

RESEARCH

Open Access



Estimation of divergence measures via weighted Jensen inequality on time scales

Iqrar Ansari^{1*}, Khuram Ali Khan¹, Ammara Nosheen², Đilda Pečarić³ and Josip Pečarić⁴

*Correspondence:

iqrar@math.qau.edu.pk

¹Department of Mathematics,
University of Sargodha, Sargodha
40100, Pakistan

Full list of author information is
available at the end of the article

Abstract

The main purpose of the presented paper is to obtain some time scale inequalities for different divergences and distances by using weighted time scales Jensen's inequality. These results offer new inequalities in h -discrete calculus and quantum calculus and extend some known results in the literature. The lower bounds of some divergence measures are also presented. Moreover, the obtained discrete results are given in the light of the Zipf–Mandelbrot law and the Zipf law.

Keywords: Time scales; Jensen's inequality; Csiszár divergence; Zipf and Zipf–Mandelbrot law

1 Introduction

Distance or divergence measures are of key importance in statistics and information theory. Depending upon the nature of the problem, different divergence measures are suitable. A number of measures of divergence that compare two probability distributions have been proposed (see [15, 16, 23, 24, 31, 37] and the references therein). Csiszár [12] introduced the f -divergence functional as follows.

Definition 1.1 Suppose that $f : \mathbb{R}^+ \rightarrow (0, \infty)$ is a convex function. Let $\tilde{\mathbf{r}} = (r_1, \dots, r_n)$ and $\tilde{\mathbf{s}} = (s_1, \dots, s_n)$ be such that $\sum_{k=1}^n r_k = 1$ and $\sum_{k=1}^n s_k = 1$. Then an f -divergence functional is stated as

$$I_f(\tilde{\mathbf{r}}, \tilde{\mathbf{s}}) := \sum_{k=1}^n s_k f\left(\frac{r_k}{s_k}\right),$$

where f bears the following requirements:

$$f(0) := \lim_{\epsilon \rightarrow 0^+} f(\epsilon); \quad 0f\left(\frac{0}{0}\right) := 0; \quad 0f\left(\frac{a}{0}\right) := \lim_{\epsilon \rightarrow 0^+} \epsilon f\left(\frac{a}{\epsilon}\right), \quad a > 0.$$

The Csiszár's f -divergence is a broad class of divergences which consists of various divergence measures used in finding out the difference between two probability densities. A significant property of Csiszár's f -divergence is that several well-known divergence measures

© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

can be deduced from this divergence measure by suitable substitutions to the convex function f . In recent years, several researchers have done a considerable work providing various kinds of bounds on the divergences and distances, see e.g. [13, 14, 25, 33]. Jensen's inequality has an important role in obtaining inequalities for divergence measures. It helps to compute useful upper bounds for several entropic measures used in information theory. In [18], Jain *et al.* established an information inequality regarding Csiszár f -divergence by utilizing the convexity condition and Jensen's inequality. This inequality is applied in comparing some well-known divergences which play a significant role in information theory. In [19], Khan *et al.* obtained new results for the Shannon and Zipf–Mandelbrot entropies. They also computed different bounds for these entropies by using some refinements of the Jensen inequality. In [21], the authors established various inequalities for convex functions and applied them to Csiszár divergence. They also obtained several results for Zipf–Mandelbrot entropy. In [27], Mehmood *et al.* obtained a new generalized form of cyclic refinements of Jensen's inequality from convex to higher order convex functions by utilizing Taylor's formula. They also computed bounds for various notable inequalities utilized in information theory. In [11], Butt *et al.* used discrete and continuous cyclic refinements of Jensen's inequality and extended them from convex to higher order convex function by using new Green functions and Abel–Gontscharoff interpolating polynomial. As an application, they established a connection between new entropic bounds for relative, Shannon, and Mandelbrot entropies. In [22], Khan *et al.* established an elegant refinement of Jensen's inequality related to two finite sequences. The obtained inequality used to compute bounds for Csiszár divergence, variational distance, Shannon entropy, and Zipf–Mandelbrot entropy. In [29], Pečarić *et al.* obtained refinements of the integral version of Jensen's inequality and the Lah–Ribarić inequality and deduced estimates for the integral form of Csiszár divergence and its important particular cases. In [2], Ahmad *et al.* utilized some results of Jensen's inequality for convex functions and obtained various estimates for Shannon and generalized Zipf–Mandelbrot entropies. In [10], Butt *et al.* proved various Jensen–Grüss type inequalities under certain conditions.

The development of the theory of time scales was initiated by Hilger in 1988. The books of Bohner and Peterson [8, 9] related to time scales are compact and resolve a lot of time scales calculus. In the past years, new developments in the theory and applications of dynamic derivatives on time scales emerged. Many results from the continuous case are carried over to the discrete one very easily, but some seem to be completely different. The study on time scales comes to reveal such discrepancies and to make us understand the difference between the two cases. The Jensen inequality has been extended to time scales by Agarwal *et al.* (see [1, 8]). Various classical inequalities and their converses for isotonic linear functionals on time scales are established in [5]. In [6], Anwar *et al.* gave the properties and applications of Jensen functionals on time scales for one variable. Further in [7], the authors obtained the Jensen inequality for several variables and deduced Jensen functionals. They also derived properties of Jensen functionals and applied them to generalized means. In recent years, the study of dynamic inequalities on time scales has been considered by several authors, see [1, 28, 30, 32, 36, 39, 40]. In [3], Ansari *et al.* obtained Shannon type inequalities on an arbitrary time scale. They also deduced bounds of differential entropy on time scale for various distributions. Further in [4], the authors established several inequalities for Csiszár f -divergence among two probability densities on

time scales. They also obtained new results for divergence measures in h -discrete calculus and quantum calculus.

Quantum calculus or q -calculus is usually called calculus without limits. In 1910, Jackson [17] described a q -analogue of derivative and integral operator along with their applications. He was the first to establish q -calculus in an organized form. It is important to note that quantum integral inequalities are more significant and constructive than their classical counterparts. It has been primarily for the reason that quantum integral inequalities can interpret the hereditary properties of the fact and technique under consideration. Recently, there has been a rapid development in q -calculus. Consequently, new generalizations of the classical approach of quantum calculus have been proposed and analyzed in various literature works. The concepts of quantum calculus on finite intervals were given by Tariboon and Ntouyas [34, 35], and they obtained certain q -analogues of classical mathematical objects, which motivated numerous researchers to explore the subject in detail. Subsequently, several new results related to the quantum counterpart of classical mathematical results have been established.

2 Preliminaries

An arbitrary nonempty closed subset of the real line is known as time scale $\mathbb{T} \subset \mathbb{R}$. The subsequent results and definitions are given in [8].

Definition 2.1 Suppose that \mathbb{T} is a time scale and $\zeta \in \mathbb{T}$, then the forward, respectively backward, jump operators $\sigma, \rho : \mathbb{T} \rightarrow \mathbb{T}$ are defined as follows:

$$\sigma(\zeta) = \inf\{v \in \mathbb{T} : v > \zeta\} \quad \text{and} \quad \rho(\zeta) = \sup\{v \in \mathbb{T} : v < \zeta\}.$$

Definition 2.2 Let \mathbb{T} be a time scale and $z : \mathbb{T} \rightarrow \mathbb{R}$ be a function, then z is known as rd-continuous or right-dense continuous if its left-sided limits exist (finite) at left-dense points in \mathbb{T} and it is continuous at right-dense points in \mathbb{T} . The set of rd-continuous functions $z : \mathbb{T} \rightarrow \mathbb{R}$ is usually denoted by C_{rd} .

