Khuram Ali Khan · Tasadduq Niaz · Dilda Pečarić · Josip Pečarić

Estimation of different entropies via Lidstone polynomial using Jensen-type functionals

Received: 5 November 2018 / Accepted: 1 February 2020 / Published online: 18 February 2020
© The Author(s) 2020

Abstract In this work, some new functional of Jensen-type inequalities are constructed using Shannon entropy, f -divergence, and Rényi divergence, and some estimates are obtained for these new functionals. Also using the Zipf–Mandelbrot law and hybrid Zipf–Mandelbrot law, we investigate some bounds for these new functionals. Furthermore, we generalize these new functionals for m -convex function using Lidstone polynomial.

Mathematics Subject Classification 94Axx · 39-XX · 41A58

1 Introduction and preliminary results

The most commonly used words, population ranks of cities in various countries, corporation sizes, income rankings can be described in terms of Zipf's law. The f -divergence measures the difference between two probability distributions by making an average value, which is weighted by a specified function. There are other probability distributions like Csiszar f -divergence [10, 11], a special case of which is Kullback–Leibler-divergence which is used to find the appropriate distance between the probability distributions (see [18, 19]). The notion of distance is stronger than divergence, because it gives the properties of symmetry and triangle inequalities. Probability theory has application in many fields and the divergence between probability distributions has many applications in these fields.

The research of fourth author was supported by the Ministry of Education and Science of the Russian Federation (the Agreement number No. 02.a03.21.0008).

K. A. Khan · T. Niaz (✉)
Department of Mathematics, University of Sargodha, Sargodha 40100, Pakistan
E-mail: tasadduq_khan@yahoo.com

K. A. Khan
E-mail: khuramsms@gmail.com

T. Niaz
Department of Mathematics, The University of Lahore, Sargodha-Campus, Sargodha 40100, Pakistan

Dilda Pečarić
Catholic University of Croatia, Ilica, 242 Zagreb, Croatia
E-mail: gildapecarica@gmail.com

Josip Pečarić
RUDN University, Miklukho-Maklaya str. 6, 117198 Moscow, Russia
E-mail: pecaric@element.hr

Many natural phenomena such as distributions of wealth and income in a society, Facebook likes, football goals, and city sizes follow power-law distributions (Zipf's Law). Auerbach [2] was the first to explore the idea that the distribution of city size can be well approximated with the help of Pareto distribution (power-law distribution). This idea was well refined by many researchers, but Zipf [27] worked significantly in this field. The distribution of city sizes is investigated by many scholars of the urban economics, like Rosen and Resnick [24], Black and Henderson [3], Ioannides and Overman [17], Soo [25], Anderson and Ge [1], and Bosker et al. [4]. Zipf's law states that: "The rank of cities with a certain number of inhabitants varies proportional to the city sizes with some negative exponent, say that is close to unit". In other words, Zipf's Law states that the product of city sizes and their ranks appears roughly constant. This indicates that the population of the second largest city is one half of the population of the largest city and the third largest city equal to one-third of the population of the largest city and the population of n -th city is $\frac{1}{n}$ of the largest city population. This rule is called rank, size rule, and also named as Zipf's Law. Hence, Zip's law shows that the city size distribution follows the Pareto distribution and, in addition, that the estimated value of the shape parameter is equal to unity.

Horváth et al. [16] introduced and obtained some estimates for new functionals based on the f -divergence functionals, and obtained some estimates for the new functionals. They obtained f -divergence and Rényi divergence by applying a cyclic refinement of Jensen's inequality. They also constructed new inequalities for Rényi and Shannon entropies and used Zipf–Madelbrot law to illustrate the results.

The inequalities involving higher order convexity are used by many physicists in higher dimension problems since the founding of higher order convexity by Popoviciu (see [22, p. 15]). It is quite interesting that some results are true for convex functions, but in higher order convexity, they are not valid anymore.

In [22, p. 16], the following criteria are given to check the m -convexity of the function. If $f^{(m)}$ exists, then f is m -convex if and only if $f^{(m)} \geq 0$.

In recent years, many researchers have generalized the inequalities for m -convex functions. For instance, Butt et al. generalized the Popoviciu's inequality for m -convex function using Taylor's formula, Lidstone polynomial, Montgomery identity, Fink's identity, Abel–Goncharov interpolation, and Hermite interpolating polynomial (see [5–9]).

For many years, Jensen's inequality has been of great interest. It was refined by defining some new functions (see [14, 15]). Horváth and Pečarić ([12, 15], see also [13, p. 26]) gave a refinement of Jensen's inequality for convex function. They defined some essential notions to prove the refinement given as follows:

Let X be a set, and: $P(X) :=$ Power set of X , $|X| :=$ Number of elements of X , $\mathbb{N} :=$ Set of natural numbers with 0. Consider $q \geq 1$ and $r \geq 2$ be fixed integers. Define the functions:

$$F_{r,s} : \{1, \dots, q\}^r \rightarrow \{1, \dots, q\}^{r-1} \quad 1 \leq s \leq r,$$

$$F_r : \{1, \dots, q\}^r \rightarrow P(\{1, \dots, q\}^{r-1}),$$

and

$$T_r : P(\{1, \dots, q\}^r) \rightarrow P(\{1, \dots, q\}^{r-1}),$$

by

$$F_{r,s}(i_1, \dots, i_r) := (i_1, i_2, \dots, i_{s-1}, i_{s+1}, \dots, i_r) \quad 1 \leq s \leq r,$$

$$F_r(i_1, \dots, i_r) := \bigcup_{s=1}^r \{F_{r,s}(i_1, \dots, i_r)\},$$

and

$$T_r(I) = \begin{cases} \phi, & I = \phi; \\ \bigcup_{(i_1, \dots, i_r) \in I} F_r(i_1, \dots, i_r), & I \neq \phi. \end{cases}$$

Next, the function

$$\alpha_{r,i} : \{1, \dots, q\}^r \rightarrow \mathbb{N} \quad 1 \leq i \leq q$$

is defined by:

$$\alpha_{r,i}(i_1, \dots, i_r) \text{ is the number of occurrences of } i \text{ in the sequence } (i_1, \dots, i_r).$$

For each $I \in P(\{1, \dots, q\}^r)$, let

$$\alpha_{I,i} := \sum_{(i_1, \dots, i_r) \in I} \alpha_{r,i}(i_1, \dots, i_r) \quad 1 \leq i \leq q.$$

(H_1) Let n, m be fixed positive integers, such that $n \geq 1, m \geq 2$, and let I_m be a subset of $\{1, \dots, n\}^m$, such that:

$$\alpha_{I_m,i} \geq 1 \quad 1 \leq i \leq n.$$

Introduce the sets $I_l \subset \{1, \dots, n\}^l$ ($m-1 \geq l \geq 1$) inductively by:

$$I_{l-1} := T_l(I_l) \quad m \geq l \geq 2.$$

Obviously, the sets $I_1 = \{1, \dots, n\}$, by (H_1), and this insures that $\alpha_{I_1,i} = 1$ ($1 \leq i \leq n$). From (H_1), we have $\alpha_{I_l,i} \geq 1$ ($m-1 \geq l \geq 1, 1 \leq i \leq n$).

For $m \geq l \geq 2$, and for any $(j_1, \dots, j_{l-1}) \in I_{l-1}$, let:

$$\mathcal{H}_l(j_1, \dots, j_{l-1}) := \{((i_1, \dots, i_l), k) \times \{1, \dots, l\} | F_{l,k}(i_1, \dots, i_l) = (j_1, \dots, j_{l-1})\}.$$

With the help of these sets, they define the functions $\eta_{I_m,l} : I_l \rightarrow \mathbb{N}$ ($m \geq l \geq 1$) inductively by:

$$\begin{aligned} \eta_{I_m,m}(i_1, \dots, i_m) &:= 1 \quad (i_1, \dots, i_m) \in I_m; \\ \eta_{I_m,l-1}(j_1, \dots, j_{l-1}) &:= \sum_{((i_1, \dots, i_l), k) \in \mathcal{H}_l(j_1, \dots, j_{l-1})} \eta_{I_m,l}(i_1, \dots, i_l). \end{aligned}$$

They define some special expressions for $1 \leq l \leq m$, as follows:

$$\begin{aligned} \mathcal{A}_{m,l} = \mathcal{A}_{m,l}(I_m, x_1, \dots, x_n, p_1, \dots, p_n; f) &:= \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m,l}(i_1, \dots, i_l) \\ &\times \left(\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m,i_j}} \right) f \left(\frac{\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m,i_j}} x_{i_j}}{\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m,i_j}}} \right) \end{aligned}$$

and prove the following theorem.

