } r \neq 0, \\ \exp\left(\frac{\int_a^b p(\varrho) \log g(\varrho)\, d\varrho}{\int_a^b p(\varrho)\, d\varrho}\right), & \text{if } r = 0. \end{cases} \tag{11}$$

In the following corollary we deduce inequalities for power means.

Corollary 3 *Let p, u, v and g be positive integrable functions defined on $[a,b]$ with $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$. Let $s, t \in \mathbb{R}$ such that $s \leq t$. Then*

$$M_t(p;g) \geq \left[M_1(u;p) M_s^t(u.p;g) + M_1(v;p) M_s^t(v.p;g)\right]^{\frac{1}{t}} \geq M_s(p;g), \quad t \neq 0, \tag{12}$$

$$M_t(p;g) \geq M_1(u;p) \log M_s(u.p;g) + M_1(v;p) \log M_s(v.p;g) \geq M_s(p;g), \quad t = 0, \tag{13}$$

$$M_s(p;g) \leq \left[M_1(u;p) M_t^s(u.p;g) + M_1(v;p) M_t^s(v.p;g)\right]^{\frac{1}{s}} \leq M_t(p;g), \quad s \neq 0, \tag{14}$$

$$M_s(p;g) \leq M_1(u;p) \log M_t(u.p;g) + M_1(v;p) \log M_t(v.p;g) \leq M_t(p;g), \quad s = 0. \tag{15}$$

Proof If $s, t \in \mathbb{R}$ and $s, t \neq 0$, then using (2) for $\psi(\varrho) = \varrho^{\frac{t}{s}}, \varrho > 0, g \to g^s$ and then taking the power $\frac{1}{t}$ we get (12). For the case $t = 0$, taking the limit $t \to 0$ in (12) we obtain (13). We have the same for $s = 0$ taking the limit.

Similarly taking (2) for $\psi(\varrho) = \varrho^{\frac{s}{t}}, \varrho > 0, s, t \neq 0, g \to g^t$ and then taking the power $\frac{1}{s}$ we get (14). For $s = 0$ or $t = 0$ we take the limit as above. □

Let p be positive integrable function defined on $[a,b]$ and g be any integrable function defined on $[a,b]$. Then, for a strictly monotone continuous function h whose domain belongs to the image of g, the quasi-arithmetic mean is defined as follows:

$$M_h(p;g) = h^{-1}\left(\frac{1}{\int_a^b p(\varrho)\, d\varrho} \int_a^b p(\varrho) h\big(g(\varrho)\big)\, d\varrho\right). \tag{16}$$

We give inequalities for the quasi-arithmetic mean.

Corollary 4 *Let u, v, p be positive integrable functions defined on $[a,b]$ such that $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$ and g be any integrable function defined on $[a,b]$. Also assume that h is a strictly monotone continuous function whose domain belongs to the image of g. If $f \circ h^{-1}$ is convex function then*

$$\frac{1}{\int_a^b p(\varrho)\, d\varrho} \int_a^b p(\varrho) f\big(g(\varrho)\big)\, d\varrho$$

$$\geq M_1(u;p)f\big(M_h(p.u;g)\big) + M_1(v;p)f\big(M_h(p.v;g)\big) \geq f\big(M_h(p;g)\big). \tag{17}$$

If the function $f \circ h^{-1}$ is concave then the reverse inequalities hold in (17).

Proof The required inequalities may be deduced by using (2) for $g \to h \circ g$ and $\psi \to f \circ h^{-1}$. □

The following refinement of the Hermite–Hadamard inequality may be given.

Corollary 5 *Let $\psi : [a,b] \to \mathbb{R}$ be a convex function defined on the interval $[a,b]$. Let $u, v : [a,b] \to \mathbb{R}$ be integrable functions such that $u(\varrho), v(\varrho) \in \mathbb{R}^+$ for all $\varrho \in [a,b]$ and $u(\varrho) + v(\varrho) = 1$. Then*

$$\begin{aligned} \frac{1}{b-a}\int_a^b \psi(\varrho)\,d\varrho &\geq \frac{1}{b-a}\int_a^b u(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b \varrho u(\varrho)\,d\varrho}{\int_a^b u(\varrho)\,d\varrho}\right) \\ &\quad + \frac{1}{b-a}\int_a^b v(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b \varrho v(\varrho)\,d\varrho}{\int_a^b v(\varrho)\,d\varrho}\right) \geq \psi\left(\frac{a+b}{2}\right). \end{aligned} \tag{18}$$

For the concave function ψ the reverse inequalities hold in (18).

