Journal of Mathematical Inequalities
Volume 14, Number 4 (2020), 1321–1331 doi:10.7153/jmi-2020-14-85

SUPERQUADRATIC FUNCTIONS IN INFORMATION THEORY

JOSIPA BARIĆ, ĐILDA PEČARIĆ AND JOSIP PEČARIĆ

I dedicate this work to Professor Shoshana Abramovich on the occasion of her 80th birthday and I want to thank her for guidance, expertise, wisdom and support in my career. (*J. Barić*)

(*Communicated by G. Sinnamon*)

Abstract. Using Jensen's inequality and the converse Jensen's inequality for superquadratic functions we obtain new estimates for Shannon's entropy of the random variable X and derive new lower and upper bounds for the Shannon entropy in the terms of the Zipf and Zipf-Mandelbrot's law.

1. Introduction and preliminaries

We start with a brief overview of the class of superquadratic functions, some recent results derived from Jensen's inequality for convex functions and basic elements of information theory.

Superquadratic functions in one variable represent a generalization of the class of convex functions. Basic definitions and properties of superqudratic functions are recently introduced by S. Abramovich, G. Jameson and G. Sinnamon in [4] and [5]. Here we quote some definitions and theorems that we use in this paper. More examples and properties of superquadratic functions can be found in [1], [2], [3], [6] and its references.

DEFINITION 1. A function $\varphi : [0,\infty) \to \mathbb{R}$ is *superquadratic* provided that for all $x \geqslant 0$ there exists a constant $C(x) \in \mathbb{R}$ such that

$$\varphi(y) - \varphi(x) - \varphi(|y-x|) \geqslant C(x)(y-x), \tag{1}$$

for all $y \geqslant 0$. We say that f is *subquadratic* if $-f$ is superquadratic.

EXAMPLE 1.

a) The function $\varphi(x) = x^p$ is superquadratic for $p \geqslant 2$ and subquadratic for $p \in (0,2]$. If the function φ is defined by $\varphi(x) = \operatorname{sign}(p-2) \cdot x^p$ then φ is superquadratic for all $p > 0$.

Mathematics subject classification (2010): 94A15, 94A17, 26D15, 26A51.
Keywords and phrases: Superquadratic function, Jensen's inequality, Shannon entropy, Zipf-Mandelbrot's law.

Paper JMI-14-85

b) Any function $\varphi : [0,\infty) \to [-2,-1]$ is superquadratic (because it satisfies inequality (1) for $C(x) = 0$).

c) If φ is superquadratic and the numbers $a,b \geqslant 0$, the function $f(x) = \varphi(x) - (ax+b)$ is superquadratic too.

d) The function $\varphi(x) = x^2 \ln x$, where $x > 0$ and $\varphi(0) = 0$, is superquadratic (it follows from [4, Lemma 3.1]).

The following lemma shows that positive superquadratic functions are also convex functions.

LEMMA 1. *Let φ be a superquadratic function with $C(x)$ as in Definition 1. Then*

(i) $\varphi(0) \leqslant 0$.

(ii) If $\varphi(0) = \varphi'(0) = 0$ then $C(x) = \varphi'(x)$ wherever φ is differentiable at $x > 0$.

(iii) If $\varphi \geqslant 0$, then φ is convex and $\varphi(0) = \varphi'(0) = 0$.

The following inequality is well known in the literature as Jensen's inequality.

Let $f : X \to \mathbb{R}$ be a convex mapping defined on the linear space X and $x_i \in X$, $p_i \geqslant 0$, $i = 1,\dots,n$, $P = \sum_{i=1}^n p_i > 0$. Then

$$f\left(\frac{1}{P}\sum_{i=1}^n p_i x_i\right) \leqslant \frac{1}{P}\sum_{i=1}^n p_i f(x_i).$$

The generalization of Jensen's inequality to superquadratic functions is given in the following theorem (see [4]).

THEOREM 1.1. *Let φ be a superquadratic function and $x_i \geqslant 0$, $i = 1,\dots,n$, $\overline{x} = \frac{1}{P_n}\sum_{i=1}^n p_i x_i$, where $p_i \geqslant 0$ and $P_n = \sum_{i=1}^n p_i > 0$. Then*

$$\frac{1}{P_n}\sum_{i=1}^n p_i \varphi(x_i) - \varphi(\overline{x}) \geqslant \frac{1}{P_n}\sum_{i=1}^n p_i \varphi\left(|x_i - \overline{x}|\right). \tag{2}$$

The following converse Jensen's inequality for superqudratic functions is proved in [6].

