

Article

New Estimations for Shannon and Zipf–Mandelbrot Entropies

Muhammad Adil Khan ^{1,2}, Zaid Mohammad Al-sahwi ³ and Yu-Ming Chu ^{4,*}¹ College of Science, Hunan City University, Yiyang 413000, China; madilkhan@uop.edu.pk² Department of Mathematics, University of Peshawar, Peshawar 25000, Pakistan³ Department of Mathematics, University of Sa'adah, Sa'adah 1872, Yemen; alsahwizaid@gmail.com⁴ Department of Mathematics, Huzhou University, Huzhou 313000, China

* Correspondence: chuyuming2005@126.com; Tel.: +86-572-2321163

Received: 9 July 2018; Accepted: 14 August 2018; Published: 16 August 2018



Abstract: The main purpose of this paper is to find new estimations for the Shannon and Zipf–Mandelbrot entropies. We apply some refinements of the Jensen inequality to obtain different bounds for these entropies. Initially, we use a precise convex function in the refinement of the Jensen inequality and then tamper the weight and domain of the function to obtain general bounds for the Shannon entropy (SE). As particular cases of these general bounds, we derive some bounds for the Shannon entropy (SE) which are, in fact, the applications of some other well-known refinements of the Jensen inequality. Finally, we derive different estimations for the Zipf–Mandelbrot entropy (ZME) by using the new bounds of the Shannon entropy for the Zipf–Mandelbrot law (ZML). We also discuss particular cases and the bounds related to two different parametrics of the Zipf–Mandelbrot entropy. At the end of the paper we give some applications in linguistics.

Keywords: Jensen inequality; convex function; Shannon entropy; Zipf–Mandelbrot entropy

PACS: 26D15

1. Introduction

The idea of the Shannon entropy [1] plays a key role in information theory, while in some cases, it is denoted as measure of uncertainty. There are basically two methods for understanding the Shannon entropy. Under one point of view, the Shannon entropy quantifies the amount of information in regard to the value of X (after measurement). Under another point of view, the Shannon entropy tells us the amount of uncertainty about the variable of X before we learn its value (before measurement) [2]. The random variable, entropy, is characterized regarding its probability distribution and it can appear as a better measure of predictability or uncertainty. SE permits the appraisal of the normal least number of bits expected to encode a series of symbols based on the letters in order of estimation and the recurrence of the symbols. The formula for SE is given by [1]

$$S(\Psi) = \sum_{i=1}^n \psi_i \log \left(\frac{1}{\psi_i} \right), \quad (1)$$

where $\psi_1, \psi_2, \dots, \psi_n \in \mathbb{R}^+$ with $\sum_{i=1}^n \psi_i = 1$.

There are many applications of the Shannon entropy in most applied sciences and in other sciences, such as biology [3], genomic geography [4], and finance [5]. Currently, the Shannon entropy is applied in the simulation of laser dynamics and as an objective measure to evaluate models and compare observational results [6,7].

In 1932, George Zipf gave the idea that the size of the r 'th largest occurrence of an event is inversely proportional to its rank. That is, this law states that $P_r = 1/r^b$, where P_r is the frequency of occurrence of the r 'th ranked and b is close to unity. As in linguistics, Zipf found that one can calculate the number of times each word appears in the text. Therefore, if the rank (γ) of the word is in accordance with the frequency of the word's appearance (ρ), then the product of these two numbers is a constant ((C): $C = \rho \cdot \gamma$, see [8,9]).

There are several applications of the Zipf law, and here we present some of them. This law has been used in city populations. Kristian Giesen and Jens Suedekum conducted a study on the city measure distributions of single German districts' reliance on researching the phenomenon [10]. They built their study based on the intuition by Gabaix (1999) which states that the Zipf law takes after an irregular development process. This means that Gabaix displays that if the districts follow the Gibrat law, they should notice the Zipf law at both the districts at a national level. By utilizing non-parametric procedures, they found that the Gibrat law holds in each German district, regardless of how "districts" are defined. To put it differently, the Gibrat and Zipf laws are inclined to hold ubiquitously in space. In geology, the Zipf law has been used with temperate prosperity in the resource estimation of extracting ores from the ground and petroleum [11]. In principle, it forecasts how many entities of a confident size can be left in a sequence of decreasing size, assuming the largest has been established. The solar flare intensity (M. E. J. Newman, 2004) [12] represents the cumulative distribution of the vertex gamma-ray density of solar flares, for which perceptions were made between 1980 and 1989 by the well-known HardX-Ray fulmination spectrometer onboard the solar maximum mission satellite launched in 1980. The spectrometer uses a CsI gleaming discloser to measure gamma-rays from solar flares. For website traffic (Shane Parkins, 2015) [13], the Zipf law seems, by all accounts, to be the control as opposed to the special case. It is available at the level of routers that transmit data from one geographic location to another and in the content of the World Wide Web. At the social and economic levels, it also determines how persons choose the sites they visit and form peer-to-peer societies. The omnipresent nature of the Zipf law in cyberspace is geared toward deeper empathy with the internet phenomena, for example, discovering the potential of prevalence proxy caches in divergent Autonomous Systems (ASes) with the purpose of reducing the costs incurred by internet service providers and pacification of the load on the internet backbone [14].

