

RESEARCH

Open Access



# Generalization of the Levinson inequality with applications to information theory

Muhammad Adeel<sup>1\*</sup>, Khuram Ali Khan<sup>1</sup>, Đilda Pečarić<sup>2</sup> and Josip Pečarić<sup>3</sup>

\*Correspondence:

adeel.uosmaths@gmail.com

<sup>1</sup>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 the 3-convex function is generalized by using two Green 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 the  $f$ -divergence, the Rényi divergence, the Rényi entropy, the Shannon entropy and the 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  $f : I = (0, 2\alpha) \rightarrow \mathbb{R}$  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

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

Working with the various 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  $f : I = [a, b] \rightarrow \mathbb{R}$  be a 3-convex function and  $x_n, y_n \in [a, b]$  for  $n = 1, 2, \dots, k$  such that*

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

and  $p_n > 0$ . Then

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

where

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

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

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

**Theorem C** Let  $f : I = [a, b] \rightarrow \mathbb{R}$  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 presented 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 function  $G_{\mathcal{F},n}(t, s)$  of the boundary value problem

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

where

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

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 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

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

where

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

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 the 3-convex function is generalized by using two Green 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 the  $f$ -divergence, the Rényi divergence, the Rényi entropy, the Shannon entropy and the Zipf–Mandelbrot law.

**2 Main results**

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

**Theorem 1** Let  $f \in C^3[\zeta_1, \zeta_2]$  such that  $f : I = [\zeta_1, \zeta_2] \rightarrow \mathbb{R}$ ,  $(p_1, \dots, p_n) \in \mathbb{R}^n$ ,  $(q_1, \dots, q_m) \in \mathbb{R}^m$  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

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

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

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

$$\begin{aligned}
 J(f(\cdot)) &= 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) \\
 &\quad + \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 \\
 &\quad - 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) \\
 &\quad - \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} \\
 &\quad - \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\varrho=1}^m q_\varrho y_\varrho, s\right) f^{(3)}(s) ds \\
 &\quad - 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) \\
 &\quad - \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 \\
 &\quad + 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) \\
 &\quad + \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} \\
 &\quad + \int_{\zeta_1}^{\zeta_2} G_1\left(\sum_{\rho=1}^n p_\rho x_\rho, s\right) f^{(3)}(s) ds, \\
 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}^n p_\rho x_\rho\right)^2 \right] f^{(2)}(\zeta_2) \\
 &\quad + \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 \\
 &\quad - \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  $f : I = [0, 2\alpha] \rightarrow \mathbb{R}$ ,  $x_1, \dots, x_n \in (0, \alpha)$ ,  $(p_1, \dots, p_n) \in \mathbb{R}^n$  such that  $\sum_{\rho=1}^n p_\rho = 1$ . Also let  $x_\rho, \sum_{\rho=1}^n p_\rho(2\alpha - x_\rho), \sum_{\rho=1}^n p_\rho x_\rho \in I$ . Then*

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

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

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

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

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

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.

*Proof* (18)  $\Rightarrow$  (19): Let (18) is 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)  $\Rightarrow$  (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) holds. Then we can represent the function  $f$  in the form (9). Now by means of some simple calculations we can write

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

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  $f : I \rightarrow \mathbb{R}$ , with  $f \in C^3[\zeta_1, \zeta_2]$ , (18) is valid.  $\square$

*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 results about generalization of Bullen’s type inequality (for real weights) given in [2] (see also [11, 16]).

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

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

and

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

Also let  $(p_1, \dots, p_n) \in \mathbb{R}^n, (q_1, \dots, q_m) \in \mathbb{R}^m$  such that  $\sum_{\rho=1}^n p_{\rho} = 1$  and  $\sum_{\varrho=1}^m q_{\varrho} = 1$  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  $x_{\rho}$  and  $y_{\varrho}$  such that conditions (20) and (21) hold in Theorem 2, we get required result. □

*Remark 2* If  $p_{\rho} = q_{\varrho}$  are positive and  $x_{\rho}, y_{\varrho}$  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 generalized form (for real weights) of Bullen’s type inequality given in [17] (see also [16]).

**Corollary 3** *Let  $f : I = [\zeta_1, \zeta_2] \rightarrow \mathbb{R}$  be a 3-convex function and  $f \in C^3[\zeta_1, \zeta_2]$ ,  $(p_1, \dots, p_n) \in \mathbb{R}^n, (q_1, \dots, q_m) \in \mathbb{R}^m$  such that  $\sum_{\rho=1}^n p_{\rho} = 1$  and  $\sum_{\varrho=1}^m q_{\varrho} = 1$ . Also let  $x_1, \dots, x_n$  and  $y_1, \dots, y_m \in I$  such that  $x_{\rho} + y_{\varrho} = 2c$  and for  $\rho = 1, \dots, n, x_{\rho} + x_{n-\rho+1}$  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 generalized form of Bullen’s inequality defined in [16, p. 32, Theorem 4].

