Innovative Journal of Mathematics
v. 1 no. 2, 99-114, (2022)
doi:10.55059/ijm.2022.1.2/30

ISSN 2790-3370 (print)
ISSN 2790-3389 (online)

Estimation of Jensen's gap through an integral identity with applications to divergences

Research Article

Muhammad Adil Khan[1]*, Shah Faisal[1]† and Shahid Khan[1]‡

1 Department of Mathematics, University of Peshawar, Peshawar 25000, Pakistan

Abstract: Jensen's inequality and results related to its gap play a crucial role in the literature of applied mathematics. In this paper, we introduce a new, simple and sharp bound for Jensen's gap in discrete form by using an integral identity in terms of a certain function. We demonstrate this bound in integral form as well. Also, we illustrate some numerical examples which exhibit the tightness of the bound. The examples show that the bound is better than the existing bound given in a previously obtained result. Furthermore, we derive a new bound for the gap of Hermite-Hadamard inequality and derive two new variants of the Hölder inequality using the main results. At the end, we obtain some new inequalities for various divergences by using the main result.

Keywords: Jensen inequality • Hermite-Hadamard inequality • Hölder inequality • Csiszár f-divergence

Received 2022-02-15; Accepted 2022-04-12; Published 2022-04-14

1. Introduction

Mathematical inequalities provide a vast area of research in mathematics. There is a strong connection between convex functions and inequalities which attracted quite a large number of researchers who got several inequalities by convexity [19, 21, 24–26]. An enormous number of inequalities has been established for convex and concave functions. Jensen's inequality is believed to be the most vibrant and powerful inequality among these inequalities. It was established by J. Jensen in 1906, which states that:

If $f:[a,b]\to\mathbb{R}$ is a convex function and $x_i\in[a,b]$, $p_i\geq 0$ for $i=1,2,\cdots,n$, with $\sum_{i=1}^{n}p_i=P_n>0$ and $\frac{1}{P_n}\sum_{i=1}^{n}p_ix_i=\bar{x}$, then

$$f(\bar{x})\leq\frac{1}{P_n}\sum_{i=1}^{n}p_if(x_i). \tag{1}$$

‡ **Corresponding author**

* *E-mail:* *adilswati@gmail.com*
† *E-mail:* *sfaisaluop@gmail.com*
‡ *E-mail:* *shahidmathematics@gmail.com*

In (1), the reverse inequality holds whenever the function f is concave. Jensen's inequality is one of the most famous and widely used inequality in the field of mathematical inequalities. This inequality is considered as a gateway to other classical inequalities such as Minkwoski's, Young's, Hölder's, arithmatic-geometric and Ky-Fan's inequalities because all these inequalities can be deduced from this inequality by considering a suitable convex function. There exists an extensive literature regarding improvements, generalizations, refinements and converses etc of Jensen's inequality [1, 4–6, 9, 11].

There are numerous applications of Jensen's inequality such as: it provides better theoretical background to differential and integral equations [15], gives estimates for Zipf-Mendelbrot entropy and various divergences [2, 10, 12]. An improved Jensen's inequality is used in constructing a proper augmented Lyapunov-krasaovskii functional for static neural network with interval time-varying delays [23]. Also, Jensen's inequality is used along with a new Lyapunov functional for the stability analysis of discrete and continuous-time systems with time-varying delay [27, 28].

Some other results around Jensen's inequality and their applications can be studied in [3, 7, 13, 14, 16, 18].

This paper is organized in the following manner: In Section 2, we present a lemma following by the main results, two remarks, numerical experiments, proposition and two corollaries. Numerical experiments show that the new bound gives quite good results in comparison to an existing bound in the literature. The Proposition 1 and Corollary 1 exhibit variants for Hölder inequality. Also, a new bound is presented in Corollary 2 for Hermite-Hadamard gap. Section 3 of this paper offers applications for various divergences in information theory using the main result. Moreover, we have mentioned concluding remarks of this paper in the last section.

2. Main Results

We give the following lemma in order to derive our main results.

Lemma 2.1.
Let $f \in C[a,b]$ be a function and $t, s \in [a,b]$, then

$$f(t) = \frac{1}{b-a}\int_a^b f(s)ds + \frac{1}{b-a}\int_a^b p(t,s)f'(s)ds, \tag{2}$$

where

$$p(t,s) = \begin{cases} s-a, & a \le s \le t, \\ s-b, & t \le s \le b. \end{cases} \tag{3}$$

Proof. Using integration by parts, we have

$$\int_a^t (s-a)f'(s)ds = (t-a)f(t) - \int_a^t f(s)ds \tag{4}$$

and

$$\int_t^b (s-b)f'(s)ds = (b-t)f(t) - \int_t^b f(s)ds. \tag{5}$$

Adding (4) and (5) we get

$$\int_a^b p(t,s)f'(s)ds = (b-a)f(t) - \int_a^b f(s)ds. \tag{6}$$

Solving (6) for f, we obtain (2). □

Theorem 2.1.
Assume a function $f \in C[a,b]$ and $c \in [a,b]$ such that $|f'(c)| = \max_{x\in[a,b]} |f'(x)|$. Also, let $x_i \in [a,b]$, $p_i > 0$ for each $i \in \{1,2,\cdots,n\}$ with $\sum_{i=1}^n p_i = 1$ and $\sum_{i=1}^n p_i x_i = \bar{x}$, then

$$\left|\sum_{i=1}^n p_i f(x_i) - f(\bar{x})\right| \leq \frac{|f'(c)|}{b-a}\left[\sum_{i=1}^n p_i x_i^2 + (\bar{x}-(a+b))^2 - 2ab\right]. \tag{7}$$

Proof. Using (2) in $\sum_{i=1}^n p_i f(x_i)$ and $f\left(\sum_{i=1}^n p_i x_i\right)$, we get

$$\sum_{i=1}^n p_i f(x_i) = \frac{1}{b-a}\int_a^b f(s)ds + \frac{1}{b-a}\int_a^b \sum_{i=1}^n p_i p(x_i,s)f'(s)ds \tag{8}$$

and

$$f(\bar{x}) = \frac{1}{b-a}\int_a^b f(s)ds + \frac{1}{b-a}\int_a^b p(\bar{x},s)f'(s)ds. \tag{9}$$

Subtracting (9) from (8), we get

$$\sum_{i=1}^n p_i f(x_i) - f(\bar{x}) = \frac{1}{b-a}\int_a^b \left(\sum_{i=1}^n p_i p(x_i,s) - p(\bar{x},s)\right) f'(s)ds. \tag{10}$$