It was determined that Zipf's law can describe the size and rank distribution of earthquakes, including those with magnitude, but it cannot predict when they will occur. In the earth-moon the crater size-frequency distribution can be represented by the Zipf law [12,15].

In 1966, Benoit Mandelbrot gave an enhancement for the Zipf law, known as ZML, which gives a generalization of the account of the low-rank words in corpus [16]:

$$g(i) = \frac{c}{(i+h)^r},$$

where $i < 1000, r, c > 0$ and if $h = 0$, we get the Zipf law.

If $n \in \mathbb{N}, h \geq 0, r > 0$ and $i \in \{1, 2, \dots, n\}$, then the Zipf–Mandelbrot law (probability mass function) is defined by

$$G(i, n, h, r) = \frac{1/(i+h)^r}{Q_{n,h,r}}.$$

The formula for ZME is given by

$$Z(Q, h, r) = \frac{r}{Q_{n,h,r}} \sum_{i=1}^n \frac{\log(i+h)}{(i+h)^r} + \log Q_{n,h,r}, \tag{2}$$

where $Q_{n,h,r} = \sum_{i=1}^n \frac{1}{(i+h)^r}$.

There are many applications of ZML which can be found in linguistics [16,17], ecological field studies [18], and information sciences [9]. Recently, the Zipf–Mandelbrot law was applied to various

types of f -divergences and distances, for example Kullback–Leibler divergence, Bhattacharyya distance (via coefficient), Hellinger distance, x^2 -divergence, etc [19].

To complete this section, we give some notions and results from ref. [20].

Let $g : \mathcal{G} \rightarrow R$ be a convex function defined on the convex set \mathcal{G} , $T_n = \{1, 2, 3, \dots, n\}$. Let s be fixed positive integer and l be all those positive integers, such that $1 \leq l \leq s \leq n$. Suppose $\mathcal{M}_1^s, \mathcal{M}_2^s, \dots, \mathcal{M}_l^s$ represents any subsets of T_n , such that $\mathcal{M}_1^s \cup \mathcal{M}_2^s \dots \cup \mathcal{M}_l^s = T_n$ and $\mathcal{M}_i^s \cap \mathcal{M}_j^s = \emptyset$ for $i \neq j$, where $i, j \in T_n$, $x_i \in \mathcal{G}$, $\psi_i > 0$ with $\sum_{i=1}^n \psi_i = 1$, and for any $\mathcal{M} \subseteq \{1, 2, \dots, n\}$, $\mathcal{M} \neq \emptyset$, we represent $\Psi_{\mathcal{M}} =: \sum_{i \in \mathcal{M}} \psi_i$ and $\Omega_{\mathcal{M}} =: \sum_{i \in \mathcal{M}} \omega_i$ for positive real numbers ψ_i, ω_i . For the convex function g and the n -tuples $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\psi = (\psi_1, \psi_2, \dots, \psi_n)$, the following functional is defined:

$$A_s = \max_{\mathcal{M}_1^s, \mathcal{M}_2^s, \dots, \mathcal{M}_l^s} \left[\Psi_{\mathcal{M}_1^s} g\left(\frac{1}{\Psi_{\mathcal{M}_1^s}} \sum_{i \in \mathcal{M}_1^s} \psi_i x_i\right) + \Psi_{\mathcal{M}_2^s} g\left(\frac{1}{\Psi_{\mathcal{M}_2^s}} \sum_{i \in \mathcal{M}_2^s} \psi_i x_i\right) + \dots + \Psi_{\mathcal{M}_l^s} g\left(\frac{1}{\Psi_{\mathcal{M}_l^s}} \sum_{i \in \mathcal{M}_l^s} \psi_i x_i\right) \right]. \tag{3}$$

Particularly, for $s = 2, 3$, we have

$$A_2 = \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left[\Psi_{\mathcal{M}_1^2} g\left(\frac{1}{\Psi_{\mathcal{M}_1^2}} \sum_{i \in \mathcal{M}_1^2} \psi_i x_i\right) + \Psi_{\mathcal{M}_2^2} g\left(\frac{1}{\Psi_{\mathcal{M}_2^2}} \sum_{i \in \mathcal{M}_2^2} \psi_i x_i\right) \right]$$

$$A_3 = \max_{\mathcal{M}_1^3, \mathcal{M}_2^3, \mathcal{M}_3^3} \left[\Psi_{\mathcal{M}_1^3} g\left(\frac{1}{\Psi_{\mathcal{M}_1^3}} \sum_{i \in \mathcal{M}_1^3} \psi_i x_i\right) + \Psi_{\mathcal{M}_2^3} g\left(\frac{1}{\Psi_{\mathcal{M}_2^3}} \sum_{i \in \mathcal{M}_2^3} \psi_i x_i\right) + \Psi_{\mathcal{M}_3^3} g\left(\frac{1}{\Psi_{\mathcal{M}_3^3}} \sum_{i \in \mathcal{M}_3^3} \psi_i x_i\right) \right].$$

Analogously, for other particular values of s with $1 \leq l \leq s \leq n$, one can obtain different functionals.

