Adeel et al. *Journal of Inequalities and Applications* (2019) 2019:212
https://doi.org/10.1186/s13660-019-2166-8

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH **Open Access**

Generalization of the Levinson inequality with applications to information theory

Muhammad Adeel[1*], Khuram Ali Khan[1], Đilda Pečarić[2] and Josip Pečarić[3]

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

Abstract

In the presented paper, Levinson's inequality for 3-convex function is generalized by using two Green's functions. Čebyšev, Grüss, and Ostrowski-type new bounds are found for the functionals involving data points of two types. Moreover, the main results are applied to information theory via f-divergence, Rényi divergence, Rényi entropy, Shannon entropy, and Zipf–Mandelbrot law.

Keywords: Levinson's inequality; Information theory

1 Introduction and preliminaries

In [12], Ky Fan's inequality is generalized by Levinson for **3**-convex functions as follows:

Theorem A *Let* $f : I = (0, 2\alpha) \to \mathbb{R}$ *with* $f^{(3)}(t) \geq 0$. *Let* $x_k \in (0, \alpha)$ *and* $p_k > 0$. *Then*

$$J_1(f) \geq 0, \tag{1}$$

where

$$J_1(f) = \frac{1}{\mathbf{P}_n}\sum_{\rho=1}^{n} p_\rho f(2\alpha - x_\rho) - f\left(\frac{1}{\mathbf{P}_n}\sum_{\rho=1}^{n} p_\rho (2\alpha - x_\rho)\right) - \frac{1}{\mathbf{P}_n}\sum_{\rho=1}^{n} p_\rho f(x_\rho) + f\left(\frac{1}{\mathbf{P}_n}\sum_{\rho=1}^{n} p_\rho x_\rho\right). \tag{2}$$

Working with the divided differences, the assumptions of differentiability on f *can be weakened.*

In [18], Popoviciu noted that (1) is valid on $(0, 2a)$ for 3-convex functions, while in [2], Bullen gave a different proof of Popoviciu's result and also the converse of (1).

Theorem B (*a*) *Let* $f : I = [a, b] \to \mathbb{R}$ *be a* 3*-convex function and* $x_n, y_n \in [a, b]$ *for* $n = 1, 2, \ldots, k$ *such that*

$$\max\{x_1 \ldots x_k\} \leq \min\{y_1 \ldots y_k\}, \quad x_1 + y_1 = \cdots = x_k + y_k \tag{3}$$

and $p_n > 0$. *Then*

$$J_2(f) \geq 0, \tag{4}$$

where

$$J_2\big(f(\cdot)\big) = \frac{1}{\mathbf{P}_k}\sum_{\rho=1}^{k} p_\rho f(y_\rho) - f\left(\frac{1}{\mathbf{P}_k}\sum_{\rho=1}^{k} p_\rho y_\rho\right) - \frac{1}{\mathbf{P}_k}\sum_{\rho=1}^{k} p_\rho f(x_\rho) + f\left(\frac{1}{\mathbf{P}_k}\sum_{\rho=1}^{k} p_\rho x_\rho\right). \tag{5}$$

(b) *If f is continuous and* $p_\rho > 0$, (4) *holds for all* x_ρ, y_ρ *satisfying* (3), *then f is* 3-*convex.*

In [17], Pečarić weakened assumption (3) and proved that inequality (1) still holds, i.e., the following result holds:

Theorem C *Let* $f : I = [a,b] \to \mathbb{R}$ *be a* 3-*convex function,* $p_k > 0$, *and let for* $k = 1,\dots,n$, x_k, y_k *be such that* $x_k + y_k = 2\check{c}$, $x_k + x_{n-k+1} \leq 2\check{c}$ *and* $\frac{p_k x_k + p_{n-k+1} x_{n-k+1}}{p_k + p_{n-k+1}} \leq \check{c}$. *Then* (4) *holds.*

In [15], Mercer made a notable work by replacing the condition of symmetric distribution of points x_i and y_i with symmetric variances of points x_i and y_i. The second condition is a weaker condition.

Theorem D *Let f be a* 3-*convex function on* $[a,b]$, p_k *be positive such that* $\sum_{k=1}^{n} p_k = 1$. *Also let* x_k, y_k *satisfy* (3) *and*

$$\sum_{\rho=1}^{n} p_\rho\left(x_\rho - \sum_{\rho=1}^{n} p_\rho x_\rho\right)^2 = \sum_{\rho=1}^{n} p_\rho\left(y_\rho - \sum_{\rho=1}^{n} p_\rho y_\rho\right)^2. \tag{6}$$

Then (1) *holds.*

On the other hand, the error function $e_{\mathcal{F}}(t)$ can be represented in terms of the Green's function $G_{\mathcal{F},n}(t,s)$ of the boundary value problem

$$z^{(n)}(t) = 0,$$
$$z^{(i)}(a_1) = 0, \quad 0 \leq i \leq p,$$
$$z^{(i)}(a_2) = 0, \quad p+1 \leq i \leq n-1,$$
$$e_F(t) = \int_{a_1}^{a_2} G_{F,n}(t,s) f^{(n)}(s)\,ds, \quad t \in [a,b],$$

where

$$G_{F,n}(t,s) = \frac{1}{(n-1)!}\begin{cases} \sum_{i=0}^{p}\binom{n-1}{i}(t-a_1)^i(a_1-s)^{n-i-1}, & a_1 \leq s \leq t; \\ -\sum_{i=p+1}^{n-1}\binom{n-1}{i}(t-a_1)^i(a_1-s)^{n-i-1}, & t \leq s \leq a_2. \end{cases} \tag{7}$$

The following result holds in [1]:

Theorem E *Let $f \in C^n[a,b]$, and let P_F be its 'two-point right focal' interpolating polynomial. Then, for $a \leq a_1 < a_2 \leq b$ and $0 \leq p \leq n-2$,*

$$\begin{aligned} f(t) &= P_F(t) + e_F(t) \\ &= \sum_{i=0}^{p} \frac{(t-a_1)^i}{i!} f^{(i)}(a_1) \\ &\quad + \sum_{j=0}^{n-p-2} \left(\sum_{i=0}^{j} \frac{(t-a_1)^{p+1+i}(a_1-a_2)^{j-i}}{(p+1+i)!(j-i)!} \right) f^{(p+1+j)}(a_2) \\ &\quad + \int_{a_1}^{a_2} G_{F,n}(t,s) f^{(n)}(s)\, ds, \end{aligned} \tag{8}$$

where $G_{F,n}(t,s)$ is the Green's function, defined by (7).

Let $f \in C^n[a,b]$, and let P_F be its 'two-point right focal' interpolating polynomial for $a \leq a_1 < a_2 \leq b$. Then, for $n = 3$ and $p = 0$, (8) becomes

$$\begin{aligned} f(t) = {}& f(a_1) + (t-a_1)f^{(1)}(a_2) + (t-a_1)(a_1-a_2)f^{(2)}(a_2) + \frac{(t-a_1)^2}{2} f^{(2)}(a_2) \\ &+ \int_{a_1}^{a_2} G_1(t,s) f^{(3)}(s)\, ds, \end{aligned} \tag{9}$$

where

$$G_1(t,s) = \begin{cases} (a_1-s)^2, & a_1 \leq s \leq t; \\ -(t-a_1)(a_2-s) + \frac{1}{2}(t-a_1)^2, & t \leq s \leq a_2. \end{cases} \tag{10}$$

For $n = 3$ and $p = 1$, (8) becomes

$$f(t) = f(a_1) + (t-a_1)f^{(1)}(a_2) + \frac{(t-a_1)^2}{2} f^{(2)}(a_2) + \int_{a_1}^{a_2} G_2(t,s) f^{(3)}(s)\, ds, \tag{11}$$

where

$$G_2(t,s) = \begin{cases} \frac{1}{2}(a_1-s)^2 + (t-a_1)(a_1-s), & a_1 \leq s \leq t; \\ -\frac{1}{2}(t-a_1)^2, & t \leq s \leq a_2. \end{cases} \tag{12}$$

The presented work is organized as follows: In Sect. 2, Levinson's inequality for 3-convex function is generalized by using two Green's functions defined by (10) and (12). In Sect. 3, Čebyšev, Grüss, and Ostrowski-type new bounds are found for the functionals involving data points of two types. In Sect. 4, the main results are applied to information theory via f-divergence, Rényi divergence, Rényi entropy, Shannon entropy, and Zipf–Mandelbrot law.

2 Main results

First we give an identity involving Jensen's difference of two different data points. Then we give an equivalent form of identity by using the Green's function defined by (10) and (12).