THEOREM 1.2. *Let $\varphi : [0,\infty) \to \mathbb{R}$ be a superquadratic function. Let $(x_1,\dots,x_n) \in [m,M]^n$, $0 \leqslant m < M < \infty$ and $p_i \geqslant 0$, $i = 1,\dots,n$, $P_n = \sum_{i=1}^n p_i > 0$. Denote $\overline{x} = \frac{1}{P_n}\sum_{i=1}^n p_i x_i$. Then*

$$\begin{aligned} &\frac{1}{P_n}\sum_{i=1}^n p_i \varphi(x_i) \\ &\leqslant \frac{M-\overline{x}}{M-m}\varphi(m) + \frac{\overline{x}-m}{M-m}\varphi(M) \\ &\quad - \frac{1}{P_n(M-m)}\sum_{i=1}^n p_i\left((M-x_i)\varphi(x_i-m) + (x_i-m)\varphi(M-x_i)\right). \end{aligned} \tag{3}$$

2. New results

Information theory is a scientific field which studies informations as an instrument of communication. It is founded in 1948 by Claude Shannon who gave a mathematical definition of the concept of information in his paper "A Mathematical Theory of Communication", published in the Bell System Technical Journal. Shannon's quantity of information, called *entropy* and measured in bits, provided quantitative analysis of the relations which couldn't be exactly mathematicaly explained before.

Let us consider the Shannon entropy

$$H(X) = -\sum_{i=1}^{n} p_i \ln p_i, \tag{4}$$

where X is a random variable with the probability distribution $\mathbf{p} = (p_1, \dots, p_n)$.

Using Jensen's inequality, we get

$$H(X) = \sum_{i=1}^{n} p_i \ln \frac{1}{p_i} \leqslant \ln \left(\sum_{i=1}^{n} p_i \frac{1}{p_i} \right) = \ln n.$$

The following estimate is obtained from Jensen's inequality for superquadratic functions given in Theorem 1.1.

THEOREM 2.1. *Let X be a random variable with the probability distribution $\boldsymbol{p} = (p_1, \dots, p_n)$, with $p_i \geqslant 0$, $i = 1, \dots, n$. Then*

$$H(X) \leqslant \sum_{i=1}^{n} \frac{1}{p_i} (p_i - H_n(\boldsymbol{p}))^2 \ln |p_i - H_n(\boldsymbol{p})|^{-1} - n \cdot H_n(\boldsymbol{p}) \ln H_n(\boldsymbol{p}), \tag{5}$$

where $H_n(\boldsymbol{p}) = \dfrac{n}{S_n}$, for $S_n = \frac{1}{p_1} + \dots + \frac{1}{p_n}$, is the harmonic mean of positive real numbers $p_1, \dots, p_n$.

Proof. Taking in Theorem 1.1 superquadratic function $\varphi(x) = x^2 \ln x$ and replacing x_i by p_i and then p_i by $\frac{1}{p_i}$, $i = 1, \dots n$, we get

$$\frac{1}{S_n} \sum_{i=1}^{n} p_i \ln p_i - \frac{n^2}{S_n^2} \ln \frac{n}{S_n} \geqslant \frac{1}{S_n} \sum_{i=1}^{n} \frac{1}{p_i} \left(p_i - \frac{n}{S_n} \right)^2 \ln \left| p_i - \frac{n}{S_n} \right|.$$

Multiplying above inequality by S_n, we get

$$\sum_{i=1}^{n} p_i \ln p_i - \frac{n^2}{S_n} \ln \frac{n}{S_n} \geqslant \sum_{i=1}^{n} \frac{1}{p_i} \left(p_i - \frac{n}{S_n} \right)^2 \ln \left| p_i - \frac{n}{S_n} \right|.$$

Since, $H(X) = -\sum_{i=1}^{n} p_i \ln p_i$ and $H_n(\mathbf{p}) = \dfrac{n}{S_n}$, the proof is complete. □

In the next result, we obtain new estimate for Shannon's entropy of the random variable X using the converse Jensen's inequality for superqudratic functions.