The following generalized refinements of the Jensen inequality were given in refs. [20,21].

Theorem 1 ([20]). Let $g : \mathcal{M} \rightarrow R$ be a convex function defined on the convex set \mathcal{M} . If $x_i \in \mathcal{M}$ and $\psi_i > 0$ for $i, k \in T_n$ with $\sum_{i=1}^n \psi_i = 1$, then we have

$$\sum_{i=1}^n \psi_i g(x_i) \geq A_k \geq A_{k-1} \geq \dots \geq A_3 \geq A_2 \geq g\left(\sum_{i=1}^n \psi_i x_i\right). \tag{4}$$

Theorem 2 ([21]). Let $g : \mathcal{M} \rightarrow R$ be a convex function defined on the convex set \mathcal{M} , $x_j \in \mathcal{M}$, and $\psi_j > 0$ for $j \in \{1, \dots, n\}$ with $\sum_{j=1}^n \psi_j = 1$. Then,

$$g\left(\sum_{j=1}^n \psi_j x_j\right) \leq \min_{k \in T_n} \left[(1 - \psi_k) g\left(\frac{\sum_{j=1}^n \psi_j x_j - \psi_k x_k}{1 - \psi_k}\right) + \psi_k g(x_k) \right]$$

$$\leq \frac{1}{n} \left[\sum_{k=1}^n (1 - \psi_k) g\left(\frac{\sum_{j=1}^n \psi_j x_j - \psi_k x_k}{1 - \psi_k}\right) + \sum_{k=1}^n \psi_k g(x_k) \right]$$

$$\leq \max_{k \in T_n} \left[(1 - \psi_k) g\left(\frac{\sum_{j=1}^n \psi_j x_j - \psi_k x_k}{1 - \psi_k}\right) + \psi_k g(x_k) \right]$$

$$\leq \sum_{j=1}^n \psi_j g(x_j). \tag{5}$$

For some other results related to the Jensen inequality and the Shannon and Zipf–Mandelbrot entropies, see refs. [8,22–24].

Due to the great importance of the Shannon and Zipf–Mandelbrot entropies, many results are devoted to these entropies in the literature. The main focus of this paper was to associate some refinements of the Jensen inequality to the Shannon and Zipf–Mandelbrot entropies. In this paper, we use the main results given in ref. [20] and obtain some estimations for these entropies. We also discuss some particular cases of these results. At the end of the paper, we give some applications in

linguistics. The idea of this paper can be applied for other results of the Jensen inequality to obtain new estimations for these entropies.

2. Estimations for the Shannon Entropy

We start by giving our first main result for the Shannon entropy.

Theorem 3. Let $\psi_i, \omega_i \in R^+$, where $i = 1, 2, \dots, n$, with $\sum_{i=1}^n \psi_i = 1$. Then, the following inequalities hold:

$$\begin{aligned}
 -S(\Psi) - \sum_{i=1}^n \psi_i \log \omega_i &\geq \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_i^k} \left\{ \Psi_{\mathcal{M}_1^k} \log \left(\frac{\Psi_{\mathcal{M}_1^k}}{\Omega_{\mathcal{M}_1^k}} \right) + \dots + \Psi_{\mathcal{M}_i^k} \log \left(\frac{\Psi_{\mathcal{M}_i^k}}{\Omega_{\mathcal{M}_i^k}} \right) \right\} \\
 &\geq \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_{i-1}^k} \left\{ \Psi_{\mathcal{M}_1^k} \log \left(\frac{\Psi_{\mathcal{M}_1^k}}{\Omega_{\mathcal{M}_1^k}} \right) + \dots + \Psi_{\mathcal{M}_{i-1}^k} \log \left(\frac{\Psi_{\mathcal{M}_{i-1}^k}}{\Omega_{\mathcal{M}_{i-1}^k}} \right) \right\} \\
 &\geq \dots \geq \log \left(\frac{1}{\Omega_{T_n}} \right).
 \end{aligned} \tag{6}$$

