The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1319-5166.htm

Estimation of different entropies via Abel–Gontscharoff Green functions and Fink's identity using Jensen type functionals

Khuram Ali Khan
Department of Mathematics, University of Sargodha, Sargodha, Pakistan
Tasadduq Niaz
Department of Mathematics, University of Sargodha, Sargodha, Pakistan and Department of Mathematics, The University of Lahore, Sargodha-Campus, Sargodha, Pakistan
Đilda Pečarić
Catholic University of Croatia, Zagreb, Croatia, and
Josip Pečarić
RUDN University, Moscow, Russia

Received 7 November 2018
Revised 15 December 2018
Accepted 18 December 2018

Abstract
In this work, we estimated the different entropies like Shannon entropy, Rényi divergences, Csiszár divergence by using Jensen's type functionals. The Zipf's–Mandelbrot law and hybrid Zipf's–Mandelbrot law are used to estimate the Shannon entropy. The Abel–Gontscharoff Green functions and Fink's Identity are used to construct new inequalities and generalized them for *m*-convex function.

Keywords *m*-convex function, Jensen's inequality, Shannon entropy, *f*- and Rényi divergence, Fink's identity, Abel–Gontscharoff Green function, Entropy

Paper type Original Article

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).

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

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

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

The publisher wishes to inform readers that the article "Estimation of different entropies via Abel–Gontscharoff Green functions and Fink's identity using Jensen type functionals" was originally published by the previous publisher of the *Arab Journal of Mathematical Sciences* and the pagination of this article has been subsequently changed. There has been no change to the content of the article. This change was necessary for the journal to transition from the previous publisher to the new one. The publisher sincerely apologises for any inconvenience caused. To access and cite this article, please use Khan, K.A., Niaz, T., Pečarić, Đ., Pečarić, J. (2018), "Estimation of different entropies via Abel–Gontscharoff Green functions and Fink's identity using Jensen type functionals" *Arab Journal of Mathematical Sciences*, Vol. 26 No. 1/2, pp. 15-39. The original publication date for this paper was 31/12/2018.

Arab Journal of Mathematical Sciences
Vol. 26 No. 1/2, 2020
pp. 15-39
Emerald Publishing Limited
e-ISSN: 2588-9214
p-ISSN: 1319-5166
DOI 10.1016/j.ajmsc.2018.12.002

1. Introduction and preliminary results

In recent years many researchers generalized different inequalities using different identities involving green functions, for example in [24] Nasir et al. generalized the Popoviciu inequality using Mongomery identity along with the new green function. Also in [25] Niaz et al. used Fink's identity along with new Abel–Gontscharoff type Green functions for 'two point right focal' to generalize the refinement of Jensen inequality.

The most commonly used words, the largest cities of countries, income of 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 Csiszár f-divergence [11,12], some special case of which is Kullback–Leibler-divergence used to find the appropriate distance between the probability distributions (see [20,21]). The notion of distance is stronger than divergence because it gives the properties of symmetry and triangle inequalities. Probability theory has application in many fields and the divergence between probability distribution has many applications in these fields.

Many natural phenomena like distribution of wealth and income in a society, distribution of face book likes, distribution of football goals follow power law distribution (Zipf's Law). Like above phenomena, distribution of city sizes also follows Power Law distribution. Auerbach [3] 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 [32] worked significantly in this field. The distribution of city sizes is investigated by many scholars of the urban economics, like Rosen and Resnick [29], Black and Henderson [4], Ioannides and Overman [19], Soo [30], Anderson and Ge [2] and Bosker et al. [5]. 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 nth 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 shows that the estimated value of the shape parameter is equal to unity.

In [18] 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–Mandelbrot 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 [27, 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 remain valid.

In [27, p. 16], the following criteria are given to check the m-convexity of the function.

If $f^{(m)}$ exists, then f is m-convex if and only if $f^{(m)} \geq 0$.

In recent years many researchers have generalized the inequalities for m-convex functions; 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–Gontscharoff interpolation and Hermite interpolating polynomial (see [6–10]).

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 [16,17]). Like many researchers L. Horváth and J. Pečarić in [14,17], see also [15, 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

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

and

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

by

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

and

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

Next let the function

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

defined by

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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* $\sum_{s=1}^{n} p_s = 1$, *then*

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

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

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

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

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.

In [26], the green function $G : [\alpha_1, \alpha_2] \times [\alpha_1, \alpha_2] \to \mathbb{R}$ is defined as

$$G(u,v) = \begin{cases} \dfrac{(u-\alpha_2)(v-\alpha_1)}{\alpha_2-\alpha_1}, & \alpha_1 \leq v \leq u; \\[2ex] \dfrac{(v-\alpha_2)(u-\alpha_1)}{\alpha_2-\alpha_1}, & u \leq v \leq \alpha_2. \end{cases} \tag{5}$$

The function G is convex with respect to v and due to symmetry also convex with respect to u. One can also note that G is continuous function.

In [31] it is given that any function $f : [\alpha_1, \alpha_2] \to \mathbb{R}$, such that $f \in C^2([\alpha_1, \alpha_2])$ can be written as

$$f(u) = \frac{\alpha_2 - u}{\alpha_2 - \alpha_1} f(\alpha_1) + \frac{u - \alpha_1}{\alpha_2 - \alpha_1} f(\alpha_2) + \int_{\alpha_2}^{\alpha_1} G(u,v) f''(v) dv. \tag{6}$$

2. Inequalities for Csiszár divergence

In [11,12] Csiszár introduced the following notion.

