A NEW UPPER BOUND FOR THE KULLBACK-LEIBLER DISTANCE AND APPLICATIONS

S.S. DRAGOMIR

Abstract. In this paper we obtain another upper bound for the Kullback-Leibler distance than the bound obtained in [2] by N.M. Dragomir and S.S. Dragomir and point out that it can be better than the first one in certain cases.

1. Introduction

The *relative entropy* is a measure of the distance between two distributions. In statistics, it arises as an expected logarithm of the likelihood ratio. The relative entropy $D(p\|q)$ is a measure of the inefficiency of assuming that the distribution is q when the true distribution is p. For example, if we knew the true distribution of the random variable, then we could construct a code with average description length $H(p)$. If, instead, we used the code for a distribution q, we would need $H(p)+D(p\|q)$ bits on the average to describe the random variable [1, p. 18].

Definition 1.1. *The relative entropy or Kullback-Leibler distance between two probability mass functions $p(x)$ and $q(x)$ is defined by*

$$\begin{aligned} D_b(p\|q) &:= \sum_{x\in\mathcal{X}} p(x)\log_b\left(\frac{p(x)}{q(x)}\right) \\ &= E_p \log_b\left(\frac{p(X)}{q(X)}\right). \end{aligned} \tag{1.1}$$

In the above definition, we use the convention (based on continuity arguments) that $0\log_b\left(\frac{0}{q}\right)=0$ and $p\log_b\left(\frac{p}{0}\right)=\infty$.

It is well-known that relative entropy is always non-negative and is zero if and only if $p=q$. However, it is not a true distance between distributions since it is not symmetric and does not satisfy the triangle inequality.

The following theorem is of fundamental importance [1, p. 26].

Theorem 1.1. *(Information Inequality) Let $p(x),\ q(x)\in\mathcal{X}$, be two probability mass functions. Then*

$$D_b(p\|q)\geq 0 \tag{1.2}$$

with equality if and only if

$$p(x)=q(x)\quad \text{for all } x\in\mathcal{X}. \tag{1.3}$$

Date: February, 1999.

1991 *Mathematics Subject Classification.* Primary 94 X xx; Secondary 26D15.

Key words and phrases. Kullback-Leibler Distance, Shannon Entropy, Mutual Information, Inequalities for Logarithms.

Actually, the inequality (1.2) can be improved as follows (see , [1, p. 300]):

Theorem 1.2. *Let* p, q *be as above. Then*

$$D_b(p\|q) \geq \frac{1}{2\ln b}\|p-q\|_1^2 \tag{1.4}$$

where $\|p-q\|_1 = \sum_{x\in\mathcal{X}} |p(x)-q(x)|$ *is the usual* 1*-norm of* $p-q$*. The equality holds iff* $p=q$*.*

We remark that the argument of (1.4) is not based on the convexity of the map $-\log(.)$.

To evaluate the relative entropy $D_b(p\|q)$ it would be interesting to establish some upper bounds.

Before we do this, let us recall some other important concepts in Information Theory.

We introduce *mutual information*, which is a measure of the amount of information that one random variable contains about another random variable. It is the reduction in the uncertainty of one random variable due to the knowledge of the other [1, p. 18].

Definition 1.2. *Consider two random variables* X *and* Y *with a joint probability mass function* $p(x,y)$ *and marginal probability mass function* $p(x)$ *and* $q(y)$*. The mutual information is the relative entropy between the joint distribution and the product distribution, i.e.*

$$\begin{aligned} I_b(X;Y) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log_b\left(\frac{p(x,y)}{p(x)q(y)}\right) \\ &= D_b\left(p(x,y)\ \|\ p(x)q(y)\right) \\ &= E_{p(x,y)}\log_b\left(\frac{p(X,Y)}{p(X)q(Y)}\right). \end{aligned} \tag{1.5}$$

The following corollary of Theorem 1.1 holds [1, p. 27].

Corollary 1.3. *(Non-negativity of mutual information): For any two random variables,* X*,* Y *we have*

$$I_b(X;Y) \geq 0 \tag{1.6}$$

with equality if and only if X *and* Y *are independent.*

An improvement of this result via Theorem 1.2 is as follows

Corollary 1.4. *For any two random variables,* X*,* Y *we have*

$$I_b(X;Y) \geq \frac{1}{2\ln b}\left[\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} |p(x,y)-p(x)q(y)|\right]^2 \geq 0 \tag{1.7}$$

with equality if and only if X *and* Y *are independent.*

Now, let $u(x)=\frac{1}{|\mathcal{X}|}$ be the uniform probability mass function on $\mathcal{X}$ and let $p(x)$ be the probability mass function for X.

