Majorization, "useful" Csiszár divergence and "useful" Zipf-Mandelbrot law

https://doi.org/10.1515/math-2018-0113
Received December 8, 2017; accepted June 20, 2018.

Abstract: In this paper, we consider the definition of "useful" Csiszár divergence and "useful" Zipf-Mandelbrot law associated with the real utility distribution to give the results for majorization inequalities by using monotonic sequences. We obtain the equivalent statements between continuous convex functions and Green functions via majorization inequalities, "useful" Csiszár functional and "useful" Zipf-Mandelbrot law. By considering "useful" Csiszár divergence in the integral case, we give the results for integral majorization inequality. Towards the end, some applications are given.

Keywords: "Useful" Csiszár divergence, "Useful" Zipf-Mandelbrot law, Majorization inequality, Convex functions, Green functions, Information theory

MSC: 94A15, 94A17, 26A51, 26D15

1 Introduction and Preliminaries

Zipf’s law [1-3] and the power laws in general [4-6] have and continue to attract considerable attention in a wide variety of disciplines from astronomy to demographics to software structure to economics to zoology, and even to warfare [7]. Typically one is dealing with integer-valued observables (number of objects, people, cities, words, animals, corpses), with \( n \in \{1, 2, 3, \ldots\} \). As given in [8], sometimes the range of values is allowed to be infinite (at least in principle), sometimes a hard upper bound \( N \) is fixed (e.g., total population if one is interested in subdividing a fixed population into sub-classes). Particularly interesting probability distributions are the probability laws of the form:

- Zipf’s law: \( p_n \propto 1/n \);
- power laws: \( p_n \propto 1/n^z \);
- hybrid geometric/power laws: \( p_n \propto w^n / n^z \).

Distance or divergence measures are of key importance in different fields like theoretical and applied statistical inference and data processing problems such as estimation, detection, classification, compression, recognition, indexation, diagnosis and model selection etc. Traditionally, the information conveyed by observing \( X \) is measured by the entropy which is defined as (see [9, p.111])

\[
H(p) := \sum_{i=1}^{n} p_i \log_2 1/p_i,
\]

*Corresponding Author: Naveed Latif: Department of General Studies, Jubail Industrial College, Jubail Industrial City 31961, Kingdom of Saudi Arabia, E-mail: naveed707@gmail.com
Đilda Pečarić: Catholic University of Croatia, Ilica 242, 10000 Zagreb, Croatia, E-mail: gildapecia@gmail.com
Josip Pečarić: Faculty of Textile Technology Zagreb, University of Zagreb, Prilaz Baruna Filipovića 28A, 10000 Zagreb, Croatia and RUDN University, 6 Miklukho-Maklay St, Moscow, 117198, Russia, E-mail: pecaric@element.hr

Unauthenticated Download Date | 6/2/19 5:59 AM
and is associated with the distribution \( p, p_i > 0 \) \((1 \leq i \leq n)\), where \( \sum_{i=1}^{n} p_i = 1 \). A generalization of this is to attach a utility \( q_i > 0 \) to the outcome \( x_i \) \((1 \leq i \leq n)\) and speak of the "useful" information measure

\[
H(p; q) := \sum_{i=1}^{n} q_i p_i \log_2 \frac{1}{p_i},
\]

which is associated with the utility distribution \( q = (q_1, \ldots, q_n) \).

Bhaker and Hooda [10] (see also [9, p.112]) introduced the measures

\[
E(p; q) := \frac{\sum_{k=1}^{n} q_k p_k \log_2 \frac{1}{p_k}}{\sum_{k=1}^{n} q_k p_k} \tag{1}
\]

and

\[
E_\alpha(p; q) := \frac{1}{1 - \alpha} \log_2 \frac{\sum_{k=1}^{n} q_k p_k^\alpha}{\sum_{k=1}^{n} q_k p_k^\alpha}, \quad 0 < \alpha \neq 1, \tag{2}
\]

which have a number of useful properties. It is readily verified that these alternations leave intact the property that (2) reduces to (1) when \( \alpha \to 1 \). Also, if \( u \equiv 0 \) so that there are effectively no utilities, (1) and (2) reduce to Rényi’s entropies of order 1 and \( \alpha \), respectively.

Csiszár introduced the functional in [11] and later discussed it in [12]. Here, we consider "useful" Csiszár divergence (see [13, p.3], [9, 14, 15]):

**Definition 1.1 ("Useful" Csiszár divergence).** Assume \( J \subset \mathbb{R} \) be an interval, and let \( f : J \to \mathbb{R} \) be a function with distribution \( p := (p_1, \ldots, p_n) \), associated with the utility distribution \( u := (u_1, \ldots, u_n) \), where \( p_i, u_i \in \mathbb{R} \) for \( 1 \leq i \leq n \), and \( q := (q_1, \ldots, q_n) \in ]0, \infty[^n \) be such that

\[
\frac{p_i}{q_i} \in J, \quad i = 1, \ldots, n, \tag{3}
\]

then we denote the "useful" Csiszár divergence

\[
I_f(p, q, u) := \sum_{i=1}^{n} u_i q_i f\left(\frac{p_i}{q_i}\right). \tag{4}
\]

**Remark 1.2.** One can easily seen that if we substitute \( u = 1 \), then (4) becomes

\[
I_f(p, q, 1) := I_f(p, q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right). \tag{5}
\]

One can see the various results in information theory in [3, 16, 17].

The following theorem is a generalization of the Classical Majorization Theorem known as Weighted Majorization Theorem and was proved by Fuchs in [19] (see also [20], [21, p.323]):

**Theorem 1.3 (Weighted Majorization Theorem).** Let \( x = (x_1, \ldots, x_n) \), \( y = (y_1, \ldots, y_n) \) be two decreasing real \( n \)-tuples such that \( x_i, y_i \in J \) for \( i = 1, \ldots, n \). Let \( w = (w_1, \ldots, w_n) \) be a real \( n \)-tuple such that

\[
\sum_{i=1}^{j} w_i y_i \leq \sum_{i=1}^{j} w_i x_i, \tag{5}
\]

for \( j = 1, 2, \ldots, n - 1 \) and

\[
\sum_{i=1}^{n} w_i y_i = \sum_{i=1}^{n} w_i x_i. \tag{6}
\]

Then for every continuous convex function \( f : J \to \mathbb{R} \), we have the following inequality

\[
\sum_{i=1}^{n} w_i f(y_i) \leq \sum_{i=1}^{n} w_i f(x_i). \tag{7}
\]
The following theorem is valid ([22, p.32]):

**Theorem 1.4.** Let \( f : J \to \mathbb{R} \) be a continuous convex function on an interval \( J \), \( w \) be a positive \( n \)-tuple and \( x, y \) \( \in J^n \) satisfying

\[
\sum_{i=1}^{k} w_i y_i \leq \sum_{i=1}^{k} w_i x_i \quad \text{for} \quad k = 1, \ldots, n - 1,
\]

and

\[
\sum_{i=1}^{n} w_i y_i = \sum_{i=1}^{n} w_i x_i.
\]

(a) If \( y \) is a decreasing \( n \)-tuple, then

\[
\sum_{i=1}^{n} w_i f(y_i) \leq \sum_{i=1}^{n} w_i f(x_i).
\]

(b) If \( x \) is an increasing \( n \)-tuple, then

\[
\sum_{i=1}^{n} w_i f(x_i) \leq \sum_{i=1}^{n} w_i f(y_i).
\]

If \( f \) is strictly convex and \( x \neq y \), then (10) and (11) are strict.

