Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 45(2), 2018, Pages 243–257
ISSN: 1223-6934

Converses of the Edmundson-Lah-Ribarič inequality for generalized Csiszár divergence with applications to Zipf-Mandelbrot law

Rozarija Mikić, Đilda Pečarić, and Josip Pečarić

Abstract. In this paper we obtain some estimates for the generalized f-divergence functional via converses of the Jensen and Edmundson-Lah-Ribarič inequalities for convex functions, and then we obtain some estimates for the Kullback-Leibler divergence. All of the obtained results are applied to Zipf-Mandelbrot law and Zipf law.

2010 Mathematics Subject Classification. Primary 26D15; Secondary 28D20, 60E05.
Key words and phrases. Edmundson-Lah-Ribarič inequalitiy, convex functions, f-divergence, Kullback-Leibler divergence, Zipf-Mandelbrot law, Zipf law.

1. Introduction and preliminaries

Let us denote the set of all probability densities by $\mathbb{P}$, i.e. $P = (p_1, ..., p_n) \in \mathbb{P}$ if $p_i \in [0, 1]$ for $i = 1, ..., n$ and $\sum_{i=1}^{n} p_i = 1$. One of the numerous applications of Probability Theory is finding an appropriate measure of distance (difference or divergence) between two probability distributions.

Consequently, many different divergence measures have been introduced and extensively studied, for example Kullback-Liebler divergence, variation distance, Hellinger distance, χ^2-divergence, α-divergence, Bhattacharyya distance etc. All of the mentioned divergences are special cases of Csiszár f-divergence.

These measures of distance between two probability distributions have an important application in a great number of fields such as: anthropology, genetics, economics and political science, biology, approximation of probability distributions, signal processing and pattern recognition, ecological studies, music etc.

A large number of papers has been written on the subject of inequalities for different types of divergences. Since the functions that are used to define most of the divergences are convex, Jensen's inequality and its converses play an important role in the mentioned inequalities.

Theorem 1.1. *Let $f : I \to \mathbb{R}$ be a convex function, $x_i \in I$ for $i = 1, ..., n$, and let $p_1, ..., p_n$ be positive real numbers. Then we have*

$$f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \leq \frac{1}{P_n}\sum_{i=1}^{n} p_i f(x_i), \tag{1}$$

Received October 3, 2017. Accepted November 30, 2018.
The publication was supported by the Ministry of Education and Science of the Russian Federation (the Agreement number No. 02.a03.21.0008.)

where $P_n = \sum_{i=1}^{n} p_i$.

The following converse of the Jensen inequality is known as the Edmundson-Lah-Ribarič inequality for convex functions.

Theorem 1.2. [13] *Let* $f\colon [m, M] \to \mathbb{R}$ *be a convex function,* $x_i \in [m, M]$ *for* $i = 1, ..., n$*, and let* $p_1, ..., p_n$ *be positive real numbers. Then*

$$\frac{1}{P_n}\sum_{i=1}^{n} p_i f(x_i) \le \frac{M-\bar{x}}{M-m} f(m) + \frac{\bar{x}-m}{M-m} f(M), \tag{2}$$

where $P_n = \sum_{i=1}^{n} p_i$ *and* $\bar{x} = \frac{1}{P_n}\sum_{i=1}^{n} p_i x_i$.

Csiszár [1]-[2] introduced the $f-$divergence functional as

$$D_f(P, Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right), \tag{3}$$

where $f\colon [0, +\infty\rangle$ is a convex function, and it represent a "distance function" on the set of probability distributions $\mathbb{P}$. Many common divergences, such as previously mentioned Kullback-Liebler divergence, variation distance, Hellinger distance, χ^2-divergence, α-divergence, Bhattacharyya distance etc. are special cases of f-divergence, coinciding with a particular choice of the function f.

In order to use nonnegative probability distributions in the f-divergence functional, Horvath et. al. ([8]) defined

$$f(0) := \lim_{t\to 0+} f(t),\ 0\cdot f\left(\frac{0}{0}\right) := 0,\ 0\cdot f\left(\frac{a}{0}\right) := \lim_{t\to 0+} tf\left(\frac{a}{t}\right)$$

and gave the following definition of a generalized f-divergence functional.

Definition 1.1. Let $J \subset \mathbb{R}$ be an interval, and let $f\colon J \to \mathbb{R}$ be a function. Let $P = (p_1, ..., p_n)$ be an n-tuple of real numbers and $Q = (q_1, ..., q_n)$ be an n-tuple of nonnegative real numbers such that $p_i/q_i \in J$ for every $i = 1, ..., n$. Then let

$$\hat{D}_f(P, Q) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right).$$

Dragomir [4] gave the following upper bound for the Csiszár divergence functional

$$D_f(P, Q) \le \frac{M-1}{M-m} f(m) + \frac{1-m}{M-m} f(M), \tag{4}$$

where f is a convex function on the interval $[m, M]$, $P = (p_1, ..., p_n), Q = (q_1, ..., q_n) \in \mathbb{P}$ and $m \le p_i/q_i \le M$ for every $i = 1, ..., n$ (then it easily follows that $1 \in [m, M]$).

The Kullback-Leibler divergence, also called relative entropy or KL divergence

$$D_{KL}(P, Q) := \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right)$$

is a measure of the non-symmetric difference between two probability distributions P and Q, but it is not a true metric because it does not obey the triangle inequality and in general $D_{KL}(P, Q) \ne D_{KL}(Q, P)$. The KullbackLeibler divergence was introduced by Kullback and Leibler in [12], and it is a special case of the Csiszár divergence for $f(t) = t\log t$.

