CrossMark

Refinement of Jensen's inequality and estimation of f- and Rényi divergence via Montgomery identity

Khuram Ali Khan[1], Tasadduq Niaz[1,2]*, Đilda Pečarić[3] and Josip Pečarić[4]

*Correspondence:
tasadduq_khan@yahoo.com
[1]Department of Mathematics,
University of Sargodha, Sargodha,
Pakistan
[2]Department of Mathematics, The
University of Lahore, Sargodha,
Pakistan
Full list of author information is
available at the end of the article

Abstract

Jensen's inequality is important for obtaining inequalities for divergence between probability distribution. By applying a refinement of Jensen's inequality (Horváth et al. in Math. Inequal. Appl. 14:777–791, 2011) and introducing a new functional based on an f-divergence functional, we obtain some estimates for the new functionals, the f-divergence, and Rényi divergence. Some inequalities for Rényi and Shannon estimates are constructed. The Zipf–Mandelbrot law is used to illustrate the result. In addition, we generalize the refinement of Jensen's inequality and new inequalities of Rényi Shannon entropies for an m-convex function using the Montgomery identity. It is also given that the maximization of Shannon entropy is a transition from the Zipf–Mandelbrot law to a hybrid Zipf–Mandelbrot law.

Keywords: m-convex function; Jensen's inequality; f- and Rényi divergence; Montgomery identity; Entropy

1 Introduction and preliminary results

The most commonly used words, the largest cities of countries, income of a billionaire can be described in terms of Zipf's law. The f-divergence means the distance between two probability distributions by making an average value, which is weighted by a specified function. As f-divergence, there are other probability distributions like Csiszar f-divergence [12, 13], some special case of which are Kullback–Leibler-divergence used to find the appropriate distance between the probability distributions (see [20, 21]). The notion of distance is stronger than that of divergence because it gives the properties of symmetry and triangle inequalities. Probability theory has applications in many fields, and the divergence between probability distributions has many applications in these fields.

Many natural phenomena, like distribution of wealth and income in a society, distribution of Facebook likes, distribution of football goals, follow the power law distribution (Zipf's law). Like above phenomena, the distribution of city sizes also follows the power law distribution. In [4] Auerbach was the first who gave the idea that the distribution of city sizes can be well approximated by using the Pareto distribution (power law distribution). Many researchers refined this idea. However, Zipf [29] has done notable work in this field. Rosen and Resnick [27], Black and Henderson [5], Ioannides and Overman [19], Soo [28], Anderson and Ge [3], and Bosker et al. [6] investigated the distribution of city

sizes of the urban economics. It means that the product of the rank appears and city sizes are roughly constant. This shows that population of the nth city is $\frac{1}{n}$ of the largest city population. This rule is named rank, size rule and is also called Zipf's law. Hence Zipf's law does not only show that the city size distribution follows the Pareto distribution.

By using an f-divergence functional, Horváth et al. in [17] introduced a new functional and obtained some estimates for that functional, the Rényi divergence, and f-divergence applying a cyclic refinement of Jensen's inequality. Also they obtained some new inequalities for Shannon and Rényi entropies; also they used the Zipf–Mandelbrot law to illustrate some 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 [25, p. 15]). It is quite an interesting fact that there are some results that are true for convex functions, but when we discuss them in higher order convexity, they do not remain valid.

In [25, p. 16], the following criterion is 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 example, Butt et al. generalized the Popoviciu inequality for an m-convex function using Taylor's formula, Lidstone polynomial, Montgomery identity, Fink's identity, Abel–Gonstcharoff interpolation, and Hermite interpolating polynomial (see [7–11]).

Since many years Jensen's inequality has received great interest. The researchers have given the refinement of Jensen's inequality by defining some new functions (see [16, 18]). Like many researchers Horváth and Pečarić in [14, 18] (see also [15, p. 26]) gave a refinement of Jensen's inequality for convex functions. They defined some essential notions to prove the refinement given as follows:

Suppose X to be a set, $P(X)$ denotes the power set of X, $|X|$ denotes the number of elements of X, and $\mathbb{N}$ denotes the nonnegative integers.

Consider $q \geq 1$ and $r \geq 2$ to be fixed integers. Define the functions

$$F_{r,s} : \{1,\ldots,q\}^r \to \{1,\ldots,q\}^{r-1}, \quad 1 \leq s \leq r,$$
$$F_r : \{1,\ldots,q\}^r \to P\big(\{1,\ldots,q\}^{r-1}\big),$$

and

$$T_r : P\big(\{1,\ldots,q\}^r\big) \to P\big(\{1,\ldots,q\}^{r-1}\big),$$

by

$$F_{r,s}(i_1,\ldots,i_r) := (i_1, i_2,\ldots,i_{s-1},i_{s+1},\ldots,i_r), \quad 1 \leq s \leq r,$$
$$F_r(i_1,\ldots,i_r) := \bigcup_{s=1}^{r} \big\{F_{r,s}(i_1,\ldots,i_r)\big\},$$

and

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

Next let the function

$$\alpha_{r,i} : \{1,\ldots,q\}^r \to \mathsf{N}, \quad 1 \le i \le q,$$

be defined by

$\alpha_{r,i}(i_1,\ldots,i_r)$ is the number of occurences of in the sequence $(i_1,\ldots,i_r)$.

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

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

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

$$\alpha_{I_m,i} \ge 1, \quad 1 \le i \le n.$$

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

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

Obviously, the sets $I_1 = \{1,\ldots,n\}$ by (H_1) and this ensures that $\alpha_{I_1,i} = 1$ ($1 \le i \le n$). From (H_1) we have $\alpha_{I_l,i} \ge 1$ ($m-1 \ge l \ge 1$, $1 \le i \le n$).

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

$$\mathsf{H}_{I_l}(j_1,\ldots,j_{l-1}) := \big\{\big((i_1,\ldots,i_l),k\big) \times \{1,\ldots,l\} | F_{l,k}(i_1,\ldots,i_l) = (j_1,\ldots,j_{l-1})\big\}.$$

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

$$\eta_{I_m,m}(i_1,\ldots,i_m) := 1, \quad (i_1,\ldots,i_m) \in I_m;$$

$$\eta_{I_m,l-1}(j_1,\ldots,j_{l-1}) := \sum_{((i_1,\ldots,i_l),k)\in\mathsf{H}_{I_l}(j_1,\ldots,j_{l-1})} \eta_{I_m,l}(i_1,\ldots,i_l).$$

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

$$A_{m,l} = A_{m,l}(I_m, x_1,\ldots,x_n, p_1,\ldots,p_n; f) := \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l),$$

$$\left(\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}\right) f\left(\frac{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}} x_{i_j}}{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}}\right)$$

and prove the following theorem.