Proof Using Theorem 1 for $p(\varrho) = 1, g(\varrho) = \varrho$ for all $\varrho \in [a,b]$, we obtain (18). □

3 Applications in information theory

In this section, we present some important applications for different divergences and distances in information theory [22] of our main result.

Definition 1 (Csiszár divergence) Let $T : I \to \mathbb{R}$ be a function defined on the positive interval I. Also let $u_1, v_1 : [a,b] \to (0,\infty)$ be two integrable functions such that $\frac{u_1(\varrho)}{v_1(\varrho)} \in I$ for all $\varrho \in [a,b]$, then the Csiszár divergence is defined as

$$\mathrm{C_d}(u_1, v_1) = \int_a^b v_1(\varrho) T\left(\frac{u_1(\varrho)}{v_1(\varrho)}\right) d\varrho.$$

Theorem 2 *Let $T : I \to \mathbb{R}$ be a convex function defined on the positive interval I. Let $u, v, u_1, v_1 : [a,b] \to \mathbb{R}^+$ be integrable functions such that $\frac{u_1(\varrho)}{v_1(\varrho)} \in I$ and $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$. Then*

$$\begin{aligned} \mathrm{C_d} &\geq \int_a^b u(\varrho)v_1(\varrho)\,d\varrho\,T\left(\frac{\int_a^b u(\varrho)u_1(\varrho)\,d\varrho}{\int_a^b u(\varrho)v_1(\varrho)\,d\varrho}\right) \\ &\quad + \int_a^b v(\varrho)v_1(\varrho)\,d\varrho\,T\left(\frac{\int_a^b v(\varrho)u_1(\varrho)\,d\varrho}{\int_a^b v(\varrho)v_1(\varrho)\,d\varrho}\right) \geq T\left(\frac{\int_a^b u_1(\varrho)\,d\varrho}{\int_a^b v_1(\varrho)\,d\varrho}\right)\int_a^b v_1(\varrho)\,d\varrho. \end{aligned} \tag{19}$$

Proof Using Theorem 1 for $\psi = T, g = \frac{u_1}{v_1}$ and $p = v_1$, we obtain (19). □

Definition 2 (Shannon entropy) If $v_1(\varrho)$ is positive probability density function defined on $[a,b]$, then the Shannon entropy is defined by

$$\mathrm{SE}(v_1) = -\int_a^b v_1(\varrho)\log v_1(\varrho)\,d\varrho.$$

Corollary 6 *Let $u, v, v_1 : [a,b] \to \mathbb{R}^+$ be integrable functions such that v_1 is probability density function and $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$. Then*

$$\begin{aligned}
&\int_a^b v_1(\varrho)\log\bigl(u_1(\varrho)\bigr)\,d\varrho + \mathrm{SE}(v_1) \\
&\quad\le \int_a^b u(\varrho)v_1(\varrho)\,d\varrho \log\left(\frac{\int_a^b u(\varrho)u_1(\varrho)\,d\varrho}{\int_a^b u(\varrho)v_1(\varrho)\,d\varrho}\right) \\
&\qquad + \int_a^b v(\varrho)v_1(\varrho)\,d\varrho \log\left(\frac{\int_a^b v(\varrho)u_1(\varrho)\,d\varrho}{\int_a^b v(\varrho)v_1(\varrho)\,d\varrho}\right) \le \log\left(\int_a^b u_1(\varrho)\,d\varrho\right). \qquad (20)
\end{aligned}$$

Proof Taking $T(\varrho) = -\log\varrho, \varrho \in \mathbb{R}^+$, in (19), we obtain (20). □

Definition 3 (Kullback–Leibler divergence) If u_1 and v_1 are two positive probability densities defined on $[a,b]$, the Kullback–Leibler divergence is defined by

$$\mathtt{KL}_{\mathtt{d}}(u_1, v_1) = \int_a^b u_1(\varrho)\log\left(\frac{u_1(\varrho)}{v_1(\varrho)}\right) d\varrho.$$

Corollary 7 *Let $u, v, u_1, v_1 : [a,b] \to \mathbb{R}^+$ be integrable functions such that u_1 and v_1 are probability density functions and $u(\varrho) + v(\varrho) = 1$ for all $\varrho \in [a,b]$. Then*

$$\begin{aligned}
\mathtt{KL}_{\mathtt{d}}(u_1, v_1) &\ge \int_a^b u(\varrho)u_1(\varrho)\,d\varrho \log\left(\frac{\int_a^b u(\varrho)u_1(\varrho)\,d\varrho}{\int_a^b u(\varrho)v_1(\varrho)\,d\varrho}\right) \\
&\quad + \int_a^b v(\varrho)u_1(\varrho)\,d\varrho \log\left(\frac{\int_a^b v(\varrho)u_1(\varrho)\,d\varrho}{\int_a^b v(\varrho)v_1(\varrho)\,d\varrho}\right) \ge 0. \qquad (21)
\end{aligned}$$