Definition 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

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

And he stated that by defining

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

we can also use the nonnegative probability distributions as well.

In [18], L. Horv́ath, et al. gave the following functional based on the previous definition.

Definition 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

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

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

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

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

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*

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

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

$$\begin{aligned}\widehat{I}_f(\mathbf{r},\ \mathbf{q}) &= \sum_{s=1}^{n} q_s\ f\left(\frac{r_s}{q_s}\right)=A_{m,1}^{[1]}\geq A_{m,2}^{[1]}\geq\cdots\geq A_{m,m-1}^{[1]}\geq A_{m,m}^{[1]}\\ &\geq f\left(\frac{\sum_{s=1}^{n} r_s}{\sum_{s=1}^{n} q_s}\right)\sum_{s=1}^{n} q_s.\end{aligned} \tag{10}$$

where

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

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

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

$$\begin{aligned}\left(\sum_{s=1}^{n} r_s\right) f\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) &\leq A_{m,m}^{[2]} \leq A_{m,m-1}^{[2]} \leq \cdots \leq A_{m,2}^{[2]} \leq A_{m,1}^{[2]} \\ &= \sum_{s=1}^{n} r_s f\left(\frac{r_s}{q_S}\right) = \widehat{I}_{idf}(\mathbf{r},\ \mathbf{q})\end{aligned} \tag{12}$$

where

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

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

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

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

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

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

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

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

$$\times\left(\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n}q_s}}{\alpha_{I_m,i_j}}\right)\left(\frac{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n}q_s}}{\alpha_{I_m,i_j}}\frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n}q_s}}{\alpha_{I_m,i_j}}}\right)f\left(\frac{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n}q_s}}{\alpha_{I_m,i_j}}\frac{r_{i_j}}{q_{i_j}}}{\sum_{j=1}^{l}\frac{\frac{q_{i_j}}{\sum_{s=1}^{n}q_s}}{\alpha_{I_m,i_j}}}\right) \tag{15}$$

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

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

3. Inequalities for Shannon Entropy

Definition 3 (*See* [*18*]). The Shannon entropy of positive probability distribution $\mathbf{r}=(r_1,\ldots,r_n)$ is defined by

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

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*

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

where

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

If the base of log *is between* 0 *and* 1*, then inequality signs in* (17) *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* $\mathbf{q}$

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

where

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

Proof. (i) Using $f := \log$ and $\mathbf{r} = (1, \ldots, 1)$ in Theorem 2.1 (i), we get (17).
(ii) It is the special case of (i). □

Definition 4 (*See* [18])
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

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

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*

$$\begin{aligned}\sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) &\le A^{[5]}_{m,m} \le A^{[5]}_{m,m-1} \le \cdots \le A^{[5]}_{m,2} \le A^{[5]}_{m,1} \\ &= \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right) = D(\mathbf{r},\ \mathbf{q}),\end{aligned} \tag{21}$$

where

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

If the base of $\log$ *is between 0 and 1, then inequality in* (21) *is reversed.*
(ii) *If* $\mathbf{r}$ *and* $\mathbf{q}$ *are positive probability distributions, and the base of* l *is greater than* 1, *then we have*

$$D(\mathbf{r},\ \mathbf{q}) = A^{[6]}_{m,1} \ge A^{[6]}_{m,2} \ge \cdots \ge A^{[6]}_{m,m-1} \ge A^{[6]}_{m,m} \ge 0, \tag{22}$$

where

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

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

Proof. (i) On taking $f := \log$ in Theorem 2.1 (ii), we get (21).

(ii) Since $\mathbf{r}$ and $\mathbf{q}$ are positive probability distributions therefore $\sum_{s=1}^{n} r_s = \sum_{s=1}^{n} q_s = 1$, so the smallest term in (21) is given as

$$\sum_{s=1}^{n} r_s \log\left(\sum_{s=1}^{n} \frac{r_s}{\sum_{s=1}^{n} q_s}\right) = 0. \tag{23}$$

Hence for positive probability distribution $\mathbf{r}$ and $\mathbf{q}$ the (21) will become (22). □

4. Inequalities for Rényi Divergence and Entropy

The Rényi divergence and entropy come from [28].

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

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

$$D_\lambda(\mathbf{r},\ \mathbf{q}) := \frac{1}{\lambda - 1}\log\left(\sum_{i=1}^{n} q_i\left(\frac{r_i}{q_i}\right)^{\lambda}\right). \tag{24}$$

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

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

The Rényi divergence and the Rényi entropy can also be extended to non-negative probability distributions. If $\lambda \to 1$ in (24), we have the Kullback–Leibler divergence, and if $\lambda \to 1$ in (25), 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* $\mathbf{r} = (r_1, \ldots, r_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *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*

$$D_\lambda(\mathbf{r},\ \mathbf{q}) \leq A^{[7]}_{m,m} \leq A^{[7]}_{m,m-1} \leq \cdots \leq A^{[7]}_{m,2} \leq A^{[7]}_{m,1} = D_\mu(\mathbf{r},\ \mathbf{q}), \tag{26}$$

where

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

The reverse inequalities hold in (26) *if the base of* log *is between 0 and 1.*
(ii) *If* $1 < \mu$ *and the base of* log *is greater than 1, then*