One can see the various generalizations of the majorization inequality and bounds for Zipf-Mandelbrot entropy in [23-25].

Benoit Mandelbrot in [26] gave generalization of Zipf’s law, now known as the Zipf-Mandelbrot law which gave improvement in account for the low-rank words in corpus where \( k < 100 \) [27]:

\[
f(k) = \frac{C}{(k + q)^s},
\]

and when \( q = 0 \), we get Zipf’s law.

For \( n \in \mathbb{N}, q \geq 0, s > 0, k \in \{1, 2, \ldots, n\} \), in a more clear form, the Zipf-Mandelbrot law (probability mass function) is defined with

\[
f(k, n, q, s) := \frac{1}{(k + q)^s}, \quad \text{where} \quad H_{n,q,s} := \sum_{i=1}^{n} \frac{1}{(i + q)^s},
\]

\( n \in \mathbb{N}, q \geq 0, s > 0, k \in \{1, 2, \ldots, n\} \).

Application of the Zipf-Mandelbrot law can also be found in linguistics [27], information sciences [28, 29] and ecological field studies [30].

In probability theory and statistics, the cumulative distribution function (CDF) of a real-valued random variable \( X \), or just distribution function of \( X \), evaluated at \( x \), is the probability that \( X \) will take a value less than or equal to \( x \) and we often denote CDF as the following ratio:

\[
\text{CDF} := \frac{H_{k,t,s}}{H_{n,t,s}}.
\]

The cumulative distribution function is an important application of majorization.

We consider the following definition of “useful” Zipf-Mandelbrot law (see [9, 11, 12, 14, 15]):

**Definition 1.5 ("Useful" Zipf-Mandelbrot law).** Assume \( J \subset \mathbb{R} \) be an interval, and \( f : J \to \mathbb{R} \) be a function with \( n \in \{1, 2, 3, \ldots\} \), \( t_1 \geq 0 \). Let also distribution \( q_i > 0 \) and associated with the utility distribution \( u_i \in \mathbb{R} \) for \( (i = 1, \ldots, n) \) such that

\[
\frac{1}{q_i(i + t_1)^{s_1}H_{n,t_1,s_1}} \in J, \quad i = 1, \ldots, n,
\]

then we denote “useful” Zipf-Mandelbrot law as

\[
I_f (i, n, t_1, s_1, q, u) := \frac{\sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_1)^{s_1}H_{n,t_1,s_1}} \right)}{H_{n,t_1,s_1}}.
\]
Remark 1.6. One can easily seen that for $u = 1$, then

$$I_f(i, n, t_1, s_1, q, 1) = I_f(i, n, t_1, s_1, q) := \sum_{i=1}^{n} q_i f \left( \frac{1}{q_i, (i + t_1)^{J_n, t_1, s_1}} \right).$$

If we substitute $q_i = \frac{1}{(i + t_1)^{J_n, t_1, s_1}}$, then

$$I_f(i, n, t_1, t_3, s_1, s_3) := \sum_{i=1}^{n} \frac{1}{(i + t_1)^{J_n, t_1, s_1}} f \left( \frac{(i + t_3)^{J_n, t_3, s_3}}{(i + t_1)^{J_n, t_1, s_1}} \right).$$

This paper is organized as follows. In section 2, we give the results as the connection between useful Csisár divergence, useful Zipf-Mandelbrot law and majorization inequality for one monotonic sequence or both of them. We obtain some corollaries for our obtained results. In section 3, we present the equivalent statements between continuous convex functions and defined Green functions. In section 4, we give the results for integral majorization inequality for considering the integral form of useful Csisár divergence. Finally, in section 5 we give some applications for obtained results.

2 Main results

Assume $p$ and $q$ be $n$-tuples such that $q_i > 0$ $(i = 1, ..., n)$ and define

$$\frac{p}{q} := \left( \frac{p_1}{q_1}, \frac{p_2}{q_2}, ..., \frac{p_n}{q_n} \right).$$

We start with the following theorem which provides the connection between "useful" Csisár divergence and weighted majorization as one sequence is monotonic:

Theorem 2.1. Assume $J \subset \mathbb{R}$ be an interval, $f : J \to \mathbb{R}$ be a continuous convex function, $p_i, r_i (i = 1, ..., n)$ be real numbers and $q_i, u_i (i = 1, ..., n)$ be positive real numbers such that

$$\sum_{i=1}^{k} u_i r_i \leq \sum_{i=1}^{n} u_i p_i \quad \text{for} \quad k = 1, ..., n - 1, \quad (14)$$

and

$$\sum_{i=1}^{n} u_i r_i = \sum_{i=1}^{n} u_i p_i, \quad (15)$$

with $\frac{p_i}{q_i}, \frac{r_i}{q_i} \in J$ $(i = 1, ..., n)$.

(a) If $\frac{r_i}{q_i}$ is decreasing, then

$$I_f(\frac{r_i}{q_i}, q, u) \leq I_f(p, q, u). \quad (16)$$

(b) If $\frac{r_i}{q_i}$ is increasing, then

$$I_f(\frac{r_i}{q_i}, q, u) \geq I_f(p, q, u). \quad (17)$$

If $f$ is a continuous concave function, then the reverse inequalities hold in (16) and (17).

Proof. (a): We use Theorem 1.4 (a) with substitutions $x_i := \frac{p_i}{q_i}, y_i := \frac{r_i}{q_i}, w_i = u_i q_i$ as $q_i > 0, (i = 1, ..., n)$ then we get (16).

We can prove part (b) with the similar substitutions in Theorem 1.4 (b).

We present the following theorem as the connection between "useful" Csisár divergence and weighted majorization theorem as both sequences are decreasing:

Theorem 2.2. Assume $J \subset \mathbb{R}$ be an interval, $f : J \to \mathbb{R}$ be a continuous convex function, $p_i, r_i, u_i (i = 1, ..., n)$ be real numbers and $q_i (i = 1, ..., n)$ be positive real numbers such that $\frac{p_i}{q_i}$ and $\frac{r_i}{q_i}$ be decreasing satisfying (14) and (15) with $\frac{p_i}{q_i}, \frac{r_i}{q_i} \in J$ $(i = 1, ..., n)$, then

$$I_f(\frac{r_i}{q_i}, q, u) \leq I_f(p, q, u). \quad (18)$$
Proof. We use Theorem 1.3 with substitutions $x_i := \frac{u_i}{q_i^2}$, $y_i := \frac{x_i}{q_i}$ and $w_i = u_i q_i$ as $q_i > 0$ ($i = 1, \ldots, n$) then we get (18).