Theorem 1 *Let $f \in C^3[\zeta_1,\zeta_2]$ such that $f : I = [\zeta_1,\zeta_2] \to \mathbb{R}$, $(p_1,\dots,p_n) \in \mathbb{R}^n$, $(q_1,\dots,q_m) \in \mathbb{R}^m$ such that $\sum_{\rho=1}^n p_\rho = 1$ and $\sum_{\varrho=1}^m q_\varrho = 1$. Also let $x_\rho, y_\varrho, \sum_{\rho=1}^n p_\rho x_\rho, \sum_{\varrho=1}^m q_\varrho y_\varrho \in I$. Then*

$$J(f(\cdot)) = \frac{1}{2}\left[\sum_{\varrho=1}^m q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^m q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^n p_\rho x_\rho^2 + \left(\sum_{\rho=1}^m p_\rho x_\rho\right)^2\right] f^{(2)}(\zeta_2) + \int_{\zeta_1}^{\zeta_2} J(G_k(\cdot,s)) f^{(3)}(s)\,ds, \tag{13}$$

where

$$J(f(\cdot)) = \sum_{\varrho=1}^m q_\varrho f(y_\varrho) - f\left(\sum_{\varrho=1}^m q_\varrho y_\varrho\right) - \sum_{\rho=1}^n p_\rho f(x_\rho) + f\left(\sum_{\rho=1}^n p_\rho x_\rho\right) \tag{14}$$

and

$$J(G_k(\cdot,s)) = \sum_{\varrho=1}^m q_\varrho G_k(y_\varrho,s) - G_k\left(\sum_{\varrho=1}^m q_\varrho y_\varrho, s\right) - \sum_{\rho=1}^n p_\rho G_k(x_\rho,s) + G_k\left(\sum_{\rho=1}^n p_\rho x_\rho, s\right), \tag{15}$$

for $G_k(\cdot,s)$ $(k = 1,2)$ defined in (10) and (12) respectively.

Proof (i) For $k = 1$.

Using (9) in (14), we have

$$\begin{aligned}
J(f(\cdot)) = {} & \sum_{\varrho=1}^m q_\varrho \bigg[f(\zeta_1) + (y_\varrho - \zeta_1) f^{(1)}(\zeta_2) + (y_\varrho - \zeta_1)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
& + \frac{(y_\varrho - \zeta_1)^2}{2} f^{(2)}(\zeta_2) + \int_{\zeta_1}^{\zeta_2} G_1(y_\varrho,s) f^{(3)}(s)\,ds\bigg] \\
& - \bigg[f(\zeta_1) + \left(\sum_{\varrho=1}^m q_\varrho y_\varrho - \zeta_1\right) f^{(1)}(\zeta_2) + \left(\sum_{\varrho=1}^m q_\varrho y_\varrho - \zeta_1\right)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
& + \frac{(\sum_{\varrho=1}^m q_\varrho y_\varrho - \zeta_1)^2}{2} f^{(2)}(\zeta_2) + \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\varrho=1}^m q_\varrho y_\varrho, s\right) f^{(3)}(s)\,ds\bigg] \\
& - \sum_{\rho=1}^n p_\rho \bigg[f(\zeta_1) + (x_\rho - \zeta_1) f^{(1)}(\zeta_2) + (x_\rho - \zeta_1)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
& + \frac{(x_\rho - \zeta_1)^2}{2} f^{(2)}(\zeta_2) + \int_{\zeta_1}^{\zeta_2} G_1(x_\rho,s) f^{(3)}(s)\,ds\bigg] \\
& + \bigg[f(\zeta_1) + \left(\sum_{\rho=1}^n p_\rho x_\rho - \zeta_1\right) f^{(1)}(\zeta_2) + \left(\sum_{\rho=1}^n p_\rho x_\rho - \zeta_1\right)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
& + \frac{(\sum_{\rho=1}^n p_\rho x_\rho - \zeta_1)^2}{2} f^{(2)}(\zeta_2) + \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\rho=1}^n p_\rho x_\rho, s\right) f^{(3)}(s)\,ds\bigg].
\end{aligned}$$

$$\begin{aligned}
J\big(f(\cdot)\big) ={}& f(\zeta_1) + \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho - \zeta_1\right) f^{(1)}(\zeta_2) + \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho - \zeta_1\right)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
&+ \frac{(\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - 2\zeta_1 \sum_{\varrho=1}^{m} q_\varrho y_\varrho + \zeta_1^2) f^{(2)}(\zeta_2)}{2} + \sum_{i=1}^{m} q_\varrho \int_{\zeta_1}^{\zeta_2} G_1(y_\varrho, s) f^{(3)}(s)\,ds \\
&- f(\zeta_1) - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho - \zeta_1\right) f^{(1)}(\zeta_2) - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho - \zeta_1\right)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
&- \frac{((\sum_{\varrho=1}^{m} q_\varrho y_\varrho)^2 - 2\zeta_1 \sum_{\varrho=1}^{m} q_\varrho y_\varrho + \zeta_1^2) f^{(2)}(\zeta_2)}{2} \\
&- \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho, s\right) f^{(3)}(s)\,ds \\
&- f(\zeta_1) - \left(\sum_{\rho=1}^{n} p_\rho x_\rho - \zeta_1\right) f^{(1)}(\zeta_2) - \left(\sum_{\rho=1}^{n} p_\rho x_\rho - \zeta_1\right)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
&- \frac{(\sum_{\rho=1}^{n} p_\rho x_\rho^2 - 2\zeta_1 \sum_{\rho=1}^{n} p_\rho x_\rho + \zeta_1^2) f^{(2)}(\zeta_2)}{2} - \sum_{\rho=1}^{n} p_\rho \int_{\zeta_1}^{\zeta_2} G_1(x_\rho, s) f^{(3)}(s)\,ds \\
&+ f(\zeta_1) + \left(\sum_{\rho=1}^{n} p_\rho x_\rho - \zeta_1\right) f^{(1)}(\zeta_2) + \left(\sum_{\rho=1}^{n} p_\rho x_\rho - \zeta_1\right)(\zeta_1 - \zeta_2) f^{(2)}(\zeta_2) \\
&+ \frac{((\sum_{\rho=1}^{n} p_\rho x_\rho)^2 - 2\zeta_1 \sum_{\rho=1}^{n} p_\rho x_\rho + \zeta_1^2) f^{(2)}(\zeta_2)}{2} \\
&+ \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\rho=1}^{n} p_\rho x_\rho, s\right) f^{(3)}(s)\,ds,
\end{aligned}$$

$$\begin{aligned}
J\big(f(\cdot)\big) ={}& \frac{1}{2}\left[\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^{n} p_\rho x_\rho^2 + \left(\sum_{\rho=1}^{n} p_\rho x_\rho\right)^2\right] f^{(2)}(\zeta_2) \\
&+ \sum_{\varrho=1}^{m} q_\varrho \int_{\zeta_1}^{\zeta_2} G_1(y_\varrho, s) f^{(3)}(s)\,ds - \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho, s\right) f^{(3)}(s)\,ds \\
&- \sum_{\rho=1}^{n} p_\rho \int_{\zeta_1}^{\zeta_2} G_1(x_\rho, s) f^{(3)}(s)\,ds + \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\rho=1}^{n} p_\rho x_\rho, s\right) f^{(3)}(s)\,ds.
\end{aligned}$$

After rearranging, we have (13).

(ii) For $k = 2$

Using (11) in (14) and following similar steps as in the proof of (i), we get (13). □

Corollary 1 *Let $f \in C^3[0, 2\alpha]$ such that $f : I = [0, 2\alpha] \to \mathbb{R}$, $x_1, \ldots, x_n \in (0, \alpha)$, $(p_1, \ldots, p_n) \in \mathbb{R}^n$ such that $\sum_{\rho=1}^{n} p_\rho = 1$. Also let x_ρ, $\sum_{\rho=1}^{n} p_\rho (2\alpha - x_\rho)$, $\sum_{\rho=1}^{n} p_\rho x_\rho \in I$. Then*

$$J\big(f(\cdot)\big) = \int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot, s)\big) f^{(3)}(s)\,ds, \quad 0 \le \zeta_1 < \zeta_2 \le 2\alpha, \tag{16}$$

where $J(f(\cdot))$ and $J(G(\cdot, s))$ are defined in (14) and (15) respectively.