Aim of this paper is to obtain an improvement of Dragomir's result (4) for f-divergence and to prove difference type converses of the obtained results. Those results will then be applied to Zipf-Mandelbrot law since there are a lot applications of Zipf and Zipf-Mandelbrot laws.

First we state an improvement of the Edmundson-Lah-Ribarič inequality (2) proved by Klaričić Bakula, Pečarić and Perić in [9], which we will utilize to obtain an improvement of Dragomir's result (4).

Theorem 1.3. [9] *Let $f\colon [m,M] \to \mathbb{R}$ be a convex function, $x_i \in [m,M]$ for $i = 1,...,n$, and let $p_1,...,p_n$ be positive real numbers. Then*

$$\begin{aligned}\frac{1}{P_n}\sum_{i=1}^{n} p_i f(x_i) \le \frac{M-\bar{x}}{M-m} f(m) + \frac{\bar{x}-m}{M-m} f(M)&\\ - \frac{1}{P_n}\sum_{i=1}^{n} p_i \left(\frac{1}{2} - \frac{1}{M-m}\left|x_i - \frac{m+M}{2}\right|\right)\delta_f,&\qquad (5)\end{aligned}$$

where $\bar{x} = \frac{1}{P_n}\sum_{i=1}^{n} p_i x_i$ and

$$\delta_f = f(m) + f(M) - 2f\left(\frac{m+M}{2}\right). \qquad (6)$$

We also need to state discrete versions of two results found in [10].

Theorem 1.4. [10] *Let f be a continuous convex function on an interval I whose interior contains $[m,M]$, $x_i \in [m,M]$ for $i = 1,...,n$, and let $p_1,...,p_n$ be positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. Then*

$$\begin{aligned}0 &\le \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right)\\ &\le (M-\bar{x})(\bar{x}-m) \sup_{t\in\langle m,M\rangle} \Psi_f(t;m,M) - \left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n} p_i \left|x_i - \frac{m+M}{2}\right|\right)\delta_f\\ &\le \frac{(M-\bar{x})(\bar{x}-m)}{M-m}(f'_-(M) - f'_+(m)) - \left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n} p_i \left|x_i - \frac{m+M}{2}\right|\right)\delta_f\\ &\le \frac{1}{4}(M-m)(f'_-(M) - f'_+(m)) - \left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n} p_i \left|x_i - \frac{m+M}{2}\right|\right)\delta_f, \qquad (7)\end{aligned}$$

where δ_f is defined in (6) and $\Psi_f(\cdot;m,M)\colon \langle m,M\rangle \to \mathbb{R}$ is the second order divided difference of the function f at the points m, t and M for any $t \in \langle m,M\rangle$

$$\Psi_f(t;m,M) = \frac{1}{M-m}\left(\frac{f(M)-f(t)}{M-t} - \frac{f(t)-f(m)}{t-m}\right). \qquad (8)$$

Theorem 1.5. [10] *Let f be a continuous convex function on an interval I whose interior contains $[m,M]$, $x_i \in [m,M]$ for $i = 1,...,n$, and let $p_1,...,p_n$ be positive real* *numbers such that* $\sum_{i=1}^{n} p_i = 1$. *Then*

$$\begin{aligned}\left(\frac{1}{2}-\frac{1}{M-m}\sum_{i=1}^{n}p_i\left|x_i-\frac{m+M}{2}\right|\right)\delta_f \leq & \frac{M-\bar{x}}{M-m}f(m)+\frac{\bar{x}-m}{M-m}f(M)-\sum_{i=1}^{n}p_if(x_i)\\ \leq &(M-\bar{x})(\bar{x}-m)\sup_{t\in\langle m,M\rangle}\Psi_f(t;m,M)\\ \leq &\frac{(M-\bar{x})(\bar{x}-m)}{M-m}(f'_-(M)-f'_+(m))\\ \leq &\frac{1}{4}(M-m)(f'_-(M)-f'_+(m)). \qquad (9)\end{aligned}$$

2. Results

Our first result in this section is an improved version of Dragomir's result (4) for the generalized f-divergence functional, and it provides us an upper bound for the mentioned functional.

Theorem 2.1. *Let* $[m, M] \subset \mathbb{R}$ *be an interval, let* $f\colon [m, M] \to \mathbb{R}$ *be a function and let* δ_f *be defined in (6). Let* $P = (p_1, ..., p_n)$ *be an* n*-tuple of real numbers and* $Q = (q_1, ..., q_n)$ *be an* n*-tuple of nonnegative real numbers such that* $p_i/q_i \in [m, M]$ *for every* $i = 1, ..., n$*. If the function* f *is convex, we have*

$$\begin{aligned}\hat{D}_f(P,Q)\leq &\frac{MQ_n-P_n}{M-m}f(m)+\frac{P_n-mQ_n}{M-m}f(M)\\ &-\Big(\frac{Q_n}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i-\frac{m+M}{2}q_i\Big|\Big)\delta_f. \qquad (10)\end{aligned}$$