Theorem 1.1 Assume (H_1), and let $f : I \rightarrow \mathbb{R}$ be a convex function where $I \subset \mathbb{R}$ is an interval. If $x_1, \dots, x_n \in I$, and $p_1, \dots, p_n$ are positive real numbers, such that $\sum_{i=1}^n p_i = 1$, then

$$f \left(\sum_{s=1}^n p_s x_s \right) \leq \mathcal{A}_{m,m} \leq \mathcal{A}_{m,m-1} \leq \dots \leq \mathcal{A}_{m,2} \leq \mathcal{A}_{m,1} = \sum_{s=1}^n p_s f(x_s). \quad (1)$$

We define the following functionals by taking the differences of refinement of Jensen's inequality given in (1):

$$\Theta_1(f) = \mathcal{A}_{m,r} - f \left(\sum_{s=1}^n p_s x_s \right), \quad r = 1, \dots, m, \quad (2)$$

$$\Theta_2(f) = \mathcal{A}_{m,r} - \mathcal{A}_{m,k}, \quad 1 \leq r < k \leq m. \quad (3)$$

Under the assumptions of Theorem 1.1, we have:

$$\Theta_i(f) \geq 0, \quad i = 1, 2. \quad (4)$$

Inequalities (4) are reversed if f is concave on I .

1.1 Lidstone polynomial

We generalize the refinement of Jensen's inequality for higher order convex function using Lidstone interpolating polynomial. In [26], Widder gives the following result.

Lemma A If $g \in C^\infty([0, 1])$, then:

$$g(u) = \sum_{l=0}^{m-1} \left[g^{(2l)}(0) \mathfrak{F}_l(1-u) + g^{(2l)}(0) \mathfrak{F}_l(t) \right] + \int_0^1 G_m(u, s) g^{(2m)}(s) ds,$$

where $\mathfrak{F}_l$ is a polynomial of degree $2l + 1$ defined by the relation:

$$\mathfrak{F}_0(u) = u, \quad \mathfrak{F}_m''(u) = \mathfrak{F}_{m-1}(u), \quad \mathfrak{F}_m(0) = \mathfrak{F}_m(1) = 0, \quad m \geq 1, \quad (5)$$

and

$$G_1(u, s) = G(u, s) = \begin{cases} (u-1)s, & \alpha_1 \leq s \leq u \leq \alpha_2; \\ (s-1)u, & \alpha_1 \leq u \leq s \leq \alpha_2, \end{cases}$$

is a homogeneous Green's function of the differential operator $\frac{d^2}{ds^2}$ on $[0, 1]$, and with the successive iterates of $G(u, s)$:

$$G_m(u, s) = \int_0^1 G_1(u, p) G_{m-1}(p, s) dp, \quad m \geq 2.$$

The Lidstone polynomial can be expressed in terms of $G_m(u, s)$ as:

$$\mathfrak{F}_m(u) = \int_0^1 G_m(u, s) s ds.$$

Lidstone series representation of $g \in C^{2m}[\alpha_1, \alpha_2]$ is given by:

$$\begin{aligned} g(u) = & \sum_{l=0}^{m-1} (\alpha_2 - \alpha_1)^{2l} g^{(2l)}(\alpha_1) \mathfrak{F}_l\left(\frac{\alpha_2 - u}{\alpha_2 - \alpha_1}\right) + \sum_{l=0}^{m-1} (\alpha_2 - \alpha_1)^{2l} g^{(2l)}(\alpha_2) \mathfrak{F}_l\left(\frac{u - \alpha_1}{\alpha_2 - \alpha_1}\right) \\ & + (\alpha_2 - \alpha_1)^{2l-1} \int_{\alpha_1}^{\alpha_2} G_m\left(\frac{u - \alpha_1}{\alpha_2 - \alpha_1}, \frac{t - \alpha_1}{\alpha_2 - \alpha_1}\right) g^{(2l)}(t) dt. \end{aligned} \quad (6)$$

2 Inequalities for Csiszár divergence

In [10, 11], Csiszár introduced the following notion.

Definition 2.1 Let $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a convex function, let $\mathbf{r} = (r_1, \dots, r_n)$, and $\mathbf{q} = (q_1, \dots, q_n)$ be positive probability distributions. Then, f -divergence functional is defined by:

$$I_f(\mathbf{r}, \mathbf{q}) := \sum_{i=1}^n q_i f\left(\frac{r_i}{q_i}\right). \quad (7)$$

And he stated that by defining:

$$f(0) := \lim_{x \rightarrow 0^+} f(x); \quad 0f\left(\frac{0}{0}\right) := 0; \quad 0f\left(\frac{a}{0}\right) := \lim_{x \rightarrow 0^+} xf\left(\frac{a}{x}\right), \quad a > 0, \quad (8)$$

we can also use the non-negative probability distributions, as well.

Horváth et al. [16] gave the following functional based on the previous definition.

Definition 2.2 Let $I \subset \mathbb{R}$ be an interval and let $f : I \rightarrow \mathbb{R}$ be a function. Let $\mathbf{r} = (r_1, \dots, r_n) \in \mathbb{R}^n$ and $\mathbf{q} = (q_1, \dots, q_n) \in (0, \infty)^n$, such that:

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

Then, they define the sum $\hat{I}_f(\mathbf{r}, \mathbf{q})$ as:

$$\hat{I}_f(\mathbf{r}, \mathbf{q}) := \sum_{s=1}^n q_s f\left(\frac{r_s}{q_s}\right). \quad (9)$$

We apply Theorem 1.1 to $\hat{I}_f(\mathbf{r}, \mathbf{q})$.

Theorem 2.3 Assume (H_1) , let $I \subset \mathbb{R}$ be an interval and let $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{q} = (q_1, \dots, q_n)$ be in $(0, \infty)^n$, such that

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

(i) If $f : I \rightarrow \mathbb{R}$ is a convex function, then:

$$\hat{I}_f(\mathbf{r}, \mathbf{q}) = \sum_{s=1}^n q_s f\left(\frac{r_s}{q_s}\right) = A_{m,1}^{[1]} \geq A_{m,2}^{[1]} \geq \dots \geq A_{m,m-1}^{[1]} \geq A_{m,m}^{[1]} \geq f\left(\frac{\sum_{s=1}^n r_s}{\sum_{s=1}^n q_s}\right) \sum_{s=1}^n q_s, \quad (10)$$

where

$$A_{m,l}^{[1]} = \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}} \right) f\left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right). \quad (11)$$

If f is a concave function, then inequality signs in (10) are reversed.

(ii) If $f : I \rightarrow \mathbb{R}$ is a function, such that $x \rightarrow xf(x)$ ($x \in I$) is convex, then:

$$\left(\sum_{s=1}^n r_s \right) f\left(\frac{\sum_{s=1}^n r_s}{\sum_{s=1}^n q_s} \right) \leq A_{m,m}^{[2]} \leq A_{m,m-1}^{[2]} \leq \dots \leq A_{m,2}^{[2]} \leq A_{m,1}^{[2]} = \sum_{s=1}^n r_s f\left(\frac{r_s}{q_s}\right) = \hat{I}_{idf}(\mathbf{r}, \mathbf{q}), \quad (12)$$

where

$$A_{m,l}^{[2]} = \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}} \right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right) f\left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right).$$

Proof (i) Consider $p_s = \frac{q_s}{\sum_{s=1}^n q_s}$ and $x_s = \frac{r_s}{q_s}$ in Theorem 1.1, we have:

$$\begin{aligned} f\left(\frac{\sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s}}\right) &\leq \dots \leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \\ &\times \left(\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m, i_j}} \right) f\left(\frac{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s} \frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}}} \right) \leq \dots \leq \sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} f\left(\frac{r_s}{q_s}\right). \end{aligned} \quad (13)$$

And taking the sum $\sum_{s=1}^n q_i$, we have (10).