Proof Choosing $I = [0, 2\alpha]$, $y_\varrho = (2\alpha - x_\rho)$, $x_1, \ldots, x_n \in (0, \alpha)$, $p_\rho = q_\varrho$, and $m = n$ in Theorem 1, after simplification we get (16). □

Theorem 2 *Let $f : I = [\zeta_1, \zeta_2] \to \mathbb{R}$ be a* **3**-*convex function. Also let $(p_1, \ldots, p_n) \in \mathbb{R}^n$, $(q_1, \ldots, q_m) \in \mathbb{R}^m$ be such that $\sum_{\rho=1}^n p_\rho = 1$ and $\sum_{\varrho=1}^m q_\varrho = 1$ and x_ρ, y_ϱ, $\sum_{\rho=1}^n p_\rho x_\rho$, $\sum_{\varrho=1}^m q_\varrho y_\varrho \in I$.*

If

$$\left[\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^{n} p_\rho x_\rho^2 + \left(\sum_{\rho=1}^{n} p_\rho x_\rho\right)^2\right] f^{(2)}(\zeta_2) \geq 0, \tag{17}$$

then the following statements are equivalent:

For $f \in C^3[\zeta_1, \zeta_2]$,

$$\sum_{\rho=1}^{n} p_\rho f(x_\rho) - f\left(\sum_{\rho=1}^{n} p_\rho x_\rho\right) \leq \sum_{\varrho=1}^{m} q_\varrho f(y_\varrho) - f\left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right). \tag{18}$$

For all $s \in I$,

$$\sum_{\rho=1}^{n} p_\rho G_k(x_\rho, s) - G_k\left(\sum_{\rho=1}^{n} p_\rho x_\rho, s\right) \leq \sum_{\varrho=1}^{m} q_\varrho G_k(y_\varrho, s) - G_k\left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho, s\right), \tag{19}$$

where $G_k(\cdot, s)$ are defined by (10) *and* (12) *for $k = 1, 2$ respectively.*

Moreover, inequality in (18) *is reversed iff inequality in* (19) *is reversed.*

Proof (18) ⇒ (19): Let (18) be valid. Then, as the function $G_k(\cdot, s)$ ($s \in I$) is also continuous and **3**-convex, it follows that also for this function (18) holds, i.e., (19) is valid.

(19) ⇒ (18): If f is **3**-convex, then without loss of generality we can suppose that there exists the third derivative of f. Let $f \in C^3[\zeta_1, \zeta_2]$ be a **3**-convex function and (19) hold. Then we can represent function f in the form (9). Now, by means of some simple calculations, we can write

$$\begin{aligned}
&\sum_{\varrho=1}^{m} q_\varrho f(y_\varrho) - f\left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right) - \sum_{\rho=1}^{n} p_\rho f(x_\rho) + f\left(\sum_{\rho=1}^{n} p_\rho x_\rho\right) \\
&\quad = \frac{1}{2}\left[\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^{n} p_\rho x_\rho^2 + \left(\sum_{\rho=1}^{n} p_\rho x_\rho\right)^2\right] f^{(2)}(\zeta_2) \\
&\qquad + \int_{\zeta_1}^{\zeta_2} \left(\sum_{\varrho=1}^{m} q_\varrho G_k(y_\varrho, s) - G_k\left(\sum_{\varrho=1}^{m} q_\varrho (y_\varrho, s)\right)\right. \\
&\qquad \left. - \sum_{\rho=1}^{n} p_\rho G_k(x_\rho, s) + G_k\left(\sum_{\rho=1}^{n} p_\rho x_\rho, s\right)\right) f^{(3)}(s)\, ds.
\end{aligned}$$

By the convexity of f, we have $f^{(3)}(s) \geq 0$ for all $s \in I$. Hence, if for every $s \in I$, (19) is valid, then it follows that for every **3**-convex function $f : I \to \mathbb{R}$, with $f \in C^3[\zeta_1, \zeta_2]$, (18) is valid. □

Remark 1 If the expression

$$\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^{n} p_\rho x_\rho^2 + \left(\sum_{\rho=1}^{n} p_\rho x_\rho\right)^2$$

and $f^{(2)}(\zeta_2)$ have different signs in (17), then inequalities (18) and (19) are reversed.

Next we have the results about generalization of Bullen-type inequality (for real weights) given in [2] (see also [16] and [11]).

Corollary 2 *Let $f : I = [\zeta_1, \zeta_2] \to \mathbb{R}$ be a* **3**-*convex function and $f \in C^3[\zeta_1, \zeta_2]$, $x_1, \ldots, x_n$, $y_1, \ldots, y_m \in I$ such that*

$$\max\{x_1, \ldots, x_n\} \leq \min\{y_1, \ldots, y_m\} \tag{20}$$

and

$$x_1 + y_1 = \cdots = x_n + y_m. \tag{21}$$

Also let $(p_1, \ldots, p_n) \in \mathbb{R}^n$, $(q_1, \ldots, q_m) \in \mathbb{R}^m$ be such that $\sum_{\rho=1}^{n} p_\rho = 1$ and $\sum_{\varrho=1}^{m} q_\varrho = 1$ and $x_\rho, y_\varrho, \sum_{\rho=1}^{n} p_\rho x_\rho, \sum_{\varrho=1}^{m} q_\varrho y_\varrho \in I$. If (17) holds, then (18) and (19) are equivalent.

Proof By choosing x_ρ and y_ϱ such that conditions (20) and (21) hold in Theorem 2, we get the required result. □

Remark 2 If $p_\rho = q_\varrho$ are positive and x_ρ, y_ϱ satisfy (20) and (21), then inequality (18) reduces to Bullen's inequality given in [16, p. 32, Theorem 2] for $m = n$.

Next we have a generalized form (for real weights) of Bullen-type inequality given in [17] (see also [16]).

Corollary 3 *Let $f : I = [\zeta_1, \zeta_2] \to \mathbb{R}$ be a* **3**-*convex function and $f \in C^3[\zeta_1, \zeta_2]$, $(p_1, \ldots, p_n) \in \mathbb{R}^n$, $(q_1, \ldots, q_m) \in \mathbb{R}^m$ be such that $\sum_{\rho=1}^{n} p_\rho = 1$ and $\sum_{\varrho=1}^{m} q_\varrho = 1$. Also let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m \in I$ be such that $x_\rho + y_\varrho = 2c$, and for $\rho = 1, \ldots, n$, $x_\rho + x_{n-\rho+1}$ and $\frac{p_\rho x_\rho + p_{n-\rho+1} x_{n-\rho+1}}{p_\rho + p_{n-\rho+1}} \leq c$. If (17) holds, then (18) and (19) are equivalent.*

Proof Using Theorem 2 with the conditions given in the statement, we get the required result. □

Remark 3 In Theorem 2, if $m = n$, $p_\rho = q_\varrho$ are positive, $x_\rho + y_\varrho = 2c$, $x_\rho + x_{n-\rho+1}$ and $\frac{p_\rho x_\rho + p_{n-\rho+1} x_{n-\rho+1}}{p_\rho + p_{n-\rho+1}} \leq c$. Then (18) reduces to a generalized form of Bullen's inequality defined in [16, p. 32, Theorem 4].

In [15], Mercer made a notable work by replacing condition (21) of symmetric distribution of points x_ρ and y_ϱ with symmetric variances of points x_ρ and y_ϱ for $\rho = 1, \ldots, n$ and $\varrho = 1, \ldots, m$.

So in the next result we use Mercer's condition (6), but for $\rho = \varrho$ and $m = n$.

Corollary 4 *Let $f : I = [\zeta_1, \zeta_1] \to \mathbb{R}$ be a* **3***-convex function and $f \in C^3[\zeta_1, \zeta_2]$, p_ρ, q_ρ be positive such that $\sum_{\rho=1}^{n} p_\rho = 1$ and $\sum_{\rho=1}^{n} q_\rho = 1$. Also let x_ρ, y_ρ satisfy* (20) *and*

$$\sum_{\rho=1}^{n} p_\rho \left(x_\rho - \sum_{\rho=1}^{n} p_\rho x_\rho \right)^2 = \sum_{\rho=1}^{n} p_\rho \left(y_\rho - \sum_{\rho=1}^{n} q_\rho y_\rho \right)^2. \tag{22}$$

If (17) *holds, then* (18) *and* (19) *are equivalent.*

Proof For positive weights, using (6) and (20) in Theorem 2, we get the required result. □

Next we have the results that lean on the generalization of Levinson-type inequality given in [12] (see also [16]).

Corollary 5 *Let $f : I = [0, 2\alpha] \to \mathbb{R}$ be a* **3***-convex function and $f \in C^3[0, 2\alpha]$, $x_1, \ldots, x_n \in (0, \alpha)$, $(p_1, \ldots, p_n) \in \mathbb{R}^n$ and $\sum_{\rho=1}^{n} p_\rho = 1$. Also let x_ρ, $\sum_{\rho=1}^{n} p_\rho (2\alpha - x_\rho)$, $\sum_{\rho=1}^{n} p_\rho x_\rho \in I$. Then the following are equivalent*:

$$\sum_{\rho=1}^{n} p_\rho f(x_\rho) - f\left(\sum_{\rho=1}^{n} p_\rho x_\rho \right) \le \sum_{\rho=1}^{n} p_\rho f(2\alpha - x_\rho) - f\left(\sum_{\rho=1}^{n} p_\rho (2\alpha - x_\rho) \right). \tag{23}$$

For all $s \in I$,

$$\begin{aligned} \sum_{\rho=1}^{n} p_\rho G_k(x_\rho, s) - G_k\left(\sum_{\rho=1}^{n} p_\rho x_\rho, s \right) \le{}& \sum_{\rho=1}^{n} p_\rho G_k(2\alpha - x_\rho, s) \\ & - G_k\left(\sum_{\rho=1}^{n} p_\rho (2\alpha - x_\rho), s \right), \end{aligned} \tag{24}$$

where $G_k(\cdot, s)$ is defined in (10) *and* (12) *for $k = 1, 2$ respectively.*

Proof Use $I = [0, 2\alpha]$, $(x_1, \ldots, x_n) \in (0, \alpha)$, $p_\rho = q_\varrho$, $m = n$, and $y_\varrho = (2\alpha - x_\rho)$ in Theorem 2 with $0 \le \zeta_1 < \zeta_2 \le 2\alpha$, we get the required result. □

Remark 4 In Corollary 5, if p_ρ are positive, then inequality (23) reduces to Levinson's inequality given in [16, p. 32, Theorem 1].

