mathematics

Article

Estimations of the Jensen Gap and Their Applications Based on 6-Convexity

Muhammad Adil Khan [1,*], Asadullah Sohail [1], Hidayat Ullah [1] and Tareq Saeed [2]

[1] Department of Mathematics, University of Peshawar, Peshawar 25000, Pakistan; asadstd@uop.edu.pk (A.S.)
[2] Financial Mathematics and Actuarial Science (FMAS)-Research Group, Department of Mathematics, Faculty of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia
* Correspondence: madilkhan@uop.edu.pk

Abstract: The main purpose of this manuscript is to present some new estimations of the Jensen gap in a discrete sense along with their applications. The proposed estimations for the Jensen gap are provided with the help of the notion of 6-convex functions. Some numerical experiments are performed to determine the significance and correctness of the intended estimates. Several outcomes of the main results are discussed for the Hölder inequality and the power and quasi-arithmetic means. Furthermore, some applications are presented in information theory, which provide some bounds for the divergences, Bhattacharyya coefficient, Shannon entropy, and Zipf–Mandelbrot entropy.

Keywords: convex function; 6-convex function; Jensen's inequality; Hölder inequality; power means; quasi-arithmetic means; information theory; Zipf–Mandelbrot entropy

MSC: 26A51; 26D15; 68P30

check for updates

Citation: Adil Khan, M.; Sohail, A.; Ullah, H.; Saeed, T. Estimations of the Jensen Gap and Their Applications Based on 6-Convexity. *Mathematics* **2023**, *11*, 1957. https://doi.org/10.3390/math11081957

Academic Editor: Xiangmin Jiao

Received: 23 March 2023
Revised: 11 April 2023
Accepted: 18 April 2023
Published: 20 April 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

There is no doubt that many real-life problems can be captured and modeled with the help of functions, and due to this fact, functions are regarded as the most fundamental and significant concepts in the fields of applied and pure mathematics as well as in other areas of science and engineering [1–3]. Convex functions are regarded as an intriguing class of functions [4–6]. Convex functions have many attractive and important properties, and due to these properties and their characteristics, convex functions play a leading role in the solutions to many complicated problems [7–9]. Moreover, convex functions are also popular because they deal with problems very smoothly [10–12]. Due to this, convex functions have attracted the attention of many researchers [13–15]. A convex function is defined as follows [5]:

A function $\Phi : [\sigma_1, \sigma_2] \to$ is called convex, if the relation

$$\Phi\big(\delta\varrho_1 + (1-\delta)\varrho_2\big) \le \delta\Phi(\varrho_1) + (1-\delta)\Phi(\varrho_2) \tag{1}$$

is valid, for all $\varrho_1, \varrho_2 \in [\sigma_1, \sigma_2]$ and $\delta \in [0,1]$. If (1) is accurate in the opposite sense, then Φ is called concave over $[\sigma_1, \sigma_2]$.

Due to the extensive applications and great importance of convex functions, this class has been generalized in various ways with the help of its characteristics and properties [16–18]. Some of the generalizations of convex functions are *P*-convex [19], *s*-convex [20], coordinates convex [21], *h*-convex [22], and 4-convex functions [23], etc. Among the above classes of convex functions, one of the significant classes is 6-convex functions. To give the classical definition of 6-convex, first, we explain the divided difference.

Recursively, one can define the divided difference $[\varrho_1, \varrho_2, \ldots, \varrho_n]\Phi$ of $\Phi(\varrho)$ at the distinct points $\varrho_1, \varrho_2, \ldots, \varrho_n$ as [5] (p. 24):

$$[\varrho_1]\Phi = \Phi(\varrho_1),$$

$$[\varrho_1, \varrho_2, \ldots, \varrho_n]\Phi = \frac{[\varrho_2, \varrho_2, \ldots, \varrho_n]\Phi - [\varrho_1, \varrho_2, \ldots, \varrho_{n-1}]\Phi}{\varrho_n - \varrho_1}. \tag{2}$$

The definition of an n-convex function is given as follows [5] (p. 23):
A function $\Phi : [\sigma_1, \sigma_2] \to \mathbb{R}$ is called n-convex, if for the distinct points $\varrho_1, \varrho_2, \ldots, \varrho_{n+1} \in [\sigma_1, \sigma_2]$, we have

$$[\varrho_1, \varrho_2, \ldots, \varrho_{n+1}]\Phi \geq 0. \tag{3}$$

If (3) is true in the reverse sense, then Φ is called n-concave.
The next theorem presents a criteria for a function to be n-convex [5] (p. 16).

Theorem 1. *Suppose that $\Phi : [\sigma_1, \sigma_2] \to \mathbb{R}$ is any function such that Φ^n (n^{th} derivatives of the function Φ) exists; then, Φ is n-convex, if and only if*

$$\Phi^n(\varrho) \geq 0, \qquad \varrho \in [\sigma_1, \sigma_2]. \tag{4}$$

If (4) reverses, then Φ is called n-concave.

Definition 1. *A function $\Phi : [\sigma_1, \sigma_2] \to \mathbb{R}$ is called 6-convex, if*

$$[\varrho_1, \varrho_2, \ldots, \varrho_7]\Phi \geq 0. \tag{5}$$

If (5) is valid in the reverse sense, then Φ is called 6-concave.

Remark 1. *Using Theorem 1, we can say that if Φ is a function such that Φ'''' exists, then Φ is 6-convex, if Φ'''' is convex.*

In recent years, the extensive applications of convex functions and their generalizations have been observed in the field of mathematical inequalities [24–27]. There are many inequalities, which it would not be possible to establish without convex functions [28–30]. Some of the interesting inequalities that have been acquired by convex functions are the majorization inequality [21], Slater's inequality [31], Hermite–Hadamard's inequality [7], Jensen–Mercer's [32] inequality, and many more [33–35]. It has been confirmed that convex functions have played an important role in the development of the field of mathematical inequalities [36–39]. Jensen's inequality is the most dynamic and important inequality in the current literature on mathematical inequalities for convex functions [40]. This inequality provides a very interesting relation between finite sequences and convex functions [40]. Furthermore, Jensen's inequality also provides criteria for a function to be convex in the given domain [41]. This inequality states that if $\varrho_\varsigma \in [\sigma_1, \sigma_2]$ and $\delta_\varsigma \geq 0$, for all $\varsigma \in \{1, 2, \ldots, m\}$ with $\delta^* = \sum\limits_{\varsigma=1}^{n} \delta_\varsigma > 0$, then the inequality

$$\Phi\left(\frac{1}{\delta^*}\sum_{\varsigma=1}^{m} \delta_\varsigma \varrho_\varsigma\right) \leq \frac{1}{\delta^*}\sum_{\varsigma=1}^{m} \delta_\varsigma \Phi(\varrho_\varsigma) \tag{6}$$

is valid for each convex function $\Phi : [\sigma_1, \sigma_2] \to \mathbb{R}$. If Φ is concave, then (6) is true in the opposite direction.

The continuous form of the Jensen inequality given in (6) can be stated as follows. Assume that $T_1, T_2 : [\sigma_1, \sigma_2] \to [c, d]$ are any integrable functions, with $T_1(\delta) \geq 0$ for all $\delta \in [\sigma_1, \sigma_2]$, and $\mathrm{T}_1 := \int_{\sigma_1}^{\sigma_2} T_1(\delta) d\delta > 0$. If $\Phi \circ T_2$ is an integrable function, then

$$\Phi\left(\frac{1}{\mathrm{T}_1}\int_{\sigma_1}^{\sigma_2} T_1(\delta) T_2(\delta) d\delta\right) \leq \frac{1}{\mathrm{T}_1}\int_{\sigma_1}^{\sigma_2} T_1(\delta) \Phi(T_2(\delta)) d\delta, \tag{7}$$

for each convex function $\Phi : [c, d] \to \mathbb{R}$. If Φ is concave, then (7) reverses.

The most interesting feature that is hidden in the structure of Jensen's inequality is that it enhances the concept of convexity [42]. Another dominant feature of the Jensen inequality is that there are many famous inequalities that are the direct outcomes of this inequality, namely Hölder's inequality, the Hermite–Hadamard inequality, Ky Fan's inequality, Young's inequality, and several others [40,41]. In the last three decades, the Jensen inequality and its applications have recorded significant growth with important consequences in several areas of science including engineering [43], optimization [44], statistics [45], finance [46], epidemiology [47], information theory [48], etc. Moreover, due to its good behavior and essential properties, Jensen's inequality has been expanded, polished, and enhanced in diverse directions [49–51]. Furthermore, many interesting concepts have also been developed with the help of Jensen's inequality related to ordinary convexity and its generalizations [52–54]. Moreover, many significant inequalities for convex functions have been generalized and improved with the support of Jensen's inequality [55,56]. Bakula and Nikodem [52] utilized the concepts of strong convexity and mild convexity and established counterparts to Jensen's inequality. Kian [57] provided an operator form of the Jensen inequality for the super quadratic functions and also discussed the applications of the acquired results. Matković et al. [58] demonstrated a variant of Jensen's inequality for operators by utilizing convexity, which is the generalization of Mercer's result. In 2021, Deng et al. [59] presented refinements of the discrete Jensen inequality by utilizing some results of majorization. Moreover, they discussed the applications of their refined Jensen's inequality for the quasi-arithmetic means, power means, and information theory. In 2022, Ullah et al. [40] provided Jensen inequality improvements in both continuous and discrete forms via 4-convexity. They acquired some bounds for the Csiszár divergence and its related forms. Additionally, they provided improvements of the Hölder and Hermite–Hadamard inequalities as the direct consequences of the main results. Moreover, they compared their main results with some other earlier established improvements of Jensen's inequality and showed that the obtained improvements provided better estimations for the Jensen difference. Saeed et al. [60] provided refinements of the celebrated integral Jensen inequality via the theory of majorization. With the help of these, refinements of the Hölder and Hermite–Hadamard inequalities were provided. Moreover, they discussed some consequences of the refined Jensen inequality for means. Furthermore, they also presented applications for information theory. Adil Khan et al. [61] presented an estimation for the Jensen gap through the notion of convexity and also discussed its applications for some classical inequalities as well as in information theory. Furthermore, they compared their results with some earlier proven results of similar type and concluded that the obtained estimations provided better estimates as compared to other results. Costarelli and Spigler [62] developed several estimations pertaining to the Jensen inequality with the support of convex functions. The usefulness of the acquired estimates were discussed in modular estimates in Orlicz spaces or L^p–estimates for linear and nonlinear integral operators. In 2008, Zhu and Yang [63] applied Jensen's inequality and discussed the stability of discrete-time delay systems.

Providing new estimates for the Jensen gap is the main focus of this research. The article is structured as follows:

- Section 2 provide estimates for the Jensen gap.
- Section 3 gives numerical estimations for the Jensen gap and their comparisons with other results.
- Section 4 provides applications for the Hölder inequality.
- Section 5 presents applications for the power and quasi-arithmetic means.
- Section 6 provides applications for information theory.
- Section 7 gives applications for the Zipf–Mandelbrot entropy.