If the function f *is concave, then the inequality sign is reversed.*

Proof. Let $f\colon [m, M] \to \mathbb{R}$ be a convex function. Since $Q = (q_1, ..., q_n)$ are nonnegative real numbers, we can put

$$p_i = q_i \quad \text{and} \quad x_i = \frac{p_i}{q_i}$$

in Theorem 1.3 and get

$$\begin{aligned}\frac{1}{Q_n}\sum_{i=1}^{n}q_if\left(\frac{p_i}{q_i}\right)\leq &\frac{M-\frac{1}{Q_n}\sum_{i=1}^{n}q_i\frac{p_i}{q_i}}{M-m}f(m)+\frac{\frac{1}{Q_n}\sum_{i=1}^{n}q_i\frac{p_i}{q_i}-m}{M-m}f(M)\\ &-\frac{1}{Q_n}\left(\frac{Q_n}{2}-\frac{1}{M-m}\sum_{i=1}^{n}q_i\left|\frac{p_i}{q_i}-\frac{m+M}{2}\right|\right)\delta_f,\end{aligned}$$

and after multiplying by Q_n we get (10). □

Remark 2.1. From $m \leq p_i/q_i \leq M$ it easily follows that (see [9])

$$-\frac{M-m}{2}q_i\leq p_i-\frac{m+M}{2}q_i\leq\frac{M-m}{2}q_i,\quad \text{i.e.}\quad \left|p_i-\frac{m+M}{2}q_i\right|\leq\frac{M-m}{2}q_i$$

which together with $\delta_f \geq 0$ for a convex function f gives us

$$\Big(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i - \frac{m+M}{2}q_i\Big|\Big)\delta_f \geq 0.$$

Remark 2.2. If in the previous theorem we take P and Q to be probability distributions, we directly get an improvement of Dragomir's result for the Csiszár f-divergence functional:

$$D_f(P,Q) \leq \frac{M-1}{M-m}f(m) + \frac{1-m}{M-m}f(M) - \Big(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i - \frac{m+M}{2}q_i\Big|\Big)\delta_f.$$

Next result is a special case of Theorem 2.1, and provides us with bounds for the Kullback-leibler divergence of two probability distributions.

Corollary 2.2. *Let $[m, M] \subset \mathbb{R}$ be an interval and let us assume that the base of the logarithm is greater than* 1*.*

- *Let $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$ be n-tuples of nonnegative real numbers such that $p_i/q_i \in [m, M]$ for every $i = 1, ..., n$. Then*

$$\begin{aligned}\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) &\leq Q_n \frac{Mm}{M-m}\log\left(\frac{m}{M}\right) + \frac{P_n}{M-m}\log\left(\frac{M^M}{m^m}\right) \\ &- \Big(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i - \frac{m+M}{2}q_i\Big|\Big)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right). \qquad (11)\end{aligned}$$

- *Let $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n) \in \mathbb{P}$ be probability distributions such that $m \leq p_i/q_i \leq M$ holds for every $i = 1, ..., n$. Then*

$$\begin{aligned}D_{KL}(P,Q) &\leq \frac{Mm}{M-m}\log\left(\frac{m}{M}\right) + \frac{1}{M-m}\log\left(\frac{M^M}{m^m}\right) \\ &- \Big(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i - \frac{m+M}{2}q_i\Big|\Big)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right). \qquad (12)\end{aligned}$$

If the base of the logarithm is less than 1*, the inequality sign in the inequalities above is reversed.*

Proof. Let $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$ be an n-tuples of nonnegative real numbers. Since the function $t \mapsto t\log t$ is convex when the base of the logarithm is greater than 1, the inequality (11) follows from Theorem 2.1, inequality (10), by setting $f(t) = t\log t$.

Inequality (12) is a special case of the inequality (11) for probability distributions P and Q. □

Remark 2.3. If in Theorem 2.1, inequality (10), we set $f(t) = -\log t$ with the base greater than 1, we get a bound for the reversed Kullback-Leibler divergence of two probability distributions:

- for n-tuples of nonnegative real numbers $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$ such that $p_i/q_i \in [m, M]$ for every $i = 1, ..., n$ we have

$$\sum_{i=1}^{n} q_i \log\left(\frac{q_i}{p_i}\right) \leq \frac{P_n}{M-m}\log\frac{m}{M} + \frac{Q_n}{M-m}\log\frac{M^m}{m^M} - \Big(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i - \frac{m+M}{2}q_i\Big|\Big)\log\frac{(m+M)^2}{4mM}; \tag{13}$$

- for probability distributions $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n) \in \mathbb{P}$ such that $m \leq p_i/q_i \leq M$ holds for every $i = 1, ..., n$ we have

$$D_{KL}(Q, P) \leq \frac{1}{M-m}\log\left(\frac{M^{m-1}}{m^{M-1}}\right) - \Big(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\Big|p_i - \frac{m+M}{2}q_i\Big|\Big)\log\frac{(m+M)^2}{4mM}. \tag{14}$$

If the base of the logarithm is less than 1, the inequality sign in the inequalities above is reversed.

Our next result is obtained by utilizing Theorem 1.4, and it also gives us bounds for the generalized f-divergence functional. Concurrently, it represents an improvement of bounds for f-divergence functional obtained by Dragomir in the paper [4].

Theorem 2.3. *Let $I \subset \mathbb{R}$ be an interval such that its interior contains the interval $[m, M]$, let $f\colon I \to \mathbb{R}$ be a continuous function and let δ_f be defined in (6). Let $P = (p_1, ..., p_n)$ be an n-tuple of real numbers and $Q = (q_1, ..., q_n)$ be an n-tuple of nonnegative real numbers such that $p_i/q_i \in [m, M]$ for every $i = 1, ..., n$. Let δ_f be defined in (6), and Ψ_f in (8). If the function f is convex, then*

$$\begin{aligned}
0 &\leq \hat{D}_f(P, Q) - Q_n f\left(\frac{P_n}{Q_n}\right) \\
&\leq Q_n\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right)\sup_{t\in\langle m,M\rangle}\Psi_f(t; m, M) \\
&\quad - \left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\delta_f \\
&\leq \frac{Q_n}{M-m}\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right)\left(f'_-(M) - f'_+(m)\right) \\
&\quad - \left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\delta_f \\
&\leq \frac{Q_n}{4}(M-m)(f'_-(M) - f'_+(m)) - \left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\delta_f.
\end{aligned} \tag{15}$$

If the function f is concave, the inequality signs are reversed.