The following two theorem gives the connection between "useful" Zipf-Mandelbrot law and weighted majorization inequality:

**Theorem 2.3.** Assume $I \subset \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ be a continuous convex function with $u_i > 0$, $n \in \{1, 2, 3, \ldots\}$, $t_1, t_2 \geq 0$ and $s_1, s_2 > 0$ such that satisfying

$$
\sum_{i=1}^{k} \frac{u_i}{(i + t_2)^{s_2}} \leq \frac{H_{n,t_2,s_2}}{H_{n,t_1,s_1}} \sum_{i=1}^{k} \frac{u_i}{(i + t_1)^{s_1}}, \quad k = 1, \ldots, n - 1,
$$

(19)

and

$$
\sum_{i=1}^{n} \frac{u_i}{(i + t_2)^{s_2}} = \frac{H_{n,t_2,s_2}}{H_{n,t_1,s_1}} \sum_{i=1}^{n} \frac{u_i}{(i + t_1)^{s_1}},
$$

(20)

and also let $q_i > 0$, ($i = 1, \ldots, n$) with

$$
\frac{1}{q_i(i + t_1)^{s_1} H_{n,t_1,s_1}}, \quad \frac{1}{q_i(i + t_2)^{s_2} H_{n,t_2,s_2}} \in f (i = 1, \ldots, n).
$$

(a) If $(\frac{i+t_2}{i+t_1})^{s_2} \leq \frac{q_1}{q_i}$, ($i = 1, \ldots, n$), then

$$
I_f (i, n, t_2, s_2, q, u) := \sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_2)^{s_2} H_{n,t_2,s_2}} \right)
$$

$$\leq I_f (i, n, t_1, s_1, q, u) := \sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_1)^{s_1} H_{n,t_1,s_1}} \right).
$$

(21)

(b) If $(\frac{i+t_1}{i+t_2})^{s_1} \geq \frac{q_1}{q_i}$, ($i = 1, \ldots, n$), then

$$
\sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_2)^{s_2} H_{n,t_1,s_1}} \right)
$$

$$\geq \sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_1)^{s_1} H_{n,t_2,s_2}} \right).
$$

(22)

If $f$ is continuous concave function, then the reverse inequalities hold in (21) and (22).

**Proof.** (a) Let us consider that $p_i := \frac{1}{(i + t_1)^{s_1} H_{n,t_1,s_1}}$ and $r_i := \frac{1}{(i + t_2)^{s_2} H_{n,t_2,s_2}}$, then

$$
\sum_{i=1}^{k} u_i p_i := \sum_{i=1}^{k} \frac{u_i}{(i + t_1)^{s_1} H_{n,t_1,s_1}} = \frac{1}{H_{n,t_1,s_1}} \sum_{i=1}^{k} \frac{u_i}{(i + t_1)^{s_1}}, \quad k = 1, \ldots, n - 1,
$$

and similarly

$$
\sum_{i=1}^{k} u_i r_i := \frac{1}{H_{n,t_2,s_2}} \sum_{i=1}^{k} \frac{u_i}{(i + t_2)^{s_2}}, \quad k = 1, \ldots, n - 1,
$$

leading to

$$
\sum_{i=1}^{k} u_i r_i \leq \sum_{i=1}^{k} u_i p_i \quad \Leftrightarrow \quad \sum_{i=1}^{k} \frac{u_i}{(i + t_2)^{s_2}} \leq \frac{H_{n,t_2,s_2}}{H_{n,t_1,s_1}} \sum_{i=1}^{k} \frac{u_i}{(i + t_1)^{s_1}}, \quad k = 1, \ldots, n - 1.
$$

One can see easily that $\frac{1}{(i + t_1)^{s_1} H_{n,t_1,s_1}}$ is decreasing over $i = 1, \ldots, n$ and similarly $r_i$ too. Now, we find the behaviour of $\frac{r_i}{q_i}$ for $q_i > 0$ ($i = 1, 2, \ldots, n$), take

$$
\frac{r_i}{q_i} = \frac{1}{q_i(i + t_2)^{s_2} H_{n,t_2,s_2}} \quad \text{and} \quad \frac{r_{i+1}}{q_{i+1}} = \frac{1}{q_{i+1}(i + 1 + t_2)^{s_2} H_{n,t_2,s_2}},
$$
which shows that \( \frac{r}{q} \) is decreasing. So, all the assumptions of Theorem 2.1 (a) are true, then by using (16) we get (21).

(b) If we switch the role of \( r_t \) to \( p_i \) in the first part (a), then by using (17) in Theorem 2.1 (b) we get (22). \( \square \)

**Theorem 2.4.** Assume \( J \subseteq \mathbb{R} \) be an interval, \( f : J \to \mathbb{R} \) be a continuous convex function with \( u_i \in \mathbb{R}, n \in \{1, 2, 3, \ldots \}, t_1, t_2 \geq 0 \) and \( s_1, s_2 > 0 \), such that satisfying (19), (20) and

\[
\frac{(i + t_1)^{s_1}}{(i + 1 + t_2)^{s_2}} \leq \frac{q_i}{q_{i+1}} \quad (i = 1, \ldots, n),
\]

\[
\frac{(i + t_2)^{s_2}}{(i + 1 + t_2)^{s_2}} \leq \frac{q_i}{q_{i+1}} \quad (i = 1, \ldots, n),
\]

hold and also let \( q_i > 0, (i = 1, \ldots, n) \) with

\[
\frac{1}{q_i(i + t_1)^{s_1}H_{n,t_1,s_1}}, \quad \frac{1}{q_i(i + t_2)^{s_2}H_{n,t_2,s_2}} \in J \quad (i = 1, \ldots, n),
\]

then the following inequality holds

\[
I_f (i, n, t_2, s_2, q, u) := \sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_2)^{s_2}H_{n,t_2,s_2}} \right)
\leq I_f (i, n, t_1, s_1, q, u) := \sum_{i=1}^{n} u_i q_i f \left( \frac{1}{q_i(i + t_1)^{s_1}H_{n,t_1,s_1}} \right).
\]

(23)

**Proof.** Let us consider that \( p_i := \frac{1}{(i + 1 + t_2)^{s_2}H_{n,t_2,s_2}} \) and \( r_t := \frac{1}{(i + t_1)^{s_1}H_{n,t_1,s_1}} \), so as given in the proof of Theorem 2.3, we get \( y = r/q \) is decreasing \( \iff \frac{(i + t_2)^{s_2}}{(i + 1 + t_2)^{s_2}} \leq \frac{q_i}{q_{i+1}} \), for \( i = 1, \ldots, n \), similarly we can prove that \( x = p/q \) is also decreasing \( \iff \frac{(i + t_1)^{s_1}}{(i + 1 + t_1)^{s_1}} \leq \frac{q_i}{q_{i+1}} \), for \( i = 1, \ldots, n \). So, all the assumptions of Theorem 2.2 are true, then by using (18) we get (23). \( \square \)

The following two corollaries obtain from Theorem 5 and Theorem 6 respectively but we use three Zipf-Mandelbrot laws for different parameters:

**Corollary 2.5.** Assume \( J \subseteq \mathbb{R} \) be an interval, \( f : J \to \mathbb{R} \) be a continuous convex function with \( u_i > 0, n \in \{1, 2, 3, \ldots \}, t_1, t_2 \geq 0 \) and \( s_1, s_2 > 0 \) such that satisfying (19) and (20) and

\[
(i + t_1)^{s_1}H_{n,t_1,s_1}, \quad (i + t_2)^{s_2}H_{n,t_2,s_2} \in J \quad (i = 1, \ldots, n).
\]