Taking absolute value of (10), we have

$$\begin{aligned}\left|\sum_{i=1}^n p_i f(x_i) - f(\bar{x})\right| &= \left|\frac{1}{b-a}\int_a^b \left(\sum_{i=1}^n p_i p(x_i,s) - p(\bar{x},s)\right) f'(s)ds\right| \\ &\leq \frac{1}{b-a}\int_a^b \left|\left(\sum_{i=1}^n p_i p(x_i,s) - p(\bar{x},s)\right)\right| |f'(s)|ds \\ &\leq \frac{1}{b-a}\int_a^b \left[\left|\sum_{i=1}^n p_i p(x_i,s)\right| + |p(\bar{x},s)|\right] |f'(s)|ds. \end{aligned} \tag{11}$$

Since $|f'(c)| = \max_{x\in[a,b]} |f'(x)|$ for $c \in [a,b]$ therefore (11) becomes

$$\left|\sum_{i=1}^n p_i f(x_i) - f(\bar{x})\right| \leq \frac{|f'(c)|}{b-a}\int_a^b \left[\left|\sum_{i=1}^n p_i p(x_i,s)\right| + |p(\bar{x},s)|\right] ds. \tag{12}$$

From (3), we have the following functions and utilizing in forthright

$$p(x_i, s) = \begin{cases} s - a, & a \leq s \leq x_i, \\ s - b, & x_i \leq s \leq b. \end{cases}$$

$$p(\bar{x}, s) = \begin{cases} s - a, & a \leq s \leq \bar{x}, \\ s - b, & \bar{x} \leq s \leq b. \end{cases}$$

$$\begin{aligned} \int_a^b \left| \sum_{i=1}^n p_i p(x_i, s) \right| ds &= \int_a^b \left(|p_1 p(x_1, s) + p_2 p(x_2, s) + \cdots + p_n p(x_n, s)| \right) ds \\ &\leq \int_a^b \left(|p_1||p(x_1, s)| + |p_2||p(x_2, s)| + \cdots + |p_n||p(x_n, s)| \right) ds \\ &= \sum_{i=1}^n p_i \int_a^b |p(x_i, s)| ds \\ &= \sum_{i=1}^n p_i \left(\int_a^{x_i} |s - a| ds + \int_{x_i}^b |s - b| ds \right) \\ &= \sum_{i=1}^n p_i \left(\int_a^{x_i} (s - a) ds - \int_{x_i}^b (s - b) ds \right) \\ &= \frac{a^2}{2} + \frac{b^2}{2} + \sum_{i=1}^n p_i x_i^2 - a\bar{x} - b\bar{x}. \end{aligned} \tag{13}$$

Also

$$\begin{aligned} \int_a^b |p(\bar{x}, s)| \, ds &= \int_a^{\bar{x}} (s - a) ds - \int_{\bar{x}}^b (s - b) ds \\ &= \frac{a^2}{2} + \frac{b^2}{2} + (\bar{x})^2 - a\bar{x} - b\bar{x}. \end{aligned} \tag{14}$$

From (12), (13) and (14), we get result (7). □

Theorem 2 is the integral version of Theorem 1 and can be proved by adopting the similar procedure.

Theorem 2.2.
Assume a function $f \in C[a, b]$ and $c \in [a, b]$ such that $|f'(c)| = \max_{x \in [a,b]} |f'(x)|$. Also, let $h_1, h_2 : [a_1, a_2] \longrightarrow \mathbb{R}$ be some integrable functions such that $h_1(x) \in [a, b]$ and $h_2(x) \geq 0$, $\forall x \in [a_1, a_2]$ with $\int_{a_1}^{a_2} h_2(x)dx = 1$ and $\bar{h} = \int_{a_1}^{a_2} h_1(x)h_2(x)dx$, then

$$\begin{aligned} &\left| \int_{a_1}^{a_2} (f \circ h_1)(x) h_2(x) dx - f\left(\bar{h}\right) \right| \\ &\qquad \leq \frac{|f'(c)|}{b - a} \left[\int_{a_1}^{a_2} h_2(x) h_1^2(x) dx + \left(\bar{h} - (a + b)\right)^2 - 2ab \right]. \end{aligned} \tag{15}$$

Example 2.1.
Let $f(x) = \frac{1}{30}x^{30}$, $h_1(x) = x$, $h_2(x) = 1$, $\forall x \in [0,1]$. Here, $[a,b] = [a_1, a_2] = [0,1]$ with $\int_0^1 h_2(x)dx = 1$ and maximum of $f'(x)$ occurs at $x = 1$ which is $|f'(1)| = 1$. Therefore, using all these values in inequality (15), we get $\left|\int_o^1 (f \circ h_1)(x)h_2(x)dx - f\left(\bar{h}\right)\right|$
$= 1.07 \times 10^{-3} - 3.10 \times 10^{-11} = 0.00106$ and its corresponding right hand side gives 0.5834. Now from inequality (15) it can be observed that

$$0.00106 < 0.5834. \tag{16}$$

We obtain the following expression from the right side of inequality (5) in [8]

$$\begin{aligned}\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times \left[||h_1 - c||^2_{L^2} + ||h_1 - c||^2_{L^1}\right] \\ = 14.5 \times \left[\frac{3c^2 - 3c + 1}{3} + \left(c^2 - c + \frac{1}{2}\right)^2\right] = 14.5 \times g(c).\end{aligned} \tag{17}$$

Since minimum of $g(c)$ occurs at $c = 0.5$ which is $g(0.5) \approx 0.1458$. Therefore (17), gives

$$\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times \left[||h_1 - c||^2_{L^2} + ||h_1 - c||^2_{L^1}\right] \approx 2.1141. \tag{18}$$

Hence from inequality (5) in [8], we get

$$0.00106 < 2.1141. \tag{19}$$

Likewise we get the following expression from the right side of inequality (8) in [8]

$$\begin{aligned}\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times ||h_1 - c||^2_{L^2} - \frac{1}{2}\inf_{[0,1]} f'' \times \left[\int_0^1 (h_1(x) - c)dx\right]^2 \\ = 14.5 \times \left[\frac{3c^2 - 3c + 1}{3}\right] = 14.5 \times l(c).\end{aligned} \tag{20}$$

Now minimum of $l(c)$ occurs at $c = 0.5$ which is $l(0.5) \approx 0.0833$ and so from (20) we conclude

$$\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times ||h_1 - c||^2_{L^2} - \frac{1}{2}\inf_{[0,1]} f'' \times \left[\int_0^1 (h_1(x) - c)dx\right]^2 \approx 1.2078. \tag{21}$$

Hence from inequality (8) in [8], we get

$$0.00106 < 1.2078. \tag{22}$$

Inequality (16), (19) and (22), show that the bound in inequality (15) improves over the bounds in inequalities (5), (8) given in [8]. Furthermore, inequality (16) illustrates that the bound in (15) is tight with regard to Jensen's gap.