In [15], Mercer made a notable work by replacing the condition (21) of symmetric distribution of points  $x_{\rho}$  and  $y_{\varrho}$  with symmetric variances of points  $x_{\rho}$  and  $y_{\varrho}$  for  $\rho = 1, \dots, n$  and  $\varrho = 1, \dots, m$ .

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

**Corollary 4** Let  $f : I = [\zeta_1, \zeta_1] \rightarrow \mathbb{R}$  be 3-convex function and  $f \in C^3[\zeta_1, \zeta_2]$ ,  $p_\rho, q_\rho$  are positive such that  $\sum_{\rho=1}^n p_\rho = 1$  and  $\sum_{\rho=1}^n q_\rho = 1$ . Also let  $x_\rho, y_\rho$  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 required result.  $\square$

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

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

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

For all  $s \in I$

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

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

*Proof* If  $I = [0, 2\alpha]$ ,  $(x_1, \dots, x_n) \in (0, \alpha)$ ,  $p_\rho = q_\rho$ ,  $m = n$  and  $y_\rho = (2\alpha - x_\rho)$  in Theorem 2 with  $0 \leq \zeta_1 < \zeta_2 \leq 2\alpha$  we get required result.  $\square$

**Remark 4** In Corollary 5 if  $p_\rho$  are positive then inequality (23) reduces to Levinson’s inequality given in [16, p. 32, Theorem 1].

### 3 New bounds for Levinson’s type functionals

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

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

where the integrals are assumed to exist.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Using the identity (13), we get (28). □

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

**Theorem 4** *Let  $f \in C^3[\zeta_1, \zeta_2]$  such that  $f : I = [\zeta_1, \zeta_2] \rightarrow \mathbb{R}, f^{(3)}(\cdot)$  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  $\sum_{\rho=1}^n p_\rho = 1, \sum_{\varrho=1}^m q_\varrho = 1, x_\rho, y_\varrho, \sum_{\rho=1}^n p_\rho x_\rho, \sum_{\varrho=1}^m q_\varrho y_\varrho \in I$ . Then identity (28) holds, where the remainder satisfies the estimation*

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

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

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

Since

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

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

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

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

$\frac{1}{r} + \frac{1}{s} = 1$ . If  $|f^{(3)}|^r : [\zeta_1, \zeta_2] \rightarrow \mathbb{R}$  be Riemann integrable function, then

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

*Proof* Rearrange identity (13) in such a way

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

Employing the classical Hölder inequality for the R.H.S. of (34) yields (33). □

### 4 Application to information theory

The idea of the Shannon entropy is the focal point of data hypothesis, now and then al- luded to as the 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 ran- domness or uncertainty. The Shannon entropy permits one 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 become familiar with a measure of the difference between them. A variety of sorts of divergences exist, for instance the  $f$ - difference (particularly, the Kullback–Leibler divergence, the Hellinger distance and the total variation distance), the Rényi divergence, the 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. The Jensen inequality assumes a crucial role in a part of these inequalities. In any case, Jensen’s inequality deals with one sort of information focus and Levinson’s inequality manages two type information points.

The Zipf law is one of the central laws in data science, and it has been utilized in lin- guistics. 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 a steady ( $C$ ) :  $C = r \times f$ . Aside from the utilization of this law in data science and linguistics, the Zipf law is utilized in city popu- lation, sun powered flare power, site traffic, earthquack magnitude, the span of moon pits, and so forth. In financial aspects this distribution is known as the Pareto law, which ana- lyzes 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 the  $f$ -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:

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

he stated that nonnegative probability distributions can also be used.

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

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

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

Then

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

**Theorem 6** Let  $\tilde{\mathbf{r}} = (r_1, \dots, r_n), \tilde{\mathbf{k}} = (k_1, \dots, k_n)$  be in  $(0, \infty)^n$  and  $\tilde{\mathbf{w}} = (w_1, \dots, w_m), \tilde{\mathbf{t}} = (t_1, \dots, t_m)$  are in  $(0, \infty)^m$  such that

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

and

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

If

$$\left[ \frac{1}{\sum_{u=1}^m t_u} \sum_{u=1}^m \frac{(w_u)^2}{t_u} - \left( \sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u} \right)^2 - \frac{1}{\sum_{v=1}^n k_v} \sum_{v=1}^n \frac{(r_v)^2}{k_v} + \left( \sum_{v=1}^n \frac{r_v}{\sum_{v=1}^n k_v} \right)^2 \right] f^{(2)}(\zeta_2) \geq 0, \tag{35}$$

then the following are equivalent.