THEOREM 2.2. *Let X be a random variable with the probability distribution $\boldsymbol{p} = (p_1,\ldots,p_n)$, with $0 \leqslant m \leqslant p_i \leqslant M < \infty$, $i = 1,\ldots,n$. Then*

$$H(X) \geqslant \frac{1}{M-m}\sum_{i=1}^{n}\frac{1}{p_i}\Big((M-p_i)(p_i-m)^2\ln(p_i-m) + (p_i-m)(M-p_i)^2\ln(M-p_i)\Big) - \frac{n}{H_n(\boldsymbol{p})}\left(\frac{M-H_n(\boldsymbol{p})}{M-m}m^2\ln m + \frac{H_n(\boldsymbol{p})-m}{M-m}M^2\ln M\right), \quad (6)$$

where $H_n(\boldsymbol{p})$ is the harmonic mean of positive real numbers $p_1,\ldots,p_n$.

Proof. Inequality (6) follows directly from Theorem 1.2 using $\varphi(x) = x^2\ln x$ and replacing x_i by p_i and then p_i by $\frac{1}{p_i}$, $i = 1,\ldots n$. $\square$

REMARK 1. Regarding the fact that the function $\varphi(x) = x^2\ln x$ is concave for $0 < x \leqslant e^{-\frac{3}{2}}$ and taking in Theorem 2.1 $p_i \leqslant e^{-\frac{3}{2}}$ for all i, we get

$$H(X) \geqslant -nH_n(\mathbf{p})\ln H_n(\mathbf{p}).$$

In the following inequality (see [11]) we have one of the basic results related to the Shannon entropy

$$\sum_{i=1}^{n} p_i \log\frac{1}{p_i} \leqslant \sum_{i=1}^{n} p_i \log\frac{1}{q_i}, \quad (7)$$

which holds for all positive real numbers p_i and q_i with $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$. Equality in (7) holds iff $q_i = p_i$, for all i. For details see [12, pp.635–650].

Inequality (7) can be written in the form

$$H(\mathbf{p}|\mathbf{q}) \geqslant 0, \quad (8)$$

where $H(\mathbf{p}|\mathbf{q})$ denotes the relative entropy of the probabilty $\mathbf{p}$ with respect to the probability $\mathbf{q}$, that is

$$H(\mathbf{p}|\mathbf{q}) = \sum_{i=1}^{n} p_i \log\frac{p_i}{q_i}.$$

Namely, by Jensen

$$H(\mathbf{p}|\mathbf{q}) = -\sum_{i=1}^{n} p_i \log\frac{q_i}{p_i} \geqslant -\log\left(\sum_{i=1}^{n} p_i\frac{q_i}{p_i}\right) = 0.$$

Now, we generalize inequality (7) using Jensen's inequality for superquadratic functions.

THEOREM 2.3. *Let $\{p_i, i=1,\dots,n\}$ be a set of positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. If $\{q_i, i=1,\dots,n\}$ is a set of positive real numbers with $\sum_{i=1}^{n} q_i = \alpha > 0$, then*

$$H(\boldsymbol{p}|\boldsymbol{q}) - \frac{\alpha^2 H_n(\boldsymbol{p},\boldsymbol{q}^2)}{\sum_{i=1}^{n} q_i^2} \ln \frac{\alpha H_n(\boldsymbol{p},\boldsymbol{q}^2)}{\sum_{i=1}^{n} q_i^2} \qquad (9)$$

$$\geqslant \sum_{i=1}^{n} \frac{1}{p_i} \left(p_i - \alpha q_i \frac{H_n(\boldsymbol{p},\boldsymbol{q}^2)}{\sum_{i=1}^{n} q_i^2} \right)^2 \ln \left| \frac{p_i}{q_i} - \frac{\alpha H_n(\boldsymbol{p},\boldsymbol{q}^2)}{\sum_{i=1}^{n} q_i^2} \right|,$$

where $H_n(\boldsymbol{p},\boldsymbol{q}^2)$ is weighted harmonic mean of positive numbers p_i with weights q_i^2.

