International Journal of Analysis and Applications
Volume 17, Number 5 (2019), 686-710
URL: https://doi.org/10.28924/2291-8639
DOI: 10.28924/2291-8639-17-2019-686

ESTIMATION OF DIFFERENT ENTROPIES VIA TAYLOR ONE POINT AND TAYLOR TWO POINTS INTERPOLATIONS USING JENSEN TYPE FUNCTIONALS

TASADDUQ NIAZ[1,2,*], KHURAM ALI KHAN[1], ĐILDA PEČARIĆ[3], JOSIP PEČARIĆ[4]

[1] *Department of Mathematics, University of Sargodha, Sargodha 40100, Pakistan*
[2] *Department of Mathematics, The University of Lahore, Sargodha-Campus, Sargodha 40100, Pakistan*
[3] *Catholic University of Croatia, Ilica 242, Zagreb, Croatia*
[4] *RUDN University, Miklukho-Maklaya str. 6, 117198 Moscow, Russia*

[*] *Corresponding author:* tasadduq_khan@yahoo.com

Abstract. In this work, we estimated the different entropies like Shannon entropy, Rényi divergences, Csiszar divergence by using the Jensen's type functionals. The Zipf's mandelbrot law and hybrid Zipf's mandelbrot law are used to estimate the Shannon entropy. Further the Taylor one point and Taylor two points interpolations are used to generalize the new inequalities for m-convex function.

1. Introduction and preliminary results

In numerical analysis, interpolation is a method of constructing new data points within the range of a discrete set of known data points for example in the situation when one obtained the number of data after experiment which actually represent the value of function for a limited number of value of the independent variable. It is usually require to interpolate which means that it has to be estimated the value of the function for an intermediate value of independent variable. There are many interpolating polynomial can be found in literature for example Taylor polynomial, Lidstone polynomial etc.

Received 2019-05-27; accepted 2019-07-16; published 2019-09-02.

2010 *Mathematics Subject Classification.* 26D07, 94A17.

Key words and phrases. m-convex function; Jensen's inequality; Shannon entropy; f- and Rényi divergence; Taylor interpolation; entropy.

The most commonly used words, the largest cities of countries income of billionare can be described in term of Zipf's law. The f-divergence which means that distance between two probability distribution by making an average value, which is weighted by a specified function. As f-divergence, there are other probabilities distributions like Csiszar f-divergence [15,16], some special case of which are Kullback-Leibler-divergence use to find the appropriate distance between the probability distribution (see [19,20]). The notion of distance is stronger than divergence because it give the properties of symmetry and triangle inequalities. Probability theory has application in many fields and the divergence between probability distribution have many application in these fields.

Many natural phenomena's like distribution of wealth and income in a society, distribution of face book likes, distribution of football goals follows power law distribution (Zipf's Law). Like above phenomena's, distribution of city sizes also follow Power Law distribution. Auerbach [2] first time gave 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 [28] worked significantly in this field. The distribution of city sizes is investigated by many scholars of the urban economics, like Rosen and Resnick [26] , Black and Henderson [3], Ioannides and Overman [14], Soo [27], 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 appear 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 the 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 not only shows that the city size distribution follows the Pareto distribution, but also show that the estimated value of the shape parameter is equal to unity.

In [17] L. Horváth et al. introduced some 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 construct some 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 T. Popoviciu (see [24, p. 15]). It is quite 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 remaind valid.

In [24, p. 16], the following criteria 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; like S. I. Butt et

al. generalized the Popoviciu inequality for m-convex function using Taylor's formula, Lidstone polynomial, montgomery identity, Fink's identity, Abel-Gonstcharoff interpolation and Hermite interpolating polynomial (see [5–9]).

In [23] T. Niaz et al generalized the refinement of Jensen's inequality for m-convex function using Abel-Gontscharoff green function and Fink's identity. In [18] K. A. Khan et al used refinement of Jensen inequality and introduced new functional based on an f-divergence functional, and estimate some bounds for the new functionals, the f-divergence and Rényi divergence. They also constructed some new inequalities for Réneyi and Shannon estimates. They also generalized the new inequality for m-convex function using Montgomery identity. Further the used hybrid Zipf Mandelbrot law to estimate the Shannon entropy.

Since many years Jensen's inequality has of great interest. The researchers have given the refinement of Jensen's inequality by defining some new functions (see [12, 13]). Like many researchers L. Horváth and J. Pečarić in ([10, 13], see also [11, 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

$$F_{r,s} : \{1, \ldots, q\}^r \to \{1, \ldots, q\}^{r-1} \qquad 1 \leq s \leq r,$$

$$F_r : \{1, \ldots, q\}^r \to P\left(\{1, \ldots, q\}^{r-1}\right),$$

and

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

by

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

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

and

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

Next let the function

$$\alpha_{r,i} : \{1, \ldots, q\}^r \to \mathbb{N} \qquad 1 \leq i \leq q$$

defined by

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

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

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

(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 $\{1,\dots,n\}^m$ such that

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

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

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

Obviously the sets $I_1 = \{1,\dots,n\}$, by (H_1) and this insures 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 $(j_1,\dots,j_{l-1}) \in I_{l-1}$, let

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

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

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

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

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

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

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 $x_1,\dots,x_n \in I$ and $p_1,\dots,p_n$ are positive real numbers such that $\sum_{i=1}^{n} p_i = 1$, then*

$$f\left(\sum_{s=1}^{n} p_s x_s\right) \le \mathcal{A}_{m,m} \le \mathcal{A}_{m,m-1} \le \dots \le \mathcal{A}_{m,2} \le \mathcal{A}_{m,1} = \sum_{s=1}^{n} p_s f(x_s). \tag{1.1}$$

We define the following functionals by taking the differences of refinement of Jensen's inequality given in (1.1).

$$\Theta_1(f) = \mathcal{A}_{m,r} - f\left(\sum_{s=1}^{n} p_s x_s\right), \qquad r = 1, \ldots, m, \tag{1.2}$$

$$\Theta_2(f) = \mathcal{A}_{m,r} - \mathcal{A}_{m,k}, \qquad 1 \le r < k \le m. \tag{1.3}$$

Under the assumptions of Theorem 1.1, we have

$$\Theta_i(f) \ge 0, \qquad i = 1, 2. \tag{1.4}$$

Inequalities (1.4) are reversed if f is concave on I.