2. Main Results

This section concerns the estimations of the Jensen gap. The proposed estimations are based on 6-convexity. For the achievement of the intended estimations, we utilize Jensen's inequality and 6-convexity. We begin with the following theorem, in which we apply the definition of a 6-convex function to derive an upper bound for the Jensen gap.

Theorem 2. *Let $\Phi : [\sigma_1, \sigma_2] \to \mathbb{R}$ be a fourth time differentiable function, $\varrho_\varsigma \in [\sigma_1, \sigma_2]$, $\delta_\varsigma \geq 0$, for all $\varsigma \in \{1, 2, \ldots, m\}$ with $\delta^* := \sum_{\varsigma=1}^{m} \delta_\varsigma > 0$ and $\sigma^* := \frac{1}{\delta^*} \sum_{\varsigma=1}^{m} \delta_\varsigma \varrho_\varsigma$. If Φ is 6-convex, then*

$$\begin{aligned}\frac{1}{\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma\Phi(\varrho_\varsigma)-\Phi(\sigma^*)&\leq\frac{1}{120\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\\&\times\Big(4\Phi''''(\sigma^*)+\Phi''''(\varrho_\varsigma)\Big)-\frac{1}{6\delta^*}\Phi'''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3\\&+\frac{1}{2\delta^*}\Phi''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2.\end{aligned}\tag{8}$$

The inequality (8) reverses for the concave function Φ.

Proof. Without loss of generality, assume that $\sigma^* \neq \varrho_\varsigma$, for every $\varsigma \in \{1, 2, \ldots, m\}$. Applying integration by parts, we obtain

$$\begin{aligned}&\frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\int_0^1 t^3\Phi''''(t\sigma^*+(1-t)\varrho_\varsigma)dt\\
=&\ \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\left(t^3\frac{\Phi'''(t\sigma^*+(1-t)\varrho_\varsigma)\Big|_0^1}{\sigma^*-\varrho_\varsigma}-\frac{\int_0^1 3t^2\Phi'''(t\sigma^*+(1-t)\varrho_\varsigma)dt}{\sigma^*-\varrho_\varsigma}\right)\\
=&\ \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\left(\frac{\Phi'''(\sigma^*)}{\sigma^*-\varrho_\varsigma}-\frac{3}{\sigma^*-\varrho_\varsigma}\left(t^2\frac{\Phi''(t\sigma^*+(1-t)\varrho_\varsigma)\Big|_0^1}{\sigma^*-\varrho_\varsigma}\right.\right.\\
&\left.\left.-\frac{2}{\sigma^*-\varrho_\varsigma}\int_0^1 t\Phi''(t\sigma^*+(1-t)\varrho_\varsigma)dt\right)\right)\\
=&\ \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\left(\frac{\Phi'''(\sigma^*)}{\sigma^*-\varrho_\varsigma}-3\frac{\Phi''(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^2}+\frac{6}{(\sigma^*-\varrho_\varsigma)^2}\int_0^1 t\Phi''(t\sigma^*+(1-t)\varrho_\varsigma)dt\right)\\
=&\ \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\left(\frac{\Phi'''(\sigma^*)}{\sigma^*-\varrho_\varsigma}-3\frac{\Phi''(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^2}\right.\\
&\left.+\frac{6}{(\sigma^*-\varrho_\varsigma)^2}\left(t\frac{\Phi'(t\sigma^*+(1-t)\varrho_\varsigma)\Big|_0^1}{\sigma^*-\varrho_\varsigma}-\frac{\int_0^1\Phi'(t\sigma^*+(1-t)\varrho_\varsigma)dt}{\sigma^*-\varrho_\varsigma}\right)\right)\\
=&\ \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\left(\frac{\Phi'''(\sigma^*)}{\sigma^*-\varrho_\varsigma}-3\frac{\Phi''(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^2}\right.\end{aligned}$$

$$
\begin{aligned}
& +6\frac{\Phi^{'}(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^3}-\frac{6}{(\sigma^*-\varrho_\varsigma)^3}\int_0^1\Phi^{'}(t\sigma^*+(1-t)\varrho_\varsigma)dt \\
= & \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\bigg(\frac{\Phi^{'''}(\sigma^*)}{\sigma^*-\varrho_\varsigma}-3\frac{\Phi^{''}(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^2} \\
& +6\frac{\Phi^{'}(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^3}-\frac{6}{(\sigma^*-\varrho_\varsigma)^4}\Phi(t\sigma^*+(1-t)\varrho_\varsigma)\Big|_0^1\bigg) \\
= & \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\bigg(\frac{\Phi^{'''}(\sigma^*)}{\sigma^*-\varrho_\varsigma}-3\frac{\Phi^{''}(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^2}+6\frac{\Phi^{'}(\sigma^*)}{(\sigma^*-\varrho_\varsigma)^3} \\
& -\frac{6}{(\sigma^*-\varrho_\varsigma)^4}(\Phi(\sigma^*)-\Phi(\varrho_\varsigma))\bigg) \\
= & \frac{1}{6\delta^*}\Phi^{'''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3-\frac{1}{2\delta^*}\Phi^{''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
& +\frac{1}{\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma\Phi(\varrho_\varsigma)-\Phi(\sigma^*).
\end{aligned}
$$

This implies that

$$
\begin{aligned}
\frac{1}{\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma\Phi(\varrho_\varsigma)-\Phi(\sigma^*) = & \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\int_0^1 t^3\Phi^{''''}(t\sigma^*+(1-t)\varrho_\varsigma)dt \\
& -\frac{1}{6\delta^*}\Phi^{'''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi^{''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2. \qquad (9)
\end{aligned}
$$

Now, utilizing the 6-convexity of Φ on the right side of (9), we gain

$$
\begin{aligned}
\frac{1}{\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma\Phi(\varrho_\varsigma)-\Phi(\sigma^*) \leq & \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\int_0^1 t^3(t\Phi^{''''}(\sigma^*)+(1-t)\Phi^{''''}(\varrho_\varsigma))dt \\
& -\frac{1}{6\delta^*}\Phi^{'''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi^{''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
= & \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\bigg(\Phi^{''''}(\sigma^*)\int_0^1 t^4dt+\Phi^{''''}(\varrho_\varsigma)\int_0^1(t^3-t^4)dt\bigg) \\
& -\frac{1}{6\delta^*}\Phi^{'''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi^{''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
= & \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\bigg(\Phi^{''''}(\sigma^*)\frac{1}{5}+\Phi^{''''}(\varrho_\varsigma)\frac{1}{20}\bigg) \\
& -\frac{1}{6\delta^*}\Phi^{'''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi^{''}(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2. \qquad (10)
\end{aligned}
$$

Instantly, by evaluating integrals in inequality (10), we arrive at inequality (8). □

The following theorem affords a lower bound for the Jensen gap, which can be acquired with the help of the integral version of Jensen's inequality.

Theorem 3. *Let the conditions of Theorem 2 be valid. Then,*

$$\frac{1}{\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma\Phi(\varrho_\varsigma)-\Phi(\sigma^*)\geq\frac{1}{24\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\Phi''''\left(\frac{4\sigma^*+\varrho_\varsigma}{5}\right)$$
$$-\frac{\Phi'''(\sigma^*)}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{\Phi''(\sigma^*)}{2\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2. \tag{11}$$

For the 6-concave function Φ, the inequality (11) changes its direction.

Proof. By utilizing the Jensen inequality for (9), we obtain

$$\begin{aligned}
\frac{1}{\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma\Phi(\varrho_\varsigma)-\Phi(\sigma^*) &= \frac{1}{6\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\int_0^1 t^3\Phi''''(t\sigma^*+(1-t)\varrho_\varsigma)dt \\
&\quad -\frac{1}{6\delta^*}\Phi'''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
&= \frac{1}{24\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\left(\frac{\int_0^1 t^3\Phi''''(t\sigma^*+(1-t)\varrho_\varsigma)dt}{\int_0^1 t^3dt}\right) \\
&\quad -\frac{1}{6\delta^*}\Phi'''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
&\geq \frac{1}{24\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\Phi''''\left(\frac{\int_0^1 t^3(t\sigma^*+(1-t)\varrho_\varsigma)dt}{\int_0^1 t^3dt}\right) \\
&\quad -\frac{1}{6\delta^*}\Phi'''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
&= \frac{1}{24\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\Phi''''\left(\frac{\sigma^*\int_0^1 t^4dt+\varrho_\varsigma\int_0^1(t^3-t^4)dt}{\int_0^1 t^3dt}\right) \\
&\quad -\frac{1}{6\delta^*}\Phi'''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2 \\
&= \frac{1}{24\delta^*}\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^4\Phi''''\left(\frac{4\sigma^*+\varrho_\varsigma}{5}\right) \\
&\quad -\frac{1}{6\delta^*}\Phi'''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^3+\frac{1}{2\delta^*}\Phi''(\sigma^*)\sum_{\varsigma=1}^{m}\delta_\varsigma(\sigma^*-\varrho_\varsigma)^2.
\end{aligned} \tag{12}$$

Clearly, (12) is equivalent to (11). □

Remark 2. *There are many functions in the literature of convex analysis, which are not 2-convex, 3-convex, 4-convex, and 5-convex; however, they are 6-convex. Moreover, there exist many interesting results, which have been proved for the classes of 2-convex, 3-convex, 4-convex, and 5-convex functions. The existing results for the classes of 2-convex, 3-convex, 4-convex, and 5-convex functions are not applicable for the 6-convex functions. Therefore, our results are important due to the fact that they make the class of 6-convex functions dynamic and applicable. Following are examples of 6-convex functions, which are not 2-convex, 3-convex, 4-convex, or 5-convex:*

- $\Phi(\delta)=-10\varrho^5+\exp\varrho,\qquad \varrho\in[0,10].$
- $\Psi(\delta)=-\varrho^5-\log\varrho,\qquad \varrho\in(0,1].$
- $f(\delta)=-\varrho^5-\sqrt{\varrho},\qquad \varrho\in[0,21].$

3. Numerical Estimations for the Jensen Gap

This section presents the numerical values of the obtained estimations for the Jensen gap. We apply the main result for a particular 6-convex function while replacing the arbitrary tuples with particular tuples and obtain specific estimations to show the importance of our results, as compared to some other earlier established results. We commence this part with the following example, which provides a numerical estimation of the Jensen gap and its comparison with other results.