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

$$f\left(\sum_{s=1}^{n} p_s x_s\right) \leq A_{m,m} \leq A_{m,m-1} \leq \cdots \leq A_{m,2} \leq A_{m,1} = \sum_{s=1}^{n} p_s f(x_s). \tag{1}$$

By using the differences of inequalities in (1), we define some new functionals as follows:

$$\Theta_1(f) = A_{m,r} - f\left(\sum_{s=1}^{n} p_s x_s\right), \quad r = 1, \ldots, m, \tag{2}$$

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

Under the assumptions of Theorem 1.1, we have

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

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

The Montgomery identity via Taylor's formula is given in [1] and [2].

Theorem 1 *Let* $m \in \mathbb{N}$, $f : I \to \mathbb{R}$ *be such that* $f^{(m-1)}$ *is absolutely continuous,* $I \subset \mathbb{R}$ *be an open interval,* $\alpha_1, \alpha_2 \in I$, $\alpha_1 < \alpha_2$. *Then the following identity holds*:

$$\psi(x) = \frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} \psi(u)\, du + \sum_{k=0}^{m-2} \frac{\psi^{(k+1)}(\alpha_1)(x - \alpha_1)^{k+2}}{k!(k+2)(\alpha_2 - \alpha_1)} - \sum_{k=0}^{m-2} \frac{\psi^{(k+1)}(\alpha_2)(x - \alpha_2)^{k+2}}{k!(k+2)(\alpha_2 - \alpha_1)}$$
$$+ \frac{1}{(m-1)!} \int_{\alpha_1}^{\alpha_2} R_m(x, u) \psi^{(m)}(u)\, du, \tag{5}$$

where

$$R_m(x, u) = \begin{cases} -\frac{(x-u)^m}{m(\alpha_2 - \alpha_1)} + \frac{x - \alpha_1}{\alpha_2 - \alpha_1}(x - u)^{m-1}, & \alpha_1 \leq u \leq x; \\ -\frac{(x-u)^m}{m(\alpha_2 - \alpha_1)} + \frac{x - \alpha_2}{\alpha_2 - \alpha_1}(x - u)^{m-1}, & x \leq u \leq \alpha_2. \end{cases} \tag{6}$$

Theorem 2 *Let* $m \in \mathbb{N}$, $f : I \to \mathbb{R}$ *be such that* $f^{(m-1)}$ *is absolutely continuous,* $I \subset \mathbb{R}$ *be an interval,* $\alpha_1, \alpha_2 \in I$, $\alpha_1 < \alpha_2$. *Then the following identity holds*:

$$\psi(x) = \frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} \psi(u)\, du + \sum_{k=0}^{m-2} \psi^{(k+1)}(x) \frac{(\alpha_1 - x)^{k+2} - (\alpha_2 - x)^{k+2}}{(k+2)!(\alpha_2 - \alpha_1)}$$
$$+ \frac{1}{(m-1)!} \int_{\alpha_1}^{\alpha_2} \hat{R}(x, u) \psi^{(m)}(u)\, du, \tag{7}$$

where

$$\hat{R}(x, u) = \begin{cases} -\frac{1}{m(\alpha_2 - \alpha_1)}(\alpha_1 - u), & \alpha_1 \leq u \leq x; \\ -\frac{1}{m(\alpha_2 - \alpha_1)}(\alpha_2 - u), & x \leq u \leq \alpha_2. \end{cases} \tag{8}$$

In case $m = 1$, the sum $\sum_{k=0}^{m-2} \ldots$ is empty, so (5) and (7) reduce to the well-known Montgomery identity (see [24])

$$f(x) = \frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} f(t)\, dt + \frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} p(x, u) f'(u)\, du,$$

where $p(x, u)$ is the Peano kernel defined by

$$p(x, u) = \begin{cases} \frac{u - \alpha_1}{\alpha_2 - \alpha_1}, & \alpha_1 \le u \le x; \\ \frac{u - \alpha_2}{\alpha_2 - \alpha_1}, & x \le u \le \alpha_2. \end{cases}$$

2 Inequalities for Csiszár divergence

In [12, 13] Csiszár introduced the following notion.

Definition 1 Let $f : \mathsf{R}^+ \to \mathsf{R}^+$ be a convex function, let $\mathbf{r} = (r_1, \ldots, r_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ be positive probability distributions. Then the f-divergence functional is defined by

$$I_f(\mathbf{r}, \mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{r_i}{q_i}\right). \tag{9}$$

And he stated that by defining

$$f(0) := \lim_{x \to 0^+} f(x); \qquad 0f\left(\frac{0}{0}\right) := 0; \qquad 0f\left(\frac{a}{0}\right) := \lim_{x \to 0^+} x f\left(\frac{a}{0}\right), \quad a > 0, \tag{10}$$

we can also use the nonnegative probability distributions.

In [17], Horváth et al. gave the following functional on the basis of previous definition.

Definition 2 Let $I \subset \mathsf{R}$ be an interval, and let $f : I \to \mathsf{R}$ be a function, let $\mathbf{r} = (r_1, \ldots, r_n) \in \mathsf{R}^n$ and $\mathbf{q} = (q_1, \ldots, q_n) \in (0, \infty)^n$ such that

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

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

$$\hat{I}_f(\mathbf{r}, \mathbf{q}) := \sum_{s=1}^{n} q_s f\left(\frac{r_s}{q_s}\right). \tag{11}$$

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

Theorem 2.1 *Assume* (H_1), *let* $I \subset \mathsf{R}$ *be an interval, and let* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *be in* $(0, \infty)^n$ *such that*

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

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

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

$$\geq A_{m,m}^{[1]} \geq f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s}\right) \sum_{s=1}^{n} q_s, \tag{12}$$

where

$$A_{m,l}^{[1]} = \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{q_{i_j}}{\alpha_{I_m,i_j}}\right) f\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{q_{i_j}}{\alpha_{I_m,i_j}}}\right). \tag{13}$$

If f is a concave function, then the inequality signs in (12) *are reversed.*

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

$$\left(\sum_{s=1}^{n} r_s\right) f\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) \leq A_{m,m}^{[2]} \leq A_{m,m-1}^{[2]} \leq \cdots \leq A_{m,2}^{[2]} \leq A_{m,1}^{[2]}$$

$$= \sum_{s=1}^{n} r_s f\left(\frac{r_s}{q_S}\right) = \hat{I}_{\mathrm{id}f}(\mathbf{r}, \mathbf{q}), \tag{14}$$

where

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

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

$$f\left(\sum_{s=1}^{n} \frac{q_s}{\sum_{s=1}^{n} q_s} \frac{r_s}{q_s}\right) \leq \cdots \leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1,\ldots,i_l),$$

$$\left(\sum_{j=1}^{l} \frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}\right) f\left(\frac{\sum_{j=1}^{l} \frac{\frac{q_{i_j}}{\sum_{i=1}^{n} q_i}}{\alpha_{I_m,i_j}} \frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^{l} \frac{\frac{q_{i_j}}{\sum_{i=1}^{n} q_i}}{\alpha_{I_m,i_j}}}\right) \leq \cdots \leq \sum_{s=1}^{n} \frac{q_s}{\sum_{i=1}^{n} q_s} f\left(\frac{r_s}{q_s}\right). \tag{15}$$

And taking the sum $\sum_{s=1}^{n} q_i$, we have (12).