Proof. Let $f\colon [m, M] \to \mathbb{R}$ be a convex function. Since $Q = (q_1, ..., q_n)$ are nonnegative real numbers, we can put

$$p_i = \frac{q_i}{\sum_{i=1}^{n} q_i} = \frac{q_i}{Q_n} \quad \text{and} \quad x_i = \frac{p_i}{q_i}$$

in Theorem 1.4 and get

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i} f\left(\frac{p_i}{q_i}\right) - f\left(\sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\frac{p_i}{q_i}\right) \\
&\leq \left(M - \sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\frac{p_i}{q_i}\right)\left(\sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\frac{p_i}{q_i} - m\right) \sup_{t\in\langle m,M\rangle} \Psi_f(t;m,M) \\
&\quad - \left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\left|\frac{p_i}{q_i} - \frac{m+M}{2}\right|\right)\delta_f \\
&\leq \frac{f'_-(M) - f'_+(m)}{M-m}\left(M - \sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\frac{p_i}{q_i}\right)\left(\sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\frac{p_i}{q_i} - m\right) \\
&\quad - \left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\left|\frac{p_i}{q_i} - \frac{m+M}{2}\right|\right)\delta_f \\
&\leq \frac{1}{4}(M-m)(f'_-(M) - f'_+(m)) - \left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n} \frac{q_i}{\sum_{i=1}^{n} q_i}\left|\frac{p_i}{q_i} - \frac{m+M}{2}\right|\right)\delta_f,
\end{aligned}
$$

and after multiplying by Q_n we get (15). $\square$

The result that follows is a special cases of Theorem 2.3. It gives us different bounds of those that we have already obtained for the Kullback-Leibler divergence of two probability distributions.

Corollary 2.4. *Let $[m,M] \subset \mathbb{R}$ be an interval and let us assume that the base of the logarithm is greater than* 1.

- *Let $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$ be n-tuples of nonnegative real numbers such that $p_i/q_i \in [m,M]$ for every $i = 1, ..., n$. Then*

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{n} p_i \log\frac{p_i}{q_i} - P_n \log\left(\frac{P_n}{Q_n}\right) \\
&\leq Q_n\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right) \sup_{t\in\langle m,M\rangle} \Psi_{id\cdot\log}(t;m,M) \\
&\quad - \left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right) \\
&\leq \frac{Q_n}{M-m}\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right)\log\frac{M}{m} \\
&\quad - \left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right) \\
&\leq \frac{Q_n}{4}(M-m)\log\frac{M}{m} \\
&\quad - \left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right).
\end{aligned} \tag{16}
$$

- *Let $P=(p_1,...,p_n)$ and $Q=(q_1,...,q_n)\in\mathbb{P}$ be probability distributions such that $m\le p_i/q_i\le M$ holds for every $i=1,...,n$. Then*

$$\begin{aligned}
0 &\le D_{KL}(P,Q)\\
&\le (M-1)(1-m)\sup_{t\in\langle m,M\rangle}\Psi_{id\cdot\log}(t;m,M)\\
&\quad-\left(\frac{1}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\left|p_i-\frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M}+M\log\frac{2M}{m+M}\right)\\
&\le \frac{1}{M-m}(M-1)(1-m)\log\frac{M}{m}\\
&\quad-\left(\frac{1}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\left|p_i-\frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M}+M\log\frac{2M}{m+M}\right)\\
&\le \frac{1}{4}(M-m)\log\frac{M}{m}\\
&\quad-\left(\frac{1}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\left|p_i-\frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M}+M\log\frac{2M}{m+M}\right). \qquad (17)
\end{aligned}$$

If the base of the logarithm is less than 1, the inequality sign in the inequalities above is reversed.

Proof. Let $P=(p_1,...,p_n)$ and $Q=(q_1,...,q_n)$ be an n-tuples of nonnegative real numbers. Function $t\mapsto t\log t$ is convex, so inequality (16) follows from Theorem 2.3, inequality (15), by setting $f(t)=t\log t$.

Inequality (17) is a special case of the inequality (16) for probability distributions P and Q. □

Remark 2.4. If in Theorem 2.3, inequality (15), we set $f(t)=-\log t$ with the base greater than 1, we get the following:

- for n-tuples of nonnegative real numbers $P=(p_1,...,p_n)$ and $Q=(q_1,...,q_n)$ such that $p_i/q_i\in[m,M]$ for every $i=1,...,n$ we have

$$\begin{aligned}
0 &\le \sum_{i=1}^{n} q_i\log\left(\frac{q_i}{p_i}\right)+Q_n\log\left(\frac{P_n}{Q_n}\right)\\
&\le Q_n\left(M-\frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n}-m\right)\sup_{t\in\langle m,M\rangle}\Psi_{-\log}(t;m,M)\\
&\quad-\left(\frac{Q_n}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\left|p_i-\frac{m+M}{2}q_i\right|\right)\log\frac{(m+M)^2}{4mM}\\
&\le \frac{Q_n}{Mm}\left(M-\frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n}-m\right)\\
&\quad-\left(\frac{Q_n}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\left|p_i-\frac{m+M}{2}q_i\right|\right)\log\frac{(m+M)^2}{4mM}\\
&\le \frac{Q_n(M-m)^2}{4Mm}-\left(\frac{Q_n}{2}-\frac{1}{M-m}\sum_{i=1}^{n}\left|p_i-\frac{m+M}{2}q_i\right|\right)\log\frac{(m+M)^2}{4mM}. \qquad (18)
\end{aligned}$$

