Refinements of Jensen’s Inequality via Majorization Results with Applications in the Information Theory

Yongping Deng, Hidayat Ullah, Muhammad Adil Khan, Sajid Iqbal, and Shanhe Wu

1Department of Mathematics, Longyan University, Longyan 364012, China
2Department of Mathematics, University of Peshawar, Peshawar 25000, Pakistan
3Department of Mathematics, Riphah International University, Faisalabad Campus, Satyana Road, Faisalabad, Pakistan

Correspondence should be addressed to Shanhe Wu; shanhely@126.com

Received 28 May 2021; Accepted 2 August 2021; Published 18 August 2021

Academic Editor: Georgios Psarrakos

Copyright © 2021 Yongping Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, we present some new refinements of the Jensen inequality with the help of majorization results. We use the concept of convexity along with the theory of majorization and obtain refinements of the Jensen inequality. Moreover, as consequences of the refined Jensen inequality, we derive some bounds for power means and quasiarithmetic means. Furthermore, as applications of the refined Jensen inequality, we give some bounds for divergences, Shannon entropy, and various distances associated with probability distributions.

1. Introduction

The Jensen inequality is one of the most significant and fundamental inequalities in the current literature of mathematical inequalities. The Jensen inequality got more importance due to the fact that many inequalities such as Hölder inequality, Minkowski’s inequality, Ky Fan’s inequality, Levinson’s inequality, and Hermite–Hadamard inequality can be obtained from this inequality. Also, many problems arising in different fields of science can be solved with the help of the Jensen inequality [1–3]. Moreover, several applications of the Jensen inequality have been observed in the estimations of several divergences and the well-known Zipf–Mandelbrot entropy [4–9]. Due to the wide significance of the Jensen inequality, many researchers gave several extensions, improvements, generalizations, and refinements of this inequality. Furthermore, this inequality has been also used as a foundation for exploring the generalized convex functions such as s-convex function, η-convex function, coordinate convex function, and h-convex function.

The discrete form of the Jensen inequality is given in the following theorem.

**Theorem 1.** Assume that $\Psi: I \rightarrow \mathbb{R}$ is a convex function ($I$ is an arbitrary interval), $\zeta_k \in I$ and $\eta_k \geq 0$, for all $k \in \{1, 2, \ldots, n\}$, with $\eta = \sum_{k=1}^{n} \eta_k > 0$, then

$$
\Psi \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \zeta_k \right) \leq \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \Psi (\zeta_k).
$$

(1)

If a function $\Psi$ is concave, then (1) is valid in the contrary direction.

One of the most attractive features of the Jensen inequality is that it generalized the definition of the convex function. Recently, many researchers devoted their attention to the Jensen inequality and contributed to it in various directions. In 2006, Dragomir [10] presented new refinements of the Jensen inequality by using a convex function defined on the linear space. He also gave some applications...
of the obtained results in the information theory. Pečarić and Perić [11] established some improvements of the converse of the Jensen inequality for linear functional, and some applications of the obtained results were also presented therein. Xiao et al. [12] obtained a new extension and refinement of the Jensen inequality for midconvex functions by using the arithmetic means of the given function.

In the remaining portion of the present section, we introduce some fundamental theory and essential results of majorization, which will help us in the obtaining our new results.

Majorization is a connection between two vectors, which shows that the components of one vector is "more nearly equal" than the components of the second vector. Due to this property, the concept of majorization becomes a focus point for the researchers of different fields. Many mathematical inequalities have been established with the help of majorization [13]. Over the years, the concept of majorization has been applied widely as a powerful tool in the area of pure and applied mathematics, as well as some other related fields. Majorization also performed an extremely significant character in the theory of matrices. Several matrix inequalities and norm inequalities have been deduced by the majorization relation among the Eigen values and singular values of matrices [14]. For some literature related to majorization theory, refer [15–18].

For fixed $n \geq 2$, let
\[ s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n), \]
\[ s_2 = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n), \]
be two $n$-tuples of real numbers, and we denote the components of $s_1$ and $s_2$ in the decreasing order by
\[ \zeta_1 \geq \zeta_2 \geq \cdots \geq \zeta_n, \]
\[ \vartheta_1 \geq \vartheta_2 \geq \cdots \geq \vartheta_n. \]

**Definition 1.** The tuple $s_2$ is said to be majorized by the tuple $s_1$, namely, $s_2 \prec s_1$, if
\[ \sum_{k=1}^{l} \vartheta_k \leq \sum_{k=1}^{l} \zeta_k, \]
\[ \text{for } l = 1, 2, \ldots, n-1, \]
\[ \sum_{k=1}^{n} \vartheta_k = \sum_{k=1}^{n} \zeta_k. \]

Symbolically, it can be written as $s_2 < s_1$.

The following theorem is very famous in the literature as majorization theorem, and it can be found in [13].

**Theorem 2.** Let $s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n)$ and $s_2 = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n)$ be two $n$-tuples, such that $\zeta_k, \vartheta_k \in I (k = 1, 2, \ldots, n)$. Then, the inequality
\[ \sum_{k=1}^{n} \zeta_k \leq \sum_{k=1}^{n} \zeta_k, \]
is true for each continuous convex function $\Psi : I \rightarrow \mathbb{R}$ if and only if $s_2 \prec s_1$.

The generalization of the above majorization inequality is given in following theorem [19, 20].

**Theorem 3.** Let a function $\Psi : I \rightarrow \mathbb{R}$ be convex and $s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n)$ and $s_2 = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n)$ be two $n$-tuples, such that $\zeta_k, \vartheta_k \in I (k = 1, 2, \ldots, n)$ and $s_3 = (\eta_1, \eta_2, \ldots, \eta_n)$ be nonnegative $n$-tuples satisfying
\[ \sum_{k=1}^{l} \eta_k \vartheta_k \leq \sum_{k=1}^{l} \eta_k \zeta_k, \]
\[ \text{for } l = 1, \ldots, n-1, \]
\[ \sum_{k=1}^{n} \eta_k \vartheta_k = \sum_{k=1}^{n} \eta_k \zeta_k. \]

If the tuple $s_2$ is decreasing, then the following inequality holds:
\[ \sum_{k=1}^{n} \eta_k \Psi(\vartheta_k) \leq \sum_{k=1}^{n} \eta_k \Psi(\zeta_k). \]

In the following theorem, Dragomir [21] obtained a weighted majorization inequality with the help of support line inequality and Chebyshev’s inequality.