(ii) Using $f := id$ (where “ id ” is the identity function) in Theorem 1.1, we have:

$$\sum_{s=1}^n p_s x_s f\left(\sum_{s=1}^n p_s x_s\right) \leq \dots \leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \\ \times \left(\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m, i_j}}\right) \left(\frac{\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m, i_j}} x_{i_j}}{\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m, i_j}}}\right) f\left(\frac{\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m, i_j}} x_{i_j}}{\sum_{j=1}^l \frac{p_{i_j}}{\alpha_{I_m, i_j}}}\right) \leq \dots \leq \sum_{s=1}^n p_s x_s f(x_s). \quad (14)$$

Now, using $p_s = \frac{q_s}{\sum_{s=1}^n q_s}$ and $x_s = \frac{r_s}{q_s}$, $s = 1, \dots, n$, we get:

$$\sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s} f\left(\sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s}\right) \leq \dots \leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \\ \times \left(\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m, i_j}}\right) \left(\frac{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s} \frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m, i_j}}}}{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m, i_j}}}\right) f\left(\frac{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s} \frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m, i_j}}}}{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m, i_j}}}\right) \leq \dots \leq \sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s} f\left(\frac{r_s}{q_s}\right). \quad (15)$$

By taking sum $\sum_{s=1}^n q_s$ on both sides, we get (12). $\square$

3 Inequalities for Shannon entropy

Definition 3.1 (See [16]) The Shannon entropy of positive probability distribution $\mathbf{r} = (r_1, \dots, r_n)$ is defined by:

$$S := - \sum_{s=1}^n r_s \log(r_s). \quad (16)$$

Corollary 3.2 Assume (H_1) .

(i) If $\mathbf{q} = (q_1, \dots, q_n) \in (0, \infty)^n$, and the base of $\log$ is greater than 1, then:

$$S \leq A_{m,m}^{[3]} \leq A_{m,m-1}^{[3]} \leq \dots \leq A_{m,2}^{[3]} \leq A_{m,1}^{[3]} = \log\left(\frac{n}{\sum_{s=1}^n q_s}\right) \sum_{s=1}^n q_s, \quad (17)$$

where

$$A_{m,l}^{[3]} = - \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}\right) \log\left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}\right). \quad (18)$$

If the base of $\log$ is between 0 and 1, then inequality signs in (17) are reversed.

(ii) If $\mathbf{q} = (q_1, \dots, q_n)$ is a positive probability distribution and the base of $\log$ is greater than 1, then we have the estimates for the Shannon entropy of $\mathbf{q}$:

$$S \leq A_{m,m}^{[4]} \leq A_{m,m-1}^{[4]} \leq \dots \leq A_{m,2}^{[4]} \leq A_{m,1}^{[4]} = \log(n), \quad (19)$$

where

$$A_{m,l}^{[4]} = - \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}\right) \log\left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}\right).$$

Proof (i) Using $f := \log$ and $\mathbf{r} = (1, \dots, 1)$ in Theorem 2.3 (i), we get (17).

(ii) It is a special case of (i). $\square$

Definition 3.3 (See [16]) The Kullback–Leibler divergence between the positive probability distribution $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{q} = (q_1, \dots, q_n)$ is defined by:

$$D(\mathbf{r}, \mathbf{q}) := \sum_{s=1}^n r_s \log \left(\frac{r_s}{q_s} \right). \quad (20)$$

Corollary 3.4 Assume (H_1) .

(i) Let $\mathbf{r} = (r_1, \dots, r_n) \in (0, \infty)^n$ and $\mathbf{q} := (q_1, \dots, q_n) \in (0, \infty)^n$. If the base of $\log$ is greater than 1, then:

$$\sum_{s=1}^n r_s \log \left(\sum_{s=1}^n \frac{r_s}{q_s} \right) \leq A_{m,m}^{[5]} \leq A_{m,m-1}^{[5]} \leq \dots \leq A_{m,2}^{[5]} \leq A_{m,1}^{[5]} = \sum_{s=1}^n r_s \log \left(\frac{r_s}{q_s} \right) = D(\mathbf{r}, \mathbf{q}), \quad (21)$$

where

$$A_{m,l}^{[5]} = \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}} \right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right).$$

If the base of $\log$ is between 0 and 1, then inequality in (21) is reversed.

(ii) If $\mathbf{r}$ and $\mathbf{q}$ are positive probability distributions, and the base of $\log$ is greater than 1, then we have:

$$D(\mathbf{r}, \mathbf{q}) = A_{m,1}^{[6]} \geq A_{m,2}^{[6]} \geq \dots \geq A_{m,m-1}^{[6]} \geq A_{m,m}^{[6]} \geq 0, \quad (22)$$

where

$$A_{m,l}^{[6]} = \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}} \right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m, i_j}}} \right).$$

If the base of $\log$ is between 0 and 1, then inequality signs in (22) are reversed.

Proof (i) On taking $f := \log$ in Theorem 2.3 (ii), we get (21).

(ii) It is a special case of (i). $\square$

4 Inequalities for Rényi divergence and entropy

The Rényi divergence and entropy come from [23].

Definition 4.1 Let $\mathbf{r} := (r_1, \dots, r_n)$ and $\mathbf{q} := (q_1, \dots, q_n)$ be positive probability distributions, and let $\lambda \geq 0$, $\lambda \neq 1$.

(a) The Rényi divergence of order λ is defined by:

$$D_\lambda(\mathbf{r}, \mathbf{q}) := \frac{1}{\lambda - 1} \log \left(\sum_{i=1}^n q_i \left(\frac{r_i}{q_i} \right)^\lambda \right). \quad (23)$$

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

$$H_\lambda(\mathbf{r}) := \frac{1}{1 - \lambda} \log \left(\sum_{i=1}^n r_i^\lambda \right). \quad (24)$$

The Rényi divergence and the Rényi entropy can also be extended to non-negative probability distributions. If $\lambda \rightarrow 1$ in (23), we have the Kullback–Leibler divergence, and if $\lambda \rightarrow 1$ in (24), then we have the Shannon entropy. In the next two results, inequalities can be found for the Rényi divergence.

Theorem 4.2 Assume (H_1) , let $\mathbf{r} = (r_1, \dots, r_n)$ be $\mathbf{q} = (q_1, \dots, q_n)$ be probability distributions.

(i) If $0 \leq \lambda \leq \mu$, such that $\lambda, \mu \neq 1$, and the base of log is greater than 1, then:

$$D_\lambda(\mathbf{r}, \mathbf{q}) \leq A_{m,m}^{[7]} \leq A_{m,m-1}^{[7]} \leq \dots \leq A_{m,2}^{[7]} \leq A_{m,1}^{[7]} = D_\mu(\mathbf{r}, \mathbf{q}), \quad (25)$$

where

$$A_{m,l}^{[7]} = \frac{1}{\mu - 1} \log \left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \right) \right. \\ \left. \times \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \left(\frac{r_{i_j}}{q_{i_j}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \right).$$

The reverse inequalities hold in (25) if the base of log is between 0 and 1.

(ii) If $1 < \mu$ and the base of log is greater than 1, then:

$$D_1(\mathbf{r}, \mathbf{q}) = D(\mathbf{r}, \mathbf{q}) = \sum_{s=1}^n r_s \log \left(\frac{r_s}{q_s} \right) \leq A_{m,m}^{[8]} \leq A_{m,m-1}^{[8]} \leq \dots \leq A_{m,2}^{[8]} \leq A_{m,1}^{[8]} = D_\mu(\mathbf{r}, \mathbf{q}), \quad (26)$$

where

$$A_{m,l}^{[8]} = \leq \frac{1}{\mu - 1} \log \left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \right) \exp \right. \\ \left. \times \left(\frac{(\mu - 1) \sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \log \left(\frac{r_{i_j}}{q_{i_j}} \right)}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}} \right) \right);$$

here the base of exp is same as the base of log, and the reverse inequalities hold if the base of log is between 0 and 1.