Example 1. *Let $\boldsymbol{p} = (1,1,1,1)$ and $\boldsymbol{x} = (-2,-1,0,1)$ be two tuples, and $\Phi(\delta) = \exp\delta$, $\delta \in (-\infty,\infty)$. Then, $\Phi''(\delta) = \exp\delta$, $\Phi''''(\delta) = \exp\delta$, and $\Phi''''''(\delta) = \exp\delta$. Clearly, Φ'', Φ'''', and Φ'''''' all are positive on $(-\infty,\infty)$, which admit the convexity, 4-convexity, and 6-convexity of $\Phi(\delta) = \exp\delta$ on $(-\infty,\infty)$. Therefore, we apply inequality (8) for the above tuples, and for the given 6-convex function, we acquire*

$$0.4489 < 0.4612. \tag{13}$$

Now, using the above values and the specific function in inequality (2) given in the article [40], we deduce

$$0.4489 < 0.5345. \tag{14}$$

From inequalities (13) and (14), it is obvious that inequality (8) gives better estimates than inequality (2) in [40].

Example 2. *Consider the tuples $\boldsymbol{p} = (1,1,1,1,1)$ and $\delta = (-0.25,-0.5,0,0.25,0.5)$ and $\Phi(\delta) = \delta^6$ defined on $(-\infty,\infty)$. Then, $\Phi''(\delta) = 30\delta^4$, $\Phi''''(\delta) = 360\delta^2$, and $\Phi''''''(\delta) = 720$; clearly, $\Phi''(\delta) = 30\delta^4$, $\Phi''''(\delta) = 360\delta^2$, and $\Phi''''''(\delta) = 720$ are nonnegative on $(-\infty,\infty)$. Thus, this confirms the convexity, 4-convexity, and 6-convexity of the given function. Therefore, using this information in inequality (8), we obtain*

$$0.0063 < 0.0190. \tag{15}$$

Now, utilizing inequality (2) in [40] for the above information, we obtain

$$0.0063 < 0.0317. \tag{16}$$

From inequalities (15) and (16), we conclude that the estimates for the Jensen gap given in (11) provide better estimates than the estimates mentioned in inequality (2) in [40].

Remark 3. *The authors compared the estimates for the Jensen gap given in inequality (2) in [40] to the estimates specified in inequalities (5) and (8) in article [62] and the estimates given in inequality (4) of article [64] and declare that the estimates in (2) of [40] are superior to the estimates (5) and (8) given in [62] and the estimate (4) in [64]. Hence, from this discussion, we can say that our estimates for the Jensen gap may provide better estimates than the estimates (5) and (8) described in [62] and the estimate (4) in [64].*

4. Applications for the Hölder Inequality

The Hölder inequality is thought to be the most advantageous inequality due to its significance and wide range of uses. By using our main results, we give some relations of the Hölder inequality. The next proposition presents a relation for the Hölder inequality as a consequence of Theorem 2.

Proposition 1. *We assume that $a_\varsigma, b_\varsigma > 0$, for $\varsigma = 1,2,\ldots,m$ and $q,p > 1$, such that $\frac{1}{p} + \frac{1}{q} = 1$.*

(i) If $p \in (1,2] \cup [3,4] \cup [5,\infty)$, then

$$\begin{aligned}
&\left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{\frac{1}{q}}\left(\sum_{\varsigma=1}^{m} b_\varsigma^p\right)^{\frac{1}{p}} - \sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma \leq \left(\frac{p(p-1)}{2}\right)^{\frac{1}{p}}\left[\frac{(p-2)(p-3)}{60}\right.\\
&\times \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^4 \times \left(4\left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{p-4} + \left(b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^{p-4}\right)\\
&-\frac{(p-2)}{3}\left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{p-3} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^3\\
&\left.+\left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{p-2} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^2\right]^{\frac{1}{p}} \left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{\frac{1}{q}}. \qquad (17)
\end{aligned}$$

(ii) *If* $p \in (2,3) \cup (4,5)$, *then* (17) *reverses.*

Proof. (i) We consider the function $\Phi(\delta) = \delta^p$ $x > 0$; then, by successive differentiation of the given function Φ, we obtain $\Phi''(\delta) = p(p-1)\delta^{p-2}$, and $\Phi''''''(\delta) = p(p-1)(p-2)(p-3)(p-4)(p-5)\delta^{p-6}$. Clearly, both Φ'' and Φ'''''' are positive on $(0,\infty)$ for $p \in (1,2] \cup [3,4] \cup [5,\infty)$, which substantiates the convexity as well as the 6-convexity of $\Phi(\delta) = \delta^p$. Therefore, utilizing (8) for $\Phi(\delta) = \delta^p$, $\delta_\varsigma = a_\varsigma^q$, $\varrho_\varsigma = b_\varsigma a_\varsigma^{-\frac{q}{p}}$, and then taking the power $\frac{1}{p}$, we obtain

$$\begin{aligned}
&\left(\left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{(p-1)}\left(\sum_{\varsigma=1}^{m} b_\varsigma^p\right) - \left(\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma\right)^p\right)^{\frac{1}{p}} \leq \left[\frac{p(p-1)(p-2)(p-3)}{120}\right.\\
&\times \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^4 \left(4\left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{p-4} + \left(b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^{p-4}\right)\\
&-\frac{p(p-1)(p-2)}{6}\left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{p-3} \times \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^3\\
&\left.+\frac{p(p-1)}{2}\left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{p-2} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - b_\varsigma a_\varsigma^{-\frac{q}{p}}\right)^2\right]^{\frac{1}{p}} \times \left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{\frac{1}{q}}. \qquad (18)
\end{aligned}$$

Since the inequality

$$\delta^r - \varrho^r \leq (\delta - \varrho)^r \qquad (19)$$

holds for all $\varrho, \delta \geq 0$ and $r \in [0,1]$, we take $\delta = \left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{(p-1)}\left(\sum_{\varsigma=1}^{m} b_\varsigma^p\right)$, $\varrho = \left(\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma\right)^p$, and $r = \frac{1}{p}$ in (19), and we receive

$$\left(\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}\right)^{\frac{1}{q}}\left(\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}\right)^{\frac{1}{p}}-\sum_{\varsigma=1}^{m} a_{\varsigma} b_{\varsigma}$$
$$\leq\left(\left(\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}\right)^{(p-1)}\left(\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}\right)-\left(\sum_{\varsigma=1}^{m} a_{\varsigma} b_{\varsigma}\right)^{p}\right)^{\frac{1}{p}}. \tag{20}$$

By comparing inequalities (18) and (20), we arrive at inequality (17).

(ii) Obviously, the function $\Phi(\delta)=\delta^{p}$ is 6-concave on $(0,\infty)$ for $p\in(2,3)\cup(4,5)$. Therefore, the reverse of (18) follows the technique utilized in (i).

□

The next corollary, which follows from Theorem 2, gives the relations for the Hölder inequality.

Corollary 1. *We presume that $a_{\varsigma}, b_{\varsigma}>0$ for $\varsigma=1,2,\ldots,m$ and $0<p<1$ with $q=\frac{p}{p-1}$.*

(i) *If $\frac{1}{p}\in(1,2]\cup[3,4]\cup[5,\infty)$, then*

$$\sum_{\varsigma=1}^{m} a_{\varsigma} b_{\varsigma}-\left(\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}\right)^{\frac{1}{q}}\left(\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}\right)^{\frac{1}{p}}$$
$$\leq \frac{(1-p)(1-2p)(1-3p)}{120p^{4}}\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}\left(\frac{\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}}{\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}}-a_{\varsigma}^{-q} b_{\varsigma}^{p}\right)^{4}\left(4\left(\frac{\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}}{\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}}\right)^{\frac{1}{p}-4}+\left(b_{\varsigma} a_{\varsigma}^{-\frac{q}{p}}\right)^{\frac{1}{p}-4}\right)$$
$$-\frac{(1-p)(1-2p)}{6p^{3}}\left(\frac{\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}}{\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}}\right)^{\frac{1}{p}-3}\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}\left(\frac{\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}}{\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}}-a_{\varsigma}^{-q} b_{\varsigma}^{p}\right)^{3}$$
$$+\frac{(1-p)}{2p^{2}}\left(\frac{\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}}{\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}}\right)^{\frac{1}{p}-2}\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}\left(\frac{\sum_{\varsigma=1}^{m} b_{\varsigma}^{p}}{\sum_{\varsigma=1}^{m} a_{\varsigma}^{q}}-a_{\varsigma}^{-q} b_{\varsigma}^{p}\right)^{2}. \tag{21}$$

(ii) *If $\frac{1}{p}\in(2,3)\cup(4,5)$, then (21) is positive in the opposite sense.*

Proof. (i) Let us take the function $\Phi(\delta)=\delta^{\frac{1}{p}}$ defined on $(0,\infty)$. Then, $\Phi''(\delta)=\frac{1}{p}\left(\frac{1}{p}-1\right)\delta^{\frac{1}{p}-2}$, $\Phi''''''(\delta)=\frac{1}{p}\left(\frac{1}{p}-1\right)\left(\frac{1}{p}-2\right)\left(\frac{1}{p}-3\right)\left(\frac{1}{p}-4\right)\left(\frac{1}{p}-5\right)\delta^{\frac{1}{p}-6}$, and obviously, $\Phi''>0$ and $\Phi''''''>0$. This confirms that the function $\Phi(\delta)=\delta^{\frac{1}{p}}$ is both convex and 6-convex on $(0,\infty)$ for $p\in(0,1)$, with $\frac{1}{p}\in(1,2]\cup[3,4]\cup[5,\infty)$. Therefore, applying inequality (8) for $\Phi(\delta)=\delta^{\frac{1}{p}}$, $\delta_{\varsigma}=a_{\varsigma}^{q}$ and $\varrho_{\varsigma}=a_{\varsigma}^{-q} b_{\varsigma}^{p}$, we obtain

$$\frac{\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - \frac{\left(\sum_{\varsigma=1}^{m} b_\varsigma^p\right)^{\frac{1}{p}}}{\left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{\frac{1}{p}}} \leq \frac{(1-p)(1-2p)(1-3p)}{120p^4 \sum_{\varsigma=1}^{m} a_\varsigma^q} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - a_\varsigma^{-q} b_\varsigma^p\right)^4$$
$$\times \left(4\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{\frac{1}{p}-4} + (b_\varsigma a_\varsigma^{-\frac{q}{p}})^{\frac{1}{p}-4}\right)$$
$$-\frac{(1-p)(1-2p)}{6p^3 \sum_{\varsigma=1}^{m} a_\varsigma^q}\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{\frac{1}{p}-3} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - a_\varsigma^{-q} b_\varsigma^p\right)^3$$
$$+\frac{(1-p)}{2p^2 \sum_{\varsigma=1}^{m} a_\varsigma^q}\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{\frac{1}{p}-2} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - a_\varsigma^{-q} b_\varsigma^p\right)^2. \tag{22}$$

To acquire (21), we multiply (22) by $\sum_{\varsigma=1}^{m} a_\varsigma^q$.

(*ii*) Clearly, $\Phi(\delta) = \delta^{\frac{1}{p}}$ is a 6-concave function with the given conditions. Therefore, to deduce the reverse inequality in (21), we follow the procedure of (*i*). □

The next corollary provides an application of Theorem 3 for the Hölder inequality.