(a) If \( \frac{(i + t_1 + t_3)^{s_1}}{(i + 1 + t_3)^{s_1}} \leq \frac{(i + t_2 + t_3)^{s_2}}{(i + 1 + t_3)^{s_2}} \quad (i = 1, \ldots, n) \), then

\[
I_f (i, n, t_2, s_2, t_3, s_3, u) := \sum_{i=1}^{n} \frac{u_i}{(i + t_2)^{s_2}H_{n,t_2,s_2}} \left( \frac{(i + t_3)^{s_3}H_{n,t_3,s_3}}{(i + t_2)^{s_2}H_{n,t_2,s_2}} \right)
\leq I_f (i, n, t_1, s_1, t_3, s_3, u) := \sum_{i=1}^{n} \frac{u_i}{(i + t_1)^{s_1}H_{n,t_1,s_1}} \left( \frac{(i + t_3)^{s_3}H_{n,t_3,s_3}}{(i + t_1)^{s_1}H_{n,t_1,s_1}} \right).
\]

(24)

(b) If \( \frac{(i + t_1 + t_3)^{s_1}}{(i + 1 + t_3)^{s_1}} \geq \frac{(i + t_2 + t_3)^{s_2}}{(i + 1 + t_3)^{s_2}} \quad (i = 1, \ldots, n) \), then

\[
\sum_{i=1}^{n} \frac{u_i}{(i + t_3)^{s_3}H_{n,t_3,s_3}} \left( \frac{(i + t_3)^{s_3}H_{n,t_3,s_3}}{(i + t_1)^{s_1}H_{n,t_1,s_1}} \right)
\geq \sum_{i=1}^{n} \frac{u_i}{(i + t_2)^{s_2}H_{n,t_2,s_2}} \left( \frac{(i + t_3)^{s_3}H_{n,t_3,s_3}}{(i + t_1)^{s_1}H_{n,t_1,s_1}} \right).
\]

(25)
If $f$ is continuous concave function, then the reverse inequalities hold in (24) and (25).

**Proof.** (a) Let $p_i := \frac{1}{(i+t_1)^{s_1}H_{n,t_1}}$, $q_i := \frac{1}{(i+t_2)^{s_2}H_{n,t_2}}$ and $r_i := \frac{1}{(i+t_3)^{s_3}H_{n,t_3}}$, here $p_i$, $q_i$ and $r_i$ are decreasing over $i = 1, \ldots, n$. Now, we investigate the behavior of $\frac{r_i}{q_i}$, take

$$
\frac{r_{i+1}}{q_{i+1}} - \frac{r_i}{q_i} = \frac{(i + t_2)^{s_2}H_{n,t_2}}{(i + t_3)^{s_3}H_{n,t_3}} - \frac{(i + t_1)^{s_1}H_{n,t_1}}{(i + t_3)^{s_3}H_{n,t_3}}
$$

the R. H. S. is non-positive by using the assumption, which shows that $\frac{r_i}{q_i}$ is decreasing, therefore using Theorem 5(a) we get (24).

(b) If we switch the role of $\frac{r_i}{q_i}$ with $\frac{p_i}{q_i}$ in the part (a) and using Theorem 5(b), we get (25).

**Corollary 2.6.** Assume $J \subset \mathbb{R}$ be an interval, $f : J \to \mathbb{R}$ be a continuous convex function with $u_i \in \mathbb{R}$, $n \in \{1, 2, 3, \ldots\}$, $t_1, t_2 \geq 0$ and $s_1, s_2 > 0$, such that satisfying (19) and (20) and

- \( (i+t_1)^{s_1} \leq \frac{(i+t_2)^{s_2}}{(i+t_1)^{s_2}} \) for $i = 1, \ldots, n$,
- \( (i+t_2)^{s_2} \leq \frac{(i+t_3)^{s_3}}{(i+t_2)^{s_3}} \) for $i = 1, \ldots, n$,

hold with

$$
\frac{(i + t_3)^{s_3}H_{n,t_3}}{(i + t_2)^{s_2}H_{n,t_2}}, \frac{(i + t_2)^{s_2}H_{n,t_2}}{(i + t_1)^{s_1}H_{n,t_1}} \in J \quad (i = 1, \ldots, n),
$$

then the following inequality holds

$$
I_f (i, n, t_2, s_2, t_3, s_3, u) := \sum_{i=1}^{n} \frac{u_i}{(i + t_3)^{s_3}H_{n,t_3}} \frac{f((i + t_3)^{s_3}H_{n,t_3})}{(i + t_2)^{s_2}H_{n,t_2}} \leq I_f (i, n, t_1, s_1, t_3, s_3, u) := \sum_{i=1}^{n} \frac{u_i}{(i + t_1)^{s_1}H_{n,t_1}} \frac{f((i + t_1)^{s_1}H_{n,t_1})}{(i + t_3)^{s_3}H_{n,t_3}}.
$$

**Proof.** (a) Let us consider that $p_i := \frac{1}{(i+t_1)^{s_1}H_{n,t_1}}$ and $r_i := \frac{1}{(i+t_2)^{s_2}H_{n,t_2}}$, so as given in the proof of Corollary 2.5 for $q_i > 0$ we get $y = r/q$ is decreasing $\Leftrightarrow \frac{(i+t_2)^{s_2}}{(i+t_1)^{s_1}} \leq \frac{(i+t_3)^{s_3}}{(i+t_2)^{s_3}}$, for $i = 1, \ldots, n$, similarly we can prove that $x = p/q$ is also decreasing $\Leftrightarrow \frac{(i+t_3)^{s_3}}{(i+t_2)^{s_2}} \leq \frac{(i+t_1)^{s_1}}{(i+t_3)^{s_3}}$, for $i = 1, \ldots, n$. Therefore, all the assumptions of Theorem 2.4 are true, then by using (23) we get (26).

**Remark 2.7.** We can give Theorem 2.1, Theorem 2.2, Theorem 2.3, Theorem 2.4, Corollary 2.5 and Corollary 2.6 for $u := 1$ as special case, some of them has been given in [14].

### 3 "Useful" information measure via Green functions

Consider the Green function $G_1$ defined on $[\vartheta_1, \vartheta_2] \times [\vartheta_1, \vartheta_2]$ by

$$
G_1(u, v) = \left\{ \begin{array}{ll}
\frac{(u-\vartheta_1)(v-\vartheta_1)}{\vartheta_2-\vartheta_1}, & \vartheta_1 \leq u \leq v; \\
\frac{(v-\vartheta_1)(u-\vartheta_1)}{\vartheta_2-\vartheta_1}, & u \leq v \leq \vartheta_2.
\end{array} \right.
$$

The function $G_1$ is convex in $v$, it is symmetric, so it is also convex in $u$. The function $G_1$ is continuous in $v$ and continuous in $u$. 

Unauthenticated
Download Date | 6/2/19 5:59 AM
For any function \( f : \var{\var{1}, \var{2}} \to \mathbb{R}, f \in C^2([\var{1}, \var{2}]) \), we can easily show by integrating by parts that the following is valid
\[
f(u) = \frac{\var{2} - u}{\var{2} - \var{1}} f(\var{1}) + \frac{u - \var{1}}{\var{2} - \var{1}} f(\var{2}) + \int_{\var{1}}^{\var{2}} G(u, v)f''(v)dv,
\]
where the function \( G_1 \) is defined as above in (27) (31).