Proof. If we take $g(x) = -\log x$, $x_i = \frac{\omega_i}{\psi_i}$, $i \in T_n$ in (4), then we obtain

$$\begin{aligned}
 \sum_{i=1}^n \psi_i g(x_i) &= -\sum_{i=1}^n \psi_i \log \left(\frac{\omega_i}{\psi_i} \right) = \sum_{i=1}^n \psi_i \log \psi_i - \sum_{i=1}^n \psi_i \log \omega_i \\
 &= -S(\Psi) - \sum_{i=1}^n \psi_i \log \omega_i
 \end{aligned}$$

and

$$\begin{aligned}
 A_k &= \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_i^k} \left\{ \Psi_{\mathcal{M}_1^k} \log \left(\frac{\Psi_{\mathcal{M}_1^k}}{\Omega_{\mathcal{M}_1^k}} \right) + \dots + \Psi_{\mathcal{M}_i^k} \log \left(\frac{\Psi_{\mathcal{M}_i^k}}{\Omega_{\mathcal{M}_i^k}} \right) \right\}. \\
 A_{k-1} &= \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_{i-1}^k} \left\{ \Psi_{\mathcal{M}_1^k} \log \left(\frac{\Psi_{\mathcal{M}_1^k}}{\Omega_{\mathcal{M}_1^k}} \right) + \dots + \Psi_{\mathcal{M}_{i-1}^k} \log \left(\frac{\Psi_{\mathcal{M}_{i-1}^k}}{\Omega_{\mathcal{M}_{i-1}^k}} \right) \right\} \\
 &\vdots \\
 A_3 &= \max_{\mathcal{M}_1^3, \mathcal{M}_2^3, \mathcal{M}_3^3} \left\{ \Psi_{\mathcal{M}_1^3} \log \left(\frac{\Psi_{\mathcal{M}_1^3}}{\Omega_{\mathcal{M}_1^3}} \right) + \Psi_{\mathcal{M}_2^3} \log \left(\frac{\Psi_{\mathcal{M}_2^3}}{\Omega_{\mathcal{M}_2^3}} \right) + \Psi_{\mathcal{M}_3^3} \log \left(\frac{\Psi_{\mathcal{M}_3^3}}{\Omega_{\mathcal{M}_3^3}} \right) \right\}. \\
 A_2 &= \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \Psi_{\mathcal{M}_1^2} \log \left(\frac{\Psi_{\mathcal{M}_1^2}}{\Omega_{\mathcal{M}_1^2}} \right) + \Psi_{\mathcal{M}_2^2} \log \left(\frac{\Psi_{\mathcal{M}_2^2}}{\Omega_{\mathcal{M}_2^2}} \right) \right\}.
 \end{aligned}$$

Therefore, from (4), we deduce (6). \square

Corollary 1. Let $\psi_i, \omega_i \in R^+$, where $i = 1, 2, \dots, n$, with $\sum_{i=1}^n \psi_i = 1$, then

$$-S(\Psi) - \sum_{i=1}^n \psi_i \log \omega_i \geq \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \Psi_{\mathcal{M}_1^2} \log \left(\frac{\Psi_{\mathcal{M}_1^2}}{\Omega_{\mathcal{M}_1^2}} \right) + \Psi_{\mathcal{M}_2^2} \log \left(\frac{\Psi_{\mathcal{M}_2^2}}{\Omega_{\mathcal{M}_2^2}} \right) \right\} \geq \log \left(\frac{1}{\Omega_{T_n}} \right). \tag{7}$$

Proof. By taking $k = 2$ in (6), we obtain (7). \square

Corollary 2. Let $\psi_i, \omega_i \in R^+$, $i \in T_n$ with $\sum_{i=1}^n \psi_i = 1$, then

$$\begin{aligned}
 -S(\Psi) - \sum_{i=1}^n \psi_i \log \omega_i &\geq \max_{k \in T_n} \left\{ \psi_k \log \left(\frac{\psi_k}{\omega_k} \right) + (1 - \psi_k) \log \left(\frac{1 - \psi_k}{\sum_{i=1}^n \omega_i - \omega_k} \right) \right\} \\
 &\geq \log \left(\frac{1}{\Omega_{T_n}} \right).
 \end{aligned} \tag{8}$$

Proof. If we take $\mathcal{M}_1^2 = \{k\}$ in (7), then obviously, $\mathcal{M}_2^2 = \{1, 2, \dots, n\} \setminus \{k\}$ and $\Psi_{\mathcal{M}_1^2} = \psi_k, \Psi_{\mathcal{M}_2^2} = 1 - \psi_k$. Hence, from (7), we deduce (8). \square

Remark 1. Note that Corollary 1 is, in fact, the application of ([25], Theorem 1) and Corollary 2 is the application of ([21], Theorem 1).

In the following corollary, we discuss another particular case of Theorem 3.