**Theorem 4.** Suppose that a function $\Psi : I \rightarrow \mathbb{R}$ is convex and $s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n)$ and $s_2 = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n)$ are any two $n$-tuples, such that $\zeta_k, \vartheta_k \in I (k = 1, 2, \ldots, n)$, and $s_3 = (\eta_1, \eta_2, \ldots, \eta_n)$ is a nonnegative real $n$-tuple. Also, suppose that condition (7) is satisfied. If $s_2$ and $s_1$-$s_2$ are monotonic in the parallel direction, then
\[ \sum_{k=1}^{n} \eta_k \Psi(\vartheta_k) \leq \sum_{k=1}^{n} \eta_k \Psi(\zeta_k). \]

**Remark 1.** Dragomir also obtained inequality (9) by using strict condition that $\Psi$ is an increasing convex function while using the relax condition that $\sum_{k=1}^{n} \eta_k \vartheta_k \leq \sum_{k=1}^{n} \eta_k \zeta_k$ instead of using (7).

The majorization inequalities are generalized, extended, and refined in many directions. Adil Khan et al. [22] extended the majorization type inequalities from the convex functions defined on the intervals to the convex functions defined on the rectangles. They used the main result to obtain some extensions for the weighted Favard’s inequality. Wu et al. [23] obtained some refinements of majorization-type inequalities with the help of Taylor theorem with the mean value form of the remainder and also presented an application. Ullah et al. [24] established some improvements and generalizations of majorization-type inequalities with the help of a strongly convex function and also gave some applications of the obtained results.

The aim of this study is to establish some refinements of the famous Jensen equality with the help of majorization theory. We use some basic and fundamental results of majorization to obtain some new refinements of the Jensen inequality. We also establish refined inequalities for different means. As applications, we obtain new bounds for the Csiszár divergence, Kullback–Leibler divergence, and
Shannon entropy. Furthermore, we also establish bounds for the various distances, Bhattacharyya coefficient, and triangular discrimination.

2. Main Results

We begin this section with the following result, in which we obtain a refinement of the Jensen inequality with the help of Theorem 2.

Theorem 5. Let \( \Psi: I \to \mathbb{R} \) be a convex function, and let \( s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n) \) and \( s_2 = (\theta_1, \theta_2, \ldots, \theta_n) \) be two \( n \)-tuples, such that \( \zeta_k, \theta_k \in I \) for all \( k \in \{1, 2, \ldots, n\} \). Also, let \( \ell \in [0, 1] \) and \( \bar{\zeta} = (1/n) \sum_{k=1}^n \zeta_k \). If \( s_2 < s_1 \), then

\[
\Psi(\bar{\zeta}) \leq \frac{1}{n} \sum_{k=1}^n \Psi((1-\ell)\zeta + \ell\theta_k)
\]

\[
\leq (1-\ell)\Psi(\bar{\zeta}) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\theta_k)
\]

\[
\leq \frac{1}{n} \sum_{k=1}^n (1-\ell)\Psi(\zeta_k) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\theta_k)
\]

\[
\leq \frac{1}{n} \sum_{k=1}^n \Psi(\zeta_k).
\]

If \( \Psi \) is a concave function, then the inequalities given in (10) are true in the opposite direction.

Proof. Since \( s_2 < s_1 \), using the second condition of majorization, we have

\[
\Psi \left( \frac{1}{n} \sum_{k=1}^n \zeta_k \right) = \Psi \left( \frac{1}{n} \sum_{k=1}^n ((1-\ell)\zeta + \ell\theta_k) \right).
\]

Now, applying Jensen’s inequality to the right hand side of (11), we get

\[
\Psi \left( \frac{1}{n} \sum_{k=1}^n ((1-\ell)\zeta + \ell\theta_k) \right) \leq \frac{1}{n} \sum_{k=1}^n \Psi((1-\ell)\zeta + \ell\theta_k).
\]

Applying the definition of the convex function to the right hand side of (12), we obtain

\[
\frac{1}{n} \sum_{k=1}^n \Psi((1-\ell)\zeta + \ell\theta_k) \leq (1-\ell)\Psi(\bar{\zeta}) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\theta_k).
\]

\[
\text{Instantly, applying Jensen’s inequality to the first term on the right hand side of (13), we acquire}
\]

\[
(1-\ell)\Psi(\bar{\zeta}) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\theta_k) \leq \frac{1-\ell}{n} \sum_{k=1}^n \Psi(\zeta_k) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\theta_k).
\]

As \( s_2 < s_1 \), using majorization theorem, we have

\[
\sum_{k=1}^n \Psi(\theta_k) \leq \sum_{k=1}^n \Psi(\zeta_k).
\]

Now, using (15) in (14), we obtain

\[
\frac{1-\ell}{n} \sum_{k=1}^n \Psi(\zeta_k) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\theta_k) \leq \frac{1-\ell}{n} \sum_{k=1}^n \Psi(\zeta_k) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\zeta_k)
\]

\[
= \frac{1}{n} \sum_{k=1}^n \Psi(\zeta_k).
\]

From (12)–(14) and (16), we obtain (10). Replacing \( \Psi \) with \( -\Psi \), we deduce the opposite inequalities for the case that \( \Psi \) is a concave function. This completes the proof of Theorem 5. \( \blacksquare \)

Corollary 1. Assume that all the assumptions of Theorem 5 are fulfilled; then, we have

\[
\Psi(\bar{\zeta}) \leq \frac{1}{n} \sum_{k=1}^n \frac{1}{\theta_k - \bar{\zeta}} \int_{\bar{\zeta}}^{\theta_k} \Psi(u) \, du \leq \frac{1}{2} \Psi(\bar{\zeta}) + \frac{1}{2n} \sum_{k=1}^n \Psi(\theta_k)
\]

\[
\leq \frac{1}{2n} \sum_{k=1}^n \Psi(\zeta_k) + \frac{1}{2n} \sum_{k=1}^n \Psi(\theta_k) \leq \frac{1}{n} \sum_{k=1}^n \Psi(\zeta_k).
\]

Proof. Integrating (10) over \( \ell \in [0, 1] \), we obtain (17). \( \blacksquare \)