Corollary 2. *We presume that the tuples $a_\varsigma, b_\varsigma > 0$ for $\varsigma = 1, 2, \ldots, m$ and $p \in (0,1)$ with $q = \frac{p}{p-1}$.*

(*i*) *If $\frac{1}{p} \in (1,2] \cup [3,4] \cup [5,\infty)$, then*

$$\sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma - \left(\sum_{\varsigma=1}^{m} a_\varsigma^q\right)^{\frac{1}{q}}\left(\sum_{\varsigma=1}^{m} b_\varsigma^p\right)^{\frac{1}{p}}$$
$$\geq \frac{(1-p)(1-2p)(1-3p)}{120p^4} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - a_\varsigma^{-q} b_\varsigma^p\right)^4$$
$$\times \left(4\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p + a_\varsigma^{-q} b_\varsigma^p \sum_{\varsigma=1}^{m} a_\varsigma^q}{5\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{\frac{1-4p}{p}}$$
$$-\frac{(1-p)(1-2p)}{6p^3}\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{\frac{1-3p}{p}} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - a_\varsigma^{-q} b_\varsigma^p\right)^3$$
$$+\frac{(1-p)}{2p^2}\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q}\right)^{\frac{1-2p}{p}} \sum_{\varsigma=1}^{m} a_\varsigma^q \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma^p}{\sum_{\varsigma=1}^{m} a_\varsigma^q} - a_\varsigma^{-q} b_\varsigma^p\right)^2. \tag{23}$$

(*ii*) *If* $\frac{1}{p} \in (2,3) \cup (4,5)$, *then* (23) *is is valid in the reverse path.*

Proof. Since the function $\Phi(\delta) = \delta^{\frac{1}{p}}$ is convex as well as 6-convex on $(0,\infty)$ for $p \in (0,1)$ such that $\frac{1}{p} \in (1,2] \cup [3,4] \cup [5,\infty)$, utilizing inequality (11) by setting $\Phi(\delta) = \delta^{\frac{1}{p}}$ and $\varrho_\varsigma = a_\varsigma^{-q} b_\varsigma^p$, we obtain (23).

(*ii*) For the aforementioned setting, $\Phi(\delta) = \delta^{\frac{1}{p}}$ is 6-concave. Consequently, if we pursue the mode (*i*), we acquire the converse of (23). □

5. Application for the Power and Quasi-Arithmetic Means

The concept of means is very important in the fields of pure and applied mathematics, statistics, economics, information theory, and many others, because they are general and unavoidable in their applications. Among the means, the power and quasi-arithmetic means are influential in the sense that they generalize the other classical means. This section of the paper provides some relations for the aforementioned means as consequences of the main results. First, we recall the definition of the power mean.

Definition 2. *Let* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *and* $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ *be two positive n-tuples with* $\overline{a} = \sum_{\varsigma=1}^{m} a_\varsigma$. *Then, the power mean of order* $\varrho \in \mathbb{R}$ *is defined by:*

$$\mathcal{P}_\varrho(\boldsymbol{a},\boldsymbol{b}) = \begin{cases} \left(\frac{1}{\overline{a}} \sum_{\varsigma=1}^{m} a_\varsigma b_\varsigma^\varrho\right)^{\frac{1}{\varrho}}, & \varrho \neq 0 \\ \left(\prod_{\varsigma=1}^{m} b_\varsigma^{a_\varsigma}\right)^{\frac{1}{\overline{a}}}, & \varrho = 0. \end{cases}$$

An application of Theorem 2 for the power mean is given in the following corollary.

Corollary 3. *Assume that* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *and* $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ *are n-tuples such that* $a_\varsigma, b_\varsigma > 0$, *for all* $\varsigma \in \{1, 2, \ldots, m\}$ *with* $\overline{a} = \sum_{\varsigma=1}^{m} a_\varsigma$. *Moreover, let* $\kappa, \tau \in \mathbb{R} \setminus \{0\}$.

(*i*) *If* $\tau > 0$ *with* $\tau \le \kappa \le 2\tau$ *or* $3\tau \le \kappa \le 4\tau$ *or* $\kappa \ge 5\tau$ *or* $\kappa<0$, *then*

$$\begin{aligned} \mathcal{P}_\kappa^\kappa(\boldsymbol{a},\boldsymbol{b}) - \mathcal{P}_\tau^\kappa(\boldsymbol{a},\boldsymbol{b}) \le{} & \frac{\kappa(\kappa-\tau)(\kappa-2\tau)(\kappa-3\tau)}{120\tau^4\overline{a}} \\ & \times \sum_{\varsigma=1}^{m} a_\varsigma \left(\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^\tau\right)^4 \left(\mathcal{P}_\tau^{\kappa-4\tau}(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^{\kappa-4\tau}\right) \\ & - \frac{\kappa(\kappa-\tau)(\kappa-2\tau)\mathcal{P}_\tau^{\kappa-3\tau}(\boldsymbol{a},\boldsymbol{b})}{6\tau^3\overline{a}} \sum_{\varsigma=1}^{m} a_\varsigma \left(\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^\tau\right)^3 \\ & + \frac{\kappa(\kappa-\tau)\mathcal{P}_\tau^{\kappa-2\tau}(\boldsymbol{a},\boldsymbol{b})}{2\tau^2\overline{a}} \sum_{\varsigma=1}^{m} a_\varsigma \left(\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^\tau\right)^2. \end{aligned} \tag{24}$$

(*ii*) *If* $\tau < 0$ *with* $2\tau \le \kappa \le \tau$ *or* $4\tau \le \kappa \le 3\tau$ *or* $\kappa \le 5\tau$ *or* $\kappa>0$, *then* (24) *holds.*

(*iii*) *If* $\tau > 0$ *with* $0 < \kappa < \tau$, $2\tau < \kappa < 3\tau$ *or* $4\tau < \kappa < 5\tau$, *then* (24) *is true in the reverse direction.*

(*iv*) *If* $\tau < 0$ *with* $\tau < \kappa < 0$, $3\tau < \kappa < 2\tau$ *or* $5\tau < \kappa < 4\tau$, *then the inequality in* (24) *is valid in the opposite sense.*

Proof. (*i*) We consider the function $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$, $\delta > 0$; then, by differentiating the given function with respect to δ, we acquire $\Phi''(\delta) = \frac{\kappa}{\tau}\left(\frac{\kappa}{\tau} - 1\right)\delta^{\frac{\kappa}{\tau}-2}$ and $\Phi''''''(\delta) = \frac{\kappa}{\tau}\left(\frac{\kappa}{\tau}-1\right)\left(\frac{\kappa}{\tau}-2\right)\left(\frac{\kappa}{\tau}-3\right)\left(\frac{\kappa}{\tau}-4\right)\left(\frac{\kappa}{\tau}-5\right)\delta^{\frac{\kappa}{\tau}-6}$. Clearly, both Φ'''''' and Φ'' are positive with the given conditions, and consequently, this confirms the convexity and 6-

convexity of the function $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$ on $(0,\infty)$. Therefore, taking $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = b_\varsigma^\tau$ in (8), we obtain (24).

(ii) For the aforementioned conditions on κ and τ, the function $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$ is 6-convex on $(0,\infty)$. Therefore, by following the procedure (i), we receive (24).

(iii) For the said values of κ and τ, the function $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$ is 6-concave on $(0,\infty)$. Therefore, the inequality (24) can easily be deduced by adopting the procedure of proof of (i).

(iv) Surely, the function $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$ is 6-concave on $(0,\infty)$ for the given values of κ and τ. Therefore, to acquire inequality (24), we proceed in the same way as in the proof of (i).

□

In the next corollary, we present a relation for the power means as an application of Theorem 3.

Corollary 4. *Suppose that $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ are positive n-tuples with $\bar{a} = \sum_{\varsigma=1}^{m} a_\varsigma$ and $\kappa, \tau \in \mathbb{R} \setminus \{0\}$.*

(i) *If $\tau > 0$ with $\tau \le \kappa \le 2\tau$ or $3\tau \le \kappa \le 4\tau$ or $\kappa \ge 5\tau$ or $\kappa<0$, then*

$$\begin{aligned}
\mathcal{P}_\kappa^\kappa(\boldsymbol{a},\boldsymbol{b}) - \mathcal{P}_\tau^\kappa(\boldsymbol{a},\boldsymbol{b}) \ge{}& \frac{\kappa(\kappa-\tau)(\kappa-2\tau)(\kappa-3\tau)}{24\tau^4\bar{a}} \\
&\times \sum_{\varsigma=1}^{m} a_\varsigma \left(\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^\tau\right)^4 \times \left(\frac{4\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) + b_\varsigma^\tau}{5}\right)^{\frac{\kappa}{\tau}-4} \\
&- \frac{\kappa(\kappa-\tau)(\kappa-2\tau)\mathcal{P}_\tau^{\kappa-3\tau}(\boldsymbol{a},\boldsymbol{b})}{6\tau^3\bar{a}} \sum_{\varsigma=1}^{m} a_\varsigma \left(\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^\tau\right)^3 \\
&+ \frac{\kappa(\kappa-\tau)\mathcal{P}_\tau^{\kappa-2\tau}(\boldsymbol{a},\boldsymbol{b})}{2\tau^2\bar{a}} \sum_{\varsigma=1}^{m} a_\varsigma \left(\mathcal{P}_\tau^\tau(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma^\tau\right)^2. \qquad (25)
\end{aligned}$$

(ii) *If $\tau < 0$ with $2\tau \le \kappa \le \tau$ or $4\tau \le \kappa \le 3\tau$ or $\kappa \le 5\tau$ or $\kappa>0$, then (25) holds.*

(iii) *If $\tau > 0$ with $0 < \kappa < \tau$, $2\tau < \kappa < 3\tau$ or $4\tau < \kappa < 5\tau$, then (25) is true in the reverse direction.*

(iv) *If $\tau < 0$ with $\tau < \kappa < 0$, $3\tau < \kappa < 2\tau$ or $5\tau < \kappa < 4\tau$, then the inequality in (25) is valid in the opposite sense.*

Proof. (i)–(ii) Since, the function $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$ is 6-convex on$(0,\infty)$ for the mentioned values of τ and κ, we assume $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = b_\varsigma^\tau$ in (11), and we obtain (25). (iii)–(iv) By utilizing (11) for $\delta_\varsigma = a_\varsigma$, $\varrho_\varsigma = b_\varsigma^\tau$, and $\Phi(\delta) = \delta^{\frac{\kappa}{\tau}}$, we deduce the converse of (25). □

Theorem 2 leads to the following corollary, which gives an interesting relationship for the power means.

