New Inequalities for Logarithmic Map and Their Application for Entropy and Mutual Information

RGMIA Research Report Collection, Vol. 2, No. 1, 1999
http://sci.vu.edu.au/~rgmia

NEW INEQUALITIES FOR LOGARITHMIC MAP AND THEIR APPLICATION FOR ENTROPY AND MUTUAL INFORMATION

S.S. Dragomir, C.E.M. Pearce and J. Pečarić

Abstract. In this paper we discuss new inequalities for logarithmic mapping and apply them in Information Theory obtaining new bounds for the entropy mapping and mutual information.

1 Introduction

Let $x = (x_1, \ldots, x_n)$, $p = (p_1, \ldots, p_n)$ be n-tuple of real numbers so that $x_i > 0$, $(i = 1, \ldots, n)$ and $p_i \geq 0$ with $P_n := \sum_{i=1}^{n} p_i > 0$. The following inequality is well known in the literature as *arithmetic mean - geometric mean inequality*

$$\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i \geq \left(\prod_{i=1}^{n} x_i^{p_i}\right)^{\frac{1}{P_n}} \tag{1.1}$$

with equality for $p_i > 0$, $(i = 1, \ldots, n)$ if and only if $x_i = x_j$ for all $i, j \in \{1, \ldots, n\}$.

This elementary inequality is closely related to entropy mapping and mutual information which are important in Information Theory and Coding.

Suppose that X is a discrete random variable whose range $R := \{x_1, \ldots, x_n\}$ is finite.

Let $p_i = P\{X = x_i\}$, $(i = 1, \ldots, n)$ and assume that $p_i > 0$ for all $i \in \{1, \ldots, n\}$.

Define the *b-entropy* (we use notation entropy in the rest of the paper) of X by

$$H_b(X) := \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$$

where, for simplicity, log denotes the logarithm of base $b > 1$.

The following theorem is well known in the literature and concerns the maximum possible value for $H_b(X)$ in terms of the size of R [3, p. 17].

Theorem 1.1. *Let X assume values in $R = \{x_1, ..., x_n\}$. Then*

$$0 \leq H_b(X) \leq \log n. \tag{1.2}$$

Furthermore, $H_b(X) = 0$ if and only if $p_i = 1$ for some i and $H_b(X) = \log n$ if and only if $p_i = \frac{1}{n}$ for all $i \in \{1, \ldots, n\}$.

Note that if in (1.1) we take $\log(\cdot)$ we get

$$\frac{1}{P_n}\sum_{i=1}^{n} p_i \log x_i \leq \log\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \tag{1.3}$$

(which is Jensen's inequality for the concave map $\log(\cdot)$) and if in (1.3) we choose $x_i = \frac{1}{p_i}$, we deduce the second inequality in (1.2).

In paper [1], S.S. Dragomir and C.J. Goh proved the following counterpart of (1.3):

Date. January 22, 1999
1991 Mathematics Subject Classification. 26D15, 26D20, 26D99, 94Axx.
Key words and phrases. Inequalities for Logarithmic Map, Entropy, Mutual Information.

Theorem 1.2. *Let $x_i \in (0,\infty)$ and $p_i, i \in \{1,\dots,n\}$, are as above. Then*

$$
\begin{aligned}
(1.4)\qquad 0 &\le \log\left(\frac{1}{P_n}\sum_{i=1}^n p_i x_i\right) - \frac{1}{P_n}\sum_{i=1}^n p_i \log x_i \\
&\le \frac{1}{\ln b}\frac{1}{P_n^2}\left(\sum_{i=1}^n \frac{p_i}{x_i}\sum_{i=1}^n p_i x_i - P_n^2\right) \\
&= \frac{1}{2\ln b}\frac{1}{P_n^2}\sum_{i,j=1}^n \frac{p_i p_j}{x_i x_j}(x_i - x_j)^2.
\end{aligned}
$$

In papers [1] and [2] they used this inequality to establish various counterpart inequalities involving entropy mapping, joint and conditional entropy, mutual information and conditional mutual information.

The main aim of this paper is to establish some other inequalities for $\log(\cdot)$ map similar to (1.4) and to apply them for some relevant mappings in Information Theory.

2 Preliminary results

The following theorem is important.