Remark 2. Since \( s_1 < s_1 \), applying Theorem 5 and Corollary 1 for \( \theta_k = \zeta_k \), where \( k \in \{1, 2, \ldots, n\} \), we have the following inequalities:

\[
\Psi(\bar{\zeta}) \leq \frac{1}{n} \sum_{k=1}^n \Psi((1-\ell)\zeta + \ell\zeta_k) \leq (1-\ell)\Psi(\bar{\zeta}) + \ell \frac{1}{n} \sum_{k=1}^n \Psi(\zeta_k).
\]

\[
\Psi(\bar{\zeta}) \leq \frac{1}{n} \sum_{k=1}^n \frac{1}{\zeta_k - \bar{\zeta}} \int_{\bar{\zeta}}^{\zeta_k} \Psi(u) \, du \leq \frac{1}{2} \Psi(\bar{\zeta}) + \frac{1}{2n} \sum_{k=1}^n \Psi(\zeta_k) \leq \frac{1}{n} \sum_{k=1}^n \Psi(\zeta_k).
\]
Theorem 6. Let $\Psi : I \rightarrow \mathbb{R}$ be a convex function; let $s_1 = (\xi_1, \xi_2, \ldots, \xi_n)$ and $s_2 = (\eta_1, \eta_2, \ldots, \eta_n)$ be two $n$-tuples, such that their entries are from the interval $I$, and let $s_3 = (\eta_1, \eta_2, \ldots, \eta_n)$ be the nonnegative $n$-tuple with $\eta = \sum_{k=1}^n \eta_k > 0$. Suppose that the conditions (6) and (7) are satisfied, $\zeta = (1/\eta)\sum_{k=1}^n \eta_k \xi_k$ and $\ell \in [0, 1]$. If $s_2$ is a decreasing $n$-tuple, then the following inequalities hold:

\[ \Psi(\zeta) \leq \frac{1}{\eta} \sum_{k=1}^n \eta_k \Psi((1 - \ell)\xi + \ell \eta_k) \leq (1 - \ell)\Psi(\zeta) + \frac{\ell}{\eta} \sum_{k=1}^n \eta_k \Psi(\xi_k) \]

From (21)–(23) and (25), it follows (19).

Replacing $\Psi$ with $-\Psi$, we deduce the opposite inequalities for the case that $\Psi$ is a concave function. The proof of Theorem 6 is complete.

Theorem 7. Under the assumptions of Theorem 4, if $\eta = \sum_{k=1}^n \eta_k > 0$ and $\ell \in [0, 1]$, then the inequalities in (19) hold.

Proof. Theorem 7 can be proved easily by using the idea of Theorem 6 along with the result of Theorem 4.

Corollary 2. Under the assumptions of Theorem 6, the following inequalities hold:

\[ \Psi(\zeta) \leq \frac{1}{\eta} \sum_{k=1}^n \eta_k \Psi((1 - \ell)\xi + \ell \eta_k) \leq (1 - \ell)\Psi(\zeta) + \frac{\ell}{\eta} \sum_{k=1}^n \eta_k \Psi(\xi_k) \leq \frac{1}{2} \sum_{k=1}^n \eta_k \Psi(\xi_k) \]

Proof. Integrating (19) over $\ell \in [0, 1]$, we obtain (26).

Remark 3. If $s_1 = s_2$, then obviously the conditions (6) and (7) are satisfied. So, in this case, the inequalities (19) and (26) become

\[ \Psi(\zeta) \leq \frac{1}{\eta} \sum_{k=1}^n \eta_k \Psi((1 - \ell)\xi + \ell \xi_k) \]

\[ \leq (1 - \ell)\Psi(\zeta) + \frac{\ell}{\eta} \sum_{k=1}^n \eta_k \Psi(\xi_k) \]

(27)

\[ \Psi(\zeta) \leq \frac{1}{\eta} \sum_{k=1}^n \eta_k \Psi((1 - \ell)\xi_k - \ell \xi_k) \]

\[ \leq \frac{1}{2} \Psi(\zeta) + \frac{\ell}{2\eta} \sum_{k=1}^n \eta_k \Psi(\xi_k) \leq \frac{1}{2} \sum_{k=1}^n \eta_k \Psi(\xi_k) \]

The inequalities (18) and (28) can also be found in Remark 2 described in [25].
3. Applications to Inequalities Involving Special Means

In this section, we give some applications of Theorem 6 for power means and quasiarithmetic means.

**Definition 2.** Let \( s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n) \) and \( s_3 = (\eta_1, \eta_2, \ldots, \eta_n) \) be two nonnegative \( n \)-tuples and \( \eta = \sum_{k=1}^{n} \eta_k > 0 \). Then, the power mean of order \( \rho \in \mathbb{R} \) is defined by

\[
\mathcal{M}_\rho(s_1; s_3) = \begin{cases} 
\left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\rho}, & \text{if } \rho \neq 0, \\
\left( \prod_{k=1}^{n} \zeta_k^{\eta_k} \right)^{1/\rho}, & \text{if } \rho = 0.
\end{cases}
\]

(29)

We establish the refinement of inequalities involving power means as follows.

**Corollary 3.** Let \( S_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n), S_2 = (\theta_1, \theta_2, \ldots, \theta_n), \) and \( S_3 = (\eta_1, \eta_2, \ldots, \eta_n) \) be the nonnegative \( n \)-tuples, such that

\[
\sum_{k=1}^{l} \eta_k \theta_k^\rho \leq \sum_{k=1}^{l} \eta_k \zeta_k^\rho, \quad \text{for } l = 1, 2, \ldots, n-1,
\]

\[
\sum_{k=1}^{n} \eta_k \theta_k^\rho = \sum_{k=1}^{n} \eta_k \zeta_k^\rho.
\]

(30)

If the \( n \)-tuple \( S_3 = (\theta_1, \theta_2, \ldots, \theta_n) \) is decreasing and \( \rho \in \mathbb{R} \setminus \{0\} \), such that \( \rho \leq \xi \), then the following inequalities hold:

\[
\mathcal{M}_\rho(s_1; s_3) \leq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \left( (1-\ell) \mathcal{M}_\rho(s_1; s_3) + \ell \mathcal{M}(s_2; s_3) \right) \right)^{1/\xi}
\]

\[
\leq \left( (1-\ell) \mathcal{M}_\xi(s_1; s_3) + \ell \mathcal{M}(s_2; s_3) \right)^{1/\xi}
\]

\[
\leq \left( (1-\ell) \mathcal{M}_\xi(s_1; s_3) + \ell \mathcal{M}_\xi(s_2; s_3) \right)^{1/\xi}
\]

\[
\leq \mathcal{M}_\xi(s_1; s_3), \quad \xi \neq 0,
\]

(31)

\[
\mathcal{M}_\rho(s_1; s_1) \leq \exp \left( \frac{1}{\rho \eta} \sum_{k=1}^{n} \eta_k \log \left( (1-\ell) \mathcal{M}_\rho(s_1; s_3) + \ell \mathcal{M}_\rho(s_2; s_3) \right) \right)
\]

\[
\leq \exp (1-\ell) \log \mathcal{M}_\rho(s_1; s_3) + \ell \log \mathcal{M}_\rho(s_2; s_3)
\]

\[
\leq \mathcal{M}_\rho(s_1; s_3), \quad \xi = 0,
\]

(32)

where \( \ell \in [0, 1] \) and \( \eta = \sum_{k=1}^{n} \eta_k > 0 \).

