An International Journal
computers & mathematics with applications

PERGAMON Computers and Mathematics with Applications 38 (1999) 11–17

An Inequality for Logarithms and Applications in Information Theory

S. S. DRAGOMIR
School of Communications and Informatics
Victoria University of Technology
P.O. Box 14428, Melbourne City MC, Victoria 8001, Australia
sever@matilda.vut.edu.au

N. M. DRAGOMIR
Department of Mathematics
University of Transkei, Private Bag X1, UNITRA
Umtata, 5117, South Africa

K. PRANESH
Department of Statistics
University of Transkei, Private Bag X1, UNITRA
Umtata, 5117, South Africa

(Received May 1998; revised and accepted March 1999)

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.

Keywords—Jensen's inequality, Analytic inequalities, Entropy mapping, Joint entropy.

1. INTRODUCTION

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

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

$$0 \le \log_b \left(\sum_{k=1}^n p_k \xi_k \right) - \sum_{k=1}^n p_k \log_b \xi_k \le \frac{1}{2\ln b} \sum_{k,i=1}^n \frac{p_k p_i}{\xi_k \xi_i} (\xi_i - \xi_k)^2 . \tag{1.1}$$

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

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

PII: S0898-1221(99)00177-7

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 (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*

$$\begin{aligned} 0 &\le \log_b\left(\sum_{k=1}^n p_k\xi_k\right) - \sum_{k=1}^n p_k \log_b \xi_k \\ &\le \frac{1}{4\ln b}\sum_{i,j=1}^n p_i p_j (\xi_i - \xi_j)^2 . \end{aligned} \tag{2.1}$$

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

PROOF. We shall use the well-known Jensen's discrete inequality for convex mappings which states that

$$f\left(\sum_{i=1}^n p_i x_i\right) \le \sum_{i=1}^n p_i f(x_i), \tag{2.2}$$

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

Now, let consider the mapping $f : [1,\infty) \to \mathbf{R}$, $f(x) = x^2/2 + \ln x$. Then

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

and

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

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

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

$$\frac{1}{2}\left(\sum_{i=1}^n p_i\xi_i\right)^2 + \ln\left(\sum_{i=1}^n p_i\xi_i\right) \le \frac{1}{2}\sum_{i=1}^n p_i\xi_i^2 + \sum_{i=1}^n p_i \ln \xi_i, \tag{2.3}$$

which is equivalent to

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

But

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

and then the above inequality becomes

$$\ln\left(\sum_{i=1}^n p_i\xi_i\right) - \sum_{i=1}^n p_i \ln\xi_i \le \frac{1}{4}\sum_{i,j=1}^n p_i p_j (\xi_i - \xi_j)^2 . \tag{2.4}$$

Now, as $\log_b x = (\ln x/\ln b)$, 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

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

and

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

and compute the difference

$$B_1 - B_2 = \frac{1}{2\ln b}\sum_{i,j=1}^{n}p_i p_j\left(\xi_i-\xi_j\right)^2\left[\frac{1}{\xi_i\xi_j}-\frac{1}{2}\right]$$
$$= \frac{1}{4\ln b}\sum_{i,j=1}^{n}\frac{p_i p_j\left(\xi_i-\xi_j\right)^2}{\xi_i\xi_j}\left(2-\xi_i\xi_j\right).$$

Consequently, if $\xi_i \in [1,\infty)$ so that $\xi_i\xi_j \le 2$, for all $i,j \in \{1,\dots,n\}$, then the bound B_2 provided by Theorem 2.1 is better than the bound B_1 provided by Theorem 1.1. If $\xi_i \in [1,\infty)$ so that $\xi_i\xi_j \ge 2$, for all $i,j \in \{1,\dots,n\}$, 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 $q_i > 0$ with $Q_n := \sum_{i=1}^{n} q_i$, the *arithmetic mean* of x_i with the weights q_i, $i \in \{1,\dots,n\}$ is

$$A_n\left(\overline{q},\overline{x}\right) := \frac{1}{Q_n}\sum_{i=1}^{n}q_i x_i \tag{A}$$

and the *geometric mean* of x_i with the weights q_i, $i \in \{1,\dots,n\}$, is

$$G_n\left(\overline{q},\overline{x}\right) := \left(\prod_{i=1}^{n}x_i^{q_i}\right)^{1/Q_n}. \tag{G}$$

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

$$A_n\left(\overline{q},\overline{x}\right) \ge G_n\left(\overline{q},\overline{x}\right) \tag{2.5}$$

with equality if and only if $x_1 = \cdots = x_n$.