(ii) Using $f := \mathrm{id}f$ (where "*id*" is the identity function) in Theorem 1.1, we have

$$\sum_{s=1}^{n} p_s x_s f\left(\sum_{s=1}^{n} p_s x_s\right) \leq \cdots \leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1,\ldots,i_l),$$

$$\left(\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}\right) \left(\frac{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}} x_{i_j}}{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}}\right) f\left(\frac{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}} x_{i_j}}{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}}\right) \leq \cdots \leq \sum_{s=1}^{n} p_s x_s f(x_s). \tag{16}$$

Now, on 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

$$
\sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s} f\left(\sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s}\right) \le \cdots \le \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l)\in I_l} \eta_{I_m,l}(i_1,\dots,i_l),
$$

$$
\left(\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m,i_j}}\right)\left(\frac{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m,i_j}} \frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m,i_j}}}\right) f\left(\frac{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m,i_j}} \frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^l \frac{\frac{q_{i_j}}{\sum_{s=1}^n q_s}}{\alpha_{I_m,i_j}}}\right) \tag{17}
$$

$$
\le \cdots \le \sum_{s=1}^n \frac{q_s}{\sum_{s=1}^n q_s} \frac{r_s}{q_s} f\left(\frac{r_s}{q_S}\right).
$$

On taking sum $\sum_{s=1}^n q_s$ on both sides, we get (14). □

3 Inequalities for Shannon entropy

Definition 3 (See [17]) Let $\mathbf{r} = (r_1, \dots, r_n)$ be a positive probability distribution, the Shannon entropy of $\mathbf{r}$ is defined by

$$
S := -\sum_{s=1}^n r_s \log(r_s). \tag{18}
$$

Corollary 3.1 *Assume* (H_1).

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

$$
S \le A_{m,m}^{[3]} \le A_{m,m-1}^{[3]} \le \cdots \le A_{m,2}^{[3]} \le A_{m,1}^{[3]} = \log\left(\frac{n}{\sum_{s=1}^n q_s}\right) \sum_{s=1}^n q_s, \tag{19}
$$

where

$$
A_{m,l}^{[3]} = -\frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l)\in I_l} \eta_{I_m,l}(i_1,\dots,i_l)\left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m,i_j}}\right)\log\left(\sum_{j=1}^l \frac{q_{i_j}}{\alpha_{I_m,i_j}}\right). \tag{20}
$$

And in case $\log$ *is between* 0 *and* 1, *then the reverse sign of inequalities holds in* (19).

(ii) *Suppose that the base of* $\log$ *is greater than* 1, *if* $\mathbf{q} = (q_1, \dots, q_n)$ *is a positive probability distribution, then*

$$
S \le A_{m,m}^{[4]} \le A_{m,m-1}^{[4]} \le \cdots \le A_{m,2}^{[4]} \le A_{m,1}^{[4]} = \log(n), \tag{21}
$$

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 2.1(i), we get (19).

(ii) It is the special case of (i). □

Definition 4 (See [17]) Let $\mathbf{r} = (r_1, \ldots, r_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ be positive probability distributions, the Kullback–Leibler divergence between $\mathbf{r}$ and $\mathbf{q}$ is defined by

$$D(\mathbf{r}, \mathbf{q}) := \sum_{s=1}^{n} r_i \log\left(\frac{r_i}{q_i}\right). \tag{22}$$

Corollary 3.2 *Assume* (H_1).

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

$$\sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) \le A_{m,m}^{[5]} \le A_{m,m-1}^{[5]} \le \cdots \le A_{m,2}^{[5]} \le A_{m,1}^{[5]}$$

$$= \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right) = D(\mathbf{r}, \mathbf{q}), \tag{23}$$

where

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

And in case $\log$ *is between* 0 *and* 1, *then the reverse sign of inequalities holds in* (23).

(ii) *Suppose that the base of* $\log$ *is greater than* 1, *if* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *are two positive probability distributions, then*

$$D(\mathbf{r}, \mathbf{q}) = A_{m,1}^{[6]} \ge A_{m,2}^{[6]} \ge \cdots \ge A_{m,m-1}^{[6]} \ge A_{m,m}^{[6]} \ge 0, \tag{24}$$

where

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

And in case $\log$ *is between* 0 *and* 1, *then the reverse sign of inequalities holds in* (24).

Proof

(i) On taking $f := \log$ in Theorem 2.1(ii), we get (23).

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

4 Inequalities for Rényi divergence and entropy

In [26] Rényi divergence and entropy is given as follows.

Definition 5 Let $\mathbf{r} := (r_1, \ldots, r_n)$ and $\mathbf{q} := (q_1, \ldots, q_n)$ be positive probability distributions, and let $\lambda \ge 0$, $\lambda \ne 1$.

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

$$D_\lambda(\mathbf{r}, \mathbf{q}) := \frac{1}{\lambda - 1} \log\left(\sum_{i=1}^{n} q_i \left(\frac{r_i}{q_i}\right)^\lambda\right). \tag{25}$$

(b) The Rényi entropy of order λ of r is defined by

$$H_\lambda(\mathbf{r}) := \frac{1}{1-\lambda} \log\left(\sum_{i=1}^n r_i^\lambda\right). \tag{26}$$

The Rényi divergence (25) and the Rényi entropy (26) can also be extended to nonnegative probability distributions. Note that $\lim_{\lambda\to 1} D_\lambda(\mathbf{r},\mathbf{q}) = D(\mathbf{r},\mathbf{q})$ and $\lim_{\lambda\to 1} H_\lambda(\mathbf{r}) = S$.

The next two results are given for Rényi divergence.