**Proof.** Note that the function \( \Psi(\nu) = \nu^{\ell/\xi} \) is convex on \((0, \infty)\) for \( \{ (\rho, \xi) | 0 < \rho < \xi \} \) or \( \{ (\rho, \xi) | 0 < \xi, \xi > 0 \} \). Therefore, utilizing (19) and substituting \( \zeta_k = \zeta_k^\rho, \theta_k = \theta_k^\rho, \) and \( \Psi(\nu) = \nu^{\ell/\xi} \), we get

\[
\left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} \leq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \left( (1-\ell) \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho + \ell \mathcal{M}_\rho(s_1; s_3) \right)^{1/\xi} \right)^{1/\ell}
\]

\[
\leq \left( 1-\ell \right) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} + \ell \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \left( \mathcal{M}_\rho(s_1; s_3) \right)^{1/\xi} \right)^{1/\ell}
\]

\[
\leq \left( 1-\ell \right) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} + \ell \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}
\]

\[
\leq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}.
\]

(33)

Taking \( \xi^{th} \) root of (33), we obtain

\[
\left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} \leq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \left( 1-\ell \right) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\xi} \right)^{1/\ell}
\]

\[
\leq \left( 1-\ell \right) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} + \ell \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}
\]

\[
\leq \left( 1-\ell \right) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} + \ell \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}
\]

\[
\leq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell},
\]

(34)

which is equivalent to (31).

Since the function \( \Psi(\nu) = \nu^{\ell/\xi} \) is concave on \((0, \infty)\) for \( \{ (\rho, \xi) | 0 \leq \xi \leq 0 \} \), utilizing Theorem 6 and choosing \( \Psi(\nu) = \nu^{\ell/\xi} \), \( \zeta_k = \zeta_k^\rho \) and \( \theta_k = \theta_k^\rho \), we obtain

\[
\left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{\ell/\xi} \geq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{\ell/\xi}
\]

\[
\geq (1-\ell) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} + \ell \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}
\]

\[
\geq (1-\ell) \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell} + \ell \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}
\]

\[
\geq \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k \zeta_k^\rho \right)^{1/\ell}.
\]

(35)

Therefore, taking power \((1/\xi)(\xi < 0)\) of (35), we get
\[
\left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k c_k^\xi \right)^{1/\xi} \leq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \left( (1 - \ell) \frac{1}{\eta} \sum_{k=1}^{n} \eta_k c_k^\xi + \ell \theta_k^\xi \right)^{1/\xi} \right) \\
\leq \left( (1 - \ell) \frac{1}{\eta} \sum_{k=1}^{n} \eta_k c_k^\xi \right)^{1/\xi} + \ell \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \theta_k^\xi \right)^{1/\xi} \\
\leq \left( (1 - \ell) \frac{1}{\eta} \sum_{k=1}^{n} \eta_k c_k^\xi \right)^{1/\xi} + \ell \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \theta_k^\xi \right)^{1/\xi} \\
\leq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k c_k^\xi \right)^{1/\xi}
\] (36)

Clearly, (36) is equivalent to (31). The inequalities in (31) are proved.

Taking the logarithm (log) of (31) and letting \( \xi \rightarrow 0 \), we have

\[
\log \mathcal{M}_\rho (s_1; s_3) \leq \lim_{\xi \rightarrow 0} \log \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \left( (1 - \ell) \mathcal{M}_\rho (s_1; s_3) + \ell \theta_k^\xi \right)^{1/\xi} \right) \\
\leq \lim_{\xi \rightarrow 0} \log \left( (1 - \ell) \mathcal{M}_\rho^\xi (s_1; s_3) + \ell \mathcal{M}_\rho^\xi (s_2; s_3) \right) \\
\leq \lim_{\xi \rightarrow 0} \log \left( (1 - \ell) \mathcal{M}_\rho^\xi (s_1; s_3) + \ell \mathcal{M}_\rho^\xi (s_2; s_3) \right) \frac{1}{\xi} \\
\leq \lim_{\xi \rightarrow 0} \log \mathcal{M}_\rho (s_1; s_3) \\
\] (37)

Applying L'Hospital’s rule for limit, we obtain

\[
\log \mathcal{M}_\rho (s_1; s_3) \leq \frac{1}{\rho \eta} \sum_{k=1}^{n} \eta_k \log \left( (1 - \ell) \mathcal{M}_\rho (s_1; s_3) + \ell \theta_k^\xi \right) \\
\leq (1 - \ell) \log \mathcal{M}_\rho (s_1; s_3) + \ell \log \mathcal{M}_\rho (s_2; s_3) \\
\leq (1 - \ell) \log \mathcal{M}_\rho (s_1; s_3) + \ell \log \mathcal{M}_0 (s_2; s_3) \\
\leq \log \mathcal{M}_0 (s_1; s_3) \\
\] (38)

Now, taking the exponent (exp) of (38), we get

\[
\mathcal{M}_\rho (s_1; s_3) \leq \exp \left( \frac{1}{\rho \eta} \sum_{k=1}^{n} \eta_k \log \left( (1 - \ell) \mathcal{M}_\rho (s_1; s_3) + \ell \theta_k^\xi \right) \right) \\
\leq \exp \left( (1 - \ell) \log \mathcal{M}_\rho (s_1; s_3) + \ell \log \mathcal{M}_\rho (s_2; s_3) \right) \\
\leq \exp \left( (1 - \ell) \log \mathcal{M}_0 (s_1; s_3) + \ell \log \mathcal{M}_0 (s_2; s_3) \right) \\
\leq \mathcal{M}_0 (s_1; s_3),
\] (39)

which is the required inequality (32). \(\square\)
Corollary 4. Assume that \( S_1 = (\zeta^1_1, \zeta^1_2, \ldots, \zeta^1_n), S_2 = (\xi^2_1, \xi^2_2, \ldots, \xi^2_n), S_3 = (\zeta_1, \zeta_2, \ldots, \zeta_n), S_4 = (\xi_1, \xi_2, \ldots, \xi_n) \), and \( s_1 = (\eta_1, \eta_2, \ldots, \eta_n) \) are nonnegative \( n \)-tuples, such that

\[
\begin{align*}
\sum_{k=1}^{l} \eta_k \xi_k^l &\leq \sum_{k=1}^{l} \eta_k \xi_k^l, \quad \text{for } l = 1, 2, \ldots, n - 1, \\
\sum_{k=1}^{n} \eta_k \xi_k^l &\leq \sum_{k=1}^{n} \eta_k \xi_k^l.
\end{align*}
\]

(40)