Let us introduce the set \mathbb{T}^k as follows:

$$\mathbb{T}^k = \begin{cases} \mathbb{T} \setminus (\rho(\sup \mathbb{T}), \sup \mathbb{T}] & \text{if } \sup \mathbb{T} < \infty, \\ \mathbb{T} & \text{if } \sup \mathbb{T} = \infty. \end{cases}$$

Definition 2.3 Consider a function $z : \mathbb{T} \rightarrow \mathbb{R}$ and $\zeta \in \mathbb{T}^k$. Then we define $z^\Delta(\zeta)$ to be the number (when it exists) with the property that given any $\epsilon > 0$, there is a neighborhood U of ζ such that

$$|z(\sigma(\zeta)) - z(v) - z^\Delta(\zeta)(\sigma(\zeta) - v)| \leq \epsilon |\sigma(\zeta) - v| \quad \text{for all } v \in U.$$

In this case, z is said to be delta differentiable at ζ .

For $\mathbb{T} = \mathbb{R}$, z^Δ becomes ordinary derivative z' , while if $\mathbb{T} = \mathbb{Z}$, then z^Δ turns into the usual forward difference operator $\Delta z(\zeta) = z(\zeta + 1) - z(\zeta)$. If $\mathbb{T} = \overline{q^{\mathbb{Z}}} = \{q^n : n \in \mathbb{Z}\} \cup \{0\}$ is

the so-called q -difference operator, with $q > 1$, then

$$z^\Delta(\zeta) = \frac{z(q\zeta) - z(\zeta)}{(q-1)\zeta}, \quad z^\Delta(0) = \lim_{\nu \rightarrow 0} \frac{z(\nu) - z(0)}{\nu}.$$

Theorem 2.1 (Existence of antiderivatives) *Every rd-continuous function has an antiderivative. If $x_0 \in \mathbb{T}$, then F is defined by*

$$F(\zeta) := \int_{x_0}^x f(\zeta) \Delta\zeta \quad \text{for } x \in \mathbb{T}^k$$

is an antiderivative of f .

For $\mathbb{T} = \mathbb{R}$, we have $\int_a^b f(\zeta) \Delta\zeta = \int_a^b f(\zeta) d\zeta$, and if $\mathbb{T} = \mathbb{N}$, then $\int_a^b f(\zeta) \Delta\zeta = \sum_{\zeta=a}^{b-1} f(\zeta)$, where $a, b \in \mathbb{T}$ with $a \leq b$.

In [38], Wong et al. gave the weighted Jensen inequality on time scales which is stated as follows.

Theorem 2.2 *Assume that $I \subset \mathbb{R}$, and let $r \in C_{rd}([a, b]_{\mathbb{T}}, \mathbb{R})$ be a positive function with*

$$\int_a^b r(\zeta) \Delta\zeta > 0,$$

where $a, b \in \mathbb{T}$. If $f \in C(I, \mathbb{R})$ is convex and $g \in C_{rd}([a, b]_{\mathbb{T}}, I)$, then

$$f\left(\frac{\int_a^b r(\zeta)g(\zeta) \Delta\zeta}{\int_a^b r(\zeta) \Delta\zeta}\right) \leq \frac{\int_a^b r(\zeta)f(g(\zeta)) \Delta\zeta}{\int_a^b r(\zeta) \Delta\zeta}. \tag{1}$$

When f is a strictly convex function, the inequality sign in (1) is strict.

3 Divergences on time scales

Consider the set of rd-continuous functions on time scale \mathbb{T} to be

$$\Omega := \left\{ r \in C_{rd}([a, b]_{\mathbb{T}}, (0, \infty)), r(\zeta) > 0, \int_a^b r(\zeta) \Delta\zeta > 0 \right\}.$$

In the sequel, we assume that $r, s \in \Omega$ and the following integrals exist:

$$R = \int_a^b r(\zeta) \Delta\zeta \quad \text{and} \quad S = \int_a^b s(\zeta) \Delta\zeta.$$

3.1 Csiszár f -divergence

Csiszár f -divergence on time scale is defined in [4] as follows:

$$D_f(s, r) := \int_a^b r(\zeta) f\left(\frac{s(\zeta)}{r(\zeta)}\right) \Delta\zeta, \tag{2}$$

where f is convex on $(0, \infty)$.

Theorem 3.1 Assume that $I \subset \mathbb{R}$, and if $f \in C(I, \mathbb{R})$ is convex, then

$$Rf\left(\frac{S}{R}\right) \leq D_f(s, r), \tag{3}$$

where $D_f(s, r)$ is given in (2).

Proof Put $g(\zeta) = \frac{s(\zeta)}{r(\zeta)}$ in (1) to get (3). □

Example 3.1 For $\mathbb{T} = \mathbb{R}$, Theorem 3.1 becomes [20, Theorem 5.2 on p. 10].

Example 3.2 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.1 to get a lower bound for Csiszár divergence in h -discrete calculus

$$\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)hf\left(\frac{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h}{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h}\right) \leq \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)hf\left(\frac{s(lh)}{r(lh)}\right).$$

Remark 3.1 Choose $h = 1$ in Example 3.2, and let $a = 0$, $b = n$, $r(l) = r_j$, and $s(l) = s_j$ to get the discrete Csiszár divergence

$$\sum_{j=1}^n r_j f\left(\frac{\sum_{j=1}^n s_j}{\sum_{j=1}^n r_j}\right) \leq I_f(s, r), \tag{4}$$

where

$$I_f(\tilde{s}, \tilde{r}) = \sum_{j=1}^n r_j f\left(\frac{s_j}{r_j}\right), \tag{5}$$

$\tilde{s} = (s_1, \dots, s_n)$ and $\tilde{r} = (r_1, \dots, r_n)$.

Example 3.3 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.1 to have a new lower bound of the Csiszár divergence in quantum calculus

$$\sum_{l=0}^{n-1} q^{l+1}r(q^l)f\left(\frac{\sum_{l=0}^{n-1} q^{l+1}s(q^l)}{\sum_{l=0}^{n-1} q^{l+1}r(q^l)}\right) \leq \sum_{l=0}^{n-1} q^{l+1}r(q^l)f\left(\frac{s(q^l)}{r(q^l)}\right).$$

3.2 Differential entropy (continuous entropy)

Consider a positive density function r on time scale \mathbb{T} to a continuous random variable X with $\int_a^b r(\zeta)\Delta\zeta = 1$, wherever the integral exists.

In [3], Ansari *et al.* defined the so-called differential entropy on time scale by

$$h_{\bar{b}}(X) := \int_a^b r(\zeta) \log \frac{1}{r(\zeta)} \Delta\zeta, \tag{6}$$

where $\bar{b} > 1$ is the base of log. In the sequel, we assume that the base of log is greater than 1.

Theorem 3.2 Suppose that $r, s \in C_{rd}([a, b]_{\mathbb{T}}, \mathbb{R})$ are Δ -integrable functions and r is a positive probability density function with $S = \int_a^b s(\zeta) \Delta\zeta > 0$. If $f \in C(I, \mathbb{R})$ is convex and $\bar{b} > 1$, then

$$h_{\bar{b}}(X) \leq \int_a^b r(\zeta) \log \frac{1}{s(\zeta)} \Delta\zeta + \log(S), \tag{7}$$

where $h_{\bar{b}}(\zeta)$ is defined in (6) and $a, b \in \mathbb{T}$.

Proof The function $f(\zeta) = -\log \zeta$ is convex. Use $f(\zeta) = -\log \zeta$ with $\int_a^b r(\zeta) \Delta\zeta = 1$ in (3) to get

$$\begin{aligned} -\log(S) &\leq \int_a^b -r(\zeta) \log\left(\frac{s(\zeta)}{r(\zeta)}\right) \Delta\zeta, \\ &= \int_a^b (r(\zeta) \log r(\zeta) - r(\zeta) \log s(\zeta)) \Delta\zeta, \\ &= \int_a^b r(\zeta) \log r(\zeta) \Delta\zeta - \int_a^b r(\zeta) \log s(\zeta) \Delta\zeta \\ &= -\int_a^b r(\zeta) \log \frac{1}{r(\zeta)} \Delta\zeta + \int_a^b r(\zeta) \log \frac{1}{s(\zeta)} \Delta\zeta, \\ &= -h_{\bar{b}}(X) + \int_a^b r(\zeta) \log \frac{1}{s(\zeta)} \Delta\zeta, \end{aligned}$$

the stated result. □

Remark 3.2 The inequality in (7) holds in the opposite direction for the base of log less than 1.

