ON AN INEQUALITY FOR LOGARITHMS AND APPLICATIONS IN INFORMATION THEORY

S.S. Dragomir, N.M. Dragomir and K. Pranesh

Abstract. A new analytic inequality for logarithms which provides a converse to arithmetic mean-geometric mean inequality and its applications in information theory are given.

1 INTRODUCTION

The present paper continues the investigations started in [1], where the main result is

Theorem 1.1. Let $\xi_k \in (0, \infty), p_k > 0, k = 1, \ldots, n$ with $\sum_{k=1}^{n} p_k = 1$ and $b > 1$. Then

$$0 \leq \log_b \left( \sum_{k=1}^{n} p_k \xi_k \right) - \sum_{k=1}^{n} p_k \log_b \xi_k$$

$$\leq \frac{1}{2 \ln b} \sum_{k=1}^{n} p_k \left( \xi_k - \xi_k \right)^2.$$ 

The equality holds in both inequalities simultaneously if and only if $\xi_1 = \ldots = \xi_n$.

2 A NEW ANALYTIC INEQUALITY FOR LOGARITHMS

We shall start to the following analytic inequality for logarithms which provides a different bound than the inequality of Dragomir-Goh (1.1):

Theorem 2.1. Let $\xi_k \in [1, \infty)$ and $p_k > 0$ with $\sum_{k=1}^{n} p_k = 1$ and $b > 1$. Then we have

$$0 \leq \log_b \left( \sum_{k=1}^{n} p_k \xi_k \right) - \sum_{k=1}^{n} p_k \log_b \xi_k$$

$$\leq \frac{1}{4 \ln b} \sum_{i,j=1}^{n} p_i p_j \left( \xi_i - \xi_j \right)^2.$$ 

The equality holds in both inequalities simultaneously if and only if $\xi_1 = \ldots = \xi_n$.

Proof. We shall use the well known Jensen’s discrete inequality for convex mappings which states that

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

for all $p_i > 0, \sum_{i=1}^{n} p_i = 1, f$ a convex mapping on a given interval $I$ and $x_i \in I (i = 1, \ldots, n).$

Date. December 14, 1998

1991 Mathematics Subject Classification. Primary 26 D 15, 26 D xx; Secondary 94 X xx.

Key words and phrases. Jensen’s Inequality, Analytic Inequalities, Entropy Mapping, Joint Entropy.
Now, let consider the mapping \( f : [1, \infty) \to \mathbb{R} \), \( f(x) = \frac{x^2}{2} + \ln x \). Then

\[
f'(x) = x + \frac{1}{x} = \frac{x^2 + 1}{x} \quad \text{for all } x \in [1, \infty)
\]

and

\[
f''(x) = 1 - \frac{1}{x^2} = \frac{x^2 - 1}{x^2} \quad \text{for all } x \in [1, \infty),
\]

i.e., \( f \) is a strictly convex mapping on \([1, \infty)\).

Applying Jensen’s discrete inequality for convex mappings, we have

\[
\frac{1}{2} \left( \sum_{i=1}^{n} p_i \xi_i \right)^2 + \ln \left( \sum_{i=1}^{n} p_i \xi_i \right) \leq \frac{1}{2} \sum_{i=1}^{n} p_i \xi_i^2 + \sum_{i=1}^{n} p_i \ln \xi_i
\]

which is equivalent to

\[
\ln \left( \sum_{i=1}^{n} p_i \xi_i \right) - \sum_{i=1}^{n} p_i \ln \xi_i \leq \frac{1}{2} \left[ \sum_{i=1}^{n} p_i \xi_i^2 - \left( \sum_{i=1}^{n} p_i \xi_i \right)^2 \right].
\]

But

\[
\sum_{i,j=1}^{n} p_i p_j (\xi_i - \xi_j)^2 = \sum_{i,j=1}^{n} p_i p_j \left[ \xi_i^2 + \xi_j^2 - 2 \xi_i \xi_j \right]
\]

\[
= 2 \left[ \sum_{i=1}^{n} p_i \sum_{i=1}^{n} p_i \xi_i^2 - \left( \sum_{i=1}^{n} p_i \xi_i \right)^2 \right] = 2 \left[ \sum_{i=1}^{n} p_i \xi_i^2 - \left( \sum_{i=1}^{n} p_i \xi_i \right)^2 \right]
\]

and then the above inequality becomes

\[
\ln \left( \sum_{i=1}^{n} p_i \xi_i \right) - \sum_{i=1}^{n} p_i \ln \xi_i \leq \frac{1}{4} \sum_{i,j=1}^{n} p_i p_j (\xi_i - \xi_j)^2.
\]

Now, as \( \log_b x = \frac{\ln x}{\ln b} \), the inequality (2.4) is equivalent to the desired inequality (2.1).