Let \([\var{1}, \var{2}] \subset \mathbb{R} \) and \( d = 2, 3, 4, 5 \). Recently in (2017), Mehmood et al. [32] (also see [33]) introduced some new types of Green functions, \( G_d : \var{1} \times \var{2} \to \mathbb{R} \) and give Lemma 1, which are defined as follows:

\[
G_2(u, v) = \begin{cases} 
(\var{1} - v), & \var{1} \leq v \leq u, \\
(\var{1} - u), & u \leq v \leq \var{2}, 
\end{cases}
\]

(28)

\[
G_3(u, v) = \begin{cases} 
(u - \var{2}), & \var{1} \leq v \leq u, \\
(v - \var{2}), & u \leq v \leq \var{2}, 
\end{cases}
\]

(29)

\[
G_4(u, v) = \begin{cases} 
(u - \var{1}), & \var{1} \leq v \leq u, \\
(v - \var{1}), & u \leq v \leq \var{2}, 
\end{cases}
\]

(30)

\[
G_5(u, v) = \begin{cases} 
(\var{2} - v), & \var{1} \leq v \leq u, \\
(\var{2} - u), & u \leq v \leq \var{2}, 
\end{cases}
\]

(31)

**Lemma 3.1.** Let \( f : \var{1} \to \mathbb{R} \) such that \( f \in C^2([\var{1}, \var{2}]) \) and \( G_d (d = 2, 3, 4, 5) \) be Green functions as defined in \((28), (29), (30) \) and \((31)\), then we have the following identities.

\[
f(u) = f(\var{1}) + (u - \var{1})f'(\var{2}) + \int_{\var{1}}^{\var{2}} G_2(u, v)f''(v)dv,
\]

(32)

\[
f(u) = f(\var{2}) + (u - \var{2})f'(\var{1}) + \int_{\var{1}}^{\var{2}} G_3(u, v)f''(v)dv,
\]

(33)

\[
f(u) = f(\var{2}) - (\var{2} - \var{1})f'(\var{2}) + (u - \var{1})f'(\var{1}) + \int_{\var{1}}^{\var{2}} G_4(u, v)f''(v)dv,
\]

(34)

\[
f(u) = f(\var{1}) + (\var{2} - \var{1})f'(\var{1}) - (\var{2} - u)f'(\var{2}) + \int_{\var{1}}^{\var{2}} G_5(u, v)f''(v)dv.
\]

(35)

The following theorem gives the equivalent statements between continuous convex functions and Green functions via majorization inequality and “useful” Csiszár divergence.

**Theorem 3.2.** Assume \( J \subset \mathbb{R} \) be an interval, \( p_i, r_i (i = 1, \ldots, n) \) be real numbers and \( q_i, u_i (i = 1, \ldots, n) \) be positive real numbers such that satisfying 

\[
\sum_{i=1}^{n} u_ir_i = \sum_{i=1}^{n} u_ip_i,
\]

(36)

with \( \frac{p_i}{q_i}, \frac{r_i}{q_i} \in J \) (i = 1, \ldots, n). If \( \frac{p_i}{q_i} \) is decreasing and \( G_d (d = 1, 2, 3, 4, 5) \) be defined as in \((27)-(31)\), then we have following equivalent statements.

(i) For every continuous convex function \( f : \var{1} \to \mathbb{R} \), we have 

\[
I_f(p, q, u) - I_f(r, q, u) \geq 0.
\]

(37)
(ii) For all \( v \in [\vartheta_1, \vartheta_2] \), we have

\[
I_{G_d} (p, q, u) - I_{G_d} (r, q, u) \geq 0, \quad d = 1, 2, 3, 4, 5. \tag{38}
\]

Moreover, if we change the sign of inequality in both inequalities (37) and (38), then the above result still holds.

**Proof.** The scheme of proof is similar for each \( d = 1, 2, 3, 4, 5 \), therefore we will only give the proof for \( d = 5 \).

(i) \( \Rightarrow \) (ii): Let statement (i) holds. As the function \( G_5 : [\vartheta_1, \vartheta_2] \times [\vartheta_1, \vartheta_2] \rightarrow \mathbb{R} \) is convex and continuous, so it will satisfy the condition (37), i.e.,

\[
I_{G_5} (p, q, u) - I_{G_5} (r, q, u) \geq 0.
\]

(ii) \( \Rightarrow \) (i): Let \( f : [\vartheta_1, \vartheta_2] \rightarrow \mathbb{R} \) be a convex function such that \( f \in C^2([\vartheta_1, \vartheta_2]) \), and further, assume that the statement (ii) holds. Then by Lemma 3.1, we have

\[
f(x_i) = f(\vartheta_1) + (\vartheta_2 - \vartheta_1)f'(\vartheta_1) - (\vartheta_2 - x_i)f'(\vartheta_2) + \int_{\vartheta_1}^{\vartheta_2} G_5(x_i, v)f''(v)dv,
\]

\[
f(y_i) = f(\vartheta_1) + (\vartheta_2 - \vartheta_1)f'(\vartheta_1) - (\vartheta_2 - y_i)f'(\vartheta_2) + \int_{\vartheta_1}^{\vartheta_2} G_5(y_i, v)f''(v)dv.
\]

From (39) and (40), we get

\[
I_f (p, q, u) - I_f (r, q, u) = \sum_{i=1}^{n} u_i q_i f \left( \frac{p_i}{q_i} \right) - \sum_{i=1}^{n} u_i q_i f \left( \frac{r_i}{q_i} \right)
= - \sum_{i=1}^{n} u_i q_i \left( \vartheta_2 - \vartheta_1 \right) f'(\vartheta_2) + \sum_{i=1}^{n} u_i q_i \left( \vartheta_2 - \vartheta_1 \right) f'(\vartheta_2)
+ \int_{\vartheta_1}^{\vartheta_2} \left[ \sum_{i=1}^{n} u_i q_i G_5 \left( \frac{p_i}{q_i}, v \right) - \sum_{i=1}^{n} u_i q_i G_5 \left( \frac{r_i}{q_i}, v \right) \right] f''(v)dv.
\]

Using (36), we have

\[
I_f (p, q, u) - I_f (r, q, u) = \int_{\vartheta_1}^{\vartheta_2} \left[ \sum_{i=1}^{n} u_i q_i G_5 \left( \frac{p_i}{q_i}, v \right) - \sum_{i=1}^{n} u_i q_i G_5 \left( \frac{r_i}{q_i}, v \right) \right] f''(v)dv.
\]

As \( f \) is convex function, therefore \( f''(v) \geq 0 \) for all \( v \in [\vartheta_1, \vartheta_2] \). Hence using (38) in (42), we get (37).

Note that the condition for the existence of second derivative of \( f \) is not necessary ([21, p.172]). As it is possible to approximate uniformly a continuous convex function by convex polynomials, so we can directly eliminate this differentiability condition.

The following theorem gives equivalent statements between continuous convex functions and Green functions via majorization inequality and "useful" Zipf-Mandelbrot law.