If the tuple \( S_2 = (\xi^2_1, \xi^2_2, \ldots, \xi^2_n) \) is decreasing and \( \xi \in \mathbb{R} \setminus \{0\} \), such that \( \rho \leq \xi \), then we have the following inequalities:

\[
\begin{align*}
\mathcal{M}_\xi(s_1; s_2) &\geq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \left( (1 - \rho) \mathcal{M}^{\rho}(s_1; s_3) + \epsilon \xi \right) \right)^{1/\rho} \\
&\geq (1 - \rho) \mathcal{M}^{\rho}(s_1; s_2) + \epsilon \mathcal{M}^{\rho}(s_2; s_3)^{1/\rho} \\
&\geq (1 - \rho) \mathcal{M}^{\rho}(s_1; s_2) + \epsilon \mathcal{M}^{\rho}(s_2; s_3)^{1/\rho} \\
&\geq \mathcal{M}_\rho(s_1; s_3), \quad \rho \neq 0,
\end{align*}
\]

(41)

\[
\begin{align*}
\mathcal{M}_\xi(s_1; s_3) &\geq \exp \left( \frac{1}{\xi} \sum_{k=1}^{n} \eta_k \log \left( (1 - \rho) \mathcal{M}^{\rho}(s_1; s_3) + \epsilon \xi \right) \right) \\
&\geq \exp \left( (1 - \rho) \log \mathcal{M}(s_1; s_3) + \epsilon \log \mathcal{M}(s_2; s_3) \right) \\
&\geq \exp \left( (1 - \rho) \log \mathcal{M}(s_1; s_3) + \epsilon \log \mathcal{M}(s_2; s_3) \right) \\
&\geq \mathcal{M}(s_1; s_3), \quad \rho = 0,
\end{align*}
\]

(42)

where \( \eta = \sum_{k=1}^{n} \eta_k > 0 \) and \( \rho \in [0, 1] \).

Proof. Since the function \( \Psi(u) = u^{\rho/\xi} \) is convex on \( (0, \infty) \) for \( (\rho, \xi)|0 < \rho \leq \xi < 0 \) or \( (\rho, \xi)|0 < \xi \leq \rho, \xi > 0 \), utilizing (19) and choosing \( \zeta_k = \xi^1_k, \delta_k = \xi^2_k \), and \( \Psi(u) = u^{\rho/\xi} \), we get

\[
\begin{align*}
\left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{\rho/\xi} &\leq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \left( (1 - \rho) \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k + \epsilon \xi \right) \right)^{\rho/\xi} \\
&\leq (1 - \rho) \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{\rho/\xi} + \epsilon \sum_{k=1}^{n} \eta_k (\xi^2_k)^{\rho/\xi} \\
&\leq \frac{1 - \rho}{\eta} \sum_{k=1}^{n} \eta_k (\xi^1_k)^{\rho/\xi} + \frac{\epsilon}{\eta} \sum_{k=1}^{n} \eta_k (\xi^2_k)^{\rho/\xi} \\
&\leq \frac{1}{\eta} \sum_{k=1}^{n} \eta_k (\xi^1_k)^{\rho/\xi}.
\end{align*}
\]

(43)

Now, taking \( \rho \)th root of (43), we acquire

\[
\begin{align*}
\left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho} &\leq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \left( (1 - \rho) \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k + \epsilon \xi \right) \right)^{1/\rho} \\
&\geq (1 - \rho) \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho} + \epsilon \frac{\eta}{\eta} \sum_{k=1}^{n} \eta_k (\xi^2_k)^{1/\rho} \\
&\geq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho}.
\end{align*}
\]

(44)

For \( \{ (\rho, \xi)|0 < \rho \leq \xi \} \), one has

\[
\begin{align*}
\left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho} &\geq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \left( (1 - \rho) \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k + \epsilon \xi \right) \right)^{1/\rho} \\
&\geq \left( 1 - \rho \right) \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho} + \epsilon \frac{\eta}{\eta} \sum_{k=1}^{n} \eta_k (\xi^2_k)^{1/\rho} \\
&\geq \left( 1 - \rho \right) \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho} + \epsilon \frac{\eta}{\eta} \sum_{k=1}^{n} \eta_k (\xi^2_k)^{1/\rho} \\
&\geq \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k \xi^1_k \right)^{1/\rho}.
\end{align*}
\]

(45)

Inequalities (45) and (44) imply the required inequality (41) as \( \rho \to 0 \). \( \square \)

Definition 3. Let \( s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n) \) and \( s_2 = (\xi_1, \xi_2, \ldots, \xi_n) \) be two nonnegative \( n \)-tuples and \( \overline{\eta} = \sum_{k=1}^{n} \eta_k > 0 \). If a function \( h \) is continuous and strictly monotonic, then the quasiarithmetic mean is defined as

\[
M_h(s_1; s_2) = h^{-1} \left( \frac{1}{\eta} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right).
\]

(46)

Corollary 5. Suppose that \( h \) is continuous and strictly monotonic on \( (0, \infty) \). Also, suppose that \( s_1 = (\zeta_1, \zeta_2, \ldots, \zeta_n), s_2 = (\xi_1, \xi_2, \ldots, \xi_n), \) and \( s_3 = (\eta_1, \eta_2, \ldots, \eta_n) \) are the positive \( n \)-tuples with

\[
\sum_{k=1}^{l} \eta_k h(\zeta_k) \leq \sum_{k=1}^{l} \eta_k h(\zeta_k), \quad l = 1, 2, \ldots, n - 1,
\]

(47)

\[
\sum_{k=1}^{n} \eta_k h(\zeta_k) = \sum_{k=1}^{n} \eta_k h(\zeta_k).
\]

If \( (h(\zeta_1), h(\zeta_2), \ldots, h(\zeta_n)) \) is decreasing and the function \( \Phi \circ h^{-1} \) is convex on \( (0, \infty) \), then
If \((h(\vartheta_1), h(\vartheta_2), \ldots, h(\vartheta_n))\) is decreasing and the function \(\Phi \circ h^{-1}\) is concave, then the inequality (48) holds in the opposite direction.

**Proof.** Using (19) for \(\zeta_k = h(\zeta_k), \vartheta_k = h(\vartheta_k),\) and \(\Psi(\upsilon) = \Phi \circ h^{-1}(\upsilon),\) we get

\[
\Phi \left( M_h(s_1; s_3) \right) \leq \frac{1}{n} \sum_{k=1}^{n} \eta_k \Phi \circ h^{-1} \left( (1 - \ell) h \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right) + \ell h(\vartheta_k) \right)
\]

\[
\leq (1 - \ell) \Phi \left( M_h(s_1; s_3) \right) + \frac{\ell}{n} \sum_{k=1}^{n} \eta_k \Phi(\zeta_k)
\]

\[
\leq \frac{1}{n} \sum_{k=1}^{n} \eta_k \Phi(\zeta_k).
\]

(48)

If \((h(\vartheta_1), h(\vartheta_2), \ldots, h(\vartheta_n))\) is decreasing and the function \(\Phi \circ h^{-1}\) is concave, then the inequality (48) holds in the opposite direction.