Corollary 5. *We assume that* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *and* $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ *are positive tuples with* $\bar{a} = \sum_{\varsigma=1}^{m} a_\varsigma$*; then,*

$$\frac{\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b})}{\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})} \leq \exp\left[\frac{1}{120\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma\big)^4\big(24\mathcal{P}_1^{-4}(\boldsymbol{a},\boldsymbol{b}) + 6b_\varsigma^{-4}\big) + \frac{\mathcal{P}_1^{-3}(\boldsymbol{a},\boldsymbol{b})}{3\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma\big)^3 + \frac{\mathcal{P}_1^{-2}(\boldsymbol{a},\boldsymbol{b})}{2\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma\big)^2\right]. \quad (26)$$

Proof. We consider the function $\Phi(\delta) = -\ln x$ defined on the interval $(0, \infty)$; then, $\Phi''(\delta) = \delta^{-2}$ and $\Phi''''''(\delta) = 12\delta^{-6}$. From the above expressions, it is clear that $\Phi'' > 0$ and $\Phi'''''' > 0$, which substantiates the convexity as well as 6-convexity of the function $\Phi(\delta) = -\ln\delta$. Therefore, using inequality (8) by substituting $\Phi(\delta) = -\ln\delta$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = b_\varsigma$, we obtain (26). □

In the following corollary, we provide an application of Theorem 3 for the power means.

Corollary 6. *We suppose that all the hypotheses of Corollary 5 are valid; then,*

$$\frac{\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b})}{\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})} \geq \exp\left[\frac{625}{4\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma\big)^4\big(4\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) + b_\varsigma\big)^{-4} + \frac{\mathcal{P}_1^{-3}(\boldsymbol{a},\boldsymbol{b})}{3\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma\big)^3 + \frac{\mathcal{P}_1^{-2}(\boldsymbol{a},\boldsymbol{b})}{2\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - b_\varsigma\big)^2\right]. \quad (27)$$

Proof. Utilizing inequality (11) for $\Phi(\delta) = -\ln\delta$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = b_\varsigma$, we obtain (27). □

The following corollary provides another relation for the power means as a consequence of Theorem 2.

Corollary 7. *We assume that the conditions of Corollary 5 are satisfied; then,*

$$\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - \mathcal{P}_0(\boldsymbol{a},\boldsymbol{b}) \leq \frac{1}{120\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\ln\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b}) - \ln b_\varsigma\big)^4\big(4\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b}) + b_\varsigma\big) - \frac{\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})}{6\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\ln\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - \ln b_\varsigma\big)^3 + \frac{\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})}{2\bar{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\ln\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b}) - \ln b_\varsigma\big)^2. \quad (28)$$

Proof. Let us take the function $\Phi(\delta) = \exp(\delta)$, $\delta \in \mathbb{R}$; then, $\Phi''(\delta) = \exp(\delta)$, and the function $\Phi''''''(\delta) = \exp(\delta)$. Obviously, the function $\Phi(\delta) = \exp(\delta)$ is convex and 6-convex because both $\Phi''(\delta) = \exp(\delta)$ and the function $\Phi''''''(\delta) = \exp(\delta)$ are positive. Therefore, by applying (8) for $\varrho_\varsigma = \ln b_\varsigma$, $\delta_\varsigma = a_\varsigma$, and $\Phi(\delta) = \exp(\delta)$, we deduce (28) . □

Theorem 3 leads to the establishment of the following relation for the means.

Corollary 8. *We assume that the conditions of Corollary 5 are satisfied. Then,*

$$\begin{aligned}\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b})-\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b}) \geq{}& \frac{1}{24\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\ln\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})-\ln b_\varsigma\big)^4\mathcal{P}_0^{\frac{4}{5}}(\boldsymbol{a},\boldsymbol{b})\\ &-\frac{\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})}{6\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\ln\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b})-\ln b_\varsigma\big)^3\\ &+\frac{\mathcal{P}_0(\boldsymbol{a},\boldsymbol{b})}{2\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(\ln\mathcal{P}_1(\boldsymbol{a},\boldsymbol{b})-\ln b_\varsigma\big)^2. \end{aligned} \tag{29}$$

Proof. Applying inequality (11) for $\Phi(\delta) = \exp(\delta)$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = \ln b_\varsigma$, we acquire (29). $\square$

The quasi-arithmetic means is defined as follows:

Definition 3. *Let g be a continuous and strictly monotonic function and $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ be positive tuples with $\overline{a} = \sum_{\varsigma=1}^{m} a_\varsigma$. Then, the quasi-arithmetic mean is given by:*

$$Q_g(\boldsymbol{a},\boldsymbol{b}) = g^{-1}\left(\frac{1}{\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma g(b_\varsigma)\right).$$

The following corollary presents a relation for the quasi-arithmetic means.

Corollary 9. *Assuming that g is a continuous and strictly monotonic function such that $\Phi\circ g^{-1}$ is 6-convex, and the tuples $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ are positive with $\overline{a} = \sum_{\varsigma=1}^{m} a_\varsigma$, then*

$$\begin{aligned}\frac{1}{\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\Phi(b_\varsigma)-\Phi\Big(Q_g(\boldsymbol{a},\boldsymbol{b})\Big) \leq{}& \frac{1}{120\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\left(\frac{1}{\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma g(b_\varsigma)-g(b_\varsigma)\right)^4\\ &\times\Big(4(\Phi\circ g^{-1})''''Q_g(\boldsymbol{a},\boldsymbol{b})-(\Phi\circ g^{-1})''''g(b_\varsigma)\Big)\\ &-\frac{(\Phi\circ g^{-1})'''\Big(g(Q_g(\boldsymbol{a},\boldsymbol{b}))\Big)}{6\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\Big(g(Q_g(\boldsymbol{a},\boldsymbol{b}))-g(b_\varsigma)\Big)^3\\ &+\frac{(\Phi\circ g^{-1})''\Big(g(Q_g(\boldsymbol{a},\boldsymbol{b}))\Big)}{2\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\Big(g(Q_g(\boldsymbol{a},\boldsymbol{b}))-g(b_\varsigma)\Big)^2. \end{aligned} \tag{30}$$

Proof. To deduce inequality (30), we assume $\Phi = \Phi\circ g^{-1}$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = g(b_\varsigma)$ in (8). $\square$

An application of Theorem 3 for the quasi-arithmetic means is presented in the following corollary.

Corollary 10. *Let the Corollary 9 hypotheses be valid. Then,*

$$\frac{1}{\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma \Phi(b_\varsigma) - \Phi\big(Q_g(\boldsymbol{a},\boldsymbol{b})\big) \geq \frac{1}{24\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma \big(g(Q_g(\boldsymbol{a},\boldsymbol{b}))) - g(b_\varsigma)\big)^4$$
$$\times\big((\Phi\circ g^{-1})''''\left(\frac{4g(Q_g(\boldsymbol{a},\boldsymbol{b})) + g(b_\varsigma)}{5}\right)$$
$$-\frac{(\Phi\circ g^{-1})'''\big(g\big(Q_g(\boldsymbol{a},\boldsymbol{b})\big)\big)}{6\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(g(Q_g(\boldsymbol{a},\boldsymbol{b})) - g(b_\varsigma)\big)\big)^3$$
$$+\frac{(\Phi\circ g^{-1})''\big(g\big(Q_g(\boldsymbol{a},\boldsymbol{b})\big)\big)}{2\overline{a}}\sum_{\varsigma=1}^{m} a_\varsigma\big(g(Q_g(\boldsymbol{a},\boldsymbol{b})) - g(b_\varsigma)\big)\big)^2. \tag{31}$$

Proof. We apply inequality (11) for $\Phi = \Phi\circ g^{-1}$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = g(b_\varsigma)$, and we receive (31). $\square$

6. Applications in Information Theory

This section of the paper is devoted to the applications of the main results in information theory. The proposed applications provide estimates for the Csiszár divergence and for its related concepts, such as the Rényi divergence, Kullback–Leibler divergence, Bhattacharyya coefficient, and the Shannon entropy. To present the applications of the main results, first, we give the definitions of some notions in information theory. The following is the formal definition of the Csiszár divergence:

Definition 4. *Let the tuples $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ be positive with $\frac{b_\varsigma}{a_\varsigma} \in [\sigma_1, \sigma_2]$ and $a_\varsigma > 0\ \forall\, \varsigma \in \{1, 2, \ldots, m\}$, and the function $g : [\sigma_1, \sigma_2] \to \mathbb{R}$ be convex. Then, the Csiszár divergence is given by:*

$$C_d(\boldsymbol{a},\boldsymbol{b}) = \sum_{\varsigma=1}^{m} a_\varsigma g\left(\frac{b_\varsigma}{a_\varsigma}\right).$$

Following are the definitions of some notions, which can be deduced from the Csiszár divergence.

Definition 5. *Let $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ be positive probability distributions. Then,*

- *The **Rényi divergence** is defined by:*

$$R_d(\boldsymbol{a},\boldsymbol{b}) = \frac{1}{c-1}\log\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c}\right).$$

- *The **Shannon Entropy** is defined as follows:*

$$S_e(\boldsymbol{a}) = -\sum_{\varsigma=1}^{m} a_\varsigma \log a_\varsigma.$$

- *The **Kullback –Leibler Divergence** is given by:*

$$K_{bl}(\boldsymbol{a},\boldsymbol{b}) = \sum_{\varsigma=1}^{m} b_\varsigma \log\left(\frac{b_\varsigma}{a_\varsigma}\right).$$

- *The **Bhattacharyya Coefficient** is defined as:*

$$B_c(\boldsymbol{a},\boldsymbol{b}) = \sum_{\varsigma=1}^{m} \sqrt{a_\varsigma b_\varsigma}.$$

An estimate for the Csiszár divergence as an application of Theorem 3 is presented in the preceding corollary.

Theorem 4. *We presume that the tuples $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ are positive, with $\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}$, $\frac{b_\varsigma}{a_\varsigma} \in [\sigma_1, \sigma_2]$ and $a_\varsigma > 0$ for all $\varsigma \in \{1, 2, \ldots, m\}$. If the function $g : [\sigma_1, \sigma_2] \to \mathbb{R}$ is 6-convex, then*

$$\begin{aligned}
C_d(\boldsymbol{a},\boldsymbol{b}) - g\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right)\sum_{\varsigma=1}^{m} a_\varsigma &\leq \frac{1}{120}\sum_{\varsigma=1}^{m} a_\varsigma \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma} - \frac{b_\varsigma}{a_\varsigma}\right)^4 \\
&\times \left(4g''''\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right) + g''''\left(\frac{b_\varsigma}{a_\varsigma}\right)\right) - \frac{1}{6}g'''\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right)\sum_{\varsigma=1}^{m} a_\varsigma \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma} - \frac{b_\varsigma}{a_\varsigma}\right)^3 \\
&+ \frac{1}{2}g''\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right)\sum_{\varsigma=1}^{m} a_\varsigma \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma} - \frac{b_\varsigma}{a_\varsigma}\right)^2.
\end{aligned} \tag{32}$$

Proof. By applying (8) for $\Phi = g$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = \frac{b_\varsigma}{a_\varsigma}$, we acquire (32). □

The following corollary presents an estimate for the Csiszár divergence as an application of Theorem 3.