- for probability distributions $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n) \in \mathbb{P}$ such that $m \leq p_i/q_i \leq M$ holds for every $i = 1, ..., n$ we have

$$\begin{aligned}
0 &\leq D_{KL}(Q, P) \\
&\leq (M-1)(1-m) \sup_{t \in \langle m, M \rangle} \Psi_{-\log}(t; m, M) \\
&\quad - \left(\frac{1}{2} - \frac{1}{M-m} \sum_{i=1}^{n} \left| p_i - \frac{m+M}{2} q_i \right| \right) \log \frac{(m+M)^2}{4mM} \\
&\leq \frac{1}{Mm} (M-1)(1-m) \\
&\quad - \left(\frac{1}{2} - \frac{1}{M-m} \sum_{i=1}^{n} \left| p_i - \frac{m+M}{2} q_i \right| \right) \log \frac{(m+M)^2}{4mM} \\
&\leq \frac{(M-m)^2}{4Mm} - \left(\frac{1}{2} - \frac{1}{M-m} \sum_{i=1}^{n} \left| p_i - \frac{m+M}{2} q_i \right| \right) \log \frac{(m+M)^2}{4mM}. \qquad (19)
\end{aligned}$$

If the base of the logarithm is less than 1, the inequality sign in the inequalities above is reversed.

By making the same substitutions in Theorem 1.5 as in the proof of Theorem 2.3, we get lower and upper bounds for the difference in the results from Theorem 2.1, and consequently in Dragomir's result (4).

Theorem 2.5. *Let $I \subset \mathbb{R}$ be an interval such that its interior contains the interval $[m, M]$, let $f \colon I \to \mathbb{R}$ be a continuous function and let δ_f be defined in (6). Let $P = (p_1, ..., p_n)$ be an n-tuple of real numbers and $Q = (q_1, ..., q_n)$ be an n-tuple of nonnegative real numbers such that $p_i/q_i \in [m, M]$ for every $i = 1, ..., n$. Let δ_f be defined in (6), and Ψ_f in (8). If the function f is convex, then we have*

$$\begin{aligned}
&\left(\frac{Q_n}{2} - \frac{1}{M-m} \sum_{i=1}^{n} \left| p_i - \frac{m+M}{2} q_i \right| \right) \delta_f \\
&\quad \leq \frac{MQ_n - P_n}{M-m} f(m) + \frac{P_n - mQ_n}{M-m} f(M) - \hat{D}_f(P, Q) \\
&\quad \leq Q_n \left(M - \frac{P_n}{Q_n} \right) \left(\frac{P_n}{Q_n} - m \right) \sup_{t \in \langle m, M \rangle} \Psi_f(t; m, M) \\
&\quad \leq Q_n \left(M - \frac{P_n}{Q_n} \right) \left(\frac{P_n}{Q_n} - m \right) \frac{f'_-(M) - f'_+(m)}{M-m} \\
&\quad \leq \frac{Q_n}{4} (M-m)(f'_-(M) - f'_+(m)). \qquad (20)
\end{aligned}$$

If the function f is concave, the inequality signs are reversed.

We can utilize Theorem 2.5 to obtain lower and upper bounds for the difference in the results from Corollary 2.2, as well as for the inequalities from Remark 2.3.

Corollary 2.6. *Let $[m, M] \subset \mathbb{R}$ be an interval and let us assume that the base of the logarithm is greater than* 1.

- *Let $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$ be n-tuples of nonnegative real numbers such that $p_i/q_i \in [m, M]$ for every $i = 1, ..., n$. Then*

$$\begin{aligned}
&\left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right) \\
&\quad\le Q_n\frac{Mm}{M-m}\log\left(\frac{m}{M}\right) + \frac{P_n}{M-m}\log\left(\frac{M^M}{m^m}\right) - \sum_{i=1}^{n} p_i\log\left(\frac{p_i}{q_i}\right) \\
&\quad\le Q_n\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right)\sup_{t\in\langle m,M\rangle}\Psi_{id\cdot\log}(t; m, M) \\
&\quad\le \frac{Q_n}{M-m}\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right)\log\left(\frac{M}{m}\right) \le \frac{Q_n}{4}(M-m)\log\left(\frac{M}{m}\right). \qquad (21)
\end{aligned}$$

- *Let $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n) \in \mathbb{P}$ be probability distributions such that $m \le p_i/q_i \le M$ holds for every $i = 1, ..., n$. Then*

$$\begin{aligned}
&\left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\left(m\log\frac{2m}{m+M} + M\log\frac{2M}{m+M}\right) \\
&\quad\le \frac{Mm}{M-m}\log\left(\frac{m}{M}\right) + \frac{1}{M-m}\log\left(\frac{M^M}{m^m}\right) - D_{KL}(P, Q) \\
&\quad\le (M-1)(1-m)\sup_{t\in\langle m,M\rangle}\Psi_{id\cdot\log}(t; m, M) \\
&\quad\le \frac{1}{M-m}(M-1)(1-m)\log\left(\frac{M}{m}\right) \le \frac{1}{4}(M-m)\log\left(\frac{M}{m}\right). \qquad (22)
\end{aligned}$$

If the base of the logarithm is less than 1, the inequality sign in the inequalities above is reversed.