3 New bounds for Levinson-type functionals

Consider the Čebyšev functional for two Lebesgue integrable functions $f_1, f_2 : [\zeta_1, \zeta_2] \to \mathbb{R}$

$$\begin{aligned} \Theta(f_1, f_2) ={}& \frac{1}{\zeta_2 - \zeta_1} \int_{\zeta_1}^{\zeta_2} f_1(x) f_2(x)\, dx \\ & - \frac{1}{\zeta_2 - \zeta_1} \int_{\zeta_1}^{\zeta_2} f_1(x)\, dx \times \frac{1}{\zeta_2 - \zeta_1} \int_{\zeta_1}^{\zeta_2} f_2(x)\, dx, \end{aligned} \tag{25}$$

where the integrals are assumed to exist.

Theorem F ([3]) *Let $f_1 : [\zeta_1, \zeta_2] \to \mathbb{R}$ be a Lebesgue integrable function and $f_2 : [\zeta_1, \zeta_2] \to \mathbb{R}$ be an absolutely continuous function with $(\cdot, -\zeta_1)(\cdot, -\zeta_2)[f_2']^2 \in L[\zeta_1, \zeta_2]$. Then*

$$\left|\Theta(f_1,f_2)\right| \leq \frac{1}{\sqrt{2}}\left[\Theta(f_1,f_1)\right]^{\frac{1}{2}} \frac{1}{\sqrt{\zeta_2-\zeta_1}}\left(\int_{\zeta_1}^{\zeta_2}(t-\zeta_1)(\zeta_2-t)\left[f_2'(t)\right]^2\,dt\right)^{\frac{1}{2}}. \tag{26}$$

$\frac{1}{\sqrt{2}}$ is the best possible.

Theorem G ([3]) *Let $f_1 : [\zeta_1, \zeta_2] \to \mathbb{R}$ be absolutely continuous with $f_1' \in L_\infty[\zeta_1, \zeta_2]$, and let $f_2 : [\zeta_1, \zeta_2] \to \mathbb{R}$ be monotonic nondecreasing on $[\zeta_1, \zeta_2]$. Then*

$$\left|\Theta(f_1,f_2)\right| \leq \frac{1}{2(\zeta_2-\zeta_1)}\left\|f'\right\|_\infty \int_{\zeta_1}^{\zeta_2}(t-\zeta_1)(\zeta_2-t)\left[f_2'(t)\right]\,df_2(t). \tag{27}$$

$\frac{1}{2}$ is the best possible.

In the next result we construct the Čebyšev-type bound for our functional defined in (5).

Theorem 3 *Let $f \in C^3[\zeta_1, \zeta_2]$ be such that $f : I = [\zeta_1, \zeta_2] \to \mathbb{R}$ and $f^{(3)}(\cdot)$ is absolutely continuous with $(\cdot - \zeta_1)(\zeta_2 - \cdot)[f^{(4)}]^2 \in L[\zeta_1, \zeta_2]$. Also let $(p_1, \ldots, p_n) \in \mathbb{R}^n$, $(q_1, \ldots, q_m) \in \mathbb{R}^m$ be such that $\sum_{\rho=1}^n p_\rho = 1$, $\sum_{\varrho=1}^m q_\varrho = 1$, x_ρ, y_ϱ, $\sum_{\rho=1}^n p_\rho x_\rho$, $\sum_{\varrho=1}^m q_\varrho y_\varrho \in I$. Then*

$$\begin{aligned} J\big(f(\cdot)\big) &= \frac{1}{2}\left[\sum_{\varrho=1}^m q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^m q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^n p_\rho x_\rho^2 + \left(\sum_{\varrho=1}^m p_\rho x_\rho\right)^2\right] f^{(2)}(\zeta_2) \\ &\quad + \frac{f^{(2)}(\zeta_2) - f^{(2)}(\zeta_1)}{(\zeta_2-\zeta_1)} \int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big) f^{(3)}(s)\,ds + \mathcal{R}_3(\zeta_1,\zeta_2;f), \end{aligned} \tag{28}$$

where $J(f(\cdot))$, $J(G_k(\cdot,s))$ are defined in (14) and (15) respectively, and the remainder $\mathcal{R}_3(\zeta_1, \zeta_2; f)$ satisfies the bound

$$\begin{aligned} \left|\mathcal{R}_3(\zeta_1,\zeta_2;f)\right| \leq{}& \frac{\zeta_2-\zeta_1}{\sqrt{2}}\left[\Theta\big(J\big(G_k(\cdot,s)\big),J\big(G_k(\cdot,s)\big)\big)\right]^{\frac{1}{2}} \times \\ & \frac{1}{\sqrt{\zeta_2-\zeta_1}}\left(\int_{\zeta_1}^{\zeta_2}(s-\zeta_1)(\zeta_2-s)\left[f^{(4)}(s)\right]ds\right)^{\frac{1}{2}} \end{aligned} \tag{29}$$

for $G_k(\cdot, s)$ $(k = 1, 2)$ defined in (10) and (12) respectively.

Proof Setting $f_1 \mapsto J(G_k(\cdot, s))$ and $f_2 \mapsto f^{(3)}$ in Theorem F, we get

$$\begin{aligned} &\left|\frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big) f^{(3)}(s)\,ds - \frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)\,ds \times \frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} f^{(3)}(s)\,ds\right| \\ &\quad\leq \frac{1}{\sqrt{2}}\left[\Theta\big(J\big(G_k(\cdot,s)\big),J\big(G_k(\cdot,s)\big)\big)\right]^{\frac{1}{2}} \frac{1}{\sqrt{\zeta_2-\zeta_1}}\left(\int_{\zeta_1}^{\zeta_2}(s-\zeta_1)(\zeta_2-s)\left[f^{(4)}(s)\right]ds\right)^{\frac{1}{2}}, \end{aligned}$$

$$\left|\frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)f^{(3)}(s)\,ds-\frac{f^{(2)}(\zeta_2)-f^{(2)}(\zeta_1)}{(\zeta_2-\zeta_1)^2}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)\,ds\right|$$
$$\leq \frac{1}{\sqrt{2}}\big[\Theta\big(J\big(G_k(\cdot,s)\big),J\big(G_k(\cdot,s)\big)\big)\big]^{\frac{1}{2}}\frac{1}{\sqrt{\zeta_2-\zeta_1}}\left(\int_{\zeta_1}^{\zeta_2}(s-\zeta_1)(\zeta_2-s)\big[f^{(4)}(s)\big]\,ds\right)^{\frac{1}{2}}.$$

Multiplying $(\zeta_2-\zeta_1)$ on both sides of the above inequality and using estimation (29), we get

$$\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)f^{(3)}\,ds=\frac{f^{(2)}(\zeta_2)-f^{(2)}(\zeta_1)}{(\zeta_2-\zeta_1)}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)\,ds+\mathcal{R}_3(\zeta_1,\zeta_1;f).$$

Using identity (13), we get (28). □

In the next result the bounds of Grüss-type inequalities are estimated.