Proof. Inequality (9) follows from inequality (2) replacing x_i by $\frac{p_i}{q_i}$ and p_i by $\frac{q_i^2}{p_i}$, $i=1,\dots n$. Then we have

$$\frac{1}{\sum_{i=1}^{n} \frac{q_i^2}{p_i}} \sum_{i=1}^{n} \frac{q_i^2}{p_i} \varphi\left(\frac{p_i}{q_i}\right) - \varphi\left(\frac{1}{\sum_{i=1}^{n} \frac{q_i^2}{p_i}} \cdot \sum_{i=1}^{n} \frac{q_i^2}{p_i} \cdot \frac{p_i}{q_i} \right)$$

$$\geqslant \frac{1}{\sum_{i=1}^{n} \frac{q_i^2}{p_i}} \sum_{i=1}^{n} \frac{q_i^2}{p_i} \varphi\left(\left| \frac{p_i}{q_i} - \frac{1}{\sum_{i=1}^{n} \frac{q_i^2}{p_i}} \cdot \sum_{i=1}^{n} \frac{q_i^2}{p_i} \cdot \frac{p_i}{q_i} \right| \right).$$

Applying this inequality to superquadratic function $\varphi(x) = x^2 \ln x$ and using the fact that $\sum_{i=1}^{n} q_i = \alpha > 0$ and $H_n(\mathbf{p},\mathbf{q}^2) = \frac{\sum_{i=1}^{n} q_i^2}{\sum_{i=1}^{n} \frac{q_i^2}{p_i}}$, we get inequality (9). □

REMARK 2. If in Theorem 2.3 we take $q_i = 1$, $i=1,\dots,n$, inequality (9) becomes equivalent to inequality (5).

In the following theorem we derive new upper bound for the relative entropy of the probabilty $\mathbf{p}$ with respect to the probability $\mathbf{q}$ using converse Jensen's inequality for superqudratic functions.

THEOREM 2.4. *Let $\{p_i, i=1,\dots,n\}$ be a set of positive real numbers such that $\sum_{i=1}^{n} p_i = 1$. If $\{q_i, i=1,\dots,n\}$ is a set of positive real numbers with $\sum_{i=1}^{n} q_i = \alpha > 0$ and*

$0 < m \leqslant \frac{p_i}{q_i} \leqslant M < \infty$, *then*

$$H(\boldsymbol{p}|\boldsymbol{q}) \leqslant \frac{\sum_{i=1}^{n} q_i^2}{H_n(\boldsymbol{p},\boldsymbol{q}^2)} \cdot \left(\frac{M - \frac{\alpha H_n(\boldsymbol{p},\boldsymbol{q}^2)}{\sum_{i=1}^{n} q_i^2}}{M-m} \cdot m^2 \ln m + \frac{\frac{\alpha H_n(\boldsymbol{p},\boldsymbol{q}^2)}{\sum_{i=1}^{n} q_i^2} - m}{M-m} \cdot M^2 \ln M \right)$$
$$- \frac{1}{M-m} \sum_{i=1}^{n} \frac{(Mq_i - p_i)(p_i - mq_i)}{p_i}$$
$$\cdot \left(\left(\frac{p_i}{q_i} - m \right) \ln \left| \frac{p_i}{q_i} - m \right| + \left(M - \frac{p_i}{q_i} \right) \ln \left| M - \frac{p_i}{q_i} \right| \right), \tag{10}$$

where $H_n(\boldsymbol{p},\boldsymbol{q}^2)$ *is weighted harmonic mean of positive numbers* p_i *with weights* q_i^2.

Proof. Inequality (10) follows from inequality (3) when we replace x_i by $\frac{p_i}{q_i}$, p_i by $\frac{q_i^2}{p_i}$, $i = 1, \ldots n$ and then aplly obtained inequality to superqudratic function $\varphi(x) = x^2 \ln x$. Using the fact that $\sum_{i=1}^{n} q_i = \alpha$ and $\frac{\sum_{i=1}^{n} q_i^2}{\sum_{i=1}^{n} \frac{q_i^2}{p_i}} = H_n(\mathbf{p},\mathbf{q}^2)$, we get inequality (10). □

REMARK 3. If in Theorem 2.4 we put $q_i = 1$, $i = 1, \ldots, n$, inequality (10) becomes equivalent to inequality (6).

2.1. Applications to the Zipf-Mendelbrot law

George Kingsley Zipf, (1902–1950), was an American linguist and philologist who investigated statistical regularity and the distribution of words frequencies in human languages. He concluded that a small number of words in text appears very often while many others occur rarely and formulated Zipf's law: if the words are ranked according to the frequency of occurance then the product of rank and frequency is a constant. Today Zipf's law is one of the basic laws in information theory and can be also found in the other scientific disciplines such as: economics (Pareto's law), physics, biology, earth and planetary sciences, computer science, demography and the social sciences (see, [7], [8]).