It is well-known that [1, p. 27]

$$\begin{aligned} D_b(p\|u) &= \sum_{x\in\mathcal{X}} p(x)\log_b \frac{p(x)}{u(x)} \\ &= \log_b|\mathcal{X}| - H_b(X). \end{aligned} \tag{1.8}$$

The following corollary of Theorem 1.1 is important [1, p. 27].

Corollary 1.5. *Let X be a random variable and $|\mathcal{X}|$ denotes the number of elements in the range of X. Then*

$$H_b(X) \le \log_b|\mathcal{X}| \tag{1.9}$$

with equality if and only if X has a uniform distribution over $\mathcal{X}$.

Using Theorem 1.2 we also can state

Corollary 1.6. *Let X be as above. Then*

$$\log_b|\mathcal{X}| - H_b(X) \ge \frac{1}{2\ln b}\left[\sum_{x\in\mathcal{X}}\left|p(x)-\frac{1}{|\mathcal{X}|}\right|\right]^2 \ge 0. \tag{1.10}$$

The equality holds iff p is uniformly distributed on $\mathcal{X}$.

In the recent paper [2], the authors proved between other the following upper bound for the relative entropy and employed it in Coding Theory in connection to Noiseless Coding Theorem:

Theorem 1.7. *Under the above assumptions for $p(x)$ and $q(x)$ we have the inequality*

$$\frac{1}{\ln b}\left[\sum_{x\in\mathcal{X}}\frac{p^2(x)}{q(x)} - 1\right] \ge D_b(p\|q) \tag{1.11}$$

with equality if and only if $p(x) = q(x)$ for all $x \in \mathcal{X}$.

The following upper bound for the mutual information holds

Corollary 1.8. *For any two random variables, X, Y we have*

$$\frac{1}{\ln b}\left[\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\frac{p^2(x,y)}{p(x)q(y)} - 1\right] \ge I_b(X;Y)$$

with equality iff X and Y are independent.

Finally, we note that the following upper bound for the difference $\log_b|\mathcal{X}| - H_b(X)$ is valid

Corollary 1.9. *We have*

$$\frac{1}{\ln b}\left[|\mathcal{X}|\sum_{x\in\mathcal{X}}p^2(x) - 1\right] \ge \log_b|\mathcal{X}| - H_b(X)$$

with equality if and only if p is uniformly distributed on $\mathcal{X}$.

2. Inequalities for Relative Entropy

Let X and Y be two discrete random variables having the same range and the probability distributions $p(x), q(x) > 0$ for all $x \in \mathcal{X}$.

The following upper bound for the relative entropy also holds.

Theorem 2.1. *Under the above assumptions, we have*

$$0 \leq D_b(p\|q) \leq \frac{1}{\ln b} \sum_{x \in \mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} \left| p(x) - q(x) \right|, b > 1 \tag{2.1}$$

with equality iff $p(x) = q(x)$ *for all* $x \in \mathcal{X}$ *.*

Proof. We recall the following well known inequality between the geometric mean $G(a,b) := \sqrt{ab}\,(a, b \geq 0)$ and the logarithmic mean $L(a,b)$, where

$$L(a,b) := \begin{cases} a \text{ if } b = a \\ \frac{b-a}{\ln b - \ln a} \text{ if } b \neq a \end{cases} (a, b > 0);$$

i.e., we recall it,

$$G(a,b) \leq L(a,b) \text{ for all } a, b > 0. \tag{2.2}$$

Note that, this inequality is equivalent to

$$|\ln b - \ln a| \leq \frac{|b-a|}{\sqrt{ab}}; a, b > 0. \tag{2.3}$$

The equality holds in (2.3) iff $a = b$.