Theorem 4 *Let $f\in C^3[\zeta_1,\zeta_2]$ be such that $f:I=[\zeta_1,\zeta_2]\to\mathbb{R}$, $f^{(3)}(\cdot)$ is absolutely continuous and $f^{(4)}(\cdot)\geq 0$ a.e. on $[\zeta_1,\zeta_2]$. Also let $(p_1,\dots,p_n)\in\mathbb{R}^n$, $(q_1,\dots,q_m)\in\mathbb{R}^m$ be such that $\sum_{\rho=1}^n p_\rho=1$, $\sum_{\varrho=1}^m q_\varrho=1$, x_ρ, y_ϱ, $\sum_{\rho=1}^n p_\rho x_\rho$, $\sum_{\varrho=1}^m q_\varrho y_\varrho\in I$. Then identity (28) holds, where the remainder satisfies the estimation*

$$\big|\mathcal{R}_3(\zeta_1,\zeta_2;f)\big|\leq(\zeta_2-\zeta_1)\big\|J\big(G_k(\cdot,s)\big)'\big\|_\infty\left[\frac{f^{(2)}(\zeta_2)+f^{(2)}(\zeta_1)}{2}-\frac{f^{(2)}(\zeta_2)-f^{(2)}(\zeta_1)}{\zeta_2-\zeta_1}\right].\tag{30}$$

Proof Setting $f_1\mapsto J(G_k(\cdot,s))$ and $f_2\mapsto f^{(3)}$ in Theorem G, we get

$$\left|\frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)f^{(3)}(s)\,ds-\frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} J\big(G_k(\cdot,s)\big)\,ds\cdot\frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2} f^{(3)}(s)\,ds\right|$$
$$\leq\frac{1}{2}\big\|J\big(G_k(\cdot,s)\big)'\big\|_\infty\frac{1}{\zeta_2-\zeta_1}\int_{\zeta_1}^{\zeta_2}(s-\zeta_1)(\zeta_2-s)\big[f^{(4)}(s)\big]\,ds.\tag{31}$$

Since

$$\begin{aligned}\int_{\zeta_1}^{\zeta_2}&(s-\zeta_1)(\zeta_2-s)\big[f^{(4)}(s)\big]\,ds\\&=\int_{\zeta_1}^{\zeta_2}[2s-\zeta_1-\zeta_2]f^3(s)\,ds\\&=(\zeta_2-\zeta_1)\big[f^{(2)}(\zeta_2)+f^{(2)}(\zeta_1)\big]-2\big(f^{(2)}(\zeta_2)-f^{(2)}(\zeta_1)\big),\end{aligned}\tag{32}$$

using (13), (31), and (32), we have (28). □

Ostrowski-type bounds for a newly constructed functional defined in (5).

Theorem 5 *Let $f\in C^3[\zeta_1,\zeta_2]$ be such that $f:I=[\zeta_1,\zeta_2]\to\mathbb{R}$ and $f^{(2)}(\cdot)$ is absolutely continuous. Also let $(p_1,\dots,p_n)\in\mathbb{R}^n$, $(q_1,\dots,q_m)\in\mathbb{R}^m$ be such that $\sum_{\rho=1}^n p_\rho=1$, $\sum_{\varrho=1}^m q_\varrho=1$, x_ρ, y_ϱ, $\sum_{\rho=1}^n p_\rho x_\rho$, $\sum_{\varrho=1}^m q_\varrho y_\varrho\in I$. Also let (r,s) be a pair of conjugate exponents, that is,*

$1 \leq r, s, \leq \infty$, $\frac{1}{r} + \frac{1}{s} = 1$. *If* $|f^{(3)}|^r : [\zeta_1, \zeta_2] \to \mathbb{R}$ *is a Riemann integrable function, then*

$$\left| J(f(\cdot)) - \frac{1}{2}\left[\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^{n} p_\rho x_\rho^2 + \left(\sum_{\varrho=1}^{m} p_\rho x_\rho\right)^2\right] f^{(2)}(\zeta_2)\right|$$
$$\leq \|f^{(3)}\|_r \left(\int_{\zeta_1}^{\zeta_2} \left|J\left(G_k(\cdot, s)\right) ds\right|^s\right)^{\frac{1}{s}}. \tag{33}$$

Proof Rearrange identity (13) in the following way:

$$\left| J(f(\cdot)) - \frac{1}{2}\left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho^2 - \left(\sum_{\varrho=1}^{m} q_\varrho y_\varrho\right)^2 - \sum_{\rho=1}^{n} p_\rho x_\rho^2 + \left(\sum_{\rho=1}^{m} p_\rho x_\rho\right)^2\right) f^{(2)}(\zeta_2)\right|$$
$$\leq \int_{\zeta_1}^{\zeta_2} J\left(G_k(\cdot, s)\right) f^{(3)}(s)\, ds. \tag{34}$$

Employing the classical Holder's inequality to R.H.S of (34) yields (33). □

4 Application to information theory

The idea of Shannon entropy is the focal job of data hypothesis once in a while alluded as measure of uncertainty. The entropy of a random variable is characterized regarding its probability distribution and can be shown to be a decent measure of randomness or uncertainty. Shannon entropy permits to evaluate the normal least number of bits expected to encode a series of images dependent on the letters in order size and the recurrence of the symbols.

Divergences between probability distributions have been acquainted with measure of the difference between them. A variety of sorts of divergences exist, for instance the f-difference (particularly, Kullback–Leibler divergence, Hellinger distance, and total variation distance), Rényi divergence, Jensen–Shannon divergence, and so forth (see [13, 21]). There are a lot of papers managing inequalities and entropies, see, e.g., [8, 10, 20] and the references therein. Jensen's inequality assumes a crucial role in a portion of these inequalities. In any case, Jensen's inequality deals with one sort of information focuses and Levinson's inequality manages two types of information points.

Zipf's law is one of the central laws in data science, and it has been utilized in linguistics. George Zipf in 1932 found that we can tally how frequently each word shows up in the content. So on the off chance that we rank (r) word as per the recurrence of word event (f), at that point the result of these two numbers is steady (C) : $C = r \times f$. Aside from the utilization of this law in data science and linguistics, Zipf's law is utilized in city population, sun powered flare power, site traffic, earthquake magnitude, the span of moon pits, and so forth. In financial aspects this distribution is known as the Pareto law, which analyzes the distribution of the wealthiest individuals in the community [6, p. 125]. These two laws are equivalent in the mathematical sense, yet they are involved in different contexts [7, p. 294].

4.1 Csiszár divergence

In [4, 5] Csiszár gave the following definition:

Definition 1 Let f be a convex function from $\mathbb{R}^+$ to $\mathbb{R}^+$. Let $\tilde{\mathbf{r}}, \tilde{\mathbf{k}} \in \mathbb{R}_+^n$ be such that $\sum_{s=1}^n r_s = 1$ and $\sum_{s=1}^n q_s = 1$. Then an f-divergence functional is defined by

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

By defining the following:

$$f(0) := \lim_{x \to 0^+} f(x); \qquad 0f\left(\frac{0}{0}\right) := 0; \qquad 0f\left(\frac{a}{0}\right) := \lim_{x \to 0^+} xf\left(\frac{a}{0}\right), \quad a > 0,$$

he stated that nonnegative probability distributions can also be used.

Using the definition of f-divergence functional, Horváth *et al.* [9] gave the following functional:

Definition 2 Let I be an interval contained in $\mathbb{R}$ and $f: I \to \mathbb{R}$ be a function. Also let $\tilde{\mathbf{r}} = (r_1, \ldots, r_n) \in \mathbb{R}^n$ and $\tilde{\mathbf{k}} = (k_1, \ldots, k_n) \in (0, \infty)^n$ be such that

$$\frac{r_s}{k_s} \in I, \quad s = 1, \ldots, n.$$

Then

$$\hat{I}_f(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) := \sum_{s=1}^n k_s f\left(\frac{r_s}{k_s}\right).$$

We apply a generalized form of Bullen's inequality (18) (for positive weights) to $\hat{I}_f(\tilde{\mathbf{r}}, \tilde{\mathbf{k}})$.

Let us denote the following set of assumptions by $\mathcal{G}$:

Let $f: I = [\alpha, \beta] \to \mathbb{R}$ be a 3-convex function. Also let $(p_1, \ldots, p_n) \in \mathbb{R}^+$, $(q_1, \ldots, q_m) \in \mathbb{R}^+$ be such that $\sum_{s=1}^n p_s = 1$ and $\sum_{s=1}^m q_s = 1$ and x_s, y_s, $\sum_{s=1}^n p_s x_s$, $\sum_{s=1}^m q_s y_s \in I$.

Theorem 6 *Assume* $\mathcal{G}$.

Let $\tilde{\mathbf{r}} = (r_1, \ldots, r_n)$, $\tilde{\mathbf{k}} = (k_1, \ldots, k_n)$ *be in* $(0, \infty)^n$, *and* $\tilde{\mathbf{w}} = (w_1, \ldots, w_m)$, $\tilde{\mathbf{t}} = (t_1, \ldots, t_m)$ *be in* $(0, \infty)^m$ *such that*

$$\frac{r_s}{k_s} \in I, \quad s = 1, \ldots, n,$$

and

$$\frac{w_u}{t_u} \in I, \quad u = 1, \ldots, m.$$

Then

$$\text{(i)} \quad \frac{1}{\sum_{s=1}^n k_s} \hat{I}_f(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - f\left(\sum_{s=1}^n \frac{r_s}{\sum_{s=1}^n k_s}\right) \le \frac{1}{\sum_{u=1}^m t_u} \hat{I}_f(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) - f\left(\sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u}\right). \tag{35}$$

(ii) *If* $x \to xf(x)$ $(x \in [a, b])$ *is* 3-*convex, then*

$$\frac{1}{\sum_{s=1}^n k_s} \hat{I}_{idf}(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - f\left(\sum_{s=1}^n \frac{r_s}{\sum_{s=1}^n k_s}\right) \le \frac{1}{\sum_{u=1}^m t_u} \hat{I}_{idf}(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) - f\left(\sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u}\right), \tag{36}$$

where

$$\hat{I}_{idf}(\tilde{\mathbf{r}},\tilde{\mathbf{k}}) = \sum_{s=1}^{n} r_s f\left(\frac{r_s}{k_s}\right)$$

and

$$\hat{I}_{idf}(\tilde{\mathbf{w}},\tilde{\mathbf{t}}) = \sum_{u=1}^{m} w_u f\left(\frac{w_u}{t_u}\right).$$