**Proof.** Using (19) for \(\zeta_k = h(\zeta_k), \vartheta_k = h(\vartheta_k),\) and \(\Psi(\upsilon) = \Phi \circ h^{-1}(\upsilon),\) we get

\[
\Phi \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right) \leq \frac{1}{n} \sum_{k=1}^{n} \eta_k \Phi \circ h^{-1} \left( (1 - \ell) \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) + \ell h(\vartheta_k) \right)
\]

\[
\leq (1 - \ell) \Phi \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right) + \frac{\ell}{n} \sum_{k=1}^{n} \eta_k \Phi \circ h^{-1}(h(\vartheta_k))
\]

\[
\leq \frac{1}{n} \sum_{k=1}^{n} \eta_k \Phi \circ h^{-1}(h(\vartheta_k))
\]

(49)

this implies that

\[
\Phi \left( h^{-1} \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right) \right) \leq \frac{1}{n} \sum_{k=1}^{n} \eta_k \Phi \circ h^{-1} \left( (1 - \ell) h \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right) \right) + \ell h(\vartheta_k)
\]

\[
\leq (1 - \ell) \Phi \left( h^{-1} \left( \frac{1}{n} \sum_{k=1}^{n} \eta_k h(\zeta_k) \right) \right)
\]

\[
+ \frac{\ell}{n} \sum_{k=1}^{n} \eta_k \Phi(\vartheta_k)
\]

(50)

which is equivalent to (48).
4. Applications to Information Theory

The information theory is the mathematical treatment of the concepts, framework, and rule governing the spread of messages through communication systems. The information theory has been used very extensively in the telecommunication systems, and it has also a great role in the study of linguistic studies such as in the sense of speed of reading, length of words, and frequency of words. In this section, we give some meaningful applications of our main results in the area of information theory. We will establish bounds for the different divergences and Shannon entropy. We also establish bounds for various distances, Bhattacharyya coefficient, and triangular discrimination.

Definition 4 (See [26]). Let \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_2 = (\tau_1, \tau_2, \ldots, \tau_n) \) be two positive \( n \)-tuples, and let \( \Phi : (0, \infty) \rightarrow \mathbb{R} \) be a convex function. Then, the Čsiszár \( \Phi \)-divergence functional is defined as

\[
M_\Phi (s_1, s_2) = \sum_{k=1}^{n} r_k \Phi \left( \frac{\theta_k}{\tau_k} \right).
\]  

(51)

Theorem 8. Suppose that \( \Phi : (0, \infty) \rightarrow \mathbb{R} \) is a convex function, suppose also that \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \), \( s_2 = (\tau_1, \tau_2, \ldots, \tau_n) \), and \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) are three positive \( n \)-tuples, such that

\[
\sum_{k=1}^{l} \theta_k \geq \sum_{k=1}^{l} \tau_k \theta_k, \quad \text{for } l = 1, 2, \ldots, n-1,  \tag{52}
\]

\[
\sum_{k=1}^{n} \theta_k = \sum_{k=1}^{n} \tau_k \theta_k. \tag{53}
\]

If \( (\theta_1, \theta_2, \ldots, \theta_n) \) is a decreasing tuple, \( \ell \in [0, 1] \), then

\[
M_\Phi (s_1, s_3) \geq (1 - \ell)M_\Phi (s_1, s_2) + \ell \sum_{k=1}^{n} r_k \Phi (\theta_k)
\]

\[
\geq (1 - \ell)\Phi \left( \frac{\sum_{k=1}^{n} \theta_k}{\sum_{k=1}^{n} \tau_k} \right) \sum_{k=1}^{n} \tau_k + \ell \sum_{k=1}^{n} r_k \Phi (\theta_k)
\]

\[
\geq \sum_{k=1}^{n} r_k \Phi \left( \frac{1 - \ell}{\sum_{k=1}^{n} \tau_k} \sum_{k=1}^{n} \tau_k + \ell \theta_k \right)
\]

\[
\geq \Phi \left( \frac{\sum_{k=1}^{n} \theta_k}{\sum_{k=1}^{n} \tau_k} \right) \sum_{k=1}^{n} \tau_k. \tag{54}
\]

Proof. Note that the function \( \Phi \) is convex, and the conditions of Theorem 6 are satisfied. Therefore, applying Theorem 6 by choosing \( \Psi = \Phi, \xi = (\theta_k/\tau_k) \), and \( \eta_k = \tau_k \), for all \( k \in \{1, 2, \ldots, n\} \), we obtain (54).

Definition 5 (See [27]). Assume that \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) is a probability distribution; then, the Shannon entropy is defined by

\[
S(s_3) = - \sum_{k=1}^{n} r_k \log (r_k). \tag{55}
\]

Corollary 6. Suppose that \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) is a probability distribution with \( \sum_{k=1}^{n} \tau_k = 1 \), \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_2 = (\tau_1, \tau_2, \ldots, \tau_n) \) are two \( n \)-tuples, and \( \tau_k, \theta_k, \theta_k > 0 (k = 1, 2, \ldots, n) \). If \( s_2 \) is a decreasing tuple and the conditions (52) and (53) are satisfied with \( \theta_k = 1 \) for all \( k \in \{1, 2, \ldots, n\} \), then

\[
S(s_3) \leq (1 - \ell)S(s_3) + \ell \sum_{k=1}^{n} r_k \log (\theta_k) \leq (1 - \ell)\log(n) + \ell \sum_{k=1}^{n} r_k \log (\theta_k)
\]

\[
\leq \sum_{k=1}^{n} r_k \log ((1 - \ell)n + \ell \theta_k) \leq \log(n). \tag{56}
\]

Corollary 7. Suppose that \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) are the probability distributions with \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \). If \( (\theta_1, \theta_2, \ldots, \theta_n) \) is a decreasing tuple, such that (52) and (53) are satisfied, and \( \tau_k, \theta_k, \theta_k > 0 (k = 1, 2, \ldots, n) \), \( \ell \in [0, 1] \), then

\[
K_\phi(s_1, s_3) = \sum_{k=1}^{n} \theta_k \log \left( \frac{\theta_k}{\tau_k} \right). \tag{57}
\]
\[ K_d(s_1, s_3) \geq (1 - \ell)K_d(s_1, s_3) + \ell \sum_{k=1}^{n} \tau_k \theta_k \log(\theta_k) \geq \ell \sum_{k=1}^{n} \tau_k \theta_k \log(\theta_k) \]
\[ \geq \sum_{k=1}^{n} \tau_k ((1 - \ell) + \ell \theta_k) \log((1 - \ell) + \ell \theta_k). \]

**Proof.** Since the function \( \Phi(v) = v \log v \) is convex on \((0, \infty)\), applying Theorem 8 by selecting \( \Phi(v) = v \log v \), with the help of \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \) given by the probability distributions, we obtain (58).