**Theorem 3.3.** Assume \( n \in \{1, 2, 3, \ldots\} \), \( t_1, t_2 \geq 0 \) and \( s_1, s_2 > 0 \) such that satisfying

\[
\sum_{i=1}^{n} \frac{u_i}{(i + t_2)^{s_2}} = \frac{H_{n, t_2, s_2}}{H_{n, t_1, s_1}} \sum_{i=1}^{n} \frac{u_i}{(i + t_1)^{s_1}},
\]

with

\[
\frac{1}{q_i(i + t_1)^{s_1}H_{n, t_1, s_1}}, \quad \frac{1}{q_i(i + t_2)^{s_2}H_{n, t_2, s_2}} \epsilon f \quad (i = 1, \ldots, n).
\]

If \( \frac{(i + t_1)^{s_1}}{(i + t_2)^{s_2}} \leq \frac{q_i}{q_i} \) \( (i = 1, \ldots, n) \) and \( G_d (d = 1, 2, 3, 4, 5) \) be defined as in (27)-(31), then we have following equivalent statements.
(i) For every continuous convex function \( f : [\vartheta_1, \vartheta_2] \to \mathbb{R} \), we have
\[
I_f(i, n, t_1, s_1, q, u) - I_f(i, n, t_2, s_2, q, u) \geq 0. \tag{44}
\]

(ii) For all \( \nu \in [\vartheta_1, \vartheta_2] \), we have
\[
I_{G_d}(i, n, t_1, s_1, q, u) - I_{G_d}(i, n, t_2, s_2, q, u) \geq 0, \quad d = 1, 2, 3, 4, 5. \tag{45}
\]
Moreover, if we change the sign of inequality in both inequalities \((44)\) and \((45)\), then the above result still holds.

**Proof.** (i) \( \Rightarrow \) (ii): The proof is similar to the proof of Theorem 3.2.

(ii) \( \Rightarrow \) (i): Let \( f : [\vartheta_1, \vartheta_2] \to \mathbb{R} \) be a convex function such that \( f \in C^2([\vartheta_1, \vartheta_2]) \), and further, assume that the statement (ii) holds. Then by Lemma 3.1, we have (39) and (40).

From (39) and (40), we get
\[
I_f(i, n, t_1, s_1, q, u) - I_f(i, n, t_2, s_2, q, u) = \sum_{i=1}^{n} u_i q_i f'(\lambda_i) - \sum_{i=1}^{n} u_i q_i f'\mu_i)
\]
\[
= -\sum_{i=1}^{n} u_i q_i (\vartheta_2 - \lambda_i) f'(\vartheta_2) + \sum_{i=1}^{n} u_i q_i (\vartheta_2 - \mu_i) f'(\vartheta_2)
\]
\[
+ \int_{\vartheta_1}^{\vartheta_2} \sum_{i=1}^{n} u_i q_i G_5(\lambda_i, \nu) - \sum_{i=1}^{n} u_i q_i G_5(\mu_i, \nu) f''(\nu) d\nu,
\]
where,
\[
\lambda_i := \frac{1}{q_i(1 + t_1)^{s_1} H_{n, t_1, s_1}}, \quad \text{and} \quad \mu_i := \frac{1}{q_i(1 + t_2)^{s_2} H_{n, t_2, s_2}}.
\]
Using (43), we have
\[
I_f(i, n, t_1, s_1, q, u) - I_f(i, n, t_2, s_2, q, u) = \int_{\vartheta_1}^{\vartheta_2} \left[ \sum_{i=1}^{n} u_i q_i G_5(\lambda_i, \nu) - \sum_{i=1}^{n} u_i q_i G_5(\mu_i, \nu) \right] f''(\nu) d\nu. \tag{46}
\]
As \( f \) is convex function, therefore \( f''(\nu) \geq 0 \) for all \( \nu \in [\vartheta_1, \vartheta_2] \). Hence using (45) in (46), we get (44). \( \square \)

4 "Useful" information measure in integral form

The following theorem is a slight extension of Lemma 2 in [34] which is proved by Maligranda et al. (also see [35]):

**Theorem 4.1.** Let \( w, x \) and \( y \) be positive functions on \( [a, b] \). Suppose that \( f : [0, \infty) \to \mathbb{R} \) is a convex function and that
\[
\int_{a}^{\nu} y(t) w(t) \, dt \leq \int_{a}^{\nu} x(t) w(t) \, dt, \quad \nu \in [a, b] \quad \text{and}
\]
\[
\int_{a}^{b} y(t) w(t) \, dt = \int_{a}^{b} x(t) w(t) \, dt.
\]
(i) If \( y \) is a decreasing function on \( [a, b] \), then
\[
\int_{a}^{b} f(y(t)) w(t) \, dt \leq \int_{a}^{b} f(x(t)) w(t) \, dt. \tag{47}
\]
(ii) If \( x \) is an increasing function on \([a, b]\), then
\[
\int_a^b f(x(t)) w(t) \, dt \leq \int_a^b f(y(t)) w(t) \, dt.
\]  
(48)

If \( f \) is strictly convex function and \( x \neq y \) (a.e.), then (47) and (48) are strict.

We consider "useful" Csiszár functional \([11, 12]\) in integral form:

**Definition 4.2** ("Useful" Csiszár divergence as integral form). Assume \( J := [\alpha, \beta] \subset \mathbb{R} \) be an interval, and let \( f : J \to \mathbb{R} \) be a function with densities \( p : [a, b] \to f, q : [a, b] \to (0, \infty) \) and associated with the utility density \( u : [a, b] \to J \) such that

\[
\frac{p(x)}{q(x)} \in J, \quad \forall x \in [a, b],
\]

then we denote "useful" Csiszár divergence in integral form as

\[
\hat{I}_f(p, q, u) := \int_a^b u(t) q(t) f \left( \frac{p(t)}{q(t)} \right) dt.
\]  
(49)

**Remark 4.3.** One can easily seen that if we substitute \( u(t) = 1 \) for all \( t \in [a, b] \), then (49) becomes

\[
\hat{I}_f(p, q, 1) := \hat{I}_f(p, q) = \int_a^b q(t) f \left( \frac{p(t)}{q(t)} \right) dt.
\]

**Theorem 4.4.** Assume \( J := [0, \infty) \subset \mathbb{R} \) be an interval, \( f : J \to \mathbb{R} \) be a convex function and \( p, q, r, u : [a, b] \to (0, \infty) \) such that

\[
\int_a^u u(t) r(t) \, dt \leq \int_a^u u(t) p(t) \, dt, \quad \forall u \in [a, b]
\]  
(50)

and

\[
\int_a^b u(t) r(t) \, dt = \int_a^b u(t) p(t) \, dt,
\]  
(51)

with

\[
\frac{p(t)}{q(t)} \frac{r(t)}{q(t)} \in J, \quad \forall t \in [a, b].
\]

(i) If \( \frac{r(t)}{q(t)} \) is a decreasing function on \([a, b]\), then

\[
\hat{I}_f(r, q, u) \leq \hat{I}_f(p, q, u).
\]  
(52)

(ii) If \( \frac{p(t)}{q(t)} \) is an increasing function on \([a, b]\), then the inequality is reversed, i.e.

\[
\hat{I}_f(r, q, u) \geq \hat{I}_f(p, q, u).\]  
(53)

If \( f \) is strictly convex function and \( p(t) \neq r(t) \) (a.e.), then strict inequality holds in (52) and (53).

If \( f \) is concave function then the reverse inequalities hold in (52) and (53). If \( f \) is strictly concave and \( p(t) \neq r(t) \) (a.e.), then the strict reverse inequalities hold in (52) and (53).

**Proof.** (i): We use Theorem 4.1 (i) with substitutions \( x(t) := \frac{p(t)}{q(t)}, y(t) := \frac{r(t)}{q(t)}, w(t) := u(t) q(t) > 0 \forall t \in [a, b] \) and also using the fact that \( \frac{r(t)}{q(t)} \) is a decreasing function then we get (52).

(ii) We can prove with the similar substitutions as in the first part by using Theorem 4.1 (ii) that is the fact that \( \frac{p(t)}{q(t)} \) is an increasing function.