Corollary 3. Let $\psi_i, \omega_i \in R^+$ for $i \in T_n$ with $\sum_{i=1}^n \psi_i = 1$, then

$$-S(\Psi) \geq \Psi_{\mathcal{M}_1^2} \log \left(\frac{\Psi_{\mathcal{M}_1^2}}{n} \right) + \Psi_{\mathcal{M}_2^2} \log \left(\frac{\Psi_{\mathcal{M}_2^2}}{n} \right) \geq 0. \tag{9}$$

Proof. By taking $\omega_i = 1$ in (7), $i = 1, 2, \dots, n$, we get (9). \square

Remark 2. It is obvious that

$$\begin{aligned} & \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \Psi_{\mathcal{M}_1^2} \log \left(\frac{\Psi_{\mathcal{M}_1^2}}{\Omega_{\mathcal{M}_1^2}} \right) + \Psi_{\mathcal{M}_2^2} \log \left(\frac{\Psi_{\mathcal{M}_2^2}}{\Omega_{\mathcal{M}_2^2}} \right) \right\} \\ & \geq \max_{k \in T_n} \left\{ \psi_k \log \left(\frac{\psi_k}{\omega_k} \right) + (1 - \psi_k) \log \left(\frac{1 - \psi_k}{\sum_{i=1}^n \omega_i - \omega_k} \right) \right\}. \end{aligned}$$

3. Estimations for the Zipf–Mandelbrot Entropy

In the following main result, we obtain some general estimations for the Zipf–Mandelbrot entropy.

Theorem 4. Let $n \in N, h \geq 0, r > 0, \omega_i > 0, i \in T_n$, then

$$\begin{aligned} -Z(Q, h, r) &= \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}} \\ &\geq \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_l^k} \left\{ \sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_1^k} \omega_i} \right) + \right. \\ &\quad \left. \dots + \sum_{i \in \mathcal{M}_l^k} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_l^k} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_l^k} \omega_i} \right) \right\} \\ &\geq \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_{l-1}^k} \left\{ \sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_1^k} \omega_i} \right) + \right. \\ &\quad \left. \dots + \sum_{i \in \mathcal{M}_{l-1}^k} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_{l-1}^k} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_{l-1}^k} \omega_i} \right) \right\} \\ &\geq \dots \geq \log \left(\frac{1}{\Omega_{T_n}} \right). \end{aligned} \tag{10}$$

Proof. If we substitute ψ_i with $\frac{1}{(i+h)^r Q_{n,h,r}}$, ($i \in T_n$), we have

$$\begin{aligned} \sum_{i=1}^n \psi_i g(x_i) &= - \sum_{i=1}^n \psi_i \log \omega_i + \sum_{i=1}^n \psi_i \log \psi_i \\ &= \sum_{i=1}^n \frac{1}{(i+h)^r Q_{n,h,r}} \log \frac{1}{(i+h)^r Q_{n,h,r}} - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}} \\ &= - \sum_{i=1}^n \frac{\log[(i+h)^r Q_{n,h,r}]}{(i+h)^r Q_{n,h,r}} - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}} \\ &= - \sum_{i=1}^n \frac{r \log(i+h)}{(i+h)^r Q_{n,h,r}} - \sum_{i=1}^n \frac{\log Q_{n,h,r}}{(i+h)^r Q_{n,h,r}} - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}} \\ &= - \frac{r}{Q_{n,h,r}} \sum_{i=1}^n \frac{\log(i+h)}{(i+h)^r} - \frac{\log Q_{n,h,r}}{Q_{n,h,r}} \sum_{i=1}^n \frac{1}{(i+h)^r} - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}}. \end{aligned}$$

Then,

$$\sum_{i=1}^n \psi_i g(x_i) = -Z(Q, h, r) - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}},$$

where $Q_{n,h,r} = \sum_{i=1}^n \frac{1}{(i+h)^r}$ and $\sum_{i=1}^n \frac{1}{(i+h)^r Q_{n,h,r}} = 1$.

Now, by applying Theorem 3 for $\psi_i = \frac{1}{(i+h)^r Q_{n,h,r}}$, we obtain the required result. \square

Corollary 4. Let $n \in \mathbb{N}$, $h \geq 0$, $r > 0$, $\omega_i > 0$ for $i \in T_n$, then

$$\begin{aligned} & -Z(Q, h, r) - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}} \\ & \geq \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_1^2} \omega_i} \right) \right. \\ & \quad \left. + \sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_2^2} \omega_i} \right) \right\} \geq \log \left(\frac{1}{\Omega_{T_n}} \right). \end{aligned} \tag{11}$$

Proof. By taking $k = 2$ in (10), we obtain (11). \square

We can use Theorem 4 to obtain the following corollary.