Proof (i) Taking $p_s = \frac{k_s}{\sum_{s=1}^{n} k_s}$, $x_\rho = \frac{r_s}{k_s}$, $q_s = \frac{t_u}{\sum_{u=1}^{m} t_u}$, and $y_s = \frac{w_u}{t_u}$ in inequality (18) (for positive weights), we have

$$\sum_{s=1}^{n} \frac{k_s}{\sum_{s=1}^{n} k_s} f\left(\frac{r_s}{k_s}\right) - f\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s}\right) \le \sum_{u=1}^{m} \frac{t_u}{\sum_{u=1}^{m} t_u} f\left(\frac{w_u}{t_u}\right) - f\left(\sum_{u=1}^{m} \frac{w_u}{\sum_{u=1}^{m} t_u}\right). \tag{37}$$

Multiplying (37) by the sum $\sum_{s=1}^{n} k_s$, we get

$$\begin{aligned} \hat{I}_f(\tilde{\mathbf{r}},\tilde{\mathbf{k}}) - f\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s}\right) \sum_{s=1}^{n} k_s \le{}& \sum_{u=1}^{m} \frac{t_u}{\sum_{u=1}^{m} t_u} f\left(\frac{w_u}{t_u}\right) \sum_{s=1}^{n} k_s \\ &- f\left(\sum_{u=1}^{m} \frac{w_u}{\sum_{u=1}^{m} t_u}\right) \sum_{s=1}^{n} k_s. \end{aligned} \tag{38}$$

Now again multiplying (38) by the sum $\sum_{u=1}^{m} t_u$, we get

$$\begin{aligned} &\sum_{u=1}^{m} t_u \hat{I}_f(\tilde{\mathbf{r}},\tilde{\mathbf{k}}) - f\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s}\right) \sum_{s=1}^{n} k_s \sum_{u=1}^{m} t_u \\ &\quad \le \sum_{s=1}^{n} k_s \hat{I}_f(\tilde{\mathbf{w}},\tilde{\mathbf{t}}) - f\left(\sum_{u=1}^{m} \frac{w_u}{\sum_{u=1}^{m} t_u}\right) \sum_{s=1}^{n} k_s \sum_{u=1}^{m} t_u. \end{aligned}$$

If we divide the above inequality with the product $\sum_{s=1}^{n} k_s \sum_{u=1}^{m} t_u$, we get (35).

(ii) Using $f := idf$ (where "*id*" is the identity function) in (18)(for positive weights), we have

$$\sum_{s=1}^{n} p_s x_s f(x_s) - \sum_{s=1}^{n} p_s x_s f\left(\sum_{s=1}^{n} p_s x_s\right) \le \sum_{u=1}^{m} q_u y_u f(y_u) - \sum_{u=1}^{m} q_u y_u f\left(\sum_{u=1}^{m} q_u y_u\right).$$

Using the same steps as in the proof of (i), we get (36). □

4.2 Shannon entropy

Definition 3 (see [9]) The $\mathcal{S}$hannon entropy of positive probability distribution $\tilde{\mathbf{r}} = (r_1, \ldots, r_n)$ is defined by

$$\boldsymbol{\mathcal{S}} := -\sum_{s=1}^{n} r_s \log(r_s). \tag{39}$$

Corollary 6 *Assume $\mathcal{G}$.*

If $\tilde{\mathbf{k}} = (k_1, \dots, k_n) \in \mathbb{R}_+^n$, $\tilde{\mathbf{t}} = (t_1, \dots, t_m) \in \mathbb{R}_+^m$ and if base of log *is greater than* 1, *then*

$$\begin{aligned} &\frac{1}{\sum_{s=1}^{n} k_s}\left[\boldsymbol{\mathcal{S}} + \sum_{s=1}^{n} r_s \log(k_s)\right] + \left[\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s} \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s}\right)\right] \\ &\quad\le \frac{1}{\sum_{u=1}^{m} t_u}\left[\tilde{\boldsymbol{\mathcal{S}}} + \sum_{u=1}^{m} w_u \log(t_u)\right] + \left[\sum_{u=1}^{m} \frac{w_u}{\sum_{u=1}^{m} t_u} \log\left(\sum_{u=1}^{m} \frac{w_u}{\sum_{u=1}^{m} t_u}\right)\right], \end{aligned} \tag{40}$$

where $\boldsymbol{\mathcal{S}}$ is defined in (39), *and*

$$\tilde{\boldsymbol{\mathcal{S}}} := -\sum_{u=1}^{m} w_u \log(w_u).$$

If base of log *is less than* 1, *then inequality* (40) *is reversed.*

Proof The function $f \mapsto -x\log(x)$ is 3-convex for base of log is greater than 1. So, using $f := -x\log(x)$ in Theorem 6(i), we get (40). □

Remark 5 If k and t are positive probability distributions, then (40) becomes

$$\begin{aligned} &\left[\boldsymbol{\mathcal{S}} + \sum_{s=1}^{n} r_s \log(k_s)\right] + \left[\sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} r_s\right)\right] \le \left[\tilde{\boldsymbol{\mathcal{S}}} + \sum_{s=1}^{m} w_s \log(t_s)\right] \\ &\qquad + \left[\sum_{s=1}^{m} w_s \log\left(\sum_{s=1}^{m} w_s\right)\right]. \end{aligned} \tag{41}$$

Definition 4 (see [9]) For $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{q}}$, where $\tilde{\mathbf{r}}, \tilde{\mathbf{q}} \in \mathbb{R}_+^n$ the Kullback–Leibler divergence is defined by

$$\boldsymbol{\mathcal{D}}(\tilde{\mathbf{r}}, \tilde{\mathbf{q}}) := \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right). \tag{42}$$

Corollary 7 *Assume $\mathcal{G}$.*

Let $\tilde{\mathbf{r}} = (r_1, \dots, r_n)$, $\tilde{\mathbf{k}} = (k_1, \dots, k_n) \in \mathbb{R}_+^n$, and $\tilde{\mathbf{w}} := (w_1, \dots, w_m)$, $\tilde{\mathbf{t}} = (t_1, \dots, t_m) \in \mathbb{R}_+^m$ be such that $\sum_{s=1}^{n} r_s$, $\sum_{s=1}^{n} k_s$, $\sum_{s=1}^{m} w_s$, and $\sum_{s=1}^{m} t_s$ be equal to 1, *then*

$$\sum_{s=1}\left(\frac{r_s}{k_s}\right)\boldsymbol{\mathcal{D}}(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - \sum_{s=1}\left(\frac{w_s}{t_s}\right)\boldsymbol{\mathcal{D}}(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) \ge 0, \tag{43}$$

where base of log *is greater than* 1.

If base of log *is less than* 1, *then the signs of inequality in* (43) *are reversed.*

Proof In Theorem 6(ii), replacing f by $-x\log(x)$, we have

$$\begin{aligned} &\frac{\sum_{s=1}^{n}(\frac{r_s}{k_s})}{\sum_{s=1}^{n} k_s}\boldsymbol{\mathcal{D}}(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - \sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s} \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} k_s}\right) \\ &\quad\ge \frac{\sum_{s=1}^{m}(\frac{w_s}{t_s})}{\sum_{s=1}^{m} t_s}\boldsymbol{\mathcal{D}}(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) - \sum_{s=1}^{m} \frac{w_s}{\sum_{s=1}^{n} t_s} \log\left(\sum_{s=1}^{m} \frac{w_s}{\sum_{s=1}^{m} t_s}\right). \end{aligned} \tag{44}$$

Now simply taking $\sum_{s=1}^{n} r_s$, $\sum_{s=1}^{n} k_s$, $\sum_{s=1}^{m} w_s$, and $\sum_{s=1}^{m} t_s$ are equal to 1 and after rearranging, we get (43). □

4.3 Rényi divergence and entropy

The Rényi divergence and Rényi entropy are given in [19].

Definition 5 Let $\tilde{\mathbf{r}}, \tilde{\mathbf{q}} \in \mathbb{R}_+^n$ be such that $\sum_1^n r_i = 1$ and $\sum_1^n q_i = 1$, and let $\delta \geq 0$, $\delta \neq 1$.

(a) The Rényi divergence of order δ is defined by

$$\mathcal{D}_\delta(\tilde{\mathbf{r}}, \tilde{\mathbf{q}}) := \frac{1}{\delta - 1} \log\left(\sum_{i=1}^{n} q_i \left(\frac{r_i}{q_i}\right)^{\delta}\right). \tag{45}$$

(b) The Rényi entropy of order δ of $\tilde{\mathbf{r}}$ is defined by

$$\mathcal{H}_\delta(\tilde{\mathbf{r}}) := \frac{1}{1 - \delta} \log\left(\sum_{i=1}^{n} r_i^{\delta}\right). \tag{46}$$

These definitions also hold for nonnegative probability distributions. If $\delta \to 1$ in (45), we have (42), and if $\delta \to 1$ in (46), then we have (39).