Example 3.4 For $\mathbb{T} = \mathbb{R}$, Theorem 3.2 becomes [26, Theorem 21(a)].

Example 3.5 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.2 to get an upper bound for entropy in h -discrete calculus

$$\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h \log\left(\frac{1}{r(lh)h}\right) \leq \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h \log\left(\frac{1}{s(lh)h}\right) + \log\left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h\right). \tag{8}$$

Remark 3.3 Put $h = 1$ in (8) to get [26, Theorem 8 (i)].

Example 3.6 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.2 to have

$$\sum_{l=0}^{n-1} q^{l+1} r(q^l) \log\left(\frac{1}{r(q^l)}\right) \leq \sum_{l=0}^{n-1} q^{l+1} r(q^l) \log\left(\frac{1}{s(q^l)}\right) + \log\left(\sum_{l=0}^{n-1} q^{l+1} s(q^l)\right). \tag{9}$$

Remark 3.4 (9) contains Shannon entropy which is new in quantum calculus up to the knowledge of authors.

3.3 Karl Pearson χ^2 -divergence

The χ^2 -divergence on time scale is defined in [4] as follows:

$$D_{\chi^2}(s, r) := \int_a^b r(\zeta) \left[\left(\frac{s(\zeta)}{r(\zeta)} \right)^2 - 1 \right] \Delta\zeta. \tag{10}$$

Theorem 3.3 Assume the conditions of Theorem 3.1 to get

$$\frac{1}{R} [S^2 - R^2] \leq D_{\chi^2}(s, r), \tag{11}$$

where $D_{\chi^2}(s, r)$ is defined in (10).

Proof Consider $f(\zeta) = \zeta^2 - 1$ in (3) to obtain

$$\left(\frac{S}{R} \right)^2 - 1 \leq \frac{1}{R} \int_a^b r(\zeta) \left[\left(\frac{s(\zeta)}{r(\zeta)} \right)^2 - 1 \right] \Delta\zeta,$$

after simplification we get

$$S^2 - R^2 \leq R \int_a^b r(\zeta) \left[\left(\frac{s(\zeta)}{r(\zeta)} \right)^2 - 1 \right] \Delta\zeta,$$

the desired result. □

Example 3.7 If $\mathbb{T} = \mathbb{R}$, then (11) takes the form

$$\frac{1}{\int_a^b r(\zeta) d\zeta} \left[\left(\int_a^b s(\zeta) d\zeta \right)^2 - \left(\int_a^b r(\zeta) d\zeta \right)^2 \right] \leq \int_a^b r(\zeta) \left[\left(\frac{s(\zeta)}{r(\zeta)} \right)^2 - 1 \right] d\zeta.$$

Example 3.8 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.3 to get a new lower bound for χ^2 -divergence in h -discrete calculus

$$\frac{1}{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h} \left[\left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h \right)^2 - \left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h \right)^2 \right] \leq \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h \left[\left(\frac{s(lh)}{r(lh)} \right)^2 - 1 \right]. \tag{12}$$

Remark 3.5 Choose $h = 1$ in (12), let $a = 0$, $b = n$, $r(l) = r_j$, and $s(l) = s_j$ to get χ^2 -divergence

$$\frac{1}{\sum_{j=1}^n r_j} \left[\left(\sum_{j=1}^n s_j \right)^2 - \left(\sum_{j=1}^n r_j \right)^2 \right] \leq \chi^2(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}),$$

where

$$\chi^2(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}) = \sum_{j=1}^n r_j \left[\left(\frac{s_j}{r_j} \right)^2 - 1 \right]. \tag{13}$$

Example 3.9 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.3 to have a new lower bound for χ^2 -divergence in quantum calculus

$$\begin{aligned} & \frac{1}{\sum_{l=0}^{n-1} q^{l+1} r(q^l)} \left[\left(\sum_{l=0}^{n-1} q^{l+1} s(q^l) \right)^2 - \left(\sum_{l=0}^{n-1} q^{l+1} r(q^l) \right)^2 \right] \\ & \leq \sum_{l=0}^{n-1} q^{l+1} r(q^l) \left[\left(\frac{s(q^l)}{r(q^l)} \right)^2 - 1 \right]. \end{aligned} \tag{14}$$

3.4 Kullback–Leibler divergence

Kullback–Leibler divergence on time scale is defined in [4] as follows:

$$D(s, r) = \int_a^b s(\zeta) \ln \left[\frac{s(\zeta)}{r(\zeta)} \right] \Delta \zeta. \tag{15}$$

Theorem 3.4 Assume the conditions of Theorem 3.1, then we have

$$S \ln \left(\frac{S}{R} \right) \leq D(s, r), \tag{16}$$

where $D(s, r)$ is defined in (15).

Proof Consider $f(\zeta) = \zeta \ln \zeta$ in (3) to get

$$\frac{S}{R} \ln \left(\frac{S}{R} \right) \leq \frac{1}{R} \int_a^b s(\zeta) \ln \left(\frac{s(\zeta)}{r(\zeta)} \right) \Delta \zeta,$$

or we have

$$S \ln \left(\frac{S}{R} \right) \leq \int_a^b s(\zeta) \ln \left(\frac{s(\zeta)}{r(\zeta)} \right) \Delta \zeta,$$

the desired result. □

Example 3.10 For $\mathbb{T} = \mathbb{R}$, (16) becomes

$$\int_a^b s(\zeta) d\zeta \ln \left(\frac{\int_a^b s(\zeta) d\zeta}{\int_a^b r(\zeta) d\zeta} \right) \leq \int_a^b s(\zeta) \ln \left(\frac{s(\zeta)}{r(\zeta)} \right) d\zeta.$$

Example 3.11 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.4 to get a new lower bound in h -discrete calculus

$$\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h \ln \left(\frac{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h}{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h} \right) \leq \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h \ln \left(\frac{s(lh)}{r(lh)} \right). \tag{17}$$

Remark 3.6 Choose $h = 1$ in (17), let $a = 0$, $b = n$, $r(l) = r_j$, and $s(l) = s_j$ to get the discrete Kullback–Leibler divergence

$$\sum_{j=1}^n s_j \ln \left(\frac{\sum_{j=1}^n s_j}{\sum_{j=1}^n r_j} \right) \leq \text{KL}(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}),$$

where

$$KL(\tilde{s}, \tilde{r}) = \sum_{j=1}^n s_j \ln\left(\frac{s_j}{r_j}\right). \tag{18}$$

Example 3.12 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.4 to have a new lower bound in quantum calculus

$$\sum_{l=0}^{n-1} q^{l+1} s(q^l) \ln\left(\frac{\sum_{l=0}^{n-1} q^{l+1} s(q^l)}{\sum_{l=0}^{n-1} q^{l+1} r(q^l)}\right) \leq \sum_{l=0}^{n-1} q^{l+1} s(q^l) \ln\left(\frac{s(q^l)}{r(q^l)}\right).$$

3.5 Hellinger discrimination

Hellinger discrimination on time scale is defined in [4] as follows:

$$h^2(s, r) = \frac{1}{2} \int_a^b [\sqrt{s(\zeta)} - \sqrt{r(\zeta)}]^2 \Delta\zeta. \tag{19}$$

Theorem 3.5 Assume the conditions of Theorem 3.1 to obtain

$$\frac{1}{2}(\sqrt{S} - \sqrt{R})^2 \leq h^2(s, r), \tag{20}$$

where $h^2(s, r)$ is defined in (19).