Proof Taking $T(\varrho) = \varrho\log\varrho, \varrho \in \mathbb{R}^+$, in (19), we obtain (20). □

Definition 4 (Variational distance) If u_1 and v_1 are positive probability density functions defined on $[a,b]$, then the variational distance is defined by

$$\mathtt{V}_{\mathtt{d}}(u_1, v_1) = \int_a^b \bigl|u_1(\varrho) - v_1(\varrho)\bigr|\,d\varrho.$$

Corollary 8 *Let u, v, u_1, v_1 be as stated in Corollary 7. Then*

$$\begin{aligned}
\mathtt{V}_{\mathtt{d}}(u_1, v_1) &\ge \left|\int_a^b u(\varrho)\bigl(u_1(\varrho) - v_1(\varrho)\bigr)\,d\varrho\right| \\
&\quad + \left|\int_a^b v(\varrho)\bigl(u_1(\varrho) - v_1(\varrho)\bigr)\,d\varrho\right|. \qquad (22)
\end{aligned}$$

Proof Using the function $T(\varrho) = |\varrho - 1|, \varrho \in \mathbb{R}^+$, in (19), we obtain (22). □

Definition 5 (Jeffrey's distance) If u_1 and v_1 are two positive probability density functions defined on $[a,b]$, then the Jeffrey distance is defined by

$$\mathtt{J}_{\mathtt{d}}(u_1, v_1) = \int_a^b \bigl(u_1(\varrho) - v_1(\varrho)\bigr)\log\left(\frac{u_1(\varrho)}{v_1(\varrho)}\right) d\varrho.$$

Corollary 9 *Let u, v, u_1, v_1 be as stated in Corollary 7. Then*

$$\mathtt{J_d}(u_1, v_1) \geq \int_a^b u(\varrho)\big(u_1(\varrho) - v_1(\varrho)\big)\, d\varrho \log\left(\frac{\int_a^b u(\varrho)u_1(\varrho)\, d\varrho}{\int_a^b u(\varrho)v_1(\varrho)\, d\varrho}\right) + \int_a^b v(\varrho)\big(u_1(\varrho) - v_1(\varrho)\big)\, d\varrho \log\left(\frac{\int_a^b v(\varrho)u_1(\varrho)\, d\varrho}{\int_a^b v(\varrho)v_1(\varrho)\, d\varrho}\right) \geq 0. \tag{23}$$

Proof Using the function $T(\varrho) = (\varrho - 1)\log \varrho, \varrho \in \mathbb{R}^+$, in (19), we obtain (23). □

Definition 6 (Bhattacharyya coefficient) If u_1 and v_1 are two positive probability density functions defined on $[a, b]$, then the Bhattacharyya coefficient is defined by

$$\mathtt{B_d}(u_1, v_1) = \int_a^b \sqrt{u_1(\varrho)v_1(\varrho)}\, d\varrho.$$

Corollary 10 *Let u, v, u_1, v_1 be as stated in Corollary 7. Then*

$$\mathtt{B_d}(u_1, v_1) \leq \sqrt{\int_a^b u(\varrho)v_1(\varrho)\, d\varrho \int_a^b u(\varrho)u_1(\varrho)\, d\varrho} + \sqrt{\int_a^b v(\varrho)v_1(\varrho)\, d\varrho \int_a^b v(\varrho)u_1(\varrho)\, d\varrho}. \tag{24}$$

Proof Using the function $T(\varrho) = -\sqrt{\varrho}, \varrho \in \mathbb{R}^+$, in (19), we obtain (24). □

Definition 7 (Hellinger distance) If u_1 and v_1 are two positive probability density functions defined on $[a, b]$, then the Hellinger distance is defined by

$$\mathtt{H_d}(u_1, v_1) = \int_a^b \left(\sqrt{u_1(\varrho)} - \sqrt{v_1(\varrho)}\right)^2 d\varrho.$$

Corollary 11 *Let u, v, u_1, v_1 be as stated in Corollary 7. Then*

$$\mathtt{H_d}(u_1, v_1) \geq \left(\sqrt{\int_a^b u(\varrho)u_1(\varrho)\, d\varrho} - \sqrt{\int_a^b u(\varrho)v_1(\varrho)\, d\varrho}\right)^2 + \left(\sqrt{\int_a^b v(\varrho)u_1(\varrho)\, d\varrho} - \sqrt{\int_a^b v(\varrho)v_1(\varrho)\, d\varrho}\right)^2 \geq 0. \tag{25}$$