2. Inequalities for Csiszár divergence

In [15,16] Csiszár introduced the following notion.

Definition 2.1. *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ 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 f-divergence functional is defined by*

$$I_f(\mathbf{r}, \mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{r_i}{q_i}\right). \tag{2.1}$$

And he stated that by defining

$$f(0) := \lim_{x \to 0^+} f(x); \qquad 0 f\left(\frac{0}{0}\right) := 0; \qquad 0 f\left(\frac{a}{0}\right) := \lim_{x \to 0^+} x f\left(\frac{a}{x}\right), \qquad a > 0, \tag{2.2}$$

we can also use the nonnegative probability distributions as well.

In [17], L. Horv́ath, et al. gave the following functional on the based of previous definition.

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

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

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

$$\hat{I}_f(\mathbf{r}, \mathbf{q}) := \sum_{s=1}^{n} q_s f\left(\frac{r_s}{q_s}\right). \tag{2.3}$$

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

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

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

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

$$\hat{I}_f(\mathbf{r},\mathbf{q}) = \sum_{s=1}^{n} q_s f\left(\frac{r_s}{q_s}\right) = A^{[1]}_{m,1} \geq A^{[1]}_{m,2} \geq \ldots \geq A^{[1]}_{m,m-1} \geq A^{[1]}_{m,m} \geq f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s}\right)\sum_{s=1}^{n} q_s. \tag{2.4}$$

where

$$A^{[1]}_{m,l} = \frac{(m-1)!}{(l-1)!}\sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l)\left(\sum_{j=1}^{l}\frac{q_{i_j}}{\alpha_{I_m,i_j}}\right) f\left(\frac{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}{\sum\limits_{j=1}^{l}\frac{q_{i_j}}{\alpha_{I_m,i_j}}}\right) \tag{2.5}$$

If f is concave function, then inequality signs in (2.4) are reversed.

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

$$\left(\sum_{s=1}^{n} r_s\right) f\left(\sum_{s=1}^{n}\frac{r_s}{\sum_{s=1}^{n} q_s}\right) \leq A^{[2]}_{m,m} \leq A^{[2]}_{m,m-1} \leq \ldots \leq A^{[2]}_{m,2} \leq A^{[2]}_{m,1} = \sum_{s=1}^{n} r_s f\left(\frac{r_s}{q_S}\right) = \hat{I}_{idf}(\mathbf{r},\mathbf{q}) \tag{2.6}$$

where

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

Proof. (*i*) 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

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

$$\left(\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}\right) f\left(\frac{\sum\limits_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{i=1}^{n} q_i}}{\alpha_{I_m,i_j}}\frac{r_{i_j}}{q_{i_j}}}{\sum\limits_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{i=1}^{n} q_i}}{\alpha_{I_m,i_j}}}\right) \leq \ldots \leq \sum_{s=1}^{n}\frac{q_s}{\sum_{i=1}^{n} q_s} f\left(\frac{r_s}{q_s}\right). \tag{2.7}$$

On multiplying $\sum_{s=1}^{n} q_s$, we have (2.4).

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

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

$$\left(\sum_{j=1}^{l}\frac{p_{i_j}}{\alpha_{I_m,i_j}}\right)\left(\frac{\sum\limits_{j=1}^{l}\frac{p_{i_j}}{\alpha_{I_m,i_j}}x_{i_j}}{\sum\limits_{j=1}^{l}\frac{p_{i_j}}{\alpha_{I_m,i_j}}}\right) f\left(\frac{\sum\limits_{j=1}^{l}\frac{p_{i_j}}{\alpha_{I_m,i_j}}x_{i_j}}{\sum\limits_{j=1}^{l}\frac{p_{i_j}}{\alpha_{I_m,i_j}}}\right) \leq \ldots \leq \sum_{s=1}^{n} p_s x_s f(x_s). \tag{2.8}$$

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

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

$$\left(\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}\right)\left(\frac{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}\frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}}\right) f\left(\frac{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}\frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n} q_s}}{\alpha_{I_m,i_j}}}\right) \leq \ldots \leq \sum_{s=1}^{n}\frac{q_s}{\sum_{s=1}^{n} q_s}\frac{r_s}{q_s} f\left(\frac{r_s}{q_S}\right). \tag{2.9}$$

On multiplying $\sum_{s=1}^{n} q_s$, we get (2.6). □

3. Inequalities for Shannon Entropy

Definition 3.1 (see [17]). *The Shannon entropy of positive probability distribution* $\mathbf{r} = (r_1, \ldots, r_n)$ *is defined by*

$$S := -\sum_{s=1}^{n} r_s \log(r_s). \tag{3.1}$$

Corollary 3.1. *Assume* (H_1).

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

$$S \leq A^{[3]}_{m,m} \leq A^{[3]}_{m,m-1} \leq \ldots \leq A^{[3]}_{m,2} \leq A^{[3]}_{m,1} = \log\left(\frac{n}{\sum_{s=1}^{n} q_s}\right) \sum_{s=1}^{n} q_s, \tag{3.2}$$

where

$$A^{[3]}_{m,l} = -\frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1, \ldots, i_l) \left(\sum_{j=1}^{l} \frac{q_{i_j}}{\alpha_{I_m,i_j}}\right) \log\left(\sum_{j=1}^{l} \frac{q_{i_j}}{\alpha_{I_m,i_j}}\right). \tag{3.3}$$

If the base of log *is between* 0 *and* 1, *then inequality signs in* (3.2) *are reversed.*

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

$$S \leq A^{[4]}_{m,m} \leq A^{[4]}_{m,m-1} \leq \ldots \leq A^{[4]}_{m,2} \leq A^{[4]}_{m,1} = \log(n), \tag{3.4}$$

where

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

Proof. (i) Using $f := \log$ and $\mathbf{r} = (1, \ldots, 1)$ in Theorem 2.1 (i), we get (3.2).