Benoit Mandelbrot in 1966 gave the improvement of Zipf's law for the count of the low ranked words (see [9], [10]).

In probability theory and statistics, the Zipf-Mendelbrot law is a discrete probability distribution defined by the probability mass function as follows.

$$f(i;N,t,s) = \frac{1}{(i+t)^s \cdot H_{N,t,s}}, \tag{11}$$

where N is an integer, $t \in [0,\infty)$, $s > 0$, $i \in \{1,2,\ldots,N\}$ and

$$H_{N,t,s} = \sum_{k=1}^{N} \frac{1}{(k+t)^s}$$

which may be thought of as a generalization of a harmonic number. In the formula (11), i is the rank of the data, and t and s are parameters of the distribution.

For finite N and for $t = 0$ the Zipf-Mandelbrot law becomes the Zipf law defined as follows.

$$f(i;N,s) = \frac{1}{i^s \cdot H_{N,s}}, \tag{12}$$

where

$$H_{N,s} = \sum_{k=1}^{N} \frac{1}{k^s},$$

which means that out of population of N elements the frequency of elements of rank i is $f(i;N,s)$, where s is the value of the exponent that characterizes the distribution.

Now, we apply previous results on the Zipf-Mandelbrot law.

If we define p_i, in the definition of the Shannon entropy (4), to be Zipf-Mandelbrot law, namely

$$p_i = f(i;N,t,s) = \frac{1}{(i+t)^s \cdot H_{N,t,s}},$$

then the Shannon entropy for the Zipf-Mandelbrot law becomes

$$\widetilde{H}(i;N,t,s) = \sum_{i=1}^{N} \frac{1}{(i+t)^s \cdot H_{N,t,s}} \ln\left((i+t)^s \cdot H_{N,t,s}\right), \tag{13}$$

where $N \in \mathbb{N}$, $t \geqslant 0$, $s > 0$.

Following statements provide new lower and upper estimates for the Shannon entropy in the terms of Zipf-Mandelbrot's law and it will be derived from previous results on the bounds for the Shannon entropy of the random variable X obtained by the properties of the superquadratic functions.

COROLLARY 2.1. *Let $\widetilde{H}(i;N,t,s)$ be Shannon's entropy in the terms of the Zipf-Mandelbrot law defined in (13) with parameters $N \in \mathbb{N}$, $t \in [0,\infty\rangle$, $s > 0$ and $i \in \{1,\ldots,N\}$. Then*

$$\begin{aligned}\widetilde{H}(i;N,t,s) \leqslant \sum_{i=1}^{N} \frac{(i+t)^s}{H_{N,t,s}} &\left(\frac{1}{(i+t)^s} - \frac{N}{\sum_{i=1}^{N}(i+t)^s}\right)^2 \\ &\cdot \ln H_{N,t,s} \cdot \left|\frac{1}{(i+t)^s} - \frac{N}{\sum_{i=1}^{N}(i+t)^s}\right|^{-1} \\ &- \frac{N^2}{H_{N,t,s}\sum_{i=1}^{N}(i+t)^s} \ln \frac{N}{H_{N,t,s}\sum_{i=1}^{N}(i+t)^s}.\end{aligned} \tag{14}$$

Proof. Inequality (14) follows directly from Theorem 2.1, replacing in inequality (5) positive real numbers p_i, of the probability distribution $\mathbf{p}$, by the probability mass function $f(i;N,t,s) = \dfrac{1}{(i+t)^s \cdot H_{N,t,s}}$ which defines Zipf-Mandelbrot's law. □