(iii) If $0 \leq \lambda < 1$, and the base of log is greater than 1, then:

$$D_\lambda(\mathbf{r}, \mathbf{q}) \leq A_{m,m}^{[9]} \leq A_{m,m-1}^{[9]} \leq \dots \leq A_{m,2}^{[9]} \leq A_{m,1}^{[9]} = D_1(\mathbf{r}, \mathbf{q}), \quad (27)$$

where

$$A_{m,l}^{[9]} = \frac{1}{\lambda - 1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \left(\frac{r_{i_j}}{q_{i_j}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}} \right). \quad (28)$$

Proof By applying Theorem 1.1 with $I = (0, \infty)$, $f : (0, \infty) \rightarrow \mathbb{R}$, $f(t) := t^{\frac{\mu-1}{\lambda-1}}$:

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

we have:

$$\begin{aligned} \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s}\right)^{\lambda}\right)^{\frac{\mu-1}{\lambda-1}} &= \left(\sum_{s=1}^n r_s \left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}} \\ &\leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}\right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}\right)^{\frac{\mu-1}{\lambda-1}} \\ &\leq \dots \leq \sum_{s=1}^n r_s \left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}}, \end{aligned} \quad (29)$$

if either $0 \leq \lambda < 1 < \beta$ or $1 < \lambda \leq \mu$, and the reverse inequality in (29) holds if $0 \leq \lambda \leq \beta < 1$. By raising to power $\frac{1}{\mu-1}$, we have from all:

$$\begin{aligned} \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s}\right)^{\lambda}\right)^{\frac{1}{\lambda-1}} &\leq \dots \leq \left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}\right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}}\right)^{\frac{\mu-1}{\lambda-1}}\right)^{\frac{1}{\mu-1}} \\ &\leq \dots \leq \left(\sum_{s=1}^n r_s \left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}}\right)^{\frac{1}{\mu-1}} = \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s}\right)^{\mu}\right)^{\frac{1}{\mu-1}}. \end{aligned} \quad (30)$$

Since log is increasing if the base of log is greater than 1, it now follows (25). If the base of log is between 0 and 1, then log is decreasing and, therefore, inequality in (25) is reversed. If $\lambda = 1$ and $\beta = 1$, we have (ii) and (iii), respectively, by taking limit. $\square$

Theorem 4.3 Assume (H_1) ; let $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{q} = (q_1, \dots, q_n)$ be probability distributions. If either $0 \leq \lambda < 1$ and the base of log is greater than 1, or $1 < \lambda$ and the base of log is between 0 and 1, then:

$$\begin{aligned} \frac{1}{\sum_{s=1}^n q_s \left(\frac{r_s}{q_s}\right)^{\lambda}} \sum_{s=1}^n q_s \left(\frac{r_s}{q_s}\right)^{\lambda} \log \left(\frac{r_s}{q_s}\right) &= A_{m,1}^{[10]} \leq A_{m,2}^{[10]} \leq \dots \leq A_{m,m-1}^{[10]} \leq A_{m,m}^{[10]} \leq D_{\lambda}(\mathbf{r}, \mathbf{q}) \leq A_{m,m}^{[11]} \\ &\leq A_{m,m}^{[11]} \leq \dots \leq A_{m,2}^{[11]} \leq A_{m,1}^{[11]} = D_1(\mathbf{r}, \mathbf{q}), \end{aligned} \quad (31)$$

where

$$A_{m,m}^{[10]} = \frac{1}{(\lambda-1) \sum_{s=1}^n q_s \left(\frac{r_s}{q_s}\right)^{\lambda}} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l)$$

$$\times \left(\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}}} \right)$$

and

$$A_{m,m}^{[11]} = \frac{1}{\lambda-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}}} \right).$$

The inequalities in (31) are reversed if either $0 \leq \lambda < 1$ and the base of log is between 0 and 1, or $1 < \lambda$ and the base of log is greater than 1.

Proof We prove only the case when $0 \leq \lambda < 1$ and the base of log is greater than 1 and the other cases can be proved similarly. Since $\frac{1}{\lambda-1} < 0$ and the function log is concave and then choose $I = (0, \infty)$, $f := \log$,

$p_s = r_s$, $x_s := \left(\frac{r_s}{q_s} \right)^{\lambda-1}$ in Theorem 1.1, we have:

$$\begin{aligned} D_\lambda(\mathbf{r}, \mathbf{q}) &= \frac{1}{\lambda-1} \log \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right) = \frac{1}{\lambda-1} \log \left(\sum_{s=1}^n r_s \left(\frac{r_s}{q_s} \right)^{\lambda-1} \right) \\ &\leq \dots \leq \frac{1}{\lambda-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}}} \right) \\ &\leq \dots \leq \frac{1}{\lambda-1} \sum_{s=1}^n r_s \log \left(\left(\frac{r_s}{q_s} \right)^{\lambda-1} \right) = \sum_{s=1}^n r_s \log \left(\frac{r_s}{q_s} \right) = D_1(\mathbf{r}, \mathbf{q}), \end{aligned} \quad (32)$$

and this gives the upper bound for $D_\lambda(\mathbf{r}, \mathbf{q})$.

Since the base of log is greater than 1, the function $x \mapsto xf(x)$ ($x > 0$) is convex; therefore, $\frac{1}{1-\lambda} < 0$, and Theorem 1.1 gives:

$$\begin{aligned} D_\lambda(\mathbf{r}, \mathbf{q}) &= \frac{1}{\lambda-1} \log \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right) = \frac{1}{\lambda-1} \frac{1}{\left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right)} \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right) \log \left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right) \\ &\geq \dots \geq \frac{1}{\lambda-1} \frac{(m-1)!}{\left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right) (l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \right) \\ &\quad \times \left(\frac{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}}} \right) \\ &= \frac{1}{\lambda-1} \frac{(m-1)!}{\left(\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^\lambda \right) (l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \\ &\quad \times \left(\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}} \left(\frac{r_{ij}}{q_{ij}} \right)^{\lambda-1}}{\sum_{j=1}^l \frac{r_{ij}}{\alpha_{I_m, i_j}}} \right) \end{aligned}$$

$$\begin{aligned} &\geq \dots \geq \\ &\quad \times \frac{1}{\lambda - 1} \sum_{s=1}^n r_s \left(\frac{r_s}{q_s} \right)^{\lambda-1} \log \left(\frac{r_s}{q_s} \right)^{\lambda-1} \frac{1}{\sum_{s=1}^n r_s \left(\frac{r_s}{q_s} \right)^{\lambda-1}} \\ &= \frac{1}{\sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^{\lambda}} \sum_{s=1}^n q_s \left(\frac{r_s}{q_s} \right)^{\lambda} \log \left(\frac{r_s}{q_s} \right), \end{aligned} \quad (33)$$

which give the lower bound of $D_{\lambda}(\mathbf{r}, \mathbf{q})$. $\square$

Using the previous results, some inequalities of Rényi entropy are obtained. Let $\frac{1}{\mathbf{n}} = (\frac{1}{n}, \dots, \frac{1}{n})$ be a discrete probability distribution.

Corollary 4.4 Assume (H_1) ; let $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{q} = (q_1, \dots, q_n)$ be positive probability distributions.