- (i) For every continuous 3-convex function  $f : I \rightarrow \mathbb{R}$ ,

$$J_f(r, w, k, t) \geq 0. \tag{36}$$

(ii)

$$J_{G_k}(r, w, k, t) \geq 0, \tag{37}$$

where

$$J_f(r, w, k, t) = \frac{1}{\sum_{u=1}^m t_u} \hat{I}_f(\tilde{w}, \tilde{t}) - f\left(\sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u}\right) - \frac{1}{\sum_{v=1}^n k_v} \hat{I}_f(\tilde{r}, \tilde{k}) + f\left(\sum_{v=1}^n \frac{r_v}{\sum_{v=1}^n k_v}\right). \tag{38}$$

*Proof* Using  $p_v = \frac{k_v}{\sum_{v=1}^n k_v}$ ,  $x_v = \frac{r_v}{k_v}$ ,  $q_u = \frac{t_u}{\sum_{u=1}^m t_u}$  and  $y_u = \frac{w_u}{t_u}$  in Theorem 2, we get the required results.  $\square$

### 4.2 Shannon entropy

**Definition 3** (See [9]) The Shannon entropy of a positive probability distribution  $\tilde{r} = (r_1, \dots, r_n)$  is defined by

$$\mathcal{S} := - \sum_{v=1}^n r_v \log(r_v). \tag{39}$$

**Corollary 6** Let  $\tilde{r} = (r_1, \dots, r_n)$  and  $\tilde{w} = (w_1, \dots, w_m)$  be probability distributions,  $\tilde{k} = (k_1, \dots, k_n)$  be in  $(0, \infty)^n$  and  $\tilde{t} = (t_1, \dots, t_m)$  be in  $(0, \infty)^m$ . If

$$\left[ \frac{1}{\sum_{u=1}^m t_u} \sum_{u=1}^m \frac{(w_u)^2}{t_u} - \left(\sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u}\right)^2 - \frac{1}{\sum_{v=1}^n k_v} \sum_{v=1}^n \frac{(r_v)^2}{k_v} + \left(\sum_{v=1}^n \frac{r_v}{\sum_{v=1}^n k_v}\right)^2 \right] \geq 0 \tag{40}$$

and

$$J_{G_k}(r, w, k, t) \leq 0, \tag{41}$$

then

$$J_s(r, w, k, t) \leq 0, \tag{42}$$

where

$$J_s(r, w, k, t) = \frac{1}{\sum_{u=1}^m t_u} \left[ \tilde{\mathcal{S}} + \sum_{u=1}^m w_u \log(t_u) \right] + \left[ \sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u} \log\left(\sum_{u=1}^m \frac{w_u}{\sum_{u=1}^m t_u}\right) \right] - \frac{1}{\sum_{v=1}^n k_v} \left[ \mathcal{S} - \sum_{v=1}^n r_v \log(k_v) \right] - \left[ \sum_{v=1}^n \frac{r_v}{\sum_{v=1}^n k_v} \log\left(\sum_{v=1}^n \frac{r_v}{\sum_{v=1}^n k_v}\right) \right]; \tag{43}$$

$\mathcal{S}$  is defined in (39) and

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

If the base of the log is less than 1, then (42) and (41) hold in reverse direction.

*Proof* The function  $f(x) \mapsto -x \log(x)$  is 3-convex for a base of the log is greater than 1. So using  $f(x) := -x \log(x)$  in (35) and (36), we get the required results by Remark 1.  $\square$

### 4.3 Rényi divergence and entropy

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

**Definition 4** 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  $\delta$  is defined by

$$\mathcal{D}_\delta(\tilde{\mathbf{r}}, \tilde{\mathbf{q}}) := \frac{1}{\delta - 1} \log \left( \sum_{i=1}^n q_i \left( \frac{r_i}{q_i} \right)^\delta \right). \tag{44}$$

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

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

These definitions also hold for non-negative probability distributions.

**Theorem 7** Let  $\tilde{\mathbf{r}} = (r_1, \dots, r_n), \tilde{\mathbf{k}} = (k_1, \dots, k_n) \in \mathbb{R}_+^n, \tilde{\mathbf{w}} = (w_1, \dots, w_m), \tilde{\mathbf{t}} = (t_1, \dots, t_m) \in \mathbb{R}_+^m$  be such that  $\sum_1^n r_v = 1, \sum_1^n k_v = 1, \sum_1^m w_u = 1$  and  $\sum_1^m t_u = 1$ .