Theorem 2.1. *Let $x_i > 0$ and $q_i \ge 0$ $(i = 1,\dots,n)$ so that $\sum_{i=1}^n q_i = 1$. Then we have the inequality:*

$$
\begin{aligned}
(2.1)\qquad 0 &\le \sum_{i=1}^n q_i x_i \log x_i - \sum_{i=1}^n q_i x_i \log\left(\sum_{i=1}^n q_i x_i\right) \\
&\le \frac{1}{2}\sum_{i,j=1}^n q_i q_j (x_i - x_j)\log\frac{x_i}{x_j} \\
&\left(= \sum_{i=1}^n q_i x_i \log x_i - \sum_{i=1}^n q_i x_i \sum_{i=1}^n q_i \log x_i\right).
\end{aligned}
$$

The equality holds in (2.1) simultaneously if and only if $x_i = x_j$ for all $i,j \in \{1,\dots,n\}$.

Proof. We use Jensen's discrete inequality:

$$
(2.2)\qquad f\left(\sum_{i=1}^n q_i x_i\right) \le \sum_{i=1}^n q_i f(x_i)
$$

where $f : I \subset R \to R$ is a convex mapping, $x_i \in I$, $q_i \ge 0$ $(i = 1,\dots,n)$ with $\sum_{i=1}^n q_i = 1$ and $n \ge 2$.

Consider now the map $f_b : (0,\infty) \to R$, $f_b(x) := x\log x$.

From

$$
f_b'(x) = \log x + \frac{1}{\ln b}, \qquad x > 0
$$

and

$$
f_b''(x) = \frac{1}{x\ln b} > 0, \qquad x > 0
$$

we can see that f_b is convex.

Applying Jensen's inequality (2.2) for f_b we get the first inequality in (2.1).

The well known identity

$$
\frac{1}{2}\sum_{i,j=1}^n q_i q_j (x_i - x_j)(y_i - y_j) = \sum_{i=1}^n q_i \sum_{i=1}^n q_i x_i y_i - \sum_{i=1}^n q_i x_i \sum_{i=1}^n q_i y_i
$$

gives equation in brackets in (2.1).

The second inequality in (2.1) is equivalent with:

$$\begin{aligned}&\sum_{i=1}^{n} q_i x_i \log x_i - \sum_{i=1}^{n} q_i x_i \log\left(\sum_{i=1}^{n} q_i x_i\right)\\ \leq\ &\sum_{i=1}^{n} q_i x_i \log x_i - \sum_{i=1}^{n} q_i x_i \sum_{i=1}^{n} q_i \log x_i\end{aligned}$$

i.e.

$$\sum_{i=1}^{n} q_i x_i \sum_{i=1}^{n} q_i \log x_i \leq \sum_{i=1}^{n} q_i x_i \log\left(\sum_{j=1}^{n} q_i x_i\right).$$

This follows by Jensen's discrete inequality (2.2) applied for the convex map $f_b(x) := -\log x$ and considering that $x_i > 0$, $q_i \geq 0$ $(i = 1, \ldots, n)$ and $\sum_{i=1}^{n} q_i = 1$. The case of equality is obvious. ∎

Corollary 2.2. *Let $x_i > 0 (i = 1, ..., n)$. Then we have the inequality*

$$\begin{aligned}0 &\leq \sum_{i=1}^{n} x_i \log x_i - \sum_{i=1}^{n} \log\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)\\ &\leq \frac{1}{2n} \sum_{i,j=1}^{n} (x_i - x_j) \log \frac{x_i}{x_j}\\ &= \frac{1}{n} \sum_{1 \leq i < j \leq n} (x_i - x_j) \log \frac{x_i}{x_j}\\ &= \sum_{i=1}^{n} x_i \log x_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} \log x_i.\end{aligned}$$

The following theorem is also useful.