Example 2.2.
Let $f(x) = \frac{1}{30}x^{30}$, $h_1(x) = x^2$, $h_2(x) = 1$, $\forall x \in [0,1]$. Here, $[a,b] = [a_1, a_2] = [0,1]$ with $\int_0^1 h_2(x)dx = 1$ and maximum of $f'(x)$ occurs at $x = 1$ which is $|f'(1)| = 1$. Therefore, using all these values in inequality (15), we get $\left|\int_o^1 (f \circ h_1)(x)h_2(x)dx - f\left(\bar{h}\right)\right|$

$= 5.46\times10^{-4} - 1.61\times10^{-16} = 0.000546$ and its corresponding right hand side gives 0.4222. Now from inequality (15) it can be observed that

$$0.000546 < 0.4222. \tag{23}$$

We deduce the following mathematical statement from the right side of inequality (5) in [8]

$$\begin{aligned}\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times \left[||h_1-c||^2_{L^2} + ||h_1-c||^2_{L^1}\right] \\ = 14.5 \times \left[c^2 - \frac{2}{3}c + \frac{1}{5} + \left(\frac{4}{3}c^{\frac{3}{2}} - c + \frac{1}{3}\right)^2\right] = 14.5 \times g(c).\end{aligned} \tag{24}$$

Since minimum of $g(c)$ occurs at $c = 0.31$ which is $g(0.31) \approx 0.1536$. Therefore (24), gives

$$\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times \left[||h_1-c||^2_{L^2} + ||h_1-c||^2_{L^1}\right] \approx 2.2272. \tag{25}$$

Hence from inequality (5) in [8], we get

$$0.000546 < 2.2272. \tag{26}$$

Likewise we obtain the following expression from the right side of inequality (8) in [8]

$$\begin{aligned}\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times ||h_1-c||^2_{L^2} - \frac{1}{2}\inf_{[0,1]} f'' \times \left[\int_0^1 (h_1(x)-c)dx\right]^2 \\ = 14.5 \times \left[c^2 - \frac{2}{3}c + \frac{1}{5}\right] = 14.5 \times l(c).\end{aligned} \tag{27}$$

Now minimum of $l(c)$ occurs at $c = 0.3$ which is $l(0.3) \approx 0.09$ and so from (27) we conclude

$$\frac{1}{2}\|f''\|_{L^\infty([0,1])} \times ||h_1-c||^2_{L^2} - \frac{1}{2}\inf_{[0,1]} f'' \times \left[\int_0^1 (h_1(x)-c)dx\right]^2 \approx 1.305 \tag{28}$$

Hence from inequality (8) in [8], we get

$$0.000546 < 1.305. \tag{29}$$

Inequality (23), (26) and (29), show that the bound in inequality (15) improves over the bounds in (5), (8) given in [8]. Furthermore, inequality (23) illustrates that the bound in (15) is tight with regard to Jensen's gap.

Example 2.3.
Let $f(x) = \frac{1}{30}x^{30}$, $h_1(x) = x^3$, $h_2(x) = 1$, $\forall x \in [0,1]$. Here, $[a,b] = [a_1,a_2] = [0,1]$ with $\int_0^1 h_2(x) = 1$ and maximum of $f'(x)$ occurs at $x = 1$ which is $|f'(1)| = 1$. Therefore, using all these values in inequality (15), we get $\left|\int_o^1 (f\circ h_1)(x)h_2(x)dx - f\left(\bar{h}\right)\right|$
$= 3.66\times10^{-4} - 2.89\times10^{-20} = 0.000366$ and its corresponding right hand side gives 0.70535. Now from inequality (15) it can be observed that

$$0.000366 < 0.70535. \tag{30}$$

We get the following expression from the right side of inequality (5) in [8]

$$\begin{aligned}&\frac{1}{2}\|f''\|_{L^{\infty}([0,1])}\times\Big[||h_1-c||_{L^2}^2+||h_1-c||_{L^1}^2\Big]\\&\qquad=14.5\times\left[c^2-\frac{c}{2}+\frac{1}{7}+\left(\frac{3}{2}c^{\frac{4}{3}}-c+\frac{1}{4}\right)^2\right]=14.5\times g(c).\end{aligned}\tag{31}$$

Since minimum of $g(c)$ occurs at $c=0.2$ which is $g(0.2)\approx 0.13368$. Therefore (31), gives

$$\frac{1}{2}\|f''\|_{L^{\infty}([0,1])}\times\Big[||h_1-c||_{L^2}^2+||h_1-c||_{L^1}^2\Big]\approx 1.93836.\tag{32}$$

Hence from inequality (5) in [8], we get

$$0.000366<1.93836.\tag{33}$$

Likewise the right side of inequality (8) in [8], gives

$$\begin{aligned}&\frac{1}{2}\|f''\|_{L^{\infty}([0,1])}\times||h_1-c||_{L^2}^2-\frac{1}{2}\inf_{[0,1]}f''\times\left[\int_0^1(h_1(x)-c)dx\right]^2\\&\qquad=14.5\times\left[c^2-\frac{c}{2}+\frac{1}{7}\right]=14.5\times l(c).\end{aligned}\tag{34}$$

Now minimum of $l(c)$ occurs at $c=0.2$ and thus $l(0.2)\approx 0.082857$, so from (34), we conclude

$$\frac{1}{2}\|f''\|_{L^{\infty}([0,1])}\times||h_1-c||_{L^2}^2-\frac{1}{2}\inf_{[0,1]}f''\times\left[\int_0^1(h_1(x)-c)dx\right]^2\approx 1.20132.\tag{35}$$

Hence from inequality (8) in [8], we get

$$0.000366<1.20132.\tag{36}$$

Inequality (30), (33) and (36), show that the bound in inequality (15) improves over the bounds in (5), (8) given in [8]. Furthermore, inequality (30) illustrates that the bound in (15) is tight with regard to Jensen's gap.

A new variant of the Hölder inequality is obtained using Theorem 1 which is given as under:

Proposition 2.1.