Now, choose in (2.3), $b = p(x), a = q(x)\ (x \in \mathcal{X})$ to get

$$|\ln p(x) - \ln q(x)| \leq \frac{|p(x) - q(x)|}{\sqrt{p(x)q(x)}}, x \in \mathcal{X}. \tag{2.4}$$

If we multiply (2.4) by $p(x) > 0$, we get

$$|p(x) \ln p(x) - p(x) \ln q(x)| \leq \sqrt{\frac{p(x)}{q(x)}} \left| p(x) - q(x) \right|, x \in \mathcal{X}.$$

Summing over $x \in \mathcal{X}$ and using the generalized triangle inequality we get

$$\begin{aligned} 0 &\leq D_b(p\|q) = |D_b(p\|q)| \\ &= \frac{1}{\ln b} \left| \sum_{x \in \mathcal{X}} p(x) \ln \frac{p(x)}{q(x)} \right| \\ &\leq \frac{1}{\ln b} \sum_{x \in \mathcal{X}} |p(x) \ln p(x) - p(x) \ln q(x)| \\ &\leq \frac{1}{\ln b} \sum_{x \in \mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} \left| p(x) - q(x) \right| \end{aligned}$$

and the inequality in (2.1) is proved.

The case of equality follows from the fact that in (2.4) we have equality if and only if $p(x) = q(x)$ for all $x \in \mathcal{X}$. ∎

The following corollary is obvious

Corollary 2.2. *Assume that p, q are as above and, in addition,*

$$R := \max_{x \in \mathcal{X}} \frac{p(x)}{q(x)} < \infty.$$

Then we have the inequality

$$0 \leq D_b(p\|q) \leq \frac{\sqrt{R}}{\ln b} \|p - q\|_1 \tag{2.5}$$

where $\|\cdot\|_1$ is the usual $1-norm$, i.e., $\|t\|_1 := \sum_{x \in \mathcal{X}} |t(x)|$.

Corollary 2.3. *Assume that p, q are as above and, in addition,*

$$\Delta := \max_{x \in \mathcal{X}} |p(x) - q(x)|.$$

Then, we have the inequality

$$0 \leq D_b(p\|q) \leq \frac{\Delta}{\ln b} \|\frac{p}{q}\|_{\frac{1}{2}} \tag{2.6}$$

where $\|t\|_{\frac{1}{2}} := \sum_{x \in \mathcal{X}} |t(x)|^{1/2}$.

Another bound in terms of the $s-norm$ $(s > 1)$ of the difference $p-q$ also holds.

Corollary 2.4. *Let p, q be as above. Then*

$$0 \leq D_b(p\|q) \leq \frac{1}{\ln b} \|p - q\|_s \left\|\frac{p}{q}\right\|_{l/2}^{1/2} \tag{2.7}$$

where $s > 1$ and $\frac{1}{s} + \frac{1}{l} = 1$ and $\|t\|_s := \left(\sum_{x \in \mathcal{X}} |t(x)|^s\right)^{1/s}$.

Proof. The proof follows by Hölder's inequality

$$\begin{aligned} & \sum_{x \in \mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} |p(x) - q(x)| \\ \leq \quad & \left(\sum_{x \in \mathcal{X}} |p(x) - q(x)|^s\right)^{\frac{1}{s}} \times \left(\sum_{x \in \mathcal{X}} \left(\sqrt{\frac{p(x)}{q(x)}}\right)^l\right)^{\frac{1}{l}} \end{aligned}$$

and we omit the details. ∎

Remark 2.1. *For $s = l = 2$, we get the bound in terms of the euclidian norm of $p - q$*

$$\begin{aligned} 0 \quad \leq \quad & D_b(p\|q) \leq \frac{1}{\ln b} \|p - q\|_2 \left\|\frac{p}{q}\right\|_1^{1/2} \\ \leq \quad & \frac{\sqrt{R\,|\mathcal{X}|}}{\ln b} \|p - q\|_2 \end{aligned}$$

where R is as in Corollary 2.2.

The following corollary also holds.

Corollary 2.5. *Under the above assumptions for p, q we have the inequality*

$$0 \leq D_b(p\|q) \leq \frac{1}{\ln b} \left[\sum_{x \in \mathcal{X}} q(x) \left(\frac{p(x)}{q(x)} - 1\right)^2\right]^{1/2}. \tag{2.8}$$

The equality holds iff $p(x) = q(x)$ for all $x \in \mathcal{X}$.