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 $\overline{q}$ and $\overline{x}$, we have*

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

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

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

PROPOSITION 2.3. *Let $\overline{q}$ be as above and $\overline{x} \in \mathbf{R}^n$ with $x_i \ge 1$, $i = 1,\dots,n$. Then we have the inequality*

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

where $b > 1$. The equality holds in both inequalities simultaneously if and only if $x_1 = \cdots = x_n$.

REMARK 2.2. As in the previous remark, if $1 \le x_i x_j \le 2$ then bound (2.7) is better than (2.6). If $x_i x_j \ge 2$, 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 $p = (p_i)$, $i = \{1, \dots, n\}$,

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

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

$$0 \le \log_b n - H_b(X) \le \frac{1}{2 \ln b} \sum_{i,j=1}^{n} (p_i - p_j)^2 \tag{3.1}$$

with equality if and only if $p_i = 1/n$, for all $i \in \{1, \dots, n\}$.

The following similar result also holds.

THEOREM 3.1. *Let X be as above. Then we have*

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

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

PROOF. As $p_i \in (0, 1]$, then $\xi_i = 1/p_i \in [1, \infty)$ and we can apply Theorem 2.1 to get

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

The equality holds iff $\xi_i = \xi_j$, for all $i, j \in \{1, \dots, n\}$ which is equivalent to $p_i = p_j$, for all $i, j \in \{1, \dots, n\}$, i.e., $p_i = 1/n$, for all $i \in \{1, \dots, n\}$. ∎

The following corollary is important in applications as it provides a sufficient condition on 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, $i = 1, \dots, n$, satisfy the conditions*

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

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

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

then we have the estimation

$$0 \le \log_b n - H_b(X) \le \varepsilon. \tag{3.5}$$

PROOF. Observe that

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

Suppose that

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

Then

$$p_i^2 - (2+k)p_ip_j + p_j^2 \le 0, \qquad \text{for } 1 \le i < j \le n.$$

Denoting $t = p_i/p_j$, the above inequality is equivalent to $t^2 - (2+k)\,t + 1 \le 0$, 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 $k = (4\varepsilon \ln b / n\,(n-1))$, then by (3.3) we have

$$\begin{aligned} 0 &\le \log_b n - H_b(X) \le \frac{1}{4\ln b} \sum_{i,j=1}^{n} \frac{(p_i - p_j)^2}{p_ip_j} \\ &= \frac{1}{2\ln b} \sum_{1 \le i < j \le n} \frac{(p_i - p_j)^2}{p_ip_j} \\ &\le \frac{1}{2\ln b} \sum_{1 \le i < j \le n} k = \frac{n\,(n-1)}{4\ln b} \cdot \frac{4\varepsilon \ln b}{n\,(n-1)} = \varepsilon, \end{aligned}$$

and the corollary is proved. ∎

Now, consider the bounds

$$M_1 := \frac{1}{2\ln b} \sum_{i,j=1}^{n} (p_i - p_j)^2, \qquad \text{(given by (3.1))}$$

and

$$M_2 := \frac{1}{4\ln b} \sum_{i,j=1}^{n} \frac{(p_i - p_j)^2}{p_ip_j}, \qquad \text{(given by (3.3))}.$$

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 $\log_b n - H_b(X)$.