Now we obtain inequalities for the Rényi divergence.

Theorem 7 *Assume* $\mathcal{G}$.

Let $\tilde{\mathbf{r}} = (r_1, \ldots, r_n)$, $\tilde{\mathbf{k}} = (k_1, \ldots, k_n) \in \mathbb{R}_+^n$, $\tilde{\mathbf{w}} = (w_1, \ldots, w_m)$, *and* $\tilde{\mathbf{t}} = (t_1, \ldots, t_m) \in \mathbb{R}_+^m$.

(i) *If base of* log *is greater than* 1 *and* $0 \leq \delta \leq \theta$ *are such that* $\delta, \theta \neq 1$, *then*

$$\mathcal{D}_\theta(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - \mathcal{D}_\delta(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) \leq \mathcal{D}_\theta(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) - \mathcal{D}_\delta(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}). \tag{47}$$

If base of log *is less than* 1, *the inequality* (47) *holds in reverse*.

(ii) *If* $\theta > 1$ *and if base of* log *is greater than* 1, *then*

$$\mathcal{D}_\theta(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - \mathcal{D}_1(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) \leq \mathcal{D}_\theta(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) - \mathcal{D}_1(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}). \tag{48}$$

(iii) *If* $\delta \in (0, 1)$ *and if base of* log *is greater than* 1, *then*

$$\mathcal{D}_1(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) - \mathcal{D}_\delta(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) \leq \mathcal{D}_1(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) - \mathcal{D}_\delta(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}). \tag{49}$$

Proof With the mapping f defined by $f : (0, \infty) \to \mathbb{R}$ by $f(t) := t^{\frac{\theta-1}{\delta-1}}$ and using

$$p_s := r_s, \qquad x_s := \left(\frac{r_s}{k_s}\right)^{\delta-1}, \quad s = 1, \ldots, n,$$

and

$$q_u := w_u, \qquad y_u := \left(\frac{w_u}{t_u}\right)^{\delta-1}, \quad u = 1, \ldots, m,$$

in (18) (for positive weights) and after simplifications, we have

$$\sum_{s=1}^{n} k_s \left(\frac{r_s}{k_s}\right)^{\theta} - \left(\sum_{s=1}^{n} k_s \left(\frac{r_s}{k_s}\right)^{\delta}\right)^{\frac{\theta-1}{\delta-1}} \leq \sum_{u=1}^{m} t_u \left(\frac{w_u}{t_u}\right)^{\theta} - \left(\sum_{u=1}^{m} t_u \left(\frac{w_u}{t_u}\right)^{\delta}\right)^{\frac{\theta-1}{\delta-1}} \tag{50}$$

if either $0 \le \delta < 1 < \gamma$ or $1 < \delta \le \theta$, and inequality (50) holds in reverse if $0 \le \delta \le \gamma < 1$. Raising the power $\frac{1}{\theta-1}$ in (50),

$$\begin{aligned}
&\left(\sum_{s=1}^{n} k_s\left(\frac{r_s}{k_s}\right)^{\theta}\right)^{\frac{1}{\theta-1}} - \left(\sum_{s=1}^{n} k_s\left(\frac{r_s}{k_s}\right)^{\delta}\right)^{\frac{1}{\delta-1}} \\
&\quad \le \left(\sum_{u=1}^{m} t_u\left(\frac{w_u}{t_u}\right)^{\theta}\right)^{\frac{1}{\theta-1}} - \left(\sum_{u=1}^{m} t_u\left(\frac{w_u}{t_u}\right)^{\delta}\right)^{\frac{1}{\delta-1}}.
\end{aligned} \tag{51}$$

For base of log is greater than 1, the log function is increasing, therefore on taking log in (51), we get (47). If base of log is less than 1, inequality in (47) is reversed. If $\delta = 1 = \theta$, and by taking the limit, we have (48) and (49) respectively. □

Theorem 8 *Assume $\mathcal{G}$.*

Let $\tilde{\mathbf{r}} = (r_1, \dots, r_n)$, $\tilde{\mathbf{k}} = (k_1, \dots, k_n) \in \mathbb{R}^n_+$, $\tilde{\mathbf{w}} = (w_1, \dots, w_m)$, and $\tilde{\mathbf{t}} = (t_1, \dots, t_m) \in \mathbb{R}^m_+$.

If either $1 < \delta$ and base of log *is greater than 1 or $\delta \in [0,1)$ and base of* log *is less than 1, then*

$$\begin{aligned}
&\frac{1}{\sum_{s=1}^{n} k_s(\frac{r_s}{k_s})^{\delta}} \sum_{s=1}^{n} k_s\left(\frac{r_s}{k_s}\right)^{\delta} \log\left(\frac{r_s}{k_s}\right) - \mathcal{D}_{\delta}(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) \\
&\quad \le \frac{1}{\sum_{s=1}^{n} k_s(\frac{r_s}{k_s})^{\delta}} \sum_{s=1}^{m} t_s\left(\frac{w_s}{t_s}\right)^{\delta} \log\left(\frac{w_s}{t_s}\right) - \frac{\sum_{s=1}^{m} t_s(\frac{w_s}{t_s})^{\delta}}{\sum_{s=1}^{n} k_s(\frac{r_s}{k_s})^{\delta}} \mathcal{D}_{\delta}(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}).
\end{aligned} \tag{52}$$

If either $1 < \delta$ and base of log *is greater than 1 or $\delta \in [0,1)$ and base of* log *is less than 1, inequality in (52) is reversed.*

Proof The proof is only for the case when $\delta \in [0,1)$ and base of log is greater than 1, and similarly the remaining cases are simple to prove.

The function $x \mapsto xf(x)$ $(x > 0)$ is 3-convex for base of log is less than 1. Also $0 > \frac{1}{1-\delta}$ and choosing $I = (0, \infty)$,

$$p_s := r_s, \qquad x_s := \left(\frac{r_s}{k_s}\right)^{\delta-1}, \quad s = 1, \dots, n,$$

and

$$q_u := w_u, \qquad y_u := \left(\frac{w_u}{t_u}\right)^{\delta-1}, \quad u = 1, \dots, m,$$

in (18) (for positive weights) and after simplifications, we have (52). □

Corollary 8 *Assume $\mathcal{G}$.*

Let $\tilde{\mathbf{r}} = (r_1, \dots, r_n)$, $\tilde{\mathbf{k}} = (k_1, \dots, k_n) \in \mathbb{R}^n_+$, $\tilde{\mathbf{w}} = (w_1, \dots, w_m)$, and $\tilde{\mathbf{t}} = (t_1, \dots, t_m) \in \mathbb{R}^m_+$ be such that $\sum_{s=1}^{n} r_s$, $\sum_{s=1}^{n} k_s$, $\sum_{u=1}^{m} w_u$, and $\sum_{u=1}^{m} t_u$ are equal to 1.

(i) *If base of* log *is greater than 1 and $0 \le \delta \le \theta$ such that $\delta, \theta \ne 1$, then*

$$\mathcal{H}_{\theta}(\tilde{\mathbf{r}}) - \mathcal{H}_{\delta}(\tilde{\mathbf{r}}) \ge \mathcal{H}_{\theta}(\tilde{\mathbf{w}}) - \mathcal{H}_{\delta}(\tilde{\mathbf{w}}). \tag{53}$$

The reverse inequality holds in (53) *if base of* log *is less than* 1.

(ii) *If* $1 < \theta$ *and base of* log *is greater than* 1, *then*

$$\mathcal{H}_{\theta}(\tilde{\mathbf{r}}) - \mathcal{S} \geq \mathcal{H}_{\theta}(\tilde{\mathbf{w}}) - \tilde{\mathcal{S}}. \tag{54}$$

The reverse inequality holds in (54) *if base of* log *is greater than* 1.

(iii) *If* $0 \leq \delta < 1$ *and base of* log *is greater than* 1, *then*

$$\mathcal{S} - \mathcal{H}_{\delta}(\tilde{\mathbf{r}}) \geq \tilde{\mathcal{S}} - \mathcal{H}_{\delta}(\tilde{\mathbf{w}}). \tag{55}$$

If base of log *is less than* 1, *the inequality in* (55) *is reversed.*

Proof (i) Suppose $\tilde{\mathbf{k}}, \tilde{\mathbf{t}} = \frac{\mathbf{1}}{\mathbf{n}}$. Then from (45) we have

$$\mathcal{D}_{\delta}(\tilde{\mathbf{r}}, \tilde{\mathbf{q}}) = \frac{1}{\delta - 1} \log\left(\sum_{s=1}^{n} n^{\delta-1} r_s^{\delta}\right) = \log(n) + \frac{1}{\delta - 1} \log\left(\sum_{s=1}^{n} r_s^{\delta}\right)$$

and

$$\mathcal{D}_{\delta}(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) = \frac{1}{\delta - 1} \log\left(\sum_{s=1}^{n} n^{\delta-1} w_s^{\delta}\right) = \log(n) + \frac{1}{\delta - 1} \log\left(\sum_{s=1}^{n} w_s^{\delta}\right).$$

We have

$$\mathcal{H}_{\delta}(\tilde{\mathbf{r}}) = \log(n) - \mathcal{D}_{\delta}\left(\tilde{\mathbf{r}}, \frac{\mathbf{1}}{\mathbf{n}}\right) \tag{56}$$

and

$$\mathcal{H}_{\delta}(\tilde{\mathbf{w}}) = \log(n) - \mathcal{D}_{\delta}\left(\tilde{\mathbf{w}}, \frac{\mathbf{1}}{\mathbf{n}}\right). \tag{57}$$

We get (53) after using Theorem 7(i), (56) and (57).