Proof. We have

$$\begin{aligned} & \sum_{x\in\mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} \left|p(x)-q(x)\right| \\ = & \sum_{x\in\mathcal{X}} q(x) \sqrt{\frac{p(x)}{q(x)}} \left|\frac{p(x)}{q(x)}-1\right|. \end{aligned}$$

Using Cauchy-Buniakowski-Schwarz's inequality, we have

$$\begin{aligned} & \sum_{x\in\mathcal{X}} q(x) \sqrt{\frac{p(x)}{q(x)}} \left|\frac{p(x)}{q(x)}-1\right| \\ \leq & \left[\sum_{x\in\mathcal{X}} q(x) \left(\sqrt{\frac{p(x)}{q(x)}}\right)^2\right]^{1/2} \left[\sum_{x\in\mathcal{X}} q(x) \left(\frac{p(x)}{q(x)}-1\right)^2\right]^{1/2} \\ = & \left(\sum_{x\in\mathcal{X}} p(x)\right)^{1/2} \left[\sum_{x\in\mathcal{X}} q(x) \left(\frac{p(x)}{q(x)}-1\right)^2\right]^{1/2} \\ = & \left[\sum_{x\in\mathcal{X}} q(x) \left(\frac{p(x)}{q(x)}-1\right)^2\right]^{1/2}. \end{aligned}$$

The case of equality is obvious. ∎

The previous corollary has a consequence which can be more useful in practical applications.

Consequence. Let $\varepsilon > 0$. If

$$\left|\frac{p(x)}{q(x)}-1\right| \leq \varepsilon \ln b \text{ for all } x \in \mathcal{X} \ , \tag{2.9}$$

then we have the inequality

$$0 \leq D_b(p\|q) \leq \varepsilon. \tag{2.10}$$

Proof. Using (2.8) and (2.9) , we have

$$\begin{aligned} D_b(p\|q) & \leq \left[\sum_{x\in\mathcal{X}} q(x) \left(\frac{p(x)}{q(x)}-1\right)^2\right]^{1/2} \\ & \leq \left[\sum_{x\in\mathcal{X}} q(x) \left(\varepsilon^2 \ln^2 b\right)^2\right]^{1/2} = \varepsilon \end{aligned}$$

and the statement is proved. ∎

Remark 2.2. *From the inequality* (1.11) *we have the upper bound*

$$D_b\left(p\|q\right) \leq \frac{1}{\ln b} \left[\sum_{x\in\mathcal{X}} \frac{p^2(x)}{q(x)} - 1\right] \tag{2.11}$$

and from the inequality (2.1) *the bound*

$$D_b(p\|q) \leq \frac{1}{\ln b} \sum_{x\in\mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} \left|p(x) - q(x)\right|. \tag{2.12}$$

It is natural to compare these two bounds.

Assume that $|\mathcal{X}| = 2,\ q(1) = q(2) = \frac{1}{2},\ p(1) = p, p(2) = 1 - p,\ p \in (0,1)$. *Define*

$$B_1(p) := \sum_{x\in\mathcal{X}} \frac{p^2(x)}{q(x)} - 1 = (2p-1)^2$$

and

$$B_2(p) := \sum_{x\in\mathcal{X}} \sqrt{\frac{p(x)}{q(x)}} \left|p(x) - q(x)\right| = \sqrt{2}\left|p - \frac{1}{2}\right|\left(\sqrt{p} + \sqrt{1-p}\right).$$

The plot of the difference

$$B_1(p) - B_2(p), p \in (0,1) \tag{2.13}$$

shows that for $p \in (0, p_0) \cup (p_1, 1)$, *the second bound is better than the first one while for* $p \in [p_0, p_1]$ *the conclusion is the other way around. Note that* p_0 *is the first intersection with the op axis and is symmetric with* p_1 *which is the second intersection.*

3. Applications for Entropy

Let X be a random variable having the probability mass function $p(x) > 0$, $x \in \mathcal{X}$. Then we have the following converse inequality

Theorem 3.1. *Under the above assumptions, we have*

$$\begin{aligned} 0 &\leq \log_b |\mathcal{X}| - H_b(X) \\ &\leq \frac{1}{\ln b}\sqrt{|\mathcal{X}|} \sum_{x\in\mathcal{X}} \sqrt{p(x)} \left|p(x) - \frac{1}{|\mathcal{X}|}\right|. \end{aligned} \tag{3.1}$$

The equality holds in both inequalities simultaneously iff p *is uniformly distributed on* $\mathcal{X}$.