Theorem 2.3. *Let $x_m := \min_{i=1,\ldots,n} \{x_i\}$, $x_M := \max_{i=1,\ldots,n} \{x_i\}$ with $x_m > 0$ and $q_i \geq 0$ with $\sum_{i=1}^{n} q_i = 1$. Then we have the inequality:*

$$\begin{aligned}0 &\leq \sum_{i=1}^{n} q_i x_i \log x_i - \sum_{i=1}^{n} q_i x_i \log\left(\sum_{i=1}^{n} q_i x_i\right)\\ &\leq \frac{1}{4}(x_M - x_m) \log \frac{x_M}{x_m}\end{aligned} \tag{2.3}$$

Proof. As $x_m \leq x_i \leq x_M$ we have $\ln x_m \leq \ln x_i \leq \ln x_M$ for all $i \in \{1, \ldots, n\}$. Now (2.3) is a simple consequence of the well-known Grüss inequality

$$\left|\frac{\sum_{i=1}^{n} q_i a_i b_i}{\sum_{i=1}^{n} q_i} - \frac{\sum_{i=1}^{n} q_i a_i}{\sum_{i=1}^{n} q_i} \frac{\sum_{i=1}^{n} q_i b_i}{\sum_{i=1}^{n} q_i}\right| \leq \frac{1}{4}(A - a)(B - b)$$

where $q_i > 0$, $a \leq a_i \leq A$, $b \leq b_i \leq B$, $i = 1, ..., n$. By Chebyshev inequality for similar ordered n-tuples the expression in absolute value is positive what is our case with $a_i = x_i$ and $b_i = \log x_i$. ∎

Using the Corollary 2.2, we can prove the following result:

Corollary 2.4. *With the above assumption for x_i, $i = 1, \ldots, n$, we have the inequality:*

$$0 \leq \sum_{i=1}^{n} x_i \log x_i - \sum_{i=1}^{n} x_i \log\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \leq \frac{1}{n}\left[\frac{n^2}{4}\right](x_M - x_m) \log \frac{x_M}{x_m}$$

where $[x]$ is the integer part of x.

Proof. We obtain the above result by using Corollary 2.2, and the following result of M. Biernacky, H. Pidek and C. Ryll-Nardzewski (see for example [4, p. 205]):

$$\left|\sum_{i=1}^{n} a_i b_i - \frac{1}{n}\sum_{i=1}^{n} a_i \sum_{i=1}^{n} b_i\right| \leq \left[\frac{n}{2}\right]\left(1-\frac{1}{n}\left[\frac{n}{2}\right]\right)(A-a)(B-b)$$

where $a \leq a_i \leq A$ and $b \leq b_i \leq B$ $(i = 1, ..., n)$ and taking into account that

$$\left[\frac{n}{2}\right]\left(1-\frac{1}{n}\left[\frac{n}{2}\right]\right) = \frac{1}{n}\left[\frac{n^2}{4}\right] \tag{2.4}$$

for all $n \in N, n \geq 1$. ∎

The following theorem also holds:

Theorem 2.5. *With the assumptions of Theorem 2.1, we have the inequality:*

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} q_i x_i \log x_i - \sum_{i=1}^{n} q_i x_i \log\left(\sum_{i=1}^{n} q_i x_i\right) \\ &\leq \frac{1}{2\ln b}\sum_{i,j=1}^{n} q_i q_j \frac{(x_i - x_j)^2}{\sqrt{x_i x_j}}. \end{aligned}$$

Proof. Consider $A(a,b) := \frac{a+b}{2}$, $G(a,b) := \sqrt{ab}$ and $L(a,b) := \frac{a-b}{\ln a - \ln b}$, $(a \neq b)$ the *arithmetic mean, geometric mean* and *logarithmic mean*, respectively. It is well known that

$$G(a,b) \leq L(a,b) \leq A(a,b)$$

thus

$$\frac{\ln x_i - \ln x_i}{x_i - x_j} \leq \frac{1}{\sqrt{x_i x_j}}$$

which gives us

$$(x_i - x_j)\ln\frac{x_i}{x_j} \leq \frac{(x_i - x_j)^2}{\sqrt{x_i x_j}}$$

for all $i, j \in \{1, \ldots, n\}$.

Consequently

$$\sum_{i,j=1}^{n} q_i q_j (x_i - x_j)\ln\frac{x_i}{x_j} \leq \sum_{i,j=1}^{n} q_i q_j \frac{(x_i - x_j)^2}{\sqrt{x_i x_j}}$$

and the theorem is thus proved. ∎

Corollary 2.6. *With the above assumptions, we get:*

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} x_i \log x_i - \sum_{i=1}^{n} x_i \log\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \\ &\leq \frac{1}{n\ln b}\sum_{1\leq i<j\leq n} \frac{(x_i - x_j)^2}{\sqrt{x_i x_j}}. \end{aligned} \tag{2.5}$$