Theorem 4.1 *Assume* (H_1), *let* $\mathbf{r} = (r_1,\ldots,r_n)$ *and* $\mathbf{q} = (q_1,\ldots,q_n)$ *be probability distributions.*

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

$$D_\lambda(\mathbf{r},\mathbf{q}) \le A_{m,m}^{[7]} \le A_{m,m-1}^{[7]} \le \cdots \le A_{m,2}^{[7]} \le A_{m,1}^{[7]} = D_\mu(\mathbf{r},\mathbf{q}), \tag{27}$$

where

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

And in case $\log$ *is between* 0 *and* 1, *then the reverse sign of inequalities holds in* (27).

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

$$D_1(\mathbf{r},\mathbf{q}) = D(\mathbf{r},\mathbf{q}) = \sum_{s=1}^n r_s \log\left(\frac{r_s}{q_s}\right) \le A_{m,m}^{[8]} \le A_{m,m-1}^{[8]} \le \cdots \le A_{m,2}^{[8]} \le A_{m,1}^{[8]}$$
$$= D_\mu(\mathbf{r},\mathbf{q}), \tag{28}$$

where

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

Here, the $\exp$ *and* $\log$ *functions have the same bases, and if the base of* $\log$ *is in the interval* $(0,1)$, *then the reverse sign of inequalities holds in* (28).

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

$$D_\lambda(\mathbf{r},\mathbf{q}) \le A_{m,m}^{[9]} \le A_{m,m-1}^{[9]} \le \cdots \le A_{m,2}^{[9]} \le A_{m,1}^{[9]} = D_1(\mathbf{r},\mathbf{q}), \tag{29}$$

where

$$A_{m,l}^{[9]} = \frac{1}{\lambda-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l)\in I_l} \eta_{I_m,l}(i_1,\dots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right)$$
$$\times \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} (\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right). \tag{30}$$

Proof By taking $I = (0,\infty), f : (0,\infty) \to \mathsf{R}, f(t) := t^{\frac{\mu-1}{\lambda-1}}$

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

in Theorem 1.1, we have

$$\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s} \right)^{\lambda} \right)^{\frac{\mu-1}{\lambda-1}}$$
$$= \left(\sum_{s=1}^{n} r_s \left(\frac{r_s}{q_s} \right)^{\lambda-1} \right)^{\frac{\mu-1}{\lambda-1}}$$
$$\leq \cdots \leq \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l)\in I_l} \eta_{I_m,l}(i_1,\dots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} (\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right)^{\frac{\mu-1}{\lambda-1}}$$
$$\leq \cdots \leq \sum_{s=1}^{n} r_s \left(\left(\frac{r_s}{q_s} \right)^{\lambda-1} \right)^{\frac{\mu-1}{\lambda-1}}, \tag{31}$$

if either $0 \leq \lambda < 1 < \beta$ or $1 < \lambda \leq \mu$, and the reverse inequality in (31) holds if $0 \leq \lambda \leq \beta < 1$. By raising to power $\frac{1}{\mu-1}$, we have from all

$$\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s} \right)^{\lambda} \right)^{\frac{1}{\lambda-1}}$$
$$\leq \cdots \leq \left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l)\in I_l} \eta_{I_m,l}(i_1,\dots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \right.$$
$$\times \left. \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} (\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \right)^{\frac{1}{\mu-1}}$$
$$\leq \cdots \leq \left(\sum_{s=1}^{n} r_s \left(\left(\frac{r_s}{q_s} \right)^{\lambda-1} \right)^{\frac{\mu-1}{\lambda-1}} \right)^{\frac{1}{\mu-1}} = \left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s} \right)^{\mu} \right)^{\frac{1}{\mu-1}}. \tag{32}$$

Since the log function is increasing for the base greater than 1, therefore on taking log in (32) we get (29). And the log function is decreasing for the base between 0 and 1, in this case on taking log in (32) we get the reverse sign in (27). If $\lambda = 1$ and $\beta = 1$, we have (ii) and (iii) respectively by taking limit. □

Theorem 4.2 *Assume* (H_1), *let* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *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*

$$\frac{1}{\sum_{s=1}^{n} q_s (\frac{r_s}{q_s})^\lambda} \sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^\lambda \log\left(\frac{r_s}{q_s}\right) = A_{m,1}^{[10]} \leq A_{m,2}^{[10]} \leq \cdots \leq A_{m,m-1}^{[10]} \leq A_{m,m}^{[10]}$$
$$\leq D_\lambda(r,q) \leq A_{m,m}^{[11]}$$
$$\leq A_{m,m}^{[11]} \leq \cdots \leq A_{m,2}^{[11]} \leq A_{m,1}^{[11]} = D_1(\mathbf{r},\mathbf{q}), \qquad (33)$$

where

$$A_{m,m}^{[10]} = \frac{1}{(\lambda - 1) \sum_{s=1}^{n} q_s (\frac{r_s}{q_s})^\lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}\right)$$
$$\times \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} (\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)$$

and

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

The inequalities in (33) 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 Here we prove for $0 \leq \lambda < 1$ and base when the base of $\log$ is greater than 1, the other case can be proved by following similar steps. Since $\frac{1}{\lambda-1} < 0$ and the function $\log$ is concave, then choosing $I = (0, \infty)$, $f := \log$, $p_s = r_s$, $x_s := (\frac{r_s}{q_s})^{\lambda-1}$ in Theorem 1.1, we have

$$D_\lambda(\mathbf{r},\mathbf{q}) = \frac{1}{\lambda - 1} \log\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^\lambda\right) = \frac{1}{\lambda - 1} \log\left(\sum_{s=1}^{n} r_s \left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)$$
$$\leq \cdots \leq \frac{1}{\lambda - 1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\right)$$
$$\times \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} (\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)$$
$$\leq \cdots \leq \frac{1}{\lambda - 1} \sum_{s=1}^{n} r_s \log\left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right) = \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right) = D_1(\mathbf{r},\mathbf{q}) \qquad (34)$$

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

Since $x \mapsto x \log(x)$ $(x > 0)$ is a convex function for base of $\log \log \log \log$ greater than 1, also $\frac{1}{1-\lambda} < 0$, therefore using Theorem 1.1, we have

$$
\begin{aligned}
D_\lambda(\mathbf{r}, \mathbf{q}) &= \frac{1}{\lambda - 1} \log\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^\lambda\right) \\
&= \frac{1}{\lambda - 1\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^\lambda\right)} \left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^\lambda\right) \log\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^\lambda\right) \\
&\geq \cdots \geq \frac{1}{\lambda - 1\left(\sum_{s=1}^{n} q_s(\frac{r_s}{q_s})^\lambda\right)} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\right) \\
&\quad \times \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}(\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right) \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}(\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right) \\
&= \frac{1}{\lambda - 1\left(\sum_{s=1}^{n} q_s(\frac{r_s}{q_s})^\lambda\right)} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \\
&\quad \times \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}\right) \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}(\frac{r_{i_j}}{q_{i_j}})^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right) \\
&\geq \cdots \\
&\geq \frac{1}{\lambda - 1} \sum_{s=1}^{n} r_s\left(\frac{r_s}{q_s}\right)^{\lambda-1} \log\left(\frac{r_s}{q_s}\right)^{\lambda-1} \frac{1}{\sum_{s=1}^{n} r_s(\frac{r_s}{q_s})^{\lambda-1}} \\
&= \frac{1}{\sum_{s=1}^{n} q_s(\frac{r_s}{q_s})^\lambda} \sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^\lambda \log\left(\frac{r_s}{q_s}\right)
\end{aligned} \tag{35}
$$