COROLLARY 2.2. *Let $\widetilde{H}(i;N,t,s)$ be Shannon's entropy in the terms of the Zipf-Mandelbrot law defined in (13) with parameters $N \in \mathbb{N}$, $t \in [0,\infty\rangle$, $s > 0$ and $i \in \{1,\ldots,N\}$. Let $0 \leqslant m \leqslant \frac{1}{(i+t)^s \cdot H_{N,t,s}} \leqslant M < \infty$, for all i. Then*

$$\begin{aligned}
&\widetilde{H}(i,N,t,s) \\
&\geqslant \frac{1}{M-m} \cdot \sum_{i=1}^{N} \Bigg[\left(M(i+t)^s \cdot H_{N,t,s} - 1\right) \left(\frac{1}{(i+t)^s \cdot H_{N,t,s}} - m \right)^2 \\
&\quad \cdot \ln \left| \frac{1}{(i+t)^s \cdot H_{N,t,s}} - m \right| \\
&\quad + \left(1 - m(i+t)^s \cdot H_{N,t,s}\right) \left(M - \frac{1}{(i+t)^s \cdot H_{N,t,s}} \right)^2 \ln \left| M - \frac{1}{(i+t)^s \cdot H_{N,t,s}} \right| \Bigg] \\
&\quad - \frac{1}{M-m} \Bigg[\left(M \cdot H_{N,t,s} \sum_{i=1}^{N} (i+t)^s - N \right) m^2 \ln m \\
&\quad + \left(N - m \cdot H_{N,t,s} \sum_{i=1}^{N} (i+t)^s \right) M^2 \ln M \Bigg]. \qquad (15)
\end{aligned}$$

Proof. Let positive real numbers p_i, $i = 1,\ldots,N$, of the probability distribution $\mathbf{p}$, in Theorem 2.2, be replaced by the probability mass function $f(i;N,t,s) = \dfrac{1}{(i+t)^s \cdot H_{N,t,s}}$. Then, inequality (15) follows directly from inequality (6). □

We can generalize previous results using the relative entropy of the probability $\mathbf{p}$ with respect to the probability $\mathbf{q}$, i.e. defining the relative entropy for the Zipf-Mandelbrot law.

If we take $\mathbf{p}$ and $\mathbf{q}$ both to be Zipf-Mandelbrot law N-tuples of the form

$$p_i = \frac{1}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}} \quad \text{and} \quad q_i = \frac{1}{(i+t_2)^{s_2} \cdot H_{N,t_2,s_2}},$$

where $N \in \mathbb{N}$, $t_1, t_2 \geqslant 0$, $s_1, s_2 > 0$ and $i = 1,\ldots,N$, then the relative entropy for Zipf-Mandelbrot law is

$$\widetilde{H}(i;N,t_1,s_1,t_2,s_2) = \sum_{i=1}^{N} \frac{1}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}} \log \frac{(i+t_2)^{s_2} \cdot H_{N,t_2,s_2}}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}}. \qquad (16)$$

Using new estimate for the relative Shannon entropy, given in inequality (9), we now derive generalized result for the relative entropy regarding the Zipf-Mandelbrot law.

COROLLARY 2.3. *Assume* $\widetilde{H}(i;N,t_1,s_1,t_2,s_2)$ *is relative entropy for the Zipf-Mandelbrot law defined in* (16)*, with* $N \in \mathbb{N}$*,* $t_1,t_2 \geqslant 0$*,* $s_1,s_2 > 0$ *and* $i = 1,\ldots,N$*. Then*

$$\widetilde{H}(i;N,t_1,s_1,t_2,s_2) - \frac{\alpha^2 \cdot H^2_{N,t_2,s_2}}{H_{N,t_1,s_1} \cdot \sum_{i=1}^{N} \frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}} \cdot \ln \frac{\alpha \cdot H^2_{N,t_2,s_2}}{H_{N,t_1,s_1} \cdot \sum_{i=1}^{N} \frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}}$$
$$\geqslant \sum_{i=1}^{N} \frac{(i+t_1)^{s_1}}{H_{N,t_1,s_1}} \left(\frac{1}{(i+t_1)^{s_1}} - \frac{\alpha \cdot H_{N,t_2,s_2}}{(i+t_2)^{s_2} \cdot \sum_{i=1}^{N} \frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}} \right)^2$$
$$\cdot \ln \frac{H_{N,t_2,s_2}}{H_{N,t_1,s_1}} \cdot \left| \frac{(i+t_2)^{s_2}}{(i+t_1)^{s_1}} - \frac{\alpha \cdot H_{N,t_2,s_2}}{\sum_{i=1}^{N} \frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}} \right|. \tag{17}$$