(i) If $0 \leq \lambda \leq \mu$, $\lambda, \mu \neq 1$, and the base of log is greater than 1, then:

$$H_{\lambda}(\mathbf{r}) = \log(n) - D_{\lambda} \left(\mathbf{r}, \frac{1}{\mathbf{n}} \right) \geq A_{m,m}^{[12]} \geq A_{m,m}^{[12]} \geq \dots \geq A_{m,2}^{[12]} \geq A_{m,1}^{[12]} = H_{\mu}(\mathbf{r}), \quad (34)$$

where

$$\begin{aligned} A_{m,l}^{[12]} &= \frac{1}{1 - \mu} \log \left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{l,m,l}(i_1, \dots, i_l) \right. \\ &\quad \times \left. \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}} \right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}^{\lambda}}{\alpha_{l,m,i_j}}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \right). \end{aligned}$$

The reverse inequalities hold in (34) if the base of log is between 0 and 1.

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

$$S = - \sum_{s=1}^n p_i \log(p_i) \geq A_{m,m}^{[13]} \geq A_{m,m-1}^{[13]} \geq \dots \geq A_{m,2}^{[13]} \geq A_{m,1}^{[13]} = H_{\mu}(\mathbf{r}), \quad (35)$$

where

$$\begin{aligned} A_{m,l}^{[13]} &= \log(n) + \frac{1}{1 - \mu} \log \left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{l,m,l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}} \right) \exp \right. \\ &\quad \times \left. \left(\frac{(\mu-1) \sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}} \log(nr_{i_j})}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}}} \right) \right); \end{aligned}$$

the base of exp is the same as the base of log. The inequalities in (35) are reversed if the base of log is between 0 and 1.

(iii) If $0 \leq \lambda < 1$, and the base of log is greater than 1, then:

$$H_{\lambda}(\mathbf{r}) \geq A_{m,m}^{[14]} \geq A_{m,m-1}^{[14]} \geq \dots \geq A_{m,2}^{[14]} \leq A_{m,1}^{[14]} = S, \quad (36)$$

where

$$A_{m,m}^{[14]} = \frac{1}{1 - \lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{l,m,l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{i_j}^{\lambda}}{\alpha_{l,m,i_j}}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{l,m,i_j}}} \right). \quad (37)$$

The inequalities in (36) are reversed if the base of log is between 0 and 1.

Proof (i) Suppose $\mathbf{q} = \frac{1}{\mathbf{n}}$; then from (23), we have:

$$D_\lambda(\mathbf{r}, \mathbf{q}) = \frac{1}{\lambda - 1} \log \left(\sum_{s=1}^n n^{\lambda-1} r_s^\lambda \right) = \log(n) + \frac{1}{\lambda - 1} \log \left(\sum_{s=1}^n r_s^\lambda \right); \quad (38)$$

therefore, we have:

$$H_\lambda(\mathbf{r}) = \log(n) - D_\lambda(\mathbf{r}, \frac{1}{\mathbf{n}}). \quad (39)$$

Now, using Theorem 4.2 (i) and (39), we get:

$$\begin{aligned} H_\lambda(\mathbf{r}) = \log(n) - D_\lambda \left(\mathbf{r}, \frac{1}{\mathbf{n}} \right) &\geq \dots \geq \log(n) - \frac{1}{\mu - 1} \log \left(n^{\mu-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \right. \\ &\quad \times \left. \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \right) \left(\frac{\sum_{j=1}^l \frac{r_{i_j}^\lambda}{\alpha_{I_m, i_j}^\lambda}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \right) \geq \dots \geq \log(n) - D_\mu(\mathbf{r}, \mathbf{q}) = H_\mu(\mathbf{r}); \end{aligned} \quad (40)$$

(ii) and (iii) can be proved similarly. $\square$

Corollary 4.5 Assume (H_1) , and let $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{q} = (q_1, \dots, q_n)$ be positive probability distributions.

If either $0 \leq \lambda < 1$ and the base of log is greater than 1, or $1 < \lambda$ and the base of log is between 0 and 1, then:

$$\begin{aligned} -\frac{1}{\sum_{s=1}^n r_s^\lambda} \sum_{s=1}^n r_s^\lambda \log(r_s) &= A_{m,1}^{[15]} \geq A_{m,2}^{[15]} \geq \dots \geq A_{m,m-1}^{[15]} \geq A_{m,m}^{[15]} \geq H_\lambda(\mathbf{r}) \geq A_{m,m}^{[16]} \geq A_{m,m-1}^{[16]} \\ &\geq \dots \geq A_{m,2}^{[16]} \geq A_{m,1}^{[16]} = H(\mathbf{r}), \end{aligned} \quad (41)$$

where

$$\begin{aligned} A_{m,l}^{[15]} &= \frac{1}{(\lambda - 1) \sum_{s=1}^n r_s^\lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}^\lambda}{\alpha_{I_m, i_j}} \right) \log \left(n^{\lambda-1} \frac{\sum_{j=1}^l \frac{r_{i_j}^\lambda}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}} \right) \\ &\quad \times \text{and} \\ A_{m,1}^{[16]} &= \frac{1}{1 - \lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}} \right) \log \left(\frac{\sum_{j=1}^l \frac{r_{i_j}^\lambda}{\alpha_{I_m, i_j}}}{\sum_{j=1}^l \frac{r_{i_j}}{\alpha_{I_m, i_j}}} \right). \end{aligned}$$

The inequalities in (41) are reversed if either $0 \leq \lambda < 1$ and the base of log is between 0 and 1, or $1 < \lambda$ and the base of log is greater than 1.

Proof The proof is similar to Corollary 4.4 using Theorem 4.3. $\square$

5 Inequalities using Zipf–Mandelbrot law

The Zipf–Mandelbrot law is defined as follows (see [20]).

Definition 5.1 Zipf–Mandelbrot law is a discrete probability distribution depending on three parameters $N \in \{1, 2, \dots\}$, $q \in [0, \infty)$ and $t > 0$, and is defined by:

$$f(s; N, q, t) := \frac{1}{(s + q)^t H_{N,q,t}}, \quad s = 1, \dots, N, \quad (42)$$

where

$$H_{N,q,t} = \sum_{j=1}^N \frac{1}{(j + q)^t}. \quad (43)$$

If the total mass of the law is taken over all $\mathbb{N}$, then for $q \geq 0$, $t > 1$, $s \in \mathbb{N}$, density function of Zipf–Mandelbrot law becomes:

$$f(s; q, t) = \frac{1}{(s + q)^t H_{q,t}}, \quad (44)$$

where

$$H_{q,t} = \sum_{j=1}^{\infty} \frac{1}{(j + q)^t}. \quad (45)$$

For $q = 0$, the Zipf–Mandelbrot law. By Corollary 4.4 (iii), we get:

Conclusion 5.2 Assume (H_1) ; let $\mathbf{r}$ be a Zipf–Mandelbrot law, by Corollary 4.4 (iii), we get. If $0 \leq \lambda < 1$, and the base of log is greater than 1, then:

$$\begin{aligned} H_{\lambda}(\mathbf{r}) &= \frac{1}{1 - \lambda} \log \left(\frac{1}{H_{N,q,t}^{\lambda}} \sum_{s=1}^n \frac{1}{(s + q)^{\lambda s}} \right) \\ &\geq \dots \geq \frac{1}{1 - \lambda} \frac{(m - 1)!}{(l - 1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \left(\sum_{j=1}^l \frac{1}{\alpha_{I_m, i_j} (i_j + q) H_{N,q,t}} \right) \log \\ &\quad \times \left(\frac{1}{H_{N,q,t}^{\lambda-1}} \sum_{j=1}^l \frac{1}{\alpha_{I_m, i_j} (i_j - q)^{\lambda s}} \right) \\ &\geq \dots \geq \frac{t}{H_{N,q,t}} \sum_{s=1}^N \frac{\log(s + q)}{(s + q)^t} + \log(H_{N,q,t}) = S. \end{aligned} \quad (46)$$

The inequalities in (46) are reversed if the base of log is between 0 and 1.