Theorem 5. *We presume that the Theorem 4 conditions are authentic; then,*

$$\begin{aligned}
C_d(\boldsymbol{a},\boldsymbol{b}) - g\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right)\sum_{\varsigma=1}^{m} a_\varsigma &\geq \frac{1}{24}\sum_{\varsigma=1}^{m} a_\varsigma \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma} - \frac{b_\varsigma}{a_\varsigma}\right)^4 g''''\left(\frac{4\sum_{\varsigma=1}^{m} b_\varsigma}{5\sum_{\varsigma=1}^{m} a_\varsigma} + \frac{b_\varsigma}{a_\varsigma}\right) \\
&- \frac{1}{6}g'''\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right)\sum_{\varsigma=1}^{m} a_\varsigma \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma} - \frac{b_\varsigma}{a_\varsigma}\right)^3 \\
&+ \frac{1}{2}g''\left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma}\right)\sum_{\varsigma=1}^{m} a_\varsigma \left(\frac{\sum_{\varsigma=1}^{m} b_\varsigma}{\sum_{\varsigma=1}^{m} a_\varsigma} - \frac{b_\varsigma}{a_\varsigma}\right)^2.
\end{aligned} \tag{33}$$

Proof. By applying inequality (11), while taking $\Phi = g$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = \frac{b_\varsigma}{a_\varsigma}$, we obtain (33). □

A bound for the Rényi divergence is derived using Theorem 2 and is presented in the following corollary.

Corollary 11. *If* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *and* $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ *are two positive probability distributions, and* $c \in [0, \infty)$, *such that* $c \neq 1$, *then*

$$\begin{aligned} R_d(\boldsymbol{a},\boldsymbol{b}) - \sum_{\varsigma=1}^{m} a_\varsigma \log\left(\frac{a_\varsigma}{b_\varsigma}\right) \leq \frac{1}{20}\sum_{\varsigma=1}^{m} a_\varsigma \left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} - \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}\right)^4 \\ \times\left(\frac{4}{c-1}\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c}\right)^{-4} + \frac{1}{c-1}\left(\frac{a_\varsigma}{b_\varsigma}\right)^{4(1-c)}\right) \\ +\frac{1}{3(c-1)}\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c}\right)^{-3}\sum_{\varsigma=1}^{m} a_\varsigma\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} - \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}\right)^3 \\ +\frac{1}{2(c-1)}\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c}\right)^{-2}\sum_{\varsigma=1}^{m} a_\varsigma\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} - \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}\right)^2. \end{aligned} \tag{34}$$

Proof. We consider $\Phi(\delta) = -\frac{1}{c-1}\ln\delta$, $\delta > 0$; then, $\Phi''(\delta) = \frac{1}{(c-1)\delta^2}$ and $\Phi''''''(\delta) = \frac{120}{(c-1)\delta^6}$. It is clear from the above expressions that both Φ'' and Φ'''''' are positive for all $\delta \in (0, \infty)$. This confirms that the function $\Phi(\delta) = -\frac{1}{c-1}\ln\delta$ is both convex and 6-convex. Therefore, by substituting $\Phi(\delta) = -\frac{1}{c-1}\ln\delta$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}$ in (8), we obtain (34). □

The accompanying corollary examines how Theorem 3 is connected to the Rényi divergence.

Corollary 12. *We suppose that the Corollary 11 assumptions are valid; then,*

$$\begin{aligned} R_d(\boldsymbol{a},\boldsymbol{b}) - \sum_{\varsigma=1}^{m} a_\varsigma \log\left(\frac{a_\varsigma}{b_\varsigma}\right) \geq \frac{1}{4(c-1)}\sum_{\varsigma=1}^{m} a_\varsigma \left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} - \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}\right)^4 \\ \times\left(\frac{4\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} + \left(\frac{a_\varsigma}{b_\varsigma}\right)^{(c-1)}}{5}\right)^{-4} \\ +\frac{1}{3(c-1)}\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c}\right)^{-3}\sum_{\varsigma=1}^{m} a_\varsigma\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} - \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}\right)^3 \\ +\frac{1}{2(c-1)}\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c}\right)^{-2}\sum_{\varsigma=1}^{m} a_\varsigma\left(\sum_{\varsigma=1}^{m} a_\varsigma^c b_\varsigma^{1-c} - \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}\right)^2. \end{aligned} \tag{35}$$

Proof. Utilizing inequality (11) for $\Phi(\delta) = -\frac{1}{c-1}\ln\delta$, $\delta_\varsigma = a_\varsigma$, and $\varrho_\varsigma = \left(\frac{a_\varsigma}{b_\varsigma}\right)^{c-1}$, we deduce (35). □

The ensuing corollary offers an estimate for the Shannon entropy as an application of Theorem 2.

Corollary 13. *Let the tuple* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *be positive, with* $\sum\limits_{\varsigma=1}^{m} a_\varsigma = 1$. *Then,*

$$\log n - S_e(\boldsymbol{a}) \leq \frac{1}{20}\sum_{\varsigma=1}^{m} a_\varsigma \left(n - \frac{1}{a_\varsigma}\right)^4 \left(4(n)^{-4} + a_\varsigma^4\right) + \frac{n^{-3}}{3}\sum_{\varsigma=1}^{m} a_\varsigma \left(n - \frac{1}{a_\varsigma}\right)^3 + \frac{n^{-2}}{2}\sum_{\varsigma=1}^{m} a_\varsigma \left(n - \frac{1}{a_\varsigma}\right)^2. \tag{36}$$

Proof. Since, the function $g(\delta) = -\ln\delta$ is convex as well as 6-convex on $(0, \infty)$. Therefore, to deduce inequality (36), we apply inequality (32) for $g(\delta) = -\ln\delta,\ \delta > 0$, and $b_\varsigma = 1$ for all $\varsigma \in \{1, 2, \ldots, n\}$. □

The next corollary presents an estimate for the Shannon entropy.

Corollary 14. *Assuming that the conditions of Corollary 13 are valid, then*

$$\log n - S_e(\boldsymbol{a}) \geq \frac{1}{4}\sum_{\varsigma=1}^{m} a_\varsigma \left(n - \frac{1}{a_\varsigma}\right)^4 \left(\frac{4n}{5} + \frac{1}{5a_\varsigma}\right)^{-4} + \frac{n^{-3}}{3}\sum_{\varsigma=1}^{m} a_\varsigma \left(n - \frac{1}{a_\varsigma}\right)^3 + \frac{n^{-2}}{2}\sum_{\varsigma=1}^{m} a_\varsigma \left(n - \frac{1}{a_\varsigma}\right)^2. \tag{37}$$

Proof. Taking $g(\delta) = -\ln\delta,\ \delta > 0$, and $b_\varsigma = 1$ for all $\varsigma \in \{1, 2, \ldots, n\}$ in (33), we obtain (37). □

The application of Theorem 2 for the Kullback–Leibler divergence is discussed in the following corollary.

Corollary 15. *Let* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *and* $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ *be any positive probability distributions. Then,*

$$D_{kl}(\boldsymbol{a}, \boldsymbol{b}) \leq \frac{1}{60}\sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^4 \left(4 + \left(\frac{a_\varsigma}{b_\varsigma}\right)^3\right) + \frac{1}{6}\sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^3 + \frac{1}{2}\sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^2. \tag{38}$$

Proof. We consider the function $g(\delta) = \delta\ln\delta,\ \delta > 0$; then, by differentiating the given with respect to δ, we acquire $g''(\delta) = \frac{1}{\delta}$ and $g''''''(\delta) = 24\delta^{-5}$. Clearly, both g'' and g'''''' are positive with the given conditions, which admits that the function $g(\delta) = \delta\ln\delta$ is convex as well as 6-convex. Therefore, putting $g(\delta) = \delta\ln\delta$ in (32), we deduce (38). □

The consequence of Theorem 3 for the Kullback–Leibler divergence is given in the following corollary.

Corollary 16. *We assume that the postulates of Corollary 15 are valid; then,*

$$D_{kl}(\boldsymbol{a}, \boldsymbol{b}) \geq \frac{1}{60}\sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^4 \left(4 + \frac{b_\varsigma}{a_\varsigma}\right)^{-3} + \frac{1}{6}\sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^3 + \frac{1}{2}\sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^2. \tag{39}$$

Proof. By applying inequality (33) for the function $g(\delta) = \delta\ln\delta,\ \delta > 0$, we obtain (39). □

The next corollary provides an estimate for the Bhattacharyya coefficient as an application of Theorem 2.

Corollary 17. *We assume that* $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ *and* $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)$ *are arbitrary probability distributions, such that* $a_\varsigma, b_\varsigma > 0$ *for all* $\varsigma \in \{1, 2, \ldots, m\}$*. Then,*

$$1 - B_c(\boldsymbol{a}, \boldsymbol{b}) \leq \frac{15}{480} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^4 \left(1 + \frac{1}{4}\left(\frac{a_\varsigma}{b_\varsigma}\right)^{\frac{7}{2}}\right) + \frac{1}{16} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^3 + \frac{1}{8} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^2. \tag{40}$$

Proof. Let us choose the function $g(\delta) = -\sqrt{\delta}$ defined on $(0, \infty)$. Then, $g''(\delta) = \frac{1}{4}\delta^{\frac{-1}{2}}$, and $g''''''(\delta) = \frac{945}{64}\delta^{\frac{-11}{2}}$. From these expressions, we conclude that the function $g(\delta) = -\sqrt{\delta}$ is convex as well as 6-convex because both g'' and g'''''' are positive on $(0, \infty)$. Therefore, utilizing (32) for $g(\delta) = -\sqrt{\delta}$, we obtain (40). □

Another estimate for the Bhattacharyya coefficient is presented in the following corollary.

Corollary 18. *We suppose that the postulates of Corollary 17 are valid; then,*

$$1 - B_c(\boldsymbol{a}, \boldsymbol{b}) \geq \frac{15}{384} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^4 \left(\frac{4}{5} + \frac{b_\varsigma}{5a_\varsigma}\right)^{-\frac{7}{2}} + \frac{1}{16} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^3 + \frac{1}{8} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^2. \tag{41}$$

Proof. Choosing the function $g(\delta) = \sqrt{\delta},\ \delta > 0$ in (33), we receive (41). □

7. Applications for the Zipf–Mandelbrot Entropy

The Zipf–Mandelbrot law is a discrete probability distribution that depends on three different factors, $n \in \mathbb{N}$, $\theta \geq 0$, and $s > 0$, which is defined as:

$$\Phi(\varsigma, m, \theta, s) = \frac{1}{(\varsigma + \theta)^s N_{m,\theta,s}}, \quad \varsigma \in \{1, 2, \ldots, m\},$$

where Φ represents the probability mass function, and $N_{m,\theta,s} = \sum\limits_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta)^s}$ is the generalized harmonic number. The famed Zipf–Mandelbrot entropy can be used to calculate the quantity of information in a given text in relation to the perspective of information theory. The traditional form of the Zipf–Mandelbrot entropy is given by:

$$Z(m, \theta, s) = \frac{s}{N_{m,\theta,s}} \sum_{\varsigma=1}^{m} \frac{\log(\varsigma + \theta)}{(\varsigma + \theta)^s} + \log N_{m,\theta,s}.$$

The following corollary presents an estimate for the Zipf–Mandelbrot entropy as a direct consequence of Theorem 2.