Proof Consider $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ in (3) to get

$$\frac{1}{2}\left(\sqrt{\frac{S}{R}} - 1\right)^2 \leq \frac{1}{2R} \int_a^b r(\zeta) \left(\sqrt{\frac{s(\zeta)}{r(\zeta)}} - 1\right)^2 \Delta\zeta, \tag{21}$$

after simplification we obtain

$$\frac{1}{2}(\sqrt{S} - \sqrt{R})^2 \leq \frac{1}{2} \int_a^b (\sqrt{s(\zeta)} - \sqrt{r(\zeta)})^2 \Delta\zeta,$$

the desired result. □

Example 3.13 For $\mathbb{T} = \mathbb{R}$, (20) becomes

$$\frac{1}{2}\left(\left[\int_a^b s(\zeta) d\zeta\right]^{\frac{1}{2}} - \left[\int_a^b r(\zeta) d\zeta\right]^{\frac{1}{2}}\right)^2 \leq \frac{1}{2} \int_a^b (\sqrt{s(\zeta)} - \sqrt{r(\zeta)})^2 d\zeta.$$

Example 3.14 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.5 to get a new lower for Hellinger discrimination in h -discrete calculus

$$\frac{1}{2}\left[\left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h\right)^{\frac{1}{2}} - \left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h\right)^{\frac{1}{2}}\right]^2 \leq \frac{1}{2} \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} (\sqrt{s(lh)h} - \sqrt{r(lh)h})^2. \tag{22}$$

Remark 3.7 Choose $h = 1$ in (22), let $a = 0, b = n, r(l) = r_j,$ and $s(l) = s_j$ to get the Hellinger distance

$$\frac{1}{2} \left[\left(\sum_{j=1}^n s_j \right)^{\frac{1}{2}} - \left(\sum_{j=1}^n r_j \right)^{\frac{1}{2}} \right]^2 \leq h^2(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}),$$

where

$$h^2(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}) = \frac{1}{2} \sum_{j=1}^n (\sqrt{s_j} - \sqrt{r_j})^2. \tag{23}$$

Example 3.15 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.5 to have a new lower for Hellinger discrimination in quantum calculus

$$\begin{aligned} & \frac{1}{2} \left[\left(\sum_{l=0}^{n-1} q^{l+1} s(q^l) \right)^{\frac{1}{2}} - \left(\sum_{l=0}^{n-1} q^{l+1} r(q^l) \right)^{\frac{1}{2}} \right]^2 \\ & \leq \frac{1}{2} \sum_{k=0}^{n-1} q^{k+1} [\sqrt{s(q^k)} - \sqrt{r(q^k)}]^2. \end{aligned} \tag{24}$$

3.6 Bhattacharyya coefficient

The Bhattacharyya coefficient on time scale is defined in [4] as follows:

$$D_B(s, r) = \int_a^b \sqrt{r(\zeta)s(\zeta)} \Delta\zeta. \tag{25}$$

Theorem 3.6 Assume the conditions of Theorem 3.1 to get

$$D_B(s, r) \leq \sqrt{RS}, \tag{26}$$

where $D_B(s, r)$ is defined in (25).

Proof Consider $f(\zeta) = -\sqrt{\zeta}$ in (3) to get

$$-\sqrt{\frac{S}{R}} \leq \frac{-1}{R} \int_a^b \sqrt{r(\zeta)s(\zeta)} \Delta\zeta,$$

after simplification we obtain

$$\int_a^b \sqrt{r(\zeta)s(\zeta)} \Delta\zeta \leq \sqrt{RS},$$

the desired result. □

Example 3.16 If $\mathbb{T} = \mathbb{R}$, then (26) takes the form

$$\int_a^b (r(\zeta)s(\zeta))^{\frac{1}{2}} d\zeta \leq \left(\int_a^b r(\zeta) d\zeta \int_a^b s(\zeta) d\zeta \right)^{\frac{1}{2}}.$$

Example 3.17 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.6 to get a new upper bound for the Bhattacharyya coefficient in h -discrete calculus

$$\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} (r(lh)hs(lh)h)^{\frac{1}{2}} \leq \left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h \right)^{\frac{1}{2}}. \tag{27}$$

Remark 3.8 Choose $h = 1$ in (27), let $a = 0$, $b = n$, $r(l) = r_j$, and $s(l) = s_j$ to get the Bhattacharyya coefficient

$$B(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}) \leq \left(\sum_{j=1}^n r_j \sum_{j=1}^n s_j \right)^{\frac{1}{2}},$$

where

$$B(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}) = \sum_{j=1}^n \sqrt{r_j s_j}. \tag{28}$$

Example 3.18 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.6 to have a new upper bound for the Bhattacharyya coefficient in quantum calculus

$$\sum_{l=0}^{n-1} q^{l+1} [r(q^l)s(q^l)]^{\frac{1}{2}} \leq \left(\sum_{l=0}^{n-1} q^{l+1} r(q^l) \sum_{l=0}^{n-1} q^{l+1} s(q^l) \right)^{\frac{1}{2}}.$$

3.7 Jeffreys distance

Jeffreys distance on time scale is defined in [4] as follows:

$$D_J(s, r) = \int_a^b (s(\zeta) - r(\zeta)) \ln \left[\frac{s(\zeta)}{r(\zeta)} \right] \Delta \zeta. \tag{29}$$

Theorem 3.7 Assume the conditions of Theorem 3.1 to get

$$(S - R) \ln \left(\frac{S}{R} \right) \leq D_J(s, r), \tag{30}$$

where $D_J(s, r)$ is defined in (29).

Proof Consider $f(\zeta) = (\zeta - 1) \ln \zeta$ in (3) to get

$$R \left(\frac{S}{R} - 1 \right) \ln \left(\frac{S}{R} \right) \leq \int_a^b r(\zeta) \left(\frac{s(\zeta)}{r(\zeta)} - 1 \right) \ln \left(\frac{s(\zeta)}{r(\zeta)} \right) \Delta \zeta,$$

or we have

$$(S - R) \ln \left(\frac{S}{R} \right) \leq \int_a^b (s(\zeta) - r(\zeta)) \ln \left(\frac{s(\zeta)}{r(\zeta)} \right) \Delta \zeta,$$

the desired result. □

Example 3.19 For $\mathbb{T} = \mathbb{R}$, (30) takes the form

$$\left(\int_a^b s(\zeta) d\zeta - \int_a^b r(\zeta) d\zeta \right) \ln \left(\frac{\int_a^b s(\zeta) d\zeta}{\int_a^b r(\zeta) d\zeta} \right) \leq \int_a^b [s(\zeta) - r(\zeta)] \ln \left(\frac{s(\zeta)}{r(\zeta)} \right) d\zeta.$$

Example 3.20 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.7 to get a new lower bound for Jeffreys distance in h -discrete calculus

$$\begin{aligned} & \left(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h - \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h \right) \ln \left(\frac{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h}{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h} \right) \\ & \leq \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} (s(lh)h - r(lh)h) \ln \left(\frac{s(lh)}{r(lh)} \right). \end{aligned} \tag{31}$$

Remark 3.9 Choose $h = 1$ in (31), let $a = 0$, $b = n$, $r(l) = r_j$, and $s(l) = s_j$ to get Jeffreys distance

$$\left(\sum_{j=1}^n s_j - \sum_{j=1}^n r_j \right) \ln \left(\frac{\sum_{j=1}^n s_j}{\sum_{j=1}^n r_j} \right) \leq D_J(\mathfrak{s}, \mathfrak{r}),$$

where

$$D_J(\mathfrak{s}, \mathfrak{r}) = \sum_{j=1}^n (s_j - r_j) \ln \left(\frac{s_j}{r_j} \right). \tag{32}$$

Example 3.21 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.7 to have a new lower bound for the Jeffreys distance in quantum calculus

$$\begin{aligned} & \left(\sum_{l=0}^{n-1} q^{l+1} s(q^l) - \sum_{l=0}^{n-1} q^{l+1} r(q^l) \right) \ln \left(\frac{\sum_{l=0}^{n-1} q^{l+1} s(q^l)}{\sum_{l=0}^{n-1} q^{l+1} r(q^l)} \right) \\ & \leq \sum_{l=0}^{n-1} q^{l+1} (s(q^l) - r(q^l)) \ln \left(\frac{s(q^l)}{r(q^l)} \right). \end{aligned}$$

3.8 Triangular discrimination

Triangular discrimination on time scale is defined in [4] as follows:

$$D_{\Delta}(r, s) = \int_a^b \frac{[s(\zeta) - r(\zeta)]^2}{s(\zeta) + r(\zeta)} \Delta\zeta. \tag{33}$$

Theorem 3.8 Assume the conditions of Theorem 3.1 to obtain

$$\frac{[S - R]^2}{S + R} \leq D_{\Delta}(r, s), \tag{34}$$

where $D_{\Delta}(r, s)$ is defined in (33).