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

$$\left[ \sum_{u=1}^m \frac{(t_u)^2}{w_u} \left( \frac{w_u}{t_u} \right)^{2\delta} - \left( \sum_{u=1}^m t_u \left( \frac{w_u}{t_u} \right)^\delta \right)^2 - \sum_{v=1}^n \frac{(k_v)^2}{r_v} \left( \frac{r_v}{k_v} \right)^{2\delta} + \left( \sum_{v=1}^n k_v \left( \frac{r_v}{k_v} \right)^\delta \right)^2 \right] \geq 0 \tag{46}$$

and

$$\sum_{v=1}^n r_v G_k \left( \left( \frac{r_v}{k_v} \right)^{\delta-1}, s \right) - G_k \left( \sum_{v=1}^n r_v \left( \frac{r_v}{k_v} \right)^{\delta-1}, s \right) \geq \sum_{u=1}^m w_u G_k \left( \left( \frac{w_u}{t_u} \right)^{\delta-1}, s \right) - G_k \left( \sum_{u=1}^m w_u \left( \frac{w_u}{t_u} \right)^{\delta-1}, s \right), \tag{47}$$

then

$$\sum_{v=1}^n r_v \log \left( \frac{r_v}{k_v} \right) - \mathcal{D}_\delta(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) \geq \sum_{u=1}^m w_u \log \left( \frac{w_u}{t_u} \right) - \mathcal{D}_\delta(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}). \tag{48}$$

If either  $1 < \delta$  and the base of the log is greater than 1 or  $\delta \in [0, 1)$  and the base of the log is less than 1, then (47) and (48) hold in reverse direction.

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

Choosing  $I = (0, \infty)$   $f = \log$  so  $f^{(2)}(x)$  is negative and  $f^{(3)}(x)$  is positive, therefore  $f$  is 3-convex. So using  $f = \log$  and the substitutions

$$p_v := r_v, \quad x_v := \left(\frac{r_v}{k_v}\right)^{\delta-1}, \quad v = 1, \dots, n,$$

and

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

in the reverse of inequality (18) (by Remark 1), we have

$$(\delta - 1) \sum_{v=1}^n r_v \log\left(\frac{r_v}{k_v}\right) - \log\left(\sum_{v=1}^n k_v \left(\frac{r_v}{k_v}\right)^\delta\right) \geq (\delta - 1) \sum_{u=1}^m w_u \log\left(\frac{w_u}{t_u}\right) - \log\left(\sum_{u=1}^m t_u \left(\frac{w_u}{t_u}\right)^\delta\right). \tag{49}$$

Dividing (49) with  $(\delta - 1)$  and using

$$\mathcal{D}_\delta(\tilde{\mathbf{r}}, \tilde{\mathbf{k}}) = \frac{1}{\delta - 1} \log\left(\sum_{v=1}^n k_v \left(\frac{r_v}{k_v}\right)^\delta\right),$$

$$\mathcal{D}_\delta(\tilde{\mathbf{w}}, \tilde{\mathbf{t}}) = \frac{1}{\delta - 1} \log\left(\sum_{u=1}^m t_u \left(\frac{w_u}{t_u}\right)^\delta\right),$$

we get (48). □

**Corollary 7** Let  $\tilde{\mathbf{r}} = (r_1, \dots, r_n) \in \mathbb{R}_+^n$ ,  $\tilde{\mathbf{w}} = (w_1, \dots, w_m) \in \mathbb{R}_+^m$  be such that  $\sum_1^n r_v = 1$  and  $\sum_1^m w_u = 1$ .

Also let

$$\left[ \sum_{u=1}^m \frac{1}{m^2 w_u} (m w_u)^{2\delta} - \left(\sum_{u=1}^m \frac{1}{m} (m w_u)^\delta\right)^2 - \sum_{v=1}^n \frac{1}{n^2 r_v} (n r_v)^{2\delta} + \left(\sum_{v=1}^n \frac{1}{n} (n r_v)^\delta\right)^2 \right] \geq 0 \tag{50}$$

and

$$\sum_{v=1}^n r_v G_k((n r_v)^{\delta-1}, s) - G_k\left(\sum_{v=1}^n r_v (n r_v)^{\delta-1}, s\right) \geq \sum_{u=1}^m w_u G_k((m w_u)^{\delta-1}, s) - G_k\left(\sum_{u=1}^m w_u (m w_u)^{\delta-1}, s\right). \tag{51}$$

If  $1 < \delta$  and the base of the log is greater than 1, then

$$\sum_{v=1}^n r_v \log(r_v) + \mathcal{H}_\delta(\tilde{\mathbf{r}}) \geq \sum_{u=1}^m w_u \log(w_u) + \mathcal{H}_\delta(\tilde{\mathbf{w}}). \tag{52}$$

The reverse inequality holds in (51) and (52) if the base of the log is less than 1.