Conclusion 5.3 Assume (H_1) ; let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the Zipf–Mandelbrot law with parameters $N \in \{1, 2, \dots\}$, $q_1, q_2 \in [0, \infty)$ and $s_1, s_2 > 0$, respectively. Then, from Corollary 3.4 (ii), we have: if the base of log is greater than 1, then:

$$\begin{aligned} \bar{D}(\mathbf{r}_1, \mathbf{r}_2) &= \sum_{s=1}^n \frac{1}{(s + q_1)^{t_1} H_{N,q_1,t_1}} \log \left(\frac{(s + q_2)^{t_2} H_{N,q_2,t_2}}{(s + q_1)^{t_1} H_{N,q_2,t_1}} \right) \geq \dots \geq \frac{(m - 1)!}{(l - 1)!} \sum_{(i_1, \dots, i_l) \in I_l} \eta_{I_m, l}(i_1, \dots, i_l) \\ &\quad \times \left(\sum_{j=1}^l \frac{1}{\frac{(i_j + q_2)^{t_2} H_{N,q_2,t_2}}{\alpha_{I_m, i_j}}} \right) \left(\sum_{j=1}^l \frac{\frac{1}{(i_j + q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m, i_j}} \right) \log \left(\frac{\sum_{j=1}^l \frac{1}{\alpha_{I_m, i_j} (i_j + q_1)^{t_1} H_{N,q_1,t_1}}}{\sum_{j=1}^l \frac{1}{\alpha_{I_m, i_j} (i_j + q_2)^{t_2} H_{N,q_2,t_2}}} \right) \geq \dots \geq 0. \end{aligned} \quad (47)$$

The inequalities in (47) are reversed if base of log is between 0 and 1.

6 Shannon entropy, Zipf–Mandelbrot law, and hybrid Zipf–Mandelbrot law

Here, we maximize the Shannon entropy using the method of Lagrange multiplier under some equations constraints and get the Zipf–Mandelbrot law.

Theorem 6.1 *If $J = \{1, 2, \dots, N\}$, for a given $q \geq 0$, a probability distribution that maximizes the Shannon entropy under the constraints:*

$$\sum_{s \in J} r_s = 1, \quad \sum_{s \in J} r_s (\ln(s + q)) := \Psi,$$

is Zipf–Mandelbrot law.

Proof If $J = \{1, 2, \dots, N\}$. We set the Lagrange multipliers λ and t and consider the expression:

$$\tilde{S} = - \sum_{s=1}^N r_s \ln r_s - \lambda \left(\sum_{s=1}^N r_s - 1 \right) - t \left(\sum_{s=1}^N r_s \ln(s + q) - \Psi \right).$$

Just for the sake of convenience, replace λ by $\ln \lambda - 1$, and thus, the last expression gives:

$$\tilde{S} = - \sum_{s=1}^N r_s \ln r_s - (\ln \lambda - 1) \left(\sum_{s=1}^N r_s - 1 \right) - t \left(\sum_{s=1}^N r_s \ln(s + q) - \Psi \right).$$

From $\tilde{S}_{r_s} = 0$, for $s = 1, 2, \dots, N$, we get:

$$r_s = \frac{1}{\lambda (s + q)^t},$$

and on using the constraint $\sum_{s=1}^N r_s = 1$, we have:

$$\lambda = \sum_{s=1}^N \left(\frac{1}{(s + 1)^t} \right),$$

where $t > 0$, concluding that:

$$r_s = \frac{1}{(s + q)^t H_{N,q,t}}, \quad s = 1, 2, \dots, N.$$

□

Remark 6.2 Observe that the Zipf–Mandelbrot law and Shannon entropy can be bounded from above (see [21]):

$$S = - \sum_{s=1}^N f(s, N, q, t) \ln f(s, N, q, t) \leq - \sum_{s=1}^N f(s, N, q, t) \ln q_s,$$

where $(q_1, \dots, q_N)$ is a positive N -tuple, such that $\sum_{s=1}^N q_s = 1$.

Theorem 6.3 *If $J = \{1, \dots, N\}$, then probability distribution that maximizes Shannon entropy under constraints*

$$\sum_{s \in J} r_s = 1, \quad \sum_{s \in J} r_s \ln(s + q) := \Psi, \quad \sum_{s \in J} s r_s := \eta$$

is hybrid Zipf–Mandelbrot law given as:

$$r_s = \frac{w^s}{(s + q)^k \Phi^*(k, q, w)}, \quad s \in J,$$

where

$$\Phi_J(k, q, w) = \sum_{s \in J} \frac{w^s}{(s + q)^k}.$$

Proof First, consider $J = \{1, \dots, N\}$; we set the Lagrange multiplier and consider the expression:

$$\tilde{S} = - \sum_{s=1}^N r_s \ln r_s + \ln w \left(\sum_{s=1}^N s r_s - \eta \right) - (\ln \lambda - 1) \left(\sum_{s=1}^N r_s - 1 \right) - k \left(\sum_{s=1}^N r_s \ln(s + q) - \Psi \right).$$

On setting $\tilde{S}_{r_s} = 0$, for $s = 1, \dots, N$, we get:

$$-\ln r_s + s \ln w - \ln \lambda - k \ln(s + q) = 0;$$

after solving for r_s , we get:

$$\lambda = \sum_{s=1}^N \frac{w^s}{(s + q)^k};$$

and we recognize this as the partial sum of Lerch's transcendent that we will denote with:

$$\Phi_N^*(k, q, w) = \sum_{s=1}^N \frac{w^s}{(s + q)^k}$$

with $w \geq 0, k > 0$. □

Remark 6.4 Observe that for Zipf–Mandelbrot law, Shannon entropy can be bounded from above (see [21]):

$$S = - \sum_{s=1}^N f_h(s, N, q, k) \ln f_h(s, N, q, k) \leq - \sum_{s=1}^N f_h(s, N, q, k) \ln q_s,$$

where $(q_1, \dots, q_N)$ is any positive N -tuple, such that $\sum_{s=1}^N q_s = 1$.

Under the assumption of Theorem 2.3 (i), define the non-negative functionals as follows:

$$\Theta_3(f) = \mathcal{A}_{m,r}^{[1]} - f \left(\frac{\sum_{s=1}^n r_s}{\sum_{s=1}^n q_s} \right) \sum_{s=1}^n q_s, \quad r = 1, \dots, m, \quad (48)$$

$$\Theta_4(f) = \mathcal{A}_{m,r}^{[1]} - \mathcal{A}_{m,k}^{[1]}, \quad 1 \leq r < k \leq m. \quad (49)$$

Under the assumption of Theorem 2.3 (ii), define the non-negative functionals as follows:

$$\Theta_5(f) = \mathcal{A}_{m,r}^{[2]} - \left(\sum_{s=1}^n r_s \right) f \left(\frac{\sum_{s=1}^n r_s}{\sum_{s=1}^n q_s} \right), \quad r = 1, \dots, m, \quad (50)$$

$$\Theta_6(f) = \mathcal{A}_{m,r}^{[2]} - \mathcal{A}_{m,k}^{[2]}, \quad 1 \leq r < k \leq m. \quad (51)$$

Under the assumption of Corollary 3.2 (i), define the following non-negative functionals:

$$\Theta_7(f) = A_{m,r}^{[3]} + \sum_{i=1}^n q_i \log(q_i), \quad r = 1, \dots, n \quad (52)$$

$$\Theta_8(f) = A_{m,r}^{[3]} - A_{m,k}^{[3]}, \quad 1 \leq r < k \leq m. \quad (53)$$

Under the assumption of Corollary 3.2 (ii), define the following non-negative functionals are given as:

$$\Theta_9(f) = A_{m,r}^{[4]} - S, \quad r = 1, \dots, m \quad (54)$$

$$\Theta_{10}(f) = A_{m,r}^{[4]} - A_{m,k}^{[4]}, \quad 1 \leq r < k \leq m. \quad (55)$$

Under the assumption of Corollary 3.4 (i), let us define the non-negative functionals as follows:

$$\Theta_{11}(f) = A_{m,r}^{[5]} - \sum_{s=1}^n r_s \log \left(\sum_{s=1}^n \log \frac{r_n}{\sum_{s=1}^n q_s} \right), \quad r = 1, \dots, m \quad (56)$$

$$\Theta_{12}(f) = A_{m,r}^{[5]} - A_{m,k}^{[5]}, \quad 1 \leq r < k \leq m. \quad (57)$$