$$D_1(\mathbf{r},\ \mathbf{q}) = D(\mathbf{r},\ \mathbf{q}) = \sum_{s=1}^{n} r_s\log\left(\frac{r_s}{q_s}\right) \leq A^{[8]}_{m,m} \leq A^{[8]}_{m,m-1} \leq \cdots \leq A^{[8]}_{m,2} \leq A^{[8]}_{m,1} = D_\mu(\mathbf{r},\ \mathbf{q}), \tag{27}$$

where

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

here the base of exp *is the same as the base of* log*, and the reverse inequalities hold if the base of* log *is between 0 and 1.*
(iii) *If* $0 \leq \lambda < 1$*, and the base of* log *is greater than 1, then*

$$D_\lambda(\mathbf{r},\ \mathbf{q}) \le A^{[9]}_{m,m} \le A^{[9]}_{m,m-1} \le \cdots \le A^{[9]}_{m,2} \le A^{[9]}_{m,1} = D_1(\mathbf{r},\ \mathbf{q}), \tag{28}$$

where

$$A^{[9]}_{m,l} = \frac{1}{\lambda-1}\frac{(m-1)!}{(l-1)!}\sum_{(i_1,\ldots,i_l)\in I_l}\eta_{I_m,l}(i_1,\ldots,i_l)\left(\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\right)\times\log\left(\frac{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right) \tag{29}$$

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

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

we have

$$\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)^{\frac{\mu-1}{\lambda-1}} = \left(\sum_{s=1}^{n} r_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)^{\frac{\mu-1}{\lambda-1}} \le \ldots \le \frac{(m-1)!}{(l-1)!}\sum_{(i_1,\ldots,i_l)\in I_l}\eta_{I_m,l}(i_1,\ldots,i_l)\left(\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\right)$$
$$\times\left(\frac{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)^{\frac{\mu-1}{\lambda-1}} \le \ldots \le \sum_{s=1}^{n} r_s\left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)^{\frac{\mu-1}{\lambda-1}} \tag{30}$$

if either $0 \le \lambda < 1 < \beta$ or $1 < \lambda \le \mu$, and the reverse inequality in (30) holds if $0 \le \lambda \le \beta < 1$. By raising to power $\frac{1}{\mu-1}$, we have from all

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

Since log is increasing if the base of log is greater than 1, it now follows (26). If the base of log is between 0 and 1, then log is decreasing and therefore inequality in (26) is 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* $\mathbf{r}=(r_1,\ldots,r_n)$ *and* $\mathbf{q}=(q_1,\ldots,q_n)$ *are 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*

$$\frac{1}{\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda}} \sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda} \log\left(\frac{r_s}{q_s}\right) = A_{m,1}^{[10]} \le A_{m,2}^{[10]} \le \cdots \le A_{m,m-1}^{[10]} \le A_{m,m}^{[10]} \le D_{\lambda}(\mathbf{r},\mathbf{q}) \le A_{m,m}^{[11]} \le A_{m,m}^{[11]} \le \cdots \le A_{m,2}^{[11]} \le A_{m,1}^{[11]} = D_1(\mathbf{r},\mathbf{q}) \tag{32}$$

where

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

and

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

The inequalities in (32) 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 l is greater than 1.

Proof. We prove only the case when $0 \le \lambda < 1$ 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

$$D_{\lambda}(\mathbf{r},\ \mathbf{q}) = \frac{1}{\lambda-1} \log\left(\sum_{s=1}^{n} q_s \left(\frac{r_s}{q_s}\right)^{\lambda}\right) = \frac{1}{\lambda-1} \log\left(\sum_{s=1}^{n} r_s \left(\frac{r_s}{q_s}\right)^{\lambda-1}\right)$$
$$\le \cdots \le \frac{1}{\lambda-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\right) \log\left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}} \left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right) \tag{33}$$
$$\le \cdots \le \frac{1}{\lambda-1} \sum_{s=1}^{n} r_s \log\left(\left(\frac{r_s}{q_s}\right)^{\lambda-1}\right) = \sum_{s=1}^{n} r_s \log\left(\frac{r_s}{q_s}\right) = D_1(\mathbf{r},\ \mathbf{q})$$

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

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

$$
\begin{aligned}
D_\lambda(\mathbf{r},\ \mathbf{q}) &= \frac{1}{\lambda-1}\log\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)\\
&= \frac{1}{\lambda-1\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)}\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)\log\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)\\
&\geq \cdots \geq \frac{1}{\lambda-1\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)}\frac{(m-1)!}{(l-1)!}\sum_{(i_1,\ldots,i_l)\in I_l}\eta_{I_m,l}(i_1,\ldots,i_l)\left(\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\right)\\
&\quad\left(\frac{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)\log\left(\frac{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)\\
&= \frac{1}{\lambda-1\left(\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\right)}\frac{(m-1)!}{(l-1)!}\sum_{(i_1,\ldots,i_l)\in I_l}\eta_{I_m,l}(i_1,\ldots,i_l)\\
&\quad\left(\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}\right)\log\left(\frac{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\left(\frac{r_{i_j}}{q_{i_j}}\right)^{\lambda-1}}{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right)\\
&\geq \cdots \geq \frac{1}{\lambda-1}\sum_{s=1}^{n} r_s\left(\frac{r_s}{q_s}\right)^{\lambda-1}\log\left(\frac{r_s}{q_s}\right)^{\lambda-1}\frac{1}{\sum_{s=1}^{n} r_s\left(\frac{r_s}{q_s}\right)^{\lambda-1}}\\
&= \frac{1}{\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}}\sum_{s=1}^{n} q_s\left(\frac{r_s}{q_s}\right)^{\lambda}\log\left(\frac{r_s}{q_s}\right)
\end{aligned}
\tag{34}
$$

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

By using Theorems 4.1, 4.2 and Definition 5, some inequalities of Rényi entropy are obtained. Let $\frac{1}{\mathbf{n}} = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ be a discrete probability distribution.