Assume a positive interval $[a,b]$ with $p,q>1$ such that $\frac{1}{p}+\frac{1}{q}=1$. Also, let $(\rho_1,\cdots,\rho_n)$, $(\varrho_1,\cdots,\varrho_n)$ be two positive n-tuples such that $\frac{\sum_{i=1}^n\rho_i\varrho_i}{\sum_{i=1}^n\varrho_i^q},\rho_i\varrho_i^{-\frac{q}{p}}\in[a,b]$ for each $i\in\{1,2,\cdots,n\}$, then

$$\begin{aligned}&\left(\sum_{i=1}^n\rho_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^n\varrho_i^q\right)^{\frac{1}{q}}-\sum_{i=1}^n\rho_i\varrho_i\\&\quad\le\left(\frac{pb^{p-1}}{b-a}\right)^{\frac{1}{p}}\left[\frac{1}{\sum_{i=1}^n\varrho_i^q}\sum_{i=1}^n\rho_i^2\varrho_i^{1-q/p}+\left(\frac{1}{\sum_{i=1}^n\varrho_i^q}\sum_{i=1}^n\rho_i\varrho_i\right)^2\right.\\&\qquad\qquad\left.+a^2+b^2-2(a+b)\left(\frac{1}{\sum_{i=1}^n\varrho_i^q}\sum_{i=1}^n\rho_i\varrho_i\right)\right]^{\frac{1}{p}}\sum_{i=1}^n\varrho_i^q.\end{aligned}\tag{37}$$

Proof. Assume a function $f(x) = x^p$ for $x \in [a, b]$, $x_i = \rho_i \varrho_i^{-q/p}$ and $p_i = \frac{\varrho_i^q}{\sum_{i=1}^n \varrho_i^q}$. Here, $|f'(x)| = px^{p-1}$ which attains its maximum value at $c = b$ i.e $|f'(c)| = pb^{p-1}$. Now using all these values in (7), we get

$$\left(\left(\sum_{i=1}^n \rho_i^p\right)\left(\sum_{i=1}^n \varrho_i^q\right)^{p-1} - \left(\sum_{i=1}^n \rho_i \varrho_i\right)^p\right)^{\frac{1}{p}}$$
$$\leq \left(\frac{pb^{p-1}}{b-a}\right)^{\frac{1}{p}} \left[\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^2 \varrho_i^{1-q/p} + \left(\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i \varrho_i\right)^2\right.$$
$$\left. + a^2 + b^2 - 2(a+b)\left(\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i \varrho_i\right)\right]^{\frac{1}{p}} \sum_{i=1}^n \varrho_i^q. \tag{38}$$

By utilizing the inequality $v^\xi - u^\xi \leq (v-u)^\xi$, $0 \leq u \leq v$, $\xi \in [0, 1]$ for
$v = \left(\sum_{i=1}^n \rho_i^p\right)\left(\sum_{i=1}^n \varrho_i^q\right)^{p-1}$, $u = \left(\sum_{i=1}^n \rho_i \varrho_i\right)^p$ and $\xi = \frac{1}{p}$, we obtain

$$\left(\sum_{i=1}^n \rho_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n \varrho_i^q\right)^{\frac{1}{q}} - \sum_{i=1}^n \rho_i \varrho_i \leq \left(\left(\sum_{i=1}^n \rho_i^p\right)\left(\sum_{i=1}^n \varrho_i^q\right)^{p-1} - \left(\sum_{i=1}^n \rho_i \varrho_i\right)^p\right)^{\frac{1}{p}}. \tag{39}$$

Now using (39) in (38), we get (37). □

Remark 2.1.
The integral version of Proposition 1 can be given as an application of Theorem 2.

Here we present some Hölder type inequalities as applications of Theorem 1.

Corollary 2.1.
Assume a positive interval $[a, b]$ and two positive n-tuples $(\rho_1, \cdots, \rho_n), (\varrho_1, \cdots, \varrho_n)$, then
(i) for $1 < p$, $q = \frac{p}{p-1}$ and $\frac{\sum_{i=1}^n \rho_i^p}{\sum_{i=1}^n \varrho_i^q}, \rho_i^p \varrho_i^{-q} \in [a, b]$ for each $i \in \{1, 2, \cdots, n\}$, the following inequality holds

$$\sum_{i=1}^n \rho_i \varrho_i - \left(\sum_{i=1}^n \rho_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n \varrho_i^q\right)^{\frac{1}{q}}$$
$$\leq \frac{a^{\frac{1}{p}-1}}{p(b-a)} \left[\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^{2p} \varrho_i^{-q} + \left(\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^p\right)^2\right.$$
$$\left. + a^2 + b^2 - 2(a+b)\left(\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^p\right)\right] \sum_{i=1}^n \varrho_i^q. \tag{40}$$

(ii) for $0 < p < 1$, $q = \frac{p}{p-1}$ and $\frac{\sum_{i=1}^n \rho_i^p}{\sum_{i=1}^n \varrho_i^q}, \rho_i^p \varrho_i^{-q} \in [a, b]$ for each $i \in \{1, 2, \cdots, n\}$, the following inequality holds

$$\sum_{i=1}^n \rho_i \varrho_i - \left(\sum_{i=1}^n \rho_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n \varrho_i^q\right)^{\frac{1}{q}}$$
$$\leq \frac{b^{\frac{1}{p}-1}}{p(b-a)} \left[\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^{2p} \varrho_i^{-q} + \left(\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^p\right)^2\right.$$
$$\left. + a^2 + b^2 - 2(a+b)\left(\frac{1}{\sum_{i=1}^n \varrho_i^q} \sum_{i=1}^n \rho_i^p\right)\right] \sum_{i=1}^n \varrho_i^q. \tag{41}$$

(iii) for $p<0$, $q=\frac{p}{p-1}$ and $\frac{\sum_{i=1}^{n}\rho_i^p}{\sum_{i=1}^{n}\varrho_i^q}, \rho_i^p\varrho_i^{-q}\in[a,b]$ for each $i\in\{1,2,\cdots,n\}$, the following inequality holds

$$\sum_{i=1}^{n}\rho_i\varrho_i-\left(\sum_{i=1}^{n}\rho_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}\varrho_i^q\right)^{\frac{1}{q}} \leq \frac{a^{\frac{1}{p}-1}}{p(a-b)}\left[\frac{1}{\sum_{i=1}^{n}\varrho_i^q}\sum_{i=1}^{n}\rho_i^{2p}\varrho_i^{-q}+\left(\frac{1}{\sum_{i=1}^{n}\varrho_i^q}\sum_{i=1}^{n}\rho_i^p\right)^2 +a^2+b^2-2(a+b)\left(\frac{1}{\sum_{i=1}^{n}\varrho_i^q}\sum_{i=1}^{n}\rho_i^p\right)\right]\sum_{i=1}^{n}\varrho_i^q. \tag{42}$$