**Definition 7** (See [29]). Let \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) be two probability distributions. Then, the variational distance is defined as
\[ V_d(s_1, s_3) = \sum_{k=1}^{n} |\theta_k - \tau_k|. \] (59)

By virtue of Theorem 8, choose the convex function \( \Phi(v) = |v - 1| \) and use the conditions of probability distributions \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \), and we get Corollary 8.

**Corollary 8.** Suppose that all conditions of Corollary 7 are satisfied. Then,
\[ V_d(s_1, s_3) \geq (1 - \ell)V_d(s_1, s_3) + \ell \sum_{k=1}^{n} \tau_k |\theta_k - 1| \]
\[ \geq \ell \sum_{k=1}^{n} \tau_k (\theta_k - 1). \] (60)

**Definition 8** (See [30]). Let \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) be two probability distributions. Then, the Jeffreys distance is defined by
\[ J_d(s_1, s_3) = \sum_{k=1}^{n} (\theta_k - \tau_k) \log \left( \frac{\theta_k}{\tau_k} \right). \] (61)

By mean of Theorem 8, take the convex function \( \Phi(v) = (v - 1) \log v \) and use the conditions of probability distributions \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \), and we get Corollary 9.

**Corollary 9.** Suppose that all conditions of Corollary 7 are satisfied. Then,
\[ J_d(s_1, s_3) \geq (1 - \ell)J_d(s_1, s_3) + \ell \sum_{k=1}^{n} \tau_k (\theta_k - 1) \log(\theta_k) \]
\[ \geq \ell \sum_{k=1}^{n} \tau_k (\theta_k - 1) \log(\theta_k) \]
\[ \geq \ell \sum_{k=1}^{n} \tau_k (\theta_k - 1) \log(1 + (\theta_k - 1)). \] (62)

**Definition 10** (See [31]). Let \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) be two probability distributions. Then, the Bhattacharyya coefficient is defined as
\[ B_d(s_1, s_3) = \sum_{k=1}^{n} \sqrt{\theta_k \tau_k}. \]

By the use of Theorem 8, choose the convex function \( \Phi(v) = -\sqrt{v}(v > 0) \) and utilize the conditions of probability distributions \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \), and we get Corollary 10.

**Corollary 10.** Suppose that all conditions of Corollary 7 are satisfied. Then,
\[ B_d(s_1, s_3) \leq (1 - \ell)B_d(s_1, s_3) + \ell \sum_{k=1}^{n} \tau_k \sqrt{\theta_k} \]
\[ \leq 1 + \ell \sum_{k=1}^{n} \tau_k \sqrt{\theta_k - 1} \]
\[ \leq \sum_{k=1}^{n} \tau_k \sqrt{1 + (\theta_k - 1) - 1} \leq 1. \] (64)

**Definition 9** (See [32]). Let \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_3 = (\tau_1, \tau_2, \ldots, \tau_n) \) be two probability distributions. Then, the Hellinger distance is defined as
\[ H_d(s_1, s_3) = \sum_{k=1}^{n} \left( \sqrt{\theta_k} - \sqrt{\tau_k} \right)^2. \] (65)

By means of Theorem 8, take the convex function \( \Phi(v) = (\sqrt{v} - 1)^2 \) and use the conditions of probability distributions \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \), and we get Corollary 11.

**Corollary 11.** Suppose that all conditions of Corollary 7 are satisfied. Then,
\[ H_d(s_1, s_3) \geq (1 - \ell)H_d(s_1, s_3) + \ell \sum_{k=1}^{n} \tau_k \left( \sqrt{\theta_k} - 1 \right)^2 \]
\[ \geq \ell \sum_{k=1}^{n} \tau_k \left( \sqrt{\theta_k - 1} \right)^2 \]
\[ \geq \sum_{k=1}^{n} \tau_k \left( 1 + (\theta_k - 1) - 1 \right)^2. \] (66)

**Definition 11** (See [33]). Let \( s_1 = (\theta_1, \theta_2, \ldots, \theta_n) \) and \( s_2 = (\tau_1, \tau_2, \ldots, \tau_n) \) be two probability distributions. Then, the triangular discrimination is defined by
Given in this study would provide a significant supplement to the theory of mathematical inequalities.

\[
T_d(s_1, s_3) = \sum_{k=1}^{n} \frac{(\theta_k - \tau_k)^2}{\theta_k + \tau_k}.
\]  

(67)

By employing Theorem 8, choose the convex function \( \Phi(v) = (v-1)^2/v+1 \) and utilize the conditions of probability distributions \( \sum_{k=1}^{n} \theta_k = 1 \) and \( \sum_{k=1}^{n} \tau_k = 1 \), and we get Corollary 12.

**Corollary 12.** Suppose that all conditions of Corollary 7 are satisfied. Then,

\[
T_d(s_1, s_3) \geq (1-\ell)T_d(s_1, s_3) + \ell \sum_{k=1}^{n} \frac{(\theta_k - 1)^2}{\theta_k + 1}
\]

\[
\geq \ell \sum_{k=1}^{n} \frac{(\theta_k - 1)^2}{\theta_k + 1} \]

\[
\geq \ell \sum_{k=1}^{n} \frac{(\theta_k - 1)^2}{\theta_k + 1} + 2.
\]

**5. Conclusion**

The Jensen inequality is one of the most powerful and attractive inequalities among the mathematical inequalities. This inequality generalized the definition of convex function and generated many new inequalities. Nowadays, it became an engaging source for many researchers. The Jensen inequality has been refined, extended, improved, and generalized for convex functions in various directions. Also, it is used for developing some generalized convex functions such as s-convex, \( \eta \)-convex, strongly convex, and coordinate convex. Aside from the mathematics, the Jensen inequality has performed a key role for solving serveral problems in multiform fields of science such as economics, statistics, medical science, and physics. In this study, we worked on the refinements of Jensen inequality. By using convex functions and certain types of tuples with majorization conditions, we obtained some refinements of Jensen inequality. Furthermore, we established some bounds for the power mean as well as quasiarithmetic means with the help of the obtained results. We also gave bounds for the Csiszár divergence, Kullback–Leibler divergence, and Shannon entropy. Moreover, we obtained bounds for various distances associated with probability distributions, Bhattacharyya coefficient, and triangular discrimination. The results given in this study would provide a significant supplement to the theory of mathematical inequalities.

**Data Availability**

No data were used to support this study.