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

$$
H_\lambda(\mathbf{r}) = \log(n) - D_\lambda\left(\mathbf{r}, \frac{1}{\mathbf{n}}\right) \geq A_{m,m}^{[12]} \geq A_{m,m}^{[12]} \geq \cdots A_{m,2}^{[12]} \geq A_{m,1}^{[12]} = H_\mu(\mathbf{r}), \tag{35}
$$

where

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

The reverse inequalities hold in (35) *if the base of* log *is between 0 and 1.*

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

$$S = -\sum_{s=1}^{n} p_i\log(p_i) \geq A^{[13]}_{m,m} \geq A^{[13]}_{m,m-1} \geq \cdots \geq A^{[13]}_{m,2} \geq A^{[13]}_{m,1} = H_{\mu}(\mathbf{r}) \tag{36}$$

where

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

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

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

$$H_{\lambda}(\mathbf{r}) \geq A^{[14]}_{m,m} \geq A^{[14]}_{m,m-1} \geq \cdots \geq A^{[14]}_{m,2} \leq A^{[14]}_{m,1} = S, \tag{37}$$

where

$$A^{[14]}_{m,m} = \frac{1}{1-\lambda}\frac{(m-1)!}{(l-1)!}\sum_{(i_1,\ldots,i_l)\in I_l}\eta_{I_m,l}(i_1,\ldots,i_l)\left(\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}\right)\times\log\left(\frac{\sum_{j=1}^{l}\frac{r^{\lambda}_{i_j}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l}\frac{r_{i_j}}{\alpha_{I_m,i_j}}}\right). \tag{38}$$

The inequalities in (37) *are reversed if the base of* log *is between 0 and 1.*

Proof. (i) Suppose $\mathbf{q}=\frac{1}{\mathbf{n}}$ then from (24), we have

$$D_{\lambda}(\mathbf{r},\ \mathbf{q}) = \frac{1}{\lambda-1}\log\left(\sum_{s=1}^{n} n^{\lambda-1}r^{\lambda}_{s}\right) = \log(n) + \frac{1}{\lambda-1}\log\left(\sum_{s=1}^{n} r^{\lambda}_{s}\right), \tag{39}$$

therefore we have

$$H_{\lambda}(\mathbf{r}) = \log(n) - D_{\lambda}\left(\mathbf{r},\frac{\mathbf{1}}{\mathbf{n}}\right). \tag{40}$$

Now using Theorem 4.1 (i) and (40), we get

$$H_\lambda(\mathbf{r}) = \log(n) - D_\lambda\left(\mathbf{r}, \frac{\mathbf{1}}{\mathbf{n}}\right) \geq \cdots \geq \log(n) - \frac{1}{\mu - 1}$$
$$\times \log\left(n^{\mu-1} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \times \left(\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}\right) \left(\frac{\sum_{j=1}^{l} \frac{r_{i_j}^\lambda}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{r_{i_j}}{\alpha_{I_m,i_j}}} \right)^{\frac{\mu-1}{\lambda-1}} \right) \tag{41}$$
$$\geq \cdots \geq \log(n) - D_\mu(\mathbf{r},\ \mathbf{q}) = H_\mu(\mathbf{r}),$$

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

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

$$-\frac{1}{\sum_{s=1}^{n} r_s^\lambda} \sum_{s=1}^{n} r_s^\lambda \log(r_s) = A_{m,1}^{[15]} \geq A_{m,2}^{[15]} \geq \cdots \geq A_{m,m-1}^{[15]} \geq A_{m,m}^{[15]}$$
$$\geq H_\lambda(\mathbf{r}) \geq A_{m,m}^{[16]} \geq A_{m,m-1}^{[16]} \geq \cdots A_{m,2}^{[16]} \geq A_{m,1}^{[16]} = H(\mathbf{r}), \tag{42}$$

where

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

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 (42) 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 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 suggests a simpler distribution called Zipf's law. The Zipf's law is defined as follows (see [32]).

Definition 6. 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

$$f(s, N, t) = \frac{\frac{1}{s^t}}{\sum_{j=1}^{N} \frac{1}{j^t}}. \tag{43}$$

The Zipf–Mandelbrot law is defined as follows (see [22]).

Definition 7. 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

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

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 $\mathbb{N}$, then for $q \geq 0$, $t > 1$, $s \in \mathbb{N}$, density function of Zipf–Mandelbrot law becomes

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

where

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

For $q = 0$, the Zipf–Mandelbrot law (44) becomes Zipf's law (43).