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

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

Corollary 4.3 *Assume* (H_1), *let* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *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*

$$
H_\lambda(\mathbf{r}) = \log(n) - D_\lambda\left(\mathbf{r}, \frac{1}{\mathbf{n}}\right) \geq A_{m,m}^{[12]} \geq A_{m,m}^{[12]} \geq \cdots \geq A_{m,2}^{[12]} \geq A_{m,1}^{[12]} = H_\mu(\mathbf{r}), \tag{36}
$$

where

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

The reverse inequalities hold in (36) *if the base of* $\log$ *is between* 0 *and* 1.

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

$$S = -\sum_{s=1}^{n} p_i \log(p_i) \geq A_{m,m}^{[13]} \geq A_{m,m-1}^{[13]} \geq \cdots \geq A_{m,2}^{[13]} \geq A_{m,1}^{[13]} = H_\mu(\mathbf{r}), \qquad (37)$$

where

$$A_{m,l}^{[13]} = \log(n) + \frac{1}{1-\mu} \log\left(\frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l)\left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\right)\right.$$

$$\left. \times \exp\left(\frac{(\mu-1)\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \log(nr_{i_j})}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)\right),$$

the exp *and* log *functios have the same bases. If the base of* log *is between* 0 *and* 1, *the sign of inequalities in* (37) *is reversed.*

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

$$H_\lambda(\mathbf{r}) \geq A_{m,m}^{[14]} \geq A_{m,m-1}^{[14]} \geq \cdots \geq A_{m,2}^{[14]} \leq A_{m,1}^{[14]} = S, \qquad (38)$$

where

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

$$\times \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}^{\lambda}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right). \qquad (39)$$

The inequalities in (38) *are reversed if the base of* log *is between* 0 *and* 1.

Proof (i) Suppose $\mathbf{q} = \frac{1}{\mathbf{n}}$, then from (25) we have

$$D_\lambda(\mathbf{r},\mathbf{q}) = \frac{1}{\lambda-1} \log\left(\sum_{s=1}^{n} n^{\lambda-1} r_s^{\lambda}\right) = \log(n) + \frac{1}{\lambda-1} \log\left(\sum_{s=1}^{n} r_s^{\lambda}\right), \qquad (40)$$

therefore we have

$$H_\lambda(\mathbf{r}) = \log(n) - D_\lambda\left(\mathbf{r}, \frac{1}{\mathbf{n}}\right). \qquad (41)$$

Now, using Theorem 4.1(i) and (41), we get

$$H_\lambda(\mathbf{r}) = \log(n) - D_\lambda\left(\mathbf{r}, \frac{1}{\mathbf{n}}\right) \geq \cdots$$

$$\geq \log(n) - \frac{1}{\mu-1} \log\left(n^{\mu-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l)\right.$$

$$\times \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}^{\lambda}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \geq \cdots$$

$$\geq \log(n) - D_\mu(\mathbf{r}, \mathbf{q}) = H_\mu(\mathbf{r}), \tag{42}$$

(ii) and (iii) can be proved similarly. □

Corollary 4.4 *Assume* (H_1) *and let* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *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*

$$-\frac{1}{\sum_{s=1}^{n} r_s^{\lambda}} \sum_{s=1}^{n} r_s^{\lambda} \log(r_s) = A_{m,1}^{[15]} \geq A_{m,2}^{[15]} \geq \cdots \geq A_{m,m-1}^{[15]} \geq A_{m,m}^{[15]}$$

$$\geq H_\lambda(r) \geq A_{m,m}^{[16]} \geq A_{m,m-1}^{[16]} \geq \cdots \geq A_{m,2}^{[16]} \geq A_{m,1}^{[16]} = H(r), \tag{43}$$

where

$$A_{m,l}^{[15]} = \frac{1}{(\lambda-1)\sum_{s=1}^{n} r_s^{\lambda}} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}^{\lambda}}{\alpha_{I_m,i_j}} \right)$$

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

and

$$A_{m,1}^{[16]} = \frac{1}{1-\lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}^{\lambda}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right).$$

The inequalities in (43) *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 4.3 by using Theorem 4.2. □

5 Inequalities by using Zipf–Mandelbrot law
In [22] the Zipf–Mandelbrot law is defined as follows.

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

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

where

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

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

$$f(s;q,t) = \frac{1}{(s+q)^t H_{q,t}}, \tag{46}$$

where

$$H_{q,t} = \sum_{j=1}^{\infty} \frac{1}{(j+q)^t}. \tag{47}$$

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

Conclusion 5.1 *Assume* (H_1), *let r be a Zipf–Mandelbrot law, by Corollary* 4.3(*iii*), *we get*: *If* $0 \leq \lambda < 1$ *and the base of* log *is greater than* 1, *then*

$$H_\lambda(\mathbf{r}) = \frac{1}{1-\lambda} \log\left(\frac{1}{H_{N,q,t}^\lambda} \sum_{s=1}^{n} \frac{1}{(s+q)^{\lambda s}} \right) \geq \cdots$$

$$\geq \frac{1}{1-\lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l) \in I_l} \eta_{I_m,l}(i_1,\dots,i_l) \left(\sum_{j=1}^{l} \frac{1}{\alpha_{I_m,i_j}(i_j+q)H_{N,q,t}} \right)$$

$$\times \log\left(\frac{1}{H_{N,q,t}^{\lambda-1}} \frac{\sum_{j=1}^{l} \frac{1}{\alpha_{I_m,i_j}(i_j-q)^{\lambda s}}}{\sum_{j=1}^{l} \frac{1}{\alpha_{I_m,i_j}(i_j-q)^s}} \right) \geq \cdots$$

$$\geq \frac{t}{H_{N,q,t}} \sum_{s=1}^{N} \frac{\log(s+q)}{(s+q)^t} + \log(H_{N,q,t}) = S. \tag{48}$$

The inequalities in (48) are reversed if the base of log is between 0 and 1.