Proof. (i): Assume that $f(x)=x^{\frac{1}{p}}$ for $x\in[a,b]$, $x_i=\rho_i^p\varrho_i^{-q}$ and $p_i=\frac{\varrho_i^q}{\sum_{i=1}^{n}\varrho_i^q}$. Also, $|f'(x)|=\frac{1}{p}x^{\frac{1}{p}-1}$ which attains its maximum value at $c=a$ i.e $|f'(c)|=\frac{1}{p}a^{\frac{1}{p}-1}$. Now using all these values in (7), we get (40).
(ii): Here, $|f'(x)|=\frac{1}{p}x^{\frac{1}{p}-1}$ attains its maximum value at $c=b$ i.e $|f'(c)|=\frac{1}{p}b^{\frac{1}{p}-1}$. Now using this value along with $x_i=\rho_i^p\varrho_i^{-q}$ and $p_i=\frac{\varrho_i^q}{\sum_{i=1}^{n}\varrho_i^q}$ in (7), we get (41).
(iii): Here, $f'(x)=\frac{1}{p}x^{\frac{1}{p}-1}\Rightarrow|f'(x)|=\frac{-1}{p}x^{\frac{1}{p}-1}$ which attains its maximum value at $c=a$ i.e $|f'(c)|=\frac{-1}{p}a^{\frac{1}{p}-1}$. Now using this value along with $x_i=\rho_i^p\varrho_i^{-q}$ and $p_i=\frac{\varrho_i^q}{\sum_{i=1}^{n}\varrho_i^q}$ in (7), we get (42). □

A bound for the Hermite-Hadamard gap is obtained using Theorem 2 which is given as under:

Corollary 2.2.
Let $\psi\in C[a_1,a_2]$ be a function, then for $c\in[a_1,a_2]$ the following inequality holds

$$\left|\frac{1}{a_2-a_1}\int_{a_1}^{a_2}\psi(x)dx-\psi\left(\frac{a_1+a_2}{2}\right)\right|\leq\frac{7}{12}|f'(c)|(a_2-a_1). \tag{43}$$

Proof. Using (15) for $\psi=f$, $[a,b]=[a_1,a_2]$ and $h_2(x)=\frac{1}{a_2-a_1}, h_1(x)=x, \forall\ x\in[a_1,a_2]$, we get (43). □

3. Applications in Information Theory

Information theory deals with the transmission, mathematical treatment, quantification and storage of information. Claude Shannon [22] proposed the idea of information theory and suggested that entropy is the key measure of information. The probability distribution can also be used as a technique to measure the information. Divergences are used as a tool to measure the separation between probability distributions. This section presents some inequalities around various divergences in information theory. A literature about the results around various divergences can be found in [10, 13, 17, 20].

Definition 3.1 (Csiszár divergence).
Assume that $\mathbf{r}=(r_1,\cdots,r_n)\in\mathbb{R}^n$, $\mathbf{w}=(w_1,\cdots,w_n)\in\mathbb{R}_+^n$ and a function $f:[a,b]\to\mathbb{R}$, then the Csiszár divergence is given by

$$\bar{D}_c(\mathbf{r},\mathbf{w})=\sum_{i=1}^{n}w_if\left(\frac{r_i}{w_i}\right),$$

where $\frac{r_i}{w_i}\in[a,b]$ for $i=1,2,\cdots,n$.

Theorem 3.1.
Assume a function $f \in C[a,b]$ and $c \in [a,b]$ such that $|f'(c)| = \max_{x\in[a,b]} |f'(x)|$. Also, assume that $\mathbf{r} = (r_1, \cdots, r_n) \in \mathbb{R}^n$ and $\mathbf{w} = (w_1, \cdots, w_n) \in \mathbb{R}^n_+$ such that $\frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i}, \frac{r_i}{w_i} \in [a,b]$ for each $i \in \{1,2,\cdots,n\}$, then

$$\left|\frac{1}{\sum_{i=1}^n w_i}\bar{D}_c(\mathbf{r},\mathbf{w}) - f\left(\frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i}\right)\right| \le \frac{|f'(c)|}{b-a}\left[\frac{1}{\sum_{i=1}^n w_i}\sum_{i=1}^n \frac{r_i^2}{w_i} + \left(\frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i} - (a+b)\right)^2 - 2ab\right]. \tag{44}$$

Proof. The result (44) can be easily deduced from (7) by choosing $x_i = \frac{r_i}{w_i}$, $p_i = \frac{w_i}{\sum_{i=1}^n w_i}$. □

Definition 3.2 (Rényi-divergence).
The following functional defined for two positive probability distributions $\mathbf{r} = (r_1, \cdots, r_n)$, $\mathbf{w} = (w_1, \cdots, w_n)$ and a nonnegative real number $\mu \neq 1$ is known as rényi-divergence

$$D_{re}(\mathbf{r},\mathbf{w}) = \frac{1}{\mu-1}\log\left(\sum_{i=1}^n r_i^{\mu} w_i^{1-\mu}\right).$$

Corollary 3.1.
Let $[a,b] \subset \mathbb{R}^+$ and $\mathbf{r} = (r_1, \cdots, r_n)$, $\mathbf{w} = (w_1, \cdots, w_n)$ be positive probability distributions and $\mu \neq 1$ then,
(i) if $0 < \mu < 1$ and $(\frac{r_i}{w_i})^{\mu-1} \in [a,b]$ for $i = 1,2,\cdots,n$, we have

$$D_{re}(\mathbf{r},\mathbf{w}) - \frac{1}{\mu-1}\sum_{i=1}^n r_i \log\left(\frac{r_i}{w_i}\right)^{\mu-1} \le \frac{1}{a(1-\mu)(b-a)}\left[\sum_{i=1}^n r_i\left(\frac{r_i}{w_i}\right)^{2(\mu-1)} + \left(\sum_{i=1}^n r_i^{\mu} w_i^{1-\mu} - (a+b)\right)^2 - 2ab\right]. \tag{45}$$

(ii) if $\mu > 1$ and $(\frac{r_i}{w_i})^{\mu-1} \in [a,b]$ for $i = 1,2,\cdots,n$, we have

$$D_{re}(\mathbf{r},\mathbf{w}) - \frac{1}{\mu-1}\sum_{i=1}^n r_i \log\left(\frac{r_i}{w_i}\right)^{\mu-1} \le \frac{1}{a(\mu-1)(b-a)}\left[\sum_{i=1}^n r_i\left(\frac{r_i}{w_i}\right)^{2(\mu-1)} + \left(\sum_{i=1}^n r_i^{\mu} w_i^{1-\mu} - (a+b)\right)^2 - 2ab\right]. \tag{46}$$

Proof. (i): Assume a function $f(x) = -\frac{1}{\mu-1}\log x$ for $x \in [a,b]$, then one has $|f'(x)| = \frac{1}{(1-\mu)x}$ and $\max|f'(x)|$ occurs at $c = a$. Now using (7) for $f(x) = -\frac{1}{\mu-1}\log x$, $p_i = r_i$ and $x_i = (\frac{r_i}{w_i})^{\mu-1}$ and $|f'(c)| = \frac{1}{(1-\mu)a}$ we derive (45).