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

$$\begin{aligned} H_\lambda(\mathbf{r}) &= \frac{1}{1-\lambda} \log\left(\frac{1}{H^{\lambda}_{N,q,t}} \sum_{s=1}^{n} \frac{1}{(s+q)^{\lambda s}}\right) \geq \cdots \geq \\ &\frac{1}{1-\lambda} \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l) \left(\sum_{j=1}^{l} \frac{1}{\alpha_{I_m,i_j}(i_j+q) H_{N,q,t}} \right) \\ &\times \log\left(\frac{1}{H^{\lambda-1}_{N,q,t}} \frac{\sum_{j=1}^{l} \frac{1}{\alpha_{I_m,i_j}(i_j-q)^{\lambda s}}}{\sum_{j=1}^{l} \frac{1}{\alpha_{I_m,i_j}(i_j-q)^{s}}} \right) \geq \cdots \geq \\ &\frac{t}{H_{N,q,t}} \sum_{s=1}^{N} \frac{\log(s+q)}{(s+q)^t} + \log(H_{N,q,t}) = S. \end{aligned} \tag{48}$$

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

Conclusion 5.2. *Assume* (H_1)*, let* $\mathbf{r}_1$ *and* $\mathbf{r}_2$ *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 l is greater than 1, then*

$$\overline{D}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{s=1}^{n} \frac{1}{(s+q_1)^{t_1} H_{N,q_1,t_1}} \log\left(\frac{(s+q_2)^{t_2} H_{N,q_2,t_2}}{(s+q_1)^{t_1} H_{N,q_2,t_1}}\right) \geq \cdots$$
$$\geq \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l)\in I_l} \eta_{I_m,l}(i_1,\ldots,i_l)$$
$$\times \left(\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}\right) \left(\frac{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}}\right) \times \log\left(\frac{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_1)^{t_1} H_{N,q_1,t_1}}}{\alpha_{I_m,i_j}}}{\sum_{j=1}^{l} \frac{\frac{1}{(i_j+q_2)^{t_2} H_{N,q_2,t_2}}}{\alpha_{I_m,i_j}}}\right) \geq \cdots \geq 0. \tag{49}$$

The inequalities in (49) are reversed if the base of l 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 maximizes the Shannon entropy under the constraints*

$$\sum_{s\in J} r_s = 1, \sum_{s\in J} r_s(\text{In}(s+q)) := \psi,$$

is Zipf–Mandelbrot law.

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

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

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

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

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

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

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

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

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 maximizes Shannon entropy under constraints*

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

is hybrid Zipf–Mandelbrot law given as

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

where

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

Proof. First consider $J = \{1,\ldots,N\}$, 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 by

$$\Phi_N^*(k,q,w) = \sum_{s=1}^{N} \frac{w^s}{(s+q)^k} \text{ with } w \ge 0, k > 0.$$

$\square$

Remark 6.4. Observe that for 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 $(q_1,\ldots,q_N)$ 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:

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

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

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

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

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

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

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

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

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

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

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

Under the assumption of Corollary 3.2 (ii), define the non-negative 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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Under the assumption of Corollary 4.4, define the following functionals

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

$$\Theta_{32} = A^{[15]}_{m,r} - A^{[15]}_{m,k}, \quad 1 \leq r < k \leq m. \tag{79}$$

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

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

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

7. Generalization of refinement of Jensen's, Rényi and Shannon type inequalities Fink's Identity and Abel–Gontscharoff Green function

In [13], A. M. Fink gave the following result.

Let $f : [\alpha_1, \alpha_2] \to \mathbb{R}$, where $[\alpha_1, \alpha_2]$ be an interval, is a function such that $f^{(n-1)}$ is absolutely continuous then the following identity holds

$$f(z) = \frac{n}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} f(\zeta) d\zeta + \sum_{\lambda=1}^{n-1} \frac{n-\lambda}{\lambda!} \left(\frac{f^{(\lambda-1)}(\alpha_2)(z-\alpha_2)^{\lambda} - f^{(\lambda-1)}(\alpha_1)(z-\alpha_1)^{\lambda}}{\alpha_2 - \alpha_1} \right)$$
$$+ \frac{1}{(n-1)!(\alpha_2-\alpha_1)} \int_{\alpha_1}^{\alpha_2} (z-\zeta)^{n-1} F_{\alpha_1}^{\alpha_2}(\zeta, z) f^{(n)}(\zeta) d\zeta, \tag{83}$$

where

$$F_{\alpha_1}^{\alpha_2}(\zeta, z) = \begin{cases} \zeta - \alpha_1, & \alpha_1 \le \zeta \le z \le \alpha_2; \\ \zeta - \alpha_2, & \alpha_1 \le z < \zeta \le \alpha_2. \end{cases} \tag{84}$$

The complete reference about Abel–Gontscharoff polynomial and theorem for 'two-point right focal' problem is given in [1].

The Abel–Gontscharoff polynomial for 'two-point right focal' interpolating polynomial for $n = 2$ can be given as

$$f(z) = f(\alpha_1) + (z - \alpha_1) f'(\alpha_2) + \int_{\alpha_1}^{\alpha_2} G_1(z, w) f''(w) dw, \tag{85}$$

where

$$G_1(z, w) = \begin{cases} \alpha_1 - w, & \alpha_1 \le w \le z; \\ \alpha_1 - z, & z \le w \le \alpha_2. \end{cases} \tag{86}$$

In [8], S. I. Butt et al. gave some new types of Green functions defined as

$$G_2(z, w) = \begin{cases} \alpha_2 - z, & \alpha_1 \le w \le z; \\ \alpha_2 - w, & z \le w \le \alpha_2, \end{cases} \tag{87}$$

$$G_3(z, w) = \begin{cases} z - \alpha_1, & \alpha_1 \le w \le z; \\ w - \alpha_1, & z \le w \le \alpha_2, \end{cases} \tag{88}$$

$$G_4(z, w) = \begin{cases} \alpha_2 - w, & \alpha_1 \le w \le z; \\ \alpha_2 - z, & z \le w \le \alpha_2, \end{cases} \tag{89}$$

Figure 1 shows the graph of Green functions $G_i(z, w), i = 1, 2, 3, 4$ defined in (86)–(89) respectively for fixed value of w. They also introduced some new Abel–Gontscharoff type identities by using these new Green functions in the following lemma.