Proof. In order to apply new estimate for relative Shannon entropy, obtained in Theorem 2.3 by using Jensen's inequality for superquadratic functions, we need to determine new form of the $H_n(\mathbf{p},\mathbf{q}^2)$ from Theorem 2.3. Since we are now replacing p_i by $\frac{1}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}}$ and q_i by $\frac{1}{(i+t_2)^{s_2} \cdot H_{N,t_2,s_2}}$, we get for $H_n(\mathbf{p},\mathbf{q}^2)$ to became

$$\frac{H_{N,t_2,2s_2}}{H_{N,t_1,s_1}} \cdot \frac{1}{\sum_{i=1}^{N} \frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}}.$$

Inequality (17) now follows directly from inequality (9) of Theorem 2.3. □

In previous corollary we gave lower bound for the relative entropy related to Zipf-Mandelbrot law, applying Theorem 2.3. In our next result we obtain new upper bound for the relative entropy related to Zipf-Mandelbrot law using the upper bound for the relative entropy of the probabilty $\mathbf{p}$ with respect to the probability $\mathbf{q}$ obtained from converse Jensen's inequality for superqudratic functions.

COROLLARY 2.4. *Assume* $\widetilde{H}(i;N,t_1,s_1,t_2,s_2)$ *is relative entropy for the Zipf-Mandelbrot law defined in* (16)*, with* $N \in \mathbb{N}$*,* $t_1,t_2 \geqslant 0$*,* $s_1,s_2 > 0$ *and* $i = 1,\ldots,N$*. Let* $0 < m \leqslant \frac{(i+t_2)^{s_2} \cdot H_{N,t_2,s_2}}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}} \leqslant M < \infty$ *for all* i*. Then*

$$\widetilde{H}(i;N,t_1,s_1,t_2,s_2) \tag{18}$$
$$\leqslant R^{-1} \left(\frac{M-\alpha R}{M-m} m^2 \ln m + \frac{\alpha R - m}{M-m} M^2 \ln M \right)$$
$$- \frac{1}{M-m} \sum_{i=1}^{N} (i+t_1)^{s_1} \cdot H_{N,t_1,s_1} \cdot \left(\frac{M}{(i+t_2)^{s_2} \cdot H_{N,t_2,s_2}} - \frac{1}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}} \right)$$
$$\cdot \left(\frac{1}{(i+t_1)^{s_1} \cdot H_{N,t_1,s_1}} - \frac{m}{(i+t_2)^{s_2} \cdot H_{N,t_2,s_2}} \right)$$

$$\cdot\left(\left(\frac{(i+t_2)^{s_2}\cdot H_{N,t_2,s_2}}{(i+t_1)^{s_1}\cdot H_{N,t_1,s_1}}-m\right)\ln\left|\frac{(i+t_2)^{s_2}\cdot H_{N,t_2,s_2}}{(i+t_1)^{s_1}\cdot H_{N,t_1,s_1}}-m\right|\right.$$
$$\left.+\left(M-\frac{(i+t_2)^{s_2}\cdot H_{N,t_2,s_2}}{(i+t_1)^{s_1}\cdot H_{N,t_1,s_1}}\right)\ln\left|M-\frac{(i+t_2)^{s_2}\cdot H_{N,t_2,s_2}}{(i+t_1)^{s_1}\cdot H_{N,t_1,s_1}}\right|\right)$$

where R denotes following expression

$$R=\frac{\frac{H_{N,t_2,2s_2}}{H_{N,t_1,s_1}}\cdot\left(\sum_{i=1}^{N}\frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}\right)^{-1}}{\sum_{i=1}^{N}\frac{1}{(i+t_2)^{2s_2}\cdot H_{N,t_2,s_2}^2}}.$$

Proof. Inequality (18) follows from inequality (10) of Theorem 2.4 replacing p_i by $\frac{1}{(i+t_1)^{s_1}\cdot H_{N,t_1,s_1}}$ and q_i by $\frac{1}{(i+t_2)^{s_2}\cdot H_{N,t_2,s_2}}$, which yields $H_n(\mathbf{p},\mathbf{q}^2)$ from Theorem 2.4 to the following expression

$$\frac{H_{N,t_2,2s_2}}{H_{N,t_1,s_1}}\cdot\frac{1}{\sum_{i=1}^{N}\frac{(i+t_1)^{s_1}}{(i+t_2)^{2s_2}}}$$

and consequently new upper bound for the relative entropy related to Zipf-Mandelbrot law is derived. □

2.2. Applications to the Zipf law

If p_i, in definition of the Shannon entropy (4), is defined via the Zipf law, i.e.

$$p_i=f(i;N,s)=\frac{1}{i^s\cdot H_{N,s}},\tag{19}$$

where N is an integer, $s>0$, $i\in\{1,2,\ldots,N\}$ and

$$H_{N,s}=\sum_{k=1}^{N}\frac{1}{k^s},$$

then Shannon's entropy for the Zipf law becomes

$$\widetilde{H}\,(i;N,s)=\sum_{i=1}^{N}\frac{1}{i^s\cdot H_{N,s}}\ln\left(i^s\cdot H_{N,s}\right),\tag{20}$$

where $N\in\mathbb{N}$, $s>0$.