Proof Consider $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ in (3) to get

$$R \frac{(\frac{S}{R} - 1)^2}{\frac{S}{R} + 1} \leq \int_a^b r(\zeta) \frac{(\frac{s(\zeta)}{r(\zeta)} - 1)^2}{\frac{s(\zeta)}{r(\zeta)} + 1} \Delta\zeta$$

or

$$\frac{[S - R]^2}{S + R} \leq \int_a^b \frac{[s(\zeta) - r(\zeta)]^2}{s(\zeta) + r(\zeta)} \Delta\zeta. \quad \square$$

Example 3.22 For $\mathbb{T} = \mathbb{R}$, (34) becomes

$$\frac{[\int_a^b s(\zeta) d\zeta - \int_a^b r(\zeta) d\zeta]^2}{\int_a^b s(\zeta) d\zeta + \int_a^b r(\zeta) d\zeta} \leq \int_a^b \frac{[s(\zeta) - r(\zeta)]^2}{s(\zeta) + r(\zeta)} d\zeta.$$

Example 3.23 Choose $\mathbb{T} = h\mathbb{Z}$, $h > 0$ in Theorem 3.8 to get a new lower bound for the triangular discrimination in h -discrete calculus

$$\frac{(\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h - \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h)^2}{\sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} s(lh)h + \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} r(lh)h} \leq \sum_{l=\frac{a}{h}}^{\frac{b}{h}-1} \frac{h(s(lh) - r(lh))^2}{s(lh) + r(lh)}. \quad (35)$$

Remark 3.10 Choose $h = 1$ in (35), let $a = 0$, $b = n$, $r(l) = r_j$, and $s(l) = s_j$ to get the triangular discrimination

$$\frac{(\sum_{j=1}^n s_j - \sum_{j=1}^n r_j)^2}{\sum_{j=1}^n s_j + \sum_{j=1}^n r_j} \leq \Delta(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}),$$

where

$$\Delta(\tilde{\mathbf{s}}, \tilde{\mathbf{r}}) = \sum_{j=1}^n \frac{(s_j - r_j)^2}{s_j + r_j}. \quad (36)$$

Example 3.24 Choose $\mathbb{T} = q^{\mathbb{N}_0}$ ($q > 1$) in Theorem 3.8 to have a new lower bound for the triangular discrimination in quantum calculus

$$\frac{(\sum_{l=0}^{n-1} q^{l+1} s(q^l) - \sum_{l=0}^{n-1} q^{l+1} r(q^l))^2}{\sum_{l=0}^{n-1} q^{l+1} s(q^l) + \sum_{l=0}^{n-1} q^{l+1} r(q^l)} \leq \sum_{l=0}^{n-1} q^{l+1} \frac{[s(q^l) - r(q^l)]^2}{s(q^l) + r(q^l)}.$$

4 Zipf–Mandelbrot law

The Zipf–Mandelbrot law is a discrete probability distribution and is defined via a probability mass function which is given as follows:

$$f(j; N, a, b) = \frac{1}{(j + b)^a H_{N,a,b}}, \quad j = 1, \dots, N, \quad (37)$$

where

$$H_{N,a,b} = \sum_{i=1}^N \frac{1}{(i + b)^a} \quad (38)$$

is a generalization of a harmonic number and $N \in \{1, 2, \dots\}$, $a > 0$ and $b \in [0, \infty)$ are parameters.

If $b = 0$ and N is finite, then the Zipf–Mandelbrot law is commonly known as the Zipf law. By expression (37), the probability mass function in connection with the Zipf law is

$$f(j; N, a) = \frac{1}{(j)^a H_{N,a}}, \quad j = 1, \dots, N, \tag{39}$$

where

$$H_{N,a} = \sum_{i=1}^N \frac{1}{(i)^a}. \tag{40}$$

Using $r_j = f(j, N, a, b)$ in (37) as a probability mass function, we observe the obtained results via the Zipf–Mandelbrot law.

For this reason, we give results concerning the Csiszár functional $\tilde{I}_f(\tilde{\mathbf{s}}, \tilde{\mathbf{r}})$ for the Zipf–Mandelbrot law.

Case-1 Define $\tilde{\mathbf{r}}$ by (37) as a Zipf–Mandelbrot law N -tuple, Csiszár functional (5) becomes

$$\tilde{I}_f(j, N, a_2, b_2, \tilde{\mathbf{s}}) = \sum_{j=1}^N \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} f(s_j(j + b_2)^{a_2} H_{N,a_2,b_2}), \tag{41}$$

where $f : I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$, and $N \in \mathbb{N}$, $a_2 > 0$, $b_2 > 0$ are such that $s_j(j + b_2)^{a_2} H_{N,a_2,b_2} \in I$, $j = 1, \dots, N$.

Case-2 When $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{r}}$ both are defined via the Zipf–Mandelbrot law for N -tuples:

$$\tilde{I}_f(j, N, a_1, a_2, b_1, b_2) = \sum_{j=1}^N \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} f\left(\frac{(j + b_2)^{a_2} H_{N,a_2,b_2}}{(j + b_1)^{a_1} H_{N,a_1,b_1}}\right), \tag{42}$$

where $f : I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$, and $N \in \mathbb{N}$, $a_1, a_2 > 0$, $b_1, b_2 > 0$ are such that $\frac{(j + b_2)^{a_2} H_{N,a_2,b_2}}{(j + b_1)^{a_1} H_{N,a_1,b_1}} \in I$, $j = 1, \dots, N$.

Case-3 If $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{r}}$ both are defined as the Zipf law for N -tuples, then Csiszár functional (5) becomes

$$\tilde{I}_f(j, N, a_1, a_2) = \sum_{j=1}^N \frac{1}{j^{a_2} H_{N,a_2}} f\left(j^{a_2 - a_1} \frac{H_{N,a_2}}{H_{N,a_1}}\right). \tag{43}$$

Start from case-1 which is for the single Zipf–Mandelbrot law $r_j, j = 1, \dots, N$.

Corollary 4.1 Assume that $I \subset \mathbb{R}$, and let $N \in \mathbb{N}$, $a_2 > 0$, $b_2 > 0$ be such that $\sum_{j=1}^N s_j(j + b_2)^{a_2} H_{N,a_2,b_2} \in I$ for $j = 1, \dots, N$. If f is a convex function, then

$$\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} f\left(\frac{\sum_{j=1}^n s_j}{\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}}}\right) \leq \tilde{I}_f(j, N, a_2, b_2, \tilde{\mathbf{s}}). \tag{44}$$

Proof Put $r_j = \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}}$ for $j = 1, \dots, N$ in (4) to get (44), where $\tilde{I}_f(j, N, a_2, b_2, \tilde{\mathbf{s}})$ is defined in (41). □

Remark 4.1 The inequality sign in (44) holds in reverse direction when f is a concave function.

The next result is for case-2 as both s_j and r_j are defined by the Zipf–Mandelbrot law.

Corollary 4.2 *Assume that $I \subset \mathbb{R}$, and let $N \in \mathbb{N}$, $a_1, a_2 > 0$, $b_1, b_2 > 0$ be such that $\frac{(j+b_2)^{a_2} H_{N,a_2,b_2}}{\sum_{j=1}^N (j+b_1)^{a_1} H_{N,a_1,b_1}} \in I$ for $j = 1, \dots, N$. If f is a convex function, then*

$$\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} f\left(\frac{\sum_{j=1}^n \frac{1}{(j+b_1)^{a_1} H_{N,a_1,b_1}}}{\sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}}}\right) \leq \tilde{I}_f(j, N, a_1, a_2, b_1, b_2). \tag{45}$$

Proof Using $r_j = \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}}$ and $s_j = \frac{1}{(j+b_1)^{a_1} H_{N,a_1,b_1}}$ for $j = 1, \dots, N$, in (4), we get (45), where $\tilde{I}_f(j, N, a_1, a_2, b_1, b_2)$ is defined in (42). □

Remark 4.2 The inequality sign in (45) holds in reverse direction when f is a concave function.