Corollary 5. Let $n \in \mathbb{N}$, $h \geq 0$, $r > 0$, then

$$\begin{aligned} -Z(Q, h, r) & \geq \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_1^2} 1} \right) \right. \\ & \quad \left. + \sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_2^2} 1} \right) \right\} \geq \log \left(\frac{1}{n} \right). \end{aligned} \tag{12}$$

Proof. By taking $\omega_i = 1$ in (11), $i = 1, 2, \dots, n$, we get (12). \square

Corollary 6. Let $n \in \mathbb{N}$, $h \geq 0$, $r > 0$, then

$$\begin{aligned}
 -Z(Q, h, r) - \sum_{i=1}^n \frac{\log \omega_i}{(i+h)^r Q_{n,h,r}} &\geq \max_{k \in T_n} \left\{ \frac{1}{(k+h)^r Q_{n,h,r}} \log \left(\frac{1}{(k+h)^r Q_{n,h,r} \omega_k} \right) \right. \\
 &\quad \left. + \left(\frac{(k+h)^r Q_{n,h,r} - 1}{(k+h)^r Q_{n,h,r}} \right) \log \left(\frac{((k+h)^r Q_{n,h,r} - 1) / (k+h)^r Q_{n,h,r}}{\Omega_{T_n} - \omega_k} \right) \right\} \\
 &\geq \log \left(\frac{1}{n} \right).
 \end{aligned}$$

Proof. Using $\mathcal{M}_1^2 = \{k\}$ in (12), we obtain Corollary 6. \square

Remark 3. Note that Corollary 4 is in fact the application of ([25], Theorem 1), and Corollary 6 is the application of ([21], Theorem 1).

Remark 4. By using Remark 2, we also have

$$\begin{aligned}
 \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_1^2} \omega_i} \right) + \sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+h)^r Q_{n,h,r}} \log \left(\frac{\sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+h)^r Q_{n,h,r}}}{\sum_{i \in \mathcal{M}_2^2} \omega_i} \right) \right\} \\
 \geq \max_{k \in T_n} \left\{ \frac{1}{(k+h)^r Q_{n,h,r}} \log \left(\frac{1}{(k+h)^r Q_{n,h,r} \omega_k} \right) \right. \\
 \left. + \left(\frac{(k+h)^r Q_{n,h,r} - 1}{(k+h)^r Q_{n,h,r}} \right) \log \left(\frac{((k+h)^r Q_{n,h,r} - 1) / (k+h)^r Q_{n,h,r}}{\Omega_{T_n} - \omega_k} \right) \right\}.
 \end{aligned}$$

In the following result, we obtain the estimation for the Zipf–Mandelbrot entropy concerning two different parameters.

Theorem 5. Let $u, v \geq 0$, $r_1, r_2 > 0$, then

$$\begin{aligned}
 -Z(Q, u, r_1) + \sum_{i=1}^n \frac{\log((i+v)^{r_2} Q_{n,v,r_2})}{(i+u)^{r_1} Q_{n,u,r_1}} \\
 \geq \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_l^k} \left\{ \sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}}{\sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}} \right) + \right. \\
 \left. \dots + \sum_{i \in \mathcal{M}_l^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{\sum_{i \in \mathcal{M}_l^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}}{\sum_{i \in \mathcal{M}_l^k} \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}} \right) \right\} \tag{13} \\
 \geq \max_{\mathcal{M}_1^k, \mathcal{M}_2^k, \dots, \mathcal{M}_{l-1}^k} \left\{ \sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}}{\sum_{i \in \mathcal{M}_1^k} \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}} \right) + \right. \\
 \left. \dots + \sum_{i \in \mathcal{M}_{l-1}^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{\sum_{i \in \mathcal{M}_{l-1}^k} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}}{\sum_{i \in \mathcal{M}_{l-1}^k} \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}} \right) \right\} \\
 \geq \dots \geq 0.
 \end{aligned}$$

Proof. Let $\psi_i = \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}$, $\omega_i = \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}$, $i \in T_n$. Then, using the proof of Theorem 4, we get

$$\begin{aligned} \sum_{i=1}^n \psi_i \log \psi_i &= \sum_{i=1}^n \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} = -Z(Q, u, r_1), \\ -\sum_{i=1}^n \psi_i \log \omega_i &= \sum_{i=1}^n \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}} \\ &= \sum_{i=1}^n \frac{\log((i+v)^{r_2} Q_{n,v,r_2})}{(i+u)^{r_1} Q_{n,u,r_1}}, \\ \sum_{i=1}^n \omega_i &= \sum_{i=1}^n \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}} = 1. \end{aligned}$$

Therefore, using (6) for $\psi_i = \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}$ and $\omega_i = \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}$, $i \in T_n$, we obtain (13). \square

Corollary 7. Let $n \in N$, $u, v \geq 0$, $r_1, r_2 > 0$, then

$$\begin{aligned} & -Z(Q, u, r_1) + \sum_{i=1}^n \frac{\log((i+v)^{r_2} Q_{n,v,r_2})}{(i+u)^{r_1} Q_{n,u,r_1}} \\ & \geq \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}}{\sum_{i \in \mathcal{M}_1^2} \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}} \right) \right. \\ & \quad \left. + \sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{\sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+u)^{r_1} Q_{n,u,r_1}}}{\sum_{i \in \mathcal{M}_2^2} \frac{1}{(i+v)^{r_2} Q_{n,v,r_2}}} \right) \right\} \geq 0. \end{aligned} \tag{14}$$