Also, if we assume that both $\mathbf{p}$ and $\mathbf{q}$ are Zipf law N-tuples of the form

$$p_i=\frac{1}{i^{s_1}\cdot H_{N,s_1}}\quad\text{and}\quad q_i=\frac{1}{i^{s_2}\cdot H_{N,s_2}},$$

where $N \in \mathbb{N}$, $i = 1, \ldots, N$ and $s_1, s_2 > 0$, then the relative entropy for the Zipf law is

$$\widetilde{H}(i; N, s_1, s_2) = \sum_{i=1}^{N} \frac{1}{i^{s_1} \cdot H_{N,s_1}} \log \frac{i^{s_2} \cdot H_{N,s_2}}{i^{s_1} \cdot H_{N,s_1}}. \tag{21}$$

Now, new estimates for Shannon's entropy and relative entropy concerning Zipf's law, can be easily obtained using estimates derived from the properties of superquadratic functions in Theorems 2.1 to 2.4.

REFERENCES

[1] S. ABRAMOVICH, J. BARIĆ, J. PEČARIĆ, *Fejer and Hermite-Hadamard type inequalities for superquadratic functions*, J. Math. Anal. Appl., 344, (2008), 1048–1056.

[2] S. ABRAMOVICH, J. BARIĆ, J. PEČARIĆ, *A variant of Jessen's inequality of Mercer's type for superquadratic functions*, JIPAM, 9 (3), (2008), article 62, 13 pp.

[3] S. ABRAMOVICH, J. BARIĆ, J. PEČARIĆ, *Superquadracity, Bohr's inequality and deviation from a mean value*, AJMAA, 7 (1), (2010), 1–9.

[4] S. ABRAMOVICH, G. JAMESON AND G. SINNAMON, *Refining Jensen's inequality*, Bull. Math. Soc. Sc. Math. Roumanie 47 (95), (2004), 3–14.

[5] S. ABRAMOVICH, G. JAMESON, G. SINNAMON, *Inequalities for averages of convex and superquadratic functions*, J. Inequal. Pure and Appl. Math., 5 (4), (2004), Article 91.

[6] S. BANIĆ, J. PEČARIĆ, S. VAROŠANEC, *Superquadratic functions and refinements of some classical inequalities*, J. Korean Math. Soc. 45 (2), (2008), 513–525.

[7] M. E. J. NEWMAN, *Power Laws, Pareto Distributions and Zipf's law*, Contemporary Physics, (2005) 46: 323–351.

[8] MARCELO A. MONTEMURRO, *A generalization of the Zipf-Mandelbrot law in linguistics*, Nonextensive Entropy, interdisciplinary applications, Oxford University Press, (2004).

[9] B. MANDELBROT, *The fractal structure of nature*, Freeman: New York, 1983.

[10] B. MANDELBROT, *An informational theory of the statistical structure of languange*, In Jackson, W. editor, Communication Theory, New Yord. Academic Press, (1953), 486–502.

[11] M. MATIĆ, C. E. PEARCE, J. PEČARIĆ, *Shannon's and related inequalities in information theory*, Survey on Classical Inequalities, Springer, Dordrecht, (2000), 127–164.

[12] D. S. MITRINOVIĆ, J. E. PEČARIĆ, A. M. FINK, *Classical and New Inequalities in Analysis*, Kluwer Academic Publishers, Dordrecht/Boston/London, (1993).

(Received December 4, 2019)

Josipa Barić
University of Split
Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture
Ruđera Boškovića 32, 21 000 Split, Croatia
e-mail: jbaric@fesb.hr

Đilda Pečarić
Catholic University of Croatia
Ilica 242, 10000 Zagreb, Croatia
e-mail: gildapeca@gmail.com

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