**Remark 4.5.** We can give Theorem 4.4 for \( u(t) := 1 \) for all \( t \in [a, b] \) as special case which has been given in [36].
5 Applications

Here, we present several special cases of the previous results as applications.

The first case corresponds to the entropy of a continuous probability density (see [18, p.506]):

**Definition 5.1 (Shannon Entropy).** Let \( p : [a, b] \to (0, \infty) \) be a positive probability density, then the Shannon entropy of \( p(x) \) is defined by

\[
H(p(x), u(x)) := - \int_a^b u(x) p(x) \log p(x) \, dx,
\]

and is associated with the utility density \( u : [a, b] \to \mathbb{R} \), whenever the integral exists.

Note that there is no problem with the definition in the case of a zero probability, since

\[
\lim_{x \to 0} x \log x = 0.
\]

**Corollary 5.2.** Assume \( p, q, r, u : [a, b] \to (0, \infty) \) be functions such that satisfying (50) and (51) with

\[
p(t) q(t) \quad \frac{r(t)}{q(t)} \quad \epsilon \in J := (0, \infty), \quad \forall t \in [a, b].
\]

(i) If \( \frac{r(t)}{q(t)} \) is a decreasing function and the base of log is greater than 1, then we have estimates for the Shannon entropy of \( q(t) \) associated with utility density \( u(t) \)

\[
\int_a^b u(t) q(t) \log \left( \frac{r(t)}{q(t)} \right) \geq H(q(t), u(t)).
\]

(ii) If \( \frac{r(t)}{q(t)} \) is an increasing function and the base of log is greater than 1, then we have estimates for the Shannon entropy of \( q(t) \) associated with utility density \( u(t) \)

\[
H(q(t), u(t)) \leq \int_a^b u(t) q(t) \log \left( \frac{p(t)}{q(t)} \right).
\]

Proof. (i): Substitute \( f(x) := -\log x \) and \( p(t) := 1, \forall t \in [a, b] \) in Theorem 4.4 (i) then we get (56).

(ii) We can prove by switching the role of \( p(t) \) with \( r(t) \) i.e., \( r(t) := 1 \forall t \in [a, b] \) and \( f(x) := -\log x \) in Theorem 4.4 (ii) then we get (57).

The second case corresponds to the relative entropy or the Kullback-Leibler divergence between two probability densities associated with the utility density \( u(t) \):

**Definition 5.3 (Kullback-Leibler Divergence).** Let \( p, q : [a, b] \to (0, \infty) \) be a positive probability densities, then the Kullback-Leibler (K-L) divergence between \( p(t) \) and \( q(t) \) is defined by

\[
L(p(t), q(t), u(t)) := \int_a^b u(t) p(t) \log \left( \frac{p(t)}{q(t)} \right) \, dt,
\]

and is associated with the utility density \( u : [a, b] \to \mathbb{R} \).

**Corollary 5.4.** Assume \( p, q, r, u : [a, b] \to (0, \infty) \) be functions such that satisfying (50) and (51) with

\[
p(t) q(t) \quad \frac{r(t)}{q(t)} \quad \epsilon \in J := (0, \infty), \quad \forall t \in [a, b].
\]
(i) If \( \frac{r(t)}{q(t)} \) is a decreasing function and the base of log is greater than 1, then
\[
\hat{I}_{(-\log x)}(r, q, u) \geq \hat{I}_{(-\log x)}(p, q, u). \tag{58}
\]

If the base of log is in between 0 and 1, then the reverse inequality holds in (58).

(ii) If \( \frac{p(t)}{q(t)} \) is an increasing function and the base of log is greater than 1, then
\[
\hat{I}_{(-\log x)}(r, q, u) \leq \hat{I}_{(-\log x)}(p, q, u). \tag{59}
\]

If the base of log is in between 0 and 1 then the reverse inequality holds in (59).

Proof. (i): Substitute \( f(x) := -\log x \) in Theorem 4.4 (i) then we get (58).

(ii) We can prove with substitution \( f(x) := -\log x \) in Theorem 4.4 (ii).

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence.

**Definition 5.5** (Variational Distance). Let \( p, q : [a, b] \to (0, \infty) \) be a positive probability densities, then variation distance between \( p(t) \) and \( q(t) \) is defined by
\[
\hat{I}_v(p(t), q(t), u(t)) := \int_a^b u(t) |p(t) - q(t)| dt,
\]
and associated with the utility density \( u : [a, b] \to \mathbb{R} \).

**Corollary 5.6.** Assume \( p, q, r, u : [a, b] \to (0, \infty) \) be functions such that satisfying (50) and (51) with
\[
\frac{p(t)}{q(t)}, \frac{r(t)}{q(t)} \in \mathcal{J} := (0, \infty), \quad \forall t \in [a, b].
\]

(i) If \( \frac{r(t)}{q(t)} \) is a decreasing function, then
\[
\hat{I}_v(r(t), q(t), u(t)) \leq \hat{I}_v(p(t), q(t), u(t)). \tag{60}
\]

(ii) If \( \frac{p(t)}{q(t)} \) is an increasing function, then the inequality is reversed, i.e.
\[
\hat{I}_v(r(t), q(t), u(t)) \geq \hat{I}_v(p(t), q(t), u(t)). \tag{61}
\]

Proof. (i): Since \( f(x) := |x - 1| \) be a convex function for \( x \in \mathbb{R}^+ \), therefore substitute \( f(x) := |x - 1| \) in Theorem 4.4 (i) then
\[
\int_a^b u(t)q(t) \left| \frac{r(t)}{q(t)} - 1 \right| dt \leq \int_a^b u(t)q(t) \left| \frac{p(t)}{q(t)} - 1 \right| dt,
\]
\[
\int_a^b u(t)q(t) \left| \frac{r(t) - q(t)}{|q(t)|} \right| dt \leq \int_a^b u(t)q(t) \left| \frac{p(t) - q(t)}{|q(t)|} \right| dt,
\]
since \( q(t) > 0 \) then we get (60).

(ii) We can prove with substitution \( f(x) := |x - 1| \) in Theorem 4.4 (ii).

**Definition 5.7** (Hellinger Distance). Let \( p, q : [a, b] \to (0, \infty) \) be a positive probability densities, then the Hellinger distance between \( p(t) \) and \( q(t) \) is defined by
\[
\hat{I}_H(p(t), q(t), u(t)) := \int_a^b u(t) \left[ \sqrt{p(t)} - \sqrt{q(t)} \right]^2 dt,
\]
and is associated with the utility density \( u : [a, b] \to \mathbb{R} \).
Corollary 5.8. Assume \( p, q, r, u : [a, b] \to (0, \infty) \) be functions such that satisfying (50) and (51) with
\[
\frac{p(t)}{q(t)} \cdot \frac{r(t)}{q(t)} \in J := (0, \infty), \quad \forall t \in [a, b].
\]

(i) If \( \frac{r(t)}{q(t)} \) is a decreasing function, then
\[
\hat{I}_H (r(t), q(t), u(t)) \leq \hat{I}_H (p(t), q(t), u(t)).
\]

(ii) If \( \frac{r(t)}{q(t)} \) is an increasing function, then the inequality is reversed, i.e.
\[
\hat{I}_H (r(t), q(t), u(t)) \geq \hat{I}_H (p(t), q(t), u(t)).
\]

Proof. (i): Since \( f(x) := (\sqrt{x} - 1)^2 \) is a convex function for \( x \in \mathbb{R}^+ \), therefore substituting \( f(x) := (\sqrt{x} - 1)^2 \) in Theorem 4.4 (i)
\[
\int_a^b u(t) q(t) \left( \sqrt{\frac{r(t)}{q(t)}} - 1 \right)^2 dt \leq \int_a^b u(t) q(t) \left( \sqrt{\frac{p(t)}{q(t)}} - 1 \right)^2 dt,
\]

since \( q(t) > 0 \) then we get (62).