Under the assumption of Corollary 3.4 (ii), define the non-negative functionals as follows:

$$\Theta_{13}(f) = A_{m,r}^{[6]} - A_{m,k}^{[6]}, \quad 1 \leq r < k \leq m. \quad (58)$$

Under the assumption of Theorem 4.2 (i), consider the following functionals:

$$\Theta_{14}(f) = A_{m,r}^{[7]} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \dots, m \quad (59)$$

$$\Theta_{15}(f) = A_{m,r}^{[7]} - A_{m,k}^{[7]}, \quad 1 \leq r < k \leq m. \quad (60)$$

Under the assumption of Theorem 4.2 (ii), consider the following functionals:

$$\Theta_{16}(f) = A_{m,r}^{[8]} - D_1(\mathbf{r}, \mathbf{q}), \quad r = 1, \dots, m \quad (61)$$

$$\Theta_{17}(f) = A_{m,r}^{[8]} - A_{m,k}^{[8]}, \quad 1 \leq r < k \leq m. \quad (62)$$

Under the assumption of Theorem 4.2 (iii), consider the following functionals:

$$\Theta_{18}(f) = A_{m,r}^{[9]} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \dots, m \quad (63)$$

$$\Theta_{19}(f) = A_{m,r}^{[9]} - A_{m,k}^{[9]}, \quad 1 \leq r < k \leq m. \quad (64)$$

Under the assumption of Theorem 4.3, consider the following non-negative functionals:

$$\Theta_{20}(f) = D_\lambda(\mathbf{r}, \mathbf{q}) - A_{m,r}^{[10]}, \quad r = 1, \dots, m \quad (65)$$

$$\Theta_{21}(f) = A_{m,k}^{[10]} - A_{m,r}^{[10]}, \quad 1 \leq r < k \leq m. \quad (66)$$

$$\Theta_{22}(f) = A_{m,r}^{[11]} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \dots, m \quad (67)$$

$$\Theta_{23}(f) = A_{m,r}^{[11]} - A_{m,k}^{[11]}, \quad 1 \leq r < k \leq m. \quad (68)$$

$$\Theta_{24}(f) = A_{m,r}^{[11]} - A_{m,k}^{[10]}, \quad r = 1, \dots, m, \quad k = 1, \dots, m. \quad (69)$$

Under the assumption of Corollary 4.4 (i), consider the following non-negative functionals:

$$\Theta_{25}(f) = H_\lambda(\mathbf{r}) - A_{m,r}^{[12]}, \quad r = 1, \dots, m \quad (70)$$

$$\Theta_{26}(f) = A_{m,k}^{[12]} - A_{m,r}^{[12]}, \quad 1 \leq r < k \leq m. \quad (71)$$

Under the assumption of Corollary 4.4 (ii), consider the following functionals:

$$\Theta_{27}(f) = S - A_{m,r}^{[13]}, \quad r = 1, \dots, m \quad (72)$$

$$\Theta_{28}(f) = A_{m,k}^{[13]} - A_{m,r}^{[13]}, \quad 1 \leq r < k \leq m. \quad (73)$$

Under the assumption of Corollary 4.4 (iii), consider the following functionals:

$$\Theta_{29}(f) = H_\lambda(\mathbf{r}) - A_{m,r}^{[14]}, \quad r = 1, \dots, m \quad (74)$$

$$\Theta_{30}(f) = A_{m,k}^{[14]} - A_{m,r}^{[14]}, \quad 1 \leq r < k \leq m. \quad (75)$$

Under the assumption of Corollary 4.5, defined the following functionals.

$$\Theta_{31} = A_{m,r}^{[15]} - H_\lambda(\mathbf{r}), \quad r = 1, \dots, m \quad (76)$$

$$\Theta_{32} = A_{m,r}^{[15]} - A_{m,k}^{[15]}, \quad 1 \leq r < k \leq m. \quad (77)$$

$$\Theta_{33} = H_\lambda(\mathbf{r}) - A_{m,r}^{[16]}, \quad r = 1, \dots, m \quad (78)$$

$$\Theta_{34} = A_{m,k}^{[16]} - A_{m,r}^{[16]}, \quad 1 \leq r < k \leq m. \quad (79)$$

$$\Theta_{35} = A_{m,r}^{[15]} - A_{m,k}^{[16]}, \quad r = 1, \dots, m, \quad k = 1, \dots, m. \quad (80)$$

7 Generalization of refinement of Jensen-, Rényi-, and Shannon-type inequalities via Lidstone polynomial

We construct some new identities with the help of generalized Lidstone polynomial (6).

Theorem 7.1 Assume (H_1) ; let $f : [\alpha_1, \alpha_2] \rightarrow \mathbb{R}$ be a function, where $[\alpha_1, \alpha_2] \subset \mathbb{R}$, such that be an interval, such that $f \in C^{2m}[\alpha_1, \alpha_2]$ for $m \geq 1$. Also let $x_1, \dots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \dots, p_n$ be positive real numbers, such that $\sum_{i=1}^n p_i = 1$, and $\mathfrak{F}_m(t)$ are the same as defined in (5), and then:

$$\begin{aligned} \Theta_i(f) &= \sum_{k=1}^{m-1} (\alpha_2 - \alpha_1)^{2k} f^{(2k)}(\alpha_1) \Theta_i \left(\mathfrak{F}_l \left(\frac{\alpha_2 - x}{\alpha_2 - \alpha_1} \right) \right) \\ &\quad + \sum_{k=1}^{m-1} (\alpha_2 - \alpha_1)^{2k} f^{(2k)}(\alpha_2) \Theta_i \left(\mathfrak{F}_l \left(\frac{x - \alpha_1}{\alpha_2 - \alpha_1} \right) \right) \\ &\quad + (\alpha_2 - \alpha_1)^{2k-1} \int_{\alpha_1}^{\alpha_2} \Theta_i \left(G_m \left(\frac{x - \alpha_1}{\alpha_2 - \alpha_1}, \frac{t - \alpha_1}{\alpha_2 - \alpha_1} \right) \right) f^{(2m)}(t) dt, \quad i = 1, 2, \dots, 35. \end{aligned} \quad (81)$$

Proof Using (6) in place of f in $\Theta_i(f)$, $i = 1, 2, \dots, 35$, we get (81). $\square$

Theorem 7.2 Assume (H_1) ; let $f : [\alpha_1, \alpha_2] \rightarrow \mathbb{R}$ be a function, where $[\alpha_1, \alpha_2] \subset \mathbb{R}$, such that be an interval, such that $f \in C^{2m}[\alpha_1, \alpha_2]$ for $m \geq 1$. Also let $x_1, \dots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \dots, p_n$ be positive real numbers, such that $\sum_{i=1}^n p_i = 1$, and $\mathfrak{F}_m(t)$ are the same as defined in (5); let for $m \geq 1$:

$$\Theta_i \left(G_m \left(\frac{x - \alpha_1}{\alpha_2 - \alpha_1}, \frac{t - \alpha_1}{\alpha_2 - \alpha_1} \right) \right) \geq 0, \quad \text{for all } t \in [\alpha_1, \alpha_2]. \quad (82)$$

If f is $2m$ -convex function, then we have:

$$\begin{aligned} \Theta_i(f) &\geq \sum_{k=1}^{m-1} (\alpha_2 - \alpha_1)^{2k} f^{(2k)}(\alpha_1) \Theta_i \left(\mathfrak{F}_l \left(\frac{\alpha_2 - x}{\alpha_2 - \alpha_1} \right) \right) \\ &\quad + \sum_{k=1}^{m-1} (\alpha_2 - \alpha_1)^{2k} f^{(2k)}(\alpha_2) \Theta_i \left(\mathfrak{F}_l \left(\frac{x - \alpha_1}{\alpha_2 - \alpha_1} \right) \right), \quad i = 1, 2, \dots, 35. \end{aligned} \quad (83)$$

Proof Since f is $2m$ -convex; therefore, $f^{(2m)} \geq 0$ for all $x \in [\alpha_1, \alpha_2]$, then using (82) in (81), we get the required result. $\square$

Theorem 7.3 Assume (H_1) ; let $f : [\alpha_1, \alpha_2] \rightarrow \mathbb{R}$ be a function, where $[\alpha_1, \alpha_2] \subset \mathbb{R}$ be an interval. Also let $x_1, \dots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \dots, p_n$ be positive real numbers, such that $\sum_{i=1}^n p_i = 1$, and also suppose that $f : [\alpha_1, \alpha_2] \rightarrow \mathbb{R}$ is $2m$ -convex. Then, the following results are valid.