Statements (ii) and (iii) are similarly proved. □

Corollary 9 *Assume* $\mathcal{G}$.

Let $\tilde{\mathbf{r}} = (r_1, \ldots, r_n)$, $\tilde{\mathbf{k}} = (k_1, \ldots, k_n)$, $\tilde{\mathbf{w}} = (w_1, \ldots, w_m)$, *and* $\tilde{\mathbf{t}} = (t_1, \ldots, t_n)$ *be positive probability distributions.*

If either $\delta \in [0, 1)$ *and base if* log *is greater than* 1, *or* $\delta > 1$ *and base if* log *is less than* 1, *then*

$$-\frac{1}{\sum_{s=1}^{n} r_s^{\delta}} \sum_{s=1}^{n} r_s^{\delta} \log(r_s) - \mathcal{H}_{\delta}(r) \geq \frac{1}{\sum_{s=1}^{m} w_s^{\delta}} \sum_{s=1}^{m} w_s^{\delta} \log(w_s) - \frac{\sum_{s=1}^{m} w_s^{\delta}}{\sum_{s=1}^{n} r_s^{\delta}} \mathcal{H}_{\delta}(w). \tag{58}$$

The inequality in (58) *is reversed if either* $\delta \in [0, 1)$ *and base if* log *is less than* 1, *or* $\delta > 1$ *and the base of* log *is greater than* 1.

Proof Proof is similar to Corollary 8 □

4.4 Zipf–Mandelbrot law

In [14] the authors gave some contribution in analyzing the Zipf–Mandelbrot law which is defined as follows:

Definition 6 The Zipf–Mandelbrot law is a discrete probability distribution depending on three parameters: $\mathcal{N} \in \{1, 2, \ldots, \}$, $\phi \in [0, \infty)$, and $t > 0$, and is defined by

$$f(s; \mathcal{N}, \phi, t) := \frac{1}{(s+\phi)^t \mathcal{H}_{\mathcal{N},\phi,t}}, \quad s = 1, \ldots, \mathcal{N},$$

where

$$\mathcal{H}_{\mathcal{N},\phi,t} = \sum_{\nu=1}^{\mathcal{N}} \frac{1}{(\nu+\phi)^t}.$$

For all values of $\mathcal{N}$, if the total mass of the law is taken, then for $0 \le \phi$, $1 < t$, $s \in \mathcal{N}$, the density function of the Zipf–Mandelbrot law becomes

$$f(s; \phi, t) = \frac{1}{(s+\phi)^t \mathcal{H}_{\phi,t}},$$

where

$$\mathcal{H}_{\phi,t} = \sum_{\nu=1}^{\infty} \frac{1}{(\nu+\phi)^t}.$$

For $\phi = 0$, the Zipf–Mandelbrot law becomes Zipf's law.

Conclusion 1 Assume $\mathcal{G}$.

Let $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{w}}$ be the Zipf–Mandelbrot laws. By Corollary 8(iii). If $\delta \in [0, 1)$ and base of log is greater than 1, then

$$\begin{aligned}\tilde{\mathcal{S}} &= -\sum_{s=1}^{n} \frac{1}{(s+k)^s \mathcal{H}_{\mathcal{N},k,\nu}} \log\left(\frac{1}{(s+k)^s \mathcal{H}_{\mathcal{N},k,\nu}}\right) - \frac{1}{1-\delta} \log\left(\frac{1}{\mathcal{H}^{\delta}_{\mathcal{N},k,\nu}} \sum_{s=1}^{n} \frac{1}{(s+k)^{\delta s}}\right) \\ &\ge \tilde{\mathcal{S}} \\ &= -\sum_{s=1}^{m} \frac{1}{(s+w)^s \mathcal{H}_{\mathcal{N},w,\nu}} \log\left(\frac{1}{(s+w)^s \mathcal{H}_{\mathcal{N},w,\nu}}\right) - \frac{1}{1-\delta} \log\left(\frac{1}{\mathcal{H}^{\delta}_{\mathcal{N},w,\nu}} \sum_{s=1}^{m} \frac{1}{(s+w)^{\delta s}}\right).\end{aligned}$$

The inequality is reversed if base of log is less than 1.

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 4th author is supported by the Ministry of Education and Science of the Russian Federation (the Agreement number No. 02.a03.21.0008).

Funding
There is no funding for this work.

Competing interests
The authors declare that there is no conflict of interests regarding the publication of this paper.

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

Author details
[1]Department of Mathematics, University of Sargodha, Sargodha, Pakistan. [2]Catholic University of Croatia, Ilica, Zagreb, Croatia. [3]Rudn University, Moscow, Russia.

Publisher's Note

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

Received: 14 March 2019 Accepted: 30 July 2019 Published online: 26 August 2019

References

1. Aras-Gazić, G., Čuljak, V., Pečarić, J., Vukelić, A.: Generalization of Jensen's inequality by Lidstone's polynomial and related results. Math. Inequal. Appl. **164**, 1243–1267 (2013)
2. Bullen, P.S.: An inequality of N. Levinson. Publ. Elektroteh. Fak. Univ. Beogr., Ser. Mat. Fiz. 109–112 (1973)
3. Cerone, P., Dragomir, S.S.: Some new Ostrowski-type bounds for the Čebyšev functional and applications. J. Math. Inequal. **8**(1), 159–170 (2014)
4. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hung. **2**, 299–318 (1967)
5. Csiszár, I.: Information measures: a critical survey. In: Tans. 7th Prague Conf. on Info. Th., Statist. Decis. Funct., Random Process and 8th European Meeting of Statist., vol. B, pp. 73–86. Academia, Prague ([illegible])
6. Diodato, V.: Dictionary of Bibliometrics. Haworth Press, New York (1994)
7. Egghe, L., Rousseau, R.: Introduction to Informetrics, Quantitative Methods in Library, Documentation and Information Science. Elsevier, New York (1990)
8. Gibbs, A.L.: On choosing and bounding probability metrics. Int. Stat. Rev. **70**, 419–435 (2002)
9. 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. Soc. **42**(3) 933–946 (2019)
10. Khan, K.A., Niaz, T., Pečarić, Đ., Pečarić, J.: Refinement of Jensen's inequality and estimation of *f* and Rényi divergence via Montgomery identity. J. Inequal. Appl. **2018**, 318 (2018)
11. Krnić, M., Lovričević, N., Pečarić, J.: Superadditivity of the Levinson functional and applications. Period. Math. Hung. **71**(2), 166–178 (2015)
12. Levinson, N.: Generalization of an inequality of Ky Fan. J. Math. Anal. Appl. **6**, 133–134 (1969)
13. Liese, F., Vajda, I.: Convex Statistical Distances. Teubner-Texte zur Mathematik, vol. 95. Teubner, Leipzig (1987)
14. Lovričević, N., Pečarić, Đ., Pečarić, J.: Zipf–Mandelbrot law, *f*-divergences and the Jensen-type interpolating inequalities. J. Inequal. Appl. **2018**, 36 (2018)
15. Mercer, A.McD.: A variant of Jensen's inequality. J. Inequal. Pure Appl. Math. **4**(4), article 73 (2003)
16. Mitrinović, D.S., Pečarić, J., Fink, A.M.: Classical and New Inequalities in Analysis, vol. 61. Kluwer Academic, Dordrecht (1992)
17. Pečarić, J.: On an inequality on N. Levinson. Publ. Elektroteh. Fak. Univ. Beogr., Ser. Mat. Fiz. 71–74 (1980)
18. Popoviciu, T.: Sur une inegalite de N. Levinson. Mathematica **6**, 301–306 (1969)
19. Rényi, A.: On measure of information and entropy. In: Proceeding of the Fourth Berkely Symposium on Mathematics, Statistics and Probability, pp. 547–561 (1960)
20. Sason, I., Verdú, S.: *f*-Divergence inequalities. IEEE Trans. Inf. Theory **62**, 5973–6006 (2016)
21. Vajda, I.: Theory of Statistical Inference and Information. Kluwer Academic, Dordrecht (1989)