(ii) It is the special case of (i). □

Definition 3.2 (see [17]). *The Kullback-Leibler divergence between the positive probability distribution* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *is defined by*

$$D(\mathbf{r}, \mathbf{q}) := \sum_{s=1}^{n} r_i \log\left(\frac{r_i}{q_i}\right). \tag{3.5}$$

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*

$$\sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) \leq A^{[5]}_{m,m} \leq A^{[5]}_{m,m-1} \leq \ldots \leq A^{[5]}_{m,2} \leq A^{[5]}_{m,1} = \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right) = D(\mathbf{r}, \mathbf{q}), \tag{3.6}$$

where

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

If the base of log *is between 0 and 1, then inequality in* (3.6) *is reversed.*

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

$$D(\boldsymbol{r},\boldsymbol{q}) = A^{[6]}_{m,1} \geq A^{[6]}_{m,2} \geq \ldots \geq A^{[6]}_{m,m-1} \geq A^{[6]}_{m,m} \geq 0, \tag{3.7}$$

where

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

If the base of log *is between* 0 *and* 1*, then inequality signs in* (3.7) *are reversed.*

Proof. (*i*) On taking $f := \log$ in Theorem 2.1 (*ii*), we get (3.6).

(*ii*) Since **r** and **q** are positive probability distributions therefore $\sum_{s=1}^{n} r_s = \sum_{s}^{n} q_s = 1$, so the smallest term in (3.6) is given as

$$\sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) = 0. \tag{3.8}$$

Hence for positive probability distribution **r** and **q** the (3.6) will become (3.7).

□

4. Inequalities for Rényi Divergence and Entropy

The Rényi divergence and entropy come from [25].

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

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

$$D_\lambda(\boldsymbol{r},\boldsymbol{q}) := \frac{1}{\lambda-1} \log\left(\sum_{i=1}^{n} q_i \left(\frac{r_i}{q_i}\right)^{\lambda}\right). \tag{4.1}$$

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

$$H_\lambda(\boldsymbol{r}) := \frac{1}{1-\lambda} \log\left(\sum_{i=1}^{n} r_i^{\lambda}\right). \tag{4.2}$$

The Rényi divergence and the Rényi entropy can also be extended to non-negative probability distributions. If $\lambda \to 1$ in (4.1), we have the Kullback-Leibler divergence, and if $\lambda \to 1$ in (4.2), then we have the Shannon entropy. In the next two results, inequalities can be found for the Rényi divergence.

Theorem 4.1. *Assume* (H_1), *let* $\boldsymbol{r} = (r_1, \ldots, r_n)$ *and* $\boldsymbol{q} = (q_1, \ldots, q_n)$ *are 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*

$$D_\lambda(\boldsymbol{r}, \boldsymbol{q}) \leq A^{[7]}_{m,m} \leq A^{[7]}_{m,m-1} \leq \ldots \leq A^{[7]}_{m,2} \leq A^{[7]}_{m,1} = D_\mu(\boldsymbol{r}, \boldsymbol{q}), \tag{4.3}$$

where

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

The reverse inequalities hold in (4.3) *if the base of* log *is between 0 and 1.*

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

$$D_1(\boldsymbol{r}, \boldsymbol{q}) = D(\boldsymbol{r}, \boldsymbol{q}) = \sum_{s=1}^{n} r_s \log \left(\frac{r_s}{q_s} \right) \leq A^{[8]}_{m,m} \leq A^{[8]}_{m,m-1} \leq \ldots \leq A^{[8]}_{m,2} \leq A^{[8]}_{m,1} = D_\mu(\boldsymbol{r}, \boldsymbol{q}), \tag{4.4}$$

where

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

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*

$$D_\lambda(\boldsymbol{r}, \boldsymbol{q}) \leq A^{[9]}_{m,m} \leq A^{[9]}_{m,m-1} \leq \ldots \leq A^{[9]}_{m,2} \leq A^{[9]}_{m,1} = D_1(\boldsymbol{r}, \boldsymbol{q}), \tag{4.5}$$

where

$$A^{[9]}_{m,l} = \frac{1}{\lambda - 1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1, \ldots, i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \log \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \left(\frac{r_{i_j}}{q_{i_j}} \right)^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right). \tag{4.6}$$

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

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

we have

$$\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda}\right)^{\frac{\mu-1}{\lambda-1}} = \left(\sum_{s=1}^{n} r_s \left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}} \leq$$

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

$$\leq \ldots \leq \sum_{s=1}^{n} r_s \left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}}, \tag{4.7}$$

if either $0 \leq \lambda < 1 < \beta$ or $1 < \lambda \leq \mu$, and the reverse inequality in (4.7) holds if $0 \leq \lambda \leq \beta < 1$. By raising to power $\frac{1}{\mu-1}$, we have from all

$$\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda}\right)^{\frac{1}{\lambda-1}} \leq$$

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

$$\leq \ldots \leq \left(\sum_{s=1}^{n} r_s \left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}}\right)^{\frac{1}{\mu-1}} = \left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\mu}\right)^{\frac{1}{\mu-1}}. \tag{4.8}$$

Since log is increasing if the base of log is greater than 1, it now follows (4.3). If the base of log is between 0 and 1, then log is decreasing and therefore inequality in (4.3) are reversed. If $\lambda = 1$ and $\beta = 1$, we have (ii) and (iii) respectively by taking limit, when λ goes to 1. □

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

$$\frac{1}{\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda}} \sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda} \log\left(\frac{r_s}{q_s}\right) = A_{m,1}^{[10]} \leq A_{m,2}^{[10]} \leq \ldots \leq A_{m,m-1}^{[10]} \leq A_{m,m}^{[10]} \leq D_\lambda(\boldsymbol{r},\boldsymbol{q}) \leq A_{m,m}^{[11]}$$
$$\leq A_{m,m}^{[11]} \leq \ldots \leq A_{m,2}^{[11]} \leq A_{m,1}^{[11]} = D_1(\boldsymbol{r},\boldsymbol{q}) \tag{4.9}$$

where

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

and

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

The inequalities in (4.9) 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 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

$$D_\lambda(\mathbf{r},\mathbf{q}) = \frac{1}{\lambda-1} \log\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda}\right) = \frac{1}{\lambda-1} \log\left(\sum_{s=1}^{n} r_s \left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)$$
$$\leq \ldots \leq \frac{1}{\lambda-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\right) \log\left(\frac{\sum\limits_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum\limits_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)$$
$$\leq \ldots \leq \frac{1}{\lambda-1} \sum_{s=1}^{n} r_s \log\left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right) = \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right) = D_1(\mathbf{r},\mathbf{q}) \tag{4.10}$$

and this give 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

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

$$= \frac{1}{\lambda-1\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)}\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)\log\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)$$

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

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

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

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

$$\geq \ldots \geq \frac{1}{\lambda-1}\sum_{s=1}^{n} r_s\left(\frac{r_s}{q_s}\right)^{\lambda-1}\log\left(\frac{r_s}{q_s}\right)^{\lambda-1}\frac{1}{\sum_{s=1}^{n} r_s\left(\frac{r_s}{q_s}\right)^{\lambda-1}}$$

$$= \frac{1}{\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}}\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\log\left(\frac{r_s}{q_s}\right) \tag{4.11}$$

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

By using the Theorem 4.1, Theorem 4.2 and Definition 4.1, some inequalities of Rényi entropy are obtained. Let $\frac{\mathbf{1}}{\mathbf{n}} = (\frac{1}{n},\ldots,\frac{1}{n})$ be a discrete probability distribution.