- (i) If m is odd integer, then for every $2m$ -convex function, (83) holds. (ii) Suppose that (83) holds. If the function

$$\lambda(u) = \sum_{l=0}^{m-1} (\alpha_2 - \alpha_1)^{2l} g^{(2l)}(\alpha_1) \mathfrak{F}_l \left(\frac{\alpha_2 - u}{\alpha_2 - \alpha_1} \right) + \sum_{l=0}^{m-1} (\alpha_2 - \alpha_1)^{2l} g^{(2l)}(\alpha_2) \mathfrak{F}_l \left(\frac{u - \alpha_1}{\alpha_2 - \alpha_1} \right)$$

is convex, then the right-hand side of (83) is non-negative and we have:

$$\Theta_i(f) \geq 0, \quad i = 1, 2, \dots, 35. \quad (84)$$

Proof (i) Note that $G_1(u, s) \leq 0$ for $1 \leq u, s \leq 1$ and also note that $G_m(u, s) \leq 0$ for odd integer m and $G_m(u, s) \geq 0$ for even integer m . As G_1 is convex function and G_{m-1} is positive for odd integer m , therefore:

$$\frac{d^2}{d^2u} (G_m(u, s)) = \int_0^1 \frac{d^2}{d^2u} G_1(u, p) G_{m-1}(p, s) dp \geq 0, \quad m \geq 2.$$

This shows that G_m is convex in the first variable u if m is convex. Similarly, G_m is concave in the first variable if m is even. Hence, if m is odd, then:

$$\Theta_i \left(G_m \left(\frac{x - \alpha_1}{\alpha_2 - \alpha_1}, \frac{t - \alpha_1}{\alpha_2 - \alpha_1} \right) \right) \geq 0;$$

therefore, (84) is valid.

(ii) Using the linearity of $\Theta_i(f)$, we can write the right-hand side of (83) in the form $\Theta_i(\lambda)$. As λ is supposed to be convex, therefore, the right-hand side of (83) is non-negative, and so $\Theta_i(f) \geq 0$. $\square$

Remark A We can investigate the bounds for the identities related to the generalization of refinement of Jensen inequality using inequalities for the Čebyšev functional and some results relating to the Grüss and Ostrowski-type inequalities can be constructed as given in Section 3 of [5]. Also we can construct the non-negative functionals from inequality (83) and give related mean value theorems and we can construct the new families of m -exponentially convex functions and Cauchy means related to these functionals as given in Section 4 of [5].

Acknowledgements The authors wish to thank the anonymous referees for their very careful reading of the manuscript and fruitful comments and suggestions.

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

Compliance with ethical standards

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

Conflict of interest The authors declares that there is no conflict of interests regarding the publication of this paper.

References

1. Anderson, G.; Ge, Y.: The size distribution of Chinese cities. *Reg. Sci. Urban Econ.* **35**(6), 756–776 (2005)
2. Auerbach, F.: Das Gesetz der Bevölkerungskonzentration. *Petermanns Geogr. Mitt.* **59**, 74–76 (1913)
3. Black, D.; Henderson, V.: Urban evolution in the USA. *J. Econ. Geogr.* **3**(4), 343–372 (2003)
4. Bosker, M.; Brakman, S.; Garretsen, H.; Schramm, M.: A century of shocks: the evolution of the German city size distribution 1925–1999. *Reg. Sci. Urban Econ.* **38**(4), 330–347 (2008)
5. Butt, S.I.; Khan, K.A.; Pečarić, J.: Generalization of Popoviciu inequality for higher order convex function via Taylor's polynomial. *Acta Univ. Apulensis Math. Inform.* **42**, 181–200 (2015)
6. Butt, S.I.; Mehmood, N.; Pečarić, J.: New generalizations of Popoviciu type inequalities via new green functions and Fink's identity. *Trans. A Razmadze Math. Inst.* **171**(3), 293–303 (2017)
7. Butt, S.I.; Pečarić, J.: *Popoviciu's Inequality For N-Convex Functions*. Lap Lambert Academic Publishing, Saarbrücken (2016)
8. Butt, S.I.; Pečarić, J.: Weighted Popoviciu type inequalities via generalized Montgomery identities. *Hrvatske akademije znanosti i umjetnosti: Matematičke znanosti* **69–89** (2015)
9. Butt, S.I.; Khan, K.A.; Pečarić, J.: Popoviciu type inequalities via Hermite's polynomial. *Math. Inequal. Appl.* **19**(4), 1309–1318 (2016)
10. 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 (1978)
11. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. *Stud. Sci. Math. Hung.* **2**, 299–318 (1967)
12. Horváth, L.: A method to refine the discrete Jensen's inequality for convex and mid-convex functions. *Math. Comput. Model.* **54**(9–10), 2451–2459 (2011)
13. Horváth, L.; Khan, K.A.; Pečarić, J.: *J. Combinatorial Improvements of Jensens Inequality/Classical and New Refinements of Jensens Inequality with Applications*, Monographs in Inequalities 8, Element, Zagreb (2014)
14. Horváth, L.; Khan, K.A.; Pečarić, J.: Refinement of Jensen's inequality for operator convex functions. *Adv. Inequal. Appl.* (2014)
15. Horváth, L.; Pečarić, J.: A refinement of discrete Jensen's inequality. *Math. Inequal. Appl.* **14**, 777–791 (2011)

16. Horváth, L.; Pečarić, Đ.; Pečarić, J.: Estimations of f-and Rényi divergences by using a cyclic refinement of the Jensens inequality. *Bull. Malays. Math. Sci. Soc.* **1–14** (2017)
17. Ioannides, Y.M.; Overman, H.G.: Zipf's law for cities: an empirical examination. *Reg. Sci. Urban Econ.* **33**(2), 127–137 (2003)
18. Kullback, S.: *Information Theory and Statistics*. Courier Corporation, Chelmsford (1997)
19. Kullback, S.; Leibler, R.A.: On information and sufficiency. *Ann. Math. Stat.* **22**(1), 79–86 (1951)
20. Lovričević, N.; Pečarić, Đ.; Pečarić, J.: Zipf–Mandelbrot law, f-divergences and the Jensen-type interpolating inequalities. *J. Inequal. Appl.* **2018**(1), 36 (2018)
21. Matic, M.; Pearce, C.E.; Pečarić, J.: Shannon's and related inequalities in information theory. In: *Survey on Classical Inequalities*, pp. 127–164. Springer, Dordrecht (2000)
22. Pečarić, J.; Proschan, F.; Tong, Y.L.: *Convex Functions, Partial Orderings and Statistical Applications*. Academic Press, New York (1992)
23. Rényi, A.: On measure of information and entropy. In: *Proceeding of the Fourth Berkely Symposium on Mathematics, Statistics and Probability*, pp. 547–561 (1960)
24. Rosen, K.T.; Resnick, M.: The size distribution of cities: an examination of the Pareto law and primacy. *J. Urban Econ.* **8**(2), 165–186 (1980)
25. Soo, K.T.: Zipf's Law for cities: a cross-country investigation. *Reg. Sci. Urban Econ.* **35**(3), 239–263 (2005)
26. Widder, D.V.: Completely convex function and Lidstone series. *Trans. Am. Math. Soc.* **51**(1942), 387–398 (1942)
27. Zipf, G.K.: *Human behaviour and the principle of least-effort*. In: Cambridge MA edn. Addison-Wesley, Reading (1949)

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