3 Some inequalities for the entropy mapping

Suppose that X is a discrete random variable whose range $R = \{x_1, \dots, x_n\}$ is finite. Let $p_i = P\{X = x_i\}$, $i = 1, \dots, n$ and assume that $p_i > 0$ for all $i \in \{1, \dots, n.\}$ Define the entropy of X by

$$H_b(X) := \sum_{i=1}^{n} p_i \log_b \frac{1}{p_i}. \tag{3.1}$$

The following theorem is well known in the literature (see the Introduction)

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

$$0 \le H_b(X) \le \log n. \tag{3.2}$$

Furthermore, $H_b(X) = 0$ if and only if $p_i = 1$ for some i and $H_b(X) = \log n$ if and only if $p_i = \frac{1}{n}$ for all $i \in \{1, ..., n\}$.

In the recent paper [1], S.S. Dragomir and C.J. Goh proved the following counterpart of (3.2).

Theorem 3.2. *Let X be defined as above, then*

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

Furthermore, equality holds simultaneously in all the above inequalities if and only if $p_i = \frac{1}{n}$ for all $i \in \{1, ..., n\}$.

Using some results obtained above for the $\log(\cdot)$ mapping we can point out some different results:

Theorem 3.3. *Let X be defined as above, then*

$$\begin{aligned} 0 &\le \log n - H_b(X) \le \frac{1}{n}\sum_{1\le i<j\le n} (p_i - p_j)\log \frac{p_i}{p_j} \\ &\le \frac{1}{n\ln b}\sum_{1\le i<j\le n} \frac{(p_i - p_j)^2}{\sqrt{p_i p_j}} \end{aligned}$$

with equality in all inequalities if and only if $p_i = \frac{1}{n}$, $i \in \{1, ..., n\}$.

Proof. Follows by Corollary 2.2 and Corollary 2.6 by choosing $x_i = p_i$. ∎

The following result also holds.

Theorem 3.4. *Let X be as above. If $p := \min\{p_i | i = 1, \dots, n\}$ and $P := \max\{p_i | i = 1, \dots, n\}$ then:*

$$\begin{aligned} 0 &\le \log n - H_b(X) \le \frac{1}{n}\left[\frac{n^2}{4}\right]\frac{1}{\ln b}(P - p)\log \frac{P}{p} \\ &\le \frac{1}{n}\left[\frac{n^2}{4}\right]\frac{1}{\ln b}\frac{(P - p)^2}{\sqrt{pP}}. \end{aligned}$$

Proof. The proof follows by Corollary 2.4 for $x_i = p_i$. We omit the details. ∎

We can give now different results for the entropy mapping.

Theorem 3.5. *Let X be as above and let $Q = \sum_{i=1}^{n} p_i^2$. Then we have the inequality:*

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

Proof. In the inequality (2.1) let choose

$$q_i = p_i^2/Q.$$

Then $\sum_{i=1}^{n} q_i = 1$ and if we put $x_i = \frac{1}{p_i}$, then

$$
\begin{aligned}
0 &\le \frac{1}{Q}\sum_{i=1}^{n} p_i \log\frac{1}{p_i} - \frac{\sum_{i=1}^{n} p_i}{Q}\log\left(\frac{\sum_{i=1}^{n} p_i}{Q}\right) \\
&\le \frac{1}{2Q}\sum_{i,j=1}^{n} p_i^2 p_j^2 \left(\frac{1}{p_i} - \frac{1}{p_j}\right)\log\frac{p_j}{p_i} \\
&= \frac{1}{2Q}\sum_{i,j=1}^{n} p_i p_j \left(p_i - p_j\right)\log\frac{p_i}{p_j}
\end{aligned}
$$

and the second inequality in (3.3) is obtained.