Conclusion 5.2 *Assume* (H_1), *let* r_1 *and* r_2 *be the Zipf–Mandelbort 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* 3.2(*ii*), *we have*: *If the base of* log *is greater than* 1, *then*

$$\bar{D}(r_1,r_2) = \sum_{s=1}^{n} \frac{1}{(s+q_1)^{t_1} H_{N,q_1,t_1}} \log\left(\frac{(s+q_2)^{t_2} H_{N,q_2,t_2}}{(s+q_1)^{t_1} H_{N,q_2,t_1}} \right) \geq \cdots$$

$$\geq \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\dots,i_l) \in I_l} \eta_{I_m,l}(i_1,\dots,i_l), \tag{49}$$

$$\left(\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}} \right) \left(\frac{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}} \right) \log\left(\frac{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}} \right)$$

$$\geq \cdots \geq 0.$$

The inequalities in (49) are reversed if the 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 equation constraints and get the Zipf–Mandelbrot law.

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

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

is the Zipf–Madelbrot law.

Proof If $J = \{1, 2, \ldots, N\}$, we set the Lagrange multipliers λ and t and consider the expression

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

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

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

From $\hat{S}_{r_s} = 0$, for $s = 1, 2, \ldots, N$, 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

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

Remark 6.2 Observe that the Zipf–Mandelbrot law and Shannon entropy can be bounded from above (see [23]).

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

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

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

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

is a hybrid Zipf–Mandelbrot law given as

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

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 by

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

with $w \geq 0$, $k > 0$. $\qquad \square$

Remark 6.4 Observe that for the Zipf–Mandelbrot law, Shannon entropy can be bounded from above (see [23]).

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

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 2.1(i), define the nonnegative functionals as follows:

$$\Theta_3(f) = A_{m,r}^{[1]} - f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s}\right) \sum_{s=1}^{n} q_s, \quad r = 1, \ldots, m, \tag{50}$$

$$\Theta_4(f) = A_{m,r}^{[1]} - A_{m,k}^{[1]}, \quad 1 \leq r < k \leq m. \tag{51}$$

Under the assumption of Theorem 2.1(ii), define the nonnegative functionals as follows:

$$\Theta_5(f) = A_{m,r}^{[2]} - \left(\sum_{s=1}^{n} r_s\right) f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s}\right), \quad r = 1, \ldots, m, \tag{52}$$

$$\Theta_6(f) = A_{m,r}^{[2]} - A_{m,k}^{[2]}, \quad 1 \leq r < k \leq m. \tag{53}$$

Under the assumption of Corollary 3.1(i), define the following nonnegative functionals:

$$\Theta_7(f) = A_{m,r}^{[3]} + \sum_{i=1}^{n} q_i \log(q_i), \quad r = 1, \ldots, n, \tag{54}$$

$$\Theta_8(f) = A_{m,r}^{[3]} - A_{m,k}^{[3]}, \quad 1 \leq r < k \leq m. \tag{55}$$

Under the assumption of Corollary 3.1(ii), define the following nonnegative functionals given as

$$\Theta_9(f) = A_{m,r}^{[4]} - S, \quad r = 1, \ldots, m, \tag{56}$$

$$\Theta_{10}(f) = A_{m,r}^{[4]} - A_{m,k}^{[4]}, \quad 1 \leq r < k \leq m. \tag{57}$$

Under the assumption of Corollary 3.2(i), let us define the nonnegative functionals as follows:

$$\Theta_{11}(f) = A_{m,r}^{[5]} - \sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \log \frac{r_n}{\sum_{s=1}^{n} q_s}\right), \quad r = 1, \ldots, m, \tag{58}$$

$$\Theta_{12}(f) = A_{m,r}^{[5]} - A_{m,k}^{[5]}, \quad 1 \leq r < k \leq m. \tag{59}$$

Under the assumption of Corollary 3.2(ii), define the nonnegative functionals as follows:

$$\Theta_{13}(f) = A_{m,r}^{[6]} - A_{m,k}^{[6]}, \quad 1 \leq r < k \leq m. \tag{60}$$

Under the assumption of Theorem 4.1(i), consider the following functionals:

$$\Theta_{14}(f) = A_{m,r}^{[7]} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m, \tag{61}$$

$$\Theta_{15}(f) = A_{m,r}^{[7]} - A_{m,k}^{[7]}, \quad 1 \leq r < k \leq m. \tag{62}$$

Under the assumption of Theorem 4.1(ii), consider the following functionals:

$$\Theta_{16}(f) = A_{m,r}^{[8]} - D_1(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m, \tag{63}$$

$$\Theta_{17}(f) = A_{m,r}^{[8]} - A_{m,k}^{[8]}, \quad 1 \leq r < k \leq m. \tag{64}$$

Under the assumption of Theorem 4.1(iii), consider the following functionals:

$$\Theta_{18}(f) = A_{m,r}^{[9]} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m, \tag{65}$$

$$\Theta_{19}(f) = A_{m,r}^{[9]} - A_{m,k}^{[9]}, \quad 1 \le r < k \le m. \tag{66}$$

Under the assumption of Theorem 4.2, consider the following nonnegative functionals:

$$\Theta_{20}(f) = D_\lambda(\mathbf{r}, \mathbf{q}) - A_{m,r}^{[10]}, \quad r = 1, \ldots, m, \tag{67}$$

$$\Theta_{21}(f) = A_{m,k}^{[10]} - A_{m,r}^{[10]}, \quad 1 \le r < k \le m. \tag{68}$$

$$\Theta_{22}(f) = A_{m,r}^{[11]} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m, \tag{69}$$

$$\Theta_{23}(f) = A_{m,r}^{[11]} - A_{m,r}^{[11]}, \quad 1 \le r < k \le m, \tag{70}$$

$$\Theta_{24}(f) = A_{m,r}^{[11]} - A_{m,k}^{[10]}, \quad r = 1, \ldots, m, k = 1, \ldots, m. \tag{71}$$

Under the assumption of Corollary 4.3(i), consider the following nonnegative functionals:

$$\Theta_{25}(f) = H_\lambda(r) - A_{m,r}^{[12]}, \quad r = 1, \ldots, m, \tag{72}$$

$$\Theta_{26}(f) = A_{m,k}^{[12]} - A_{m,r}^{[12]}, \quad 1 \le r < k \le m. \tag{73}$$

Under the assumption of Corollary 4.3(ii), consider the following functionals:

$$\Theta_{27}(f) = S - A_{m,r}^{[13]}, \quad r = 1, \ldots, m, \tag{74}$$

$$\Theta_{28}(f) = A_{m,k}^{[13]} - A_{m,r}^{[13]}, \quad 1 \le r < k \le m. \tag{75}$$