Proof Using the function $T(\varrho) = (\sqrt{\varrho} - 1)^2, \varrho \in \mathbb{R}^+$, in (19), we obtain (25). □

Definition 8 (Triangular discrimination) If u_1 and v_1 are two positive probability density functions defined on $[a, b]$, then the triangular discrimination between u_1 and v_1 is defined by

$$\mathtt{T_d}(u_1, v_1) = \int_a^b \frac{(u_1(\varrho) - v_1(\varrho))^2}{u_1(\varrho) + v_1(\varrho)}\, d\varrho.$$

Corollary 12 *Let u, v, u_1, v_1 be as stated in Corollary 7. Then*

$$\mathrm{T_d}(u_1, v_1) \geq \frac{(\int_a^b u(\varrho)(u_1(\varrho) - v_1(\varrho))\,d\varrho)^2}{\int_a^b u(\varrho)(u_1(\varrho) + v_1(\varrho))\,d\varrho} + \frac{(\int_a^b (u_1(\varrho) - v_1(\varrho))v(\varrho)\,d\varrho)^2}{\int_a^b (u_1(\varrho) + v_1(\varrho))v(\varrho)\,d\varrho} \geq 0. \tag{26}$$

Proof Since the function $\phi(\varrho) = \frac{(\varrho-1)^2}{\varrho+1}, \varrho \in \mathbb{R}^+$, is convex, using the function $T(\varrho) = \phi(\varrho)$, in (19), we obtain (26). □

4 Further generalization

In the following theorem we present further refinement of the Jensen inequality concerning n functions whose sum is equal to unity.

Theorem 3 *Let $\psi : \mathrm{G} \to \mathbb{R}$ be a convex function defined on the interval G. Let $p, g, u_l \in L[a, b]$ such that $g(\varrho) \in \mathrm{G}, p(\varrho), u_l(\varrho) \in \mathbb{R}^+$ for all $\varrho \in [a, b]$ $(l = 1, 2, \dots, n)$ and $\sum_{l=1}^n u_l(\varrho) = 1$, $P = \int_a^b p(\varrho)\,d\varrho$. Assume that L_1 and L_2 are non-empty disjoint subsets of $\{1, 2, \dots, n\}$ such that $L_1 \cup L_2 = \{1, 2, \dots, n\}$. Then*

$$\begin{aligned}
&\frac{1}{P}\int_a^b p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho \\
&\geq \frac{1}{P}\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)g(\varrho)\,d\varrho}{\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)\,d\varrho}\right) \\
&\quad + \frac{1}{P}\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)g(\varrho)\,d\varrho}{\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)\,d\varrho}\right) \\
&\geq \psi\left(\frac{1}{P}\int_a^b p(\varrho)g(\varrho)\,d\varrho\right).
\end{aligned} \tag{27}$$

If the function ψ is concave then the reverse inequalities hold in (27).

Proof Since $\sum_{l=1}^n u_l(\varrho) = 1$, we may write

$$\begin{aligned}
&\int_a^b p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho \\
&= \int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho + \int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho.
\end{aligned} \tag{28}$$

Applying integral Jensen's inequality on both terms on the right hand side of (28) we obtain

$$\begin{aligned}
&\frac{1}{P}\int_a^b p(\varrho)\psi\bigl(g(\varrho)\bigr)\,d\varrho \\
&\geq \frac{1}{P}\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)g(\varrho)\,d\varrho}{\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)\,d\varrho}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{P}\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)\,d\varrho\,\psi\left(\frac{\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)g(\varrho)\,d\varrho}{\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)\,d\varrho}\right)\\
&\geq \psi\left(\frac{1}{P}\int_a^b \sum_{l\in L_1} u_l(\varrho)p(\varrho)g(\varrho)\,d\varrho + \frac{1}{P}\int_a^b \sum_{l\in L_2} u_l(\varrho)p(\varrho)g(\varrho)\,d\varrho\right)\\
&\quad \text{(by the convexity of } \psi)\\
&= \psi\left(\frac{1}{P}\int_a^b p(\varrho)g(\varrho)\,d\varrho\right). \qquad (29)
\end{aligned}
$$

□

Remark 2 If we take $n = 2$, in Theorem 3, we deduce Theorem 1. Also, analogously to the previous sections we may give applications of Theorem 3 for different means, the Hölder inequality and information theory.

Acknowledgements
The publication was supported by the Ministry of Education and Science of the Russian Federation (Agreement number No. 02.a03.21.0008.)

Funding
There is no funding for this work.