(ii): Let $f(x) = -\frac{1}{\mu-1}\log x$, $x \in [a,b]$, then $|f'(x)| = \frac{1}{(\mu-1)x}$ and $\max|f'(x)|$ occurs at $c = a$. Now using (7) for $f(x) = -\frac{1}{\mu-1}\log x$, $p_i = r_i$ and $x_i = (\frac{r_i}{w_i})^{\mu-1}$ and $|f'(c)| = \frac{1}{(\mu-1)a}$ we derive (46). □

Definition 3.3 (Shannon-entropy).
The Shanon-entropy of a positive probability distribution $\mathbf{w} = (w_1, \cdots, w_n)$ is defined to be

$$E_s(\mathbf{w}) = -\sum_{i=1}^{n} w_i \log w_i.$$

Corollary 3.2.
Let $[a, b] \subset \mathbb{R}^+$ and $\mathbf{w} = (w_1, \cdots, w_n)$ be a positive probability distribution with the condition $\frac{r_i}{w_i} \in [a, b]$ for each $i \in \{1, 2, \cdots, n\}$, then

$$\log n - E_s(\mathbf{w}) \leq \frac{1}{a(b-a)} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + n^2 + a^2 + b^2 - 2n(a+b) \right]. \tag{47}$$

Proof. Suppose that $f(x) = -\log x$ for $x \in [a, b]$ then $|f'(x)| = \frac{1}{x}$ and $\max |f'(x)|$ occurs at $c = a$. Now using (44) for $f(x) = -\log x$, $|f'(c)| = \frac{1}{a}$ and $r_i = 1$ for each $i \in \{1, 2, \cdots, n\}$, we get (47). $\square$

Definition 3.4 (Kullback-Leibler divergence).
Let $\mathbf{r} = (r_1, \cdots, r_n)$ and $\mathbf{w} = (w_1, \cdots, w_n)$ be two positive probability distributions then the kullback-Leibler divergence from $\mathbf{r}$ to $\mathbf{w}$ is defined to be

$$D_{kl}(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^{n} r_i \log \frac{r_i}{w_i}.$$

Corollary 3.3.
Let $[a, b] = [a, \frac{1}{e}] \cup [\frac{1}{e}, b] \subset \mathbb{R}^+$ and $\mathbf{r} = (r_1, \cdots, r_n)$, $\mathbf{w} = (w_1, \cdots, w_n)$ be positive probability distributions then,
(i) if $\frac{r_i}{w_i} \in [\frac{1}{e}, b]$ for $i = 1, 2, \cdots, n$, we have

$$D_{kl}(\mathbf{r}, \mathbf{w}) \leq \frac{|1 + \log b|}{b-a} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + 1 + a^2 + b^2 - 2(a+b) \right]. \tag{48}$$

(ii) if $\frac{r_i}{w_i} \in [a, \frac{1}{e}]$ for $i = 1, 2, \cdots, n$, we have

$$D_{kl}(\mathbf{r}, \mathbf{w}) \leq \frac{|1 + \log a|}{b-a} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + 1 + a^2 + b^2 - 2(a+b) \right]. \tag{49}$$

Proof. (i): Suppose the function $f(x) = x \log x$ for $x \in [\frac{1}{e}, b]$ then $|f'(x)| = |1 + \log x|$. Also, $|f'(x)|$ attains its maximum value at $c = b$ i.e $|f'(c)| = |1 + \log b|$. Now using all these values in (44), we get (48) .

(ii): Assume the function $f(x) = x \log x$ for $x \in [a, \frac{1}{e}]$ then $|f'(x)| = |1 + \log x|$. Also, $|f'(x)|$ attains its maximum value at $c = a$ i.e $|f'(c)| = |1 + \log a|$. Now using all these values in (44), we get (49) . $\square$

For any interval $[a, b]$ with $a > \frac{1}{e}$, we have inequality (48) while for interval $[a, b]$ with $b < \frac{1}{e}$, we have inequality (49).

Definition 3.5 (χ^2-divergence).
The χ^2-divergence of the two positive probability distributions $\mathbf{r} = (r_1, \cdots, r_n)$ and $\mathbf{w} = (w_1, \cdots, w_n)$ is defined to be

$$D_{\chi^2}(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^{n} \frac{(r_i - w_i)^2}{w_i}.$$

Corollary 3.4.
If $[a, b] = [a, 1] \cup [1, b] \subset \mathbb{R}^+$ and $\mathbf{r} = (r_1, \cdots, r_n)$, $\mathbf{w} = (w_1, \cdots, w_n)$ are two positive probability distributions then,
(i) if $\frac{r_i}{w_i} \in [1, b]$ for $i = 1, 2, \cdots, n$, then we have

$$D_{\chi^2}(\mathbf{r}, \mathbf{w}) \leq \frac{|2b-2|}{b-a} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + 1 + a^2 + b^2 - 2(a+b) \right]. \tag{50}$$

(ii) if $\frac{r_i}{w_i} \in [a, 1]$ for $i = 1, 2, \cdots, n$, then we have

$$D_{\chi^2}(\mathbf{r}, \mathbf{w}) \leq \frac{|2a-2|}{b-a} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + 1 + a^2 + b^2 - 2(a+b) \right]. \tag{51}$$

Proof. (i): Assume the function $f(x) = (x-1)^2$ for $x \in [1, b]$ then $|f'(x)| = |2(x-1)|$. Also, $|f'(x)|$ attains its maximum value at $c = b$ i.e $|f'(c)| = |2b - 2|$. Now using all these values in (44), we get (50).
(ii): Suppose the function $f(x) = (x-1)^2$ for $x \in [a, 1]$ then $|f'(x)| = |2(x-1)|$. Also, $|f'(x)|$ attains its maximum value at $c = a$ i.e $|f'(c)| = |2a - 2|$. Now using all these values in (44), we get (51). □

For any interval $[a, b]$ with $a > 1$, we have inequality (50) while for interval $[a, b]$ with $b < 1$, we have inequality (51).