The last inequality follows by the elementary inequality to $L^{-1} \le G^{-1}$, and we shall omit the details. ∎

Another estimation in terms of P and p (defined above) is the following:

Theorem 3.6. *Let X be as above and $P := \max\{p_i | i = 1, \dots, n\}$, $p = \min\{p_i | i = 1, \dots, n\}$ and $Q = \sum_{i=1}^{n} p_i^2$. Then we have the inequality:*

$$0 \le H_b(X) - \log\frac{1}{Q} \le \frac{Q}{4pP}(P - p)\log\frac{P}{p} \le \frac{(P-p)^2 Q}{4pP\sqrt{pP}\ln b} \tag{3.4}$$

Proof. If we choose $q_i = \frac{p_i^2}{Q}$, $x_i = \frac{1}{p_i}$ in Theorem 4, we get:

$$
\begin{aligned}
0 &\le \frac{1}{Q}\sum_{i=1}^{n} p_i \log\frac{1}{p_i} - \frac{\sum_{i=1}^{n} p_i}{Q}\log\frac{\sum_{i=1}^{n} p_i}{Q} \\
&\le \frac{\left(\frac{1}{p} - \frac{1}{P}\right)\log\frac{P}{p}}{4}
\end{aligned}
$$

which is equivalent with

$$0 \le H_b(X) - \log\frac{1}{Q} \le Q \cdot \frac{(P-p)}{4pP}\log\frac{P}{p}.$$

Now, as $L^{-1} \le G^{-1}$, we get the last part of (3.4) and the theorem is proved. ∎

4 Some new bounds for conditional entropy

For a pair of random variables X and Y, with respective ranges $\{x_1, x_2, ..., x_n\}$ and $\{y_1, y_2, ..., y_n\}$, the conditional entropy of X and Y is defined by [3, p. 22]:

$$H_b(X|Y) := \sum_{i,j} p(x_i, y_j)\log\left(\frac{1}{p(x_i|y_j)}\right) = \sum_{i,j} p(x_i, y_j)\log\left(\frac{p(y_j)}{p(x_i, y_j)}\right)$$

where

$$p(x_i, y_j) := P\{X = x_i, Y = y_j\}$$

and

$$p(x_i|y_j) := P\{X = x_i|Y = y_j\} = \frac{p(x_i, y_j)}{p(y_j)}.$$

Without loss of generality we need to define these quantities only for those (i, j) for which $p(x_i, y_j) > 0$. There will be n $(\leq rs)$ such pairs.

One can interpret the conditional entropy as the amount of uncertainty remaining about X after Y has been observed.

The following theorem holds.

Theorem 4.1. *For $1 \leq j \leq s$ define $V_j := \{i : p(x_i, y_j) > 0\}$ and $U := \{(i, j) : i \in V_j\}$ and $r' = \sum_{j=1}^{s} p(y_j)|V_j|$. Then we have the inequality:*

(4.1)

$$\begin{aligned} 0 &\leq \log r' - H_b(X|Y) \\ &\leq \frac{1}{2r'} \sum_{(i,j)\in U} \sum_{(u,v)\in U} \frac{p(y_j)p(v_y)}{p(x_i|y_j)p(x_u|y_v)} \left(p(x_i|y_j) - p(x_u|y_v)\right) \log \frac{p(x_i|y_j)}{p(x_u|y_v)} \\ &\leq \frac{1}{2r' \ln b} \sum_{(i,j)\in U} \sum_{(u,v)\in U} \frac{p(y_j)p(y_v) \left(p(x_i|y_j) - p(x_u|y_v)\right)^2}{p(x_i|y_j)p(x_u|y_v)\sqrt{p(x_i|y_j)p(x_u|y_v)}} . \end{aligned}$$

Proof. We shall apply the Theorem 2.1 for

$$q_i = \frac{p(y_j)}{r'}, x_i = \frac{p(x_i|y_j)}{p(y_j)}.$$

Note that

$$\sum_{(i,j)\in U} \frac{p(y_j)}{r'} = \frac{1}{r'} \sum_{j} p(y_j) |V_j| = 1.$$

By Theorem 2.1 we have the inequality:

$$\begin{aligned} 0 &\leq \sum_{(i,j)\in U} \frac{p(y_j)}{r'} \frac{p(x_i, y_j)}{p(y_j)} \log \frac{p(x_i, y_j)}{p(y_j)} - \\ &\quad - \sum_{(i,j)\in U} \frac{p(y_j)}{r'} \frac{p(x_i, y_j)}{p(y_j)} \log \left(\sum_{(i,j)\in U} \frac{p(y_j)}{r} \frac{p(x_i, y_j)}{p(y_j)} \right) \\ &\leq \frac{1}{2} \sum_{(i,j)\in U} \sum_{(u,v)\in U} \frac{p(y_j)}{r'} \frac{p(y_v)}{r'} \left(\frac{p(x_i, y_v)}{p(y_j)} - \frac{p(x_u, y_v)}{p(y_v)} \right) \log \frac{p(x_i, y_j)p(y_v)}{p(x_u, y_v)p(y_j)} \end{aligned}$$