Proof. Choose in Theorem 2.1 $q(x)=\frac{1}{|\mathcal{X}|}, x\in\mathcal{X}$ to get

$$\begin{aligned} 0 &\leq \log_b|\mathcal{X}| - H_b(X) \\ &\leq \frac{1}{\ln b}\sum_{x\in\mathcal{X}}\sqrt{\frac{p(x)}{\frac{1}{|\mathcal{X}|}}}\left|p(x)-\frac{1}{|\mathcal{X}|}\right|. \end{aligned}$$

which is equivalent to (3.1).

The case of equality is obvious. ∎

If we assume that we know more information about the probability distribution p, i.e.,

$$\max_{x\in\mathcal{X}} p(x) \leq \frac{\tilde{R}}{|\mathcal{X}|};\left(\tilde{R}\geq 1\right) \tag{3.2}$$

then we can state the following corollary

Corollary 3.2. *If the probability distribution satisfies the condition* (3.2), *then we have the inequality*

$$0 \leq \log_b|\mathcal{X}| - H_b(X) \leq \frac{\sqrt{\tilde{R}}}{\ln b}\|p-u\|_1 \tag{3.3}$$

where u *is the uniform distribution on* $\mathcal{X}$ *. The equality holds in both inequalities simultaneously iff* $p=u$.

Corollary 3.3. *Let* X *be as above and assume that*

$$\tilde{\Delta} := \max_{x\in\mathcal{X}}\left|p(x)-\frac{1}{|\mathcal{X}|}\right|.$$

Then we have the inequality

$$0 \leq \log_b|\mathcal{X}| - H_b(X) \leq \frac{\tilde{\Delta}}{\ln b}\sqrt{|\mathcal{X}|}\,\|p\|_{1/2}. \tag{3.4}$$

The equality holds iff $p=u$.

The following estimation in terms of $p-norm$ also holds.

Corollary 3.4. *Let* p *and* u *be as above. Then*

$$\begin{aligned} 0 &\leq \log_b|\mathcal{X}| - H_b(X) \\ &\leq \frac{\sqrt{|\mathcal{X}|}}{\ln b}\|p-u\|_s\|p\|_{l/2}^{1/2} \end{aligned} \tag{3.5}$$

where $s>1$ *and* $\frac{1}{s}+\frac{1}{l}=1$, *and, particularly,*

$$\begin{aligned} 0 &\leq \log_b|\mathcal{X}| - H_b(X) \\ &\leq \frac{\sqrt{|\mathcal{X}|}}{\ln b}\|p-u\|_2\|p\|_1^{1/2}. \end{aligned} \tag{3.6}$$

The equality holds in both inequalities iff $p=u$.

The following corollary is useful in practice by its consequence which provides a sufficient condition for p so that $H_b(X)$ should be close enough to $\log_b|\mathcal{X}|$.

Corollary 3.5. *Let p be as above. Then*

$$\begin{aligned} 0 &\leq \log_b |\mathcal{X}| - H_b(X) \qquad (3.7) \\ &\leq \frac{\sqrt{|\mathcal{X}|}}{\ln b}\left[\sum_{x\in\mathcal{X}} (|\mathcal{X}|\, p(x) - 1)^2\right]^{1/2}. \end{aligned}$$

The equality holds iff $p = u$.

The following practical criterion holds.

Criterion. Let $\varepsilon > 0$. If

$$\left|p(x) - \frac{1}{|\mathcal{X}|}\right| \leq \frac{\varepsilon \ln b}{|\mathcal{X}|} \text{ for all } x \in \mathcal{X}$$

then we have the inequality

$$0 \leq \log_b |\mathcal{X}| - H_b(X) \leq \varepsilon. \qquad (3.8)$$

4. Applications for Mutual Information

Consider the random variables X and Y with the joint probability mass function $p(x, y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information is the relative entropy between the joint distribution and the product distributions, i.e.,

$$\begin{aligned} I_b(X;Y) &:= D_b(p(x,y)\|p(x)p(y)) \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log_b\left[\frac{p(x,y)}{p(x)p(y)}\right]. \end{aligned}$$

The following upper bound for the mutual information holds.

Theorem 4.1. *Under the above assumptions for X and Y, we have*

$$\begin{aligned} 0 &\leq I_b(X;Y) \qquad (4.1) \\ &\leq \frac{1}{\ln b}\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\sqrt{\frac{p(x,y)}{p(x)p(y)}}\,|p(x,y) - p(x)p(y)| \end{aligned}$$

with equality iff X and Y are independent.