EXAMPLE 3.1. Consider the probability distribution

$$\begin{aligned} &p_1 = 0.3475, \quad p_2 = 0.2398, \quad p_3 = 0.1654, \\ &p_4 = 0.1142, \quad p_5 = 0.0788, \quad p_6 = 0.0544. \end{aligned}$$

In this case,

$$\overline{M_1} = 6.5119, \qquad \overline{M}_2 = 12.1166,$$

where

$$\overline{M}_1 := \frac{1}{2} \sum_{i,j=1}^{n} (p_i - p_j)^2, \qquad \overline{M}_2 := \frac{1}{4} \sum_{i,j=1}^{n} \frac{(p_i - p_j)^2}{p_ip_j}, \qquad \text{and} \qquad n = 6.$$

EXAMPLE 3.2. Consider the probability distribution

$$\begin{aligned} &p_1 = 0.2468, \quad p_2 = 0.2072, \quad p_3 = 0.1740, \\ &p_4 = 0.1461, \quad p_5 = 0.1227, \quad p_6 = 0.1031. \end{aligned}$$

In this case,

$$\overline{M}_1 = 5.2095, \qquad \overline{M}_2 = 2.3706.$$

4. BOUNDS FOR JOINT ENTROPY

Consider the joint entropy of two random variable X and Y [2, 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 [3], Dragomir and Goh have proved the following result using Theorem 1.1.

THEOREM 4.1. *With the above assumptions, we have that*

$$0 \le \log_b(rs) - H_b(X,Y) \le \frac{1}{2\ln b} \sum_{x,y}\sum_{u,v} \left(p(x,y) - p(u,v)\right)^2, \tag{4.1}$$

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*

$$\max_{(x,y),(u,v)} |p(x,y) - p(u,v)| \le \sqrt{\frac{2\varepsilon \ln b}{rs}}, \qquad \varepsilon > 0,$$

then we have the estimation

$$0 \le \log_b(r,s) - H_b(X,Y) \le \varepsilon.$$

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*

$$0 \le \log_b(rs) - H_b(X,Y) \le \frac{1}{4\ln b} \sum_{x,y}\sum_{u,v} \frac{\left(p(x,y) - p(u,v)\right)^2}{p(x,y)\,p(u,v)}, \tag{4.2}$$

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 $\xi_i = (1/p(x,y))$,

$$\begin{aligned} 0 &\le \log_b\left(\sum_{x,y} p(x,y) \cdot \frac{1}{p(x,y)}\right) - \sum_{x,y} p(x,y) \log_b \frac{1}{p(x,y)} \\ &\le \frac{1}{4\ln b} \sum_{x,y}\sum_{u,v} p(x,y)\,p(u,v) \left(\frac{1}{p(x,y)} - \frac{1}{p(u,v)}\right)^2 \\ &= \frac{1}{4\ln b} \sum_{x,y}\sum_{u,v} \frac{\left(p(x,y) - p(u,v)\right)^2}{p(x,y)\,p(u,v)}, \end{aligned}$$

which is clearly equal to the desired result. The case of equality is obvious by Theorem 2.1. ∎

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*

$$\frac{P}{p} \leq 1 + k + \sqrt{k\,(k+2)}, \tag{4.3}$$

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

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

This inequality is clearly equivalent to

$$a^2 - 2\,(1+k)\,ab + b^2 \leq 0$$

or denoting $t := a/b$, 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

$$1 + k - \sqrt{k\,(k+2)} \leq \frac{p\,(x,y)}{p\,(u,v)} \leq 1 + k + \sqrt{k\,(k+2)}, \tag{4.4}$$

for all (x,y) and (u,v) and $k := (2\varepsilon \ln b/(rs)^2)$. Then by (4.2), we have

$$\begin{aligned} 0 \leq \log_b (rs) - H_b\,(X,Y) &\leq \frac{1}{4\ln b} \sum_{x,y} \sum_{u,v} \frac{\left(p\,(x,y) - p\,(u,v)\right)^2}{p\,(x,y)\,p\,(u,v)} \\ &\leq \frac{1}{2\ln b} \cdot (rs)^2\,k = \frac{(rs)^2}{2\ln b} \cdot \frac{2\varepsilon \ln b}{(rs)^2} = \varepsilon. \end{aligned}$$

Now, let observe that 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 $p/P \geq 1 + k - \sqrt{k\,(k+2)}$ 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. ∎

REFERENCES

1. S.S. Dragomir and C.J. Goh, A counterpart of Jensen's discrete inequality for differentiable convex mapping and application in information theory, *Mathl. Comput. Modelling* **24** (2), 1–4, (1996).
2. S. Roman, *Coding and Information Theory*, Springer-Verlag, New York.
3. S.S. Dragomir and C.J. Goh, Further counterparts of some inequalities in information theory, (submitted).