(ii) We can prove with substitution \( f(x) := (\sqrt{x} - 1)^2 \) in Theorem 4.4 (ii). \qed

Definition 5.9 (Bhattacharyya Distance). Let \( p, q : [a, b] \to (0, \infty) \) be a positive probability densities, then the Bhattacharyya distance between \( p(t) \) and \( q(t) \) is defined by
\[
\hat{I}_B (p(t), q(t), u(t)) := \int_a^b u(t) \sqrt{p(t) q(t)} dt,
\]

and associated with the utility density \( u : [a, b] \to \mathbb{R} \).

Corollary 5.10. Assume \( p, q, r, u : [a, b] \to (0, \infty) \) be functions such that satisfying (50) and (51) with
\[
\frac{p(t)}{q(t)} \cdot \frac{r(t)}{q(t)} \in J := (0, \infty), \quad \forall t \in [a, b].
\]

(i) If \( \frac{r(t)}{q(t)} \) is a decreasing function, then
\[
\hat{I}_B (p(t), q(t), u(t)) \leq \hat{I}_B (r(t), q(t), u(t)).
\]

(ii) If \( \frac{r(t)}{q(t)} \) is an increasing function, then the inequality is reversed, i.e.
\[
\hat{I}_B (p(t), q(t), u(t)) \geq \hat{I}_B (r(t), q(t), u(t)).
\]

Proof. (i): Since \( f(x) := -\sqrt{x} \) be a convex function for \( x \in \mathbb{R}^+ \), therefore substitute \( f(x) := -\sqrt{x} \) in Theorem 4.4 (i) then
\[
\int_a^b u(t) q(t) \left( -\sqrt{\frac{r(t)}{q(t)}} \right) dt \leq \int_a^b u(t) q(t) \left( -\sqrt{\frac{p(t)}{q(t)}} \right) dt,
\]

we get (64).

(ii) We can prove with substitution \( f(x) := -\sqrt{x} \) in Theorem 4.4 (ii). \qed

Definition 5.11 (Jeffreys Distance). Let \( p, q : [a, b] \to (0, \infty) \) be a positive probability densities, then the Jeffreys distance between \( p(t) \) and \( q(t) \) is defined by
\[
\hat{I}_J (p(t), q(t), u(t)) := \int_a^b u(t) \left[ p(t) - q(t) \right] \ln \left( \frac{p(t)}{q(t)} \right) dt,
\]

and associated with the utility density \( u : [a, b] \to \mathbb{R} \).  

Unauthenticated  
Download Date | 6/2/19 5:59 AM
Corollary 5.12. Assume $p, q, r, u : [a, b] \to (0, \infty)$ be functions such that satisfying (50) and (51) with
$$\frac{p(t)}{q(t)}, \frac{r(t)}{q(t)} \in J := (0, \infty), \quad \forall t \in [a, b].$$

(i) If $\frac{r(t)}{q(t)}$ is a decreasing function, then
$$\hat{I}_f(r(t), q(t), u(t)) \leq \hat{I}_f(p(t), q(t), u(t)).$$  \hspace{1cm} (66)

(ii) If $\frac{p(t)}{q(t)}$ is an increasing function, then the inequality is reversed, i.e.
$$\hat{I}_f(r(t), q(t), u(t)) \geq \hat{I}_f(p(t), q(t), u(t)).$$  \hspace{1cm} (67)

Proof. (i): Since $f(x) := (x - 1) \ln x$ be a convex function for $x \in \mathbb{R}^+$, therefore substituting $f(x) := (x - 1) \ln x$ in Theorem 4.4 (i)
$$\int_a^b u(t) q(t) \left( \frac{r(t)}{q(t)} - 1 \right) \ln \left( \frac{r(t)}{q(t)} \right) dt$$
$$\leq \int_a^b u(t) q(t) \left( \frac{p(t)}{q(t)} - 1 \right) \ln \left( \frac{p(t)}{q(t)} \right) dt,$$
we get (66).

(ii) We can prove with substitution $f(x) := (x - 1) \ln x$ in Theorem 4.4 (ii). \hfill \qed

Definition 5.13 (Triangular Discrimination). Let $p, q : [a, b] \to (0, \infty)$ be a positive probability densities, then the triangular discrimination between $p(t)$ and $q(t)$ is defined by
$$\hat{I}_\Delta(p(t), q(t), u(t)) = \int_a^b u(t) \left[ \frac{p(t) - q(t)}{p(t) + q(t)} \right]^2 dt,$$
and is associated with the utility density $u : [a, b] \to \mathbb{R}$.

Corollary 5.14. Assume $p, q, r, u : [a, b] \to (0, \infty)$ be functions such that satisfying (50) and (51) with
$$\frac{p(t)}{q(t)}, \frac{r(t)}{q(t)} \in J := (0, \infty), \quad \forall t \in [a, b].$$

(i) If $\frac{r(t)}{q(t)}$ is a decreasing function, then
$$\hat{I}_\Delta(r(t), q(t), u(t)) \leq \hat{I}_\Delta(p(t), q(t), u(t)).$$  \hspace{1cm} (68)

(ii) If $\frac{p(t)}{q(t)}$ is an increasing function, then the inequality is reversed, i.e.
$$\hat{I}_\Delta(r(t), q(t), u(t)) \geq \hat{I}_\Delta(p(t), q(t), u(t)).$$  \hspace{1cm} (69)

Proof. (i): Since $f(x) := \frac{(x - 1)^2}{x + 1}$ be a convex function for $x \geq 0$, therefore substitute $f(x) := \frac{(x - 1)^2}{x + 1}$ in Theorem 4.4 (i) then
$$\int_a^b u(t) q(t) \frac{(r(t) - q(t))^2}{r(t) + 1} dt \leq \int_a^b u(t) q(t) \frac{(p(t) - q(t))^2}{p(t) + 1} dt,$$
$$\int_a^b u(t) q(t) \frac{(r(t) - q(t))^2}{r(t) + q(t)} dt \leq \int_a^b u(t) q(t) \frac{(p(t) - q(t))^2}{p(t) + q(t)} dt,$$
we get (68).

(ii) We can prove with substitution $f(x) := \frac{(x - 1)^2}{x + 1}$ in Theorem 4.4 (ii). \hfill \qed

Remark 5.15. We can give all the results of section 5 for $u(t) = 1$ for all $t \in [a, b]$ as a special case, which has been given in [36].
Author's contribution
All authors contributed equally. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Acknowledgement: The publication was supported by the Ministry of Education and Science of the Russian Federation (the Agreement number No. 02.a03.21.0008.) This publication is partially supported by Royal Commission (RC) Jubail Industrial College, Jubail, Kingdom of Saudi Arabia.

References

Majorization, "useful" Csiszár divergence and "useful" Zipf-Mandelbrot law