Corollary 4.3. *Assume* (H_1)*, let* $\boldsymbol{r} = (r_1,\ldots,r_n)$ *and* $\boldsymbol{q} = (q_1,\ldots,q_n)$ *are positive probability distributions.*
(i) If $0 \leq \lambda \leq \mu$, $\lambda,\mu \neq 1$*, and the base of* log *is greater than 1, then*

$$H_\lambda(\boldsymbol{r}) = \log(n) - D_\lambda\left(\boldsymbol{r},\frac{1}{\boldsymbol{n}}\right) \geq A_{m,m}^{[12]} \geq A_{m,m}^{[12]} \geq \ldots A_{m,2}^{[12]} \geq A_{m,1}^{[12]} = H_\mu(\boldsymbol{r}), \tag{4.12}$$

where

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

The reverse inequalities holds in (4.12) if the base of log *is between 0 and 1.*

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

$$S = -\sum_{s=1}^{n} p_i \log(p_i) \geq A_{m,m}^{[13]} \geq A_{m,m-1}^{[13]} \geq \ldots \geq A_{m,2}^{[13]} \geq A_{m,1}^{[13]} = H_\mu(\boldsymbol{r}) \tag{4.13}$$

where

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

the base of exp *is same as the base of* log*. The inequalities in (4.13) 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*

$$H_\lambda(\boldsymbol{r}) \geq A_{m,m}^{[14]} \geq A_{m,m-1}^{[14]} \geq \ldots \geq A_{m,2}^{[14]} \leq A_{m,1}^{[14]} = S, \tag{4.14}$$

where

$$A_{m,m}^{[14]} = \frac{1}{1-\lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \log \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}^\lambda}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right). \tag{4.15}$$

The inequalities in (4.14) are reversed if the base of log *is between 0 and 1.*

Proof. (i) Suppose $\mathbf{q} = \frac{\mathbf{1}}{\mathbf{n}}$ then from (4.1), we have

$$D_\lambda(\mathbf{r}, \mathbf{q}) = \frac{1}{\lambda - 1} \log \left(\sum_{s=1}^{n} n^{\lambda-1} r_s^\lambda \right) = \log(n) + \frac{1}{\lambda-1} \log \left(\sum_{s=1}^{n} r_s^\lambda \right), \tag{4.16}$$

therefore we have

$$H_\lambda(\mathbf{r}) = \log(n) - D_\lambda(\mathbf{r}, \frac{\mathbf{1}}{\mathbf{n}}). \tag{4.17}$$

Now using Theorem 4.1 (i) and (4.17), we get

$$H_\lambda(\mathbf{r}) = \log(n) - D_\lambda \left(\mathbf{r}, \frac{\mathbf{1}}{\mathbf{n}} \right) \geq \ldots \geq \log(n) - \frac{1}{\mu-1} \log \left(n^{\mu-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \times \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \right) \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}^\lambda}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \right) \geq \ldots \geq \log(n) - D_\mu(\mathbf{r}, \mathbf{q}) = H_\mu(\mathbf{r}), \tag{4.18}$$

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

Corollary 4.4. *Assume* (H_1) *and let* $\boldsymbol{r} = (r_1, \ldots, r_n)$ *and* $\boldsymbol{q} = (q_1, \ldots, q_n)$ *are positive probability distributions.*

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

$$\begin{aligned} -\frac{1}{\sum_{s=1}^{n} r_s^{\lambda}} \sum_{s=1}^{n} r_s^{\lambda} \log(r_s) = A_{m,1}^{[15]} \ge A_{m,2}^{[15]} \ge \ldots \ge A_{m,m-1}^{[15]} \ge A_{m,m}^{[15]} \ge H_{\lambda}(\boldsymbol{r}) \ge A_{m,m}^{[16]} \\ \ge A_{m,m-1}^{[16]} \ge \ldots A_{m,2}^{[16]} \ge A_{m,1}^{[16]} = H(\boldsymbol{r}), \end{aligned} \tag{4.19}$$

where

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

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 (4.19) *are reversed if either* $0 \le \lambda < 1$ *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 the Corollary 4.3 by using Theorem 4.2. □

5. Inequalities by Using Zipf-Mandelbrot Law

In probability theory and statistics, the Zipf-Mandelbrot law is a distribution. It is a power law distribution on ranked data, named after the linguist G. K. Zipf who suggest a simpler distribution called Zipf's law. The Zipf's law is defined as follow (see [28]).

Definition 5.1. *Let* N *be a number of elements,* s *be their rank and* t *be the value of exponent characterizing the distribution. Zipf's law then predicts that out of a population of* N *elements, the normalized frequency of element of rank* s*,* $f(s, N, t)$ *is*

$$f(s, N, t) = \frac{\frac{1}{s^t}}{\sum_{j=1}^{N} \frac{1}{j^t}}. \tag{5.1}$$

The Zipf-Mandelbrot law is defined as follows (see [21]).

Definition 5.2. *Zipf-Mandelbrot law is a discrete probability distribution depending on three parameters* $N \in \{1, 2, \ldots, \}, q \in [0, \infty)$ *and* $t > 0$*, and is defined by*

$$f(s; N, q, t) := \frac{1}{(s+q)^t H_{N,q,t}}, \qquad s = 1, \ldots, N, \tag{5.2}$$

where

$$H_{N,q,t} = \sum_{j=1}^{N} \frac{1}{(j+q)^t}. \tag{5.3}$$

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*

$$f(s;q,t) = \frac{1}{(s+q)^t H_{q,t}}, \tag{5.4}$$

where

$$H_{q,t} = \sum_{j=1}^{\infty} \frac{1}{(j+q)^t}. \tag{5.5}$$

For $q = 0$*, the Zipf-Mandelbrot law* (5.2) *becomes Zipf's law* (5.1).