Lemma A. *Let $f : [\alpha_1, \alpha_2]$ be a twice differentiable function and G_k $(k = 2, 3, 4)$ be the 'two-point right focal problem'-type Green functions defined by (87)–(89). Then the following identities hold:*

$$f(z) = f(\alpha_2) - (\alpha_2 - z) f'(\alpha_1) - \int_{\alpha_1}^{\alpha_2} G_2(z, w) f''(w) dw, \tag{90}$$

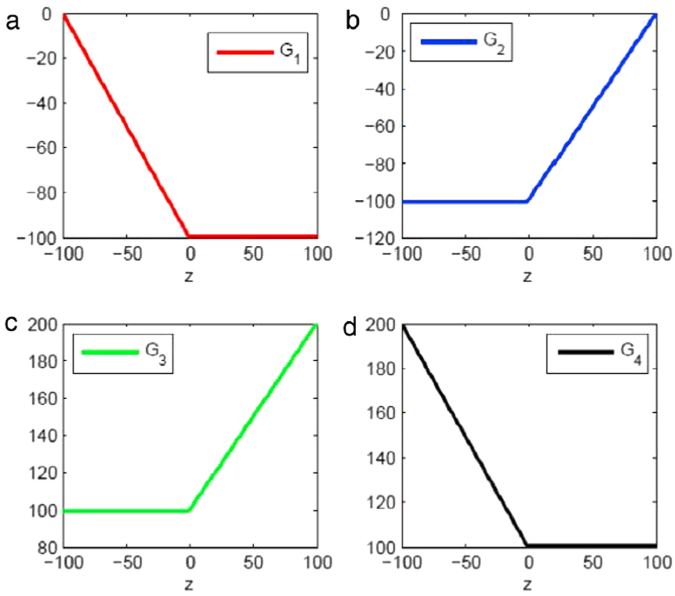

Figure 1.
Graph of Green functions for fix w.

$$f(z) = f(\alpha_2) - (\alpha_2 - \alpha_1)f'(\alpha_2) + (z - \alpha_1)f'(\alpha_1) + \int_{\alpha_1}^{\alpha_2} G_3(z, w)f''(w)dw, \tag{91}$$

$$f(z) = f(\alpha_1) + (\alpha_2 - \alpha_1)f'(\alpha_1) - (\alpha_2 - z)f'(\alpha_2) + \int_{\alpha_1}^{\alpha_2} G_4(z, w)f''(w)dw. \tag{92}$$

Theorem 7.1. *Assume* (H_1)*, and let* $f : I = [\alpha_1, \alpha_2] \to \mathbb{R}$ *be a function such that for* $m \geq 3$ *(an integer)* $f^{(m-1)}$ *is absolutely continuous. Also, let* $x_1, \ldots, x_n \in I$*,* $p_1, \ldots, p_n$*, be positive real numbers such that* $\sum_{i=1}^{n} p_i = 1$*. Assume that* $F_{\alpha_1}^{\alpha_2}$*,* G_k $(k = 1, 2, 3, 4)$ *and* Θ_i $(i = 1, \ldots, 35)$ *are the same as defined in (84), (86)–(89), (2), (3), (50)–(82) respectively.*
Then:

(1) *For* $k = 1, 3, 4$ *we have the following identities:*

$$\begin{aligned}\Theta_i(f) = (m-2)\left(\frac{f'(\alpha_2) - f'(\alpha_1)}{\alpha_2 - \alpha_1}\right)\int_{\alpha_1}^{\alpha_2} \Theta_i(G_k(\cdot, w))dw + \frac{1}{\alpha_2 - \alpha_1}\int_{\alpha_1}^{\alpha_2} \Theta_i(G_k(\cdot, w))\\ \times \sum_{\lambda=1}^{m-3}\left(\frac{m-2-\lambda}{\lambda!}\right)\left(f^{(\lambda+1)}(\alpha_2)(w-\alpha_2)^{\lambda} - f^{(\lambda+1)}(\alpha_1)(w-\alpha_1)^{\lambda}\right)dw\\ + \frac{1}{(m-3)!(\alpha_2 - \alpha_1)}\int_{\alpha_1}^{\alpha_2} f^{(m)}(\zeta)\\ \times\left(\int_{\alpha_1}^{\alpha_2} \Theta_i(G_k(\cdot, w))(w-\zeta)^{m-3}F_{\alpha_1}^{\alpha_2}\alpha_2(\zeta, w)dw\right)d\zeta, \quad i = 1, \ldots, 35.\end{aligned} \tag{93}$$

(2) *For $k = 2$ we have*

$$\begin{aligned}\Theta_i(f) &= (-1)(m-2)\left(\frac{f'(\alpha_2) - f'(\alpha_1)}{\alpha_2 - \alpha_1}\right)\int_{\alpha_1}^{\alpha_2} \Theta_i(G_2(\cdot, w))dw \\ &+ \frac{(-1)}{\alpha_2 - \alpha_1}\int_{\alpha_1}^{\alpha_2} \Theta_i(G_2(\cdot, w)) \times \sum_{\lambda=1}^{m-3}\left(\frac{m-2-\lambda}{\lambda!}\right)\left(f^{(\lambda+1)}(\alpha_2)(w-\alpha_2)^{\lambda}\right. \\ &\left. - f^{(\lambda+1)}(\alpha_1)(w-\alpha_1)^{\lambda}\right)dw + \frac{(-1)}{(m-3)!(\alpha_2 - \alpha_1)}\int_{\alpha_1}^{\alpha_2} f^{(m)}(\zeta) \\ &\times \left(\int_{\alpha_1}^{\alpha_2} \Theta_i(G_2(\cdot, w))(w-\zeta)^{m-3} F_{\alpha_1}^{\alpha_2}(\zeta, w)dw\right)d\zeta,\end{aligned} \tag{94}$$