The next result is for case-3 as both s_j and r_j are defined by the Zipf law.

Corollary 4.3 *Assume that $I \subset \mathbb{R}$, and let $N \in \mathbb{N}$, $a_1, a_2 > 0$ be such that $\frac{(j)^{a_2} H_{N,a_2}}{(j)^{a_1} H_{N,a_1}} \in I$ for $j = 1, \dots, N$. If f is a convex function, then*

$$\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}} f\left(\frac{\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}}}{\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}}}\right) \leq \tilde{I}_f(j, N, a_1, a_2). \tag{46}$$

Proof Using $r_j = \frac{1}{(j)^{a_2} H_{N,a_2}}$ and $s_j = \frac{1}{(j)^{a_1} H_{N,a_1}}$ for $j = 1, \dots, N$ in (4), we get (46), where $\tilde{I}_f(j, N, a_1, a_2)$ is defined in (43). □

Remark 4.3 The inequality sign in (46) holds in reverse direction when f is a concave function.

To give certain results related to the particular cases of f -divergences, we begin with the well-known Kullback–Leibler divergence (18).

Corollary 4.4 *Let $N \in \mathbb{N}$ and $a_2 > 0$, $b_2 > 0$. Then*

$$\sum_{j=1}^n s_j \ln\left(\frac{\sum_{j=1}^n s_j}{\sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}}}\right) \leq \tilde{\text{KL}}(j, N, a_2, b_2, \tilde{\mathbf{s}}). \tag{47}$$

Proof The function $f(\zeta) = \zeta \ln(\zeta)$ is convex. Use $f(\zeta) = \zeta \ln(\zeta)$ in (44) to obtain (47), where

$$\tilde{\text{KL}}(j, N, a_2, b_2, \tilde{\mathbf{s}}) = \sum_{j=1}^N s_j \ln(s_j (j + b_2)^{a_2} H_{N,a_2,b_2}). \tag{47}$$
□

If s_j and r_j are defined by the Zipf–Mandelbrot law.

Corollary 4.5 *Suppose that $N \in \mathbb{N}$ and $a_1, a_2 > 0, b_1, b_2 > 0$. Then*

$$\sum_{j=1}^n \frac{1}{(j + b_1)^{a_1} H_{N,a_1,b_1}} \ln \left(\frac{\sum_{j=1}^n \frac{1}{(j+b_1)^{a_1} H_{N,a_1,b_1}}}{\sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}}} \right) \leq \tilde{K}L(j, N, a_1, a_2, b_1, b_2). \tag{48}$$

Proof The function $f(\zeta) = \zeta \ln(\zeta)$ is convex. Use $f(\zeta) = \zeta \ln(\zeta)$ in (45) to get (48), where

$$\tilde{K}L(j, N, a_1, a_2, b_1, b_2) = \sum_{j=1}^N \frac{1}{(j + b_1)^{a_1} H_{N,a_1,b_1}} \ln \left(\frac{(j + b_2)^{a_2} H_{N,a_2,b_2}}{(j + b_1)^{a_1} H_{N,a_1,b_1}} \right). \quad \square$$

The following result holds as both s_j and r_j are defined by the Zipf law.

Corollary 4.6 *Let $N \in \mathbb{N}, a_1, a_2 > 0$. Then*

$$\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}} \ln \left(\frac{\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}}}{\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}}} \right) \leq \tilde{K}L(j, N, a_1, a_2). \tag{49}$$

Proof The function $f(\zeta) = \zeta \ln(\zeta)$ is convex. Use $f(\zeta) = \zeta \ln(\zeta)$ in (46) to have (49), where

$$\tilde{K}L(j, N, a_1, a_2) = \sum_{j=1}^N \frac{1}{j^{a_1} H_{N,a_1}} \ln \left(j^{a_2-a_1} \frac{H_{N,a_2}}{H_{N,a_1}} \right). \quad \square$$

Analogous results for the Hellinger distance (23) are given as follows.

Corollary 4.7 *Let $N \in \mathbb{N}, a_2 > 0, b_2 > 0$. Then*

$$\frac{1}{2} \left[\left(\sum_{j=1}^n s_j \right)^{\frac{1}{2}} - \left(\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} \right)^{\frac{1}{2}} \right]^2 \leq \tilde{h}^2(j, N, a_2, b_2, \tilde{\mathbf{s}}). \tag{50}$$

Proof Since $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ is a convex function, therefore we use $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ in (44) to get (50), where

$$\tilde{h}^2(j, N, a_2, b_2, \tilde{\mathbf{s}}) = \sum_{j=1}^N \frac{1}{2} \left[(s_j)^{\frac{1}{2}} - \left(\frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} \right)^{\frac{1}{2}} \right]^2. \quad \square$$

The following result holds as both s_j and r_j are defined by the Zipf–Mandelbrot law.

Corollary 4.8 *Let $N \in \mathbb{N}, a_1, a_2 > 0, b_1, b_2 > 0$. Then*

$$\frac{1}{2} \left[\left(\sum_{j=1}^n \frac{1}{(j + b_1)^{a_1} H_{N,a_1,b_1}} \right)^{\frac{1}{2}} - \left(\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} \right)^{\frac{1}{2}} \right]^2 \leq \tilde{h}^2(j, N, a_1, a_2, b_1, b_2). \tag{51}$$

Proof Since $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ is a convex function, therefore we use $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ in (45) to get (51), where

$$\tilde{h}^2(j, N, a_1, a_2, b_1, b_2) = \frac{1}{2} \sum_{j=1}^N \left[\left(\frac{1}{(j + b_1)^{a_1} H_{N, a_1, b_1}} \right)^{\frac{1}{2}} - \left(\frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right)^{\frac{1}{2}} \right]^2. \quad \square$$

The following result holds as both s_j and r_j are defined by the Zipf law.

Corollary 4.9 *Let $N \in \mathbb{N}$, $a_1, a_2 > 0$. Then*

$$\frac{1}{2} \left[\left(\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N, a_1}} \right)^{\frac{1}{2}} - \left(\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N, a_2}} \right)^{\frac{1}{2}} \right]^2 \leq \tilde{h}^2(j, N, a_1, a_2). \quad (52)$$

Proof Since $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ is a convex function, therefore we use $f(\zeta) = \frac{1}{2}(\sqrt{\zeta} - 1)^2$ in (46) to get (52), where

$$\tilde{h}^2(j, N, a_1, a_2) = \sum_{j=1}^N \frac{1}{2} \left[\left(\frac{1}{(j)^{a_1} H_{N, a_1}} \right)^{\frac{1}{2}} - \left(\frac{1}{(j)^{a_2} H_{N, a_2}} \right)^{\frac{1}{2}} \right]^2. \quad \square$$

Similarly, corresponding results for the Karl Pearson divergence (13) and the Jeffrey distance (32) are given below.

Corollary 4.10 *Let $N \in \mathbb{N}$ and $a_2 > 0, b_2 > 0$. Then*

$$\frac{1}{\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}}} \left[\left(\sum_{j=1}^n s_j \right)^2 - \left(\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right)^2 \right] \leq \tilde{\chi}^2(j, N, a_2, b_2, \mathfrak{s}). \quad (53)$$

Proof Since $f(\zeta) = \zeta^2 - 1$ is a convex function, therefore we use $f(\zeta) = \zeta^2 - 1$ in (44) to obtain (53), where

$$\tilde{\chi}^2(j, N, a_2, b_2, \mathfrak{s}) = \sum_{j=1}^N \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \left[(s_j (j + b_2)^{a_2} H_{N, a_2, b_2})^2 - 1 \right]. \quad \square$$

If s_j and r_j are defined via the Zipf–Mandelbrot law.