Proof. By taking $k = 2$ in (13), we obtain (14). \square

Corollary 8. Let $u, v \geq 0$, $r_1, r_2 > 0$, then

$$\begin{aligned} & -Z(Q, u, r_1) + \sum_{i=1}^n \frac{\log((i+v)^{r_2} Q_{n,v,r_2})}{(i+u)^{r_1} Q_{n,u,r_1}} \\ & \geq \max_{k \in T_n} \left\{ \frac{1}{(k+u)^{r_1} Q_{n,u,r_1}} \log \left(\frac{(k+v)^{r_2} Q_{n,v,r_2}}{(k+u)^{r_1} Q_{n,u,r_1}} \right) \right. \\ & \quad + \left(\frac{(k+u)^{r_1} Q_{n,u,r_1} - 1}{(k+u)^{r_1} Q_{n,u,r_1}} \right) \\ & \quad \left. \log \left(\frac{((k+u)^{r_1} Q_{n,u,r_1} - 1) / (k+u)^{r_1} Q_{n,u,r_1}}{((k+v)^{r_2} Q_{n,v,r_2} - 1) / (k+v)^{r_2} Q_{n,v,r_2}} \right) \right\} \\ & \geq 0. \end{aligned} \tag{15}$$

Proof. Using $\mathcal{M}_1^2 = \{k\}$ in (14), we obtain (15). \square

Now we give applications of the above results in linguistics.

In ref. [26], Gelbukh and Sidorov observed the difference between the coefficients r_1 and r_2 in the Zipf law for the English and Russian languages. 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 the coefficients for each of the mentioned texts and as a result, they obtained an average of $r_1 = 0.973863$ for the English language and $r_2 = 0.892869$ for the Russian language.

In the following results, we give the application of inequality (11) for the English language.

Application 1. Let $n \in N, \omega_i > 0$ for $i \in T_n$ Then, we have

$$\begin{aligned}
 & -Z(Q, 0, 0.973863) - \sum_{i=1}^n \frac{\log \omega_i}{i^{0.973863} Q_{n,0,0.973863}} \\
 & \geq \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \sum_{i \in \mathcal{M}_1^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}}}{\sum_{i \in \mathcal{M}_1^2} \omega_i} \right) \right. \\
 & \quad \left. + \sum_{i \in \mathcal{M}_2^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}} \log \left(\frac{\sum_{i \in \mathcal{M}_2^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}}}{\sum_{i \in \mathcal{M}_2^2} \omega_i} \right) \right\} \geq \log \left(\frac{1}{\Omega_{T_n}} \right). \tag{16}
 \end{aligned}$$

Similarly, we can give an application for the Russian language.

Now we give an application for the result of the related two parameters: $r_1 = 0.973863$ for the English language and $r_2 = 0.892869$ for the Russian language, which is in fact application of the inequality (14).

Application 2. Let $n \in N$. Then, we have

$$\begin{aligned}
 & -Z(Q, 0, 0.973863) + \sum_{i=1}^n \frac{\log(i^{0.892869} Q_{n,0,0.892869})}{i^{0.973863} Q_{n,0,0.973863}} \\
 & \geq \max_{\mathcal{M}_1^2, \mathcal{M}_2^2} \left\{ \sum_{i \in \mathcal{M}_1^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}} \log \left(\frac{\sum_{i \in \mathcal{M}_1^2} \frac{1}{i^{0.892869} Q_{n,0,0.892869}}}{\sum_{i \in \mathcal{M}_1^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}}} \right) \right. \\
 & \quad \left. + \sum_{i \in \mathcal{M}_2^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}} \log \left(\frac{\sum_{i \in \mathcal{M}_2^2} \frac{1}{i^{0.892869} Q_{n,0,0.892869}}}{\sum_{i \in \mathcal{M}_2^2} \frac{1}{i^{0.973863} Q_{n,0,0.973863}}} \right) \right\} \geq 0.
 \end{aligned}$$

Remark 5. In a similar way, applications of the remaining results from Section 3 can be given.

Author Contributions: All authors contributed equally to the final manuscript.

Funding: The research was supported by the Natural Science Foundation of China (Grants No. 61673169, No. 11601485, No. 11701176) and the Natural Science Foundation of the Department of Education of Zhejiang Province (Grant No. Y201635325).