Remark 2.5. As in Remark 2.4, we can set $f(t) = -\log t$ with the base greater than 1 in Theorem 2.5, inequality (20), and obtain the following inequalities for the reversed Kullback-Leibler divergence:

- for n-tuples of nonnegative real numbers $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n)$ such that $p_i/q_i \in [m, M]$ for every $i = 1, ..., n$ we have

$$\begin{aligned}
&\left(\frac{Q_n}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\log\frac{(m+M)^2}{4mM} \\
&\quad\le \frac{Q_n}{M-m}\log\left(\frac{M^m}{m^M}\right) + \frac{P_n}{M-m}\log\left(\frac{m}{M}\right) - \sum_{i=1}^{n} q_i\log\left(\frac{q_i}{p_i}\right) \\
&\quad\le Q_n\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right)\sup_{t\in\langle m,M\rangle}\Psi_{\log}(t; m, M) \\
&\quad\le -\frac{Q_n}{Mm}\left(M - \frac{P_n}{Q_n}\right)\left(\frac{P_n}{Q_n} - m\right) \le -\frac{Q_n}{4Mm}(M-m)^2. \qquad (23)
\end{aligned}$$

- for probability distributions $P = (p_1, ..., p_n)$ and $Q = (q_1, ..., q_n) \in \mathbb{P}$ such that $m \leq p_i/q_i \leq M$ holds for every $i = 1, ..., n$ we have

$$\begin{aligned}
&\left(\frac{1}{2} - \frac{1}{M-m}\sum_{i=1}^{n}\left|p_i - \frac{m+M}{2}q_i\right|\right)\log\frac{(m+M)^2}{4mM} \\
&\leq \frac{1}{M-m}\log\left(\frac{M^{m-1}}{m^{M-1}}\right) - D_{KL}(Q,P) \\
&\leq (M-1)(1-m)\sup_{t\in\langle m,M\rangle}\Psi_{\log}(t;m,M) \\
&\leq -\frac{1}{Mm}(M-1)(1-m) \leq -\frac{1}{4Mm}(M-m)^2. \qquad (24)
\end{aligned}$$

If the base of the logarithm is less than 1, the inequality sign in the inequalities above is reversed.

3. Applications to Zipf-Mandelbrot law

ZipfMandelbrot law is a discrete probability distribution with parameters $N \in \mathbb{N}$, $q, s \in \mathbb{R}$ such that $q \geq 0$ and $s > 0$, possible values $\{1, 2, ..., N\}$ and probability mass function

$$f(i;N,q,s) = \frac{1/(i+q)^s}{H_{N,q,s}}, \quad \text{where} \quad H_{N,q,s} = \sum_{i=1}^{N}\frac{1}{(i+q)^s}. \qquad (25)$$

It is used in various scientific fields: linguistics [17], information sciences [5, 21], ecological field studies [16] and music [14]. Benoit Mandelbrot in 1966 gave improvement of Zipf law for the count of the low-rank words. Various scientific fields use this law for different purposes, for example information sciences use it for indexing [5, 21], ecological field studies in predictability of ecosystem [16], in music is used to determine aesthetically pleasing music [14].

Let P and Q be Zipf-Mandelbrot laws with parameters $N \in \mathbb{N}$, $q_1, q_2 \geq 0$ and $s_1, s_2 > 0$ respectively. We can use Corollary 2.4 and Corollary 2.6 in a similar way as described above in order to obtain inequalities for the Kullback-Leibler divergence. Let us denote

$$\begin{aligned}
m_{P,Q} &:= \min\left\{\frac{p_i}{q_i}\right\} = \frac{H_{N,q_2,s_2}}{H_{N,q_1,s_1}}\min\left\{\frac{(i+q_2)^{s_2}}{(i+q_1)^{s_1}}\right\} \\
M_{P,Q} &:= \max\left\{\frac{p_i}{q_i}\right\} = \frac{H_{N,q_2,s_2}}{H_{N,q_1,s_1}}\max\left\{\frac{(i+q_2)^{s_2}}{(i+q_1)^{s_1}}\right\} \qquad (26)
\end{aligned}$$

Corollary 3.1. *Let P and Q be Zipf-Mandelbrot laws with parameters $N \in \mathbb{N}$, $q_1, q_2 \geq 0$ and $s_1, s_2 > 0$ respectively. If the base of the logarithm is greater than one, we have*

$$\begin{aligned}
0 &\leq D_{KL}(P,Q) \\
&\leq (M_{P,Q}-1)(1-m_{P,Q})\sup_{t\in\langle m_{P,Q},M_{P,Q}\rangle}\Psi_{id\cdot\log}(t;m_{P,Q},M_{P,Q}) - \Delta_{P,Q}
\end{aligned}$$

$$\leq \frac{1}{M_{P,Q}-m_{P,Q}}\left(M_{P,Q}-1\right)\left(1-m_{P,Q}\right)\log\frac{M_{P,Q}}{m_{P,Q}}-\Delta_{P,Q}$$
$$\leq \frac{1}{4}(M_{P,Q}-m_{P,Q})\log\frac{M_{P,Q}}{m_{P,Q}}-\Delta_{P,Q} \tag{27}$$

and

$$\begin{aligned}\Delta_{P,Q} &\leq \frac{M_{P,Q}m_{P,Q}}{M_{P,Q}-m_{P,Q}}\log\left(\frac{m_{P,Q}}{M_{P,Q}}\right)+\frac{1}{M_{P,Q}-m_{P,Q}}\log\left(\frac{M_{P,Q}^{M_{P,Q}}}{m_{P,Q}^{m_{P,Q}}}\right)-D_{KL}(P,Q)\\ &\leq \left(M_{P,Q}-1\right)\left(1-m_{P,Q}\right)\sup_{t\in\langle m_{P,Q},M_{P,Q}\rangle}\Psi_{id\cdot\log}(t;m_{P,Q},M_{P,Q})\\ &\leq \frac{1}{M_{P,Q}-m_{P,Q}}\left(M_{P,Q}-1\right)\left(1-m_{P,Q}\right)\log\left(\frac{M_{P,Q}}{m_{P,Q}}\right)\\ &\leq \frac{1}{4}(M_{P,Q}-m_{P,Q})\log\left(\frac{M_{P,Q}}{m_{P,Q}}\right),\end{aligned} \tag{28}$$