Conclusion 5.1. *Assume* (H_1)*, let* $\boldsymbol{r}$ *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*

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

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

$$\geq \ldots \geq \frac{t}{H_{N,q,t}} \sum_{s=1}^{N} \frac{\log(s+q)}{(s+q)^t} + \log(H_{N,q,t}) = S. \tag{5.6}$$

The inequalities in (5.6) *are reversed if the base of* log *is between 0 and 1.*

Conclusion 5.2. *Assume* (H_1)*, let* $\boldsymbol{r}_1$ *and* $\boldsymbol{r}_2$ *be the Zipf-Mandelbort law with parameters* $N \in \{1,2,\ldots\}$*,* $q_1, q_2 \in [0,\infty)$ *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*

$$\bar{D}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \sum_{s=1}^{n} \frac{1}{(s+q_1)^{t_1} H_{N,q_1,t_1}} \log\left(\frac{(s+q_2)^{t_2} H_{N,q_2,t_2}}{(s+q_1)^{t_1} H_{N,q_2,t_1}} \right) \geq \ldots \geq \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l)$$

$$\left(\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}} \right) \left(\frac{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}} \right) \log\left(\frac{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}} \right) \geq \ldots \geq 0. \tag{5.7}$$

The inequalities in (5.7) *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 method of Lagrange multiplier under some equations constraints and get the Zipf-Mandelbrot law.

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

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

is Zipf-Madelbrot law.

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

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

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

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

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

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

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

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

□

Remark 6.2. *Observe that the Zipf-Mandelbrot law and Shannon Entroy can be bounded from above (see [22]).*

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

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

Theorem 6.3. *If $J = \{1, \ldots, N\}$, then probability distribution that maximize 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 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 $J = \{1, \ldots, N\}$, 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 $s = 1, \ldots, N$, 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 [22]).*

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

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

Under the assumption of Theorem 2.1 (i), define the non-negative functionals as follows.

$$\Theta_3(f) = \mathcal{A}_{m,r}^{[1]} - f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s} \right) \sum_{s=1}^{n} q_s, \qquad r = 1, \ldots, m, \tag{6.1}$$

$$\Theta_4(f) = \mathcal{A}_{m,r}^{[1]} - \mathcal{A}_{m,k}^{[1]}, \qquad 1 \leq r < k \leq m. \tag{6.2}$$

Under the assumption of Theorem 2.1 (ii), define the non-negative functionals as follows.

$$\Theta_5(f) = \mathcal{A}^{[2]}_{m,r} - \left(\sum_{s=1}^{n} r_s\right) f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s}\right), \qquad r = 1, \ldots, m, \tag{6.3}$$

$$\Theta_6(f) = \mathcal{A}^{[2]}_{m,r} - \mathcal{A}^{[2]}_{m,k}, \qquad 1 \leq r < k \leq m. \tag{6.4}$$

Under the assumption of Corollary 3.1 (i), define the following non-negative functionals.

$$\Theta_7(f) = A^{[3]}_{m,r} + \sum_{i=1}^{n} q_i \log(q_i), \quad r = 1, \ldots, n \tag{6.5}$$

$$\Theta_8(f) = A^{[3]}_{m,r} - A^{[3]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.6}$$

Under the assumption of Corollary 3.1 (ii), define the following non-negative functionals give as.

$$\Theta_9(f) = A^{[4]}_{m,r} - S, \quad r = 1, \ldots, m \tag{6.7}$$

$$\Theta_{10}(f) = A^{[4]}_{m,r} - A^{[4]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.8}$$

Under the assumption of Corollary 3.2 (i), let us define the non-negative functionals as follows.

$$\Theta_{11}(f) = A^{[5]}_{m,r} - \sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \log \frac{r_n}{\sum_{s=1}^{n} q_s}\right), \quad r = 1, \ldots, m \tag{6.9}$$

$$\Theta_{12}(f) = A^{[5]}_{m,r} - A^{[5]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.10}$$

Under the assumption of Corollary 3.2 (ii), define the non-negative functionals as follows.

$$\Theta_{13}(f) = A^{[6]}_{m,r} - A^{[6]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.11}$$

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

$$\Theta_{14}(f) = A^{[7]}_{m,r} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m \tag{6.12}$$

$$\Theta_{15}(f) = A^{[7]}_{m,r} - A^{[7]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.13}$$

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

$$\Theta_{16}(f) = A^{[8]}_{m,r} - D_1(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m \tag{6.14}$$

$$\Theta_{17}(f) = A^{[8]}_{m,r} - A^{[8]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.15}$$

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

$$\Theta_{18}(f) = A^{[9]}_{m,r} - D_\lambda(\mathbf{r}, \mathbf{q}), \quad r = 1, \ldots, m \tag{6.16}$$

$$\Theta_{19}(f) = A^{[9]}_{m,r} - A^{[9]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{6.17}$$

Under the assumption of Theorem 4.2 consider the following non-negative functionals.

$$\Theta_{20}(f) = D_\lambda(\mathbf{r},\mathbf{q}) - A^{[10]}_{m,r}, \quad r=1,\dots,m \tag{6.18}$$

$$\Theta_{21}(f) = A^{[10]}_{m,k} - A^{[10]}_{m,r}, \;\; 1 \le r < k \le m. \tag{6.19}$$

$$\Theta_{22}(f) = A^{[11]}_{m,r} - D_\lambda(\mathbf{r},\mathbf{q}), \;\; r=1,\dots,m \tag{6.20}$$

$$\Theta_{23}(f) = A^{[11]}_{m,r} - A^{[11]}_{m,r}, \;\; 1 \le r < k \le m. \tag{6.21}$$

$$\Theta_{24}(f) = A^{[11]}_{m,r} - A^{[10]}_{m,k}, \;\; r=1,\dots,m, \;\; k=1,\dots,m \tag{6.22}$$

Under the assumption of Corollary 4.3 (i), consider the following non-negative functionals.