Corollary 4.11 *Suppose that $N \in \mathbb{N}$ and $a_1, a_2 > 0, b_1, b_2 > 0$. Then*

$$\frac{1}{\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}}} \left[\left(\sum_{j=1}^n \frac{1}{(j + b_1)^{a_1} H_{N, a_1, b_1}} \right)^2 - \left(\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right)^2 \right] \leq \tilde{\chi}^2(j, N, a_1, a_2, b_1, b_2). \quad (54)$$

Proof Since $f(\zeta) = \zeta^2 - 1$ is a convex function, therefore we use $f(\zeta) = \zeta^2 - 1$ in (45) to get (54), where

$$\tilde{\chi}^2(j, N, a_1, a_2, b_1, b_2) = \sum_{j=1}^N \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \left[\left(\frac{(j + b_2)^{a_2} H_{N, a_2, b_2}}{(j + b_1)^{a_1} H_{N, a_1, b_1}} \right)^2 - 1 \right]. \quad \square$$

The following result holds as both s_j and r_j are defined by the Zipf law.

Corollary 4.12 *Let $N \in \mathbb{N}$, $a_1, a_2 > 0$. Then*

$$\frac{1}{\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N, a_2}}} \left[\left(\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N, a_1}} \right)^2 - \left(\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N, a_2}} \right)^2 \right] \leq \tilde{\chi}^2(j, N, a_1, a_2). \quad (55)$$

Proof Since $f(\zeta) = \zeta^2 - 1$ is a convex function, therefore we use $f(\zeta) = \zeta^2 - 1$ in (46) to have (55), where

$$\tilde{\chi}^2(j, N, a_1, a_2) = \sum_{j=1}^N \frac{1}{j^{a_2} H_{N, a_2}} \left[\left(j^{a_2 - a_1} \frac{H_{N, a_2}}{H_{N, a_1}} \right)^2 - 1 \right]. \quad \square$$

Corollary 4.13 *Let $N \in \mathbb{N}$ and $a_2 > 0, b_2 > 0$. Then*

$$\left(\sum_{j=1}^n s_j - \sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right) \ln \left(\frac{\sum_{j=1}^n s_j}{\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}}} \right) \leq \tilde{D}_J(j, N, a_2, b_2, \tilde{\mathbf{s}}). \quad (56)$$

Proof The function $f(\zeta) = (\zeta - 1) \ln(\zeta)$ is convex. Use $f(\zeta) = (\zeta - 1) \ln(\zeta)$ in (44) to obtain (56), where

$$\tilde{D}_J(j, N, a_2, b_2, \tilde{\mathbf{s}}) = \sum_{j=1}^N \left(s_j - \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right) \ln(s_j (j + b_2)^{a_2} H_{N, a_2, b_2}). \quad \square$$

If s_j and r_j are defined via the Zipf–Mandelbrot law.

Corollary 4.14 *Suppose that $N \in \mathbb{N}$ and $a_1, a_2 > 0, b_1, b_2 > 0$. Then*

$$\left(\sum_{j=1}^n \frac{1}{(j + b_1)^{a_1} H_{N, a_1, b_1}} - \sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right) \ln \left(\frac{\sum_{j=1}^n \frac{1}{(j + b_1)^{a_1} H_{N, a_1, b_1}}}{\sum_{j=1}^n \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}}} \right) \leq \tilde{D}_J(j, N, a_1, a_2, b_1, b_2). \quad (57)$$

Proof The function $f(\zeta) = (\zeta - 1) \ln(\zeta)$ is convex. Use $f(\zeta) = (\zeta - 1) \ln(\zeta)$ in (45) to get (57), where

$$\begin{aligned} \tilde{D}_J(j, N, a_1, a_2, b_1, b_2) &= \sum_{j=1}^N \left(\frac{1}{(j + b_1)^{a_1} H_{N, a_1, b_1}} - \frac{1}{(j + b_2)^{a_2} H_{N, a_2, b_2}} \right) \\ &\quad \times \ln \left(\frac{(j + b_2)^{a_2} H_{N, a_2, b_2}}{(j + b_1)^{a_1} H_{N, a_1, b_1}} \right). \quad \square \end{aligned}$$

The following result holds as both s_j and r_j are defined by the Zipf law.

Corollary 4.15 *Let $N \in \mathbb{N}$, $a_1, a_2 > 0$. Then*

$$\left(\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}} - \sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}} \right) \ln \left(\frac{\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}}}{\sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}}} \right) \leq \tilde{D}_J(j, N, a_1, a_2). \tag{58}$$

Proof The function $f(\zeta) = (\zeta - 1) \ln(\zeta)$ is convex. Use $f(\zeta) = (\zeta - 1) \ln(\zeta)$ in (46) to have (58), where

$$\tilde{D}_J(j, N, a_1, a_2) = \sum_{j=1}^N \left(\frac{1}{j^{a_1} H_{N,a_1}} - \frac{1}{j^{a_2} H_{N,a_2}} \right) \ln \left(j^{a_2 - a_1} \frac{H_{N,a_2}}{H_{N,a_1}} \right). \quad \square$$

In addition to all, similar findings for triangular discrimination are given as follows.

Corollary 4.16 *Let $N \in \mathbb{N}$, $a_2 > 0$, $b_2 > 0$. Then*

$$\frac{(\sum_{j=1}^n s_j - \sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}})^2}{\sum_{j=1}^n s_j + \sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}}} \leq \tilde{\Delta}(j, N, a_2, b_2, \tilde{s}). \tag{59}$$

Proof Since $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ is a convex function, therefore we use $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ in (44) to obtain (59), where

$$\tilde{\Delta}(j, N, a_2, b_2, \tilde{s}) = \sum_{j=1}^N \frac{1}{(j + b_2)^{a_2} H_{N,a_2,b_2}} \frac{[s_j(j + b_2)^{a_2} H_{N,a_2,b_2} - 1]^2}{s_j(j + b_2)^{a_2} H_{N,a_2,b_2} + 1}. \quad \square$$

If s_j and r_j are defined via the Zipf–Mandelbrot law.

Corollary 4.17 *Suppose that $N \in \mathbb{N}$ and $a_1, a_2 > 0$, $b_1, b_2 > 0$. Then*

$$\frac{(\sum_{j=1}^n \frac{1}{(j+b_1)^{a_1} H_{N,a_1,b_1}} - \sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}})^2}{\sum_{j=1}^n \frac{1}{(j+b_1)^{a_1} H_{N,a_1,b_1}} + \sum_{j=1}^n \frac{1}{(j+b_2)^{a_2} H_{N,a_2,b_2}}} \leq \tilde{\Delta}(j, N, a_1, a_2, b_1, b_2). \tag{60}$$

Proof Since $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ is a convex function, therefore we use $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ in (45) to get (60), where

$$\tilde{\Delta}(j, N, a_1, a_2, b_1, b_2) = \sum_{j=1}^N \frac{[(j + b_2)^{a_2} H_{N,a_2,b_2} - (j + b_1)^{a_1} H_{N,a_1,b_1}]^2}{(j + b_2)^{a_2} H_{N,a_2,b_2} + (j + b_1)^{a_1} H_{N,a_1,b_1}}. \quad \square$$

The following result holds when both s_j and r_j are defined via the Zipf law.

Corollary 4.18 *Let $N \in \mathbb{N}$, $a_1, a_2 > 0$. Then*

$$\frac{(\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}} - \sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}})^2}{\sum_{j=1}^n \frac{1}{(j)^{a_1} H_{N,a_1}} + \sum_{j=1}^n \frac{1}{(j)^{a_2} H_{N,a_2}}} \leq \tilde{\Delta}(j, N, a_1, a_2). \tag{61}$$

Proof Since $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ is a convex function, we use $f(\zeta) = \frac{(\zeta-1)^2}{\zeta+1}$ in (46) to have (61), where

$$\tilde{\Delta}(j, N, a_1, a_2) = \sum_{j=1}^N \frac{1}{j^{a_2}} \frac{(j^{a_2-a_1} H_{N,a_2} - H_{N,a_1})^2}{j^{a_2-a_1} H_{N,a_2} + H_{N,a_1}}. \quad \square$$

Acknowledgements

The authors wish to thank the anonymous referees for their very careful reading of the manuscript and fruitful comments and suggestions. The research of the 5th author (Josip Pečarić) is supported by the Ministry of Education and Science of the Russian Federation (Agreement number 02.a03.21.0008).