*Proof* Suppose  $\tilde{\mathbf{k}} = (\frac{1}{n}, \dots, \frac{1}{n})$  and  $\tilde{\mathbf{t}} = (\frac{1}{m}, \dots, \frac{1}{m})$ . Then from (44), we have

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

and

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

This implies

$$\mathcal{H}_\delta(\tilde{\mathbf{r}}) = \log(n) - \mathcal{D}_\delta\left(\tilde{\mathbf{r}}, \frac{\mathbf{1}}{\mathbf{n}}\right) \tag{53}$$

and

$$\mathcal{H}_\delta(\tilde{\mathbf{w}}) = \log(m) - \mathcal{D}_\delta\left(\tilde{\mathbf{w}}, \frac{\mathbf{1}}{\mathbf{m}}\right). \tag{54}$$

It follows from Theorem 7,  $\tilde{\mathbf{k}} = \frac{1}{n}$ ,  $\tilde{\mathbf{t}} = \frac{1}{m}$ , (53) and (54), that

$$\sum_{v=1}^n r_v \log(nr_v) - \log(n) + \mathcal{H}_\delta(\tilde{\mathbf{r}}) \geq \sum_{u=1}^m w_u \log(mw_u) - \log(m) + \mathcal{H}_\delta(\tilde{\mathbf{w}}). \tag{55}$$

After some simple calculations we get (52). □

#### 4.4 Zipf–Mandelbrot law

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

**Definition 5** Zipf–Mandelbrot law is a discrete probability distribution depending on the three parameters:  $\mathcal{N} \in \{1, 2, \dots\}$ ,  $\phi \in [0, \infty)$  and  $t > 0$ , and is defined by

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

where

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

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

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

where

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

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

**Theorem 8** *Let  $\tilde{r}$  and  $\tilde{w}$  be the Zipf–Mandelbrot laws.*

*If (50) and (51) hold for  $r_v = \frac{1}{(v+k)^v \mathcal{H}_{\mathcal{N},k,v}}$ ,  $w_u = \frac{1}{(u+w)^u \mathcal{H}_{\mathcal{N},w,u}}$ , and if the base of the log is greater than 1, then*

$$\begin{aligned} & \sum_{v=1}^n \frac{1}{(v+k)^v \mathcal{H}_{\mathcal{N},k,v}} \log\left(\frac{1}{(v+k)^v \mathcal{H}_{\mathcal{N},k,v}}\right) + \frac{1}{1-\delta} \log\left(\frac{1}{\mathcal{H}_{\mathcal{N},k,v}^\delta} \sum_{v=1}^n \frac{1}{(v+k)^{\delta v}}\right) \\ & \geq \sum_{u=1}^m \frac{1}{(u+w)^u \mathcal{H}_{\mathcal{N},w,u}} \log\left(\frac{1}{(u+w)^u \mathcal{H}_{\mathcal{N},w,u}}\right) \\ & \quad + \frac{1}{1-\delta} \log\left(\frac{1}{\mathcal{H}_{\mathcal{N},w,u}^\delta} \sum_{u=1}^m \frac{1}{(u+w)^{\delta u}}\right). \end{aligned} \tag{56}$$

*The inequality is reversed in (51) and (56) if the base of the log is less than 1.*

*Proof* The proof is similar to Corollary 7; by using Definition 5 and the hypothesis given in the statement we get the required result. □

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

**Author details**

<sup>1</sup>Department of Mathematics, University of Sargodha, Sargodha, Pakistan. <sup>2</sup>Catholic University of Croatia, Ilica, Zagreb, Croatia. <sup>3</sup>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: 27 July 2019 Published online: 29 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. Csizsár, I.: Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hung.* **2**, 299–318 (1967)
5. Csizsár, 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 (1978)
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. Sci. Soc.* **1**(14) (2017)
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., Lovrićević, 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 Kay 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. Lovrićević, 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), 73 (2003)
16. Mitrinović, D.S., Pečarić, J., Fink, A.M.: *Classical and New Inequalities in Analysis*, vol. 61. Kluwer Academic, Norwell (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)

Submit your manuscript to a SpringerOpen<sup>®</sup> 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)

---