Definition 3.6 (Bhattacharya-coefficient).
Bhattacharya-coefficient C_b is given by the following formula

$$C_b(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^{n} \sqrt{r_i w_i},$$

where $\mathbf{r} = (r_1, \cdots, r_n)$ and $\mathbf{w} = (w_1, \cdots, w_n)$ are two positive probability distributions. The Bhattacharya-distance is given by $D_b(\mathbf{r}, \mathbf{w}) = -\log C_b(\mathbf{r}, \mathbf{w})$.

Corollary 3.5.
Let $[a, b] \subset \mathbb{R}^+$ and $\mathbf{r} = (r_1, \cdots, r_n)$, $\mathbf{w} = (w_1, \cdots, w_n)$ be some positive probability distributions with $\frac{r_i}{w_i} \in [a, b]$ for each $i \in \{1, 2, \cdots, n\}$, then

$$1 - C_b(\mathbf{r}, \mathbf{w}) \leq \frac{1}{2\sqrt{a}(b-a)} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + 1 + a^2 + b^2 - 2(a+b) \right]. \tag{52}$$

Proof. Let $f(x)=-\sqrt{x}$, $x\in[a,b]$ then $|f'(x)|=\frac{1}{2\sqrt{x}}$ and $\max|f'(x)|$ occurs at $c=a$. Now using (44) for $f(x)=-\sqrt{x}$ and $|f'(c)|=\frac{1}{2\sqrt{a}}$ we get (52) . □

Definition 3.7 (Hellinger-distance).
The Hellinger-distance is given by the following formula

$$D_h^2(\mathbf{r},\mathbf{w})=\frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{r_i}-\sqrt{w_i}\right)^2,$$

where $\mathbf{r}=(r_1,\cdots,r_n)$ and $\mathbf{w}=(w_1,\cdots,w_n)$ are two positive probability distributions.

Corollary 3.6.
Let $[a,b]=[a,1]\cup[1,b]\subset\mathbb{R}^+$ and $\mathbf{r}=(r_1,\cdots,r_n)$, $\mathbf{w}=(w_1,\cdots,w_n)$ be positive probability distributions then,
(i) if $\frac{r_i}{w_i}\in[1,b]$ for $i=1,2,\cdots,n$, we have

$$D_h^2(\mathbf{r},\mathbf{w})\le\frac{|\sqrt{b}-1|}{2\sqrt{b}(b-a)}\left[\sum_{i=1}^{n}\frac{r_i^2}{w_i}+1+a^2+b^2-2(a+b)\right]. \tag{53}$$

(ii) if $\frac{r_i}{w_i}\in[a,1]$ for $i=1,2,\cdots,n$, we have

$$D_h^2(\mathbf{r},\mathbf{w})\le\frac{\sqrt{a}-1}{2\sqrt{a}(b-a)}\left[\sum_{i=1}^{n}\frac{r_i^2}{w_i}+1+a^2+b^2-2(a+b)\right]. \tag{54}$$

Proof. (i): Assume the function $f(x)=\frac{1}{2}(1-\sqrt{x})^2$ for $x\in[1,b]$ then $|f'(x)|=\frac{|\sqrt{x}-1|}{2\sqrt{x}}$. Also, $|f'(x)|$ attains its maximum value at $c=b$ i.e $|f'(c)|=\frac{|\sqrt{b}-1|}{2\sqrt{b}}$. Now using all these values in (44), we get (53).
(ii): Suppose the function $f(x)=\frac{1}{2}(1-\sqrt{x})^2$, $x\in[a,1]$ then $|f'(x)|=\frac{|\sqrt{x}-1|}{2\sqrt{x}}$. Also, $|f'(x)|$ attains its maximum value at $c=a$ i.e $|f'(c)|=\frac{|\sqrt{a}-1|}{2\sqrt{a}}$. Now using all these values in (44), we get (54). □

For any interval $[a,b]$ with $a>1$, we have inequality (53) while for interval $[a,b]$ with $b<1$, we have inequality (54).

Definition 3.8 (Triangular-discrimination).
The Triangular-discrimination is given by the following formula

$$D_{\triangle}(\mathbf{r},\mathbf{w})=\sum_{i=1}^{n}\frac{(r_i-w_i)^2}{r_i+w_i},$$

where $\mathbf{r}=(r_1,\cdots,r_n)$ and $\mathbf{w}=(w_1,\cdots,w_n)$ are two positive probability distributions.

Corollary 3.7.
Let $[a,b]=[a,1]\cup[1,b]\subset\mathbb{R}^+$ and $\mathbf{r}=(r_1,\cdots,r_n)$, $\mathbf{w}=(w_1,\cdots,w_n)$ be positive probability distributions then,
(i) if $\frac{r_i}{w_i}\in[1,b]$ for $i=1,2,\cdots,n$, we have

$$D_{\triangle}(\mathbf{r},\mathbf{w})\le\frac{|b^2+2b-3|}{(b+1)^2(b-a)}\left[\sum_{i=1}^{n}\frac{r_i^2}{w_i}+1+a^2+b^2-2(a+b)\right]. \tag{55}$$

(ii) if $\frac{r_i}{w_i} \in [a, 1]$ *for* $i = 1, 2, \cdots, n$, *we have*

$$D_{\triangle}(\mathbf{r}, \mathbf{w}) \leq \frac{|a^2 + 2a - 3|}{(a+1)^2(b-a)} \left[\sum_{i=1}^{n} \frac{r_i^2}{w_i} + 1 + a^2 + b^2 - 2(a+b) \right]. \tag{56}$$

Proof. (i): Let $f(x) = \frac{(x-1)^2}{x+1}$, $x \in [1, b]$ then $|f'(x)| = \frac{|x^2+2x-3|}{(x+1)^2}$. Also, $|f'(x)|$ attains its maximum value at $c = b$ i.e $|f'(c)| = \frac{|b^2+2b-3|}{(b+1)^2}$. Now using all these values in (44), we get (55).

(ii): Assume the function $f(x) = \frac{(x-1)^2}{x+1}$ for $x \in [a, 1]$ then $|f'(x)| = \frac{|x^2+2x-3|}{(x+1)^2}$. Also, $|f'(x)|$ attains its maximum value at $c = a$ i.e $|f'(c)| = \frac{|a^2+2a-3|}{(a+1)^2}$. Now using all these values in (44), we get (56).