where $D_{KL}(P,Q)$ is the Kullback-Leibler divergence of distributions P and Q, $m_{P,Q}$ and $M_{P,Q}$ are defined in (26), and

$$\begin{aligned}&\Delta_{P,Q}\\ &=\left(\frac{1}{2}-\frac{1}{M_{P,Q}-m_{P,Q}}\sum_{i=1}^{N}\left|\frac{1}{H_{N;q_1,s_1}(i+q_1)^{s_1}}-\frac{m_{P,Q}+M_{P,Q}}{2}\cdot\frac{1}{H_{N;q_2,s_2}(i+q_2)^{s_2}}\right|\right)\\ &\quad\times\left(m_{P,Q}\log\frac{2m_{P,Q}}{m_{P,Q}+M_{P,Q}}+M_{P,Q}\log\frac{2M_{P,Q}}{m_{P,Q}+M_{P,Q}}\right)\end{aligned}$$

Remark 3.1. If we utilize Remark 2.4 and Remark 2.5 in the same way as described above, we can obtain companion inequalities for the reversed Kullback-Leibler divergence $D_{KL}(Q,P)$ of these distributions.

For finite N and $q=0$ the Zipf-Mandelbrot law becomes Zipf's law. I is one of the basic laws in information science and bibliometrics, but it is also often used in linguistics. George Zipf's in 1932 found that we can count how many times each word appears in the text. So if we ranked (r) word according to the frequency of word occurrence (f), the product of these two numbers is a constant $C=r*f$. Same law in mathematical sense is also used in other scientific disciplines, but name of the law can be different, since regularities in different scientific fields are discovered independently from each other. In economics same law or regularity are called Pareto's law which analyze and predicts the distribution of the wealthiest members of the community [3]. The same type of distribution that we have in Zipf's and Pareto's law, also known as the Power law, can be found in wide variety of scientific disciplines, such as: physics, biology, earth and planetary sciences, computer science, demography and the social sciences [18] and many others. At this point of time we will not explain usage and their importance of this law in each scientific field, but we will retain on frequency of the word usage. Since, words are one of basic properties in human communication system. That frequency of used word and human communication system can be explained with plain mathematical formula is extremely interesting and useful in analysis of language and their usage. Since this law is be applicable in indexing and text mining, it is quite useful in today's world in which we use Internet to retrive most of the information that we need.

Probability mass function of Zipf's law is:

$$f(k;N,s)=\frac{1/k^s}{H_{N,s}},\quad \text{where}\quad H_{N,s}=\sum_{i=1}^{N}\frac{1}{i^s}. \tag{29}$$

Since Zipf's law is a special case of the ZipfMandelbrot law, all of the results from above hold for $q=0$.

Gelbukh and Sidorov in [6] observed the difference between the coefficients s_1 and s_2 in Zipf's law for the russian and english language. They processed 39 literature texts for each language, chosen randomly from different genres, with the requirement that the size be greater than 10,000 running words each. They calculated coefficients for each of the mentioned texts and as the result they obtained the average of $s_1 = 0,892869$ for the russian language, and $s_2 = 0,973863$ for the english language.

If we take $q_1 = q_2 = 0$, we can use the results from the above regarding the Kullback-Leibler divergence of two Zipf-Mandelbrot distributions in order to give estimates for the Kullback-Leibler divergence of the distributions associated to the russian and english language. For those experimental values of s_1 and s_2 we have

$$m_N=\min\left\{\frac{p_i}{q_i}\right\}=\frac{H_{N,s_2}}{H_{N,s_1}}\min\left\{\frac{i^{s_2}}{i^{s_1}}\right\}=\frac{H_{N,s_2}}{H_{N,s_1}}\min\left\{i^{s_2-s_1}\right\}=\frac{H_{N,s_2}}{H_{N,s_1}}$$

and

$$M_N=\max\left\{\frac{p_i}{q_i}\right\}=\frac{H_{N,s_2}}{H_{N,s_1}}\max\left\{\frac{i^{s_2}}{i^{s_1}}\right\}=\frac{H_{N,s_2}}{H_{N,s_1}}\max\left\{i^{s_2-s_1}\right\}=\frac{H_{N,s_2}}{H_{N,s_1}}N^{0,080994}.$$

Hence the following bounds for the mentioned divergence, depending only on the parameter N, hold.

$$\begin{aligned}
0&\le D_{KL}(P,Q)\\
&\le (M_N-1)(1-m_N)\sup_{t\in\langle m_N,M_N\rangle}\Psi_{id\cdot\log}(t;m_N,M_N)-\Delta_N\\
&\le \frac{0,080994}{M_N-m_N}(M_N-1)(1-m_N)\log N-\Delta_N\\
&\le 0,020249(M_N-m_N)\log N-\Delta_N
\end{aligned}$$