Availability of data and materials
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details
[1]Department of Mathematics, University of Peshawar, Peshawar, Pakistan. [2]Catholic University of Croatia, Zagreb, Croatia. [3]RUDN University, Moscow, Russia.

Publisher's Note

Received: 9 January 2020 Accepted: 11 March 2020 Published online: 19 March 2020

References

1. Adil Khan, M., Al-Sahwi, Z.M., Ming Chu, Y.: New estimations for Shannon and Zipf–Mandelbrot entropies. Entropy **20**(608), 1–10 (2018)
2. Adil Khan, M., Pečarić, Đ., Pečarić, J.: On Zipf–Mandelbrot entropy. J. Comput. Appl. Math. **346**, 192–204 (2019)
3. Adil Khan, M., Pečarić, Đ., Pečarić, J.: Bounds for Shannon and Zipf–Mandelbrot entropies. Math. Methods Appl. Sci. **40**(18), 7316–7322 (2017)
4. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc., Ser. B **39**, 1–38 (1977)
5. Adil Khan, M., Hanif, M., Khan, Z.A., Ahmad, K., Chu, Y.-M.: Association of Jensen inequality for s-convex function. J. Inequal. Appl. **2019**, Article ID 162 (2019)
6. Adil Khan, M., Pečarić, Đ., Pečarić, J.: Bounds for Csiszár divergence and hybrid Zipf–Mandelbrot entropy. Math. Methods Appl. Sci. **42**, 7411–7424 (2019)
7. Khan, S., Adil Khan, M., Chu, Y.-M.: Converses of the Jensen inequality derived from the Green functions with applications in information theory. Math. Methods Appl. Sci. **43**(5), 2577–2587 (2020). https://doi.org/10.1002/mma.6066
8. Pečarić, J., Perić, J.: New improvement of the converse Jensen inequality. Math. Inequal. Appl. **21**(1), 217–234 (2018)
9. Bakula, M.K., Nikodem, K.: Converse Jensen inequality for strongly convex set-valued maps. J. Math. Inequal. **12**(2), 545–550 (2018)
10. Choi, D., Krnić, M., Pečarić, J.: More accurate classes of Jensen-type inequalities for convex and operator convex functions. Math. Inequal. Appl. **21**(2), 301–321 (2018)
11. Sababheh, M.: Improved Jensen's inequality. Math. Inequal. Appl. **20**(2), 389–403 (2017)
12. Hot, J.M., Seo, Y.: An interpolation of Jensen's inequality and its converses with applications to quasi-arithmetic mean inequalities. J. Math. Inequal. **12**(2), 303–313 (2018)
13. Lu, G.: New refinements of Jensen's inequality and entropy upper bounds. J. Math. Inequal. **12**(2), 403–421 (2018)

14. Mikić, R., Pečarić, Ð., Pečarić, J.: Inequalities of the Jensen and Edmundson–Lah–Ribarić type for 3-convex functions with applications. J. Math. Inequal. **12**(3), 677–692 (2018)
15. Song, Y.-Q., Adil Khan, M., Zahir Ullah, S., Ming-Chu, Y.: Integral inequalities for strongly convex functions. J. Funct. Spaces **2018**, Article ID 6595921 (2018)
16. Khan, J., Adil Khan, M., Pečarić, J.: On Jensen's type inequalities via generalized majorization inequalities. Filomat **32**(16), 5719–5733 (2018)
17. Adil Khan, M., Ullah, S.Z., Chu, Y.: The concept of coordinate strongly convex functions and related inequalities. Rev. R. Acad. Cienc. Exactas Fís. Nat., Ser. A Mat. **113**, 2235–2251 (2019)
18. Adil Khan, M., Khan, J., Pečarić, J.: Generalization of Jensen's and Jensen–Steffensen's inequalities by generalized majorization theorem. J. Math. Inequal. **11**(4), 1049–1074 (2017)
19. Moradi, H.R., Omidvar, M.E., Adil Khan, M., Nikodem, K.: Around Jensen's inequality for strongly convex functions. Aequ. Math. **92**(1), 25–37 (2018)
20. Zaheer Ullah, S., Adil Khan, M., Chu, Y.-M.: A note on generalized convex functions. J. Inequal. Appl. **2009**, Article ID 291 (2009)
21. Işcan, I.: New refinements for integral and sum forms of Hölder inequality. J. Inequal. Appl. **2019**, Article ID 304 (2019)
22. Latif, N., Pečarić, Ð., Pečarić, J.: Majorization, useful Csiszar divergence and useful Zipf–Mandelbrot law. Open Math. **16**, 1357–1373 (2018)