The case of equality follows by the strict convexity of \( f \) and we omit the details. □

**Remark 2.1.** Define

\[
B_1 := \frac{1}{2 \ln b} \sum_{i,j=1}^{n} p_i p_j \frac{\xi_i - \xi_j}{\xi_i \xi_j}^2 \quad \text{(as in Theorem 1.1)}
\]

and

\[
B_2 := \frac{1}{4 \ln b} \sum_{i,j=1}^{n} p_i p_j (\xi_i - \xi_j)^2 \quad \text{(as in Theorem 2.1)}
\]

and compute the difference

\[
B_1 - B_2 = \frac{1}{2 \ln b} \sum_{i,j=1}^{n} p_i p_j \frac{\xi_i - \xi_j}{\xi_i \xi_j}^2 \left[ \frac{1}{\xi_i \xi_j} - 2 \right]
\]

\[
= \frac{1}{4 \ln b} \sum_{i,j=1}^{n} p_i p_j \frac{\xi_i - \xi_j}{\xi_i \xi_j} \left( 2 - \xi_i \xi_j \right).
\]

Consequently, if \( \xi_i, \xi_j \in [1, \infty) \) so that \( \xi_i \xi_j \leq 2 \) for all \( i, j \in \{1, \ldots, n\} \), then the bound \( B_2 \) provided by Theorem 2.1 is better than the bound \( B_1 \) provided by Theorem 1.1. If \( \xi_i, \xi_j \in [1, \infty) \) so that \( \xi_i \xi_j \geq 2 \) for all \( i, j \in \{1, \ldots, n\} \), then Theorem 1.1 provides a better result than Theorem 2.1.
We give now some applications of the above results for arithmetic mean-geometric mean inequality.

Recall that for \( q \geq 0 \) with \( Q_n := \sum_{i=1}^{n} q_i \), the arithmetic mean of \( x \) with the weights \( q, i \in \{1, \ldots, n\} \) is

\[
A_n \left( \bar{q}, \bar{x} \right) := \frac{1}{Q_n} \sum_{i=1}^{n} q_i x_i
\]

and the geometric mean of \( x \) with the weights \( q, i \in \{1, \ldots, n\} \), is

\[
G_n \left( \bar{q}, \bar{x} \right) := \left( \prod_{i=1}^{n} x_i^{q_i} \right)^{\frac{1}{Q_n}}.
\]

It is well known that the following inequality so called arithmetic mean-geometric mean inequality, holds

\[
(2.5) \quad A_n \left( \bar{q}, \bar{x} \right) \geq G_n \left( \bar{q}, \bar{x} \right)
\]

with equality if and only if \( x_1 = \ldots = x_n \).

Now, using Theorem 1.1, we can state the following proposition containing a counterpart of the arithmetic mean-geometric mean inequality (2.5):

**Proposition 2.2.** With the above assumptions for \( \bar{q} \) and \( \bar{x} \), we have

\[
(2.6) \quad 1 \leq \frac{A_n \left( \bar{q}, \bar{x} \right)}{G_n \left( \bar{q}, \bar{x} \right)} \leq \exp_b \left[ \frac{1}{2Q_n^2} \ln b \sum_{i,j=1}^{n} \frac{q_i q_j}{x_i x_j} (x_i - x_j)^2 \right]
\]

where \( \exp_b (x) = b^x \) (\( b > 1 \)). The equality holds in both inequalities simultaneously if and only if \( x_1 = \ldots = x_n \).

Also, using Theorem 2.1, we have another converse inequality for (2.5).