$i = 1, \ldots, 35.$

Proof. (i) Using Abel–Gontsharoff-typeidentities (85), (91), (92) in $\Theta_i(f)$, $i = 1, \ldots, 35$, and using properties of $\Theta_i(f)$, we get

$$\Theta_i(f) = \int_{\alpha_1}^{\alpha_2} \Theta_i(G_k(\cdot, w))f''(w)dw, \quad i = 1, 2. \tag{95}$$

From identity (83), we get

$$\begin{aligned}f'(w) &= (m-2)\left(\frac{f'(\alpha_2) - f'(\alpha_1)}{\alpha_2 - \alpha_1}\right) + \sum_{\lambda=1}^{m-3}\left(\frac{m-2-\lambda}{\lambda!}\right) \\ &\times \left(\frac{f^{(\lambda)}(\alpha_2)(w-\alpha_2)^{\lambda-1} - f^{(\lambda)}(\alpha_2)(w-\alpha_2)^{\lambda-1}}{\alpha_2 - \alpha_1}\right) \\ &+ \frac{1}{(m-3)!(\alpha_2 - \alpha_1)}\int_{\alpha_1}^{\alpha_2} (w-\zeta)^{m-3} F_{\alpha_1}^{\alpha_2}(\zeta, w) f^{(m)}(\zeta)d\zeta.\end{aligned} \tag{96}$$

Using (95) and (96) and applying Fubini's theorem we get the result (93) for $k = 1, 3, 4$.

(ii) Substituting Abel–Gontscharoff-typeinequality (90) in $\Theta_i(f)$, $i = 1, \ldots, 35$, and following similar steps to (i), we get (94). □

Theorem 7.2. *Assume (H_1), and let $f : I = [\alpha_1, \alpha_2] \to \mathbb{R}$ be a function such that for $m \geq 3$ (an integer) $f^{(m-1)}$ is absolutely continuous. Also, let $x_1, \ldots, x_n \in I$, $p_1, \ldots, p_n$ are positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. Assume that $F_{\alpha_1}^{\alpha_2}$, G_k $(k = 1, 2, 3, 4)$ and Θ_i $(i = 1, 2)$ are the same as defined in (84), (86)–(89), (2), (3), (50)–(82) respectively. For $m \geq 3$ assume that*

$$\int_{\alpha_1}^{\alpha_2} \Theta_i(G_k(\cdot, \zeta))(w-\zeta)^{m-3} F_{\alpha_1}^{\alpha_2}(\zeta, w)dw \geq 0, \zeta \in [\alpha_1, \alpha_2], \quad i = 1, \ldots, 35, \tag{97}$$

for $k = 1, 3, 4$. If f is an m-convex function, then

(i) *For $k = 1, 3, 4$, the following holds:*

$$\begin{aligned}\Theta_i(f) &\geq (m-2)\left(\frac{f'(\alpha_2) - f'(\alpha_1)}{\alpha_2 - \alpha_1}\right)\int_{\alpha_1}^{\alpha_2}\Theta_i(G_k(\cdot, w))dw \\ &+ \frac{1}{\alpha_2 - \alpha_1}\int_{\alpha_1}^{\alpha_2}\Theta_i(G_k(\cdot, w)) \times \sum_{\lambda=1}^{m-3}\left(\frac{m-2-\lambda}{\lambda!}\right)\Big(f^{(\lambda+1)}(\alpha_2)(w-\alpha_2)^{\lambda} \\ &- f^{(\lambda+1)}(\alpha_1)(w-\alpha_1)^{\lambda}\Big)dw, \\ &i = 1, \ldots, 35.\end{aligned} \tag{98}$$

(ii) *For $k = 2$, we have*

$$\begin{aligned}\Theta_i(f) &\leq (-1)(m-2)\left(\frac{f'(\alpha_2) - f''(\alpha_1)}{\alpha_2 - \alpha_1}\right)\int_{\alpha_1}^{\alpha_2}\Theta_i(G_2(\cdot, w))dw \\ &+ \frac{(-1)}{\alpha_2 - \alpha_1}\int_{\alpha_1}^{\alpha_2}\Theta_i(G_2(\cdot, w)) \times \sum_{\lambda=1}^{m-3}\left(\frac{m-2-\lambda}{\lambda!}\right)\Big(f^{(\lambda+1)}(\alpha_2)(w-\alpha_2)^{\lambda} \\ &- f^{(\lambda+1)}(\alpha_1)(w-\alpha_1)^{\lambda}\Big)dw, \\ &i = 1, \ldots, 35.\end{aligned} \tag{99}$$

Proof. (i) Since $f^{(m-1)}$ is absolutely continuous on $[\alpha_1, \alpha_2]$, $f^{(m)}$ exists almost everywhere. Also, since f is m-convex therefore we have $f^{(m)}(\zeta) \geq 0$ for a.e. on $[\alpha_1, \alpha_2]$. So, applying Theorem 1.1, we obtain (98).