$$\Theta_{25}(f) = H_\lambda(\mathbf{r}) - A^{[12]}_{m,r}, \;\; r=1,\dots,m \tag{6.23}$$

$$\Theta_{26}(f) = A^{[12]}_{m,k} - A^{[12]}_{m,r}, \;\; 1 \le r < k \le m. \tag{6.24}$$

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

$$\Theta_{27}(f) = S - A^{[13]}_{m,r}, \;\; r=1,\dots,m \tag{6.25}$$

$$\Theta_{28}(f) = A^{[13]}_{m,k} - A^{[13]}_{m,r}, \;\; 1 \le r < k \le m. \tag{6.26}$$

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

$$\Theta_{29}(f) = H_\lambda(\mathbf{r}) - A^{[14]}_{m,r}, \;\; r=1,\dots,m \tag{6.27}$$

$$\Theta_{30}(f) = A^{[14]}_{m,k} - A^{[14]}_{m,r}, \;\; 1 \le r < k \le m. \tag{6.28}$$

Under the assumption of Corollary 4.4, defined the following functionals.

$$\Theta_{31} = A^{[15]}_{m,r} - H_\lambda(\mathbf{r}), \;\; r=1,\dots,m \tag{6.29}$$

$$\Theta_{32} = A^{[15]}_{m,r} - A^{[15]}_{m,k}, \;\; 1 \le r < k \le m. \tag{6.30}$$

$$\Theta_{33} = H_\lambda(\mathbf{r}) - A^{[16]}_{m,r}, \;\; r=1,\dots,m \tag{6.31}$$

$$\Theta_{34} = A^{[16]}_{m,k} - A^{[16]}_{m,r}, \;\; 1 \le r < k \le m. \tag{6.32}$$

$$\Theta_{35} = A^{[15]}_{m,r} - A^{[16]}_{m,k}, \;\; r=1,\dots,m, \;\; k=1,\dots,m. \tag{6.33}$$

7. Generalization of refinement of Jensen's, Rényi and Shannon type inequalities via Taylor one point and Taylor two points interpolations

In [5], the following functions are consider to generalized the Popoviciu's inequality, defined as

$$(u-v)_+ = \begin{cases} (u-v), & v \le u; \\ 0, & v > u, \end{cases}$$

and the well known Taylor formula is as follows.

Let m be a positive integer and $f : [\alpha_1, \alpha_2] \to \mathbb{R}$ be such that $f^{(m-1)}$ is absolutely continuous, then for all $u \in [\alpha_1, \alpha_2]$ the Taylor's formula at point $c \in [\alpha_1, \alpha_2]$ is

$$f(u) = T_{m-1}(f; c; u) + R_{m-1}(f; c; u), \tag{7.1}$$

where

$$T_{m-1}(f; c; u) = \sum_{l=0}^{m-1} \frac{f^{(l)}(c)}{l!}(u-c)^l,$$

and the remainder is given by

$$R_{m-1}(f; c; u) = \frac{1}{(m-1)!} \int_c^u f^{(m)}(t)(u-t)^{m-1} dt.$$

The Taylor's formula at point α_1 and α_2 is given by:

$$f(u) = \sum_{l=0}^{m-1} \frac{f^{(l)}(\alpha_1)}{l!}(u-\alpha_1)^l + \frac{1}{(m-1)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(t)\left((u-t)_+^{m-1}\right) dt. \tag{7.2}$$

$$f(u) = \sum_{l=0}^{m-1} \frac{(-1)^l f^{(l)}(\alpha_2)}{l!}(\alpha_2-u)^l + \frac{(-1)^{m-1}}{(m-1)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(t)\left((t-u)_+^{m-1}\right) dt. \tag{7.3}$$

We construct some new identities with the help of Taylor polynomial (7.1).

Theorem 7.1. *Assume (H_1), let $f : [\alpha_1, \alpha_2] \to \mathbb{R}$ be a function where $[\alpha_1, \alpha_2] \subset \mathbb{R}$ be an interval. Also let $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \ldots, p_n$ are positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. Then we have the following identities:*

(i)

$$\Theta_i(f) = \sum_{l=2}^{m-1} \frac{f^{(l)}(\alpha_1)}{l!}\Theta_i\left((u-\alpha_1)^l\right) + \frac{1}{(m-1)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(t)\Theta_i\left((u-t)_+^{m-1}\right) dt, \quad i = 1, 2, \ldots, 35. \tag{7.4}$$

(ii)

$$\Theta_i(f) = \sum_{l=2}^{m-1} \frac{(-1)^l f^{(l)}(\alpha_2)}{l!}\Theta_i\left((\alpha_2-u)^l\right) + \frac{(-1)^{m-1}}{(m-1)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(t)\Theta_i\left((t-u)_+^{m-1}\right) dt, \quad i = 1, 2, \ldots, 35. \tag{7.5}$$

Proof. Using (7.2) and (7.3) in (1.3), we get the required result. □

Theorem 7.2. *Assume (H_1), let $f : [\alpha_1, \alpha_2] \to \mathbb{R}$ be a function where $[\alpha_1, \alpha_2] \subset \mathbb{R}$ be an interval. Also let $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \ldots, p_n$ are positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. Let f is m-convex function such that $f^{(m-1)}$ is absolutely continuous. Then we have the following results:*

(i) If

$$\Theta_i\left((u-t)_+^{m-1}\right) \geq 0 \qquad t \in [\alpha_1, \alpha_2], \;\; i = 1, 2, \ldots, 35,$$

then

$$\Theta_i(f(u)) \geq \sum_{l=2}^{m-1} \frac{f^{(l)}(\alpha_1)}{l!}\Theta_i\left((u-\alpha_1)^l\right), \quad i = 1, 2, \ldots, 35. \tag{7.6}$$

(ii) If

$$(-1)^{m-1}\Theta_i\left((t-u)_+^{m-1}\right) \leq 0 \qquad t \in [\alpha_1, \alpha_2], \quad i = 1, 2, \ldots, 35,$$

then

$$\Theta_i(f(u)) \geq \sum_{l=2}^{m-1} \frac{(-1)^l f^{(l)}(\alpha_2)}{l!}\Theta_i\left((\alpha_2-u)^l\right), \quad i = 1, 2, \ldots, 35. \tag{7.7}$$