**Proposition 2.3.** Let \( \bar{q} \) be as above and \( \bar{x} \in \mathbb{R}^n \), so that \( x_i \geq 1, i = 1, \ldots, n \). Then we have the inequality:

\[
(2.7) \quad 1 \leq \frac{A_n \left( \bar{q}, \bar{x} \right)}{G_n \left( \bar{q}, \bar{x} \right)} \leq \exp_b \left[ \frac{1}{4Q_n^2} \ln b \sum_{i,j=1}^{n} \frac{q_i q_j}{x_i x_j} (x_i - x_j)^2 \right]
\]

where \( b > 1 \). The equality holds in both inequalities simultaneously if and only if \( x_1 = \ldots = x_n \).

**Remark 2.2.** As in the previous remark, if \( 1 \leq x_i x_j \leq 2 \) then the bound (2.7) is better than (2.6). If \( x_i x_j \geq 2 \), then (2.6) is better than (2.7).

### 3 Applications For The Entropy Mapping

Let us consider now, the \( b \)-entropy mapping of the discrete random variable \( X \) with \( n \) possible outcomes and having the probability distribution \( p = (p_i) \), \( i = \{1, \ldots, n\} \):

\[
H_b (X) = \sum_{i=1}^{n} p_i \log_b \left( \frac{1}{p_i} \right).
\]

We know (see [1]) that the following counterpart inequality holds:

\[
(3.1) \quad 0 \leq \log_b n - H_b (X) \leq \frac{1}{2 \ln b} \sum_{i,j=1}^{n} (p_i - p_j)^2
\]

with equality if and only if \( p_i = \frac{1}{n} \) for all \( i \in \{1, \ldots, n\} \).

The following similar result also holds:
Theorem 3.1. Let $X$ be as above. Then we have

$$0 \leq \log_b n - H_b(X) \leq \frac{1}{4 \ln b} \sum_{i,j=1}^{n} \left( \frac{p_i - p_j}{p_i p_j} \right)^2. \tag{3.2}$$

The equality holds if and only if $p_i = \frac{1}{n}$ for all $i \in \{1, \ldots, n\}$.

Proof. As $p_i \in (0, 1]$, then $\xi = \frac{1}{p_i} \in [1, \infty)$ and we can apply Theorem 2.1 to get

$$0 \leq \log_b n - H_b(X) \leq \frac{1}{4 \ln b} \sum_{i,j=1}^{n} p_i p_j \left( \frac{1}{p_i} - \frac{1}{p_j} \right)^2 \leq \frac{1}{4 \ln b} \sum_{i,j=1}^{n} \left( \frac{p_i - p_j}{p_i p_j} \right)^2.\tag{3.3}$$

The equality holds iff $\xi_i = \xi_j$ for all $i, j \in \{1, \ldots, n\}$ which is equivalent to $p_i = p_j$ for all $i, j \in \{1, \ldots, n\}$, i.e., $p_i = \frac{1}{n}$ for all $i \in \{1, \ldots, n\}$. \hspace{1cm} []

The following corollary is important in applications as it provides a sufficient condition for the probability $p$ so that $\log_b n - H_b(X)$ is small enough.

Corollary 3.2. Let $X$ be as above and $\varepsilon > 0$. If the probabilities $p_i, i = 1, \ldots, n$, satisfy the conditions:

$$\frac{1}{2} \left[ 2 + k - \sqrt{k(k+4)} \right] \leq \frac{p_i}{p_j} \leq \frac{1}{2} \left[ 2 + k + \sqrt{k(k+4)} \right] \tag{3.4}$$

for all $1 \leq i < j \leq n$, where

$$k = \frac{4\varepsilon \ln b}{n(n-1)} \quad (n \geq 2),$$

then we have the estimation

$$0 \leq \log_b n - H_b(X) \leq \varepsilon.$$

Proof. Let observe that

$$\frac{1}{4 \ln b} \sum_{i,j=1}^{n} \left( \frac{p_i - p_j}{p_i p_j} \right)^2 = \frac{1}{2 \ln b} \sum_{1 \leq i < j \leq n} \left( \frac{p_i - p_j}{p_i p_j} \right)^2. \tag{3.3}$$

Suppose that

$$\frac{(p_i - p_j)^2}{p_i p_j} \leq k, \quad \text{for } 1 \leq i < j \leq n.$$

Then

$$p_i^2 - (2 + k) p_i p_j + p_j^2 \leq 0 \quad \text{for } 1 \leq i < j \leq n.$$