Under the assumption of Corollary 4.3(iii), consider the following functionals:

$$\Theta_{29}(f) = H_\lambda(\mathbf{r}) - A_{m,r}^{[14]}, \quad r = 1, \ldots, m, \tag{76}$$

$$\Theta_{30}(f) = A_{m,k}^{[14]} - A_{m,r}^{[14]}, \quad 1 \le r < k \le m. \tag{77}$$

Under the assumption of Corollary 4.4, define the following functionals:

$$\Theta_{31} = A_{m,r}^{[15]} - H_\lambda(r), \quad r = 1, \ldots, m, \tag{78}$$

$$\Theta_{32} = A_{m,r}^{[15]} - A_{m,k}^{[15]}, \quad 1 \le r < k \le m, \tag{79}$$

$$\Theta_{33} = H_\lambda(\mathbf{r}) - A_{m,r}^{[16]}, \quad r = 1, \ldots, m, \tag{80}$$

$$\Theta_{34} = A_{m,k}^{[16]} - A_{m,r}^{[16]}, \quad 1 \le r < k \le m, \tag{81}$$

$$\Theta_{35} = A_{m,r}^{[15]} - A_{m,k}^{[16]}, \quad r = 1, \ldots, m, k = 1, \ldots, m. \tag{82}$$

7 Generalization of the refinement of Jensen's, Rényi, and Shannon type inequalities via Montgomery identity

We construct some new identities with the help of the generalized Montgomery identity (5).

Theorem 7.1 *Assume (H_1), let $f : [\alpha_1, \alpha_2] \to \mathsf{R}$ be a function where $[\alpha_1, \alpha_2] \subset \mathsf{R}$ is an interval. Also let $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \ldots, p_n$ be positive real numbers such that $\sum_{i=1}^{n} p_i = 1$, and $R_m(x, u)$ be the same as defined in (6), then the following identity holds*:

$$
\Theta_i(f) = \frac{1}{\alpha_2 - \alpha_1} \sum_{k=0}^{m-2} \left(\frac{1}{k!(k+2)} \right) \left(f^{(k+1)}(\alpha_1) \Theta_i\big((x - \alpha_1)^{k+1} \big) - f^{(k+1)}(\alpha_2) \right.
$$
$$
\left. \times \, \Theta_2\big((x - \alpha_2)^{k+1} \big) \right) \frac{1}{(m-1)!} \int_{\alpha_1}^{\alpha_2} \Theta_i\big(R_m(x, u) \big) f^{(m)}(u) \, du, \quad i = 1, \ldots, 35. \quad (83)
$$

Proof Using (5) in (2), (3), and (50)–(82), we get the result. □

Theorem 7.2 *Assume (H_1), let $f : [\alpha_1, \alpha_2] \to \mathsf{R}$ be a function where $[\alpha_1, \alpha_2] \subset \mathsf{R}$ is an interval. Also let $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \ldots, p_n$ be positive real numbers such that $\sum_{i=1}^{n} p_i = 1$, and $R_m(x, u)$ be the same as defined in (6). Let, for $m \geq 2$,*

$$
\Theta_i\big(R_m(x, u) \big) \geq 0 \quad \text{for all } u \in [\alpha_1, \alpha_2], i = 1, \ldots, 35.
$$

If f is m-convex such that $f^{(m-1)}$ is absolutely continuous, then

$$
\Theta_i(f) \geq \frac{1}{\alpha_2 - \alpha_1} \sum_{k=0}^{m-2} \left(\frac{1}{k!(k+2)} \right)
$$
$$
\times \, (f^{(k+1)}(\alpha_1) \Theta_i\big((x - \alpha_1)^{k+1} \big) - f^{(k+1)}(\alpha_2) \Theta_i\big((x - \alpha_2)^{k+1} \big), \quad i = 1, \ldots, 35. \quad (84)
$$

Proof As $f^{(m-1)}$ is absolutely continuous on $[\alpha_1, \alpha_2]$, therefore $f^{(m)}$ exists almost everywhere. As f is m-convex, so $f^{(m)}(u) \geq 0$ for all $u \in [\alpha_1, \alpha_2]$ (see [25, p. 16]). Hence, using Theorem 7.1, we get (84). □

Theorem 7.3 *Assume (H_1), let $f : [\alpha_1, \alpha_2] \to \mathsf{R}$ be a function where $[\alpha_1, \alpha_2] \subset \mathsf{R}$ is an interval. Also let $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \ldots, p_n$ be positive real numbers such that $\sum_{i=1}^{n} p_i = 1$, let $f : [\alpha_1, \alpha_2] \to \mathsf{R}$ be a convex function.*
  (i) *If $m \geq 2$ is even, then (84) holds.*
  (ii) *Let (84) be valid. If the function*

$$
\lambda(x) = \frac{1}{\alpha_2 - \alpha_1} \sum_{l=0}^{m-2} \left(\frac{f^{(l+1)}(\alpha_1)(x - \alpha_1)^{l+2} - f^{(l+1)}(\alpha_2)(x - \alpha_2)^{l+2}}{l!(l+2)} \right)
$$

is convex, then the right-hand side of (84) is nonnegative and

$$
\Theta_i(f) \geq 0, \quad i = 1, \ldots, 35.
$$

Proof
  (i) The function $R_m(\cdot, v)$ is convex (see [10]). Hence, for an even integer $m \geq 2$,

$$
\Theta_i\big(R_m(u, v) \big) \geq 0,
$$

therefore from Theorem 7.2, we have (84).

(ii) By using the linearity of $\Theta_i(f)$, we can write the right-hand side of (84) in the form $\Theta_i(\lambda)$. As λ is supposed to be convex, therefore the right-hand side of (84) is nonnegative, so $\Theta_i(f) \geq 0$. □

Theorem 7.4 *Assume* (H_1), *let* $f : [\alpha_1, \alpha_2] \to \mathsf{R}$ *be a function where* $[\alpha_1, \alpha_2] \subset \mathsf{R}$ *is an interval. Also let* $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ *and* $p_1, \ldots, p_n$ *be positive real numbers such that* $\sum_{i=1}^{n} p_i = 1$, *and* $\hat{R}_m(x, u)$ *be the same as defined in* (8), *then the following identity holds*:

$$\Theta_i(f) = \frac{1}{\alpha_2 - \alpha_1} \sum_{k=0}^{m-2} \left(\frac{1}{k!(k+2)} \right) (\Theta_i\big(f^{(k+1)}(x)(\alpha_1 - x)^{k+1}\big) - \Theta_i\big(f^{(k+1)}(x)(\alpha_2 - x)^{k+1}\big)$$

$$+ \frac{1}{(m-1)!} \int_{\alpha_1}^{\alpha_2} \Theta_i\big(\hat{R}_m(x, u)\big)f^{(m)}(u)\, du, \quad i = 1, \ldots, 35. \tag{85}$$