(ii) Similar to (i). □

Remark A. We can investigate the bounds for the identities related to the generalization of refinement of Jensen inequality using inequalities for the Čebyšev functional and some results relating to the Grüss and Ostrowski type inequalities can be constructed as given in Section 3 of [6]. Also we can construct the non-negative functionals from inequalities (98)–(99) 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 [6].

References

[1] R.P. Agarwal, P.J.Y. Wong, Error Inequalities in Polynomial Interpolation and their Applications, Kluwer Academic Publishers, Dordrecht/Boston/London, 1993.

[2] G. Anderson, Y. Ge, The size distribution of Chinese cities, Reg. Sci. Urban Econ. 35 (6) (2005) 756–776.

[3] F. Auerbach, Das Gesetz der Bevölkerungskonzentration, Petermanns Geographische Mitteilungen 59 (1913) 74–76.

[4] D. Black, V. Henderson, Urban evolution in the USA, J. Econ. Geogr. 3 (4) (2003) 343–372.

[5] M. Bosker, S. Brakman, H. Garretsen, M. Schramm, A century of shocks: the evolution of the German city size distribution 1925-1999, Reg. Sci. Urban Econ. 38 (4) (2008) 330–347.

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

[7] S.I. Butt, K.A. Khan, J. Pečarić, Popoviciu type inequalities via Hermite's polynomial, Math. Inequal. Appl. 19 (4) (2016) 1309–1318.

[8] S.I. Butt, N. Mehmood, J. Pečarić, New generalizations of Popoviciu type inequalities via new green functions and Fink's identity, Trans A. Razmadze Math. Inst. 171 (3) (2017) 293–303.

[9] S.I. Butt, J. Pečarić, Weighted Popoviciu type inequalities via generalized Montgomery identities, Hrvat. Akad. Znan. I Umjet.: Mat. Znan. 19 (523) (2015) 69–89.

[10] S.I. Butt, J. Pečarić, Popoviciu'S Inequality for *N*-Convex Functions, Lap Lambert Academic Publishing, 2016.

[11] I. Csiszár, Information measures: a critical survey, in: Tans. 7th Prague Conf. on Info. Th. Statist. Decis. Funct. Random Process and 8th European Meeting of Statist. Vol. B, Academia Prague, 1978, pp. 73–86.

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

[13] A.M. Fink, Bounds on the deviation of a function from its averages, Czechoslovak Math. J. 42 (2) (1992) 289–310.

[14] L. Horváth, A method to refine the discrete Jense's inequality for convex and mid-convex functions, Math. Comput. Modelling 54 (9–10) (2011) 2451–2459.

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

[16] L. Horváth, K.A. Khan, J. Pečarić, Refinement of Jensen's inequality for operator convex functions, Advances in Inequalities and Applications (2014).

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

[18] L. Horváth, Đ. Pečarić, J. Pečarić, Estimations of f-and Rényi divergences by using a cyclic refinement of the Jensen's inequality, Bull. Malays. Math. Sci. Soc. (2017) 1–14.

[19] Y.M. Ioannides, H.G. Overman, Zipf's law for cities: an empirical examination, Reg. Sci. Urban Econ. 33 (2) (2003) 127–137.

[20] S. Kullback, Information Theory and Statistics, Courier Corporation, 1997.

[21] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Statist. 22 (1) (1951) 79–86.

[22] N. Lovričević, Đ. Pečarić, J. Pečarić, Zipf-Mandelbrot law, f-divergences and the Jensen-type interpolating inequalities, J. Inequal. Appl. 2018 (1) (2018) 36.

[23] M. Matic, C.E. Pearce, J. Pečarić, Shannon's and related inequalities in information theory, in: Survey on Classical Inequalities, Springer, Dordrecht, 2000, pp. 127–164.

[24] N. Mehmood, R.P. Agarwal, S.I. Butt, J. Pečarić, New generalizations of Popoviciu-type inequalities via new Green's functions and Montgomery identity, J. Inequal. Appl. 2017 (1) (2017) 108.

[25] T. Niaz, K.A. Khan, J. Pečarić, On generalization of refinement of Jensen's inequality using Fink's identity and Abel-Gontscharoff Green function, J. Inequal. Appl. 2017 (1) (2017) 254.

[26] J. Pečarić, K.A. Khan, I. Perić, Generalization of Popoviciu type inequalities for symmetric means generated by convex functions, J. Math. Comput. Sci. 4 (6) (2014) 1091–1113.

[27] J. Pečarić, F. Proschan, Y.L. Tong, Convex Functions, Partial Orderings and Statistical Applications, Academic Press, New York, 1992.

[28] A. Rényi, On measure of information and entropy, in: Proceeding of the Fourth Berkely Symposium on Mathematics, Statistics and Probability, 1960, pp. 547–561.

[29] K.T. Rosen, M. Resnick, The size distribution of cities: an examination of the Pareto law and primacy, J. Urban Econ. 8 (2) (1980) 165–186.

[30] K.T. Soo, Zipf's Law for cities: a cross-country investigation, Reg. Sci. Urban Econ. 35 (3) (2005) 239–263.

[31] D.V. Widder, Completely convex function and Lidstone series, Trans. Amer. Math. Soc. 51 (1942) (1942) 387–398.

[32] G.K. Zipf, Human Behaviour and the Principle of Least-Effort, Addison-Wesley, Cambridge MA edn. Reading, 1949.

Corresponding author
Tasadduq Niaz can be contacted at: tasadduq_khan@yahoo.com