Corollary 19. *Let $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ be any probability distribution with positive entries, such that $\sum_{\varsigma=1}^{m} a_\varsigma = 1$ and $\theta \geq 0$, $s > 0$. Then,*

$$-Z(m,\theta,s) - \frac{1}{N_{m,\theta,s}} \sum_{\varsigma=1}^{m} \frac{\log a_\varsigma}{(\varsigma+\theta)^s}$$
$$\leq \frac{1}{60} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s)N_{m,\theta,s}}\right)^4 \left(4 + \left(a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}\right)^3\right)$$
$$+\frac{1}{6} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s)N_{m,\theta,s}}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^2. \tag{42}$$

Proof. We consider $b_\varsigma = \frac{1}{(\varsigma+\theta)^s N_{m,\theta,s}}, \quad \varsigma \in \{1, 2, \ldots, m\}$; then,

$$\begin{aligned}
\sum_{\varsigma=1}^{m} b_\varsigma \log \frac{b_\varsigma}{a_\varsigma} &= \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta)^s N_{m,\theta,s}} \log \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}} \\
&= \sum_{i=1}^{n} \frac{1}{(\varsigma+\theta)^s N_{m,\theta,s}} (-s \log(\varsigma+\theta) - \log N_{m,\theta,s} - \log a_i) \\
&= -Z(m,\theta,s) - \frac{1}{N_{m,\theta,s}} \sum_{\varsigma=1}^{m} \frac{\log a_\varsigma}{(\varsigma+\theta)^s},
\end{aligned} \tag{43}$$

and

$$\frac{1}{60} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^4 \left(4 + \left(\frac{a_\varsigma}{b_\varsigma}\right)^3\right) + \frac{1}{6} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^2$$
$$= \frac{1}{60} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^4 \left(4 + \left(a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}\right)^3\right)$$
$$+\frac{1}{6} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^2. \tag{44}$$

Now, utilizing (43) and (44) in (38), we acquire (42). □

An application of Theorem 3 for the Zipf–Mandelbrot entropy is provided in the following corollary.

Corollary 20. *Let $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ be any positive probability distribution, such that $a_\varsigma > 0$ and $\theta \geq 0$, $s > 0$. Then,*

$$-Z(m,\theta,s) - \frac{1}{N_{m,\theta,s}} \sum_{\varsigma=1}^{m} \frac{\log a_\varsigma}{(\varsigma+\theta)^s}$$
$$\geq \frac{1}{60} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^4 \left(4 + \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^{-3}$$
$$+\frac{1}{6} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{1}{a_\varsigma(\varsigma+\theta)^s N_{m,\theta,s}}\right)^2. \tag{45}$$

Proof. By utilizing inequality (39) for $b_\varsigma = \frac{1}{(\varsigma+\theta)^s N_{m,\theta,s}}$ while adopting the proof of method of Corollary 19, we deduce (45). □

An estimate for the Zipf–Mandelbrot entropy is given in the following corollary as an application of Theorem 2.

Corollary 21. *Let us assume that $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$ is any probability distribution with positive entries, such that $\sum_{\varsigma=1}^{m} a_\varsigma = 1$ and $\theta_1, \theta_2 \geq 0$, $s_1, s_2 > 0$. Then,*

$$\begin{aligned}
&-Z(m,\theta_1,s_1) - \frac{1}{N_{m,\theta_1,s_1}} \sum_{\varsigma=1}^{m} \frac{\log(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1}} \\
&\leq \frac{1}{60} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^4 \\
&\times \left(4 + \left(\frac{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}\right)^3\right) + \frac{1}{6} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \\
&\times \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \\
&\times \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^2.
\end{aligned} \tag{46}$$

Proof. We consider $b_\varsigma = \frac{1}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}$, and $a_\varsigma = \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}$; then, we have

$$\sum_{\varsigma=1}^{m} b_\varsigma \log \frac{b_\varsigma}{a_\varsigma} = -Z(m,\theta_1,s_1) - \sum_{\varsigma=1}^{m} \frac{\log(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}, \tag{47}$$

and

$$\begin{aligned}
&\frac{1}{60} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^4 \left(4 + \left(\frac{a_\varsigma}{b_\varsigma}\right)^3\right) + \frac{1}{6} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} a_\varsigma \left(1 - \frac{b_\varsigma}{a_\varsigma}\right)^2 \\
&= \frac{1}{60} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^4 \\
&\times \left(4 + \left(\frac{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}\right)^3\right) + \frac{1}{6} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \\
&\times \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \\
&\times \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^2 \quad .
\end{aligned} \tag{48}$$

Now, utilizing (47) and (48) in (38), we obtain (46). □

The subsequent corollary employs Theorem 3 to offer a further estimate for the Zipf–Mandelbrot entropy.

Corollary 22. *We assume that $\theta_1, \theta_2 \geq 0$, and $s_1, s_2 > 0$; then,*

$$
\begin{aligned}
-Z(m,\theta_1,s_1) - \frac{1}{N_{m,\theta_1,s_1}} \sum_{\varsigma=1}^{m} \frac{\log(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1}} \\
\geq \frac{1}{60} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^4 \\
\times \left(4 + \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^{-3} + \frac{1}{6} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \\
\times \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^3 + \frac{1}{2} \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}} \\
\times \left(1 - \frac{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}\right)^2.
\end{aligned}
\tag{49}
$$

Proof. To obtain (49), we use (39) for $b_\varsigma = \frac{1}{(\varsigma+\theta_1)^{s_1} N_{m,\theta_1,s_1}}$ and $a_\varsigma = \sum_{\varsigma=1}^{m} \frac{1}{(\varsigma+\theta_2)^{s_2} N_{m,\theta_2,s_2}}$ while following the proof of the procedure of Corollary 21. □

8. Conclusions

Mathematical inequalities have received a very strong response from researchers working in different areas because of their attractive features and characteristics. In this regard, the Jensen inequality is a popular inequality among mathematical inequalities. This inequality has some important properties and also has a deep relation with the notion of convexity; due to these facts, it has an important position in areas of pure and applied mathematics. In this paper, we developed some interesting relations that provide new estimates for the Jensen gap. The proposed estimations were made possible with the help of the generalized convexity known as 6-convex functions. Some examples were provided to show the accuracy and ameliorations of the acquired estimates. In these examples, we confirmed that our results provided better estimations for the Jensen gap as compared to other recent results of a similar type. We also discussed some consequences of the main results for the Hölder inequality. We also provided some applications of the main outcomes for the well-known power means and quasi-arithmetic means. Moreover, we presented some more applications of the main results in information theory that offer bounds for the Csiszár and Kullback–Leibler divergences, Bhattacharyya coefficient, and Shannon entropy. Some estimates for the Zipf–Mandelbrot entropy were also granted as additional applications of the obtained results. The techniques, which were applied for the derivations of the main results, are the definition of the convex function and prominent Jensen's inequality in the integral version. The importance of our work is also highlighted in Remark 2. The idea and techniques utilized in this manuscript may also be applied for the integral Jensen's inequality and some other inequalities such as Slater's and the majorization inequalities. Furthermore, this idea may also be applied for convexity of higher order than 6-convexity.

Author Contributions: M.A.K. provided the main idea of the main results. M.A.K. and A.S. worked on Section 2. H.U. worked on Section 3. H.U. and T.S. worked on Sections 4 and 5. A.S. and H.U. worked on Sections 6 and 7. M.A.K., A.S. and T.S. worked on Sections 7 and 8. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research work was funded by Institutional Fund Projects under grant no. (IFPIP:322-130-1443). The authors gratefully acknowledge the technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Conflicts of Interest: The authors declare that there are no conflict of interest regarding the publication of this paper.