which is equivalent with

$$\begin{aligned} 0 &\leq \frac{1}{r'} \sum_{(i,j)\in U} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(y_j)} - \frac{1}{r} \sum_{(i,j)\in U} p(x_i, y_j) \log \left(\frac{1}{r} \sum_{(i,j)\in U} p(x_i, y_j) \right) \\ &\leq \frac{1}{2r'^2} \sum_{(i,j)\in U} \sum_{(u,v)\in U} p(y_j)p(y_v) \left(\frac{1}{p(x_i, y_j)} - \frac{1}{p(x_u, y_v)} \right) \log \frac{p(x_u, y_v)}{p(x_i, y_j)} \end{aligned}$$

i.e.,

$$\begin{aligned} 0 &\leq \frac{1}{r'} \left(\log r' - H_b(X|Y) \right) \\ &\leq \frac{1}{2r'^2} \sum_{(i,j)\in U} \sum_{(u,v)\in U} \frac{p(y_j)p(y_v)}{p(x_i, y_j)p(x_u, y_v)} \left(p(x_i|y_j) - p(x_u, y_v)\right) \log \frac{p(x_i, y_j)}{p(x_u, y_v)} \end{aligned}$$

and the first inequality in (4.1) is proved. The last inequality follows from

$$\frac{\ln p(x_i|y_j) - \ln p(x_u|y_v)}{p(x_i|y_j) - p(x_u|y_v)} \le \frac{1}{\sqrt{p(x_i|y_j)p(x_u|y_v)}}$$

for $(i,j) \neq (u,v)$. ∎

In a similar way, by using Theorem 2.3, we get the inequality:

Theorem 4.2. *Let assumptions of Theorem 4.1 be satisfied. Suppose that* $k := \min\{p(x_i|y_j)/p(y_j)\}$, $K := \max\{p(x_i|y_j)/p(y_j)\}$.

Then we have the inequality:

$$0 \le \log r' - H_b(X|Y) \le \frac{r'(K-k)}{4kK} \log \frac{K}{k} \le \frac{r'(K-k)^2}{4kK\sqrt{kK}\ln b}.$$

Remark 4.1. *For some related results see [1] and [2]*

5 Results for mutual information

Consider the mutual information between two random variables X and Y defined by:

$$I(X;Y) := \sum_{i,j} p(x_i, y_j) \log \left[\frac{p(x_i, y_j)}{p(x_i)p(y_j)} \right].$$

The following theorem concerning the mutual information is known in the literature *[3, p. 23]*.

Theorem 5.1. *For any pair of discrete random variable X and Y, $I(X;Y) \ge 0$. Moreover, $I(X;Y) \ge 0$ if and only if X and Y are independent.*

Now we shall prove the following result.

Theorem 5.2. *Let* $V := \{(i,j) : p(x,y) > 0\}$ *and* $S = \sum_{(i,j)\in V} p(x_i)p(y_j)$

(5.1)

$$\begin{aligned} 0 &\le I(X;Y) + \log S \\ &\le \frac{1}{2S} \sum_{(i,j)\in V} \sum_{(u,v)\in V} p(x_i)p(y_j)p(x_u)p(y_v) \left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)} - \frac{p(x_u,y_v)}{p(x_u)p(y_v)} \right) \\ &\quad \times \log \frac{p(x_i,y_j)p(x_u)p(y_v)}{p(x_i)p(y_j)p(x_u,y_v)} \\ &= \frac{1}{2S\ln b} \sum_{(i,j)\in V} \sum_{(u,v)\in V} p(x_i)p(y_j)p(x_u)p(y_v) \sqrt{\frac{p(x_i)p(y_j)p(x_u)p(y_v)}{p(x_i,y_j)p(x_u,y_v)}} \\ &\quad \times \left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)} - \frac{p(x_u,y_v)}{p(x_u)p(y_v)} \right)^2. \end{aligned}$$