Funding

There is no funding for this work.

Availability of data and materials

Data sharing is not applicable to this paper as no datasets were generated or analyzed during the current study.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors jointly worked on the results and they read and approved the final manuscript.

Author details

¹Department of Mathematics, University of Sargodha, Sargodha 40100, Pakistan. ²Department of Mathematics, University of Lahore (Sargodha Campus), Sargodha 40100, Pakistan. ³Department of Media and Communication, University North, Trg dr. Žarka Dolinara 1, Koprivnica, Croatia. ⁴RUDN University, Miklukho-Maklaya str. 6, 117198 Moscow, Russia.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 18 December 2020 Accepted: 21 May 2021 Published online: 28 May 2021

References

1. Agarwal, R., Bohner, M., Peterson, A.: Inequalities on time scales: a survey. *Math. Inequal. Appl.* **4**, 535–557 (2001)
2. Ahmad, K., Khan, M.A., Khan, S., Ali, A., Chu, Y.M.: New estimates for generalized Shannon and Zipf–Mandelbrot entropies via convexity results. *Results Phys.* **18**, 103305 (2020)
3. Ansari, I., Khan, K.A., Nosheen, A., Pečarić, Đ., Pečarić, J.: Shannon type inequalities via time scales theory. *Adv. Differ. Equ.* **2020**, 135 (2020)
4. Ansari, I., Khan, K.A., Nosheen, A., Pečarić, Đ., Pečarić, J.: Some inequalities for Csiszár divergence via theory of time scales. *Adv. Differ. Equ.* **2020**, 698 (2020)
5. Anwar, M., Bibi, R., Bohner, M., Pečarić, J.: Integral inequalities on time scales via the theory of isotonic linear functionals. *Abstr. Appl. Anal.* **2011**, Article ID 483595 (2011)
6. Anwar, M., Bibi, R., Bohner, M., Pečarić, J.: Jensen's functionals on time scales. *J. Funct. Spaces* **2012**, Article ID 384045 (2012)
7. Anwar, M., Bibi, R., Bohner, M., Pečarić, J.: Jensen functionals on time scales for several variables. *Int. J. Anal.* **2014**, Article ID 126797 (2014)
8. Bohner, M., Peterson, A.: *Dynamic Equations on Time Scales*. Birkhäuser, Boston (2001)
9. Bohner, M., Peterson, A.: *Advances in Dynamic Equations on Time Scales*. Birkhäuser, Boston (2003)
10. Butt, S.I., Klaričić Bakula, M., Pečarić, Đ., Pečarić, J.: Jensen–Grüss inequality and its applications for the Zipf–Mandelbrot law. *Math. Methods Appl. Sci.* **44**(2), 1664–1673 (2021)
11. Butt, S.I., Mehmood, N., Pečarić, Đ., Pečarić, J.: New bounds for Shannon, relative and Mandelbrot entropies via Abel–Gontscharoff interpolating polynomial. *Math. Inequal. Appl.* **22**(4), 1283–1301 (2019)
12. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hung.* **2**, 299–318 (1967)
13. Dragomir, S.S.: Other inequalities for Csiszár divergence and applications. *RGMI Res. Rep. Collect.* (2000). Preprint
14. Dragomir, S.S.: An upper bound for the Csiszár f -divergence in terms of the variational distance and applications. *Panam. Math. J.* **12**(4), 43–54 (2002)
15. Gibbs, A.L.: On choosing and bounding probability metrics. *Int. Stat. Rev.* **70**(3), 419–435 (2002)
16. 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.* **42**(3), 933–946 (2019)
17. Jackson, H.: On q -definite integrals. *Q. J. Pure Appl. Math.* **41**, 193–203 (1910)
18. Jain, K.C., Saraswat, R.N.: A new information inequality and its application in establishing relation among various f -divergence measures. *J. Appl. Math. Stat. Inform.* **8**(1), 17–32 (2012)
19. Khan, M.A., Al-Sahwi, Z.M., Chu, Y.M.: New estimations for Shannon and Zipf–Mandelbrot entropies. *Entropy* **20**(8), 608 (2018)
20. Khan, M.A., Anwar, M., Jakšetić, J., Pečarić, J.: On some improvements of the Jensen inequality with some applications. *J. Inequal. Appl.* **2009**, 323615 (2009)

21. Khan, M.A., Pečarić, Đ., Pečarić, J.: On Zipf–Mandelbrot entropy. *J. Comput. Appl. Math.* **346**, 192–204 (2019)
22. Khan, M.A., Pečarić, Đ., Pečarić, J.: A new refinement of the Jensen inequality with applications in information theory. *Bull. Malays. Math. Sci. Soc.* **44**, 267–278 (2021)
23. Liese, F., Vajda, I.: *Convex Statistical Distances*. Teubner, Leipzig (1987)
24. Lin, J.: Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **37**(1), 145–151 (1991)
25. Lovričević, N., Pečarić, Đ., Pečarić, J.: Zipf–Mandelbrot law, f -divergences and the Jensen-type interpolating inequalities. *J. Inequal. Appl.* **2018**(1), 1 (2018)
26. Matić, M., Pearce, C.E.M., Pečarić, J.: *Shannon's and Related Inequalities in Information Theory. Survey on Classical Inequalities* pp. 127–164. Springer, Dordrecht (2000)
27. Mehmood, N., Butt, S.I., Pečarić, Đ., Pečarić, J.: Several new cyclic Jensen type inequalities and their applications. *J. Inequal. Appl.* **2019**(1), 1 (2019)
28. Özkan, U.M., Sarikaya, M.Z., Yildirim, H.: Extensions of certain integral inequalities on time scales. *Appl. Math. Lett.* **21**(10), 993–1000 (2008)
29. Pečarić, J., Perić, J.: Refinements of the integral form of Jensen's and the Lah–Ribarič inequalities and applications for Csiszár divergence. *J. Inequal. Appl.* **2020**, 108 (2020)
30. Saker, S.H.: Some nonlinear dynamic inequalities on time scales. *Math. Inequal. Appl.* **14**(3), 633–645 (2011)
31. Sason, I., Verdú, S.: f -divergence inequalities. *IEEE Trans. Inf. Theory* **62**(11), 5973–6006 (2016)
32. Sun, Y.G., Hassan, T.: Some nonlinear dynamic integral inequalities on time scales. *Appl. Math. Comput.* **220**(4), 221–225 (2013)
33. Taneja, I.J., Kumar, P.: Relative information of type s , Csiszár's f -divergence, and information inequalities. *Inf. Sci.* **166**(1–4), 105–125 (2004)
34. Tariboon, J., Ntouyas, S.K.: Quantum calculus on finite intervals and applications to impulsive difference equations. *Adv. Differ. Equ.* **2013**, 1 (2013)
35. Tariboon, J., Ntouyas, S.K.: Quantum integral inequalities on finite intervals. *J. Inequal. Appl.* **2014**, 1 (2014)
36. Tuna, A., Kutukcu, S.: Some integral inequalities on time scales. *Appl. Math. Mech.* **29**(1), 23–29 (2008)
37. Vajda, I.: *Theory of Statistical Inference and Information*. Kluwer Academic, Boston (1989)
38. Wong, F., Yeh, C., Lian, W.: An extension of Jensen's inequality on time scales. *Adv. Dyn. Syst. Appl.* **2**(2), 113–120 (2006)
39. Xu, R., Meng, F., Song, C.: On some integral inequalities on time scales and their applications. *J. Inequal. Appl.* **2010**, 1 (2010)
40. Yin, L., Luo, Q.M., Qi, F.: Several integral inequalities on time scales. *J. Math. Inequal.* **6**(3), 419–429 (2012)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