References

1. Anikin, A.S.; Gasnikov, A.V.; Dvurechensky, P.E.; Tyurin, A.I.; Chernov, A.V. Dual approaches to the minimization of strongly convex functionals with a simple structure under affine constraints. *Comput. Math. Math. Phys.* **2017**, *57*, 1262–1276. [CrossRef]
2. Ghadimi, S.; Lan, G.; Zhang, H. Generalized uniformly optimal methods for non-linear programming. *J. Sci. Comput.* **2019**, *79*, 1854–1881. [CrossRef]
3. Noor, M.A.; Noor, K.I.; Awan, M.U. Generalized convexity and integral inequalities. *Appl. Math. Inf. Sci.* **2015**, **9**, 233–243. [CrossRef]
4. Niculescu, C.P.; Persson, L.E. *Convex Functions and Their Applications, A Contemporary Approach, CMS Books in Mathematics*; Springer: New York, NY, USA, 2006.
5. Pečarić, J.; Persson, L.E.; Tong, Y.L. *Convex Functions, Partial Ordering and Statistical Applications*; Academic Press: New York, NY, USA, 1992.
6. Chu, Y.-M.; Zhao, T.-H. Concavity of the error function with respect to Hölder means. *Math. Inequal. Appl.* **2016**, *19*, 589–595. [CrossRef]
7. Dragomir, S.S.; Pearce, E.E.M. *Selected Topics on Hermite–Hadamard Inequalities and Applications*; Victoria University: Victoria, Australia, 2000.
8. Horváth, L.; Pečarić, Đ.; Pečarić, J. Estimations of f-and Rényi divergences by using a cyclic refinement of the Jensen's inequality. *Bull. Malays. Math. Sci. Soc.* **2019**, *42*, 933–946. [CrossRef]
9. Marshall, A.W.; Olkin, I.; Arnold, B. *Inequalities: Theory of Majorization and Its Applications*, 2nd ed.; Springer Series in Statistics; Springer: New York, NY, USA, 2011.
10. Ali, S.; Saif, M.; Khan, K.A.; Shah, N.A.; Weera, W. A note on varying G and L in Chern–Simons modified gravity. *Symmetry* **2022**, *14*, 1430. [CrossRef]
11. Butt, S.I.; Yousaf, S.; Khan, K.A.; Mabela, R.M.; Alsharif, A.M. Fejer–Pachpatte–Mercer–type inequalities for harmonically convex functions involving exponential function in kernel. *Math. Probl. Eng.* **2022**, *2022*, 7269033. [CrossRef]
12. Zhao, T.-H.; Wang, M.-K.; Chu, Y.-M. Concavity and bounds involving generalized elliptic integral of the first kind. *J. Math. Inequal.* **2021**, *15*, 701–724. [CrossRef]
13. Butt, S.I.; Rashid, S.; Javed, I.; Khan, K.A.; Mabela, R.M. New fractional estimates of Simpson–Mercer type for twice differentiable mappings pertaining to Mittag–Leffler kernel. *J. Funct. Space* **2022**, *2022*, 4842344. [CrossRef]
14. Chu, Y.-M.; Zhao, T.-H. Convexity and concavity of the complete elliptic integrals with respect to Lehmer mean. *J. Inequal. Appl.* **2015**, *396*, 1–6. [CrossRef]
15. Zhao, T.-H.; Wang, M.-K.; Chu, Y.-M. Monotonicity and convexity involving generalized elliptic integral of the first kind. *Revista de la Real Academia de Ciencias Exactas Físicas y Naturales Serie A Matemáticas* **2021**, *115*, 46. [CrossRef]
16. Dragomir, S.S. Some Slater's type inequalities for convex functions defined on linear spaces and applications. *Abstr. Appl. Anal.* **2012**, *2012*, 168405. [CrossRef]
17. Adeel, M.; Khan, K.A.; Pečarić, Đ.; Pečarić, J. Levinson type inequalities for higher order convex functions via Abel–Gontscharoff interpolation. *Adv. Differ. Equ.* **2019**, *430*, 1–13. [CrossRef]
18. Adeel, M.; Khan, K.A.; Pečarić, Đ.; Pečarić, J. Estimation of f–divergence and Shannon entropy by Levinson type inequalities for higher–order convex functions via Taylor polynomial. *J. Math. Comput. Sci.* **2020**, *21*, 322–334. [CrossRef]
19. Sezer, S.; Eken, Z.; Tinaztepe, G.; Adilov, G. p-convex functions and some of their properties. *Numer. Funct. Anal. Optim.* **2021**, *42*, 443–459. [CrossRef]
20. Hudzik, H.; Maligranda, L. Some remarks on s–convex functions. *Aequationes Math.* **1994**, *48*, 100–111. [CrossRef]
21. Adil Khan, M.; Wu, S.-H.; Ullah, H.; Chu, Y.-M. Discrete majorization type inequalities for convex functions on rectangles. *J. Inequal. Appl.* **2019**, *16*, 1–18. [CrossRef]
22. Varošanec, S. On h-convexity. *J. Math. Anal. Appl.* **2007**, *326*, 303–311. [CrossRef]
23. You, X.; Khan, M.A.; Ullah, H.; Saeed, T. Improvements of Slater's inequality by means of 4-convexity and its applications. *Matheatics* **2022**, *10*, 1274.
24. Dragomir, S.S.; Agarwal, R.P.; Cerone, P. On Simpson's inequality and applications. *J. Inequal. Appl.* **2000**, *5*, 533–579. [CrossRef]
25. Ullah, H.; Adil Khan, M.; Saeed, T.; Sayed, Z.M.M.M. Some improvements of Jensen's inequality via 4–convexity and applications. *J. Funct. Space* **2022**, *2022*, 2157375. [CrossRef]
26. Youness, E.A. E–convex sets, E–convex functions, and E-convex programming. *J. Optim. Theory Appl.* **1999**, *102*, 439–450. [CrossRef]
27. Zhao, C.-J.; Cheung, W.-S. On improvements of the Rozanova's inequality. *J. Inequal. Appl.* **2011**, *33*, 1–7. [CrossRef]

28. Furuichi, S.; Moradi, H.R.; Zardadi, A. Some new Karamata type inequalities and their applications to some entropies. *Rep. Math. Phys.* **2019**, *84*, 201–214. [CrossRef]
29. Zhao, T.-H.; Yang, Z.-H.; Chu, Y.-M. Monotonicity properties of a function involving the psi function with applications. *J. Inequal. Appl.* **2015**, *193*, 1–10. [CrossRef]
30. Zhao, T.-H.; Shi, L.; Chu, Y.-M. Convexity and concavity of the modified Bessel functions of the first kind with respect to Hölder means. *Revista de la Real Academia de Ciencias Exactas Físicas y Naturales Serie A Matemáticas* **2020**, *114*, 96. [CrossRef]
31. Adil Khan, M.; Ullah, H.; Saeed, T.; Alsulami, H.H.; Sayed, Z.M.M.M.; Alshehri, A.M. Estimations of the Slater gap via convexity and its applications in information theory. *Math. Probl. Eng.* **2022**, *2022*, 1750331. [CrossRef]
32. Mercer, A.M. A variant of Jensen's inequality. *J. Inequal. Pure Appl. Math.* **2003**, *4*, 73.
33. Latif, M.A.; Dragomir, S.S.; Momoniat, E. Some Fejer type integral inequalities for geometrically–arithmetically–convex functions with applications. *Filomat* **2018**, *32*, 2193–2206. [CrossRef]
34. Mihai, M.V.; Noor, M.A.; Noor, K.I.; Awan, M.U. Some integral inequalities for harmonic h–convex functions involving hypergeometric functions. *Appl. Math. Comput.* **2015**, *252*, 257–262. [CrossRef]
35. Zhao, T.-H.; He, Z.-Y.; Chu, Y.-M. On some renfements for inequalities involving zero–balanced hypergeometric function. *AIMS Math.* **2020**, *5*, 6479–6495. [CrossRef]
36. Wang, M.-K.; Hong, M.-Y.; Xu, Y.-F.; Shen, Z.-H.; Chu, Y.-M. Inequalities for generalized trigonometric and hyperbolic functions with one parameter. *J. Math. Inequal.* **2020**, *14*, 1–21. [CrossRef]
37. Liu, W.; Shi, F.; Ye, G.; Zhao, D. The properties of harmonically cr-h-convex function and its applications. *Mathematics* **2022**, *10*, 2089. [CrossRef]
38. Zhao, T.-H.; Wang, M.-K.; Chu, Y.-M. A sharp double inequality involving generalized complete elliptic integral of the first kind. *AIMS Math.* **2020**, **5**, 4512–4528. [CrossRef]
39. Davis, C. A Schwarz inequality for convex operator functions. *Proc. Am. Math. Soc.* **1957**, *8*, 42–44. [CrossRef]
40. Shor, N. *Minimization Methods for Non-Differentiable Functions*; Springer: Berlin/Heidelberg, Germany, 1985.
41. Ullah, H.; Adil Khan, M.; Saeed, T.; Determination of bounds for the Jensen gap and its applications. *Mathematics* **2021**, *9*, 3132. [CrossRef]
42. Dragomir, S.S. A converse result for Jensen's discrete inequality via Gruss inequality and applications in Information Theory. *An. Univ. Oradea Fasc. Mat* **1999**, *7*, 178–189.
43. Cloud, M.J.; Drachman, B.C.; Lebedev, L.P. *Inequalities with Applications to Engineering*; Springer: Cham, Switzerland; Heidelberg, Germany, 2014. [CrossRef]
44. White, C.C.; Harrington, D.P. Application of Jensen's inequality to adaptive suboptimal design. *J. Optim. Theory Appl.* **1980**, *32*, 89–99. [CrossRef]
45. Mukhopadhyay, N. On sharp Jensen's inequality and some unusual applications, communications in statistics. *Theor. Methods* **2011**, *40*, 1283–1297. [CrossRef]
46. Azar, S.A. Jensen's inequality in finance. *Int. Adv. Econ. Res.* **2008**, *14*, 433–440. [CrossRef]
47. Ullah, H.; Adil Khan, M.; Pečarić, J. New bounds for soft margin estimator via concavity of Gaussian weighting function. *Adv. Differ. Equ.* **2020**, *644*, 1–10. [CrossRef]
48. Tapus, N.; Popescu, P.G. A new entropy upper bound. *Appl. Math. Lett.* **2012**, *25*, 1887–1890. [CrossRef]
49. Bakula, M.K.; Pečarić, J. On the Jensen's inequality for convex functions on the co–ordinates in a rectangle from the plane. *Taiwan. J. Math.* **2006**, *10*, 1271–1292. [CrossRef]
50. Dragomir, S.S. A further improvement of Jensen's inequality. *Tamkang J. Math.* **1994**, *25*, 29–36. [CrossRef]
51. Dragomir, S.S.; Ionescu, N.M. Some converse of Jensen's inequality and applications. *Rev. Anal. Numér. Théor. Approx.* **1994**, *23*, 71–78.
52. Bakula, M.K.; Nikodem, K. On the converse Jensen inequality for strongly convex functions. *J. Math. Anal. Appl.* **2016**, *434*, 516–522. [CrossRef]
53. Dragomir, S.S. Some refinements of Jensen's inequality. *J. Math. Anal. Appl.* **1992**, *168*, 518–522. [CrossRef]
54. Dragomir, S.S. Two mappings associated with Jensen's inequality. *Extr. Math.* **1993**, *8*, 102–105.
55. Micić, J.; Pečarić, J.; Jurica, P. Refined Jensen's operator inequality with condition on spectra. *Oper. Matrices* **2013**, *7*, 293–308.
56. Vasić, P.M.; Mijajlović, Ž. On an index set function connected with Jensen inequality. *Publikacije Elektrotehničkog Fakulteta Serija Matematika i Fizika* **1976**, *544/576*, 110–112.
57. Kian, M. Operator Jensen inequality for superquadratic functions. *Linear Algebra Appl.* **2014**, *456*, 82–87. [CrossRef]
58. Matković, A.; Pečarić, J.; Perić, I. A variant of Jensen's inequality of Mercer's type for operators with applications. *Linear Algebra Appl.* **2006**, *418*, 551–564. [CrossRef]
59. Deng, Y.; Ullah, H.; Adil Khan, M.; Iqbal, S.; Wu, S. Refinements of Jensen's inequality via majorization results with applications in the information theory. *J. Math.* **2021**, *2021*, 1951799. [CrossRef]
60. Saeed, T.; Adil Khan, M.; Ullah, H. Refinements of Jensen's inequality and applications. *AIMS Math.* **2022**, *7*, 5328–5346. [CrossRef]
61. Adil Khan, M.; Ullah, H.; Saeed, T. Some estimations of the Jensen difference andapplications. *Math. Meth. Appl. Sci.* **2022**, *46*, 5863–5892. [CrossRef]
62. Costarelli, D.; Spigler, R. How sharp is the Jensen inequality? *J. Inequal. Appl.* **2015**, *69*, 1–10. [CrossRef]

63. Zhu, X.L.; Yang, G.H. Jensen inequality approach to stability analysis of discrete-time systems with time-varying delay. *Proc. Amer. Control Conf.* **2008**, 1644–1649.
64. Adil Khan, M.; Khan, S.; Chu, Y. A new bound for the Jensen gap with applications in information theory. *IEEE Access* **2020**, *18*, 98001–98008. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.