Acknowledgments: The authors express their sincere thanks to the referees for careful reading of the manuscript and very helpful suggestions that improved the current manuscript substantially.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656. [CrossRef]
2. Latif, N.; Pečarić, Đ.; Pečarić, J. Majorization, Csiszár divergence and Zipf-Mandelbrot law. *J. Inequal. Appl.* **2017**, *2017*, 197. [CrossRef] [PubMed]
3. Quastler, H. (Ed.) *Essays on the Use of Information Theory in Biology*; University of Illinois: Urbana, IL, USA, 1953.
4. Sherwin, W.B. Entropy and information approaches to genetic diversity and its expression: Genomic geography. *Entropy* **2010**, *12*, 1765–1798. [CrossRef]

5. Zhou, R.-X.; Cai, R.; Tong, G.-Q. Applications of entropy in finance: A review. *Entropy* **2013**, *15*, 4909–4931. [CrossRef]
6. Guisado, J.L.; Jiménez-Morales, F.; Guerra, J.M. Application of Shannon's entropy to classify emergent behaviors in a simulation of laser dynamics. *Math. Comput. Model.* **2005**, *42*, 847–854. [CrossRef]
7. Wellmann, J.F.; Regenauer-Lieb, K. Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. *Tectonophysics* **2012**, *526–529*, 207–216. [CrossRef]
8. Adil Khan, M.; Pečarić, Đ.; Pečarić, J. Bounds for Shannon and Zipf-Mandelbrot entropies. *Math. Methods Appl. Sci.* **2017**, *40*, 7316–7322. [CrossRef]
9. Silagadze, Z.K. Citations and the Zipf-Mandelbrot's law. *Complex Syst.* **1997**, *11*, 487–499.
10. Kristian, G.; Jens, S. Zipf's law for cities in the regions and the country. *J. Econ. Geogr.* **2010**, *11*, 667–686.
11. Merriam, D.F.; Drew, L.J.; Schuenemeyer, J.H. Zipf's law: A viable geological paradigm? *Nat. Resour. Res.* **2004**, *13*, 265–271. [CrossRef]
12. Newman, M.E.J. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **2005**, *46*, 323–351. [CrossRef]
13. Parkins, S. Website Traffic and Zipf's Law. 2015. Available online: <http://www.linkedin.com/pulse/website-traffic-zipfs-law-shane-parkins?articleId=6064338455452807169> (accessed on 26 October 2015).
14. Hefeeda, M.; Saleh, O. Traffic modeling and proportional partial caching for peer-to-peer systems. *IEEE/ACM Trans. Netw.* **2008**, *16*, 1447–1460. [CrossRef]
15. Neukum, G.; Ivanov, B.A. Crater size distributions and impact probabilities on earth from lunar, terrestrial-planet, and asteroid cratering data. In *Hazards Due to Comets & Asteroids*; University Arizona Press: Tucson, AZ, USA, 1994.
16. Montemurro, M.A. Beyond the Zipf-Mandelbrot law in quantitative linguistics. *Phys. A Stat. Mech. Its Appl.* **2001**, *300*, 567–578. [CrossRef]
17. Manin, D.Y. Mandelbrot's model for Zipf's law: Can Mandelbrot's model explain Zipf's law for language? *J. Quant. Ling.* **2009**, *16*, 274–285. [CrossRef]
18. Mouillot, D.; Lepretre, A. Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity. *Environ. Monit. Assess.* **2000**, *63*, 279–295. [CrossRef]
19. Lovričević, N.; Pečarić, Đ.; Pečarić, J. Zipf-Mandelbrot law, f -divergences and the Jensen-type interpolating inequalities. *J. Inequal. Appl.* **2018**, *2018*, 36.
20. Adil Khan, M.; Ali Khan, G.; Alia, T.; Kilicman, A. On the refinement of Jensen's inequality. *Appl. Math. Comput.* **2015**, *262*, 128–135.
21. Dragomir, S.S. A refinement of Jensen's inequality with applications for f -divergence measures. *Taiwan. J. Math.* **2010**, *14*, 153–164. [CrossRef]
22. Adil Khan, M.; Pečarić, Đ.; Pečarić, J. On Zipf-Mandelbrot entropy. *J. Comput. Appl. Math.* **2019**, *346*, 192–204. [CrossRef]
23. Dragomir, S.S. Bounds for the normalised Jensen functional. *Bull. Aust. Math. Soc.* **2006**, *74*, 471–478. [CrossRef]
24. Abbaszadeh, S.; Gordji, M.E.; Pap, E.; Szakal, A. Jensen-type inequalities for Sugeno integral. *Inf. Sci.* **2017**, *376*, 148–157. [CrossRef]
25. Dragomir, S.S. A new refinement of Jensen's inequality in linear spaces with application. *Math. Comput. Model.* **2010**, *52*, 1497–1505. [CrossRef]
26. Gelbukh, A.; Sidorov, G. Zipf and Heaps laws' coefficients depend on language. *Lect. Notes Comput. Sci.* **2001**, *2004*, 332–335.