Denoting $t = \frac{p_i}{p_j}$, the above inequality is equivalent to $t^2 - (2 + k) t + 1 \leq 0$, i.e., $t \in [t_1, t_2]$, where

$$t_1 = \frac{2 + k - \sqrt{k(k+4)}}{2} \quad \text{and} \quad t_2 = \frac{2 + k + \sqrt{k(k+4)}}{2}.$$

If we choose $k = \frac{4\varepsilon \ln b}{n(n-1)}$, then by (3.2) we have

$$0 \leq \log_b n - H_b(X) \leq \frac{1}{4 \ln b} \sum_{i,j=1}^{n} \left( \frac{p_i - p_j}{p_i p_j} \right)^2 \leq \frac{1}{2 \ln b} \sum_{1 \leq i < j \leq n} \left( \frac{p_i - p_j}{p_i p_j} \right)^2 \leq \frac{1}{2 \ln b} \sum_{1 \leq i < j \leq n} k = \frac{n(n-1)}{4 \ln b} \cdot \frac{4\varepsilon \ln b}{n(n-1)} = \varepsilon,$$

and the corollary is proved. \hspace{1cm} []
Now, consider the bounds
\[ M_1 := \frac{1}{2 \ln b} \sum_{i,j=1}^{n} (p_i - p_j)^2 \quad \text{(given by (3.1))} \]
and
\[ M_2 := \frac{1}{4 \ln b} \sum_{i,j=1}^{n} \frac{(p_i - p_j)^2}{p_i p_j} \quad \text{(given by (3.2)).} \]

We give an example for which \( M_1 \) is less than \( M_2 \) and another example for which \( M_2 \) is less than \( M_1 \) which will suggest that we can use both of them to estimate the above difference \( \log_b n - H_b(X) \).

Consider the probability distribution:
\[ p_1 = 0.3475, \quad p_2 = 0.2398, \quad p_3 = 0.1654 \]
\[ p_4 = 0.1142, \quad p_5 = 0.0788, \quad p_6 = 0.0544. \]

In this case
\[ \tilde{M}_1 = 6.5119, \quad \tilde{M}_2 = 12.1166, \]
where
\[ \tilde{M}_1 := \frac{1}{2} \sum_{i,j=1}^{n} (p_i - p_j)^2, \quad \tilde{M}_2 := \frac{1}{4} \sum_{i,j=1}^{n} \frac{(p_i - p_j)^2}{p_i p_j} \quad \text{and} \quad n = 6. \]

Consider the probability distribution
\[ p_1 = 0.2468, \quad p_2 = 0.2072, \quad p_3 = 0.1740 \]
\[ p_4 = 0.1461, \quad p_5 = 0.1227, \quad p_6 = 0.1031. \]

In this case,
\[ \tilde{M}_1 = 5.2095, \quad \tilde{M}_2 = 2.3706. \]

4 Bounds for Joint Entropy

Consider the joint entropy of two random variables \( X \) and \( Y \) [3, p. 25]:
\[ H_b(X, Y) := \sum_{x,y} p(x, y) \log_b \frac{1}{p(x, y)} \]
where the joint probability \( p(x, y) = P\{X = x, Y = y\} \).

In paper [2], S.S. Dragomir and C. J. Goh have proved the following result using Theorem 1.1:

**Theorem 4.1.** With the above assumptions, we have that
\[ 0 \leq \log_b (r, s) - H_b(X, Y) \leq \frac{1}{2 \ln b} \sum_{x,y} \sum_{u,v} (p(x, y) - p(u, v))^2 \]
where the range of \( X \) contains \( r \) elements and the range of \( Y \) contains \( s \) elements. Equality holds in both inequalities simultaneously if and only if \( p(x, y) = p(u, v) \) for all \((x, y), (u, v)\).

The following corollary is useful in practice:

**Corollary 4.2.** With the above assumptions and if
\[ \max_{(x,y),(u,v)} |p(x, y) - p(u, v)| \leq \sqrt{\frac{2\varepsilon \ln b}{rs}}, \quad \varepsilon > 0 \]
then we have the estimation
\[ 0 \leq \log_b (r, s) - H_b(X, Y) \leq \varepsilon. \]
Now, using the second converse inequality embodied in Theorem 2.1, we are able to prove another upper bound for the difference \( \log_b (rs) - H_b (X, Y) \).