**Conflicts of Interest**

The authors declare that there are no conflicts of interest.

**Authors’ Contributions**

All authors contributed equally and significantly in this study. All authors read and approved the final manuscript.

**Acknowledgments**

This work was supported by the Natural Science Foundation of Fujian Province of China (2020J01365).

**References**

[1] M. J. Cloud, B. C. Drachman, and L. P. Lebedev, *Inequalities with Applications to Engineering*, Springer, Heidelberg, Germany, 2014.

[2] J. G. Liao and A. Berg, “Sharpening jensen’s inequality,” *The American Statistician*, vol. 73, no. 3, pp. 278–281, 2019.

[3] H. Ullah, M. Adil Khan, and J. Pečarić, “New bounds for soft margin estimator via concavity of Gaussian weighting function,” *Advances in Difference Equations*, vol. 2020, Article ID 2003103, 10 pages, 2020.

[4] M. Adil Khan, M. Hanif, Z. A. Khan, K. Ahmad, and Y.-M. Chu, “Association of Jensen’s inequality for s-convex function with Csiszár divergence,” *Journal of Inequalities and Applications*, vol. 2019, Article ID 19-2112-9, 14 pages, 2019.

[5] M. Adil Khan, Z. M. Al-Sahwi, and Y.-M. Chu, “New estimations for Shannon and zipf-mandelbrot entropies,” *Entropy*, vol. 20, no. 608, pp. 1–10, 2018.

[6] M. Adil Khan, D. Pečarić, and J. Pečarić, “On Zipf-mandelbrot entropy,” *Journal of Computational and Applied Mathematics*, vol. 346, pp. 192–204, 2019.

[7] M. Adil Khan, D. Pečarić, and J. Pečarić, “Bounds for Shannon and zipf-mandelbrot entropies,” *Mathematical Methods in the Applied Sciences*, vol. 40, no. 18, pp. 7316–7322, 2017.

[8] K. Ahmad, M. Adil Khan, S. Khan, A. Ali, and Y.-M. Chu, “New estimation of Zipf-Mandelbrot and Shannon entropies via refinements of Jensen’s inequality,” *AIP Advances*, vol. 11, pp. 1–10, Article ID 015147, 2021.

[9] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via theEMAlgorithm,” *Journal of the Royal Statistical Society: Series B*, vol. 39, no. 1, pp. 1–22, 1977.

[10] S. S. Dragomir, “Bounds for the normalised Jensen function,” *Bulletin of the Australian Mathematical Society*, vol. 74, no. 3, pp. 471–478, 2006.

[11] J. Pečarić and J. Perić, “New improvement of the converse Jensen inequality,” *Mathematical Inequalities & Applications*, vol. 21, no. 1, pp. 217–234, 2018.

[12] Z.-G. Xiao, H. M. Srivastava, and Z.-H. Zhang, “Further refinements of the Jensen inequalities based upon samples with repetitions,” *Mathematical and Computer Modelling*, vol. 51, no. 5–6, pp. 592–600, 2010.

[13] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, NY, USA, 1979.

[14] T. Ando, “Majorizations and inequalities in matrix theory,” *Linear Algebra and its Applications*, vol. 199, pp. 17–67, 1994.

[15] M. Adil Khan, N. Latif, and J. Pečarić, “Generalization of majorization theorem,” *Journal of Mathematical Inequalities*, vol. 9, no. 3, pp. 847–872, 2015.

[16] N. Latif, D. Pečarić, and J. Pečarić, “Majorization, “useful” csiszár divergence and “useful” zipf-mandelbrot law,” *Open Mathematics*, vol. 16, no. 1, pp. 1357–1373, 2018.
[17] M. Sababheh, "Improved Jensen’s inequality," *Mathematical Inequalities and Applications*, vol. 20, no. 2, pp. 389–403, 2017.

[18] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 379-423, pp. 623–656, 1948.

[19] N. Latif, J. Pečarić, and I. Perić, "On discrete favard’s and berwald’s inequalities," *Communications in Mathematical Analysis*, vol. 12, no. 2, pp. 34–57, 2012.

[20] C. P. Niculescu and L.-E. Persson, *Convex Functions and Their Applications*, Springer, New York, NY, USA, 2006.

[21] S. S. Dragomir, "Some majorisation type discrete inequalities for convex functions," *Mathematical Inequalities and Applications*, vol. 7, no. 2, pp. 207–216, 2004.

[22] M. Adil Khan, S.-H. Wu, H. Ullah, and Y.-M. Chu, "Discrete majorization type inequalities for convex functions on rectangles," *Journal of Inequalities and Applications*, vol. 2019, Article ID 1, 18 pages, 2019.

[23] S. Wu, M. Adil Khan, and H. U. Haleemzai, "Refinements of majorization inequality involving convex functions via taylor’s theorem with mean value form of the remainder," *Mathematics*, vol. 7, no. 8, p. 663, 2019.

[24] S. Z. Ullah, M. Adil Khan, and Y.-M. Chu, "Majorization theorem for strongly convex function," *Journal of Inequalities and Applications*, vol. 2019, Article ID 1920079, 13 pages, 2019.

[25] H. R. Moradi and S. Furuichi, "Improvement and generalization of some Jensen-Mercer-type inequalities," *Journal of Mathematical Inequalities*, vol. 14, no. 2, pp. 377–383, 2020.

[26] I. Csiszár, "Information–type measures of difference of probability functions and indirect observations," *Studia Scientiarum Mathematicarum Hungarica*, vol. 2, pp. 299–318, 1967.

[27] C. E. Shannon, "Prediction and entropy of printed English," *Bell System Technical Journal*, vol. 30, no. 1, pp. 50–64, 1951.

[28] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[29] S. S. Dragomir, "Some inequalities for (m, M)-Convex mappings and applications for the Csiszár φ-divergence in information theory," *Mathematical Journal of Ibaraki University*, vol. 33, pp. 35–50, 2001.

[30] M. Adil Khan, D. Pečarić, and I. Pečarić, "A new refinement of the Jensen inequality with applications in information theory," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 44, no. 1, pp. 267–278, 2021.

[31] A. Bhattacharyya, "Some analogues to the amount of information and their uses in statistical estimation," *Sankhya*, vol. 8, pp. 1–14, 1946.

[32] E. Hellinger, "Neue Begründung der Theorie quadratischer formen von unendlichvielen veränderlichen," *Journal für die Reine und Angewandte Mathematik*, vol. 1909, no. 136, pp. 210–271, 1909.

[33] I. J. Taneja, "Bounds on triangular discrimination, harmonic mean and symmetric chi-square divergences," *Journal of Computational and Applied Mathematics*, vol. 4, pp. 91–111, 2006.