For any interval $[a, b]$ with $a > 1$, we have inequality (55) while for interval $[a, b]$ with $b < 1$ we have inequality (56). □

Remark 3.1.
The integral form of the above corollaries can be acquired by utilizing Theorem 2.

4. Conclusion

Jensen's inequality is one of the most widely used inequality in the field of mathematical inequalities. This inequality plays a vital role in various fields of science such as physics, computer science, financial economics and information theory etc. It is used in identification of the problems which arise in the modeling of some physical phenomenon. Researchers have obtained many results associated to the gap of Jensen's inequality. We have also determined a new simple and sharp bound for its gap. We have discussed some numerical examples and then made a comparison with an existing bound [8] in the literature, which verifies the sharpness of our new bound. By utilizing our new results, we have also suggested two new variants of the Hölder inequality and a new estimate for the gap of Hermite-Hadamard inequality. Moreover, we have given some inequalities for various divergences and for Shannon entropy, which are some applications of the main results. The work suggested in this article may motivate the researchers to extend this idea to some other inequalities.

References

[1] M. Adil Khan, G. Ali Khan, T. Ali, A. Kilicman, On the refinement of Jensen's inequality, Appl. Math. Comput., **262** (2015) 128–135.

[2] M. Adil Khan, S. Khan, Y.-M. Chu, A new bound for the Jensen gap with applications in information theory, IEEE Access, **8** (2020), 98001–98008.

[3] M. Adil Khan, S. Khan, Y.-M. Chu, New estimates for the Jensen gap using s-convexity with applications, Front. Phys., **8** (2020), Article No. 313.

[4] M. Adil Khan, T. Ali, Q. Din, A. Kilicman, Refinements of Jensen's inequality for convex functions on the co-ordinates of a rectangle from the plane, Filomat., **30** (2016), no. 3, 803–814.

[5] M. Adil Khan, M. Anwar, J. Jaksetic, J. Pečarić, On some improvements of the Jensen inequality with some applications, J. Inequal. Appl., **2009** (2009), Article ID 323615.

[6] M. Adil Khan, J. Pečarić, Pečarić, New refinement of the Jensen inequality associated to certain functions with applications, J. Inequal. Appl., **2020** (2020), Article ID 76.

[7] K. Ahmed, M. Adil Khan, S. Khan, A. Ali, Y.-M. Chu, New estimates for generalized Shannon and Zipf-Mandelbrot entropies via convexity results, Results Phys., **18** (2020), Article ID 103305.

[8] D. Costarelli, R. Spigler, How sharp is the Jensen inequality?, J. Inequal. Appl., **2015** (2015), Article ID 69.

[9] S. S. Dragomir, Some refinements of Jensen's inequality, J. Math. Anal. Appl., **168** (1992),no. 2, 518–522.

[10] L. Horváth, D. Picaric, J. Picaric, Estimations of f- and Rényi divergences by using a cyclic refinement of Jensen's Inequality, Bull. Malays. Math. Sci. Soc., **42** (2019), no. 3, 933–946.

[11] S. Khan, M. Adil Khan, Y.-M. Chu, Converses of the Jensen's Inequality derived from the Green functions with applications in infromation theory, Math. Methods Appl. Sci., **43** (2020), 2577–2587.

[12] S. Khan, M. Adil Khan, S. I. Butt, Y.-M, Chu, A new bound for the Jensen gap pertaining twice differentiable functions with applications, Adv. Difference Equ., **2020** (2020), Article ID 333.

[13] S. Khan, M. Adil Khan, Y.-M. Chu, New converses of the Jensen's Inequality via Green functions with applications, RACSAM., **114** (2020), no. 3, Article ID 114.

[14] K. A. Khan, T. Niaz, J. Pečarić, J. Pečarić, Refinement of Jensen's inequality and estimation of f-and Rényi divergence via Montgomery identity, J. Inequal. Appl., **2018** (2018), Article ID 318.

[15] V. Lakshmikantham, A. S. Vatsala, Theory of Differential and Integral Inequalities with Initial Time Difference and Applications, Springer, Berlin, (1999).

[16] J. G. Lio, A. Berg, Sharpening Jensen's inequality, Amer. Statist., **73** (2019), no. 3, 278–281.

[17] N. Lovricevic, J. Pečarić, J. Pečarić, Zipf-Mandelbrot law, f-divergences and the Jensen type interpolating inequalities, J. Inequal. Appl., **2018** (2018), Article ID 36.

[18] T. Niaz, K. A. Khan, J. Pečarić, On refinement of Jensen's inequality for 3-convex function at a point, Turkish J. Ineq., **4** (2020), no. 1, 70–80.

[19] C. E. M. Pearce, J. Pečarić, An inequality for convex functions, J. Math. Anal. Appl., **183** (1994), no. 3, 523–527.

[20] J. Pečarić, J. Pečarić, M. Rodić, About the sharpness of the Jensen inequality, J. Inequal. Appl., **2018** (2018), Article ID 337.

[21] J. Pečarić, K. K. Smoljak, New Steffensen type inequalities involving convex functions, Results Math., **67** , no. 1–2, 217–234.

[22] C. E. Shannon, A Mathematical Theory of Communication, Bell Sys Tech J., **27** (1948), no. 3, 379–423.

[23] J.-A. Wang, X.-Y. Wen, B.-Y. Hou, Advance stability criteria for static neural networks with interval time-varying delays via the improved Jensen inequality, Neurocomputing **377** (2020), 49–56.

[24] Z. H. Yang, W. M. Qian, Y. -M. Chu, W. Zhang, On approximating the error function, Math. Inequal. Appl.,

21 (2018), no. 2, 469–479.

[25] Z. H. Yang, W. M. Qian, Y. -M. Chu, W. Zhang, On approximating the arithmetic-geometric mean and complete elliptic integral of the first kind, J. Math. Anal. Appl., **462** (2018), no. 2, 1714–1726.

[26] Z. H. Yang, W. Zhang, Y. -M. Chu, Sharp Gautschi inequality for parameter $0 < p < 1$ with applications, Math. Inequal. Appl., **20** (2017), no. 4, 1107–1120.

[27] X.-L. Zhu, G.-H. Yang, Jensen's inequality approach to stability analysis of discrete-time systems with time-varying delay, IEEEXplore., DOI:10.1109/ACC.2008.4586727.

[28] X.-L. Zhu, G.-H. Yang, Jensen's integral inequality approach to stability analysis of continuous-time systems with time-varying delay, IET Control Theory Appl., **2** (2008), no. 6, 524–534.