The proof follows by Theorem 2.1 applied for the probability mass functions $p(x, y)$ and $p(x)p(y), (x, y) \in \mathcal{X} \times \mathcal{Y}$.

Corollary 4.2. *Assume that X and Y are as above and, in addition,*

$$1 \leq \xi = \max_{(x,y)} \frac{p(x,y)}{p(x)p(y)}.$$

Then we have the inequality

$$\begin{aligned} 0 &\leq I_b(X;Y) \qquad (4.2) \\ &\leq \frac{\sqrt{\xi}}{\ln b}\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}|p(x,y) - p(x)p(y)|. \end{aligned}$$

The equality holds in both inequalities simultaneously iff X and Y are independent.

We also have

Corollary 4.3. *Let X and Y be as above. If*

$$\Gamma := \max_{(x,y)} |p(x,y) - p(x)p(y)|,$$

then we have the inequality

$$0 \leq I_b(X;Y) \leq \frac{\Gamma}{\ln b} \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \sqrt{\frac{p(x,y)}{p(x)p(y)}}.$$

Another bound for the mutual information in terms of the $s - norm, s > 1$, is the following one.

Corollary 4.4. *Let X and Y be as above. Then*

$$\begin{aligned} 0 &\leq I_b(X;Y) \leq \frac{1}{\ln b} \left(\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} |p(x,y) - p(x)p(y)|^s \right)^{1/s} \\ &\times \left(\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \left(\sqrt{\frac{p(x,y)}{p(x)p(y)}} \right)^l \right)^{1/2} \end{aligned} \tag{4.3}$$

and, particularly, for $p = q = 2$

$$\begin{aligned} 0 &\leq I_b(X;Y) \leq \frac{1}{\ln b} \left(\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} |p(x,y) - p(x)p(y)|^2 \right)^{1/2} \\ &\times \left(\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \frac{p(x,y)}{p(x)p(y)} \right)^{1/2}. \end{aligned} \tag{4.4}$$

Finally, the following corollary holds

Corollary 4.5. *Under the above assumptions, we have*

$$0 \leq I_b(X;Y) \leq \frac{1}{\ln b} \left[\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x)p(y) \left(\frac{p(x,y)}{p(x)p(y)} - 1 \right)^2 \right]^{1/2}$$

The equality holds iff X and Y are independent.

The following consequence is useful in practice
Consequence. Let $\varepsilon > 0$. If

$$\left| \frac{p(x,y)}{p(x)p(y)} - 1 \right| \leq \varepsilon \ln b \text{ for all } (x,y) \in \mathcal{X} \times \mathcal{Y}$$

then

$$0 \leq I_b(X;Y) \leq \varepsilon.$$

References

[1] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons Inc., 1991.

[2] S.S. Dragomir and N.M. Dragomir, An inequality for logarithms and its applications in coding theory, *Indian Journal of Mathematics* (in press).

[3] S. S. Dragomir and C. J. Goh, Some bounds on entropy measures in information theory, *Appl. Math. Lett.,* **10**(3) (1997) 23–28.

[4] 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.
[5] M. Matić, C. E. M. Pearce and J. Pečarić, Improvements of some bounds on entropy measures in information theory, *Math. Inequal. and Appl.*, 1(1998) 295–304.
[6] M. Matić, C. E. M. Pearce and J. Pečarić, Refinements of some bounds in information theory, *J. Austr. Math. Soc. Ser. B,* to appear.
[7] S.S. Dragomir, N.M. Dragomir and K. Pranesh, Some Estimations of Kraft number and related results, *Soochow Journal of Mathematics*, **24**(4), 291-296.
[8] S.S. Dragomir, M. Sholz and J. Sunde, Some upper bounds for relative entropy and applications, *submitted.*
[9] S.S. Dragomir, Some remarks on the noiseless coding theorem, *submitted.*
[10] C. Calude and C. Grozea, Kraft-Chaitin inequality revisited, *J. UCS*, **5**(1996), 306-310.

School of Communications and Informatics, Victoria University of Technology, PO Box 14428, Melbourne City MCMC, 8001 Victoria, Australia
E-mail address: `sever@@matilda.vut.edu.au`