**Theorem 4.3.** With the above assumptions, we have

\[
0 \leq \log_b (rs) - H_b (X, Y) \leq \frac{1}{4 \ln b} \sum_{x,y} \sum_{u,v} \frac{(p(x,y) - p(u,v))^2}{p(x,y) p(u,v)}
\]

where the range of \( X \) and \( Y \) are as above. Equality holds in both inequalities simultaneously iff \( p(x,y) = p(u,v) \) for all \( (x,y) \) and \( (u,v) \).

**Proof.** Using Theorem 2.1, we have for \( p_i = p(x,y) \) and \( \xi_i = \frac{1}{p(x,y)} \),

\[
0 \leq \log_b \left( \sum_{x,y} p(x,y) \cdot \frac{1}{p(x,y)} \right) - \sum_{x,y} p(x,y) \log_b \left( \frac{1}{p(x,y)} \right)
\]

\[
\leq \frac{1}{4 \ln b} \sum_{x,y} \sum_{u,v} p(x,y) p(u,v) \left( \frac{1}{p(x,y)} - \frac{1}{p(u,v)} \right)^2
\]

\[
= \frac{1}{4 \ln b} \sum_{x,y} \sum_{u,v} \frac{(p(x,y) - p(u,v))^2}{p(x,y) p(u,v)}
\]

which is clearly equivalent to the desired result. The case of equality is obvious by Theorem 2.1. \( \blacksquare \)

The following corollary is important in practical applications:

**Corollary 4.4.** Let \( X \) and \( Y \) be as above and \( \varepsilon > 0 \). Denote \( P = \max p(x,y) \) and \( p = \min p(x,y) \). If

\[
\frac{P}{p} \leq 1 + k + \sqrt{k(k + 2)}
\]

where

\[
k := \frac{2 \varepsilon \ln b}{(rs)^2},
\]

then we have the bound

\[
0 \leq \log_b (rs) - H_b (X, Y) \leq \varepsilon.
\]

**Proof.** At the beginning, let us consider the inequality

\[
\frac{(a - b)^2}{2ab} \leq k, \quad \text{for } a, b > 0 \text{ and } k \geq 0.
\]

This inequality is clearly equivalent to

\[
a^2 - 2 (1 + k) ab + b^2 \leq 0
\]
or, denoting \( t := \frac{a}{b} \), to

\[
t^2 - 2 (1 + k) t + 1 \leq 0,
\]
i.e.,

\[
1 + k - \sqrt{k(k + 2)} \leq t \leq 1 + k + \sqrt{k(k + 2)}.
\]

Now, let suppose that

\[
1 + k - \sqrt{k(k + 2)} \leq \frac{p(x,y)}{p(u,v)} \leq 1 + k + \sqrt{k(k + 2)}
\]
for all \((x, y)\) and \((u, v)\) and \(k := \frac{2\ln b}{1 - b}\). Then by (4.2), we have

\[
0 \leq \log_b (rs) - H_b(X, Y) \leq \frac{1}{4 \ln b} \sum_{x,y} \sum_{u,v} \frac{(p(x, y) - p(u, v))^2}{p(x, y) p(u, v)}
\]

\[
\leq \frac{1}{2 \ln b} (rs)^2 k = \frac{(rs)^2}{2 \ln b} \frac{2 \ln b}{(rs)^2} = \varepsilon.
\]

Now, let observe that the inequality (4.4) is equivalent to:

\[
1 + k - \sqrt{k(k + 2)} \leq \frac{p}{P} \leq \frac{P}{p} \leq 1 + k + \sqrt{k(k + 2)}.
\]

But \(\frac{P}{p} \geq 1 + k - \sqrt{k(k + 2)}\) is equivalent to

\[
\frac{P}{p} \leq \frac{1}{1 + k - \sqrt{k(k + 2)}} = k + 1 + \sqrt{k(k + 2)}
\]

and the corollary is proved. 

**Acknowledgment.** The authors are grateful to the referee for his valuable suggestions.

**References**