We also have

$$\begin{aligned}
\Delta_N \le&\frac{0,080994N^{0,080994}}{N^{0,080994}-1}\left(1-\frac{H_{N;0,973863}}{H_{N;0,892869}}\right)\log N+\log\left(\frac{H_{N;0,973863}}{H_{N;0,892869}}\right)-D_{KL}(P,Q)\\
\le&(M_N-1)(1-m_N)\sup_{t\in\langle m_N,M_N\rangle}\Psi_{id\cdot\log}(t;m_N,M_N)\\
\le&\frac{0,080994}{M_N-m_N}(M_N-1)(1-m_N)\log N\le 0,020249(M_N-m_N)\log N,
\end{aligned}$$

where

$$\begin{aligned}
\Delta_N=&\left(\frac{1}{2}-\frac{1}{H_{N;0,973863}(N^{0,080994}-1)}\sum_{i=1}^{N}\left|\frac{1}{i^{0,892869}}-\frac{N^{0,080994}+1}{2i^{0,973863}}\right|\right)\\
&\times\left(\log\frac{2}{N^{0,080994}+1}+N^{0,080994}\log\frac{2N^{0,080994}}{N^{0,080994}+1}\right)\frac{H_{N;0,973863}}{H_{N;0,892869}}
\end{aligned}$$

By calculating the above results for the Kullback-Leibler divergence of the distributions associated to the russian (P) and english (Q) language for different values of the parameter N, we obtained the following bounds:

- from the first series of inequalities:

N	5000	10000	50000	100000
$D_{KL}(P,Q) \leq$	0,0862934	0,100855	0,138862	0,157016

- from the second series of inequalities:

N	5000	10000	50000	100000
$D_{KL}(P,Q) \leq$	0,00106	0,001274	0,0018269	0,002091

The base of the logarithm used in our calculations is 2.

Acknowledgements

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

References

[1] I. Csiszár, Information measures: A critical survey, *Trans. 7th Prague Conf. on Info. Th. Statist. Decis. Funct., Random Processes and 8th European Meeting of Statist., Volume B*, Academia Prague (1978), 73–86.

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

[3] V. Diodato, *Dictionary of Bibliometrics*, Haworth Press, New York, 1994.

[4] S.S. Dragomir, Other inequalities for Csiszár divergence and applications, *RGMIA url: https://rgmia.org/papers/Csiszar/ICDApp.pdf*

[5] L. Egghe, R. Rousseau, *Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science*, Elsevier Science Publishers, New York, 1990.

[6] A. Gelbukh, G. Sidorov, Zipf and Heaps Laws Coefficients Depend on Language, *Proc. CICLing-2001, Conference on Intelligent Text Processing and Computational Linguistics*, Mexico City, February 18-24, 2001.

[7] G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities 1st ed. and 2nd ed.* Cambridge University Press, Cambridge, England, 1934, 1952.

[8] L. Horvath, G. Pečarić, J. Pečarić, Estimations of f- and Rènyi divergences by using a cyclic refinement of the Jensen's inequality, *J. Bull. Malays. Math. Sci Soc.* (2017), https://doi.org/10.1007/s40840-017-0526-4.

[9] M. Klaričić Bakula, J.E. Pečarić, J. Perić, On the converse Jensen inequality, *Applied Mathematics and Computation* **218** (2012), no. 11, 6566–6575.

[10] M. Krnić, R. Mikić, J. Pečarić, Strengthened converses of the Jensen and Edmundson-Lah-Ribarič inequalities, *Adv. Oper. Theory* **1** (2016), no. 1, 104–122.

[11] S. Kullback, *Information Theory and Statistics*, J. Wiley, New York, 1959.

[12] S. Kullback, R.A. Leibler, On information and sufficiency, *Annals of Mathematical Statistics* **22** (1951), no. 1, 79–86.

[13] P. Lah, M. Ribarič, Converse of Jensen's inequality for convex functions, *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz.* **412-460** (1973), 201–205.

[14] B. Manaris, D. Vaughan, C.S. Wagner, J. Romero, R.B. Davis, Evolutionary Music and the Zipf-Mandelbrot Law: Developing Fitness Functions for Pleasant Music, *Proceedings of 1st European Workshop on Evolutionary Music and Art (EvoMUSART2003)*, 522–534.

[15] D.S. Mitrinović, J.E. Pečarić, A.M. Fink, *Classical and New Inequalities in Analysis, Mathematics and its Applications (East European Series)*, Kluwer Academic Publishers Group, **61**, Dordrecht, 1993.

[16] D. Mouillot, A. Lepretre, Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity, *Environmental Monitoring and Assessment, Springer* **63** (2000), no. 2, 279–295.
[17] M.A. Montemurro, Beyond the Zipf-Mandelbrot law in quantitative linguistics, (2001) `arXiv:cond-mat/0104066v2` (22.2.2017.)
[18] M.E.J. Newman, Power laws, Pareto distributions and Zipfs law, `http://arxiv.org/pdf/cond-mat/0412004.pdf?origin=publication_detail` (6.4.2014.)
[19] J.E. Pečarić, F. Proschan, Y.L. Tong, *Convex functions, Partial orderings and statistical applications*, Academic Press Inc., San Diego, 1992.
[20] A. Renyi, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Statist. Prob., Vol. 1,* University of California Press, Berkeley, 1961.
[21] Z.K. Silagadze, Citations and the ZipfMandelbrot Law, *Complex Systems* **11** (1997), 487-499.

(R. Mikić) University of Zagreb, Faculty of Textile Technology, Prilaz baruna Filipovića 28a, 10 000 Zagreb, Croatia
E-mail address: `rozarija.jaksic@ttf.hr`

(Đ. Pečarić) Catholic University of Croatia, Ilica 242, 10 000 Zagreb, Croatia
E-mail address: `gildapeca@gmail.com`

(J. Pečarić) RUDN University, Miklukho-Maklaya str. 6, 117198 Moscow, Russia
E-mail address: `pecaric@element.hr`