Proof. Since $f^{(m-1)}$ is absolutely continuous on $[\alpha_1, \alpha_2]$, $f^{(m)}$ exists almost everywhere. As f is m-convex therefore $f^{(m)}(u) \geq 0$ for all $u \in [\alpha_1, \alpha_2]$. Hence using Theorem 7.1 we obtain (7.6) and (7.7). □

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

(i) If f is m-convex, then (7.6) holds. Also if $f^{(l)}(\alpha_1) \geq 0$ for $l = 2, \ldots, m-1$, then the right hand side of (7.6) will be non-negative.
(ii) If m is even and f is m-convex, then (7.7) holds. Also if $f^{(l)}(\alpha_1) \leq 0$ for $l = 2, \ldots, m-1$ and $f^{(l)} \geq 0$ for $l = 3, \ldots, m-1$, then right hand side of (7.7) will be non-negative.
(iii) If m is odd and f is m-convex function then (7.7) is valid. Also if $f^{(l)}(\alpha_2) \geq 0$ for $l = 2, \ldots, m-1$ and $f^{(l)}(\alpha_2) \leq 0$ for $l = 2, \ldots, m-2$, then right hand side of (7.7) will be non positive.

In [7, p.20] the Green function $G : [\alpha_1, \alpha_2] \times [\alpha_1, \alpha_2] \to \mathbb{R}$ is defined as

$$G(u, v) = \begin{cases} \frac{(u-\alpha_2)(v-\alpha_1)}{\alpha_2-\alpha_1}, & \alpha_1 \leq v \leq u; \\ \frac{(v-\alpha_2)(u-\alpha_1)}{\alpha_2-\alpha_1}, & u \leq v \leq \alpha_2. \end{cases} \tag{7.8}$$

The function G is convex and continuous with respect to v, since G is symmetric therefore it is also convex and continuous with respect to variable u.
Let $\psi \in C^2([\alpha_1, \alpha_2])$, then

$$\psi(t) = \frac{\alpha_2 - t}{\alpha_2 - \alpha_1}\psi(\alpha_1) + \frac{t - \alpha_1}{\alpha_2 - \alpha_1}\psi(\alpha_2) + \int_{\alpha_1}^{\alpha_2} G(t, v)\,\psi''(v)dv. \tag{7.9}$$

Theorem 7.4. *Assume (H_1), let $f : [\alpha_1, \alpha_2] \to \mathbb{R}$ be a function where $[\alpha_1, \alpha_2] \subset \mathbb{R}$ be an interval. Also let $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ and $p_1, \ldots, p_n$ are positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. Then we have the* *following results:*

(*i*) *For* $i = 1, 2, \ldots, 35$,

$$\Theta_i(f) = \int_{\alpha_1}^{\alpha_2} \Theta_i(G(t,v)) \left(\sum_{l=1}^{n-1} \frac{f^{(l)}(\alpha_1)(v-\alpha_1)^{l-2}}{(l-2)!} \right) dv + \frac{1}{(n-3)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(s) \left(\int_{\alpha_1}^{\alpha_2} \Theta_i(G(t,v))(v-s)^{n-3} dv \right) ds. \quad (7.10)$$

(*ii*) *For* $i = 1, 2, \ldots, 35$,

$$\Theta_i(f) = \int_{\alpha_1}^{\alpha_2} \Theta_i(G(t,v)) \left(\sum_{l=1}^{n-1} \frac{f^{l}(\alpha_2)(v-\alpha_2)^{l-2}}{(l-2)!} \right) dv - \frac{1}{(n-3)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(s) \left(\int_{\alpha_1}^{\alpha_2} \Theta_i(G(t,v))(v-s)^{n-3} dv \right) ds \quad (7.11)$$

Proof. Using (7.9) in $\Theta_i, i = 1, 2, \ldots, 35$, we get

$$\Theta_i(f) = \int_{\alpha_1}^{\alpha_2} \Theta_i\left(G(t,v)\right) f''(v) dv. \quad (7.12)$$

Differentiate (7.2) twice, we get

$$f''(v) = \sum_{l=2}^{n-1} \frac{f^{(l)}(\alpha_1)}{(l-2)!} (v-\alpha_1)^{l-2} + \frac{1}{(m-3)!} \int_{\alpha_1}^{\alpha_2} f^{(m)}(v-u)^{m-3} du. \quad (7.13)$$

Using (7.13) in (7.12) and using Fubini's theorem, we get (7.10). Similarly use second derivative of (7.3) in (7.12) and apply Fubini's theorem, we get (7.11). □

Now we obtain generalization of refinement of Jensen's inequality for n-convex function.

Theorem 7.5. *Assume* (H_1), *let* $f : [\alpha_1, \alpha_2] \to \mathbb{R}$ *be a function where* $[\alpha_1, \alpha_2] \subset \mathbb{R}$ *be an interval. Also let* $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ *and* $p_1, \ldots, p_n$ *are positive real numbers such that* $\sum_{i=1}^{n} p_i = 1$. *Let* f *is* m-*convex function such that* $f^{(m-1)}$ *is absolutely continuous. Then we have the following results:*

(*i*) *If*

$$\int_{u}^{\alpha_2} \Theta_i\left(G(t,v)\right)(v-u)^{n-3} dv \geq 0 \quad u \in [\alpha_1, \alpha_2], \quad i = 1, 2, \ldots, 35, \quad (7.14)$$

then

$$\Theta_i(f) \geq \int_{\alpha_1}^{\alpha_2} \Theta_i\left(G(t,v)\right) \left(\sum_{l=2}^{n-2} \frac{f^{(l)}(\alpha_1)(v-\alpha_1)^{l-2}}{(l-2)!} \right) dv, \quad i = 1, 2, \ldots, 35, \quad (7.15)$$

and if

$$\int_{\alpha_1}^{u} \Theta_i\left(G(t,v)\right)(v-u)^{n-3} dv \leq 0 \quad u \in [\alpha_1, \alpha_2], \quad i = 1, 2, \ldots, 35, \quad (7.16)$$

then

$$\Theta_i(f) \geq \int_{\alpha_1}^{\alpha_2} \Theta i\left(G(t,v)\right) \left(\sum_{l=2}^{n-2} \frac{f^{(l)}(\alpha_2)(v-\alpha_2)^{l-2}}{(l-2)!} \right) dv \quad i = 1, 2, \ldots, 35. \quad (7.17)$$