Proof. Consider $q_i = \frac{p(x_i)p(y_j)}{S}$. Then, obviously, $\sum_i q_i = \sum_{(i,j)\in V} \frac{p(x_i)p(y_j)}{S} = 1$. If we choose now $x_i = \frac{p(x_i,y_j)}{p(x_i)p(y_j)}$ and $q_i = \frac{p(x_i)p(y_j)}{S}$ in Theorem 2.1 we deduce:

$$\begin{aligned} 0 &\le \sum_{(i,j)\in V} \frac{p(x_i)p(y_j)}{S} \frac{p(x_i,y_j)}{p(x_i)p(y_j)} \log \left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)} \right) \\ &\quad - \sum_{(i,j)\in V} p(x_i)p(y_j) \frac{p(x_i,y_j)}{p(x_i)p(y_j)} \log \left(\sum_{(i,j)\in V} \frac{p(x_i,y_j)}{p(x_i)p(y_j)} \frac{p(x_i)p(y_j)}{S} \right) \\ &\le \frac{1}{2} \sum_{(i,j)\in V} \sum_{(u,v)\in V} \frac{p(x_i)p(y_j)p(u)p(v)}{S} \left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)} - \frac{p(x_u,y_v)}{p(x_u)p(y_v)} \right) \\ &\quad \times \log \frac{p(x_i,y_j)p(x_u)p(y_v)}{p(x_i)p(y_j)p(x_u,y_v)} \end{aligned}$$

which is equivalent with

$$\begin{aligned} 0 &\leq I(X;Y)+\log S \\ &\leq \frac{1}{2S}\sum_{(i,j)\in V}\sum_{(u,v)\in V} p(x_i)p(y_j)p(x_u)p(y_v)\left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)}-\frac{p(x_u,y_v)}{p(x_u)p(y_v)}\right) \\ &\quad \times \log\frac{p(x_i,y_j)p(x_u)p(y_v)}{p(x_i)p(y_j)p(x_u,y_v)} \end{aligned}$$

and the second inequality in (5.1) is proved.

Now, as $L^{-1}\leq G^{-1}$, we get that

$$\begin{aligned} &\left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)}-\frac{p(x_u,y_v)}{p(x_u)p(y_v)}\right)\left(\ln\frac{p(x_i,y_j)}{p(x_i)p(y_j)}-\ln\frac{p(x_u,y_v)}{p(x_u)p(y_v)}\right) \\ &\qquad \leq \sqrt{\frac{p(x_i)p(y_j)p(x_u)p(y_v)}{p(x_i,y_j)p(x_u,y_v)}}\left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)}-\frac{p(x_u,y_v)}{p(x_u)p(y_v)}\right)^2 \end{aligned}$$

and the theorem is proved. ∎

Using Theorem 2.3, we get another result:

Theorem 5.3. *With the above assumptions and if* $t=\min_{(i,j)\in V}\left\{\frac{p(x_i,y_j)}{p(x_i)p(y_j)}\right\}$ *and* $T=\max_{(i,j)\in V}\left\{\frac{p(x_i,y_j)}{p(x_i)p(y_j)}\right\}$, *we have the following:*

$$0\leq I(X;Y)+\log S\leq\frac{1}{4S\ln b}(T-t)\log\frac{T}{t}\leq\frac{(T-t)^2}{4S\sqrt{tT}\ln b}.$$

Remark 5.1. *For some related results see [1] and [2] .*

References

[1] S.S. Dragomir and C.J. Goh, A counterpart of Jensen's discrete inequality for differentiable convex mappings and applications in Information Theory, *Math. Comput. Modelling,* **24** (2) (1996), 1-11.

[2] M. Matić, C.E.M. Pearce and J.Pečarić, Refinements of some bounds in Information Theory, J. Austral. Math. Soc. Ser. B, to appear.

[3] R.J. McEliece, *The Theory of Information and Coding,* Addison-Wesley Publishing Company, 1977.

[4] D.S.Mitrinović, *Analytic Inequalities*, Heidelberg, New York, 1980.

(S.S. Dragomir) School of Communications and Informatics, Victoria University of Technology, Melbourne, Australia.

(C.E.M. Pearce, J. Pečarić) Applied Mathematics Department, University of Adelaide

E-mail address: S.S. Dragomir `sever@matilda.vut.edu.au`

C.E.M. Pearce `cpearce@spam.adelaide.edu.au`

J. Pečarić `jpecaric@spam.adelaide.edu.au`