Proof Using (7) in (2), (3), and (50)–(82), we get identity (85). □

Theorem 7.5 *Assume* (H_1), *let* $f : [\alpha_1, \alpha_2] \to \mathsf{R}$ *be a function where* $[\alpha_1, \alpha_2] \subset \mathsf{R}$ *is an interval. Also let* $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ *and* $p_1, \ldots, p_n$ *be positive real numbers such that* $\sum_{i=1}^{n} p_i = 1$, *and* $\hat{R}_m(x, u)$ *be the same as defined in* (8). *Let, for* $m \geq 2$,

$$\Theta_i\big(\hat{R}_m(x, u)\big) \geq 0 \quad \text{for all } u \in [\alpha_1, \alpha_2], i = 1, \ldots, 35.$$

If f *is* m-*convex such that* $f^{(m-1)}$ *is absolutely continuous, then*

$$\Theta_i(f) \geq \frac{1}{\alpha_2 - \alpha_1} \sum_{k=0}^{m-2} \left(\frac{1}{k!(k+2)} \right)$$

$$\times (\Theta_i\big(f^{(k+1)}(x)(\alpha_1 - x)^{k+1}\big) - \Theta_i\big(f^{(k+1)}(x)(\alpha_2 - x)^{k+1}\big), \quad i = 1, \ldots, 35. \tag{86}$$

Proof As $f^{(m-1)}$ is absolutely continuous on $[\alpha_1, \alpha_2]$, therefore $f^{(m)}$ exists almost everywhere. As f is m-convex, so $f^{(m)}(u) \geq 0$ for all $u \in [\alpha_1, \alpha_2]$ (see [25, p. 16]). Hence, using Theorem 7.4, we get (86). □

Remark 7.6 We can get a similar result as that given in Theorem 7.3.

Remark 7.7 We can give related mean value theorems, also construct the new families of m-exponentialy convex functions and Cauchy means related to the functionals Θ_i, $i = 1, \ldots, 35$, as given in [7].

Authors' contributions
This is a joint work of all authors. All authors read and approved the final manuscript.

Author details
[1]Department of Mathematics, University of Sargodha, Sargodha, Pakistan. [2]Department of Mathematics, The University of Lahore, Sargodha, Pakistan. [3]Catholic University of Croatia, Zagreb, Croatia. [4]Rudn University, Moscow, Russia.

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

References

1. Aljinobić, A.A., Pečarić, J., Vukelić, A.: On some Ostrowski type inequalities via Montgomery identity and Taylor's formula II. Tamkang J. Math. **36**(4), 279–301 (2005)
2. Aljinović, A.A., Pečarić, J.: On some Ostrowski type inequalities via Montgomery identity and Taylor's formula. Tamkang J. Math. **36**(3), 199–218 (2005)
3. Anderson, G., Ge, Y.: The size distribution of Chinese cities. Reg. Sci. Urban Econ. **35**(6), 756–776 (2005)
4. Auerbach, F.: Das Gesetz der Bevölkerungskonzentration. Petermanns Geogr. Mitt. **59**, 74–76 (1913)
5. Black, D., Henderson, V.: Urban evolution in the USA. J. Econ. Geogr. **3**(4), 343–372 (2003)
6. 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)
7. Butt, S.I., Khan, K.A., Pečarić, J.: Generaliztion of Popoviciu inequality for higher order convex function via Taylor's polynomial. Acta Univ. Apulensis, Mat.-Inform. **42**, 181–200 (2015)
8. Butt, S.I., Khan, K.A., Pečarić, J.: Popoviciu type inequalities via Hermite's polynomial. Math. Inequal. Appl. **19**(4), 1309–1318 (2016)
9. 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)
10. Butt, S.I., Pečarić, J.: Weighted Popoviciu type inequalities via generalized Montgomery identities. Rad Hrvat. Akad. Znan. Umjet. Mat. Znan. **19**(523), 69–89 (2015)
11. Butt, S.I., Pečarić, J.: Popoviciu's Inequality for *N*-Convex Functions. Lap Lambert Academic Publishing, Saarbrücken (2016)
12. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hung. **2**, 299–318 (1967)
13. Csiszár, I.: Information measures: a critical survey. In: Trans. 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)
14. 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)
15. Horváth, L., Khan, K.A., Pečarić, J.: Combinatorial Improvements of Jensens Inequality/Classical and New Refinements of Jensens Inequality with Applications, Monographs in Inequalities, vol. 8. Element, Zagreb (2014)
16. Horváth, L., Khan, K.A., Pečarić, J.: Refinement of Jensen's inequality for operator convex functions. Adv. Inequal. Appl. (2014)
17. Horváth, L., Pečarić, Đ., Pečarić, J.: Estimations of *f*- and Rényi divergences by using a cyclic refinement of the Jensen's inequality. In: Bulletin of the Malaysian Mathematical Sciences Society, pp. 1–14. Springer, Singapore (2017)
18. Horváth, L., Pečarić, J.: A refinement of discrete Jensen's inequality. Math. Inequal. Appl. **14**, 777–791 (2011)
19. Ioannides, Y.M., Overman, H.G.: Zipf's law for cities: an empirical examination. Reg. Sci. Urban Econ. **33**(2), 127–137 (2003)
20. Kullback, S.: Information Theory and Statistics. Dover, New York (1997)
21. Kullback, S., Leibler, R.A.: On information and sufficiency. Ann. Math. Stat. **22**(1), 79–86 (1951)
22. Lovričević, N., Pečarić, Đ., Pečarić, J.: Zipf–Mandelbrot law, *f*-divergences and the Jensen-type interpolating inequalities. J. Inequal. Appl. **2018**(1), 36 (2018)
23. 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)
24. Mitrinović, D.S., Pečarić, J., Fink, A.M.: Classical and New Inequalities in Analysis. Kluwer Academic, Dordrecht (1993)
25. Pečarić, J., Proschan, F., Tong, Y.L.: Convex Functions, Partial Orderings and Statistical Applications. Academic Press, New York (1992)
26. Rényi, A.: On measure of information and entropy. In: Proceeding of the Fourth Berkely Symposium on Mathematics, Statistics and Probability, pp. 547–561 (1960)
27. 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)
28. Soo, K.T.: Zipf's law for cities: a cross-country investigation. Reg. Sci. Urban Econ. **35**(3), 239–263 (2005)
29. Zipf, G.K.: Human Behaviour and the Principle of Least-Effort. Addison-Wesley, Reading (1949)