Proof. Similar to the proof of Theorem 7.2. □

Corollary 7.6. *Assume* (H_1)*, let* $f : [\alpha_1, \alpha_2] \to \mathbb{R}$ *be a function where* $[\alpha_1, \alpha_2] \subset \mathbb{R}$ *be an interval. Also let* $x_1, \ldots, x_n \in [\alpha_1, \alpha_2]$ *and* $p_1, \ldots, p_n$ *are positive real numbers such that* $\sum_{i=1}^{n} p_i = 1$*. Then the following results are valid.*

(*i*) *If* f *is* m*-convex, then* (7.15) *holds. Also if*

$$\sum_{l=2}^{n-1} \frac{f^{(l)}(\alpha_1)(v-\alpha_1)^{l-2}}{(l-2)!} \geq 0, \tag{7.18}$$

then

$$\Theta_i(f) \geq 0, \qquad i = 1, 2, \ldots, 35. \tag{7.19}$$

(*ii*) *If* m *is even and* f *is* m*-convex, then* (7.17) *holds. Also if*

$$\sum_{l=2}^{n-1} \frac{f^{(l)}(\alpha_2)(v-\alpha_2)^{l-2}}{(l-2)!} \geq 0, \tag{7.20}$$

then (7.19) *holds.*

Remark 7.7. *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 Gr̆uss and Ostrowski type inequalities can be constructed as given in Section 3 of [5]. Also we can construct the non-negative functionals from inequalities* (7.6)*,* (7.7)*,* (7.15) *and* (7.17) *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].*

Funding

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

Competing interests

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

Authors contribution

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

Acknowledgements

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

References

[1] Anderson, G., & Ge, Y. The size distribution of Chinese cities. Reg. Sci. Urban Econ., 35 (2005), 756-776.

[2] Auerbach, F. (1913). Das Gesetz der Bevlkerungskonzentration. Petermanns Geographische Mitteilungen, 59 (2005), 74-76.

[3] Black, D., & Henderson, V. Urban evolution in the USA. J. Econ. Geogr., 3 (2003), 343-372.

[4] Bosker, M., Brakman, S., Garretsen, H., & Schramm, M. A century of shocks: the evolution of the German city size distribution 19251999. Reg. Sci. Urban Econ., 38 (2008), 330-347.

[5] Butt, S. I., Khan, K. A., & Pečarić, J. Generaliztion of Popoviciu inequality for higher order convex function via Tayor's polynomial, Acta Univ. Apulensis Math. Inform., 42 (2015), 181-200.

[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 (2017), 293-303.

[7] Butt, S. I., & Pečarić, J. Popoviciu's Inequality For N-convex Functions. Lap Lambert Academic Publishing, (2016).

[8] Butt, S. I., & Pečarić, J. Weighted Popoviciu type inequalities via generalized Montgomery identities. Rad Hazu. Mat. Znan., 19 (2015), 69-89.

[9] Butt, S. I., Khan, K. A., & Pečarić, J. Popoviciu type inequalities via Hermite's polynomial. Math. Inequal. Appl., 19 (2016), 1309-1318.

[10] Horváth, L. A method to refine the discrete Jensen's inequality for convex and mid-convex functions. Math. Computer Model., 54 (2011), 2451-2459.

[11] Horváth, L., Khan, K. A., & Pečarić, J. Combinatorial Improvements of Jensens Inequality / Classical and New Refinements of Jensens Inequality with Applications, Monographs in inequalities 8, Element, Zagreb. (2014).

[12] Horváth, L., Khan, K. A., & Pečarić, J. Refinement of Jensen's inequality for operator convex functions. Adv. Inequal. Appl., 2014 (2014), Art. ID 26.

[13] Horváth, L., Pečarić, J. A refinement of discrete Jensen's inequality, Math. Inequal. Appl. 14 (2011), 777-791.

[14] Ioannides, Y. M., & Overman, H. G. Zipf's law for cities: an empirical examination. Reg. Sci. Urban Econ., 33 (2003), 127-137.

[15] Csiszár, I. Information measures: a critical survey. In: Tans. 7th Prague Conf. on Info. Th., Statist. Decis. Funct., Rand. Proc. 8th Eur. Meeting Stat., Vol. B (1978), 73-86.

[16] Csiszár. I. . Information-type measures of difference of probability distributions and indirect observations. Stud. Sci. Math. Hungar. 2 (1967), 299-318.

[17] Horváth, L., Pecaric, D. & Pečarić, J. Estimations of f-and Rényi divergences by using a cyclic refinement of the Jensen's inequality. Bull. Malaysian Math. Sci. Soc., 42 (2019). 933-946 .

[18] Khan, K. A., Niaz, T., Pečarić, Đ., Pečarić, J. Refinement of Jensen's Inequality and Estimation of f- and Renyi Divergence via Montgomery identity. J. Inequal. Appl., 2018 (2018), Art. ID 318.

[19] Kullback, S. Information theory and statistics. Courier Corporation.

[20] Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. Anna. Math. Stat., 22 (1997), 79-86. Math. Dokl. **4**(1963), 121-124.

[21] Lovricevic, N., Pecaric, D. & Pecaric, J. ZipfMandelbrot law, f-divergences and the Jensen-type interpolating inequalities. J. Inequal. Appl., 2018 (2018), Art. ID 36.

[22] 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).

[23] Niaz, T., Khan, K. A., & Pečarić, J. On generalization of refinement of Jensen's inequality using Fink's identity and Abel-Gontscharoff Green function. J. Inequal. Appl., 2017 (2017), Art. ID 254.

[24] Pečarić, J., Proschan, F., & Tong, Y. L. Convex functions, Partial Orderings and Statistical Applications, Academic Press, New York. (1992).

[25] Rényi, A. On measure of information and entropy. In: Proceeding of the Fourth Berkely Symposium on Mathematics, Statistics and Probability, pp. 547-561. (1960).

[26] Rosen, K. T., & Resnick, M. The size distribution of cities: an examination of the Pareto law and primacy. J. Urban Econ., 8 (1980), 165-186.

[27] Soo, K. T. Zipf's Law for cities: a cross-country investigation. Reg. Sci. Urban Econ., 35 (2005), 239-263.

[28] Zipf, G. K. Human behaviour and the principle of least-effort. Cambridge MA edn. Reading: Addison-Wesley. (1949).